# PROCEEDINGS OF THE

# INTERNATIONAL CONFERENCE ON
# MEDICAL PHYSICS & BIOMEDICAL ENGINEERING

## Vol. 1

## MPBE '94

### 5 - 7 May 1994
### University of Cyprus
### Nicosia - Cyprus

*Organised and Co-Sponsored by*

The Cyprus Association of Medical
Physics and Biomedical Engineering
(CAMPBE)
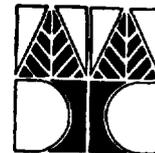
The Department of Computer Science,
University of Cyprus

# Proceedings of the

# International Conference on
# Medical Physics & Biomedical Engineering

## Vol. 1

## MPBE '94

5 - 7 May 1994, Nicosia, Cyprus

**Edited by**

*S. Spyrou*
*S. Christofides*
*C.S. Pattichis*
*E. Keravnou*
*C.N. Schizas*
*G. Christodoulides*

Organised and Co-Sponsored by

The Cyprus Association of Medical Physics and Biomedical Engineering
The Department of Computer Science, University of Cyprus

**Venue**

University of Cyprus

**Under the auspices of His Excellency
the President of the Republic of Cyprus
Mr. Glafcos Clerides**

The aim of the International Conference on Medical Physics and Biomedical Engineering (MPBE '94) is to provide an international forum for the presentation and discussion of the latest scientific and technical achievements by medical physicists, computer scientists, engineers, cognitive scientists and physicians whose work shares the goal of enhancing the provision of Health Care.

In all, about 100 papers will be presented in 18 sessions covering the topics of Education, Health Information Systems, Critical Care and Monitoring, Medical Imaging, Rehabilitation Engineering, Signal Processing and Analysis, Image Processing and Analysis, Artificial Intelligence in Medicine, Radiotherapy and Dosimetry, Cardiovascular Systems, Medical and Physiological Measurements and Control, Radiation Protection and Quality in Diagnostic Systems and Biomechanics and Implants. Amongst them there is a number of Invited Papers to be presented by distinguished scientists of international reputation in their respective fields.

The meeting is expected to be truly international since the selected papers to be presented come from 25 countries: Algeria, Austria, Belgium, Canada, Czech Republic, Finland, France, Germany, Greece, Israel, Italy, Poland, Portugal, ROC, Romania, Russia, Slovenia, Sweden, Switzerland, The Netherlands, UK, USA, Yugoslavia and Cyprus.

Both the Cyprus Association of Medical Physics and Biomedical Engineering and the Department of Computer Science of the University of Cyprus aim to organize international technical and scientific events of high standards. It is furthermore hoped that the Conference will establish itself as a regular international event for fostering and promoting developments within its scope.

It is hoped that the delegates from all these countries will find the Conference informative and professionally beneficial, will enjoy their stay in Cyprus, and that the informal contacts which always form a valued part of a conference, will endure and perhaps enable the interchange of views and experiences to continue long into the future.

Finally we would like to express our sincere thanks to C. Hadjiyianni, E. Sevasti, and E. Polykarpou for their valuable assistance during the preparation of this conference.

*S. Spyrou*
*S. Christofides*
*C.S. Pattichis*
*E. Keravnou*
*C.N. Schizas*
*G. Christodoulides*                                            *Nicosia, May 1994*

## Conference Co-Chairs

S. Spyrou
Higher Technical Institute
Nicosia, Cyprus
Tel: +357-2-305030
Fax: +357-2-494953

E. Keravnou
Department of Computer Science
Kallipoleos 75, P.O. Box 537
Nicosia, Cyprus
Tel: +357-2-366186/360589
Fax: +357-2-360881
email: elpida@jupiter.cca.ucy.cy

## Organising Committee

G. Christodoulides
Medical Physics Department
Nicosia General Hospital
Nicosia, Cyprus
Tel: +357-2-301350
Fax: +357-2-369170

S. Christofides
Medical Physics Department
Nicosia General Hospital
Nicosia, Cyprus
Tel: +357-2-301306
Fax: +357-2-369170

C.S. Pattichis
Department of Computer Science
Kallipoleos 75, P.O. Box 537
Nicosia, Cyprus
Tel: +357-2-366186/360589
Fax: +357-2-360881
email: pattichi@jupiter.cca.ucy.cy

C.N. Schizas
Department of Computer Science
Kallipoleos 75, P.O. Box 537
Nicosia, Cyprus
Tel: +357-2-360589
Fax: +357-2-360881
e-mail: schizas@jupiter.cca.ucy.cy

## International Scientific Committee

A. Achillides, Cyprus

K. Athanasiou, USA

E. Coiera, UK

F. Dams, UK

Ch. Kattamis, Greece

I. Krekule, Czechia

A. Lena, Italy

A. Nicolaides, UK

V.C. Roberts, UK

M. Tarata, Romania

H. Witte, Germany

J. Campbell, UK

A.G. Constantinides, UK

G. Harris, USA

C. Kazasis, Greece

N. Lavrac, Slovenia

M. Lyra, Greece

V. Patel, Canada

A. Schizas, Cyprus

G. Tzanakos, USA

K. Zadeh, Germany

B. Chandrasekaran, USA

C. Constantinou, USA

L. Ironi, Italy

K. Kouris, Cyprus

S. Laxminarayan, USA

I. Mozetic, Austria

B. Proimos, Greece

K. Steriopoulos, Greece

E.M. Tzanakou, USA

# Volume 1

## Medical Imaging

## Signal Processing and Analysis I

# Cardiovascular Systems

# Rehabilitation Engineering

# Artificial Intelligence and Expert Systems I

# Artificial Intelligence and Expert Systems II

# Image Processing and Analysis

# Signal Processing and Analysis II

# Education and Health Information Systems I

# Instrumentation

## Education and Health Information Systems II

# Volume 2

## Radiotherapy and Dosimetry

## Medical and Physiological Measurement and Control

## Radiation Protection and Quality in Diagnostic Systems

# Biomechanics/Implants

# Critical Care and Monitoring

# Round Table Discussion on Orthopaedic Biomechanics of the Spine, Hip and Knee

# Round Table Discussion on Artificial Intelligence in Medicine

# Poster Papers

## Additional Papers

**Magnetic Resonance Angiography:**

**Principles, Status and Perspectives**

**N. Papanikolaou BSc , J.P. Grant PhD**
**PHILIPS GREECE MEDICAL SYSTEMS**
**15, March 25th Street, Tavros, Athens 17778**

**Abstract:**

Magnetic Resonance (MR) techniques have for a long time provided a standard method for

qualitative and quantitative studies of flow [1,2]. More recently, MR has been attracting considerable

interest as a possible new tool which promises to replace certain invasive X-Ray Angiography

(XRA) procedures. The new fast developing field of Magnetic Resonance Angiography (MRA) will

be a well established clinical tool which will help together with other modalities (US-Doppler) in the

near future for the non-invasive assessment of vascular diseases.

**Introduction**
The MR signal from flowing fluid, such as blood, depends on several factors including: (i) Flow factors including the direction of flow, motion characteristics (velocity, acceleration, pulsatility) and flow mode (plug,laminar,turbulent) and (ii) MR factors such as the type and parameters of the pulse sequence and the spatial direction and type of imaging gradients.
These factors tend to combine in a rather complicated manner to influence the MR image contrast.

**MR Flow Principles**
We may consider that flow influences the MR image in two principal ways :
i) Time of Flight effects (Inflow-TOF) and ii) Spin-Phase Shift effects.
These effects will be here examined seperately.
  Flow induced (or paradoxical) signal enhancement results from the replacement of partially-saturated spins by fully magnetised spins in the imaged volume. Depending on the pulse sequence, the velocity of the spins can affect the final flow signal [3]. In the case of a Spin-Echo sequence, rapid flow will cause a flow induced signal loss, due to the fact that a proportion of spins within the imaged volume will not receive both 90 and 180 degree RF pulses and/or leave before they contribute an echo signal. In the case of slow flow, the spins contribute to the final signal because they receive both 90 and 180 degree RF pulses. On the other hand, flowing spins in GRadient Echo (GRE) pulse sequences with short Tr and Te using first order moment nulling (velocity compensation), give a high signal which depends mainly on the velocity of spins but also on other parameters; the higher the velocity of the spins the less are the saturation effects, resulting in a higher signal from flowing spins [4] (Figure 1).
  All the previous analysis refers to laminar or plug flow. In the case of turbulent flow the so called "dephasing effects" reduce dramatically the signal intensity of flowing spins and produce flow voids[5], a phenomenon first reported in 1982 [3]. These dephasing effects can now be suppressed by implementing various strategies, which mainly include the use of the shortest possible echo times

2

as well as the reduction of voxel size in order to minimize intravoxel dephasing. In this way, possible artifacts caused by turbulence are minimised.

Spin-Phase shifts occur whenever protons move with respect to a magnetic field gradient [6]; in MR, gradients are responsible for slice selection (convertially termed the Gz gradient) and position encoding (convertially termed Gy, the phase encoding gradient and Gx, the frequency encoding gradient). Individual protons precess with a frequency directly proportional to the strength of the magnetic field they experience. The magnitude of the magnetic field gradient with respect to which a group of protons move, and the time, tp, during which they interact with the gradient, define the total phase shift they accumulate compared to protons at the nominal resonant frequency [7]. A group of protons which remain in phase with each other as they move is termed an isochromat. In the rotating frame, stationary spins at a point X' gain during the time, tp, a phase angle,φ , given by:

$$\Phi = \gamma \ Gx \ X' \ dt \qquad (1)$$

where is the gyromagnetic constant. In the case where the spins move at a constant velocity Vx, in the x-direction, the coordinate X' becomes time-depented:

$$X' = X + Vx \ t \qquad (2)$$

Substitution shows that the phase change due to flow along the gradient Gx is:

$$\Phi(t) = 1/2 \ \gamma GxVxt \qquad (3)$$

The conclusion is that the phase shift due to flow is *directly* proportional to the velocity of the flowing spins. This conclusion is very important for quantifing blood flow with MR. It is for this reason that considerable effort has been expended to develop phase sensitive sequences, which can divided into three main categories:

i) Rephase-Dephase
ii) Phase Contrast-Complex Difference
iii)Phase Contrast-Phase Difference

All off them are based on the subtraction of two different image sets [8]. In the **Rephase-Dephase** technique we subtract one flow compensated set of images (Rephased) and one dephased set of images - in both of them the stationary tissue cancels out and only the vessels are highlighted in the subtracted images. In the second kind of sequences we have a flow encoding gradient applied in one, two or three spatial directions, which means that flow can be encoded in all spatial directions. Interleaved during acquisition between the previous sets is a flow compensated image set which acts as a reference. After that, subtraction takes place between the three flow encoded data sets and the reference set. In the case of **Complex Difference** the pixel intensity is proportional to the sine of the phase angle φ :

$$\text{Signal Intensity} = 2M|\sin(\varphi)| \qquad (4)$$

The **Complex Difference Imaging** leads to morphological MR Angiograms (qualitative information). In the case of **Phase Difference** pixel intensity represents velocity and flow direction information which results in quantitative velocity information in the reconstructed images.

## Established Clinical Status of MRA

MRA is presently considered to be a clinical tool mainly for the head and neck vessels. More specifically in the area of neuro applications there are a number of well established MRA applications including:
- Cerebral Aneurysms
- Arteriovenous Malformations, Venous Angiomas (Figure 2)
- Dural Sinus Thrombosis
- Hypervascular Tumours - Meningiomas (Figure 3)
- Carotid Diseases: Dissection,Stenosis,Atheromatosis

In all these pathologies there have been extensive clinical evaluations of MRA in comparison with the "gold standard" of XRA. MRA seems from these results to be capable of replacing XRA in all

the previous applications and may also be used as a screening tool by virtue of the non-invasiveness of the method [9].

## Perspectives and Potential of MRA

MRA applications in other parts of the body such as the thorax [10], liver [11], kidneys [12],and extremities [13] are currently undergoing clinical evaluation and the first results seems very promising,e.g. in the cases of portal vein thrombosis resulting from liver chirrosis, renal artery stenosis (Figure 4). Deep vein thrombosis also gives a very high percentage of sensitivity and specificity in MRA [14]. Another very important future application of MRA is the non-invasive qualitative and quantitative study of Coronary Artery Disease (CAD). MRA still has some way to go in order to realize MR Coronary Angiograms but in the near future the non-invasive assessment of CAD may be possible [15]. The use of intravenous blood-pool contrast agents will also increase drastically the usefullness of MRA in the near future [16].

In conclusion, MRA, together with US-Doppler and recently CT-Angio, may eventually replace today's gold standard of XRA for many purposes. MR scanners are still improving technologicaly with consequences that are expected to be of especial significance to MRA. Thus, better gradients in terms of strength (mT/m), and significant shorter rise times are currently in use and even higher and faster gradients will be used in the near future. Also, more sensitive RF coils are being designed, such as the so-called "Synergy or phased array" coils which realize a tremendous improvement in Signal to Noise ratio. In the near future it will be daily clinical practice to use 512X512 (Figure 5) or even 1024X1024 matrices for certain applications,which means that MRA images can be expected to improve dramatically in terms also of spatial resolution, which may eventually equal that of XRA.

## REFERENCES:
1.Carr H.Y. and Purcell E.M. Effects of diffusion on free precession in nuclear magnetic resonance experiments. Phys.Rev. 94 630, 1954

2. Singer J.R. Blood flow rates by NMR measurements. Science 130 1652, 1959

3. J.P. Grant , C.Back. NMR Rheotomography: Feasibility and clinical potential. Med.Phys 9(2) March/April 1982 188-193

4. RR. Edelman. Basic principles of magnetic resonance angiography. Cardiovasc IntervRadiol 1992;15:3

5. J.C. Gatenby, J.C. Gore. Measurement of turbulent intensity using partial echo tchniques. Procc SMRM 1992 Berlin p. 2914

6. Van Dijk P. Direct cardiac NMR imaging of heart wall and blood flow velocity. JCAT 1984;8:429-436

7. D.G. Nishimura, A. Macovski, J.M. Pauly. Magnetic Resonance Angiography.IEEE Transaction on Med Imaging Vol MI-5 No 3 Sept 1986

8. Dumoulin C.L. Souza S.P., Darrow R.D. Simultaneous acquisition of phase-contrast angiograms and stationary tissue with Hadamard encoding of flow induced phase shifts. J Magn. Reson. Imaging 1991;1:399-404

9. D.Barnes. MRA and Fast Spin Echo ready for clinical use. Diagnostic Imaging Jan 1993;97-103

10.Wielopolski P, Haacke E, Adler L. Improved diagnosis of pulmonary embolisms with a proton density and T1W 3D imaging protocol. Proc SMRM New York 1993 p. 335

11.Finn P.J. Goldmann A. , Edelman R.R. Magnetic resonance angiography in the body. Magn Reson Q 1992; 1:1-22

12.Kim D, Edelman R.R., Ken K.C. et al. Abdominal Aorta and renal artery stenosis: evaluation with MRA. Radiology 1990; 174:727

13.Yucel E.K.,.Dumoulin C.L., Waltman A.C. MRA of lower-extremity arterial disease: preliminary experience. J. Magn Reson Imaging 1992;2:303-309

14.Erdman W.A., Jayson H.T., Redman H.C., Miller G.L., Parkey R.W., Peshok R.W. Deep Venous Thrombosis of the extremities: role of MR Imaging in the diagnosis. Radiology 1990; 174:425-431

15.Manning W.J. , Li W. , Edelman R.R. A preliminary report comparing MR Coronary Angiography with conventional angiography. N Engl J Med 1993;328:828-832
16.V.M. Runge, J.E. Kirsch, C. Lee. Contrast-enhanced MRA. JMRI 1993;3:233-239

Figure legends

Figure 1:
Coronal projection MR Angiogram in the region of the neck where MRA gives satisfactory results. Inflow M2D mode in axial orientation (perpendicular to the flow in order to maximize flow-stationary tissue contrast - Courtesy of Dr. J. Heintz)
Figure 2:
Arteriovenous Malformation evaluated by MRA 3D Phase Contrast technique (Courtesy of Dr. L.Lobez Ibor).
Figure 3:
Combination of MR Imaging and MR Angiography: Hypervascular meningioma studied by 3D Phase Contrast MRA (Three orthogonal projections) . Axial MRI image, Turbo Spin Echo T2W of the same meningioma (Courtesy of Dr. L. Lopez Ibor).
Figure 4:
3D Phase Contrast application for renal artery stenosis screening
(Courtesy of Dr. J. Heintz)
Figure 5:
Axial projection MR Angiogram of the cerebral arteries using high-resolution
3D-Inflow technique (Scan Matrix 256X512, 1mm thin slices, Courtesy of Philips Medical Systems , Best , The Netherlands)



Figure 1
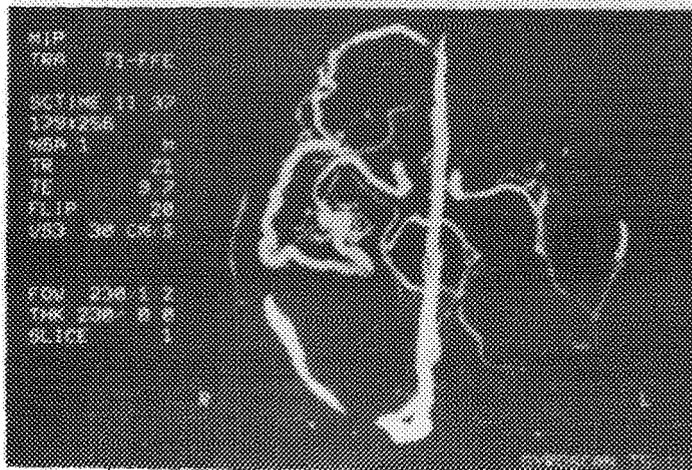


Figure 2
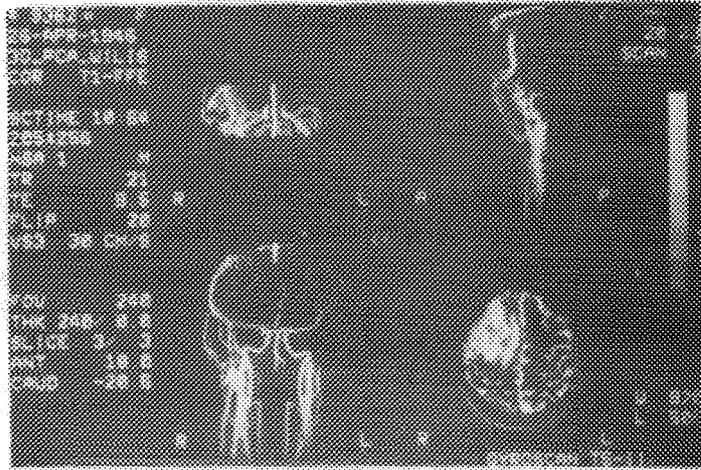
Figure 3
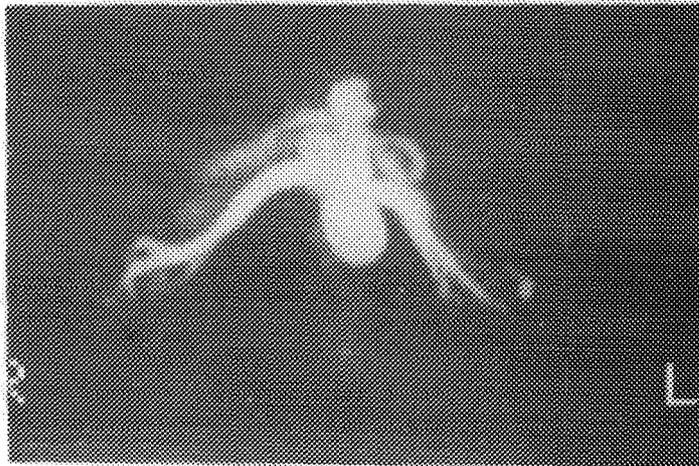


Figure 4
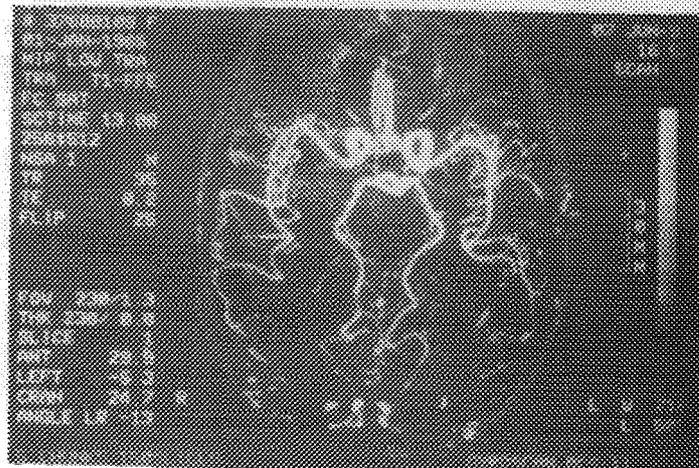


Figure 5

# Examination of an Egyptian Mummy
## - STEREOLITHOGRAPHY APPLIED

Henrik Hjalgrim, M.D. (1), Niels Lynnerup, M.D. (1), Martin Liversage, graduate student (2)

(1) Laboratory of Biological Anthropology, Institute of Anatomy, and (2) Department of Pediatric Dentistry, School of Dentistry, Faculty of Health Sciences, The Panum Institute, University of Copenhagen, Blegdamsvej 3 . DK - 2200 Copenhagen N.

## ABSTRACT

This paper describes the techniques of three dimensional imaging and stereolithography based on serial CAT-scans applied to the examination of the skull of an Egyptian mummy. Both the three dimensional image and the polymeric cast of the mummy skull presented finer details. It was confirmed that the subject was a male, approximately 30 - 35 years old. Fracturing of the ethmoid bone, sequelae to the removal of the brain, was observed in both types of presentations. Apart from this and signs of parodontitis, no pathology was observed. Stereolithography is a most powerfull, non-destructive approach to the study of mummies. It might solve some of the problems of reburials, and further be of value in forensic medicine and paleo-ontology.

## INTRODUCTION

Until the 1970s thorough anthropological investigations of mummies could only be conducted at the cost of the destruction, not only of the embalmed individual itself, but also of the wrappings. Due to the cultural and historical value of the wrappings and the associated decorations, this price has in many cases been considered too high (e.g., Vahey and Brown, 1984; Joergensen, 1989). Thus, until Hounsfield's introduction of Computer Assisted Tomography (CAT) (Hounsfield, 1973), the anthropologist had to aquiesce with the comparatively crude X-ray examinations (Harwood-Nash, 1979).

The continous evolution of the techniques have made investigations based on CAT-scannings realistic alternatives to conventional, destructive investigations. This paper presents an investigation of the skull of an Egyptian mummy with by the means of CAT-scanning, three dimensional imaging and Stereolithography (SLA).

## Material and Methods

### The mummy
At the Carlsberg Glyptotek, Copenhagen, Denmark, an Egytian mummy is exhibited. It is unique in possessing what is considered to be one of the finest mummy portraits known, presumably depicting the embalmed. Little is known about the mummy apart from that it originates from the Roman cemetry of Hawara in El-Faiyum, Egypt. The characteristics of the portraited person (hairstyle and dress), the style of the portrait together with its origin, suggest that the mummy can be dated to the first century A. D. (Shore, 1962). The individual was most likely of a wealthy family of either Greek, Roman or mixed ancestry (Shore, 1962). The possibility of the individual having been Egyptian is remote, jugded on historical facts as well as on the portrait itself (Shore, 1962).

## Methodology
### CAT-scanning
Still in its original wrappings the mummy was CAT-scanned at Gentofte County Hospital (courtesy of Dr. Rosenklint). Scans of the head were made with a 1,5 mm interval with a gantry angle of 0 degrees, and a scanning time of 3 seconds (140 kV peak, 140 mA-seconds). In all 129 exposures were made.

### Three Dimensional Imaging
The data from the scanning were stored on ordinary magnetic tapes and transmitted to a Silicon Graphics IRIS 4D/20 graphic workstation from Silicon Graphics equipped with Analyze™ software package (Liversage, 1992). On the computer, each picture was editted, i. e., all soft tissues and bandages were removed, leaving only the bony tissue.

After the editting the 129 images were stacked, and using the resizing and volume rendering functions of ANALYZE™, three dimensional images or rather projections of the mummy skull, viewed from angles specified by the programmer, could be produced.

### Stereolithography
Whereas the computerized images in reality are three dimensional projections, based on conditions specified by the programmer, the data from the CAT-scans can also be used to produce a polymeric cast of the head. To do so, the editted dataset underwent a series of transformations, somewhat the equivalent of interpolations. The original dataset could thus indirectly be used to guide two laser canons aming into a basin of acrylic monomers. For the monomers to bond a minimum of $1,2 \times 10^{-18}$ J is required. Suffcient heat to cause bonding will arise where the two beams cross (Hunziker, 1992). This part of the project was carried out by The Danish Institute of Technology.

# Results

### CAT-scans
When reviewing the CT-scans it was observed that the brain- as expected- had been removed. This had seemingly been done by aspiration transnasally through a hole in the ethmoid bone, a destruction which could be observed. Apart from that the third upper molar was missing, no signs of pathology were observed. However, it was noted that the skin of the head in some way had been retracted toward the back of the head.

### Three dimensional image
The resolution of the image was remarkable. Even the finest structures, e.g. the mental canals and foramina could be observed. Generally, in the three dimensional projections, it was obvious that the temporal bones were much thinner than the rest of the cranial vault, as holes in the bones became apparent after the editting. Also, most of the sutures could be observed, appearing mostly open. The molar was again observed missing, but it became appearent, that this may be due to severe parodontitis in this part of the right maxillary bone. Apart from this and the lesions of the ethmoid bone no other signs of pathology were seen.

### SLA model
The model demonstrated even the finest details. For instance impressions from the medial menigeal arteries and the transverse sinuses could be seen on the inside of the

calvarium. In the cranial base all foramens were open, and structures like the styloid proces were present. In the facial skeleton details like the frontal sinus, the sutures between the nasal bone and the mental canals could readily be seen.

The signs of parodontitis were observed again, and the lost tooth were rediscovered, lying loose in the neck. Hence, it must have been lost post mortem. Apart from that the teeth appeared with onle slight attrition.

## Discussion

Examination of the stereolithographic model confirmed that the subject was male. Based on the sutures and the dental attrition, he was assumed to be adult to mature, approximately 30 - 35 years old. The finding of parodontitic changes are not unusual for an individual from this period and culture (Jakobsen, 1994), and similar observations in three dimensional reconstructions have been reported (Lewin et al., 1990).

The fracturing of the ethmoid bone is a known sequelae to the techniques for removal of the brain and in agreement with previous observations involving CAT-scanning (e.g., Harwood-Nash, 1979; Vahey et Brown, 1984; Marx et D'Auria, 1986; Pickering, 1990; Taylor et Hughes, 1993).

The defects in the temporal bones in the three dimensional image as well as in the polymeric cast are artifacts, resulting from the editting in combination with the thin quality of bone; similar defects have been observed previously (Künzel et al., 1992). The method of sterelithography in itself is accurate within 0,1 mm, and thus the limitting proces is presently the CAT-scan minimum slicing interval.

The use of CAT-scanning in mummy studies are well known, but the involvement of three dimensional reconstructions is a fairly new approach (Lewin, 1987; Marx et D'Áuria, 1988; Magid et al., 1989; Lewin et al., 1990; Pickering et al., 1990; Pahl et al., 1992; Germer et al., 1993; Taylor et Hughes, 1993). Even more new are the applications of stereolithography in this field of anthropology. It might be the solution to some of the problems of reburials the subjects are scanned before reburial and the data stored. Then in case of interest, casts of e. g. the bones can be made. Also, there may be forensic applications (Lewin et al., 1990), especially in the field of craniofacial reconstruction. Hence, it should be possible to reconstruct partly destructed crania by means of mirror imaging.

## Conclusion

To our knowledge, stereolithographic casts of mummy skulls have been made only twice previously (z. Nedden et Wicke, 1992; Bohen, 1992), but the prospects seem promissing. Although the method presently is somewhat time consuming, it is a powerfull, non-destructive approach to the study of mummies, with promissing prospects in several other fields.

## REFERENCES

Bohen BE 1992
Collaborative Investigation of the UI Egypto-Roman Mummy
Abstr., Congreso Internacional de Estudios sobre, 134

Harwood-Nash DCF 1979
Computed Tomography of Ancient Egyptian Mummies
J Comp Ass Tomogr, 3(6), 768-773

Hounsfield GH 1973
Computerized transverse axial scanning (tomography), part 1. Description of the System
Br. J. Radiol., 46, 1016-22

Hughes S, Taylor J 1993
Wer war Tjentmutengebitu?
Interview in: Siemens Zeitschrift, 67(6), 12-7

Hunziker M 1992
Chemistry for Stereolithography
Proceeding of International Workshop on Stereolithography in Medicine, 10-3

Jakobsen J 1994
Personal communication

Joergensen, J. B., 1989
Anthropology of the Qilakitsoq Eskimos,
in, The mummies from Qilakitsoq - Eskimos in the 15th century
Ed. J. P. Hart Hansen, H. C. Gulloev
Meddelelser om Groenland, Man and Society, 12

Künzel KH, Steinlecher M, Gaber O, Platzer W 1992
Morphologische Vergleichstudie an Schädlen Zur Schädel-CT-Rekonstruktion des Eismannnes
Der Mann im Eis; Veröffentlichungen der Universität Innsbruck, 187; 117-131

Lewin PK 1987
Stereoscan imaging from composite 'CAT'scans
Paleopathol Newsletter (Suppl on 14th annual Meeting, April 1987)

Lewin PK, Trogadis JE, Stevens JK, 1990
Three Dimensional Reconstructions From Serial X-Ray Tomography of an Egyptian Mummified Head
Clinical Anatomy, 3, 215-8

Liversage M 1992
Industrial Case Studies, Dental College of Copenhagen
in: Layer Manufacturing- a Challenge of the Future; The Final Report from the NOR-SLA Project; Ed. Øyvind Bjørke
Tapir Publishers, Trondheim Norway

Marx M, D'Auria SH 1986
CT examination of Eleven Egyptian Mummies
Radiographics, 6(2), 321-30

Magid D, Bryan B, Drebin R, Ney D, Fishman EK
Three Dimensional Imaging of an Egyptian Mummy
Clinical Imaging, 13, 239-40

Marx M, D'Auria SH 1988
Three Dimensional CT Reconstructions of an Ancient Human Egyptian Mummy
Am J Roentg, 150, 147-9

Nedden D zur, Wicke K 1992
Der Eismann aus der Sicht der radiologischen und computertomographischen Daten
Der Mann im Eis; Veröffentlichungen der Universität Innsbruck, 187, 131-48

Pahl WM, Höhne KH, Grodd W, Lindemann W 1992
High-tech in Egyptian Mummies: 3D-reconstruction and 3D-CAD
Abstr., Congreso Internacional de Estudios sobre, 203

Pickering RB, Conces Jr DJ, Braunstein EM, Yuorco F . 1990
Three Dimensional Computed Tomography of the Mummy Wenuhotep
Am J Phys Anthr, 83, 49-55

Shore AF 1962
Portrait Painting from Roman Egypt
The Trustees of The British Museum, London

Vahey T, Brown D 1984
Comely Wenuhotep: Computed Tomography of an Egyptian Mummy
J Comp Ass Tomogr, 8(5), 992-7

# MR IMAGING OF PRIMARY AND SECONDARY ANEURISMAL BONE CYSTS

Lj. Poleksić[1], M. Atanacković[2], S. Slavković[3] and Đ. Zdravković[1]

[1]Magnetic Resonance Center, University Clinical Center, Pasterova 2, Belgrade
[2]Institute of Pathology, Faculty of Medicine, University of Belgrade
[3]Special Orthopedic Hospital "Banjica", Belgrade

## ABSTRACT

Five patients with primary or secondary aneurismal bone cysts (ABC) were examined on magnetic resonance imaging (MRI). Purpose was to evaluate the role of MRI in tissue characterization of these lesions and its accuracy in determination of ABC extent. All patients were surgically treated. The following diagnostic parameters were evaluated on MR images: extent of the lesion, margins, septations, fluid-fluid levels. Posttherapeutical MRI findings were analyzed according to presence or absence of rest/recurrence. All lesions had expansile appearance and were extracompartmental. Both, primary and secondary ABC, were well defined, with presence of internal septations and fluid-fluid levels within the cavities. MRI signal intensities, within the loculations, reflected different ages of blood and its degradation products. In secondary ABC differentiation between osteoblastoma and ABC was not possible, although contrast agent was administrated. After the therapy rest/recurrences were seen in 2 of 3 MRI controlled patients. We conclude that described MRI characteristics can be suggestive of ABC but they are not sufficient for establishing the diagnosis because some other bone tumors can have same MR appearance. Histologic confirmation is required for definite diagnosis and for differentiation of primary from secondary ABC. At the same time MRI is the method of choice for surgical approach planning.

## INTRODUCTION

Aneurismal bone cyst (ABC) is an expansile, litic tumor-like lesion of the bone, which usually occurs in patients under 20 years of age. Histologically it consists of thin wall, blood-filled cavities (cavernomatous spaces) which lack the normal features of blood vessels. ABC could be a) primary or simple and b) secondary, when it is associated with another bone lesion, either benign or malignant. Their clinical, radiological and computed tomography features have been already described (1,2,3). Magnetic resonance imaging (MRI) characteristics have been evaluated on small numbers of cases (4,5). The purpose of this article is to compare MRI with histological and surgical findings in patients with primary and secondary ABC.

## MATERIALS AND METHODS

Five patients with histologically confirmed, 4 primary and one secondary ABC were preoperatively examined on 1.5T MR imager (Magnetom, Siemens). Secondary ABC coexisted with osteoblastoma of cervical vertebra. Patients were 11 to 19 years old (mean 15). Lesions were localized in cervical spine (3), sacrum and scapula. MRI protocol consisted of T1W (TR/TE 380-500/15 ms) and T2W (TR/TE 1600-2200/90 ms) spin-echo images in at least two orthogonal planes. Slice thickness was 4-8 mm. All patients were surgically treated and three of them underwent control MRI examinations, with same protocol, after the treatment. Paramagnetic contrast agent (Magnevist, Schering) was administrated in a patient with secondary ABC on both examinations. The following diagnostic parameters were evaluated on MR images: extent of the lesion, margins, septations, fluid-fluid levels. Posttherapeutical MRI findings were analyzed according to presence or absence of rest/recurrence.

## RESULTS

All lesions had expansile appearance and were extracompartmental. MRI depicted intra and extraosseous extension precisely in all cases. Affected vertebral bodies were fractured in both primary ABC of the cervical spine. Posterior elements of cervical vertebra were involved in all (3) cases. On T1W images ABC were isointense or slightly hyperintense relative to surrounding muscles, with areas of different levels of hyperintensity within the lesion. On T2W images they were hyperintense (Fig.1a and b). Extraosseous margins of soft tissue mass were lobulated and well delineated. Thin low signal intensity (SI) rim, on both T1W and T2W images, completely or partially, surrounded all ABC. Although MRI vs. pathologic specimen correlation was not available in most cases, according to histologic examination of the ABC periphery these areas related to fibrous tissue. Different number of internal septations, characterized by low SI was observed in both primary and secondary ABC. Septa defined margins of numerous cavities. In some of them fluid-fluid levels were seen, better or only, on T2W images. Inferior fluid layers were usually isointense or slightly hyperintense, relative to surrounding muscles, while superior were hyperintense. On T1W images superior layers, when present, were hyperintense or isointense. These distributions of SI indicated presence of extracellular methemoglobin or serous fluid in the superior layer. Neighbouring loculations could have different SI but in general it increases with increasing T2 Weighted (Fig. 2). Contrast agent was administrated once. Lesion enhanced contrast diffusely indicating its good vascularization. Surrounding tissue also enhanced contrast due to presence of edema (Fig. 3 ).

Three patients were controlled on MRI after the surgical therapy. Two of them underwent partial reduction of the lesion while in secondary ABC, curettage and irradiation was done. Minimal changes were detected after the surgical treatment in patients in whom partial reduction of the lesion was performed. Patient with secondary ABC didn't show signs of rest/recurrence. In remaining two patients vertebral body resection and bone grafting was performed in one and Tikhor-Linberg resection of the shoulder girdle in the other.
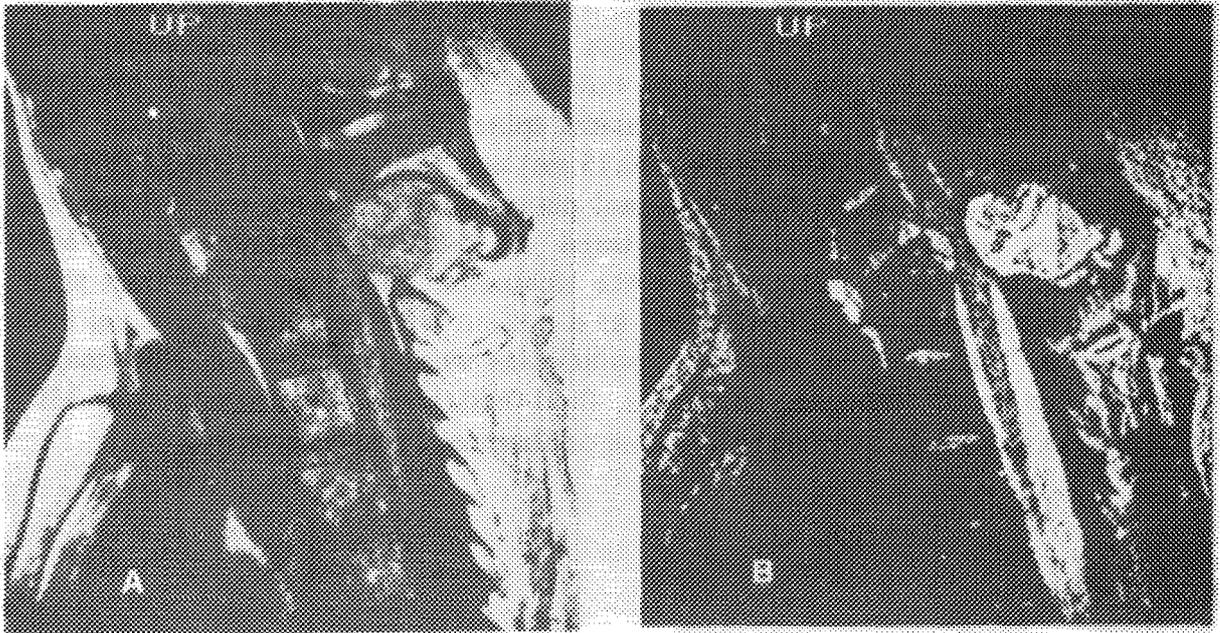
Fig.1. ABC OF CERVICAL SPINE: a) Sagittal T1W image of cervico-thoracic spine. Large, expansile lobulated lesion of posterior elements expanding into posterior aspects of fractured vertebral bodies and spinal channel. Lesion is slightly hyperintense with hardly visible fluid-fluid level. b) T2W image at the same level. Lesion is hyperintense, with fluid-fluid levels.
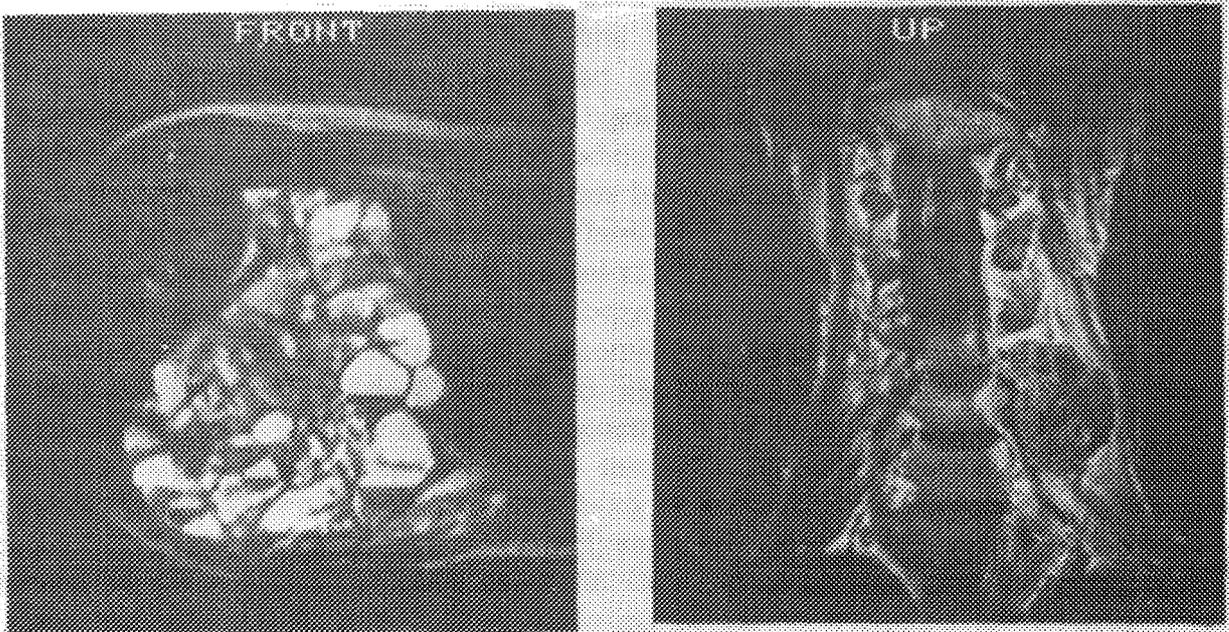


Fig.2. ABC OF SCAPULA: Axial T2W image of the shoulder girdle. Expansile, with low SI rim well delineated lesion. Fluid-fluid levels within numerous cavities.

Fig 3. SECONDARY ABC OF CERVICAL SPINE: Postcontrast T1W coronal image. Diffuse contrast enhancement of the soft tissue mass, left lateral side of vertebral bodies and surroundings muscles. Differentiation between osteoblastoma and ABC is not possible.

# DISCUSSION

MRI is now accepted as the examination of choice for estimation of musculoskeletal tumors extent. Its ability to recognize histological type or to differentiate benign from malignant lesions is limited. We evaluated the MRI characteristics of ABC and compared them with histological and intraoperative findings. ABC most frequently occur in the long bones, but in approximately 20% of cases they involve vertebral bodies (1,2,6). In our material in four patients they were localized in cervical and sacral vertebral bodies. This site requires more serious therapeutical approach because of lesion proximity to the spinal cord and nerve roots. Hence, it requires detailed preoperative visualization of the entire lesion. Examined, both primary and secondary lesions were well delineated as it was previously reported (4,5), with a low SI rim which is believed to represent fibrous tissue (pseudocapsula). However, in patients with ABC of the spine neurologic deficits were present because of irritation and compression of nerve roots and/or central compression of spinal cord caused by spinal extent of pseudotumor. Soft tissue component rather tended to compress surrounding structures than to infiltrate them indicating its slow growth and benign nature. Internal septations were in continuity with the low SI margins. Fluid-fluid levels were seen within well defined cystic cavities in all patients. Unfortunately fluid-fluid levels are nonspecific finding for ABC since they can occur in various bone and soft tissue tumors (7). Distribution of MRI signal intensities, within the loculation, reflected different ages of blood and its degradation products. Commonly different SI between neighbouring cavities is probably due to the same cause. Differentiation between osteoblastoma and ABC according to MR image, although contrast agent was administrated, was not possible in secondary ABC.

Intraoperative and MRI findings correlated very well. Excellent spatial and contrast resolution on multiple plane images enabled adequate preoperative planning, which was especially important in patients with ABC of the spine.

We conclude that described MRI characteristics can be suggestive of ABC but they are not sufficient for establishing the diagnosis because some other bone tumors can have the same MR appearance. Histologic confirmation is required for definite diagnosis and for differentiation of primary from secondary ABC, which is not possible on the basis of MR image. At the same time MRI is the method of choice for surgical approach planning.

## LITERATURE:

1. Dahlin DC, McLeod RA: Aneurismal bone cyste and other nonneoplastic conditions.Skeletal Radiol (1982) 8: 243-250

2. Bonakdarpour A, Levy WM, Aegerter E: Primary and secondary aneurismal bone cyst: a radiological study of 75 cases. Radiology (1978) 126: 75-83

3.Hudson TM: Fluid levels in aneurismal bone cyst: a CT feature. AJR (1984) 142: 1001-1004

4. Munk PL, Helms CA, Holt RG, Johnston J, Steinbach L. Neumann C: MR imaging of aneurismal bone cysts. AJR (1989) 153: 99-101

5. Beltran J, Simon DC, Levy M, Herman L, Weis L, Mueller CF: Aneurismal bone cysts: MR imaging at 1.5T. Radiology (1986) 158: 689- 690

6. Capanna R, Albisinni U, Picci P, Calderoni P, Campanacci M, Springfield DS: Aneurismal bone cyst of the spine. J. Bone and Joint Surg. (1985) 67-A: 527-531

7. Tsai JC, Dalinka MK, Fallon MD, Zlatkin MB, Kressel HY: Fluid-fluid level: a nonspecific finding in tumors of bone and soft tissue. Radiology (1990) 175: 779-782

# Cornea profile from Keratoscopic and Scheimpflug images

**F. Silva [1], H. Araújo [2], E. Leite [3], J. Murta [3], J. Cunha-Vaz [3]**

[1]Instituto Superior de Engenharia de Coimbra, Quinta da Nora, 3000 Coimbra, Portugal.

[2] Dept. Electrical Engineering University of Coimbra 3000 Coimbra, Portugal.

[3] IBILI-CAMTO, Praça Mota Pinto, 3000 Coimbra, Portugal

## ABSTRACT

Topographic analysis systems used to analyse the cornea profile have some limitations due to the aspheric nature of the cornea. In many cases the cornea is modelled by a spherocylindrical surface. However the normal cornea is aspheric and flattened from the center to the periphery. Moreover diseased and postsurgical corneas rarely approximate a spherocylindrical surface. Topographic analysis systems based on Keratoscopic images provide quantitative information by comparing the mire diameters from corneas to those reflected from standard reference spheres and by measuring hemichord lengths. In this paper we describe an aproach we have developed to improve the topographic reconstruction of the cornea by combining Keratoscopic and Scheimpflug images.

## 1. Introduction

Current topographic systems are unable to analyse severely distorted Keratoscope mires. Therefore a number of problems prevent the parametric descriptors of the corneal topography from achieving high precision and accuracy in all cases[1]. We propose to overcome this problem by combining Scheimpflug and Keratoscopic images. Scheimpflug images are obtained from a plan that is a meridian containing the optical axis of the eye. We are using four Scheimpflug images spaced 45º apart, enabling the characterization of the corneal asphericity[2]. We are developing methods to calibrate both Keratoscopic and Scheimpflug images. These methods of calibration use a reference eye model and take into account the optical characteristics of the Scheimpflug imaging device and of the Keratoscope.

## 2. Keratoscopic images

The photokeratoscope is a device that enables the acquisition of images of the cornea's anterior surface. The photokeratoscope projects several concentric circular mires on the cornea's anterior surface. The shape of those mires, reflected by the cornea, are imaged (Fig.1a).



a) Imaged          b) Detected circles
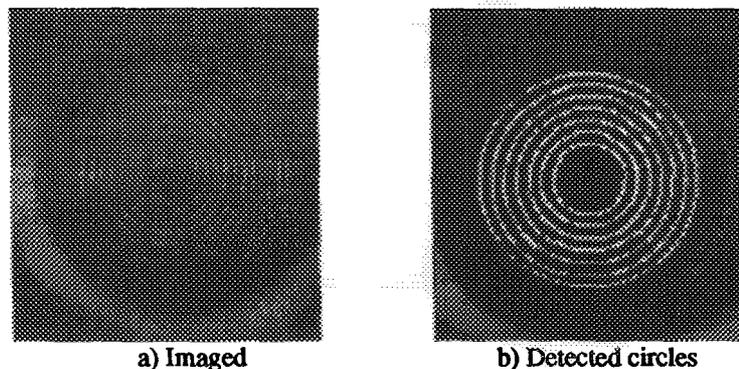
Fig.1

A normal cornea is modelled by a spherocylindrical surface. The image reflected by the cornea gives us information regarding the optical power of the anterior corneal surface[3]. If the cornea is spherical, the reflected rings of light will appear circular, concentric and equally spaced from each other. If the cornea is distorted, then the reflected rings will present deviations from that shape[4]. The image rings will appear elliptical on astigmatic corneas, will be closer to each other on steeper corneas and will be spread out on

flattened corneas. By measuring the radii of the several circles (generated by the rings) along different meridians it is possible to compute the corneal optical power at different points as well as to have an estimate of its topography[3][5].

## 3. Analysis of the keratoscopic images

From a keratoscopic image it is possible to estimate the tridimensional configuration of the cornea, by using the coordinates of several points located on the rings reflected by the cornea.

Fig.2 represents a schematic cuting, showing the incidence on the cornea of the light emmited by the first keratoscopic ring ($r_i$) followed by its reflection ($r_r$), which is imaged by a camera whose optical axis is oriented along the cornea axis.
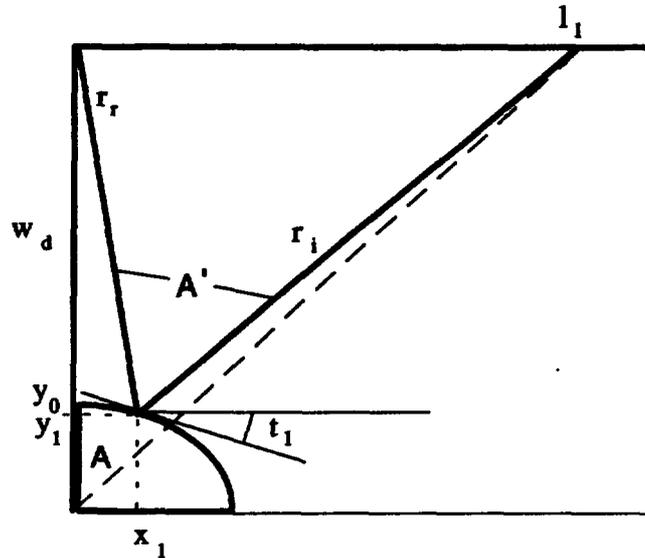


Fig.2

$X_i$ represents the circle radius reflected in that point, while $y_i$ is the cornea elevation in the same point, relatively to the plan that passes through the cornea's curvature center and is orthogonal to the cornea axis. The distance $l_i$ is measured on the instrument, and the distance $w_d$ has a fixed value, assuming that the phototokeratoscope is focused to the middle of the patient's cornea. By inspection of the Fig.3 we can establish the equations[5]:

$$y_i = y_{i-1} - \frac{(x_{i-1} - x_i)(\cos t_{i-1} - \cos t_i)}{\sin t_{i-1} - \sin t_i} \quad ; \quad t_i = \frac{\pi}{2} - \frac{A'}{2} - \text{atan} \ (\frac{d_0 - y_i}{l_i - x_i})$$

in which $t_i$ is the angle between the horizontal and the tangent to the cornea surface in the point $(x_i, y_i)$.

To calculate $y_i$ it is necessary to know the angle $t_i$. On the other side $t_i$ depends on $y_i$. So, the value $y_i$ is calculated by iteration of both equations, initializing $t_i$ to $A/2$.

Knowing the $(x_i, y_i)$ points and the angles of their meridians, we have the tridimensional characterization of the cornea. The curvature radii ($rc_i$) in the $(x_i, y_i)$ points are calculated by the expression $rc_i = x_i / \cos(\pi / 2 - t_i)$. Those radii allow us to get the dioptric power on those points by the expression $D_i = 337.5 / rc_i$, in which $rc_i$ is expressed in millimeters, and the numerator is the gradient in refractive index between air and the eye.

The measures obtained from the photokeratoscope images are affected by several sources of error, that we can point out:

- Earlier analyses[3] have considered the cornea's radius of curvature at the center ($Y_0$) as having a standard value of 7.8 mm. This biases the calculations on markedly flattened or steeped corneas. More recent analyses[5] estimate the parameter $Y_0$ by computing the average radius ($\overline{x}_1$) from the first keratoscope cornea ring image, and then using the equations:

$$\overline{y}_0 = \frac{\overline{x}_1}{\sin\theta} \ ; \ \theta = \frac{\text{atan}\left[\frac{l_1 - \overline{x}_1}{w_d}\right] - \text{atan}\left[\frac{\overline{x}_1}{w_d}\right]}{2}.$$

17

This estimated elevation is not very accurate when used with keratoscopes having the innermost ring too large to cover the cornea's central zone.

- Another source of inaccuracy is the incertitude in how well the focusing of the instrument was done by the operator. This originates two sources of lack of precision:
  - A wrong value of $w_d$ in the equations, biasing all the results;
  - The instrument optical axis is not collinear relatively to the cornea optical axis, because there is a radial decentration, or a tilting between the cornea and the orthogonal plane to the instrument optical axis. This causes the appearence of images with elliptical mires, suggesting a false astigmatism.

We propose to extract information from Scheimpflug images to correct the inaccuracy of keratoscopic images.

## 4. Scheimpflug images

The Scheimpflug is a device that acquires images from a plane containing the optical axis of the eye. The Scheimpflug principle allows us to get images of the ocular structures of the eye's anterior segment, namely cornea, anterior chamber and lens. With this instrument it is possible to get images in any meridian containing the optical axis (Fig.3).
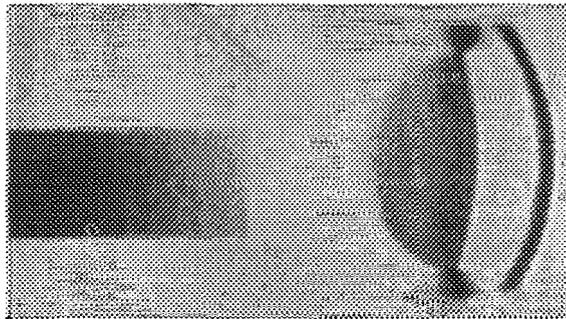


Fig.3

From one single scheimpflug image we can obtain the cornea thickness and the radii of curvature in every point of that meridian with a good accuracy. It is possible too to estimate the location of the optical axis. To obtain this estimate the eye's lens is used. By combining two or more images, we can get more information. Using a set of Scheimpflug images it is possible to verify whether or not the cornea's axis is coincident with the optical axis of the eye's lens and also to compute the angular deviation between the direction of the two axes.

## 5. Combining data from images

By using the keratoscopic image it is possible to predict the profile of the cornea's external surface along the meridian corresponding to a Scheimpflug image. If any deviations are found out a procedure is then used to decrease the error in the estimate of the center's location as well as of the different radii calculated in the keratoscopic image. These parameters were used to improve the accuracy and precision of the various topographic parameters computed from the Keratoscopic image.

Using four scheimpflug images we have calculated in each one:
- The cornea axis,
- The radii and dioptric power in several points of that image's meridian. Those points are selected so that their distances to the cornea axis ($x_i$) are the same as the distances of the points selected in the keratoscopic images.

Considering that the measures obtained from scheimpflug images have a better accuracy than the measurements obatined from keratoscopic images, they were used to replace the measurements done in the corresponding locations of the keratoscopic map (Table 1).

The points on the keratoscopic map that do not have corresponding locations on the scheimpflug map, are corrected by interpolating the measurement differences found on both maps, in the neighborhood of each position[6][7][8]. The computation is performed by means of a weighted average of the six closest neighbors measurements. The weights are inversely proportional to the distances between the neighbors and the point whose measurements are to be computed.

| Radius Direction (°) | Image radius (x; mm) | Cornea elevation (y; mm) | Dioptric power (Keratosc.) | Corrected Dioptric power |
|---|---|---|---|---|
| 000 | 2.482 | 7.412 | 43.13 | 43.04 |
| 045 | 2.482 | 7.428 | 43.14 | 43.18 |
| 090 | 2.482 | 7.430 | 43.14 | 43.97 |
| 135 | 2.480 | 7.418 | 43.18 | 43.42 |
| 180 | 2.482 | 7.414 | 43.14 | 43.00 |
| 225 | 2.482 | 7.410 | 43.13 | 43.06 |
| 270 | 2.420 | 7.424 | 44.33 | 45.01 |
| 315 | 2.447 | 7.421 | 43.81 | 44.04 |

Table 1. Dioptric power corrected with data from Keratoscopic and Scheimpflug images
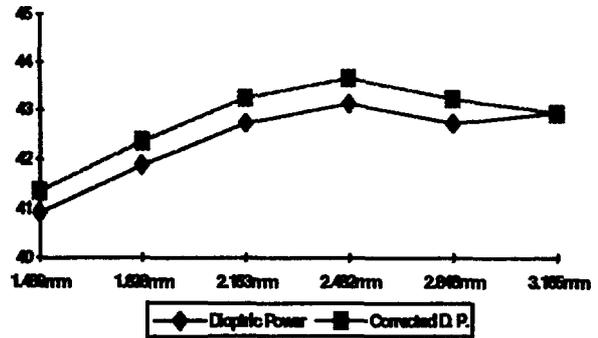


Fig. 4 - Dioptric Power throuth one direction

## 6. Conclusions

By combining cornea images obtained from different devices, we can improve the accuracy in the three-dimensional reconstruction of the cornea. One of the devices provides measurements with good accuracy whereas the other one (the keratoscope) provides dense measurements (but affected by inaccuracies). Using only Scheimpflug images a 3D reconstruction of the cornea would not be very accurate because from each patient only a maximum of 4 images (in normal clinical practice) are acquired. On the other hand a single keratoscopic image provides much more data points, but the measurements are affected by several sources of error that may severely distort any 3D cornea reconstruction performed by using only keratoscopic information. By using the scheimpflug information it it possible to obtain a better quality and better accuracy in the 3D reconstruction of the cornea.

Another approach we are using is based on a computed keratoscopic map. From this map the steepest and flattest meridians and other representative directions are computed and localized. That information is used to define the orientation of the Scheimpflug images. Based on these Scheimpflug images a new reconstruction of the corneal surface is performed with higher precision and accuracy.

The 3D reconstruction of the cornea surface is extremely useful for several medical procedures namely to the manual or laser excimer surgery of the cornea. It is also important to evaluate the evolution of the cornea after surgery as well as when subjected to different types of treatment.

We are developing models and methods for a more accurate and precise calibration of the images, to achieve a 3D reconstruction of the cornea that may be considered as a safe medical information.

REFERENCES:

[1] Steven E. Wilson, Stephen D. Klyce: Advances in the Analysis of Corneal Topography, Survey of Ophthalmology, Vol. 35, p. 269-277,1991.

[2] W.E.L. Grimson: From Images to Surfaces. Cambridge:MIT Press, 1981.

[3] J. Doss, R. Hutson, J. Rowsey, R. Brown: Method for calculation of corneal profile and power distribution, Arch. Ophthalmol, Vol 99, p.1261, 1981

[4] J.M. Beck, P.T. Farouki, J. Hinds: Surface Analysis Methods, IEEE Computer Graphics and Applications, 12:13-36, 1986

[5] Stephen D. Klyce: Computer-Assisted Corneal Topography, Invest. Ophthalmology & Visual Science, Vol. 25, p. 1426, 1984.

[6] T. Boult, J.R. Kender: Visual Surface Reconstruction Using Sparses Depth Data, Proc. Computer Vision and Pattern Recognition Conf., pp. 68-76, Miami, Fla, 1986.

[7] A. Blake: Reconstructing a Visible Surface, Proc. of Nat'l Conf. on Artificial Intelligence, pp. 23-26, Austin, Tx, 1984.

[8] E.J.Farrel and R.A. Zappulla: Three-dimensional data visualization and Biomedical applications, CRC Critical Rev. Biomed. Eng., vol 16, pp. 323-363, 1989.

# RELATIONSHIP BETWEEN FRACTURES OF THE VERTEBRAE AND PRESENCE OF SPINAL HAEMATOMA

Borota Lj, Urgent Centre, University Clinical Centre, Pasterova 2, Belgrade, Serbia, Yugoslavia,

Drndarski B, Institute of Oncology and Radiology, Pasterova 14, Belgrade, Serbia, Yugoslavia,

Stankovic D, Urgent Centre, University Clinical Centre, Pasterova 2, Belgrade, Serbia, Yugoslavia,

Bajic R, Department of Radiology, Institute of Neurosurgery, University Clinical Centre, Visegradska, Belgrade, Serbia, Yugoslavia.

## ABSTRACT

Thirty-eight patients with fractured vertebrae were involved in our study. Patients were evaluated by standard plain radiographs of the spine, myelographies and post-myelographic computed tomographies of the damaged part of the spine. The aim of the study was to give an answer to the question if there is difference between frequency of the appearance of spinal haematoma in patients with fractured vertebral bodies and in patients with fractured vertebral arches. For that reason all patients were divided into two groups. Trauma of the spine was followed by spinal haematoma in the first group of the patients, whereas in the second group of patients spinal haematoma was not evidenced. The number of the fractured vertebral bodies and fractured vertebral arches in both groups of the patients were determined. Differences between these numbers were tested by chi square test.

We concluded that fractures of the vertebral bodies and of the vertebral arches are followed by spinal haematoma in the same degree.

## INTRODUCTION

Trauma of the spine may be associated with injuries of the blood vessels resulting in epidural or subdural haematoma within the spinal canal. Contrary to extracerebral haematomas which usually occur with injuries of arterial vessels, haematomas within the spinal canal usually result from trauma of the venous plexus (1). Haematomas extending along the spinal canal have to be removed since they may lead to irreversible lesions of the medullas. It is, therefor, necessary to diagnose haematoma immediately upon hospitalization. Clinically, spinal, extramedullar haematomas are characterized with progressive neurological deficit which cannot be explained by diagnosed skeletal injury. Development of clinical picture in this way in patients with spinal injury is the main indication for myelography, i.e. post myelographic CT scan.

The aim of the study was to show wether spinal, extramedullary haematomas occur more frequently in patients with fractured vertebral bodies or in patients with fractured vertebral arches.

It was our opinion that eventual positive result may be of prognostic importance for development of spinal haematoma in patients with fractures of certain parts of the vertebra.

## MATERIAL AND METHODS

The study included 38 patients of both sexes hospitalized for spinal trauma at the Emergency Center, University Clinical Center, Belgrade, in the period 1991 to 1993. Discrepancy of neurological findings and extend of spinal lesion indicated myelography and post-myelographic CT of the injured part of the spine in all those cases.

Standard X-rays (AP, LL and oblique projections) of the injured parts of the spine were made in all patients after admission and neurological and orthopedic examination. Contrast medium Omnipaque 240 was injected into subarachnoid space in amount of 20 ccm. CT scans were performed of the injured and adjoining unaffected vertebrae using Somatom DR Siemens or General Electric 8800 equipment after standard myelograms (AP and LL projections) had been made.

## RESULTS

According to the results of standard radiographies of the spine, myelograms and CT scans all patients were divided into two groups. The first group comprise patients with diagnosed intraspinal, extramedullary haematomas, while the second included patients without haematoma (Table 1).

| | 1st group (spinal haematoma present) | 2nd group (spinal haematoma absent) |
|---|---|---|
| number of patients with fractured vert.bodies | 3 | 6 |
| number of patients with fractured vert.bodies and pedicle/lamina | 2 | 5 |
| number of patients with fractured vert.bodies, pedicle and lamina | 13 | 5 |
| number of patients with fractured lamina/pedicle | 1 | 3 |

**TABLE 1**

Since only pedicles were broken in some of them, and in some only laminae, for the purpose of the study the vertebral arch implied lamina and pedicle separately (Figure 1 and Figure 2). Association of spinal haematoma and number of fractured vertebral bodies and fractured laminae and pedicles was analyzed by chi square test. The results have shown that occurrence of spinal haematoma is associated with fractures of vertebral bodies and fractures of the pedicles and laminae to the same degree.

## Figure 1

**1st GROUP (SPINAL HAEMATOMA PRESENT)**
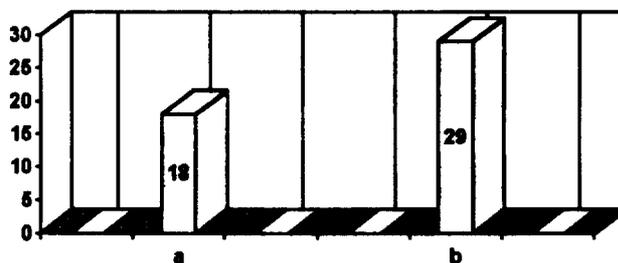


**a - number of fractured vertebral bodies**
**b - number of fractured pedicles and laminae**

## DISCUSSION

Referential literature suggests that spinal extramedullary haematoma is relatively uncommon complication of spinal trauma, although studies based on results of MR images reveal intraspinal haematoma in about 61% of indi-viduals with spinal injuries (2).

## Figure 2

**2nd GROUP (SPINAL HAEMATOMA ABSENT)**
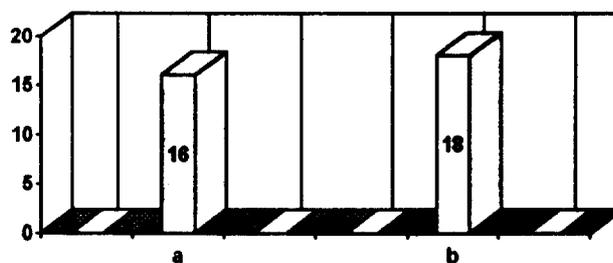


**a - number of fractured vertebral bodies**
**b - number of fractured pedicles and laminae**

Intraspinal extramedullary haematoma may develop slowly but very quickly, as well, necessitating emergency surgery (3). Prompt detection of haematoma

in the spinal canal is therefore of utmost importance. It has already been established that prognosis of injuries associated with spinal haematoma depends on severity of neurological picture and time elapsed from diagnose establishing to surgical removal (4).

Since the posterior and anterior venous plexuses, the injury of which results in haematoma, are localized along the laminae and pedicles and in the corner between the pe-dicles and vertebral bodies (1), we proposed that probability of occurrence of intraspinal haematoma will be higher if pedicles and/or laminae are injured. The conclusion was based on preliminary observations in our patients. Besides, available literature failed to support causal correlation of spinal haematoma and fractures of certain parts of the vertebrae. We also wanted to establish whether fractures of laminae and pedicles were associated with spinal haematoma mere commonly than fractures of vertebral bodies. However, statistical analysis of the results showed that fractures of vertebral bodies and fractures of the pedicles and laminae are associated with spinal haematoma to the same degree. It is therefore obvious that occurrence of haematoma may be expected irrespectively of the site of vertebral injury. It should always be kept in mind, especially in the light of the fact that spinal haematoma may develop very slowly and need not be associated with dramatic clinical presentation, particularly not at the beginning (5, 6).

We also believe that the study should be extended by introduction of additional parameters such as the number of fractured fragments, their size, shape and localization for establishing of their possible association with occurrence of spinal haematoma and other lesions accompanying spinal trauma.

## LITERATURE

1. Taveras JM, Ferrucci JT: RADIOLOGY - DIAGNOSIS, IMAGING, INTERVENTION. rv ed. Vol 3, chapt 104, Philadelphia: J.B. Lippincott Company, 1992

2. Karlslake RW, Jaspan T: MAGNETIC RESONANCE IMAGING OF SPINAL TRAUMA. Br J Radiol 64: 386-402, 1991

3. Gruszkiewicz J, Doron Y, Lamberger A, Bo-rovich B, Feinsod M: ACUTE SPINAL EXTRADURAL HAEMATOMA. Neurochirurgia (Stuttg) 30: 88-90, 1987

4. Calliauw L, Dhara M, Martens F, Vannerem L: SPINAL EPI DURAL HAEMATOMA WITHOUT LESION OF THE SPINE. REPORT OF FOUR CASES. Clin Neurol Neurosurg 90: 131-136, 1988

5. Laursen J, Fode K, Dahlerup B: SPINAL EPIDURAL HAEMATOMAS. Clin Neurol Neurosurg 89: 247-253, 1987

6. Olshaker JS, Barish RA: ACUTE TRAUMATIC CERVICAL EPIDURAL HEMATOMA FROM A STAB WOUND. Ann Emerg Med 1991 20: 662-664, 1991

# Visual Pathways Properties Characterization by Simultaneous Transient and Steady-State Stimulation

A.Tzelepi, A. Bezerianos, P. Papathanasopoulos*
Department of Medical Physics, *Neurological Clinic
School of Medicine, University of Patras, Patras 26500, Greece

## Abstract

This work presents a novel method to study the properties of the visual pathway, using simultaneous transient and steady-state hemifield (left-right, upper-lower) stimulation. The multiple information obtained (half-field testing, transient and steady-state responses), together with the time saving and the easy application of the protocol, suggests the clinical application of this method as a practical clinical tool for the assesment of visual dysfunctions.

## Introduction

Transient Pattern Reversal Visual Evoked Potentials (PRVEPs) have been widely used for assessing dysfunctions of the visual pathway. With increasing stimulation rates, faster than 4rev/sec, the transient waveform begins to deteriorate, and a sine wave of the frequency of stimulation, the so-called steady-state VEP is obtained. Although steady-state response lacks of individual peaks and latencies, it allows to present stimuli in a much less testing time and it provides sensitive estimates of the VEP amplitude. Furthermore, steady-state stimulation has been proved very-sensitive in detecting diseases mainly affecting cells which respond preferentially to high temporal frequencies e.g. glaucoma.

Presenting both types of stimuli in hemifields simultaneously, complementary information from transient and steady-state stimulation is extracted in less testing time and under the same conditions. Also, by tagging the two hemifields with different temporal frequencies which individually lead to distinct responses, it is possible to examine the interaction and distribution of the hemifiled reponses, as well as, the processing of fast and slow temproral frequencies in the visual cortex.

It has previously been reported that the potentials recorded anywhere on the head following full-field stimulation, are an algebraic sum of activities generated by the individual hemispheres. With lateral simultaneous stimulation, the transient and steady-state responses follow different routes and end at different sites in the visual cortex and consequently, it is expected to record more prominently: a) on one lateral electrode, the transient VEP response, b) on the other lateral electrode, the steady-state VEP response, c) in the midline ($O_Z$), the algebraic sum of the activities generated by the individual half-fields. Besides, EEG and EMG underlying activity, which contaminates the signal and results to distorted waveforms, is added, the same to all recording sites. Thus the subtraction of the signal recorded laterally from the signal recorded in Oz, will discard any underlying, unwanted activity and reveal the original waveform. Furthermore, the detection of differences will be emphasized.

Unlike PRVEPs from lateral hemifields, those from upper (UVF) and lower visual fields (LVF) have not been generally used as they did not demonstrate clinical applications. However, there is clear experimental support that P100 is largely generated by lower visual field stimulation (Skrandies, 1984).

This paper will present a novel method to study the properties of the visual pathway using simultaneous half-field transient and half-field steady-state stimulation for right-left and upper-lower hemifields.

## Materials and Methods

Five healthy subjects (4 male, 1 female), their ages ranging from 21 to 27 years, with normal visual acuity, and normal visual history, were tested.

Ag-AgCl electrodes, 10mm in diameter, were fixed to the scalp with collodion, and placed at $O_z$ (midline), $O_1$ and $O_2$, referred to $F_z$. Electrode impedance was kept <3k$\Omega$.

The stimulus used was a checkerboard pattern presented in a 12in diagonal Nicolet Visual Display Unit (VDU). Each check subtended 25.8'. The luminance of the white checks was 164cd/m$^2$ and the contrast of the pattern was 0.75 (Lmax-Lmin/Lmax+Lmin).

Subjects were comfortably seated in a darkened room, in 1m distance from the screen subtending $14^o$x$11^o$. They were instructed to concentrate and keep fixation on a small light spot in the center of the screen. All recordings were made monocularly, with the non-stimulated eye covered, and repeated twice to ensure fixation.

A Nicolet Compact Four electrodiagnostic system was used to acquire, amplify and filter the signals. The band pass filter was set to 1-100 Hz. The 3-channel analog output was fed to a Metrabyte 16F ADC card with sampling frequency 512Hz and stored in a PC DOS for further off-line processing. Artifact rejection was performed by Nicolet before averaging.

A programmable video pattern generator (Bezerianos and Papathanasopoulos, 1991) was used to produce the stimulation protocol. The simultaneous stimulation was achieved through the following procedure: first of all four patterns are produced. The basic unit of each pattern is the hemifield. Each hemifield can be either 0 or 1, i.e. the checkerboard pattern or its reversal. Full field patterns are then constructed by the different combinations of the hemifields and can be one of 11, 10, 01, or 00. Patterns 11 and 10 are the "reversals" of 00 and 01 respectively. The patterns are projected in the sequence:

$$00 \; 01 \; 00 \; 01 \; 00.... \; 00 \; 11 \; 10 \; 11 \; 10 \; 11.... \; 11$$



In this way, two parameters are introduced: a) appearance time of each pattern, b) number of patterns in the sequence. Altering these two parameters produces the desired couple of temporal frequencies, and selecting appropriate hemifields defines which pattern hemifield is to be reversed with each frequency (the above sequence demonstrates a slowly reversing left hemifield). In the above sequence of patterns two sweeps are included. Consequently, each sweep consists of one transient and n/2 steady-state responses.

The appearance time of each pattern was set to 50ms and n=34 patterns were used; thus one hemifield was stimulated with 20Hz (1/50ms) reversal rate and the other hemifield with 1.18Hz (1/ [(34/2)x50ms]) reversal rate. The sweep duration was 850 ms, and the recording time was 500 ms. The extra 350 ms were allowed for monitoring and saving to disk. The total number of the sweeps was 100.

In order to prevent the interaction of the hemifields, the central strip of the pattern subtending $1^o43'$ of the visual field was not reversing. This technique, previously been reported by Bodis-Wollner et al (1992), reduces the possibility of contamination of the two hemifields as they are simultaneously stimulated. This is of primary consideration as the interaction of the two hemispheres may lead to erroneous results.

## Results

*Right-left hemifield stimulution*

All the subjects appeared to have the transient VEP recorded most prominently on the side ipsilateral to the transient half-field (Fig. 1). When the subtraction of the signal

L - Eye

R- Eye

**Figure 1.** VEP recordings with simultaneous left half-field transient and right half-field steady-state stimulation, elicited monocularly from subject PS. The transient and steady-state responses seem to lateralize ipsilaterally to the respective stimulating field $(O_1, O_2)$, while in the midline $(O_Z)$ a compound response is recorded. After the subtractions $O_Z$-$O_2$ and $O_Z$-$O_1$ clear transient and steady-state waveforms are revealed. No shift on the latency of P100 is observed.

L-Eye

R- Eye
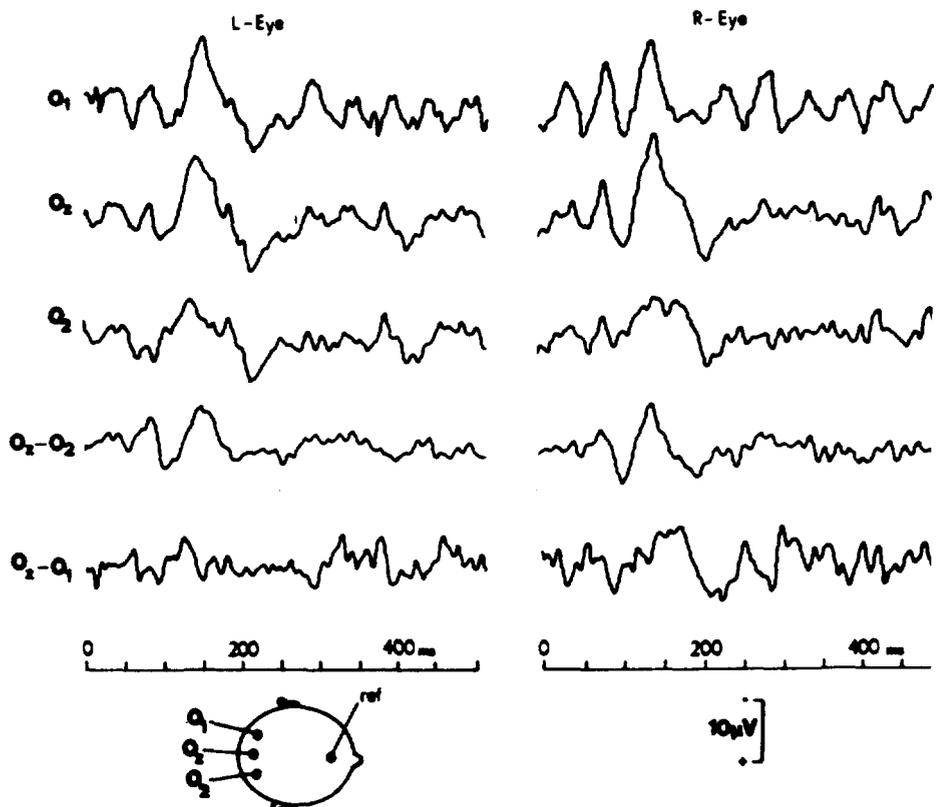
**Figure 2.** VEP recordings with simultaneous left half-field transient and right half-field steady-state stimulation, elicited monocularly from subject BN. The transient VEP emerges more prominently, ipsilaterally to the slowly reversing hemifield $(O_1)$. After the subtraction $O_Z$-$O_2$, the transient response is clearly present. The steady-state response appears contralaterally to the fast reversing hemifield, after the transient response $(O_1)$. The subtraction $O_Z$-$O_1$ gives no distinguishable response.

30

contralateral to the transient stimulation, from the one in $O_Z$ took place, it revealed consistently a dominating P100 peak, clearer, better-shaped and sharper. No shift on P100 latency was observed (Fig. 1).

Our results show the steady-state response either ipsilaterally or contralaterally to the steady-state stimulation. Three subjects showed clear evidence of a steady-state response (Fig. 1) ipsilaterally to the fast reversing hemifield, the other two on the contralateral side, after the transient response. Then, the subtraction of the signal recorded contralaterally to the steady-state stimulation, from the one recorded in Oz was performed. The three above mentioned subjects revealed a clear steady-state response (Fig. 1), although amplitude was reduced; the other two did not present a distinguishable waveform (Fig. 2).

*Upper-lower hemifield stimulation*

As with upper and lower hemifield stimulation there is no vertical symmetry, responses were recored only in $O_Z$. Upper half-filed steady-state stimulation and lower half-field transient stimulation resulted in a dominant transient response. Upper transient and lower steady-state stimulation gave a clear steady-state reponse, and although the steady-state response is dominating, the transient VEP cannot be considered absent. All the subjects tested, showed evidence of the transient response but with an attenuated P100 peak. Fig. 3 demonstrates a representative example from subject KG.

Figure 3. VEP recordings of subject KG with simultaneous (A) upper half-field transient and lower half-field steady-state stimulation which resulted in a dominating steady-state response (B) upper half-field steady-state and lower half-field transient. which resulted in a clear P100 peak. Recordings were made in $O_Z$.
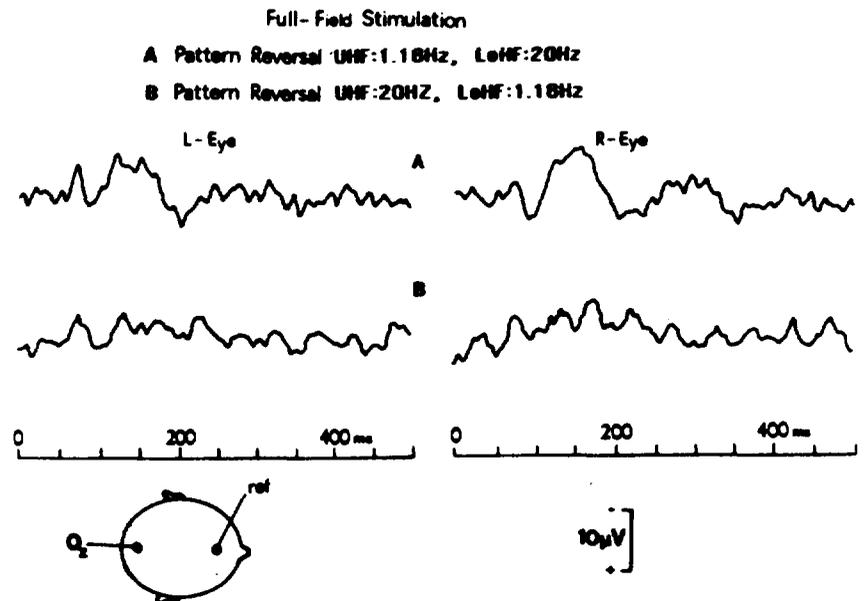


## Discussion

The idea that different types of information from common sensory receptors can be processed along separate pathways is far from new. There is clear evidence that the mammalian visual system consists of at least two major processing streams, the M- and P- pathways which consist of magnocellular and parvocellular groups respectively (Livingstone, 1988). They originate within the retina, and they maintain their anatomical segregation through to the visual cortex. It has been shown that P-cells respond preferentially to slow temporal stimulation, high spatial frequencies, high contrasts and colour. M- cells are more sensitive to fast temporal stimulation, low spatial frequencies, low contrasts and they are insensitive to colour.

Previous studies imply the existence of a magnocellular dominated lower visual field and a parvocellular dominated upper visual field (Previc, 1988). With upper half-field transient and lower half-field steady-state stimulation, the P-dominated UVF is

31

stimulated with a slow temporal frequency and the M-dominated LVF with a fast temporal frequency. In this respect, both subsystems are stimulated preferentially and, consequently, it is expected to acquire both signals, transient and steady-state, but as LVF stimulation results in an attenuated P100 peak, the transient response is expected to have a low P100 peak. With upper half-field steady-state and lower half-field transient stimulation, the UVF and LVF are not preferentially stimulated. However, the LVF being excited with transient stimulation is expected to give a prominent P100 peak. Our results support the above implications.

With left-right hemifield stimulation the transient response was recorded ipsilaterally to the slowly reversing hemifield. This is in agreement with the reports of other researchers about half-field transient stimulation ("paradoxical distribution" Barett et al, 1976). The steady-state response was recorded either ipsilaterally or contralaterally to the rapidly reversing hemifield. In the midline, the algebraic sum of the resonses in the two hemispheres, was recorded. Lateral hemifield stimulation may provoke overlapping responses from both subsystems and explain the spreading steady-state activity over the different leads. Further research on the distribution of the steady-state response has to be made, and in this respect explicit conclusions about magnocellular and parvocellular components cannot be drawn.

The method of stimulation proposed in this study has also clinical implications as it combines the benefits of hemifield stimulation, with the different information contained in transient and steady-state response in less testing time and under the same conditions. Prolonged VEP testing leads to failure of concentration and fatigue which changes the evoked potential and results in intra-individual variability of the responses and inevitably in deteriorated and less accurate responses (Skuse and Burke, 1991). With the method proposed, a big part of the visual system can be assessed more accurately. It is well known that specific lesions of the visual pathways can be better detected with the one or the other stimulation rate. For example, glaucomatous field defects are related to damages of cells which preferentially respond to higher temporal frequencies. Moreover, the hemifield stimulation allows better detection of effects masked by full-field stimulation. Application of this stimulation technique in patients, will better allow to evaluate the usefulness of this method as a practical clinical tool.

References
[1] Barrett G., Blumhardt L.D., Halliday A.M., Halliday E. and Kriss A. (1976) A paradox in the lateralization of the visual evoked response. Nature(Lond.). 261: 253-255.
[2] Bezerianos A.G. and Papathanasopoulos P.G. (1991) A programmable digital video pattern generator controlled by microprocessor for basic research and clinical applications. Vis. Res. Vol. 31, No 9: 1563-1572.
[3] Bodis-Wollner I., Brannan J.R., Nicoll J., Frkovic S. and Mylin L.H. (1992) A short latency cortical component of the foveal VEP is revealed by hemifield stimulation. Electroencephalogr. Clin. Neurophysiol. 84: 201-208.
[4] Livingstone M.S. (1988) Art, illusion and the visual system. Sci. Am. 256: 78-85.
[5] Previc F.H. (1988) The neurophysiological significance of the N1 and P1 components of the Visual Evoked Potential. Clin. Vision Sci. Vol. 3 No 3, 195-202.
[6] Sherman S.M. (1985) Parallel W-, X-, and Y- cell pathways in the cat: a model for visual function. In Models of the Visual Cortex, edited by Rose D. and Dobson V.G., Wiley J. and Sons, London.
[7] Skrandies W. (1985) Scalp potential fields evoked by grating stimuli: effects of spatial frequency and orientation. Electroenceph. clin. Neurophysiol. 58: 325-332.
[8] Skuse N.F., and Burke D. (1992) Sequence dependent deterioration in the visual evoked potential in the absence of drowsiness. Electroencephalogr. Clin. Neurophysiol. 84: 20-25.

# A VECTOR NONLINEAR FILTERING APPROACH TO EVOKED POTENTIALS

by N.Laskaris, G. Economou, S. Fotopoulos and A. Bezerianos*
Department of Physics,Electronics Lab.,
*Faculty of Medicine, Medical Physics Lab.,
University of Patras, 261 10 Patras, Greece.

## Abstract
Vector filtering processes are of special importance when multichannel data are of interest. In this case the data sequence is treated as a whole. The idea of using vectors in place of scalars has been used extensively. This work presents a method where data depended coefficients are found at each time instant, upon which the filter output is computed. This nonlinear filter has properties in between the vector median and other nonlinear vector filters and to linear ones. This filtering process is very well suited and applied to the Visual Evoked Potential signals.

## Introduction
Nonlinear filters of the median or more generally of order statistic type, have been used as an alternative to linear ones when sharp transitions or heavy tailed noise characterize the input signal. The additional requirement for vector processing arises in highly correlated data. Such cases are TV colour images and different types of multichannel data [1, 2]. These cases can not be treated in a component-wise manner but vector techniques should be employed.
A vector valued signal $\underline{x}(n)$ with l components has the form:

$$x(n) = [x_1(n), x_2(n), \ldots x_l(n)] \tag{1}$$

In the vector filtering process $\underline{F}$ the output vector $\underline{y}(n)$ is found by applying $\underline{F}$ to $\underline{x}(n)$:

$$\underline{y}(n) = \underline{F}(\underline{x}(n)) \tag{2}$$

This process is different from the corresponding component-wise version where F is applied to each component $x_i(n)$ separately:

$$\underline{y}(n) = [F(x_1(n)), F(x_2(n)), \ldots F(x_l(n))] \tag{3}$$

In (2) the correlation among the l components is taken into account and is utilized. This is the main difference between the two processes described by (2) and (3).
The Vector Median Filters (VMF) [1,3] has been recently proposed and is proven an efficient vector filtering process. The Vector Nonlinear Filters (VNF) presented in this work try to retain most of the properties of VMF while reducing their complexity. It also shares some of the properties of Linear filters.
Other approaches [5,6] utilizing the vector filtering concept have also been proposed. e.g. the Vector Directional Filters [5] have been recently presented and applied to color image processing.

## The method
As a measure of the distance between the vector $\underline{x}_i$ and the other vectors $\underline{x}_j$ of the set we use

the following :

$$D_i = \sum_{j=1}^{N} \| \vec{x}_i - \vec{x}_j \|$$

(4)

The norm used in (1) could be the $L_1$ or $L_2$. i.e the summation of the absolute value of the components or the Euclidean distance.

$$|x| = \sum_{i=1}^{l} |x_i|$$

(5)

$$|x| = \sqrt{\sum_{i=1}^{l} x_i^2}$$

Thus given a set of N vector value signals $\{x_i, i = 1,..N,\}$ the N scalars $D_i$ are a measure of their mutual distances. It is important that to each vector $x_i$ a single scalar $D_i$ is attached corresponding to the distance of the vector $x_i$ to the others. The following two observations summarize the results coming out from the maximum likelihood estimate approach.

a)The particular $x_i$ for which $D_i$ is minimum has been defined [2,3] as the Vector Median of the vector set $\{x_i, i = 1,..N.\}$

b)The distance $D_i$ is utilized as a measure of the similarity or closeness of $x_i$ to the other vectors. In that sense an ordering relation is emerging.

This relation is further exploited by defining a coefficient $C_i$ as follows:

$$C_i = \frac{\dfrac{1}{D_i^2}}{\sum_{j=1}^{N} \dfrac{1}{D_j^2}}$$

(6)

The above defining formula is becoming from (4) by inversion and is normalized for unit amplification in DC level. Several other similar relations could be employed. This inversion of $D_i$ guarantees that vectors close to the Vector Median are attached higher values than those with longer distances $D_i$. In that sense the VNF could be considered as an imitation of the Vector Median Filter, like its scalar counterpart [4]. Or in a more general view it is very close to the least error estimate of location which in this case is the vector median. It should be noticed also at this point that the VNF concept could be generalized for other parameter estimation processes. e.g the angle location estimation is one of them.

The overall VNF output is easily found as:

$$\vec{y} = \sum_{j=1}^{N} C_j \vec{x}_j$$

(7)

The above equation declares the linear character of the VNF.

For a convenient visualization of the VNF concept as well as for a rough comparison to the other Vector filtering processes we present an example in the two dimensional space. The three input vectors numbered 1,2,3 are filtered by VMF(No 2) and VNF (No 4). The averaging vector (No5) is also given. As it is seen the VNF is very close to VMF and differs significantly from the vector averager.

## Application to VEP signals

The proposed method has been found suitable for several application areas. The Visual Evoked Potentials (VEP) signals have been selected here. These are EEG signals produced by a sequence of visual stimulus and overlaid to the usual brain activity. The set of data consists of sequences, each one corresponding to a single response evoked by the visual stimulus and are strongly correlated. Five of these responses, each one consisted of 250 sample points, are given in figure2. Several methods for enhancement of the signal have been proposed the simplest and most effective one being the averager. The VNF is suitable for processing of VEP signals when each data sequence of figure2 is treated as a vector 1x250. It should be noticed that the signal No 3 which is an artifact, differs substantially from the others and could be considered as an outlier in the set of 5 vectors. Such signals deform the output if the averager is employed as the filtering process. Removing the outliers improves the process but it implies an ordering relation before averaging and somehow increases the complexity of the algorithm. Use of discriminant analysis could also be employed. VNF overcomes the problem of outliers by attaching a very small coefficient to them. For that specific case the c(3) coefficient shown in the above figure (right) is extremely small. Using a set of 30 such vectors of raw VEP data, we receive the output which is presented in figure3 (curve 3). In the same figure the average of the whole set of 110 such single sweeps that is considered as the "ideal" signal is also presented (curve 1). As it is seen in this figure the VNF is very close to the "ideal" output although the number of vector samples (30) is much smaller than those used for the average (110). The average of the same 30 signals after removing outliers is also shown (curve2). The average of the same set of 30 signals is also given in curve4 for comparison.

## Evaluation of the method

The previously used data are also utilized for a quantified evaluation of VNF. 110 single sweeps i.e. VEP signals, are the set of 110 input vectors and the VNF is the output we are seeking. The Integrated Mean Squared Error (IMSE) as it is described in [7] is used for evaluation of our method regarding robustness and for comparison to the average before and after removing outliers.

$$IMSE(\hat{x}_N) = \frac{1}{T}\sum_{n=1}^{T} [\hat{x}_N(n) - \bar{x}_{N_0}(n)]^2 - \frac{P(\epsilon)}{N_0}$$

$$where: \qquad N=4,6\ldots N_1$$

$$P(\epsilon) = noise\ power = \frac{1}{(N_0-1)T}\sum_{i=1}^{N_0}\sum_{n=1}^{T} [x_i(n) - \bar{x}_{N_0}(n)]^2 \tag{8}$$

$$\bar{x}_{N_0} = the\ average\ of\ N_0\ signals$$

$$\hat{x}_N\ the\ estimator\ of\ N\ signals\ i.e.\ \bar{y}\ for\ VNF\ (7)$$

Application of the above formula is achieved in the following steps:
a) The set of 110 sweeps has to be divided into a criterion set of $N_0 = 80$ sweeps and into an evaluation set of $N_1 = 30$ sweeps, in a disjoint way.
b) Signal averaging of $N_0$ provides us with the "true" curve
More specifically in our case N is given the values 5, 10, 20 and 30. Also T stands for the total number of sample points (250).
Computations yield the following results expressed in **percentage of decrease of IPSE in respect to ensemble average** and given in TABLE I (second column) for the four different values of N

(first column). For comparison purposes the average after rejecting the outliers is also given in the same table (third column) .

In all four cases the VNF is superior to the averager and very close to the case where outliers have been removed. When the number of data is increased the behaviour of VNF is improving, indicating the robustness of the estimator.

**TABLE I**

| N | VNF | after rejection of outliers |
|---|---|---|
| 5 | 65% | 67% |
| 10 | 71% | 76% |
| 20 | 80% | 88% |
| 30 | 83% | 90% |

## Conclusions

The nonlinear algorithm presented here is derived by using the multidimensional maximum likelihood estimate approach. The relative "distance" between each vector and the others is used to define the coefficient of the reference vector point. The VNF does not need any comparisons or swap operations and their complexity does not increase very much with the filter order i.e with the number of vectors. Application of the VNF to VEP signals improve the output by suppressing very much the artifacts and smoothing gaussian noise. Application to other multidimensional data seems promising.

## References

[1]    J. Astola, P.Haavisto and Y.Neuvo "Vector Median Filters" Proc. of the IEEE, Vol. 78, No.4, April 1990

[2]    N.P. Galatsanos and R.T. Chin, "Digital restoration of multichannel images," IEEE Trans. Acoust., Speech and Signal Processing, Vol. ASSP-37, pp.415-421, Mar.1989.

[3]    C.A. Pomalaza-Raez and Y.Fong,"Estimation of the location parameter of multispectral distribution by a median operation", in Proc. 11th Int. Symp. on Machine Processing of remotely Sensed Data, West Lafayette, IN, pp.41-48, 1985

[4]    S. Fotopoulos and G. Economou, "Median approximating algorithm" , Electronics letters, 1st August 1991, Vol.27 No.16 pp.1441-1442

[5]    P.E. Trahanias and A.N.Venetsanopoulos, "Vector Directional Filters for Color Image Processing", Proceedings of ECCTD '93, Circuit Theory and Design, 1993 Elsevier Sc. Publishers, pp.1645-1650

[6]    R.Machura and K.Phillips, "Applications of vector fields to image processing", IEEE Trans. Pattern Anal. Mach. Intell., vol.PAMI-5,pp.316-329

[7]    T. Gasser, J.Mocks and W. Kohler, "Amplitude Probability Distribution of Noise for Flash-Evoked Potentials and Robust Response Estimates", IEEE Trans. on Biomedical Engineering, Vol. BME-33, No.6, June 1986.
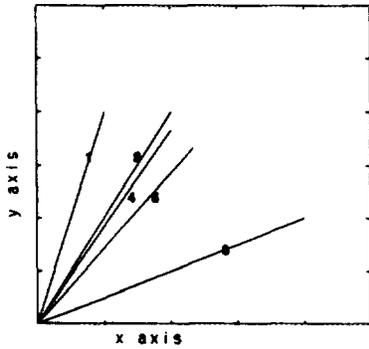
**Figure1.**
No 1,2,3:The three input vectors. No4:VNF output. No5:Vector average. No2 is the VMF output



$c(5)=.2575$

$c(4)=.2583$

$c(3)=.0472$

$c(2)=.2290$

$c(1)=.2083$
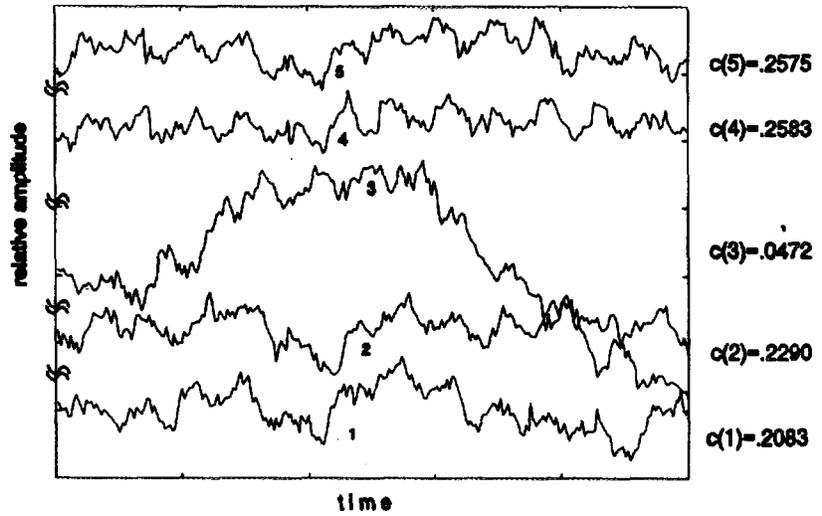
**Figure 2**
A sample of 5 vectors (VEP). The no3 which differs substantially from the others has the smallest coefficient.
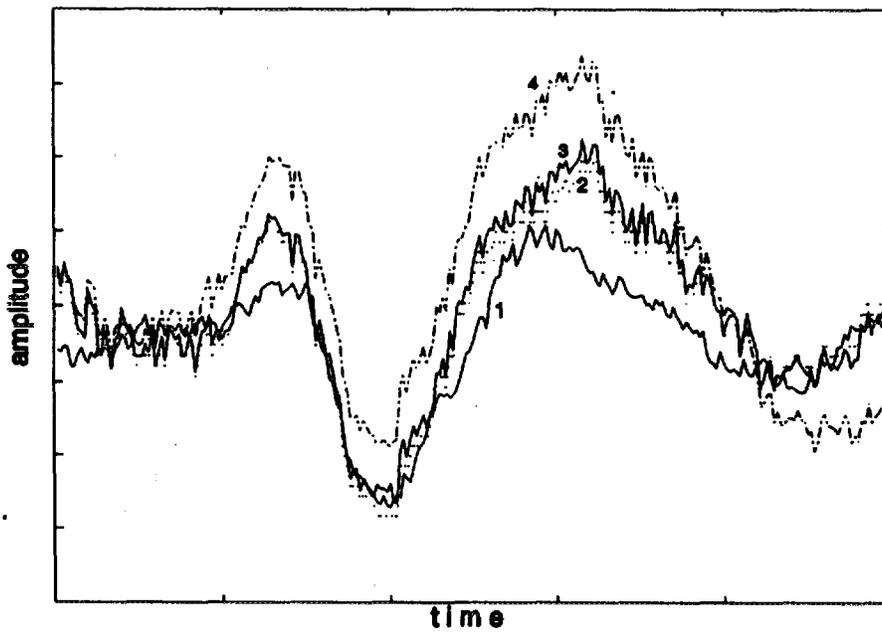


**Figure 3**
Curve1: The average of 80 vectors ("ideal")
curve2: The average of 30 vectors after removing the outliers
curve3: VNF
curve4: The average of 30 vectors

# Automated Photochromic Tracer Flow Extraction

D. Androutsos[†], P.E. Trahanias[‡] and A.N. Venetsanopoulos[†]

[†]Department of Electrical & Computer Engineering
University of Toronto
Toronto, Ontario, Canada M5S 1A4

[‡]Department of Computer Science
University of Crete
714 09 Heraklion, Crete, Greece

## Abstract

An automated method that accurately extracts trace flows from images obtained with the photochromic tracer technique is presented in this paper. This technique is used in studies of pulsatile flow fields created by modeled arterial stenoses [1,2]. The method involves three steps: image restoration, flow trace extraction, and automation of the visualization procedure. Image restoration is accomplished using nonlinear filtering operations. Active contours (snakes) are then applied to accurately extract flow traces. Snakes are energy-minimizing splines which accurately conform to inherent image forces. To achieve automation we rely on the fact that the shape of one flow trace is based on the previous trace and in turn acts as a basis for successive traces. Experimental results obtained from real data demonstrate the accuracy and robustness of the method. Consequently, this method can be employed instead of the more tedious and error prone manual determination of the flow traces.

## 1 Introduction

Photochromic tracer flow visualization is a valuable technique which has been implemented for in vitro studies of atherogenesis, the hardening of arteries from plaque and fatty deposits [1,2]. It has been used to study the downstream flow field of various levels of arterial constriction and disease progression. Furthermore, it is used for extensive study of axial wall shear stress downstream of stenoses which cause turbulence. The details of the experimental set-up of this technique and accompanying results can be found in [1]. For our purposes it suffices to mention that the flow traces are photographed by a 35 mm SLR camera, photographically enlarged and then digitized to produce standard grey-level digital images which are then used for analysis. Such an image is shown in Fig.1(a).

The study of the axial wall shear stress is conducted by analyzing the behavior of the trace flows close to the wall region. Presently, manual curve-fitting via B-splines is done to obtain a best-fit representation of the flow traces close to the wall region. This curve is then used to obtain the angle between the flow traces and the wall for calculation of shear stress [2]. This method, although conceptually simple, is very time consuming and subject to error. Furthermore, no attempt is made to use inherent image information to aid in the extraction of the trace flows. Also, film-grain noise which is prevalent in these images, is left unfiltered.

The purpose of this paper is to introduce an automated method that has been developed to aid in the analysis of pulsatile flow via the photochromic tracer flow visualization technique. This method employs standard techniques from the image processing and vision fields as well as specific techniques tailored to the images at hand. The technical aspects of this method are
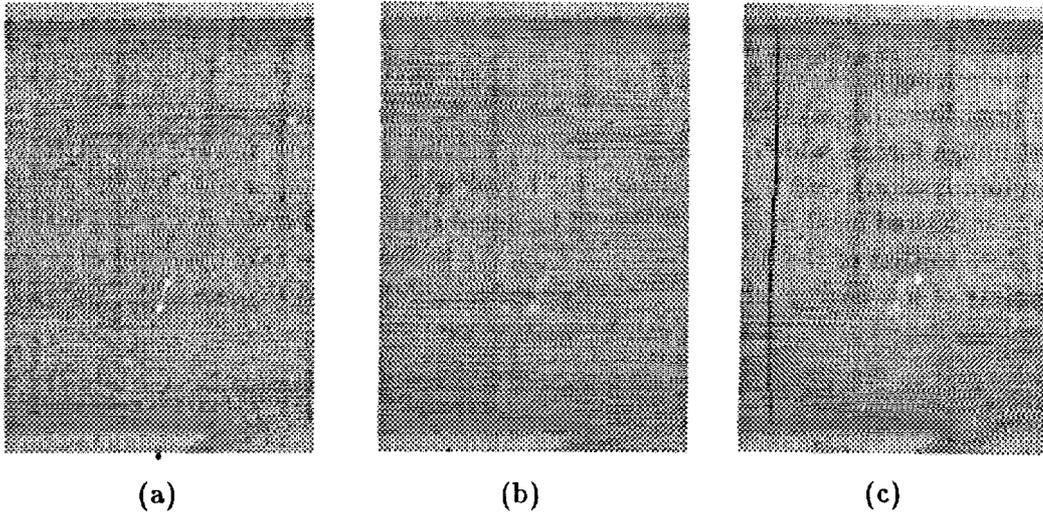
Figure 1: Image obtained with the photochromic tracer flow visualization technique. (a) Initial (digitized) image, (b) filtered image, (c) determination of the first trace flow in (b) using a snake.

analyzed in greater extent in [3], whereas, this paper focuses more on its applications in the analysis of the aforementioned images.

In what follows, image restoration as applied to the photochromic tracer images is first presented in Section 2. The determinations of trace flows using active contours is explained in Section 3 and the method of automation is described in Section 4. Finally, experimental results are presented and discussed in Section 5.

## 2 Image Restoration

In our case the prevailing degradation is film-grain noise [8]. It is primarily a result of the random location, shape and size of the silver halide grains present in photographic film. Film-grain noise can be regarded as a two-dimensional Gaussian random process. Furthermore, it can be modeled as a signal dependent multiplicative noise:

$$I(x,y) = S(x,y) + \alpha S(x,y)^{\beta} N(x,y) \tag{1}$$

where $I(x,y)$ is the observed noisy image, $S(x,y)$ is the noise-free image, $N(x,y)$ is a two-dimensional Gaussian process with zero mean and unit variance, and $\alpha$, $\beta$ are constants [9].

To reduce this degradation, a nonlinear mean filter was employed, namely the Geometric Mean [10]. It is a homomorphic filter which nonlinearly transforms multiplicative signal dependent noise, such as film-grain noise, into an additive space where the noise is uncoupled from the signal, allowing the result to be treated as additive noise. With a window size of $(2N+1) \times (2N+1)$, the filtered image is expressed as:

$$I_G(x,y) = 10^{\left[\frac{1}{(2N+1)^2} \sum_{k=-N}^{N} \sum_{l=-N}^{N} \log I(x+k,y+l)\right]} \tag{2}$$

where $p = 1 - \beta$. Overall, the geometric mean filter performs well in removing film-grain noise [11]. Furthermore, it tends to preserve edges in the images, which, in our case, is of great importance in localizing the flow traces. The application of the Geometric Mean filter to the image of Fig.1(a) is shown in Fig.1(b). As can be observed, most of the degradation due to the film grain noise has been removed.

# 3 Trace Flow Extraction

Active contours (*snakes*) are used to perform the trace flow extraction task. Snakes are dynamic energy-minimizing splines influenced by image forces that move them towards inherent image features (e.g. lines, edges) where they accurately conform [4]. Snakes rely on two types of forces: low-level image forces, which are responsible for moving the snake toward salient features (i.e energy minimization), and high-level external forces which are responsible for the placement of a snake near desired local minima. It should be noted that initial snake shape and location need only be close to that of the desired minima since the image forces take over and draw the snake to the points of low energy which ultimately determines the final shape and location.

The energy functional of a snake can be written as:

$$E_{snake} = \sum_{s=0}^{n} E_{int}(x(s), y(s)) + E_{image}(x(s), y(s)) + E_{ext}(x(s), y(s)) \qquad (3)$$

where $E_{int}$ represents internal snake energy, $E_{image}$ the image forces, $E_{ext}$ the external forces and $n$ the number of snake nodes. Minimization of this functional coincides with the localization of a salient feature in the neighborhood of the initial snake location. Additional details on snakes and the minimization procedure, which is performed iteratively, can be found in the literature [4,5]. The snake approximation of the first trace flow of Fig.1(b) is illustrated in Fig.1(c), where minimization of Eq.(3) has been obtained after 75 iterations.

# 4 Automation

The automation of the detection of the flow traces is needed in order to make the method completely independent of user intervention and increase efficiency and productivity and reduce error. The method of automation is based on the fact that the flow traces obtained via the photochromic technique are essentially interdependent samples of the continuously flowing liquid. This provides a premise for a knowledge-based system since we know that the shape of one flow trace is based on the previous one and this in turn acts as a basis for successive traces. Also, we make use of some a priori information available from the images. We note that the walls of the tube are the outermost significant characteristics in the vertical direction and that all the flow traces exist between the walls and have a generally vertical shape, especially the traces at the left end of the tube.

Now, with this information we proceed to find the traces. First, the top and bottom walls are detected by placing snakes close to them to correctly localize them. Next, an approximate location of the first flow trace is found at the left end of the tube by searching for the first strong vertical edge between the walls. We use this position as an initial placement for a snake which upon application minimizes its energy functional and correctly localizes the trace.

Next, the shape of the detected trace is used as a template to find an approximate location for the next trace. To accomplish this we move in a horizontal direction to the right and calculate the cross-correlation [6,7] of the most recently localized trace and the new position. This calculation is repeated until a position with a maximum value, representing a *best match* is found. A snake, with initial shape that of the previously localized trace, is placed at the *best match* position where it also proceeds to minimize its energy functional and localize the new flow trace. This assures that the snake will be placed in a close vicinity to the desired minima (i.e. flow trace) and since the shape of two successive traces does not vary greatly, computing time is significantly reduced.

40

# 5 Experimental Results

In order to assess the proposed method, a software system that implements it has been developed and evaluated using real images provided to us by the University of Toronto, Biomedical Engineering Laboratory [3]. These images have been obtained using the technique mentioned in the introduction and present various trace flow patterns typically encountered in photochromic tracer images. The evaluation procedure employed was as follows: each image has been independently processed manually and by the system, and the results were then compared for discrepancies. Other than the visual appearance of the snakes placed on the images, the measured angles between the flow traces and the wall were compared. This is illustrated in Fig.2; Fig.2(a) shows a flow trace and in Fig.2(b) the placement of the snake on it, along with the



(a)                                                 (b)
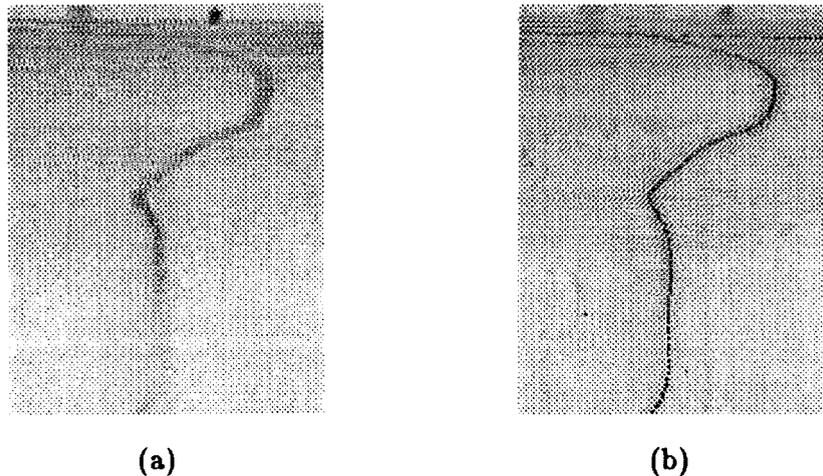
Figure 2: Calculation of angle between flow traces and the wall. (a) A flow trace pattern, (b) snake placed on the pattern of (a) and the wall. The angle is calculated at the intersection of the two snakes.

detection of the wall, is presented. The angle is computed at the point of intersection between these two curves, using information as of their tangents. In summary, all the experiments concerning angle measurements have been very accurate and, consequently, they can safely be used for subsequent analysis (e.g. calculation of shear stress).

A complete result obtained from the application of the proposed method to a typical image is shown in Fig.3. The original, unprocessed image, is shown in Fig.3(a) and our result in Fig.3(b). As can be verified, the flow traces are correctly localized by the active contours and the assertion on the continuity of the flow traces can also be verified from this figure. Thus, the method can be feasibly employed for real data analysis instead of the currently used manual procedure.

# References

1. M. Ohja, R. Hummel, R.S.C. Cobbold, and K. Johnston, "Development and Evaluation of a High Resolution Photochromic Dye Tracer Method for Pulsatile Flow Studies," *J. Phys. E: Sci. Instr.*, vol. 21, pp. 998-1004, 1988.
2. M. Ohja, R.S.C. Cobbold, and R. Hummel, "Photochromic Tracer Visualization of Pulsatile Flow Through Modeled Arterial Stenoses," *FED*, vol. 85, pp. 49-54, 1989.
3. D. Androutsos, P.E. Trahanias and A.N. Venetsanopoulos, "Feature Extraction in Photochromic Tracer Flow Visualization Using Active Contours," submitted for presentation, European Signal Processing Conference '94, Edinburgh, UK, 13-16 Sept. 1994.
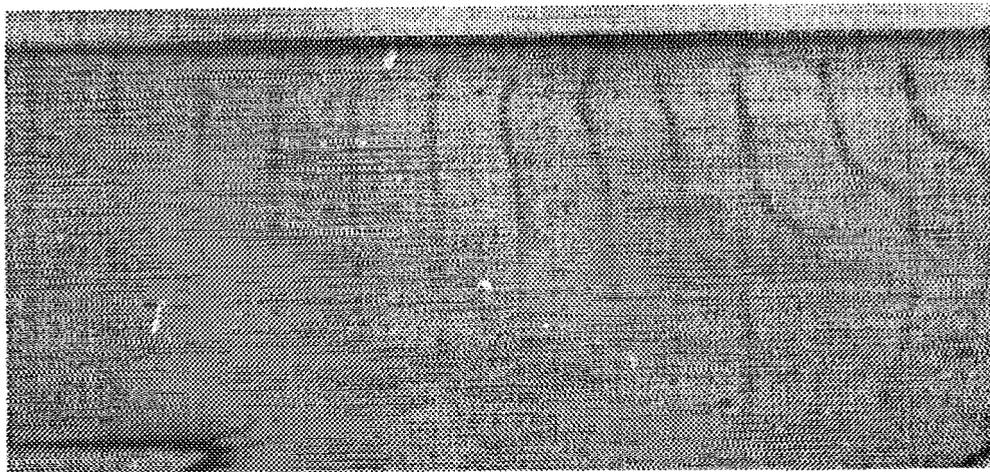
(a)
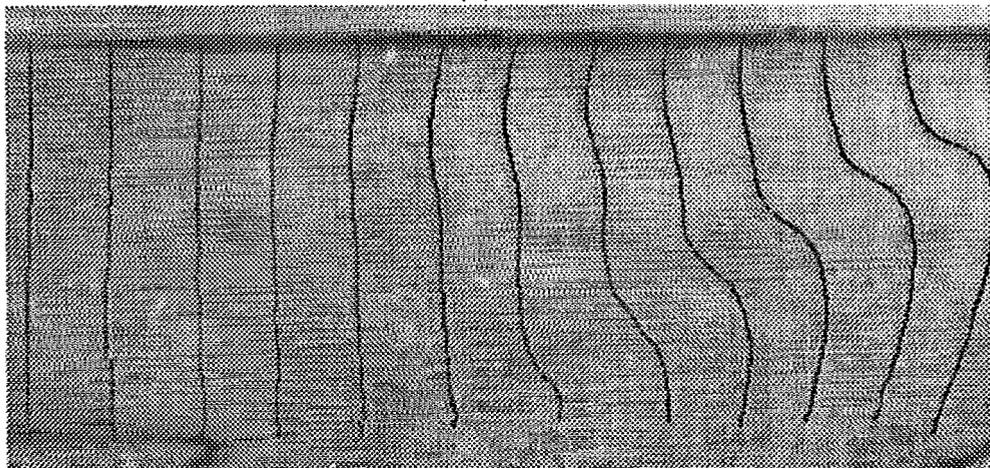


(b)

Figure 3: Image obtained with the photochromic tracer flow visualization technique. (a) Initial (unprocessed) image, (b) trace flow detection results.

4. M. Kass, A. Witkin and D. Terzopoulos, "Snakes: Active Contour Models," *International Journal of Computer Vision*, pp. 321-331, 1988.

5. D. Terzopoulos, "On Matching Deformable Models To Images," *Topical Meeting on Machine Vision, Technical Digest Series*, vol. 12, pp. 160-163, 1987.

6. A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 2nd ed. McGraw Hill, New York, 1984.

7. B. Jähne, *Digital image processing: concepts, algorithms, and scientific applications*, Berlin, 1991,

8. H.C. Andrews and B.R. Hunt, *Digital Image Restoration*. Prentice-Hall, 1977.

9. H.H. Arsenault and M. Denis, "Image processing in signal-dependent noise," *Canadian Journal of Physics*, vol. 61, pp. 309-317, 1983.

10. I. Pitas and A.N. Venetsanopoulos, *Nonlinear digital filters: principles and applications*, Kluwer Academic, 1990.

11. I. Pitas and A.N. Venetsanopoulos, "Order statistics in digital image processing," *IEEE Proceedings*, Dec. 1992.

# ON THE ERROR INDUCED BY THE ACQUISITION OF THE NEEDLE ELECTROMYOGRAM THROUGH ITS ESSENTIAL POINTS

**Mihai Tarata**
**Informatica Medicala, Faculty of Medicine, The University of Craiova, Romania**
**Bul. Maresal Antonescu 62, Craiova, RO-1100, Romania**

*To overcome or at least to alleviate the constraints implied by the classic acquisition of the needle electromyogram ( large memory, long processing time, high frequency noise, high sampling rate ) I have developed a method and an acquisition interface. According to this method, the signal is analogically preprocessed and acquired through its essential points ( zero crossings and turning points ), the result being an approximation of the signal, namely the polygonal line thus defined. The analogic processing preceeding the acquisition, uses low pass differentiators to provide front end filtering to enhance the signal/noise ratio and to generate the first and second derivative, which are subsequently used to detect the essential points of the signal. I have analyzed the original signal ( needle EMG activity from the first interosseus dorsalis muscle), acquired at a rate of 20 μs/sample, and its polygonal approximation, in the time and frequency domains, with the intention to estimate the error induced by such a truncation. In the following, the paper presents the acquisition method and interface, details and results on the error analysis, conclusions. The results have confirmed the method and the interface as valuable tools for a more efficient acquisition of the EMG signal.*

## Introduction

As the needle EMG has specific high frequency components, a large amount of memory is required due to the high acquisition rate which is highly necessary, and a long time is needed to process the data; these are the constraints of a continuous acquisition of the EMG signal. To overcome these constraints a lower acquisition rate might be used, but a major problem caused by a moderate rate sampling is the lack of information (Fig.1) resulting from the possibility of missing the peaks. An analog / digital interface has been developed, with new modes of acquisition, to accelerate the overall process.

## The Method. The acquisition of the EMG signal through its essential points

To avoid the above disadvantages without loosing information , the essential points of the signal have been choosen to define the signal i.e. the peaks (turning points) and the zero crossings.
The basic idea (Nestianu, Tarata, 1981) has been to assimilate the signal defined by its essential points , the crossings of the isoelectric line and the turning points, to a polygonal line (Fig.2) and to sample and store only the amplitudes and time of occurence of these points, instead of a continuous sampling. Due to the preservation of the time relationships within the signal and to the unsignificant loss of signal energy, the approximation holds, and the method results in avoiding the mentioned constraints.
According to the above, a new acquisition structure and interface have been developed. The interface has specific blocks to analogically preprocess the signal, to detect the essential points, to start conversion and time counting and to transfer the data in memory. Due to the qualities of lowpass differentiators to enhance the signal / noise ratio where large frequency

bandwidth noise is present (McGill et al., 1985), specific analog blocks are added , to generate the first and the second derivative of the original signal. The blocks are used

Different Sampling Rates EMG Acquisition



X 266 microsec.
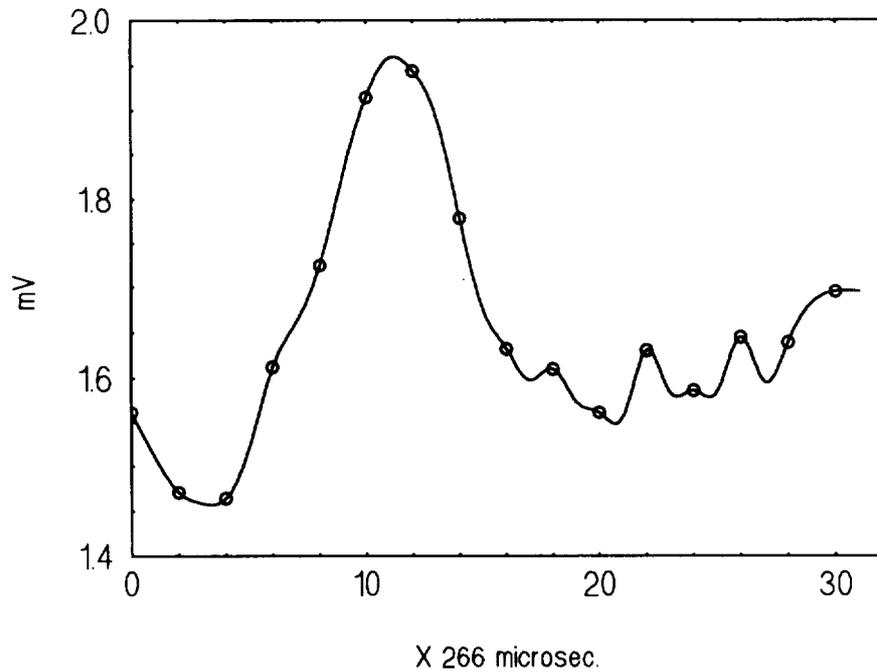
**Figure 1.** Constant rate sampling at low and high rates.
The peaks may be missed at a moderate rate.

to select the turning points and to generate the second derivative. As a major problem of the derivators is a high noise (Graeme,1972) , a special structure has been used to process the



X 266 microsec.
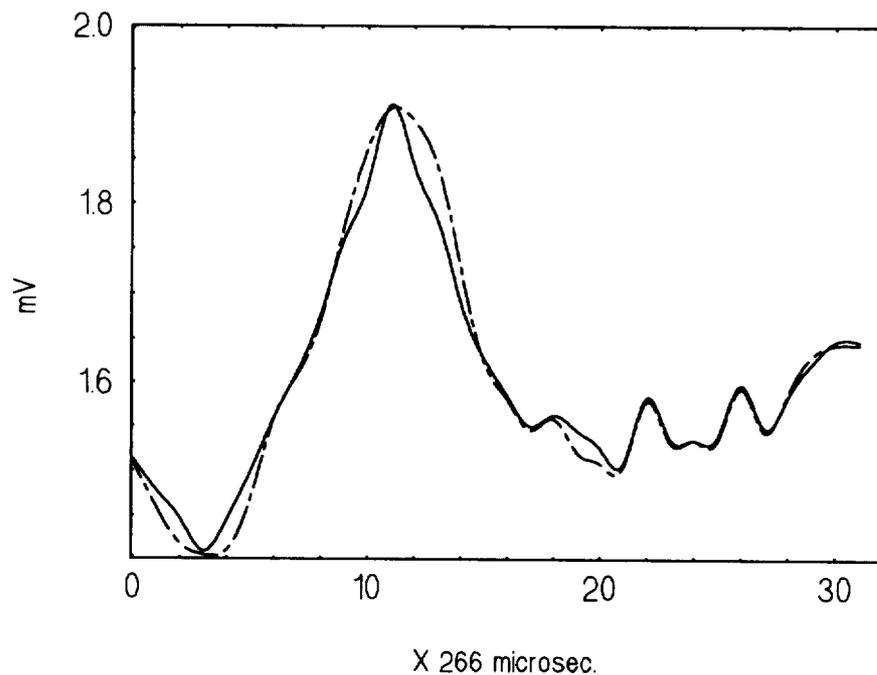
**Figure 2.** The approximation of the signal by its essential points

first derivative from the input signal and its integral. This results in a lower noise as the frequency increases, preserving in the same time the properties of a derivator.

A digital block of the interface provides the dialogue with the host computer (any PC) and runs a time counter which keeps the proper timing of every acquired sample. The interface may also continuously acquire the signal at a constant rate, as any classical A/D acquisition board.

The actual interface processes two distinct EMG channels and has offered complete satisfaction in that the processing time has shortened at least five times comparing to the method of plain acquisition and processing.

## Results. The error induced by the polygonal approximation of the EMG signal

Following the above acquisition method the EMG has been acquired with a needle monopolar electrode from the prime dorsalis interosseus muscle of a healthy subject, at a moderate steady contraction, the signal was amplified and prefiltered with an MG42 Medicor electromyograph and the data acquisition performed with the previously described interface card via a PC 386 microcomputer, continuously at a rate of 20 microsec./sample, and through its essential points . From the rough data 20 segments of 8 msec. each have been selected at random with the only condition to include action potentials ( EMG spikes); in Fig.2 an example is given, simultaneously showing the signal and its polygonal approximation .
To rebuild the signal from its essential points, two methods are possible, the hardware approach - a counter to rebuild the slope by integrating between each two consecutive essential points, or the software approach. The hardware approach is likely to be preferred, as it proceeds in real time.

Rebuilding the EMG signal thus acquired has been necessary for evaluation purposes, namely to appreciate the errors induced by such an approximation. An acceptable error may imply
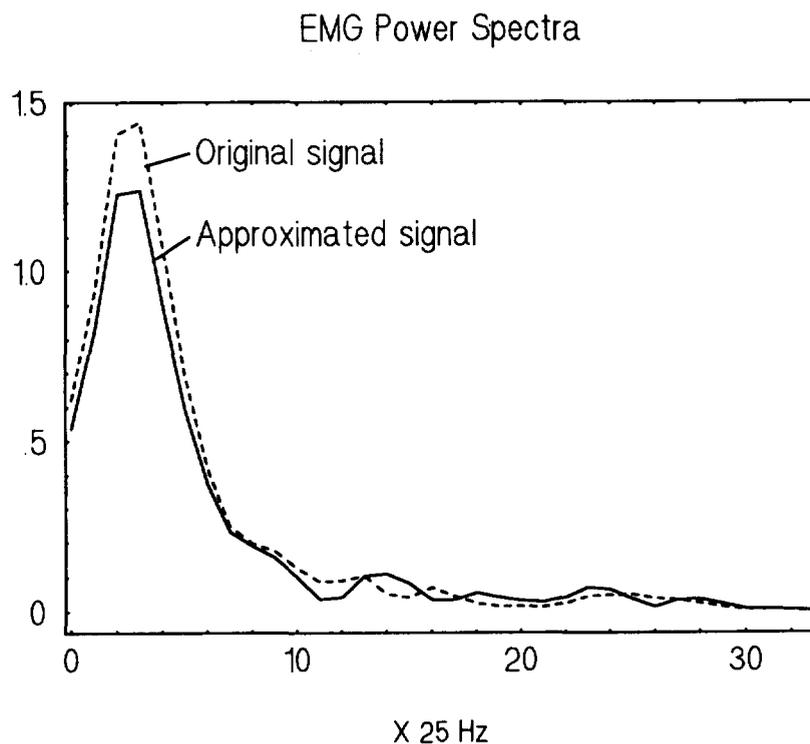


Figure 3. The power spectra

45

reconsidering the processing of the EMG with no reconstruction needed, with the major advantages above mentioned.

The method of acquiring the EMG through its essential points definitely preserves the time relationships within the signal , which is extremely important, because the major peak in e.g. the motor teritory action potential is related in time to the activity of the major part of muscular fibers within the motor territory.

As the time relationships are preserved, the second problem to be answered at, concerns the loss of energy caused by the approximation, comparing to the continuous acquisition. Consequently, for the two signals, the original and its approximation, the power spectra (Fig.3) have been computed via Hamming windowing and Fast Fourier Transforms.

Figure 4 displays the procentual difference between the two spectra for one segment of data, the detailed results laying in Table 1.

|  | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| correlation | .998 | .997 | .997 | .994 | .995 |
| % max. error | 17 | 15 | 17 | 17 | 16 |
| orig. fft terms | 12 | 12 | 10 | 16 | 10 |
| appr. fft terms | 15 | 12 | 13 | 11 | 12 |

**Table 1**

Finally the power spectrum has been integrated with the intention to assess which is the maximum order of the Fourier coefficients to reach 90% of the signal energy; the results are depicted in the last two rows of Table 1.

Normalized Spectra Difference
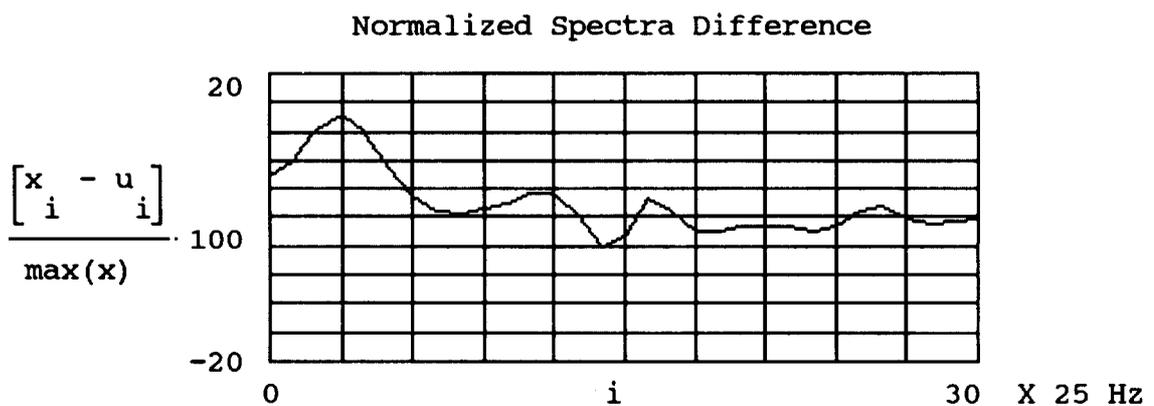


**Figure 4.** Spectral differences for one data segment

## Conclusions

The method of acquiring the EMG signal through its essential points preserves the time relationships within the waveforms. The average correlation coefficient for the signal and its approximation spectra has been minimum 0.995, which guarantees similar frequency composition. Considering the normalized difference between the power density spectra, it has

not exceeded 17 % at most centred on the spectral maximum; this difference is due mainly to a higher frequency content because of the angles at the peaks. A further low pass filtering would reduce the high frequency content, increasing in the same time the low frequency content, thus diminishing the difference. More, subtracting between 1 % and 5 % of the high frequency components from the power density spectrum of the approximated signal, the subsequently rebuilt signal and the original are practically identical, with a correlation coefficient of 0.999, which confirms that the approximated signal has a slightly higher content in higher frequencies components, which is comfortable to cancel through a low pass filtering in the frequency or time domains.

## References

Graeme JG (1972). *Applications of operational amplifiers. Third generation techniques.* McGraw Hill Book Co.

McGill KC, Cummins KL & dorfman LJ (1985). Automatic decomposition of the clinical electromyogram. *IEEE Transactions on Biomedical Engineering*, Vol. BME-32, No. 7, July 1985, 470-477

Nestianu V., Tarata M & Balsianu F (1981). Anew method for the quantitative analysis of the myoelectric spikes synchronization. *Neurology and Psychiatry* 4, Tome 19, Ed. Academiei, 283-288

Nestianu, V., Tarata, M., Ivanescu, M. - Metoda si aparat pentru Studiul cantitativ al electromiogramei, *Brevet OSIM 70610*, 1978

Nestianu, V., Tarata, M. - Metoda si circuit electronic pentru analiza cantitativa a electromiogramei, *Brevet OSIM 77109*, 1981

Tarata, M. - A New Architecture for an Efficient Acquisition of the EMG Signal, *The VIIth International Symposium on Motor Control*, Borovets, Bulgaria, June 21-25, 1993, Acta Physiologica Bulgarica, Vol.19, Nr. 1/2, 1993, p.52

# Reverse Engineering- and Rapid Prototyping Techniques in Medicine

**Bernhard Geiger**
Institut National de Recherche en
Informatique et Automatique (INRIA)
BP 93 - 06902 Sophia Antipolis
Cedex, France
E-Mail: geiger@sophia.inria.fr

**Marinos Ioannides**
University of Stuttgart
Institute of Control Technology for
Machine Tools and Manufacturing Units
(ISW 1.2)
Seidenstr. 36
D-70174 Stuttgart, Germany
E-Mail: jes@isww85.isw.uni-stuttgart.de

## Abstract

This Paper presents a new software functionality based on the recent Computer Aided Design, Computer Aided Manufacturing (CAD/CAM) and Rapid Prototyping technology , which has succeeded not only in visualising precise 3D-images of a human body on workstations, but also in producing resin prototypes and simulating a medical surgery.

A major problem in medical imaging is the three dimensional reconstruction from parallel cross sections (Reverse Engineering). Such 3D models are used for surgery planning, prosthesis milling, radiation therapy planning and volumetric measurements. In the preparation of complex bone surgery, it becomes a common practice to make solid models of bone structures. Surface generation, contour lines generation, triangulation and volume generation, will be explained for the different operation levels with respect to CAD/CAM-, Finite Element Systems (FEM) and stereolithography (Rapid Prototyping) /1-5/.

## Introduction

The utilization of the Computer Graphics technologies in order to visualize, analyze and reproduce human body is becoming crucial to the present medical field. 2D-Data which are generated by medical imaging systems, such as X-Ray Computerized Tomography (CT), Magnetic Resonance Imaging (MRI) and Ultrasonic Systems are often not comprehensive enough for the surgeons to analyze the conditions of their patients effectively /1,2,3/. To overcome the disandvantage we propose a solution to the reconstruction problem, the *Delaunay Reconstruction*, which is based on the Delaunay triangulation, a well studied data structure in computational geometry /6/.

## Three Dimensional Reconstruction by Delaunay-Triangulation

Our method produces a polyhedral model from a set of parallel contours. Unlike the voxel method /3/, we do not use equally shaped volume elements to fill the space between adjacent cross-sections, but tetrahedra that are adapted to the contour shape.

The Delaunay Reconstruction has the following properties:

- It gets directly to a 3D polyhedral representation composed of tetrahedra.
- The property of connecting contours on adjacent planes directly by triangles avoids the need for anti-aliasing or interpolation steps, especially for large cross-section distances.
- Unlike surface oriented methods, we are able to treat complicated objects presenting multiple contour branching and holes automatically (Fig.3 and Fig. 4).
- We achieve a considerable data reduction compared to other volume oriented methods (voxel). Real time display of reconstructed human organs is therefore possible on standard graphic workstations. This feature may be interesting for the design of models used in virtual reality, where rendering speed is crucial.
- The tetrahedral structure can be used for applications like simulation of motion or finite element methods.

## Surface Generation for CAD/CAM-Systems

The tooling and molding industry currently models the molding workpieces mathematically through CAD/CAM-Systems. Unfortunately the analytical description of such objects is not sufficient for practical applications. Here the space and time complexity must be increased to achieve the approximation and interpolation necessary. The computation performance of existing CAD/CAM-Systems is insufficient for the processing of such data. A data reduction is required to solve this problem. A data reduction of more than 80% can be achieved depending on the complexity of the workpiece by computing splines curves and spline surfaces and using the following algorithms:

- bicubic Bézier,

- polynomial representation (Coons),

- B-Spline,
- Non Uniform Rational B-Splines (NURBS) representation.

The output of this software functionality can be processed by all 2D/3D-oriented CAD/CAM or FEM-Systems /7,8/. For the data transfer within CAD/CAM-Systems the following standard interfaces are available: IGES, STEP, DXF and VDAFS /9/.
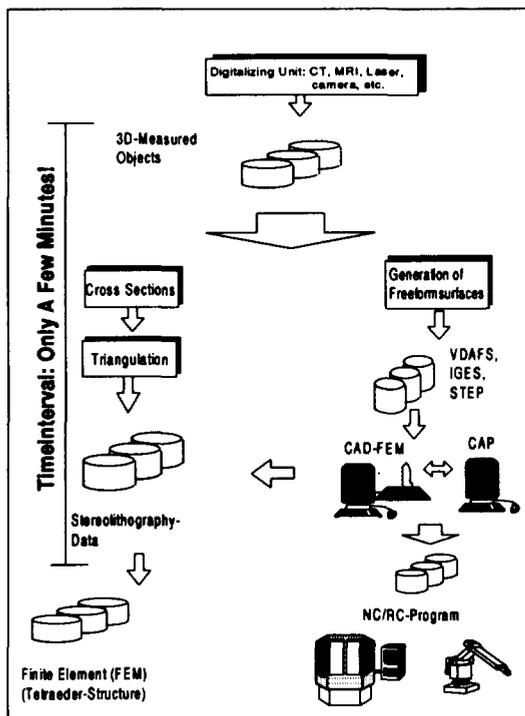
Fig. 1:  Rapid Prototyping, Reverse Engineering

## Impementation:

The entire software functionality (Fig. 1) is implemented in the programming language "C" and can be used on UNIX/X11 Motif Workstations or PC's. It is, at this time, an ideal prototype for the medicin, modell, design and milling industry /6-8,10/.

Fig. 2 shows a reconstructed skull from a CT Cadaver study (From the Chapel Hill Volume Rendering Test Data Set available at cs.unc.edu.). The contours on 113 cross-sections were extracted by simple thresholding, so that the process from CT-data to the 3D model was fully automatic. The model contains about 70000 surface triangles. Fig.3 and 4 shows a human

pelvis from a set of 23 MR images. The contours were extracted manually. In spite of the large cross-section distance (8mm vs. 1mm in-slice resolution), the model is relatively smooth. All of these objects have been processed using stereolithography.

The usefulness of the tetrahedra structure has been shown in a computer simulation of childbirth /10/.
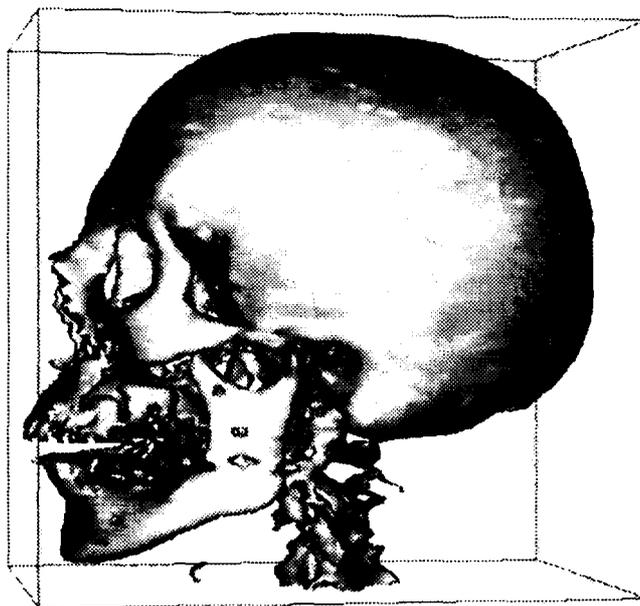


Fig. 2:     A skull from 113 CT images
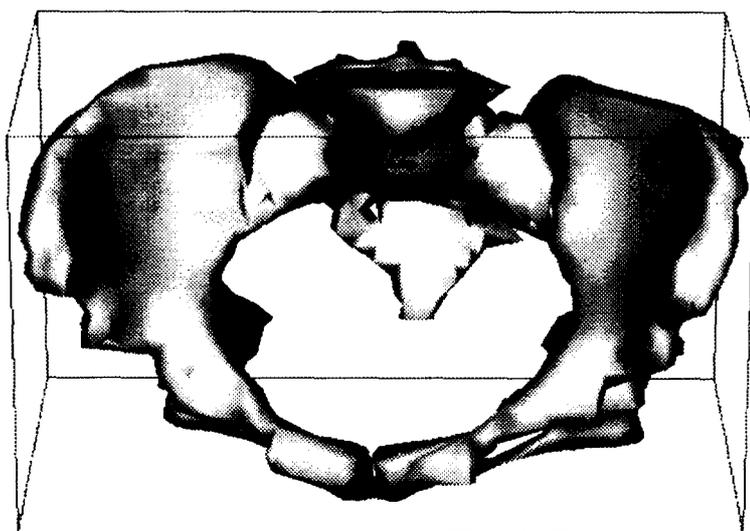


Fig 3:     A human pelvis from 23 MR images
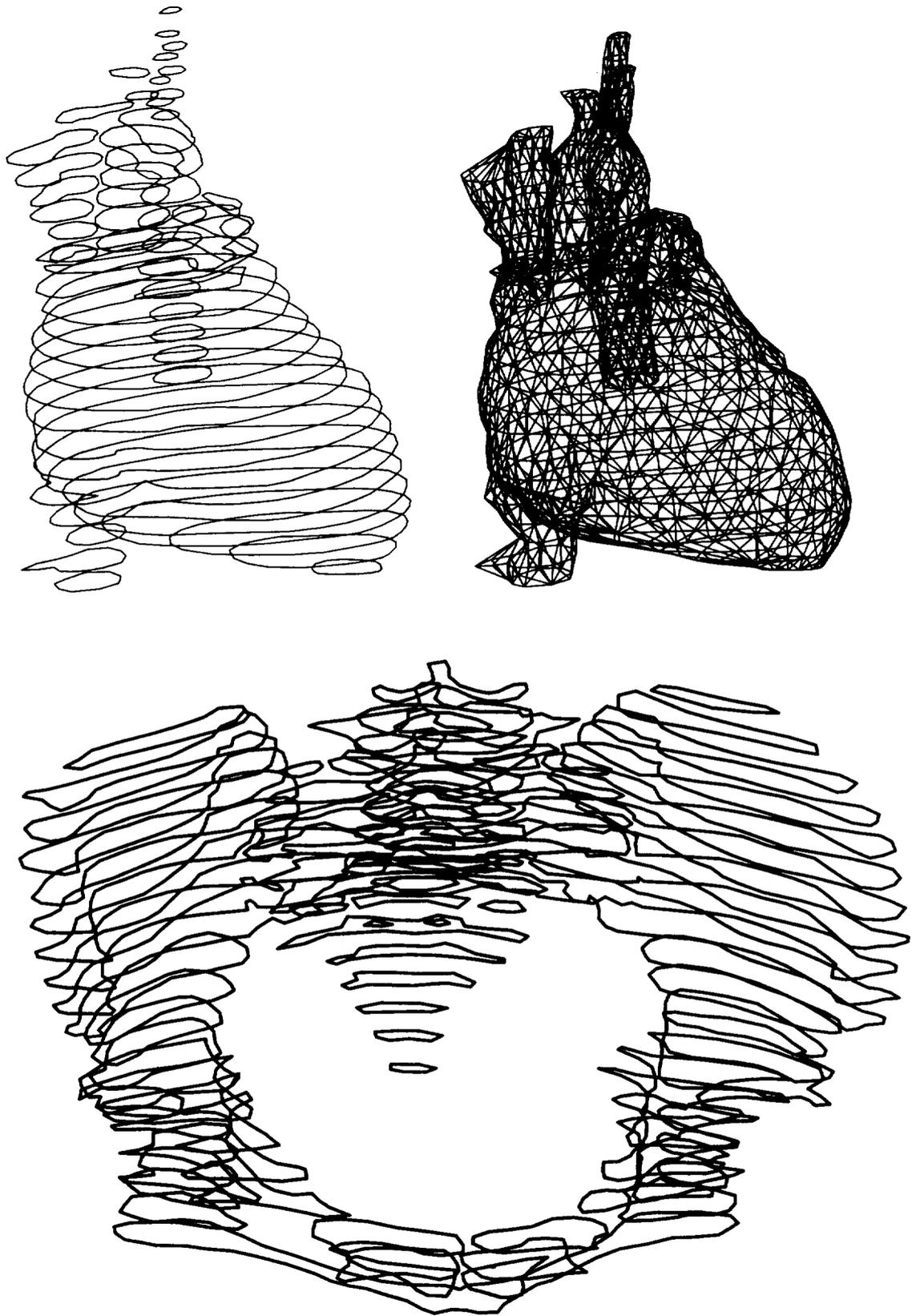
Fig. 4:    The generated contours of human heart and pelvis. The triangulated structure
(Volume-structure) of the heart is illustrated too.

# References

/1/   E. Vinarub, N. Kapoor:      Rapid Prototyping/Reverse Engineering: "New Manufacturing Tools for the 90´s". Third International Conference on Rapid Prototyping, University of Dayton, Ohio 1992. Pg. 231.

/2/   A. Jacob, et al.:      First Experience in the Use of Stereolithography in Medicine. Proceedings of the Fourth International Conference on Rapid Prototyping. University of Dayton, Ohio 1993. Pg. 121.

/3/   J. Adachi, et al.:      Surgical Simulation Using Rapid Prototyping. Proceedings of the Fourth International Conference on Rapid Prototyping. University of Dayton, Ohio 1993. Pg. 135.

/4/   K. Harding:      Overview of 3D Contouring Systems For Reverse Engineering Applications. Proceedings of the Fourth International Conference on Rapid Prototyping. University of Dayton, Ohio 1993. Pg. 143.

/5/   A. Breedveld:      State of the Art: Rapid Prototyping. EcCADENCE, Nov. 1992. Pg. 38.

/6/   B. Geiger:      Tree dimensional simulation of delivery for cephalopelvic dispropotion. First International workshop on mechatonics in medicine and surgery. Costa del Sol, 1992.

/7/   M. Ioannides:      3D Object Reconstruction Using 4D-Laser Digitalizing. Proceedings in Interfaces in Industrial Systems for Production and Engineering. Proceedings of the International Federation for Information Processing (IFIP)-Workshop, Darmstadt, Germany Elsevier Sc. Publ. B.V. (North-Holland) 1993, Pg. 245.

/8/   G. Pritschow, M. Ioannides, F. Krauß:      Schnelle und effiziente Digitalisierung, Modellierung und Fertigungsunterstützung verschiedenster Bearbeitungstechnologien. Proceedings of the First International User Congress in Solid Freeform Manufacturing (SFM), Dresden Germany, NC-Gesellschaft 1993.

/9/   H. Grabowsky, R. Anderl, X. Li, et al.:      Exchange of Freeform Surfaces using Standard Interfaces. Proceedings of the International Federation for Information Processing (IFIP)-Workshop, Darmstadt, Germany Elsevier Sc. Publ. B. V. (North-Holland) 1993, Pg.191

/10/   B. Geiger:      Three dimensional modeling of human organs and its application to diagnosis and surgical planning. Report 2105, 1994, INRIA Sophia-Antipolis, France.

# THE AVERAGE AREA / AMPLITUDE RATIO ( Raa ), A CONSISTENT PARAMETER IN THE QUANTITATIVE ANALYSIS OF THE ELECTROMYOGRAM

**Mihai Tarata**
**Informatica Medicala, Faculty of Medicine, The University of Craiova, Romania**
**Bul. Maresal Antonescu 62, Craiova, RO-1100, Romania**

*Within the effort to find more powerful descriptors for the study of the evolution of the electromyographic signal, mainly in the case of strong contractions which lead to interference in the normal muscle, I have defined the Average Area/Amplitude Ratio ( Raa ) parameter, which is computed from the EMG signal, and after the quantitative studies performed until now proves itself a consistent indicator. It is calculated as an average between consecutive zero crossings. The average of these ratios has been computed over the entire segment of data (400 msec.). Raa has been studied under strong maximum fatiguing contractions on the Biceps Brachii muscle, on a lot of 5 experiments on the same subject, with the double intention, to validate it as a tool for further investigations, and to effectively study the phenomenon of muscular fatigue, in order to assess Raa consistency, comparing with other known parameters.*

## Introduction

The present study started with the observation that the EMG signal has a great amplitude variability, strongly modulated by the distance between the recording electrode and the nearest muscular fiber or, in a larger sense, the nearest motor teritory. Thorough research on concentric needle EMG motor unit action potentials - MUAP - (Nandedkar, 1988) proved that the amplitude and the area change in parallel and as a conclusion, the ratio of MUAP area and amplitude, which measures the 'thickness' of the waveform is much less sensitive to the modifications due to changes in the electrode position. More, as the area and the amplitude respectively are robust features, much less dependent on the human operator skill, their ratio seems to be a consistent parameter comparing to the duration of the potential. The above mentioned study relates a 16% increase in the are/amplitude ratio comparing with much greater changes in both the area and amplitude alone, in concentric needle recordings from Biceps Brachii. More, this ratio proved to be significantly smaller in myopathies than in neuropathies, this assessing a diagnostic value to this parameter.

## The global Area/Amplitude Ratio Raa

I extended the idea to the interference EMG acquired with surface electrodes, and generated global area/amplitude ratio Raa, computed on a 400 msec. segment of EMG, as follows:

$$Raa = S\ med\ /\ A\ med$$

where:
**S med** - the sum of the rectified signal phases areas divided by the total number of phases ( phase - the signal between successive zero line crossings);

**A med** - the sum of the rectified signal phases maximum amplitudes divided by the total number of phases

As concerning Raa, it is obviously calculated as the ratio between the two sums, with a dimension of time.

## The method

Succesive recordings have been performed with surface electrodes on a healthy subject on two channels, with the electrodes placed on the Biceps Brachii muscle, under a strong contraction with a constant resistant force, sustained until a maximum fatiguing sensation was reported.

The signals were amplified and filtered ( 50 Hz, 20 KHz) through a Medicor MG-42 electromyograph and acquired at a rate of 200 $\mu$sec./sample via a DAP acquisition card and a 386 PC in successive 400 msec. data segments repeated each 5 sec.

All the data segments have been processed and for each, a series of specific parameters were calculated: the average value of the rectified signal (Vmr), the root mean square of the rectified signal (Rms), the integral of the rectified signal (Isr), the number of turning points ( Npi), the number of zero crossings (Ntz), the area/amplitude ratio (Raa) and its standard deviation (Dsraa) - in the time domain , and the median (Fmed) and average (Fav) frequencies, the frequency corresponding to the spectral maximum ( Fsmax), and the spectral maximum ( Smax) - in the frequency domain, respectively . Finally, the successive values of these parameters, each corresponding to a data segment, describe the evolution of the phenomenon in time.

It is to be noticed here that the evolutions traced for these parameters are similar to those described until now, so our study is a further confirmation.
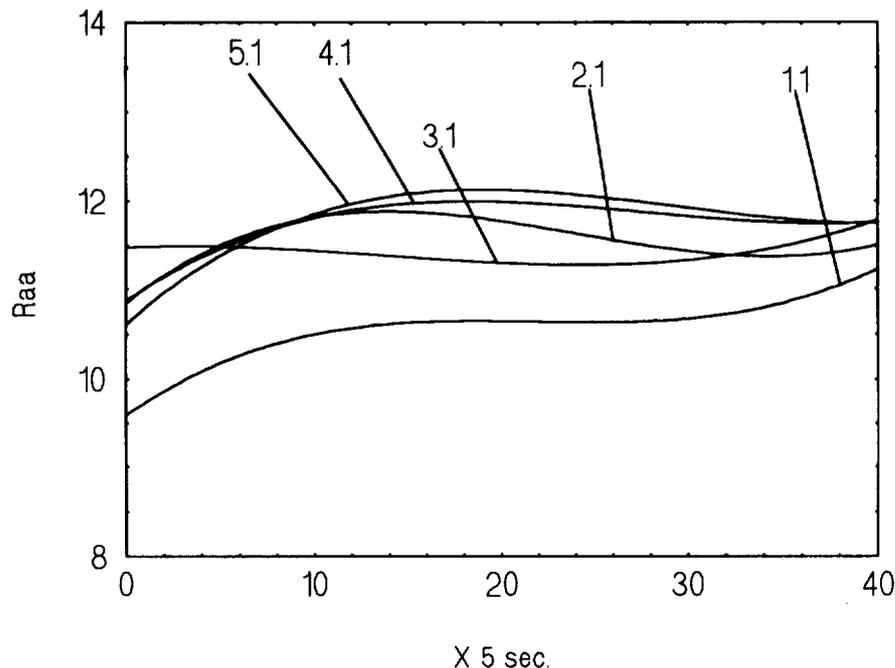


**Figure 1.** Average Raa evolution

A special attention has been focused on the Raa parameter which is of a special interest to us, so the mean rectified value (the average value of the rectified signal) and the average area corresponding to each specific Raa have been taken into consideration, with the intention to track the behaviour of Raa. For each of the curves describing these evolutions, the means and standard deviations were calculated, as displayed later in the tables, after the normalization of the values.

## Results

The graph in Fig.1 presents the evolutions of Raa for all the 5 experiments, 1.1 to 5.1, the second figure denoting the pair of electrodes, as two distinct channels were recorded.. On the y axis, the actual values of Raa are represented.



X 5 sec.
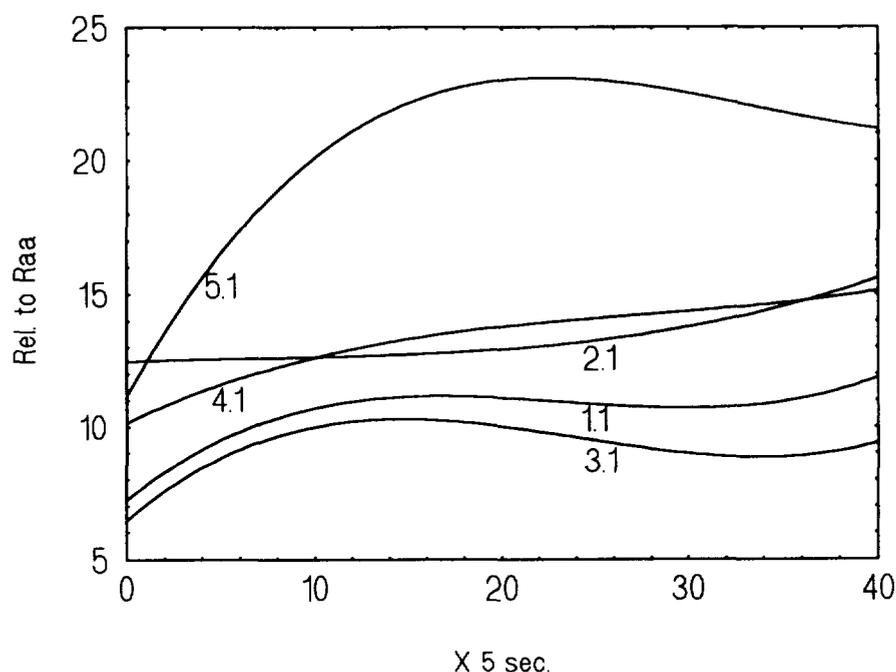
**Figure 2.** Mean rectified value evolution

The figures 2 and 3 present the evolutions of the mean rectified values and average area respectively and the table bellow the means and standard deviations for the parameters under study. For purposes of comparation only the values have been recalculated to meet the average of 10.56 specific to Raa in experiment 1; the same can be noticed on the graphs ( Fig. 2,3).

| Experiment | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| **Raa** | 10.56+/-.31 | 11.57+/-.24 | 11.4+/-.12 | 11.76+/-.26 | 11.8+/-.35 |
| **Vmr** | 10.56+/-.93 | 13.3+/-.85 | 9.29+/-.81 | 13.4+/-1.27 | 20.7+/-3.02 |
| **Area** | 10.56+/-2.091 | 17.23+/-2.21 | 9.63+/-1.1 | 16.24+/-2.89 | 47.63+/-13.95 |

**Table 1**

## Average Area Evolution
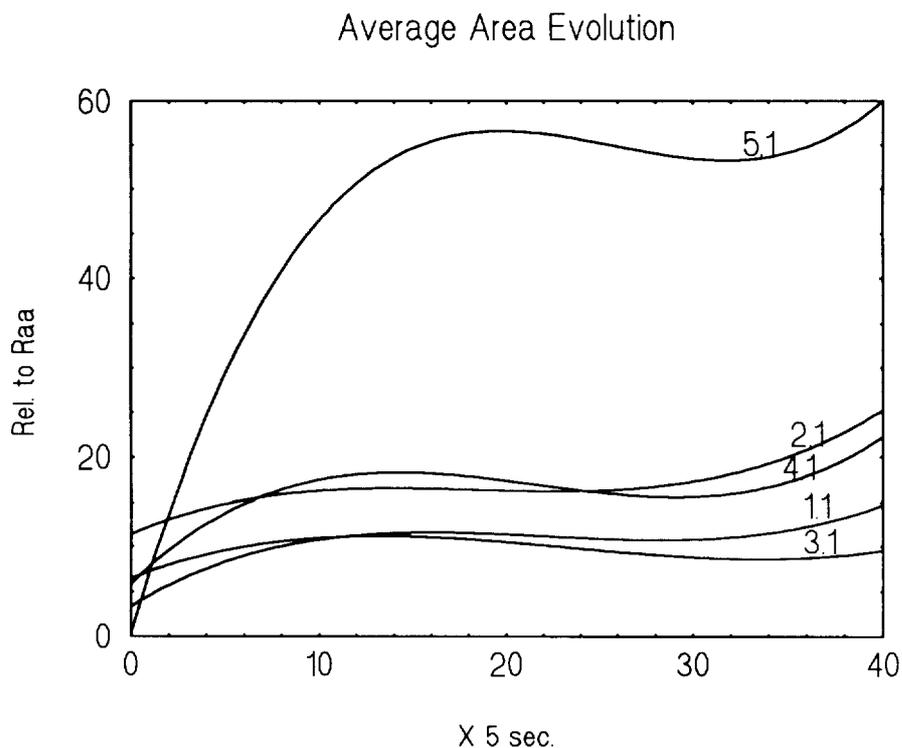


**Figure 3.** Average area evolution

It can be clearly traced that Vmr and Area have a much greater variability than Raa, again confirming the previous statements on the area/amplitude ratio calculated for the needle electromyogram on singular motor unit action potentials.

## Median Frequency Evolution
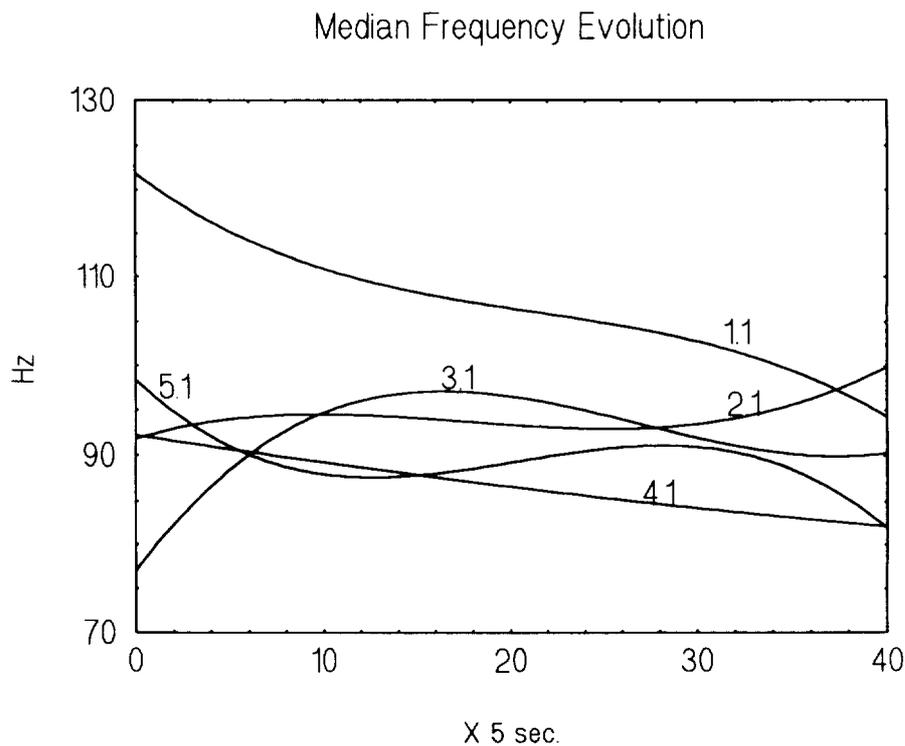


**Figure 4.** Median frequency evolution

As a new parameter ( Raa) has been defined, calculated for long segments of EMG data (400 msec. to meet the stationarity constraints), the previous results confirm a much smaller variability of the global area/ amplitude ratio ( Raa) comparing to those of the area and amplitude taken alone, which is encouraging in using it to more reliable describe the EMG, together with other consistent parameters.

It is interesting to note the Raa evolution during the sustained contraction comparing with the evolution of the median frequency Fmed which is now consacrated as a strong, consistent and reliable parameter ( Fig.4); they are reversely behaving in a certain sense, but, as the meaning of Raa is in terms of waveform "thickness" , this behaviour is likely to be normal, and would be interesting to study how strongly the two are correlated. This would be important, considering the only fact that it is much simpler to calculate Raa. In conclusion Raa has a smoother consistent evolution than other parameters, well correlated with the overall evolution of the phenomenon of fatigue, it is much less dependent on the experimental conditions, all these confirming it as a powerful descriptor.

## References

Nandedkar, S. D., Barkhaus, P. E., Sanders, D. B., Stalberg, E. V. - Analysis of amplitude and area of concentric needle EMG motor unit action potentials, *Electroencephalography and clinical Neurology*, 1988, 69: 561-567

Nestianu, V., Tarata, M. - Metoda si circuit electronic pentru analiza cantitativa a electromiogramei, Brevet OSIM 77109, 1981

Nestianu, V., Tarata, M., Balsianu, F. - A New Method for the Quantitative Analysis of the Myoelectric Spikes Synchronisation, *Rev. Roum. Med. Neurol. Psychiat.* 1981, 19, 4, p.283-288

Tarata, M. - A New Architecture for an Efficient Acquisition of the EMG Signal, *The VIIth International Symposium on Motor Control*, Borovets, Bulgaria, June 21-25, 1993, Acta Physiologica Bulgarica, Vol.19, Nr. 1/2, 1993, p.52

# ELECTRO-IMPEDANCE METHODS AND ARRANGEMENT FOR SYSTEMIC AND LIMB BLOOD FLOW ASSESSMENT

Tadeusz Pałko & Grzegorz Pawlicki,
Warsaw University of Technology
Institute of Precision and Biomedical Engineering
Chodkiewicza 8, 02-525 Warsaw, Poland

## INTRODUCTION

The electro-impedance methods makes possible detection of the mechanical action of the cardio-circulatory system and estimation of many hemodynamic parameters that are related to this action. The most popular is impedance cardipgraphy based on Kubicek method [1]. This method is often used in cardiology for evaluation of systemic circulation, particularly for determination of basic hemodynamic parameters such as stroke volume, total ystemic blood flow (SBF), contractility of the heart, several cardiac intervals, etc. Limb blood flow (LBF) is also of great clinical significance. Nowadays among of the many method of LBF examination most important are ultrasonography and impedance angiography. The presented paper deal with the last one.

The purpose of the study is presentation of modified Kubicek electroimpedance method for assessment of the total systemic blood flow (SBF) and the for detection of aortic stenosis, method of rheoangiography for examination of total or segmental limb blood flow (LBF), the rheographic arrangement developed by authors and results of selected clinical application.

## METHODS

The principle of electro-impedance tissue measurement is based on tetrapolar current method.

The SBF evaluation is based on simplified equivalent scheme of the thorax. In this method (Fig. 1) the alternating current (1 mA, 100 kHz) is injected longitudinally to around the high part of neck and the waist. The impedance signals are detected by two disc electrodes, one located on the neck base and the another near to xiphoid. This method [3] compare to the Kubicek`s one is modified by replacing the band detection electrodes with more convenient disc electrodes.

The stroke volume in ml (SV) is determined by Kubicek formula:

$$SV = \rho \left( \frac{l}{Z} \right)^2 \left| \frac{dZ}{dt} \right|_{max} T$$

where: $\rho$    - blood resistivity in $\Omega$cm, $\rho=52,3 \exp (0,022 \, Ht)$
     $l$    - distance between the detecting electrodes (Fig. 1)

Z        - base impedance in $\Omega$

$\left|\dfrac{dz}{dt}\right|max$     - maximum value of first derivate of the impedance change during heart systole period in $\Omega$/s (fig. 3)

T        - left ventricular ejection time in s (fig. 5)
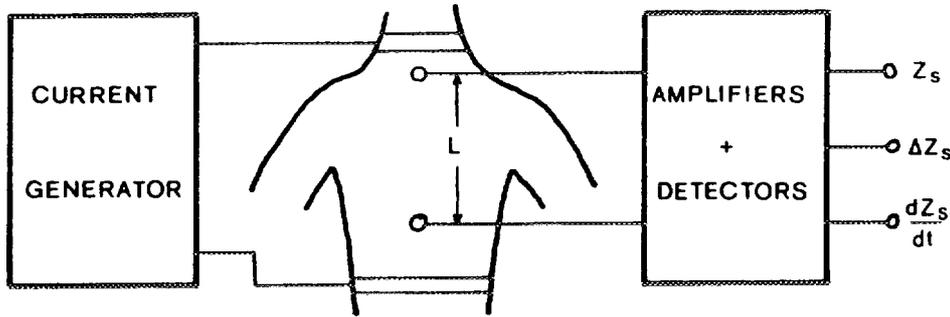


Fig. 1. The principle of SBF study.

The SBF is determined by multiplication of SBF by the heart rate (HR)
For the study of SBF was used impedance cardiometr (rheometer) type RM-23 (fig.2) developed and designed by us in Warsaw University of Technology. Impedance cardiograms (fig. 5) are recorded by using of multichannel jet pen recorder type Mingograph developed by Elema-Siemens Company.

The total or segmental LBF evaluation is based on segmental limb model. The determination of LBF take into account the relation betwen the volume change due to blood flow in the measured region of the limb and the electrical impedance change of this region. The alternating current (1 mA 50 Hz) is injected to the legs by two band electrodes - Ea and the impedance signal is detected from inner two band electrodes - Ed (Fig. 3)
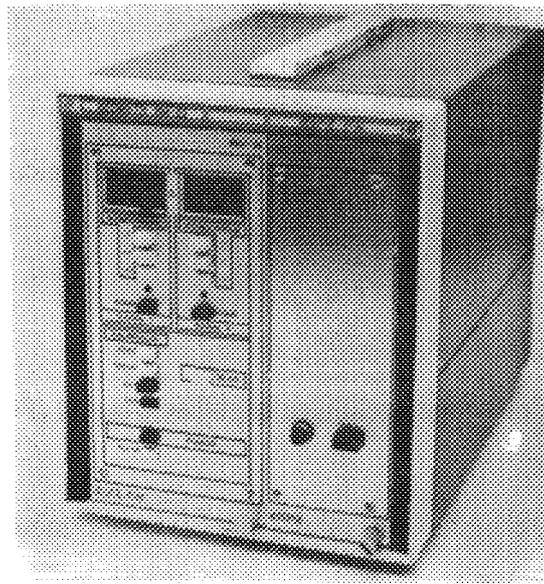


Fig. 2. Impedance cardiometer (rheometer) type RM-23 used for SBF study.

63
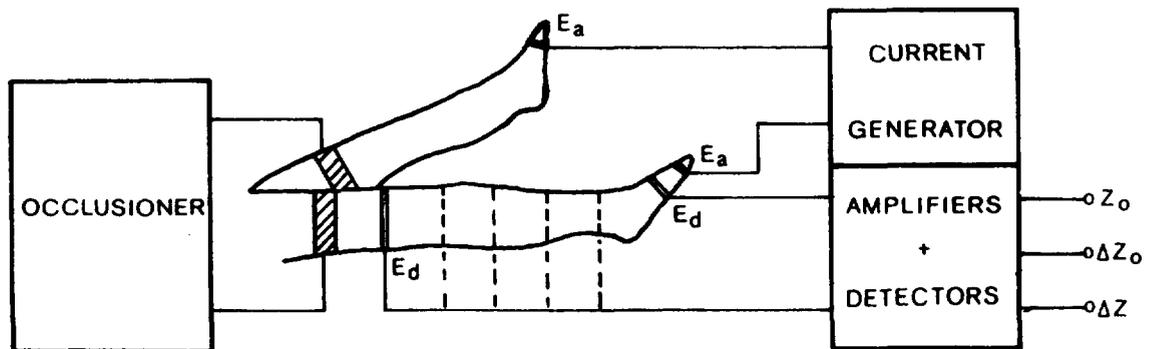
Fig. 3. The principle of LBF study.

For the measurement of total and segmental blood flow either the pulsation of arteries or the vein occlusion technique can be used. In this method [4], total LBF is measured on the segment from the end of extremity till the level at which the blood flow is to be determined. Segmental LBF is defined as the difference between total LBF determined at the proximal limit of the segment and at the distal one (Fig. 3). Using pulsation method of impedance signal, the LBF (Qp) is determined according to Nyboer [2] formula:

$$Qp = \rho \left( \frac{l}{Zo} \right)^2 |\Delta Z| HR$$

where: $\rho$ - blood resistivity,
       l - length of limb segment from the end of extremity to the level at which the blood flow is determined,
       Zo - base of impedance,
       $|\Delta Z|$ - change of impedance,
       HR - heart rate (Fig. 6).

Using the vein occlusion technique, the LBF (Qoc) is calculated according to formula:

$$Qoc = \rho \left( \frac{l}{Zo} \right)^2 \left| \frac{\Delta Zo}{\Delta t} \right|$$

where: $\left| \dfrac{\Delta Zo}{\Delta t} \right|_{max}$   - slope of the occluisive impedance signal (Fig. 6)

Standarized LBF related to 100 ml of tisue (q) is determined as:

$$q = \frac{\Delta Zo}{\Delta t \cdot Zo}$$
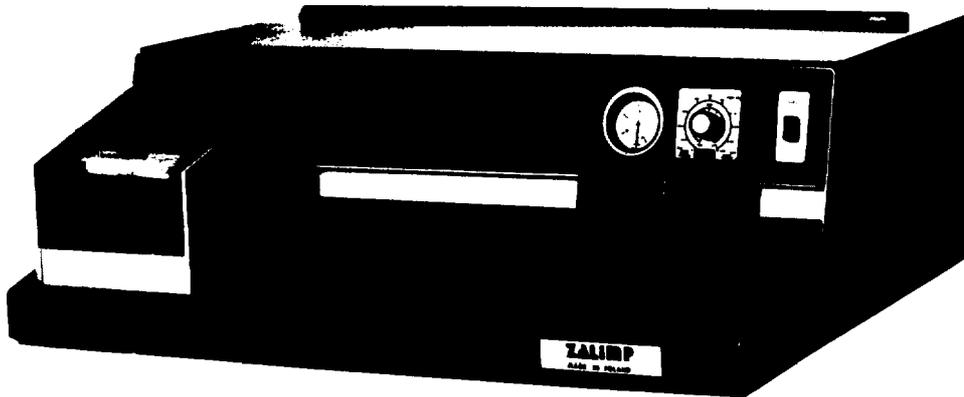
Fig. 4. Rheoangiograph type RANG-1 used for LBF study

For the study of LBF was used rheoangiograph type RANG 1 (fig. 4) with possibility of automatic occlusion performance. The device was developed by us in Warsaw University of Technology.

RESULTS
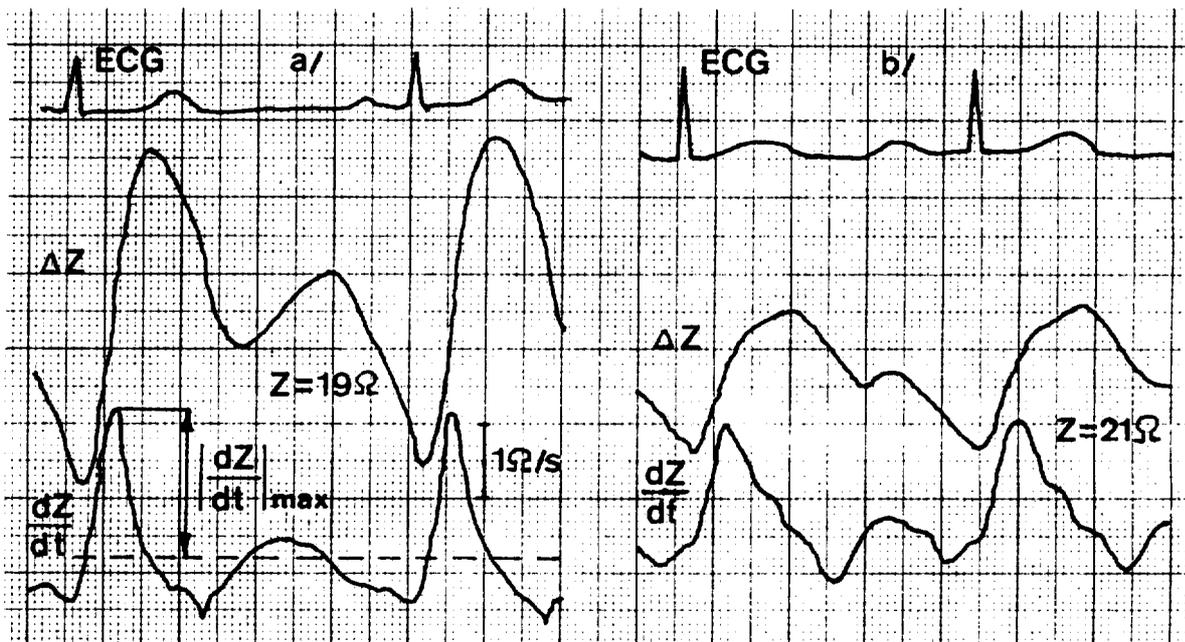


Fig. 5. Impedance cardiograms $\Delta Z$ and dZ/dt with ECG
a/ for healthy, b/ for patient with aortic stenosis

The investigation were carried out on 20 healthy persons and 40 patients divided into two groups. The SBF examination was performed 20 healthy and 14 patients with aortic stenosis. The shape of impedance cardiographic curve differs very significantly for patients

and for healthy (Fig. 5). Also SBF calculated per square meter of the body surface (SBFi) is a very important parameter for assessment of systemic circulation. For healthy: SBFi = 3,34 $\pm$ 1,2 l/min/m$^2$, for patients: SBFi = 1,64 $\pm$ 0,96 l/min/m$^2$.
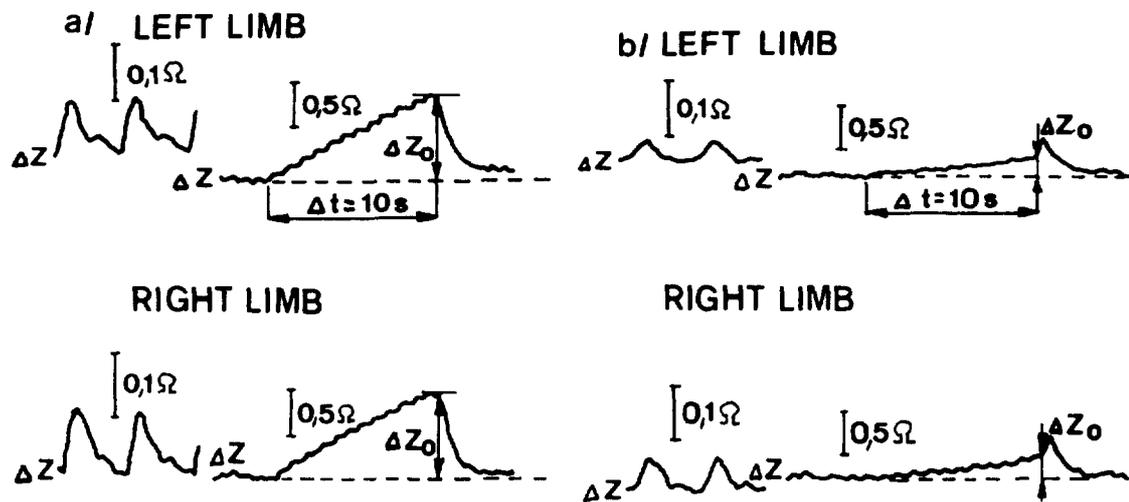


Fig. 6. Pulsation and occlusion impedance angiograms
a/ for healthy, b/ for patient with atherosclerosis

The LBF was examined on 20 healthy and 26 patients with atherosclerosis or intermittent claudication. Also on this case the shape of pulsation of impedance angiographic courve differs very significantly for healthy and patients (Fig. 6). The value of total LBF related to 100 ml of tissue (q) was appeared to be significant parameter for diagnosis of leg circulaytion abnormality. Everage value of this parameter was: for healthy q = 2,84 $\pm$ 0,58 ml/min100 ml, and for patients with atherosclerosis or intermittent claudication: q = 1,16 $\pm$ 0,79 ml/min100ml.

## CONCLUSION

1. The presented methods are noninvasive, very convenient and useful for clinical application.
2. The impedance method of SBF examination is convenient for evaluation of the hemodynamic cardiac activity and sensitive to detection of aortic stenosis.
3. The impedance method of LBF examination is helpful for recognition of blood circulation abnormality in the limbs.

## REFERENCES

1. Kubicek W.G. et al.: The Minnesota impedance cardiograph theory and application. Biomed.Eng., 9, 410- 416, 1974.
2. Nyboer J.: Electrical Impedance Plethysmography, Springfield, Charles C. Thomas, 1959.
3. Palko T.: Rheoimpedance methods for aortic and pulmonary blood flow. Med.Biol.Eng.Comp. Vol. 23, Supp. 1, 119, 1985.
4. Pawlicki W.G.: Electroimpedance rheography. Post.Fiz.med., 21, 1-57, 1986.

# ESTIMATION OF REYNOLDS SHEAR STRESSES AND TURBULENT SPECTRAL ENERGY DISTRIBUTION IN VICINITY OF MECHANICAL HEART VALVE PROSTHESIS IN VITRO

František KLIMEŠ and Josef KOŘENÁŘ
Institute of Hydrodynamics, Academy of Sciences
of the Czech Republic, Podbabská 13, 166 12 Praha 6

## 1. ABSTRACT

Laser-Doppler anemometry (LDA) was used for velocity measurements downstream of artificial valve prostheses (Björk--Shiley disc type). The velocity signals were subjected to analog signal processing prior to digital turbulence analysis. It is known that blood corpuscules damage can be related to the degree of flow disturbances especially with elevated levels of turbulent shear stresses.
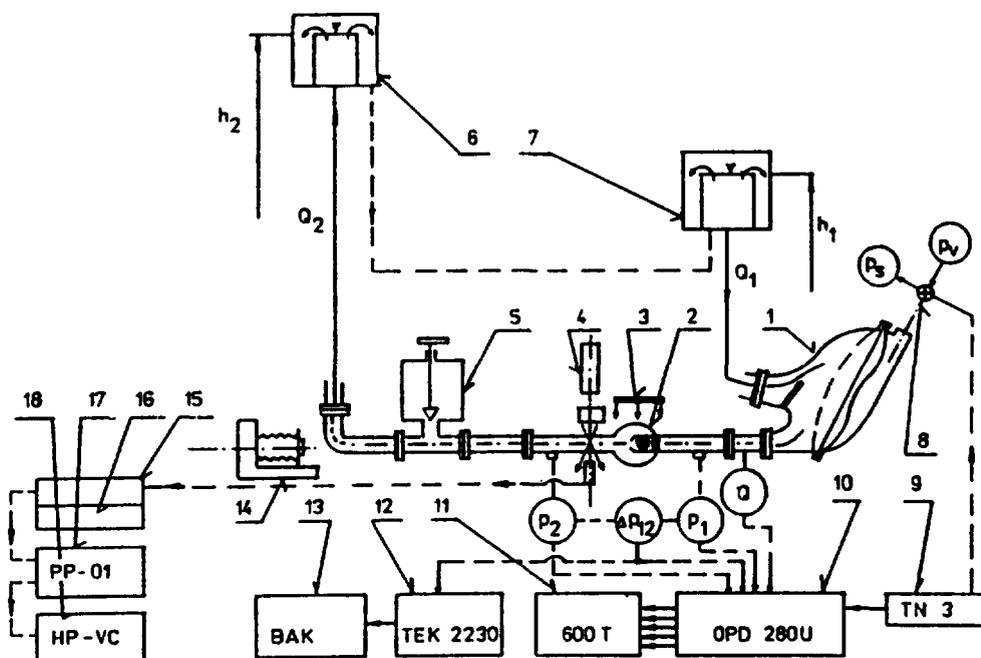
Fig. 1    A schematic sketch of the hydraulic model
1-the artificial heart (left half), 2-the test valve, 3-the source of light, 4-the laser, 5-the wind chamber (incorporation of a capacity member), 6 and 7-the freely overflowing vessels of adjustable heights, 8-the three way air valve, 9-the electropneumatic synchronizing unit, 10-the oscilloscope, 11-the six channel oscillograph, 12-portable analog oscilloscope, 13-plotter, 14-camera, 15-the frequency synthetizer, 16-counters, 17 and 18-calculators.
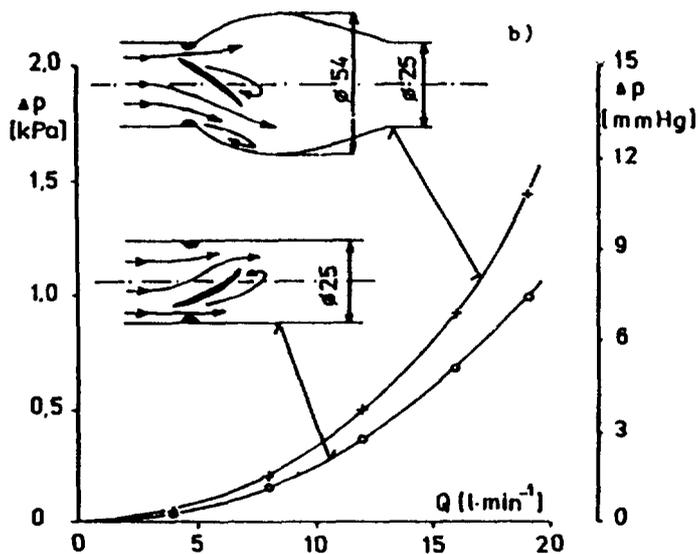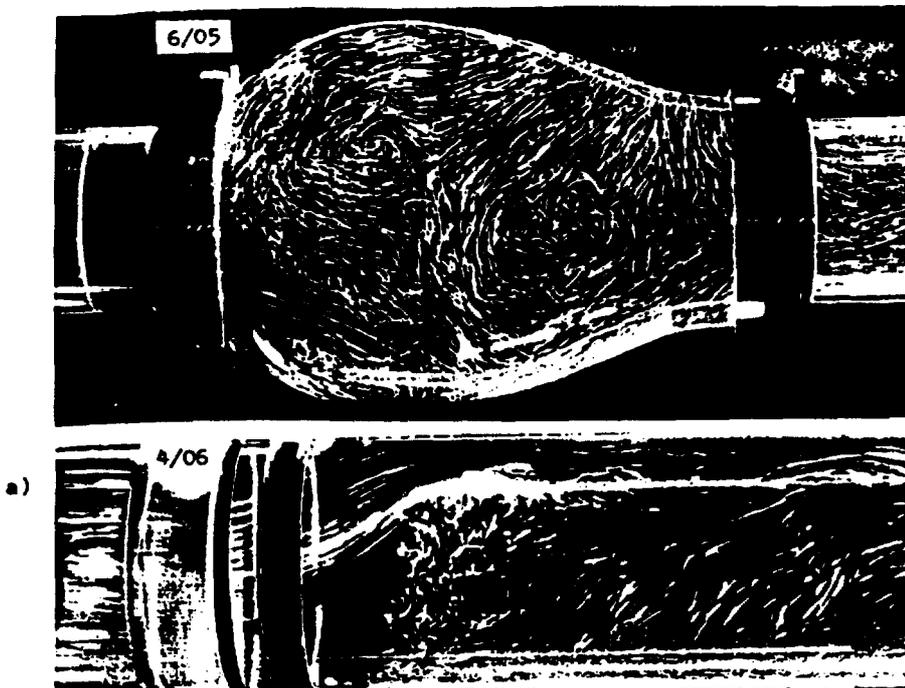
Fig. 2

a) Visualization of the flow through a disc valve located either in an axially symmetrical chamber (simulation of the aortic sinus - photo 6/05) with the flow rate 6 litres per minute or in a straight circular tube (photo 4/06 with the flow rate 5 litres per minute)

b) The corresponding shapes of the pressure losses for the stationary flow

## 2. Methods and results

The measurements were performed in a pulsatile mock circulation as seen in Fig. 1 and described in more detailed previously [1]. The source of pulsations is one half of the total-replacement heart, in which the frequency, the flow rate and the duration of the systolic phase can be changed by means of the electropneumatic unit. Constant values of the pressure simulation the level of the aortal pressure (i.e. approx. 13 kPa) and of that simulating the pressure in the pulmonary artery (approx. 1 kPa) were maintained during the experiments by the hydrostatic heights, $h_2$, $h_1$, of freely overflowing vessels. Incorporation of a capacity element into the system influences
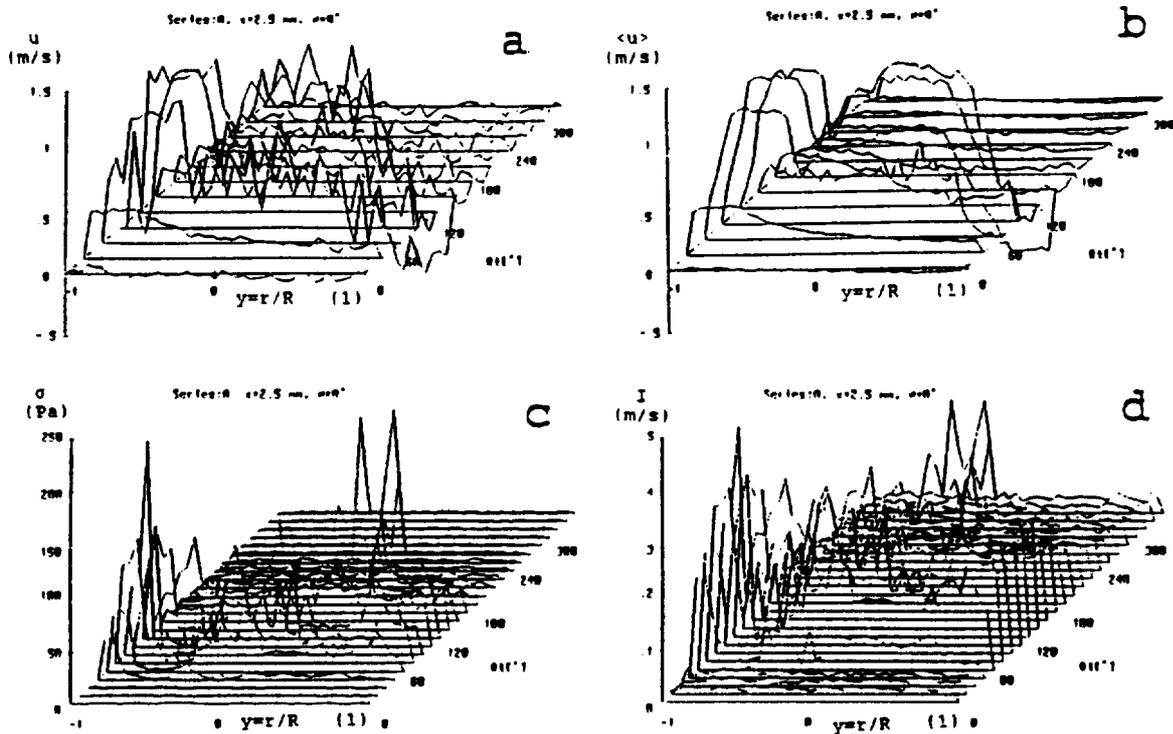
Fig. 3   The pulsatile velocity profiles measured by the LDA
2.5 mm downstream from the disc valve and some other
characteristic values calculated from the LDA data
(y=R/r - nondimensional radius, y=0 - tube axis, y =
= $\pm$ 1 - tube wall, $\Omega$ t = $\omega$t - circular frequency)
a) The measured velocity profiles (u).
   (Only 12 velocity profiles we presented for example,
   but 48 velocity profiles we measured through one
   period)
b) The mean velocity profiles <u>
c) The Reynolds normal stress ($\sigma$).
   (Only 24 profiles we presented for example)
d) The modified turbulence intensity (I)

positively the shapes of the time dependences of the pressure
and flow distributions. Instead of blood, a solution glycerol
in water was used, containing 45.7 weight % of glycerol. The
viscosity and density of this substitute solution are close
to those of blood.

In pulsatile flow model, point fluid velocities were
measured in two axial locations downstream from the valve (2.5
and 30.0 mm from the opened disc) and in two orientations of
the tilting disc valve with respect to the LDA traverse.

Fig. 2 shows an example of visualization of stationary
flow through a disc valve. Two cases are shown, one for a valve
located in the area of transition from a smaller to larger dia-
meter of the tube, the other being that of a straight tube.
Tubes of circular cross-section were used in both cases. The
higher values of the pressure losses corresponding to the for-
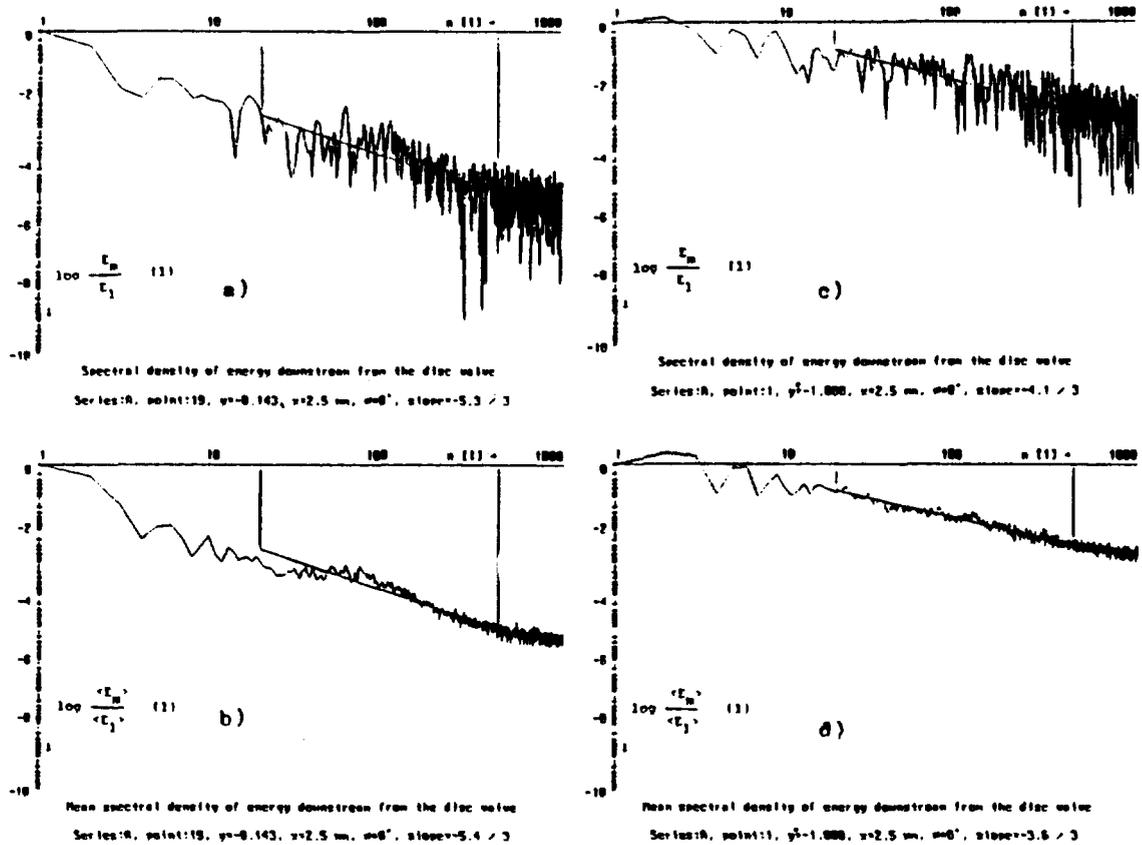
Fig. 4 The normalized spectral density of energy 2.5 mm down-
stream from the disc valve
a) The energy density spectrum (the slope between two
frequency limits is -5.3/3 , y = r/R = -0.143 = the
tube axis)
b) The mean energy density spectrum (the slope is
-5.4/3, y = -0.143)
c) The energy density spectrum (the slope is -4.1/3,
y ≐ -0.99 ≐ the tube wall)
d) The mean energy density spectrum (the slope is
-3.6/3, y ≐ -0.99 ≐ the tube wall)

mer configuration result from the higher degree of flow distur-
bances beyond the disc of the open valve in this case.

Measuring turbulent shear stresses is of major impor-
tance in artificial heart valve evaluation. According to se-
veral investigators a correlation factor C was found between
the maximum Reynolds shear stress, $\tau_{ij} = \rho \overline{u_i' u_j'}$ and Reynolds
normal stress, $\sigma = \rho \overline{u'^2}$, where $u_i'$ and $u_j'$ are the turbulent
velocity components acting normal to each other and u' is
the turbulent velocity component in the axial direction. There-
fore in vitro estimation of turbulent shear stresses downstream
of artificial valves, based on a single axial velocity component
u' alone, seems possible. If the flow is considered to be iso-
tropic than relation between $\tau_{ij}$ and $\sigma$ can be established
as $\tau_{ij} = C \sigma$ , where C = 0.4 for steady two-dimensional homo-
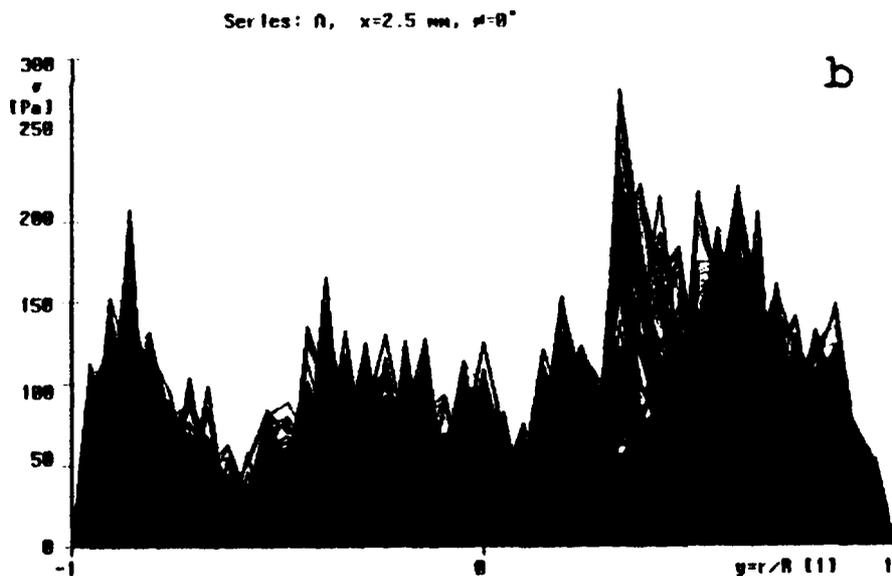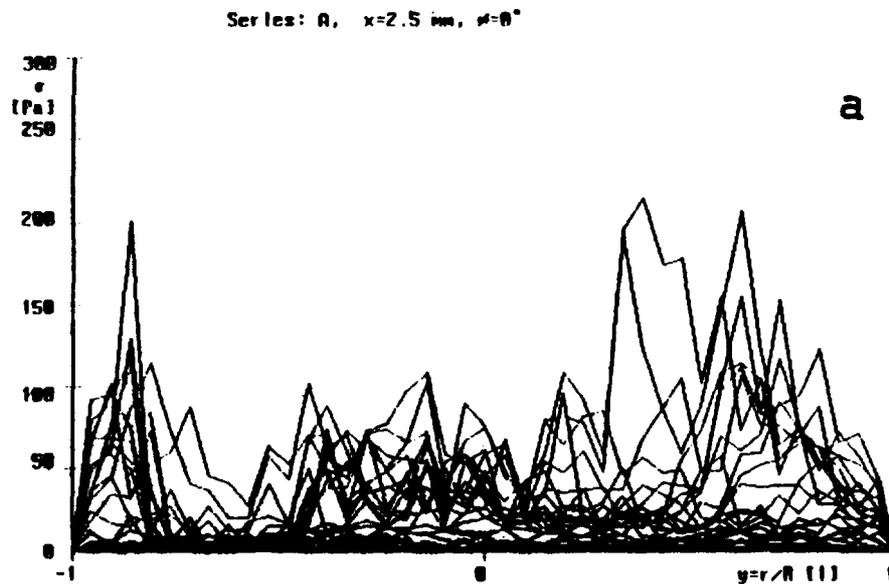
Fig. 5    The Reynolds normal stresses 2.5 mm downstream from
the disc valve (see Fig. 3c)
a) The Reynolds normal stresses are plotted together
for 24 profiles through one period
(y = r/R - nondimensional radius)
b) The Reynolds normal stresses are plotted together
for 48 profiles through one period)

geneous turbulent shear flow and C = 0.5 for large values of
Reynolds stresses. For pulsatile flow field downstream of
aortic valves the conditions of isotropy cannot be fulfilled,
because any flow having a mean velocity variation is noniso-
tropic. However there is region with small scale turbulence de-
fined as "local isotropy" [3] where the turbulence energy
density spectrum plotted in a double logarithmic scale have
the slope -5/3. This region is known as the Kolmogoroff inertial
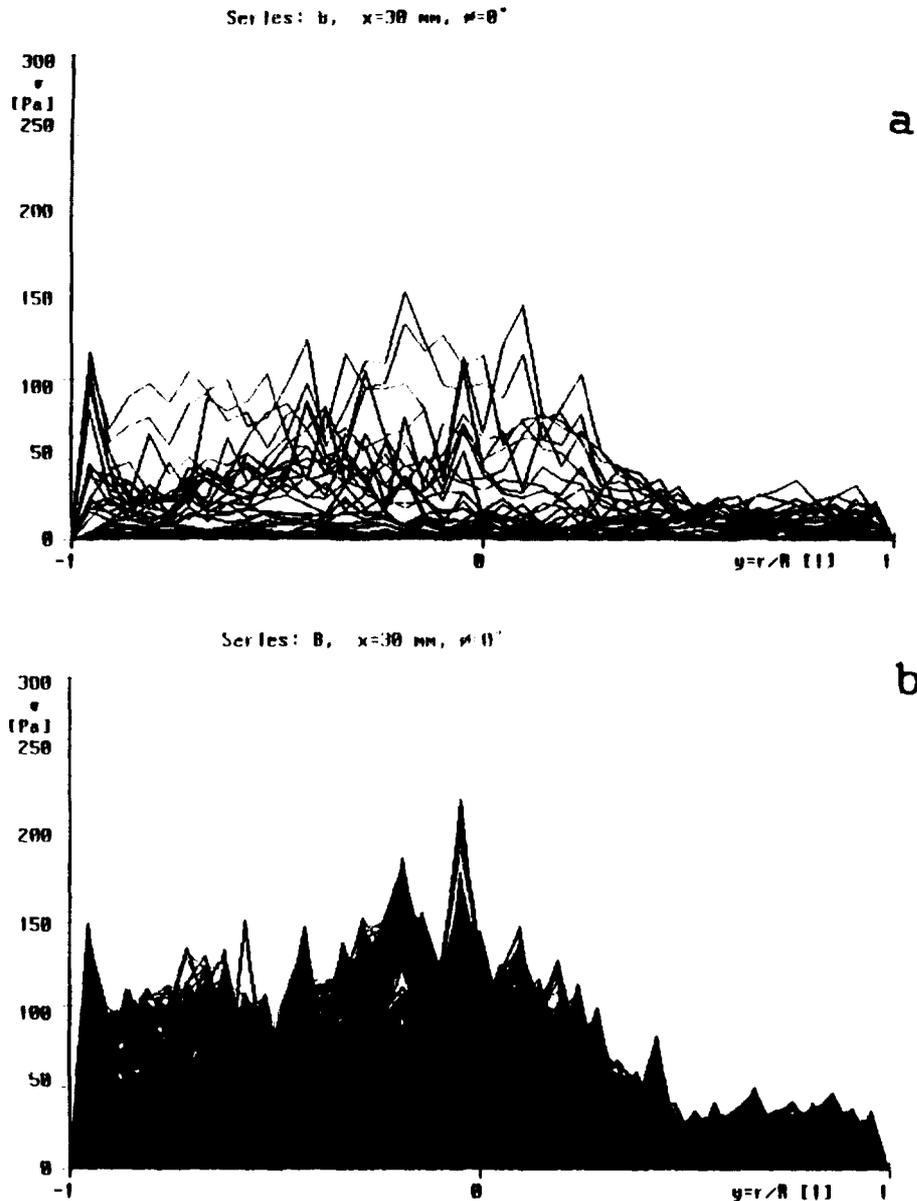
Fig. 6   The Reynolds normal stresses 30.0 mm downstream from
the disc valve
a) The Reynolds normal stresses are plotted together
for 21 profiles through one period
b) The Reynolds normal stresses are plotted together
for 48 profiles through one period

sub-range.

A typical obtained results are shown in Fig. 3 and
Fig. 4, where the following nomenclature is used [2]:

a) u - the continuous (instantaneous) record of velocity
obtained from the LDA measurements,

b) $\bar{u} = \frac{1}{T} \int_{0}^{T} u \, dt$ - the mean velocity, where T is the period of
each cycle,

$<u> = <u_k> = \dfrac{1}{P} \sum\limits_{i=1}^{P} u_{ik}$ — ensemble average of velocity

$u = u_{ik}$ , where $u_{ik}$ is the value
of velocity in the n-th period and in the k-th time of
period, P is a number of measured periods (in our case P=27)

c) $u'_{ik} = u_{ik} - <u_k>$ — the fluctuating component of velocity

d) $\sigma = \rho <u_k'^2> = \dfrac{\rho}{P} \sum\limits_{i=1}^{P} u_{ik}'^2$ — the Reynolds normal stress, where
$\rho$ is the fluid density and

$$<u_k'^2> = \dfrac{1}{P} \sum\limits_{i=1}^{P} u_{ik}'^2$$

e) $I = (<u_k'^2>)^{1/2}$ — the mean turbulence intensity. The con-
vential form $I = (<u_k'^2>)^{1/2}/<u_k>$ is not used, because the
definition gives nonphysical interpretation for pulsatile
flow, when the average velocity goes through zero value

f) $E_{mi} = a_{mi}^2 + b_{mi}^2$ — the spectral density of energy, where
i means the period and m the index of harmonic order.
For the spectral density of energy calculations we used the
following Fourier series:

$$u_i(t) = \overline{u} + \sum\limits_{m=1}^{M} a_{mi} \cos(2\pi mt/T) + b_{mi}\sin(2\pi mt/T),$$

where M is the index correspondign to Nyquist frequency
(the higher indexes have no physical interpretation in re-
lation to the limited number of obtained data).
Coefficients $a_{mi}$ and $b_{mi}$ are given as:

$$a_{mi} = \dfrac{2}{T} \int_{0}^{T} u(t) \cos(2\pi mt/T)\, dt,$$

$$b_{mi} = \dfrac{2}{T} \int_{0}^{T} u(t) \sin(2\pi mt/T)dt, \text{ for } m=1 \text{ to } m=M, \ a_{oi} = \overline{u},$$

$$b_{oi} = 0.$$

In consideration to the discret data character we used the
fast Fourier transform method (FFT) to obtained values of

coefficients which follow from the direct Fourier transform.
The turbulence energy density function can be calculated as
a ensemble averaged and as a fluctuating spectral density
of energy

$$<E_m> = \frac{1}{P} \sum_{i=1}^{P} E_{mi} \quad , \quad E_{mi}' = E_{mi} - <E_m> \quad ,$$

where P is a number of measured periods (P=27). The total
measured values in one period is 2400.

## 3. Conclusions

From the flow visualization we observe flow reversal and
separation regions through the cardiac cycle. The LDA velocity
profiles measurements also show relatively large velocity gra-
dients which may induce higher turbulent stress, especially at
peak systole, and therefore large shearing action on the formed
elements of blood.

The maximum Reynolds normal stresses were found in the
cross-section 2.5 mm immediately downstream from the opened
disc, from 200 to 250 Pa. According to several investigators
the critical shear level for lethal erythrocyte and thrombocyte
damage is 200 to 400 Pa for an exposure time of 1 to 10 ms,
which is the estimated time for a blood cell to pass a heart
valve prostheses. The endothelium is deteriorated by stresses
of 40 Pa. As can be seen from Fig. 5 and Fig. 6, towards the
far cross-section behind the valve, the Reynolds normal stres-
ses decrease and are without larger peaks.

The effect of pulse rate on the flow dynamics past valve
prosthesis is yet to be investigated.

## References

[1] Klimeš F., Kořenář J.: Turbulent stress measurements in
    pulsatile flow downstream of the heart valve prosthesis.
    Proceed. 14th Intern. Congress of Biomechanics, Paris
    1993, pp. 694-695.
[2] Hasenkam J.M., Nygaard H., Giersiepen M., Reul H.:
    Turbulent stress measurements downstream of six mechanical
    aortic valves in pulsatile flow model. J. Biomechanics,
    21, 8, 1988, pp. 631-645.
[3] Hinze J.O.: Turbulence. McGraw Hill, New York 1975.
[4] Severa M.,Chára Z.,Gardavský J.: Some results of LDA investi-
    gation of nonhomogeneous suspension flow. 7th Int.Conf.
    "Transport and sedimentation of solid particles",Wroclav 1992

# A HEART AND CIRCULATION MODEL TO SIMULATE RATE ADAPTATION OF ANS-CONTROLLED PACEMAKERS

Max Schaldach, Albrecht Urbaszek, Helmut Hutten*

Zentralinstitut für Biomedizinische Technik, Universität Erlangen-Nürnberg, Turnstr. 5, 91054 Erlangen, Germany

*Institut für Elektro- und Biomedizinische Technik, Universität Graz, Inffeldgasse 18, 8010 Graz, Austria

*Abstract* — *The comprehensive understanding of intrinsic and extrinsic control processes of the heart and circulatory system is a prerequisite for the design of rate-responsive pacemakers. To establish a mathematical description, a numerical model of the human cardiovascular system was developed for simulating the short-term regulatory processes that adapt the cardiac output of the heart to changing circulatory demands. Since control processes relevant to cardiac performance occur on different levels, e. g. myocardial cell, heart mechanics, peripheral changes and autonomic nervous system (ANS), all these levels were combined in this approach. Both normal and pathological states, e.g. sick sinus or AV block, are considered. The purpose of the simulations is to investigate the control aspects of rate responsive pacing and to develop and test algorithms for rate adaptation based on cardiac control signals.*

## INTRODUCTION

The classical form of physiological rate-adaptive pacing is dual-chamber pacing (DDD). With an intact sino-atrial node, improved hemodynamic results are obtained by restoring AV synchrony. However, in the presence of insufficient chronotropic response to exercise the benefits provided to the patients are limited. In these cases, corporeal and cardiac control parameters have to be employed for rate adaptation. Since in most patients inotropic information is still preserved, contractility-dependent parameters can provide ANS information. An elegant approach to detect inotropic changes in the mechanical contraction pattern is intracardiac impedance measurement, since the signal is determined primarily by changes in shape and geometry of the myocardium during contraction.

The ANS-controlled pacemaker, which uses the stimulating electrode after the stimulus•for intracardiac impedance measurement, was considered for simulating rate adaptation under different load, posture and pathologic conditions. Since the relationships are complex and nonlinear, analytical solutions cannot be obtained in most cases and, thus, numerical simulations are the only way to describe the system's behaviour.

## METHODS

Due to the multi-level structure of the regulatory processes to be considered the model description has to be implemented in the appropriate structure. Compared to other models of the circulation, e. g. "HUMAN" by T. G. Coleman [1], the presented model is especially designed for simulating short-term regulatory processes. Only those physiological subsystems and relationships directly or indirectly contributing to chronotropic and inotropic regulation are considered. The model has a hierarchical and modular structure, it includes three levels: the circulation, the heart muscle mechanics and the myocardial cell (fig. 1).

Circulatory System

The peripheral circulation is represented by a few segments only, this part of the model has to reflect the following effects:
- changes in preload,
- changes in afterload,
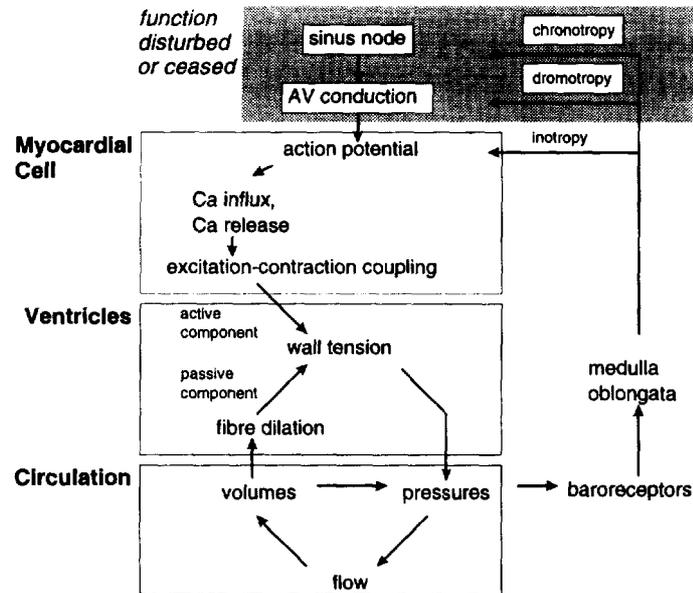
Fig. 1. Data flow graph of the heart and circulation model across the levels circulation, ventricles and myocardial cell.

- effects of physical exercise and temperature,
- effects of postural changes.

Therefore, the major vessels of systemic and pulmonary circulation (aorta, venae cavae, pulmonary arteries and veins) are considered to be compartments with elastic walls, that develop a certain pressure p for a given volume V, dependent on the compliance C and the unstressed Volume $V_0$. The flow between two compartments depends on the pressure gradient and the flow resistance between the compartments. Finally, the net inflow of each compartment equals the volume change $dV/dt$; these relationships applied for each (circulatory) compartment form a set of first order differential equations.

This simple representation describes the circulation sufficiently for simulating the load conditions for heart filling and ejection. The inertia of the flowing blood had been included in a previous version, but since it had only little effects on the system's behaviour but increased the computation times dramatically, it is ignored at present.

Ventricles

Since the impedance signal measured by the pacemaker is determined mainly by the geometrical changes of the ventricles during myocardial contraction, the mechanical action is considered in far more detail. The evaluation is based on a geometrical hypothesis, which allows to calculate the relationships between ventricular volumes, ventricular dimensions, resulting wall stress and, via the law of Laplace in a more general form, resulting ventricular pressures. Cardiac dilation and altered tissue properties, e.g. resulting from a myocardial infarct, can be represented here.
The ventricular volumes determine the geometrical dimensions according to the underlying shape hypothesis (spherical-ellipsoidal approximation, see fig. 2).
The filling of the ventricles stretches the fibres and causes a certain wall tension according to the passive length-tension curve of the sarcomere (diastole). During systole, an active tension component resulting from activation of the contractile filaments during action potential is added and leads to the opening of the valves and ejection of blood. Additionally, the force-velocity relationship (decreasing force of contraction with increasing velocity of shortening) is considered.
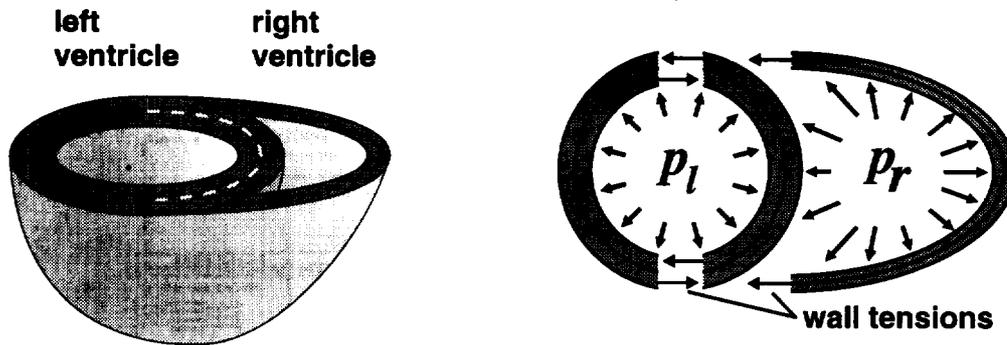
Fig. 2. Spherical-ellipsoidal approximation of the left and right ventricular geometries (left), relationship between wall tensions and ventricular pressures via the law of Laplace (right).

The ventricular pressure finally results from wall tensions and thicknesses according to the law of Laplace, which is applied here in a more general form (see fig. 2). Thus, mutual influences of both ventricles and Starling's law are incorporated [2].

Myocytes

Changes of the contractile status of the heart resulting from adrenergic influences on the myocardial cell [3, 4, 7] alter the contraction pattern and thus are mapped in the time course of the intracardiac impedance. Since these effects, together with pharmacological influences, e.g. calcium antagonists, play an important role, the dynamics of $Ca^{2+}$-uptake, -storage and -release by the sarcoplasmic reticulum, altered ion channel properties and pump activities, action potential and excitation-contraction coupling are simulated by a myocardial cell module. Figure 3 shows the different $Ca^{2+}$-currents and storage compartments considered for the simulation.

The evaluation of the transmembrane potential is based on an equivalent circuit representation of the sarcolemma, it depends on the ion conductivities of $Ca^{2+}$, $Na^+$ and $K^+$, their diffusion potentials and the membrane capacity. Diffusion potentials are determined of the respective ion concentrations according to the NERNST equation; passive ion currents result from the electro-chemical gradient and conductivity, while the pump currents depend only on the concentration
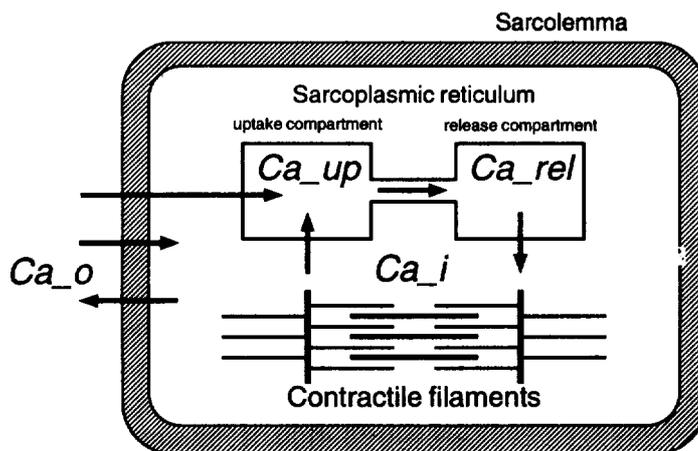


Fig. 3. Intracellular $Ca^{2+}$ concentration, which determines force of contraction, results from passive (along the electrochemical gradient) and active (via ion pumps) ion movements between extracellular and intracellular space, uptake and release compartments of the sarcoplasmic reticulum.

77
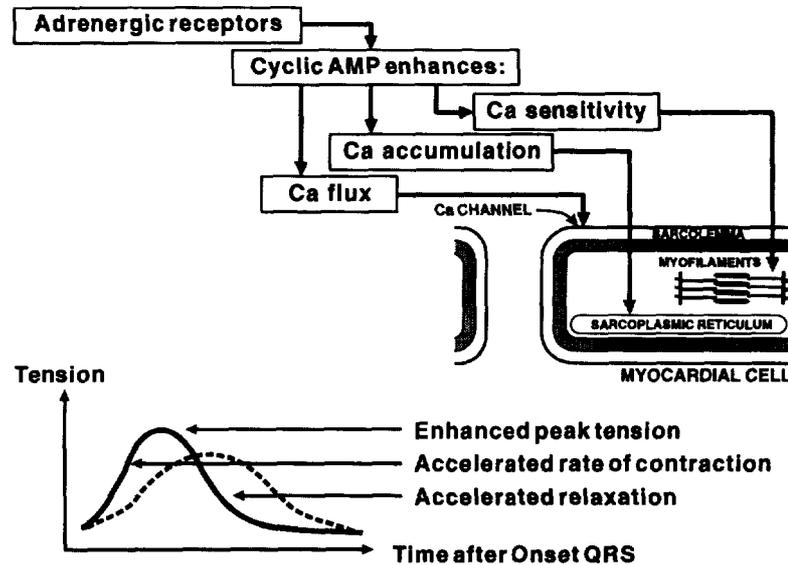
Fig. 4. Adrenergic effects on contractility.

on the input side of the pump according to first order kinetics (MICHAELIS-MENTEN-theory). When sympathetic activity is increased, a neurotransmitter, norepinephrine, is released, which exerts $\alpha$- and $\beta$- adrenergic actions. Via a second messenger system (cyclic AMP), the $Ca^{2+}$ dynamics are influenced in a way that the typical inotropic effects: higher peak tension, higher rate of rise in tension and accelerated relaxation, occur (see fig. 4).

Autonomic nervous system control

Since control aspects of the heart function are here of special interest, the autonomic nervous system (ANS) was included in the simulation. ANS-controlled adjustment of the heartbeat is considered as a closed-loop system, with mean arterial blood pressure being the controlled quantity [5, 6]. For this reason the model contains a procedure simulating the baroreceptor reflex: Symbolic control variables representing the sympathetic and parasympathetic tone are calculated as a function of mean arterial blood pressure. A mathematical description of the static and dynamic system behavior of the baroreceptors is used [7]. The control variables representing the efferent pathways in the model control the heart's pumping performance by adapting heart rate (chronotropy), AV conduction time (dromotropy) and ventricular contractility (inotropy). Furthermore, the increasing peripheral vascular tone as a consequence of sympathetic activity is also considered.

**RESULTS**

The performed simulations demonstrate, that a realistic system behaviour could be well approximated in many aspects, such as simulated pressure time courses, pressure-volume loops and ionic concentration curves. Fig. 5 shows measured time courses of left ventricular wall tension and thickness ( from [8], left), compared to simulated tensions of left ventricular wall, septum and right ventricular free wall (right). The modular structure allows the model to be easily modified or extended and facilitates the substitution of the sinus node module by a pacemaker: Either a real device is connected via an AD-converter or a software representation is used.

**CONCLUSIONS**

Evidence has been obtained that the intracardiac unipolar impedance signal, which is measured by the ANS-controlled pacemaker and from which a contractility parameter is extracted, is

Normal
Left Ventricle

Simulated
Wall Tensions
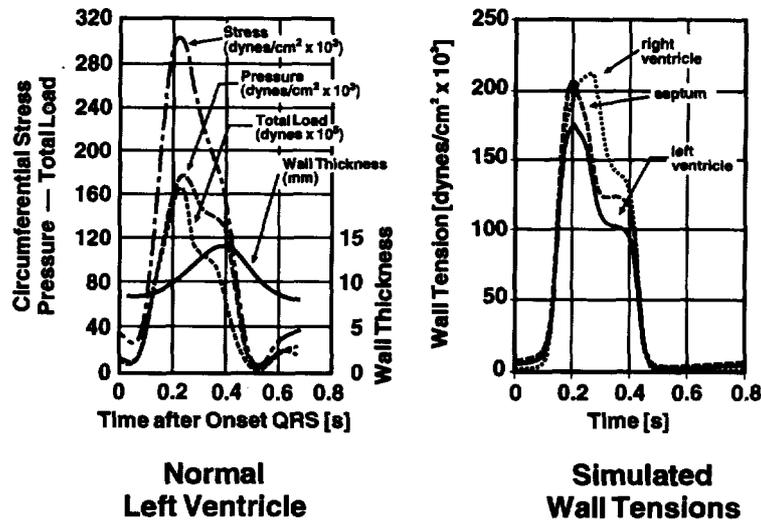
(After I. Mirsky, "Handbook of Physiology - Section 2", P.506, Fig. 8)

Fig. 5. Comparison of measured and simulated parameters. Left: measured left ventricular wall stress (tension), thickness, pressure and total load (after [7]). Right: simulated wall tension of left ventricular wall, septum and right ventricular free wall.

mainly modulated by ventricular wall tension and wall thickness time courses. These parameters depend directly on the inotropic state, so the measured impedance signal contains ANS information, that allows to determine the appropriate pacing frequency. Thus, adjusting the cardiac output adequately to circulatory demands, natural closed loop control is re-established.

## REFERENCES

[1] A. C. GUYTON, C. E. JONES, T. G. COLEMAN, *Circulatory Physiology: I. Cardiac Output and its Regulation*, Philadelphia: W. B. Saunders, 1973.

[2] A. URBASZEK, H. HUTTEN, M. SCHALDACH, "A heart and circulation model with emphasis on short-term regulation", *Proc. 14. Ann. Conf. IEEE Eng. Med. Biol. Soc. Paris 1992*, pp. 429-430.

[3] J. R. BROBECK, *Best & Taylor's Physiological Basis of Medical Practice*, 10th. ed., Baltimore: Williams & Wilkins, 1979.

[4] M. F. WENDT-GALLITELLI, "Excitation-contraction-coupling, intracellular calcium shifts and contraction of the heart muscle", in *Evaluation of Cardiac Contractility*, R. JACOB, Ed. Stuttgart, New York: Gustav Fischer, 1990.

[5] M. SCHALDACH, *Electrotherapy of the Heart*, Berlin, Heidelberg, New York: Springer, 1992.

[6] M. SCHALDACH, H. HUTTEN, "Intracardiac impedance to determine sympathetic activity in rate responsive pacing", *PACE*, vol. 15, pp. 1778-1786, November 1992.

[7] R. POLLEY, A. URBASZEK, M. SCHALDACH, "A mathematical approach to short-term baroreceptor behaviour", *Proc. 14. Ann. Conf. IEEE Eng. Med. Biol. Soc. Paris 1992*, pp. 2697-2698.

[8] I. MIRSKY, "Elastic properties of the myocardium: a quantitative approach with physiological and clinical applications", in *Handbook of physiology - Section 2: The cardiovascular system*, R. M. BERNE, N. SPERELAKIS, S. R. GEIGER, Eds. Bethesda, Maryland: American Physiological Society, 1979.

# ANS MONITORING USING RIGHT VENTRICULAR IMPEDANCE

H. Hutten, M. Schaldach*

Institut für Elektro- und Biomedizinische Technik, Technische Universität Graz,
A-8010 Graz, Austria

* Zentralinstitut für Biomedizinische Technik, Friedrich-Alexander-Universität,
Erlangen-Nürnberg, D-91054 Erlangen, Germany

## Abstract

The aim was to achieve rate-responsive pacing in patients with chronotropic insufficiency by adapting the pacing rate to the hemodynamic requirements. It has been demonstrated that the activity of the Autonomous Nervous System (ANS) on cardiac contractility is reflected by the Ventricular Inotropic Parameter (VIP) that can be measured by the unipolar impedance method. ANS monitoring allows restoration of the physiological closed-loop system for blood pressure control with appropriate rate adjustment. This concept, that has been realized in the Biotronik Neos-PEP VVI/R pacemaker, has been evaluated in more than 200 patients using spontaneous sinus rhythm (where present) for comparison, standardized exercise protocols, Holter monitoring during daily routine activities, and echocardiography. Rate response shows the typical physiological behaviour. This result is supported by reports of the patients on improvement of life quality and, additionally, by an improved NYHA ranking.

## Introduction

Proper adjustment of the pumping performance of the heart to short-term challenges, such as physical exercise, postural changes and temperature impacts, is accomplished by a closed-loop control system with a negative feedback from the baroreceptors via the medullary circulatory centers. The Autonomous Nervous System (ANS) with its controlling information mainly affects the chronotropic and inotropic behaviour of the heart. In patients with sick sinus syndrom, only the information for adjusting myocardial contractility is available in the heart. The basic idea has been to utilize that information for re-establishing another closed-loop system for rate-responsive pacing (fig. 1).

## Method

Among the different variables (fig. 2) that depend on the inotropic ANS information (e.g. systolic time intervals like PEP or LVET, stroke volume) none with the exception of VIP has met all requirements:

- any risk of positive feedback must be excluded;
- the relation between ANS activity and the respective parameter must be unequivocal:
- that parameter must be easily measurable, and the measuring method must not impair the safety of cardiac pacing or any relevant feature of the pacemaker including battery life time;
- the parameter must be applicable to rate adaptation during rest and different states of hemodynamic challenges despite individual differences in its actual value.

It has been shown [3,6,7,8] that the shape of the impedance curve that can be measured in the unipolar method between an electrode in the right-ventricular cavity and the pacemaker housing significantly changes in cases of inotropic misadjustment as a consequence of failing chronotropic adjustment. Although the shape of that curve depends on individual features (e.g. cardiac geometry, electrode position, type of electrode), its changes are well reproducible with regard to different typical hemodynamic challenges [3,8]. Furthermore, the shape of the curve is very stable with the exception of a short period (days to weeks) after implantation of the electrode. Impedance measurement is accomplished with the same electrode that is used for sensing and stimulation. The measurement period is restricted to a very short interval (24 ms) within the whole cardiac cycle. The most relevant signal is obtained if the measurement interval coincides with the begin of the ejection period as has been proven by

echocardiography. Measurement is performed by injecting a 4096 Hz square wave constant current of 40 μA.

Since the obtained signal is superposed by artifacts from ventilation, motion, and other causes, thorough signal processing, including phase demodulation and filtering with a Bessel band-pass filter (corner frequencies 0.3 Hz and 40 Hz) is utilized, resulting in a signal-to-noise ratio above 36 dB [7]. For further processing, the signal is digitised by an 11-bit analogue-to-digital converter with a resolution time of 8 ms and a minimum signal resolution of at least 6 bits. Among 28 different algorithms for isolating the ANS-related information from the VIP signal, the regional effective slope quantity (RQ) has proven to be the most qualified. This algorithm is easily adjustable to individual shapes of the impedance curve and its ANS-dependent changes, whereas it is rather insensitive against ANS-independent influences (e.g. spontaneous beat-to-beat fluctuations). Individual adaptation that considers both the pacing rate during rest and the maximum performance limit is achieved by the easily adustable Inotropic Index (II).
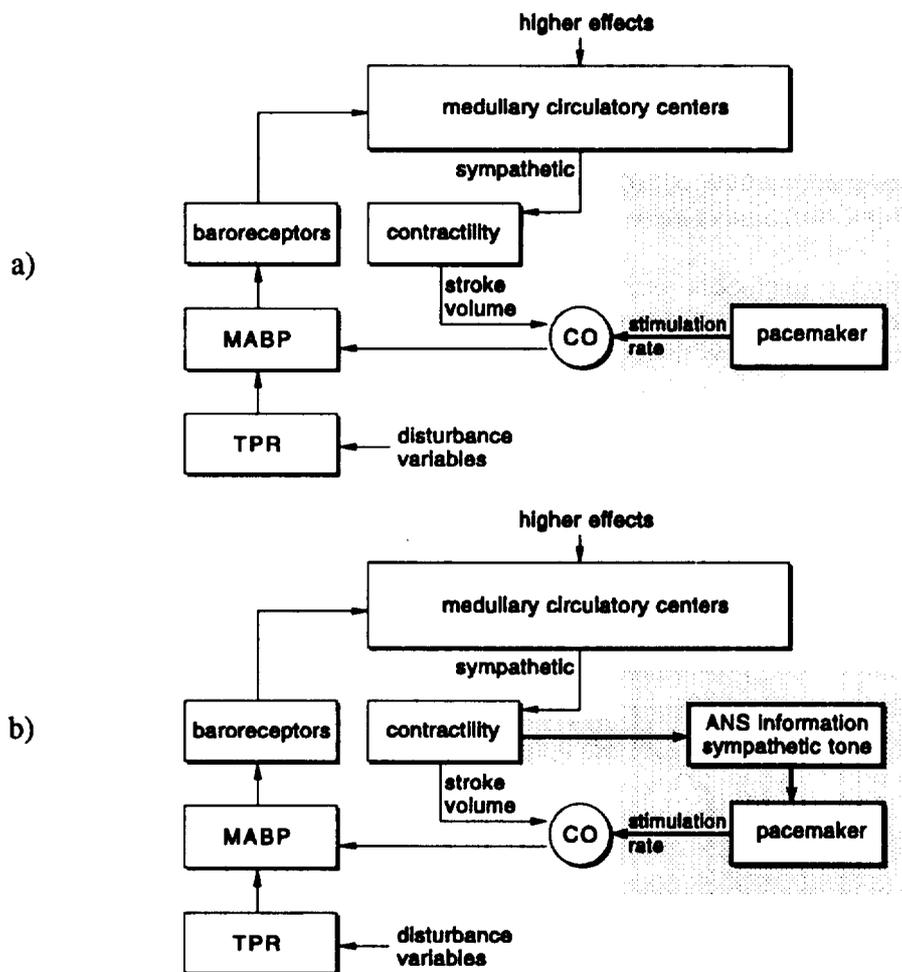


*Fig. 1   Schematic diagramm for ANS-controlled rate-adaptive pacing in patients with chronotropic incompetence: a) the control system is closed only with regard to SV, whereas HR is controlled by the pacemaker in an open-loop behaviour; b) by utilising for rate adaption the ANS information that is available in the contractility (VIP), the pacemaker becomes part of a restored closed-loop system.*

## Results

The therapeutical efficiency of the ANS-controlled rate-responsive pacemaker has been evaluated in a multi-center clinical study in more than 200 patients [4]. Most of those patients received the single chamber Biotronik Neos-PEP VVI/R pacemaker, but additionally patients

81

were included that received the dual chamber Biotronik Diplos-PEP DDD/R pacemaker. In 28 patients that received the single chamber version as a replacement for previously implanted VVI pacemakers and with clinical indication against the implantation of another electrode, the spontaneous sinus rhythm has been used for comparison with the rate adapted stimulation.
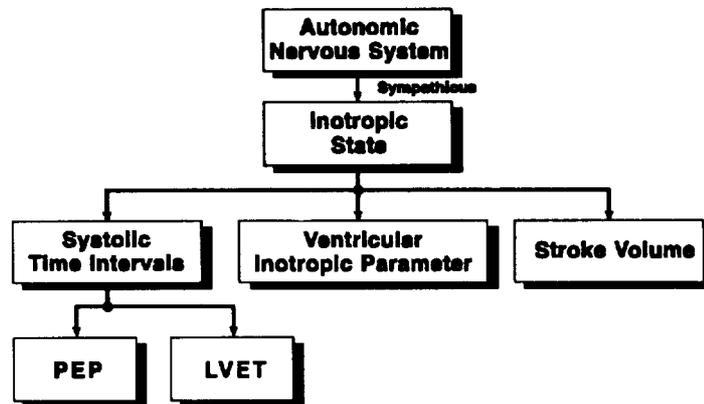


**Fig. 2** *The efferent part of the sympathicus with its inotropic influence on the myocardial muscle modulates the contractility which is reflected in the stroke volume, in systolic time intervals (e.g. LVET, PEP or QT-interval), and in the ventricular inotropic parameter (VIP).*

Standardized exercise protocols (e.g. bicycle or treadmill ergometry) and orthostatic challenges have been applied, supplemented by blood pressure measurements and echocardiographic monitoring, especially with regard to the opening phase of the right ventricular valves. A typical example that illustrates the physiological behaviour during increasing exercise is shown in fig. 3.
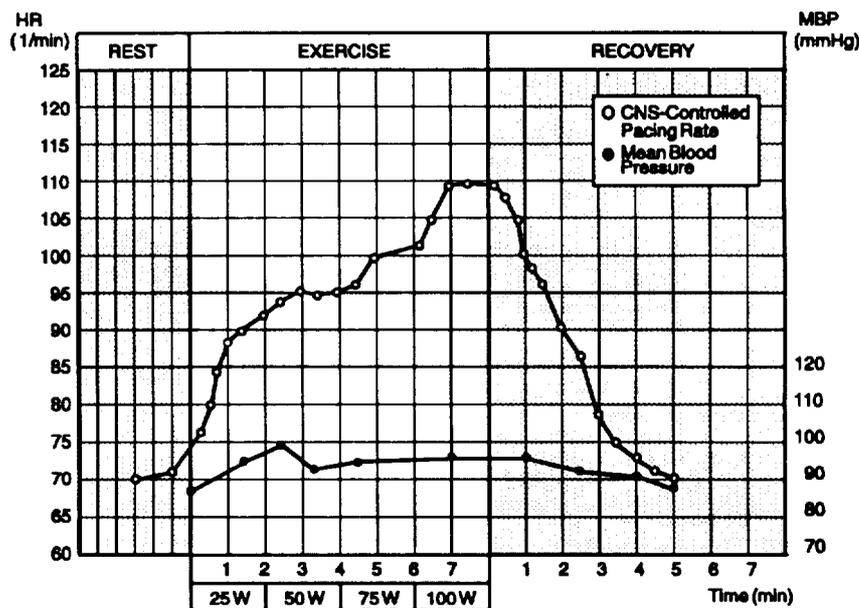


**Fig. 3** *Time courses of ANS-controlled pacing rate and mean arterial pressure, during bicycle ergometry and recovery period.*

Holter monitoring has been used for recording during daily routine activities. Periods of increased physical activity clearly coincides with periods of increased pacemaker activity (fig. 4). Individual calibration is accomplished by a rather simple procedure that takes only about half an hour. For a single chamber pacemaker, it consists of the measurement of two impedance curves under well-defined conditions, usually at rest with a pacing rate of about 70 - 75 bpm

and at about 80% of the patient's individual maximum performance capacity with a near maximal pacing rate that is tolerable for the respective patient. Those impedance curves are interrogated with the pacemaker programmer which then automatically calculates the optimal time position for the measurement interval as well as the corresponding program parameters. Care should be taken, however, that additional emotional stress during the calibration process is avoided since it might effect the ANS response.
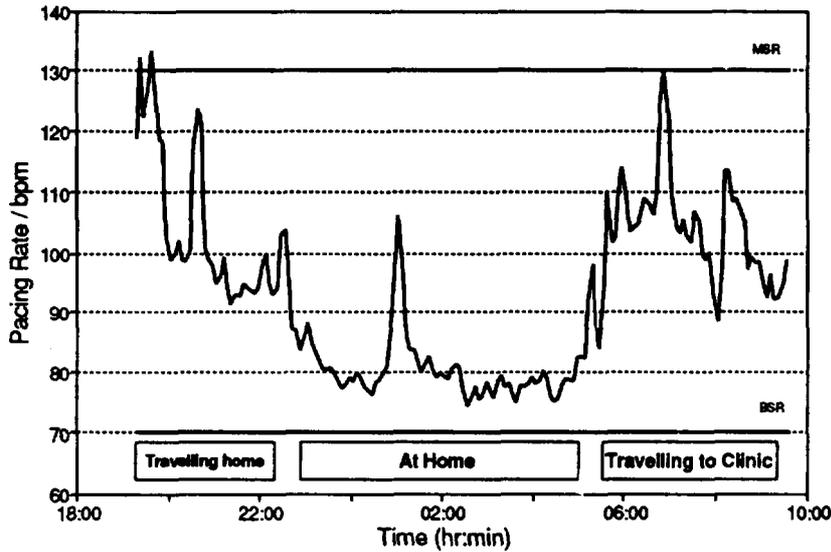


*Fig. 4*  *A 24-hour registration of the pacing rate of a patient with an implanted ANS-controlled pacemaker in the VVIR mode.*

After some time, recalibration might be necessary in cases of improved heart performance due to rate-responsive pacing [5]. Long-term monitoring of patients that are supplied with the VVI/R pacemaker has revealed such an improvement as indicated by a change in the NYHA classification index (fig. 5).
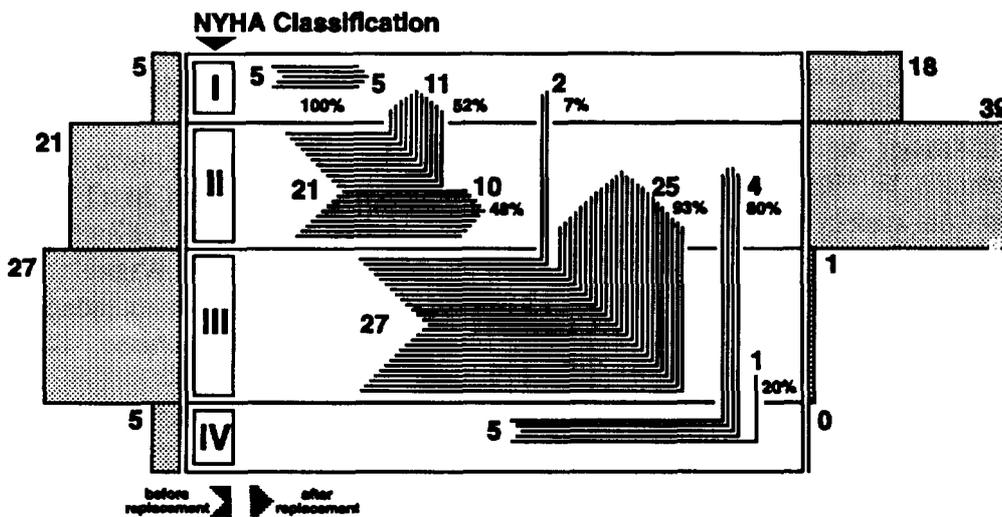


*Fig. 5*  *NYHA classification of those patients where a conventional single-chamber pacemaker was replaced by an ANS system. The average NYHA index clearly decreases, indicating an improved quality of life.*

83

Experience has shown that the impedance curves become stable within six months after the first calibration, provided the patient's condition is stable. Although a small number of patients did not significantly benefit by the ANS-controlled rate-responsive system as demonstrated by a constant NYHA classification index, nearly all patients report on a significant improvement in quality of daily life.

## Discussion

Modern pacemaker technology makes the adaptation of pacing rate to hemodynamic requirements possible. Among the different concepts, the methodological approach of restoring the closed-loop control is superior with regard to its physiological control quality [1,6]. Neither the forward-consideration of disturbance variables like motion or acceleration nor the open-loop control using such internal, non-cardiac state variables like ventilation, oxygen uptake or body temperature is really a physiological approach. They are single-variable concepts and based on control characteristics that describe the relation between the respective variable and heart rate. In a closed-loop control system with an unequivocal relation between the measured variable and the pacing rate as controlled quantity, it is not necessary to know the characteristics but to adjust the pacing rate in accordance with the target function. The actual target function is obtained for each patient by a simple calibration procedure and defining the Inotropic Index.

The multi-center study has proven the clinical efficiency of this approach. It is possible to use the same electrode for sensing, stimulation, and unipolar impedance measurement. Many electrodes with some of them implanted for years from different producers have been tested. Difficulties may arise only in those electrodes with a very long time constant for repolarisation after stimulation. With regard to this problem, further progress can be expected from new electrodes with a very short repolarisation time constant [2].

Experience has shown that improvement in the pumping performance of the heart due to properly adjusted rate-responsive pacing may require recalibration. Readjustment of the program parameters is possible without any difficulty by the same calibration procedure. Such an improvement has been revealed by long-term monitoring of patients with the VVI/R pacemaker and is in accordance with a change in the NYHA classification index.

## Literature

[1] Alt E, Barold S, Stangl K: Rate adaptive cardiac pacing. Berlin - Heidelberg - New York, Springer-Verlag (1993).

[2] Bolz A, Hubmann M, Hardt R, Riedmüller J, Schaldach M (1993): Low polarization pacing lead for detecting the ventricular evoked response. Medical Progress through Technology 19: 129 - 137.

[3] Hutten H, Schaldach M (1993): Rate-responsive pacing based on sympathetic activity. Med.&Biol. Eng. & Comput. 13: 108 - 114.

[4] Pichlmaier A, Braile D, Ebner E, Greco O, Hutten H, Von Knorre G, Niederlag W, Rentsch W, Volkmann H, Weber D, Wunderlich E, Schaldach M (1992): Autonomic nervous system controlled closed loop cardiac pacing. PACE 15: 1787 - 1791.

[5] Pichlmaier A, Hutten H, Ebner E, Göhl K, Merkely B, Res J, Weber D, Witte J, Van Woersem J, Schaldach M (1993): Physiologisch geregelte Frequenzadapatation mittels unipolarer intrakardialer Impedanzmessung - Klinische Ergebnisse. Biomedizinische Technik 38 (Ergänzungsband): 405 - 406.

[6] Schaldach M, Hutten H (1990): A physiological approach to different concepts of rate adaptive pacing. Medical Progress through Technology 16: 235 - 246.

[7] Schaldach M (1991): Microelectronics in implantable cardiovascular devices. Proc. 13th Ann. Int. Conf. IEEE Eng. in Med.& Biol. Soc., Orlando, Florida, 31st Oct.- 3rd Nov., 5/5: 2120 - 2122.

[8] Schaldach M: Electrotherapy of the heart. Berlin - Heidelberg - New York, Springer-Verlag (1992).

# A Large Arterial Vessels Investigation by Haemodynamic Analysis of Pressure - Diameter Relationships

Alexander Gutman , Ph.D.[1)]   and  Isaak Manevich , Ph.D.[2)]

1)  ALVIM R&D Ltd,  P.O.B. 801   Jerusalem  91007   Israel
2)  Jerusalem College of Technology,  21  Havaad  Haleumi  St.
P.O.B.  16031   Jerusalem  91160   Israel

The equations for the blood flow in a distensible tube are considered. Tapering and branching nature of the arterial vessel is taken into account. The cross-sectional area - pressure dependence is suggested in the form of a logistic function which well describes experimental curves. The 'haemodynamic' correction to the above dependence is shown to play a considerable role when collapse of the blood vessel emerges.

Atherosclerosis and its consequences, such as myocardial infarction and peripheral vascular obstraction remains a grave unsolved challenge in modern medicine. Non-invasive and quantitative determination of the mechanical properties of the vascular system may provide physicians with the key information for the diagnosis, treatment and prevention of these diseases.

A simple and reliable techique allowing an ambulatory measuring of the blood pressure, of the elastic properties of the arterial walls, and of the internal diameters of the large arteries for the upper and lower limbs is done by using a plethysmograph and an inflatable limb-circling cuff. From the practical point of view, there is a serious draw-back to this method, namely that the mechanical properties of arteries vary markedly with transmural pressure.

The knowledge of the analytical dependence for pressure-diameter relationships of large arterial vessels is therefore of crucial importance for the further development of the approach mentioned above.

With this in mind we have tried to solve the equations for the flow of an incompressible fluid in a distensible tube. We also take the blood to be homogeneous with density $\rho$ and invisced. The blood viscosity can be neglected since in the large arterial vessels the velocity profiles are approximately flat. In these circumstances a one-dimensional theory can be applied [1]. This means that the motion of the blood can be represented by averaged over the cross-section velocity $U(x,t)$, and that the excess pressure is $p = p(x,t)$ where $x$ is the distance along the tube and $t$ is the time.

A blood vessel is modelled as a straight tube whose cross-sectional area $A = A(x,t)$ varies continuosly with distance along it. With reference to Fig. 1, we consider the geometry of an artery to be defined by frustum of a cone with the cross-sectional area $A(x) = \pi r^2(x)$ where $r(x)$ is the radius.

$$r(x) = r_2 - \alpha x \quad , \qquad \alpha = \frac{r_2 - r_1}{l} \tag{1}$$

Here $l$ is the length of the tube, $r_1$ and $r_2$ are the radii of the cross-sectional areas of the tube at downstream and upstream ends of it, respectively.
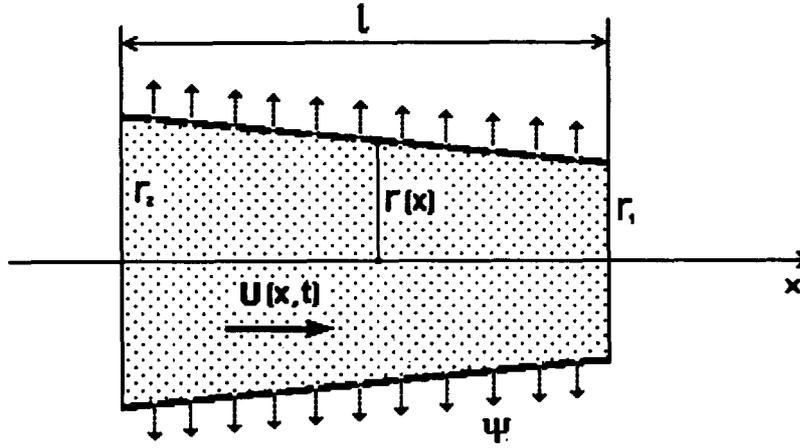
Fig. 1. Tapered vessel with distributed outflow $\psi$ across porous walls.

The continuous change in cross-sectional area $A(x)$ is natural to blood vessels and is related to the fact that a significant portion of the blood flow leaves the artery through the branches in a distributed manner. This situation is simulated by letting the tube have 'permeable' walls with a voumetric outflow along the vessel [2].

The contnuity equation in that tube is

$$\frac{\partial A}{\partial t}+\frac{\partial (AU)}{\partial x}+\psi=0 \qquad (2)$$

where $\psi$ denotes the outflow across walls per unit length of the vessel. The momentum equation for the blood flow in the tube

$$\frac{\partial U}{\partial t}+U\frac{\partial U}{\partial x}+\frac{1}{\rho}\frac{\partial p}{\partial x}=G(x,t) \qquad (3)$$

Here $G$ is the body force resulting from reaction forces of the tube's wall, as well as superposed forces of gravity, acceleration and/or inflatable cuff.

As was noted above, we assume that incompressibility condition is applied: $\mathrm{div}\,U = 0$, that is $\frac{\partial U}{\partial x} = 0$ and $U$ is function only of $A$. In these circumstances equation (2) reduces to

$$\frac{\partial A}{\partial t}+U\frac{\partial A}{\partial x}+\psi=0 \qquad (4)$$

Multiplying (4) by $U'(A)$ we arrive at

$$\frac{\partial U}{\partial t}+U(A)\cdot\frac{\partial U}{\partial x}+\psi\cdot U'(A)=0 \qquad (5)$$

Combining (2) and (5), we have

$$\frac{1}{\rho}\frac{\partial p}{\partial x}=G(x,t)+\psi\cdot U'(A) \qquad (6)$$

If we ignore the outflow term $\psi$, equation (4) reduces to a homogeneous differential equation

$$\frac{\partial U}{\partial t}+U(A)\frac{\partial U}{\partial x}=0 \qquad (7)$$

which may be solved by perturbation theory. Suppose $U(x,t)$ is expanded in a power series :

$$U = U_0 + \varepsilon\,U_1(x,t) + \varepsilon^2 U_2(x,t) + \ .\ .\ .\ + \varepsilon^n U_n(x,t) + \ .\ .\ . \qquad (8)$$

Substitute (8) into (7) and equate like powers of $\varepsilon$. Thus we have

$$\frac{\partial U_1}{\partial t} + U_0 \frac{\partial U_1}{\partial x} = 0$$

$$\frac{\partial U_2}{\partial t} + U_0 \frac{\partial U_2}{\partial x} = -U_1 \frac{\partial U_1}{\partial x} \qquad (9)$$

$$\frac{\partial U_3}{\partial t} + U_0 \frac{\partial U_3}{\partial x} = -U_2 \frac{\partial U_1}{\partial x} - U_1 \frac{\partial U_2}{\partial x}$$

. . . . . . . . . . . . . . . . . . . . . . . .

These equations can be solved successively, noting that at every step we obtain the equation of the type

$$\frac{\partial U}{\partial t} + U_0 \frac{\partial U}{\partial x} = \varphi (x, t) \qquad (10)$$

and the function $\varphi(x,t)$ may be calculated by functions $U(x,t)$, which have been found in the antecedent steps.

Now we proceed to equation (5). Let us first assume that $\psi$ is equal to zero. Then equation (5) yields the function $G(x,t)$ assuming that the area-pressure relation, $A(p)$, is established.

Healthy blood vessels are known to possess non-linear pressure response, that is static area-pressure curves are normally $S$-shaped. The artery becomes less distensible and $\frac{\partial A}{\partial p}$ decreases as $p$ increases. On condition that external pressure, $p_c$, is superposed on the vessel and $p_c$ is more than $p$, the collapse of the tube emerges with $A \to 0$.

Several analytical dependences have been suggested [3-5] to simulate experimental $A(p)$ curves of human arteries. We confined ourselves to the selection of the desired $A(p)$ dependence by means of two criteria:

1. presence of the $S$-shaped form for the $A(p)$ curve with an inflection point at transmural pressure $p_{tm} = 0$ where $p_{tm} = p - p_c$ ;
2. the force $G(x,t)$ should answer the experimental stress-strain dependences, namely that the mechanical behavior of the vessel for small strain is essentially determined by the elastin and a maximum diameter of the tube is controlled by muscle contraction.

The above criteria prove to be satisfied by the dependence $A(p)$ of hyperbolic tangent type [3]. This dependence may be represented in the form of logistic function

$$A = \frac{a}{1 + be^{-cp}} - d \qquad (11)$$

where $a, b, c, d$ are parameters and $c > 0$, $a > 2d > 0$ for $A > 0$.
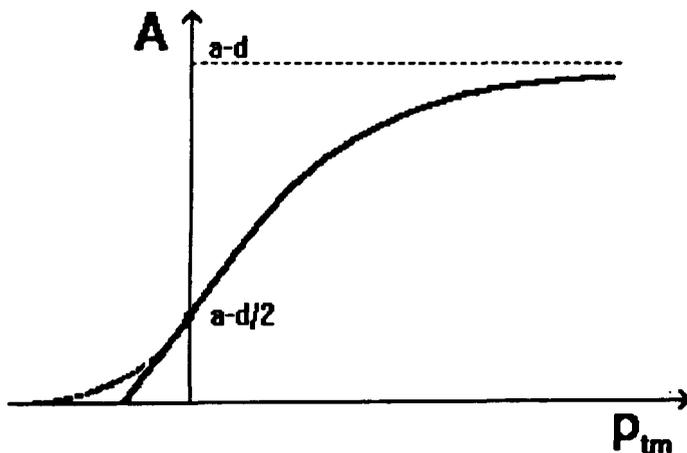


Fig. 2. Dependence of the cross-sectional area, $A$, on transmural pressure, $p_{tm}$, for the blood vessel: continuous curve, logistic function given by (11) with $A > 0$; $b = 1$; dashed curve, correction taking into account outflow.

87

If excess pressure, $p$, turns into transmural pressure, $p_{tm}$ , parameter $b$ goes to unity. This sort of logistic function is shown in Fig. 2 by a continuous curve.
Then for $G(x,t)$ we obtain

$$G(x,t) = \frac{2\pi a\, \alpha r(x)}{c[d + \pi r^2(x)][a - d - \pi r^2(x)]}$$ (12)

This shows that $G \to \infty$ as $r$ goes to its maximum value, $r_0 = \sqrt{\dfrac{a-d}{\pi}}$ .

If we take into consideration the outflow term in equation (5), we arrive at an additive correction to the pressure

$$p_{add} = \rho \cdot \int \psi U'(A)\, dx$$ (13)

We use for $\psi$ the expression which has been calculated in [6] by analysis of branching networks

$$\psi = \frac{k\alpha A U}{r_2 - \alpha x}$$ (14)

Here $k$ denotes a coefficient which depends on the ratio of the cross-sectional area of a daughter tube to that of the parent. If we take in the expansion (8) for $U(x,t)$ the two first terms and substitute it into (13) taking account of (14), we obtain $P_{add}$ of the type

$$p_{add} = O(A^{-\frac{1}{2}})$$ (15)

This 'haemodynamic' correction to the $A(p)$ relationship is likely to play a considerable role when cross-sectional area becomes small, and hence the velocity of the flow is large. It is illustrated in Fig. 2, dashed curve, as a 'tail' to the $A(p)$ curve when transmural pressure is close to zero or negative, that is when collapse of the blood vessel emerges.

## REFERENCES

1. Pedley T J 1980 *The Fluid Mechanics of Large Blood Vessels* (Cambridge: Cambridge University Press) pp. 72 - 125.
2. Streeter D D , Keitzer W F and Bohr F F 1963
   *Pulsatile Pressure and Flow through Distensible Vessels*
   Circ. Res. 13 3 - 20.
3. Vander Werff T J 1974 *Significant Parameters in Arterial Pressure and Velocity Development* J. Biomech. 7 437 - 47 .
4. van Loon P , Klip W and Bradley E L 1977
   *Length - Force and Volume - Pressure Relationships of Arteries* Biorcheology 14 181 - 201.
5. Langewouters G J , Zwart A , Busse R and Wesseling K H 1988 *Pressure - Diameter Relationships of Segments of Human Finger Arteries* Clin. Phys. Physiol. Meas. 7 43 - 55.
6. Skalak R and Stathis T 1966 *A Porous Tapered Elastic Tube Model of a Vascular Bed* In Biomechanics Symposium,
   ed. Y C Fung, American Society of Mechanical Engineering , NY.

# ACCURACY OF EPICARDIAL MAPS

Alina Czerwińska
Institute of Biocybernetics and Biomedical Engineering PAN
Twarda 55, 00-818 Warsaw, Poland

ABSTRACT: In the paper the problem of appointing of epicardial maps has been presented. The sources of errors of constructed model of torso and sources of errors of computer simulation have been discussed. On the example of unitary distribution method and "layer by layer" method the formulas for estimation of epicardial potential errors depending on applied discretization of torso region have been derived. Above formulas have been confirmed by results of numerical experiments.

## I. INTRODUCTION

Currently one observes intensive works on non-invasive diagnostic methods concerning social diseases, in particular heart diseases. One of research tendencies giving new information in cardiac diagnostics domain are the investigations, which result in appointing epicardial maps from maps of potentials measured on torso surface (i.e. Body Surface Potential Maps BSPM). The epicardial maps give additional information about electrical activity of heart. In particular, this information is useful in diagnosis of such pathologies as arhythmy and preinfarct conditions.

With the problem of appointing the epicardial maps are engaged many researches in the world, e.g. Rudy [1], Yamashita [2] and others.

To find the epicardial maps on basis of BSPM the inverse boundary - value problem is being solved. This problem is formulated as problem of electrical cardiac field. In field classification the investigated field is the source less field described by following equations (meaning of symbols according to Fig.1):

$$\text{div}\left[\alpha(x) \ \text{grad} \ \phi(x)\right] = 0 \qquad \text{for } x \in \Omega \tag{1}$$

$$n \ \alpha(x) \ \text{grad} \ \phi(x) = 0 \tag{2}$$

$$\phi(x)\Big|_S = \phi(S) \qquad\qquad \text{on } S \tag{3}$$

$$\phi(x)\Big|_H = \phi(H) \qquad \text{on } H \tag{4}$$

where: $\phi(x)$ cardiac potential function,
$\alpha(x)$ conductivity function,
$\Omega$ chest region,
$S$ torso surface,
$H$ heart surface.
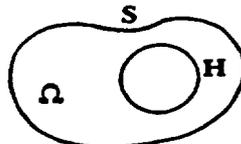


Fig.1 Geometry of investigation region.

## II. SOURCES OF ERRORS OF EPICARDIAL POTENTIALS PROBLEM

Error of calculated epicardial maps depends on:

1. accuracy of construction of chest model,
2. accuracy of measurement of BSPM,
3. accuracy of numerical experiment realized by computer simulation.

Ad 1. To solve the inverse problem the torso model is needed. To construct the torso model it is necessary to know dimensions of chest and organs inside it and locations of individual organs inside the chest. Besides, the conductivity parameters of individual organs should be known. Measurements of dimensions of chest and body organs are made by means of CT, but conductivity parameters are taken from respective tables (e.g. Rush tables). Perfect solution of the problem of finding the conductivity parameters will be applying of Electrical Impedance Tomography (EIT).

Ad 2. Measurements of BSPM are. made by special system; currently very often the system HP-7100 of Fukuda Denshi is used. The error of measured potentials on the chest by it is several microvolts.

Ad 3. There are two sources of error introduced by numerical realization:
- the replacement of partial differential equations by finite difference equations, it means the replacement of continuous region by discrete region,
- computer numerical calculations.

In this paper the sources of above errors will be considered. It will be disscused relation between final results – epicardial potential maps and assumed discretization of investigated region. This problem will be performed on example of two elaborated methods: unitary distributions method and layer by layer method.

## III. METHODS

In our investigations of electrical heart activity two methods have been applied: unitary distribution method, and layer by layer method.

The idea of the unitary distribution method will be presented on the example of homogeneous, two-dimensional region. Assume the region $\Omega$ with boundaries H and S (Fig.2a), in which the equations (1,2,3,4) are fulfilled.
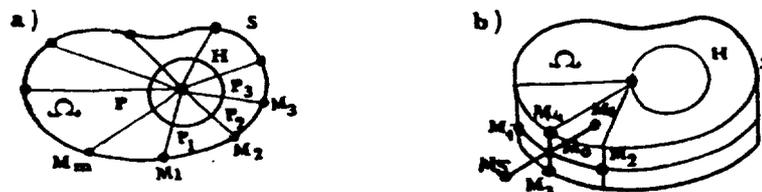


Fig.2 Division of region

For data $\phi(S)$ and $\delta\phi(S)=0$ the $\phi(H)$ have to be appointed. To solve this task and to carry the numerical experiment it is necessary to approximate the continuous region with the discrete elements. At that time the equation (1) will be approximated with the adequate difference equation.

After fitting assumed discrete net to the region $\Omega$, the knots $P_i$ on the boundary H and knots $M_i$ on the boundary S are obtained.

In the model created such way the unitary distribution has to be calculated, i.e. the boundary value problem with following conditions for n points belonging to the boundary H has to be solved:

$$V\Big|_H = \begin{cases} 1 & \text{for } P_0 \\ 0 & \text{for } P_i^0/P_0 \end{cases} \quad \text{and} \quad \frac{\delta V}{\delta n}\Big|_S = 0 \qquad (5)$$

where $i = 1,2\ldots n$

The response for above distributions are potentials $V_{mn}$ in points $M_j$, where $j=1,2,\ldots,m$. The matrix $V$ is built from potentials $V_{mn}$.

Applying the superposition rule one can express the model value of function $\phi_j(M)$ in vector form:

$$V \cdot \phi(P) = \phi(M) \qquad (6)$$

After calculation of above algebraic set we obtain the model values $\phi(P)$. These values, realized as the boundary conditions on the boundary H of investigated model will involve creating of potentials $\phi(M)$ on the boundary S.

The idea of "layer by layer" method works on the principle of solution, at the same time, the equations for knots lying on one layer. The first layer is the exterior layer of surface S, for which the normal derivative equals zero. For each knot of individual layer the following equation is calculated, assuming the homogeneous net (Fig.2b):

$$\sum_{n=1}^{6} \alpha_{0i}\Big[\phi(M_i) - \phi(M_0)\Big] = 0 \qquad (7)$$

In the process of calculating of above equation for individual knots and for consecutive layers, the surface H is achieved. The values of potentials obtained on this surface are the responses of inverse problem.

## IV. ESTIMATION OF ERROR OF EPICARDIAL POTENTIALS

In the process of application of unitary distribution method the algebraic set (6) is being solved. If potentials of torso surface map are measured with error $\delta\phi(M)$ then right side of equation (6) is disturbed and this equation takes a following form:

$$V \cdot \phi(P) = \phi(M) + \delta\phi(M) \qquad (8)$$

Solving above equation we obtain the solution $\phi(P)$. The error of this solution is estimated as follows:

$$\frac{\|\delta\phi(P)\|}{\|\phi(P)\|} \leq \frac{\|V\| \cdot \|V^{-1}\| \cdot \|\delta\phi(M)\|}{\|\phi(M)\|} = \gamma \frac{\|\delta(M)\|}{\|\phi(M)\|} \qquad (9)$$

where: $\delta\phi(P)$ - the error of vector $\phi(P)$
$\delta\phi(M)$ - the error of vector $\phi(M)$
$\gamma = \|V\| \cdot \|V^{-1}\|$ - index of condition of algebraic set.

The formula (9) has been derived in [3]. It shows, that error of epicardial potentials depends on error of torso surface potentials $\delta\phi(M)$ and on value $\gamma$ which is growing when number of elements of matrix $V$ is growing. However, one has to remember, that number $\upsilon$ of matrix elements equals the number of elements the investigated region has been divided on during the process of discretization.

In the case of application layer by layer method the estimation of error of epicardial potentials evoked by $\delta\phi(M)$

can be formulated as follows (the formula derived in [3]):

$$\frac{\delta\phi(P)}{\phi(P)} = \frac{\delta\phi(M)}{\phi(M)} \left[\frac{\sum\limits_{n=1}^{6}\alpha_n}{\alpha_6}\right]^k = \beta^k \frac{\delta\phi(M)}{\phi(M)} \qquad (10)$$

where: $\beta = \left[\sum\limits_{n=1}^{6}\alpha_n\right]/\alpha_6$ and k - number of layers.

Because of $\beta>1$ when the error of epicardial potentials $\delta\phi(M)$ increases, the number of layers is growing .

## V. EXPERIMENTAL DATA

To test the formulas (9) and (10) the numerical experiment has been carried on. As the data, the BSPM measured on human torso surface by Heart Potential Mapping system HPM-7100 have been used. Distribution of measuring electrodes is shown in Fig.3. The chest model has been constructed applying seven CT cross-sections of chest taken from anatomy atlas. In Fig.4. is shown one of them.
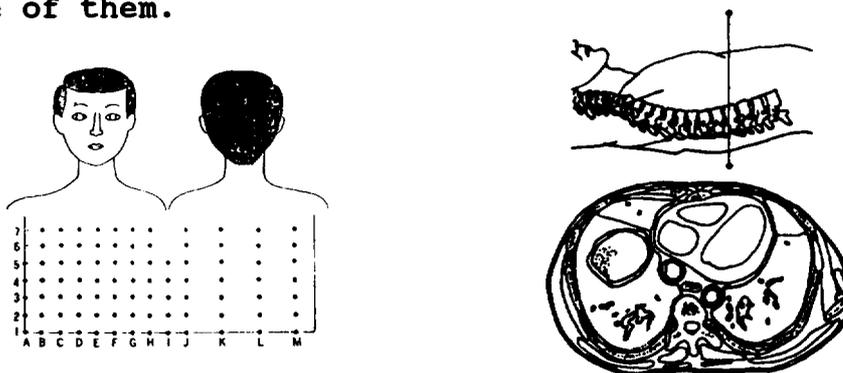


Fig.3 Distribution of electrodes   Fig.4 Cross-section of chest

## VI. EXPERIMENTS AND RESULTS

The following experiments have been done:
- for BSPM shown in Fig.5. and for chest model constructed on the basis of anatomy atlas the inverse problem has been solved for two cases of discretization of investigating region: dividing on 91 elements and dividing on 273 elements; furthermore, three cases of disturbance $\delta\phi(P)$, simulating error of measurement

$$\varepsilon = \frac{\delta\phi(M)}{\phi M)} = 1\%, \ 10\%, \ 20\%$$

have been introduced. The point of disturbance is shown in Fig.5. Results of above simulation are shown in Fig.6 and 7. These results are the same for each of two applied methods.

## VII. CONCLUSIONS

Obtained results of numerical experiments show the influence of discretization on error of epicardial potentials. Discretization of investgated region on large number of elements invokes large errors. In the case of larger disturbance on epicardial maps the additional extremes appear (see Fig.6d., Fig.7c. and Fig.7d.). It points out that models used in simulation process should be composed of small number of elements. However, this condition doesn't allow properly to take into considerations the physical parameters of chest

92

(shapes and conductivity parameters of organs of chest). For that reason it is right to use two models in iterative process of simulation. Preliminary simulation is carried on applying model consisting of smaller number of elements, while in the final simulation process is applied the model which allows to take into considerations exact shapes and conductivity parameters of simulated region. Practical application of this idea appeared advantageous.
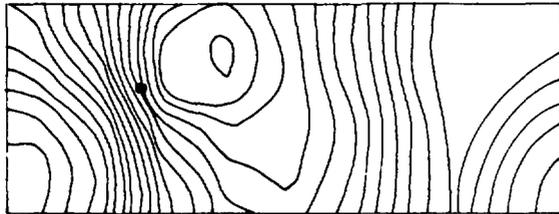


Rys.5 BSPM, ● - the point of disturbance $\delta\phi$(M)
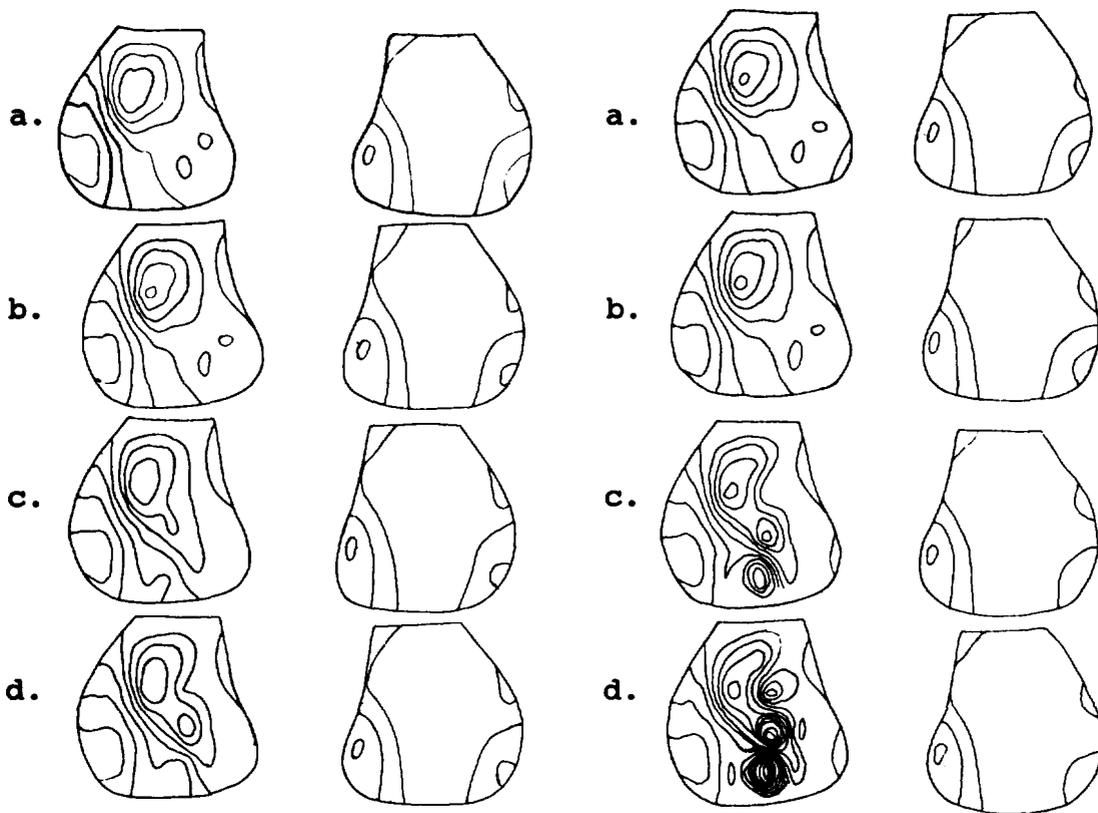


Anterior     Posterior
Fig.6 Epicardial potential
maps of 91 model elements

Anterior     Posterior
Fig.7 Epicardial potential
maps of 273 model elements

a) ε=0, b) ε=1%, c) ε=10%, d) ε=20%

## VIII. REFERENCES

1. Rudy Y.,Oster H.,"The electrocardiografic inverse problem", Critical Reviuws in Biomedical Engineering 1992, v.20.
2. Yamashita Y.,"Accuracy of epicardial potentials inversely reconstructed from body surface measurements. Proceedings of Sympozium on Body Surface Mapping, Nijmegen, The Netherlands 1988.
3. Czerwińska A.,"The methods of solving of inverse boundary problems of cardiology". Biocybernetics and Biomedical Engineering v.14 n.1,1994.

# A Thermal Image Generator Used as A Vision Prosthesis for The Blind.

*De Baetselier E., De Mey G., Van der Goten K.\*, Vandierendonck A.\**

*Department of Electronics and Information systems, University of Gent*
*St Pietersnieuwstraat 41, 9000 Gent, Belgium*
*\*Department of General Psychology, University of Gent*
*Henri Dunantlaan 2, 9000 Gent, Belgium*

## Summary.
Our research project aims the construction of a "Thermal Image Generator"(TIG). Such a tool can easily be constructed using thick film resistors, printed on a ceramic substrate. By selectively heating the resistors (hence by applying power to them) we can create a thermal pattern or temperature distribution on the substrate (text and figures). The blind can sense this pattern using his fingertips.
The development of a TIG is not only a problem from a technical but also from a psychological point of view. In this article, we will give both an introduction into the problems analysing the thermal behaviour of the system and introductory answers to the most preliminary psychological questions such as whether and to what extent information can be transmitted through a pattern of heated dots.

## Introduction.
Our research project aims the construction of a "Thermal Image Generator"(TIG). Such a tool can easily be constructed using thick film resistors, printed on a ceramic substrate. By selectively heating the resistors (hence by applying power to them) we can create a thermal pattern or temperature distribution on the substrate (text and figures). The blind can sense this pattern using his fingertips.
Linking the TIG with a Personal Computer is considered easy. Gearing electronics for the "thermal" screen can even be included on the same substrate as the TIG, also using thick film technology.
Of course, similar tools, such as Braille Reading Rulers, already exist. These are constructed using piezoelectric devices, pushing up a little pushrod as voltage is applied to them. In this way they form a mechanical pattern, which the blind can sense using his tactile senses. However, the state of the art electromechanical construction of these Braille Reading Rulers, makes them expensive and diminishes the reliability of these tools.
The TIG can be constructed using solid state technology only, which could highly reduce its price and improve its reliability. It would also be possible to produce a whole screen of text, even combined with graphics, instead of one line of text-only by means of the classical Rulers. The importance of this feature is proved by recent software evolutions making use of GUI's.

# Realisation.
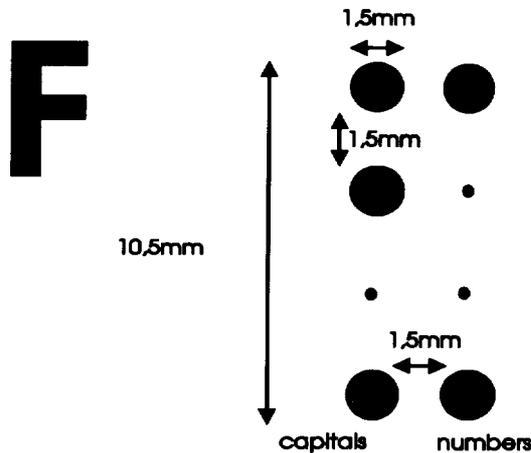


**Fig. 1** An Octobraille Character, with characteristic dimensions.
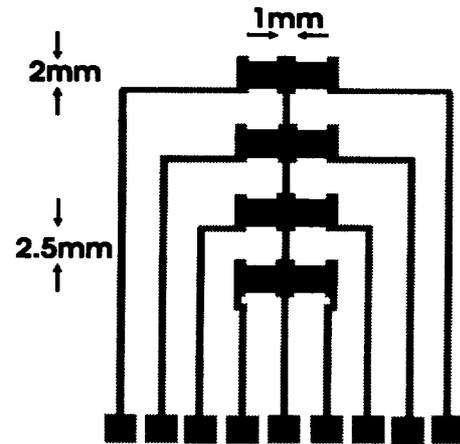


**Fig. 2** A test pattern for a TIG developed in our group. Note: Their are 8 resistors indeed, the central lead divides the rectangles into two parts.

The normal Braille system [9] is constructed using signs composed of six dots or less. Because it is impossible to construct a complete character set in this way (only 64 characters) two types of solutions exist. The first solution states that the size of the sign is limited by the dimensions of the fingertips and thus one sign can only be composed of six dots, extra characters can only be composed using a combination of two adjacent Braille signs (eg. for capitals, numbers). An other system enlarges the Braille sign at the bottom with two additional dots, thus constructing a Braille sign with a matrix of eight dots, the so called "Octobraille" (fig. 1).

Clearly the latter system has a slight disadvantage, making it difficult sensing the whole sign with a fingertip. A major advantage however is its possibility to create a set of 256 characters, which makes it particularly useful combined with computers (ASCII). The Braille sign inspired us, developing a test pattern of eight thick film printed resistors (fig. 2). While applying power to a resistor (eg. 1 W), theoretical analysis and measurements [1-5] have shown that temperatures can rise with 5°C in a few msecs (fig. 5). A provisional gearing for this device was constructed using an external power source and a 8-bit ad-on computer card (Keithly Metrabyte PIO-24). Using this device [6], we are as well able to gear the power dissipated in the thermal pixels and to detect reaction time of the person under test for a given signal.

## Thermal analysis of a TIG prototype.

Because of thick film printing technique, the thermal pixel consists of four layers (fig. 3): A resistor layer(1) in which heat is produced and underneath a thermal insulator layer(2), both printed on a ceramic substrate(3). This substrate is mounted on a heat sink(4).

Knowledge of the thermal behaviour of the system is a key problem. The aim is modelling the TIG as such, that a pixel reaches equilibrium temperature as fast as possible, and that the heat generated in this pixel undergoes a minimal spreading in the horizontal plane of the ceramic substrate.These conditions allowing optimal temperature control of the thermal pixel and high contrast of the thermal screen, are met by the model described above. Further on pixel-finger contact is studied: the thermal behaviour of a pixel printed on a ceramic substrate, covered
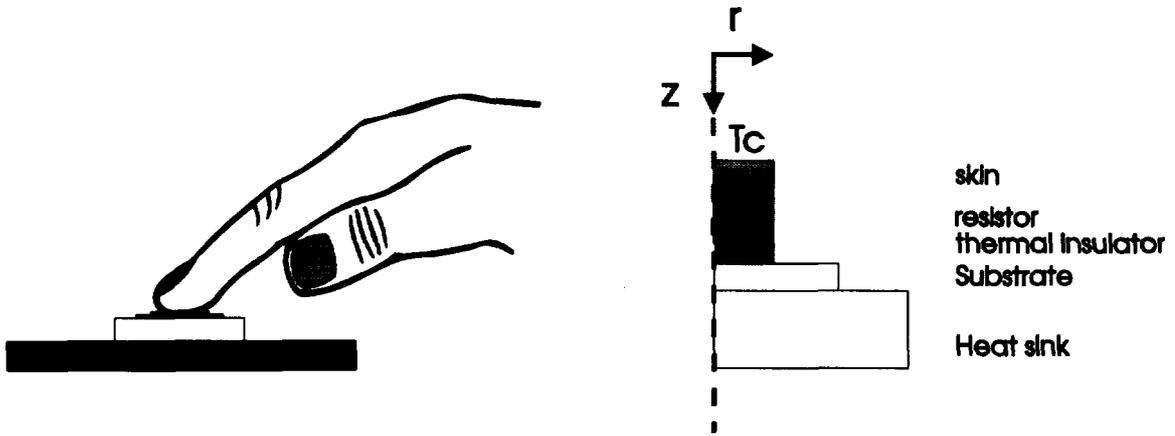
**Fig. 3** Thermal pixel making contact with finger: Thermal model.
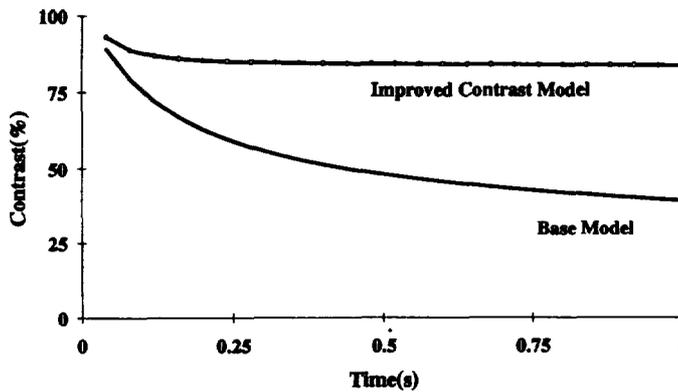


**Fig. 4** Contrast evolution versus time for two different TIG prototypes.

with a finger is different from that of a pixel submitted to natural convection [6]. While the latter system can be described theoretically (fig. 5, fig. 6), and can easily be observed with the aid of an Infrared Thermographic Camera, the first system is intricate to describe. A finger can, because of body thermoregulation, impose partially its temperature upon the thermal pixel (instead of vice versa).

The data presented here are simulation results, solving the time dependent heat equation in cylindrical coordinates, assuming radial symmetry. They were confirmed by earlier measurements [5, 6].

At first we aimed optimisation of the prototype, not taking into account any TIG-finger contact. The study compares contrast evolution versus time after a DC power onset in the pixel for the *optimised* four-layer *model* and the two-layer *base model* having only a resistor layer printed on a ceramic substrate (fig. 4).

We defined $Contrast\% = \dfrac{T(t)_{pixel\ center} - T(t)_{ref.\ point}}{T(t)_{pixel\ center} + T(t)_{ref.\ point}} \times 100\%$ , using the temperature at the

center of the pixel and the temperature of a reference point besides the pixel as input.

While the *Improved Model* reaches a stable contrast after 0.1s, the *Base Model* still shows a contrast decrease after 1s (fig. 4). The thermal insulator layer and the heat sink account for the contrast amelioration. For short time intervals after power onset (<0.1s) the thermal insulator localises heat, while for longer time intervals (0.1s) heat is removed in a direction perpendicular to the thermal screen surface. Because of the high thermal capacitance of the heat sink, an *improved* pixel also reaches thermal equilibrium relatively fast ($\approx$0.1s) (fig. 6) compared to the time needed for human heat signal detection [8] (the pixels are heated up to a maximum temperature of 45°C).
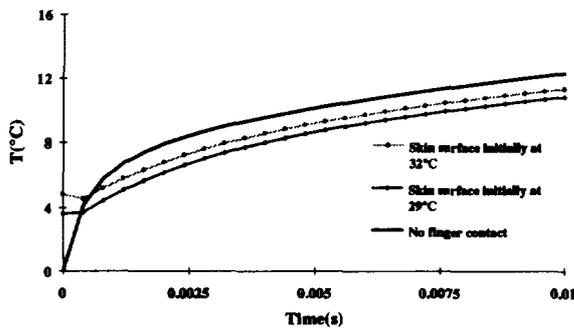
**Fig. 5** Pixel surface temperature* evolution vs. time: Showing finger-pixel contact for different external skin temperatures and a thermal pixel without finger contact. (0.01s after onset of finger contact and DC power)
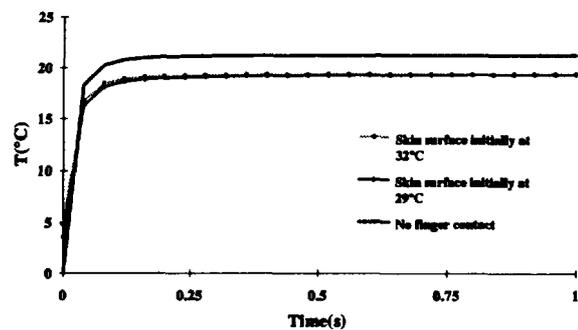
**Fig. 6** Pixel surface temperature* evolution versus time: Showing finger-pixel contact for different external skin temperatures and a thermal pixel without finger contact. (1s after onset of finger contact and DC power)

\*
All temperature evolutions relate
to ambient temperature (25°C)

When a finger touches the thermal pixel, the systems behaviour changes: Over a short time interval after finger-pixel contact the temperature evolution is not only dependent on the pixels' power dissipation but also on external skin temperature and body thermoregulation (fig. 5). At longer time intervals (fig. 6) only the influence of power dissipation and thermoregulation is left. The external skin temperature shows strong variations according to different climatological and physiological conditions [8]. Body thermoregulation, although dependent upon the same conditions varies much slower. This last feature allows at least an individual adaptation of the TIG usable over a longer time period (≈hours).

## Psychological questions concerning the use of the TIG.

The usefulness of a TIG depends upon the ability of people to interpret a thermal pattern. Therefore, questions have to be asked: What are the critical conditions for people to detect* heat? What is the distance between two pixels to be recognized as separated areas of stimulation? What role is played by the intensity of stimulation and the area of the pixel?

Answers from literature being quite poor, own fundamental research is needed. Until now this has provided us with the following [7]:

All psychological tests were performed with the device described above (fig. 1, fig. 2, fig. 3). We used printed patterns with pixels having dimensions from $0.5x0.5mm^2$ till $2x2mm^2$, with interpixel spacing from 0.5mm till 2mm.

Sensitivity for a given temperature signal increases with increasing intensity of the stimulus (power dissipation in pixel). Of course without activating pain perception!

People can discriminate between single and double stimulations on a rather high performance level. The rather odd finding that people can more easy discriminate between a single and a double stimulation as the interstimulus distance increases, can be explained by the phenomenon of *spatial summation*, which means that an increase in areal extent of the stimulus surface results in a feeling of increased intensity.

People are able to discriminate between left and right stimulations. People perform better on larger surfaces of stimulation.

Performance improves as the experiment continues. Latencies (reaction time) decrease with the

experiment in progress. These results suggest that training could be important for thermal pattern recognition.

## Conclusions.

Extensive thermal analysis is needed for further improvement of a TIG prototype. Research on thermal behaviour will emphasize both the TIG pixel and the pixel finger interaction. This last topic will need scrutinizing thermal behaviour of the human skin.

Part of the technological innovations of the TIG should come from psychological research. eg. Recently, indications exist that the thermal senses could in a way work synergistically with tactile senses. This knowledge can result in the development of a new pattern, with thermal pixels placed upon a slight elevation.

From a technological point of view, the Thermal Image Generator is a new application. Already international interest exists [10-12]. One of the major merits of the system is its inherent simplicity, which will lower its production costs and improve its reliability.

What makes our device also important is its ability to produce graphics, Icons and text simultaneously, which, considering recent software evolutions, making more use of GUI's, would enable the blind to use more recent versions of popular software packages.

## References.

1    E. Boone, G. De Mey and L. Rottiers:"Temperature distribution on ceramic substrates", Electronics Letters, 1986, vol.22, p.442-443.

2    G. Casselman and G. De Mey: "A thermal model for hybrid circuits", Hybrid Circuits, 1986, vol.9,p.9-13.

3    G. De Mey and S. Demolder: "A test substrate for thermal analysis of hybrid circuits", Hybrid Circuits, 1987, vol.12,P.43-45.

4    M. Driscart, G. De Mey and L. Rottiers:, "Simulation of the transient thermal behaviour in hybrid circuits", Sixth European Hybrid Microelectronics Conference, Bournemouth, 3-5/6/87, proceedings p.212-217.

5    G. De Mey, S. De Smet, Du Mao, E. De Baetselier, K. Vandergoten, A. Vandierendonck: "Thermal Analysis of Thick Film Hybrid Circuits", 37th International Scientific Colloqium, Technische Hochschule Ilmenau, 21-24/09/1992

6    E. De Baetselier, Du Mao, G. De Mey: "A Thick Film Resistor Circuit As A Prosthesis For Blind Persons", 9th European Hybrid Microelectronics Conference, Nice, 2-4/07/1993, proceedings p.269

7    A. Vandierendonck, K. Van der Goten: "Sensation of resistor induced warmth in blind persons", Perceptual and Motor skills , accepted for publication

8    Guyton: "Textbook of Medical Physiology 8th ed.", Saunders 1991.

9    J. Cnops, H. Lips, J. De Vilder, T. Verhulst:, "Handboek voor Brailleerders", Uitg. "De vereniging voor het Nederlandse Blindenwezen" en "Het Nationaal Comité van het Belgisch Blindenwezen".

10   "Belgische Generator voor warme Braille", NRC Handelsblad (in Dutch)

11   J. Beard:, "Hot Tips for Braille Readers", New Scientist, 16/05/92, p. 92

12   "Warme Braille", Gent-Universiteit, february 1992, (in Dutch)

# Estimation of a Simple Model of the Human Arm Kinematics

Andreja Umek, Valentina Filova, Jadran Lenarčič
Institute Jožef Stefan
Jamova 39, 61111 Ljubljana
Slovenija
*e-mail: andreja.umek@ijs.si*

### Abstract

An experimental evaluation of the mathematical model of the human arm kinematics and the human arm reachability is reported. A comparison is made between the measured human arm reachable workspace and the computed workspace. Deviations are detected especially in backward regions. It is observed that they are due to inaccuracies in representing the dependency of the joint limits on the actual values of the neighbouring joints in the shoulder complex. The analysis showed that deviations are also due to inaccuracies in specifying the path of the instantaneous center of rotation in the shoulder. To improve the model we introduced two additional dependent translational degrees of freedom in the shoulder link. They describe the variable lenght of the shoulder link that is influenced by the elevation of the arm.

## 1  Introduction

A thorough analysis of the human arm kinematics is providing information that helps toward better understanding of proportions of links and ranges of joints, as well as its motion characteristics. Investigations are implying new directions in the developement of mechanical arms and similar devices in rehabilitation and robotics, and can provide mathematical basis for the evaluation and control of human motion.

Investigations carried out in recent years in the area of mathematical modelling of the human arm kinematics motivated the development of a simplified kinematic structure which takes into acount that the calculation of the reachability of the arm heavily depends on the degrees of freedom. It must be stressed, however, that the model of the human arm kinematics is not an anathomical one. Therefore, it does not describe the motion of singular bones, but the spatial motion of the reference point on the wrist. The aim of the paper is to estimate recent results that are interpreting the human arm motion, as well as methods and devices that have been developed specially for monitoring the motion characteristics.
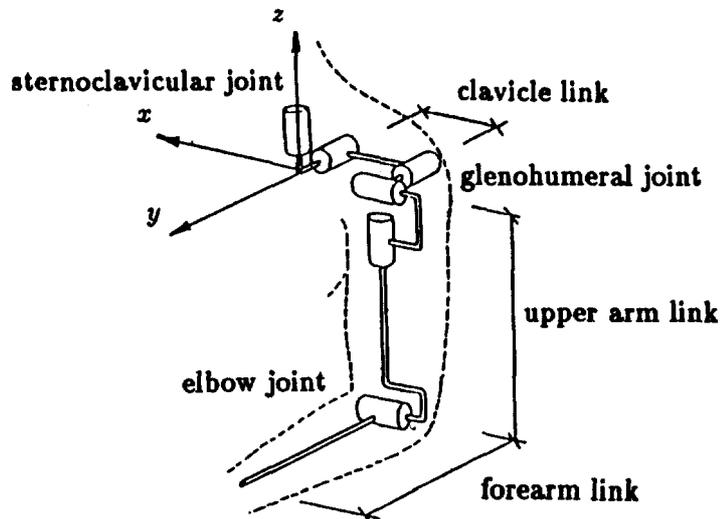
Figure 1: Kinematical model of the human arm

# 2 Mathematical Model of the Human Arm Kinematics

To monitor and evaluate the human arm characteristics, The Vicon System (Oxford Metrics) was used in this investigation. Three TV cameras were used to observe passive markers which were attached to the arm. Some selective movements such as flexion/extionsion, abduction/adduction ect., were measured.. Special mathematical treatment was applied to obtain 3D trajectories of the markers. Analysing the path of the instantaneous center of rotation of each selective movement we established that these movements do not correspond to a simple rotation about a given fixed axis in space[1]. They are composed of the rotation about a fixed axis and a corresponding translation. These translations were replaced with two rotations in the sternoclavicular joint. Considering the results the kinematical model of the human arm presented on Figure 1 was obtained.

The proposed model of the human arm is composed of six rotations arranged in the sternoclavicular joint, the glenohumeral joint, and in the elbow joint. They are connected by three links: the clavicle link, the upper arm link, and the forearm link. The wrist rotations are not included, since they do not affect the positioning of the arm in the space. The calculation of the reachable workspace of the human arm is numerically very expensive and also includes the collision detection between arm links and the body.

The calculated reachable workspace of the mechanism was compared to measurements that were carried on a real subject in selected horizontal planes (Figure 2). These planes were combined into a 3D picture of the workspace by a PC computer. The position of the human body inside the workspace is also shown on the Figure 2. It is shown from backward so that axis $\bar{y}$ lies in the anterior direction. Both workspaces were found to be quite similar. Some differencies were observed in the posterior part
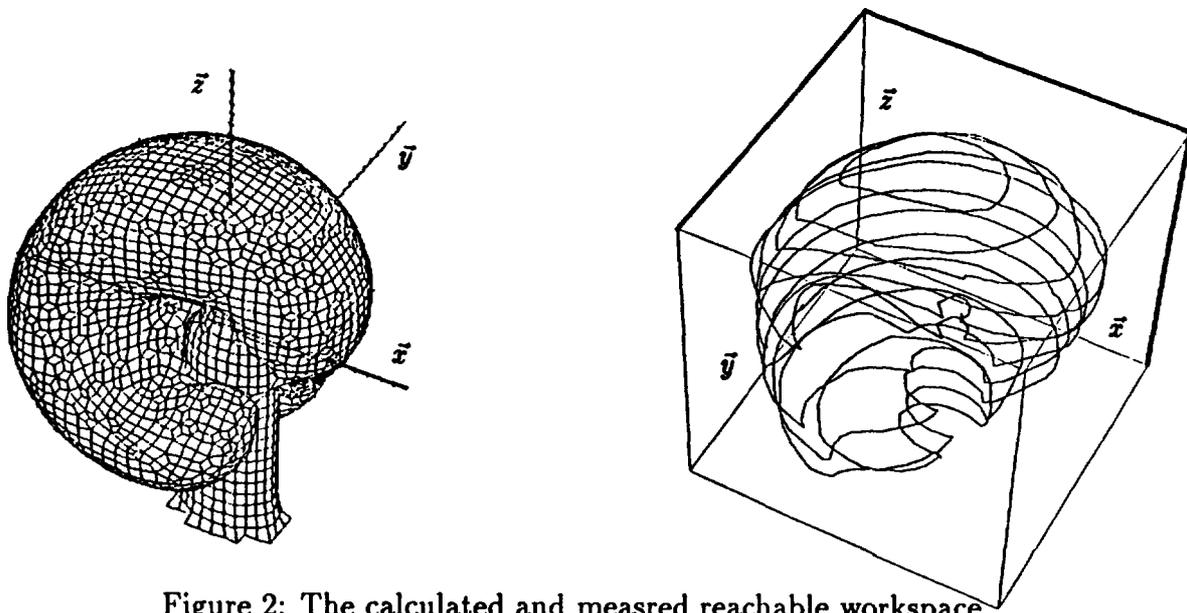
Figure 2: The calculated and measred reachable workspace

of the workspace. These are still to be corrected.

# 3 Dependancies between degrees of freedom in the glenohumeral joint

The study of the shoulder complex and its rythm is a very challenging task in the endeavor to understand the kinematics of the human arm and the human body[2, 3, 4, 5]. Since the numerical complexity of the model heavily depends on the number of degrees of freedom, we wanted to simplify the model as much as possible to make it applicable. Therefore the model is composed of only six rotations which are in mutual relationship.

Dependencies between ranges of motion of rotations in the glenohumeral joint were experimentally evaluated. The range of motion of each rotation depends on the actual values of previous rotations looking from the origin of the reference coordinate frame setteled in the sternoclavicular joint. These dependencies were modeled with quadratic functions. To test the mathematical model of the human arm kinematics such passive mechanical mechanism was designed. To each of the six rotations the potenciometer was attached and connected to PC computer to observe the motion in joints. It was shown that differencies between the calculated and the measured reachable workspace come from the simplified describtion of dependencies between rotations in the shoulder complex.

According to our findings we introduced another two translations into proposed model; along clavicle and upper arm link functionally connected with angle of elevation of the arm. These modifications gave more accurate results (Figure 3) and the model of the human arm kinematics was still kept sufficiently simple.
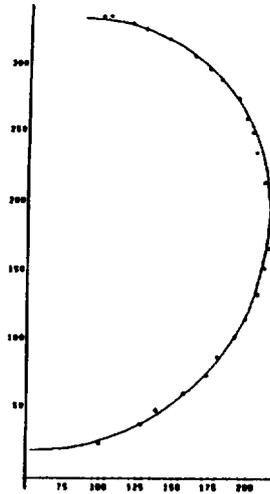
Figure 3: Comparison between calculated and measured trajectory of the abduction

# 4 Improved 3D Monitoring of Motion

A traditional approach to reconstruct motion of the human body in 3-dimensional space is to analyse trajectories of markers attached at the body. The main disadvantages of these measurements are in the deformation of the skin at which markers are attached resulting significant inaccuracies. Therefore we are trying to adopt in our investigation a new approach consisted of two levels[6]. In the first level, links of the arm are detected by fitting the arm shape with selected mathematically defined bodies, such as ellipsoids. The axes of these bodies, represented by the axes of links between which the joint angles are measured and the link length, are defined. In the second level, the human arm motion (the trajectories of joint angles) is reconstructed from the orthographic projection of the joint points on the image plane. This method assume that the structure of the kinematic model, the corresponding link lengths, as well as the starting position of the arm (in order to exclude multiple solutions) is known.

# 5 Conclusion

Through the analysis of some selective movements the mathematical model of the human arm kinematics was developed. By comparison of the measured and calculated reachable workspace of the arm it was shown that the model can be considered as a good approximation of human motion characteristics. Main disadvantages of the presented mechanical approach are related to the numerical complexity of computations that request simplifications in the mathematical model, and on the other hand to the difficulties of the proper measuring of the human motion, both resulting in inaccuracies and approximations. However, conclusions obtained by these studies are very close to the real motion of the human arm in performing different movements. Applications can be foreseen in rehabilitation, particularly in developing quantitative methods for the evaluation of motion capabilities during rehabilitation procedure.

Reconstructed 3D motion
of the human arm
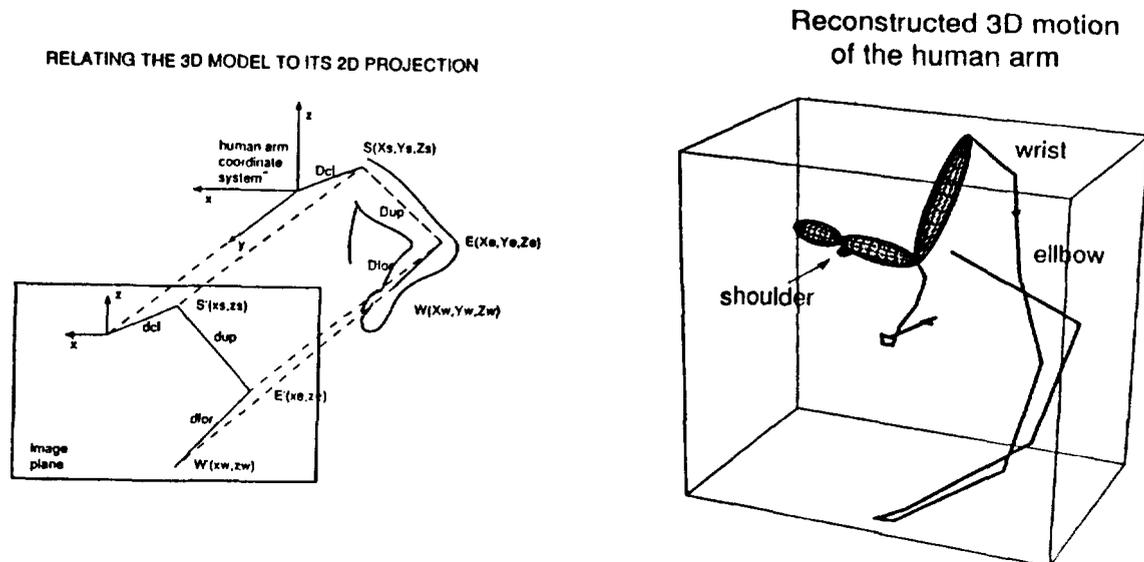


Figure 4: An orthographic projection of the 3D model and a given sequence of motion

# References

[1] A. Umek, J. Lenarčič, Recent Results in Evaluation of Human Arm Workspace, *Proc. of ICAR '91*, Pisa, Italy 1991

[2] V.T. Inman, M. Saunders, L.C. Abbott, Observation on the Function of the Shoulder Joint, *J. of Bone and Joint Surgery*, vol. 42, no. 1, pp. 1-30, 1944

[3] A.E. Engin, Kinematics of Human Shoulder Motion, in V.C. Mow, A. Rotcliffe, S.L.-Y. Woo(ed) *Biomechanics of Diarthrotrial Joints*, Springer-Verlag, New York, 1990

[4] Z. Dvir, N. Berme, The Shoulder Complex in Elevation on the Arm: A Mechanism Approach, *J. of Biomechanics*, vol. 11, pp. 219-225, 1987

[5] C. Högfors, B. Peterson, G. Sigholm, P. Herberts, Biomechanical Model of the Human Shoulder Joint - II.The Shoulder Rhytem, *Journal of Biomechanics*, vol. 24, no. 8, pp. 699-709, 1991

[6] V. Filova, Franc Solina, Jadran Lenarčič, Model-based reconstruction of 3D human arm motion from a monocular image sequence, *Proc. 1st Czech Pattern Recognition Workshop CPRW'93*, Temešvar u Pisku,1993, str.137-143.

# Artificial Intelligence and Expert Systems I

# A Fuzzy Logic Expert System for the Diagnosis of Thyroid Gland Function

Felix Grimm, Xavier Fábregas, Horst Bunke, Thomas Hänni

Institut für Informatik und angewandte Mathematik, Universität Bern

Länggassstrasse 51, CH-3012 Bern, Switzerland

{grimm, fabregas, bunke}@iam.unibe.ch, Fax: ++41-31-6313965

## Abstract

An expert system for the interpretation of the results of various laboratory tests carried out in order to determine the function of the thyroid gland is presented. The medical knowledge is modeled in such a way that a computer is able to infer conclusions from it. A particular feature of this expert system is the use of fuzzy set theory in order to handle uncertain knowledge and inaccurate input data. First, we show a solution in which the laboratory results are modeled by means of fuzzy sets. Every laboratory result is considered independently. Then, we present an alternative solution based on a decision tree model combined with methods of fuzzy logic. Both approaches have been applied to a series of real patient cases. While the pure fuzzy logic approach showed correct results in about 81% of the cases, the quality of the expert system has been significantly improved in the combined approach where about 92% of the cases were correctly diagnosed.

## 1 Introduction

Expert systems are computer programs that are able to perform the task of an expert in a specific domain, e.g. medical diagnosis. As the expert system technology was born in the early seventies as a main application of artificial intelligence (AI), medical expert systems were among the first systems of greater interest [1, 2, 3].

One great challenge in the expert systems area and especially in the medical field is the handling of uncertain knowledge. Many symptoms cannot be observed objectively, and the relation between symptoms and diseases is often unknown or ambiguous. Special techniques are needed in order to treat this kind of knowledge. A well-known approach is the use of certainty factors [1] that were first applied in a medical expert system. Various other methods have also been investigated, for example, the Dempster-Shafer theory [4, 5] and fuzzy logic [6, 7, 8].

This paper describes a technique for reasoning with uncertain facts in an expert system. This technique is applied to a component of a larger diagnosis system for thyroid gland diseases that is being developed at our institute in cooperation with the nuclear medicine department of the university hospital in Bern [9].

In order to determine the function of the thyroid gland, laboratory results of hormone concentrations in the blood are used. Crisp borders are often used for the interpretation of the numerical laboratory results which may be problematic if values near a border are observed. In order to overcome this problem and represent various possible interpretations of the same range of a test, fuzzy logic is used to deal with the numerical laboratory results. Fuzzy logic supports non-crisp sets. In this way, observed values (especially such ones near a border) may belong to a certain degree $\mu_A$ to a set $A$ (e.g. normal function) and to a degree $\mu_B$ to a set $B$ (e.g. malfunction).

The next section briefly presents the medical knowledge about the function of the thyroid gland with respect to hormone concentration in the blood. Section 3 describes the fuzzy expert system for the diagnosis of the function of the thyroid gland. In Section 4 test results of the system are reported and commented. Concluding remarks are given in Section 5.

# 2 Medical Knowledge: Thyroid Function

The function of the thyroid gland is an important part of an overall diagnosis which considers also the morphology, histology, possible tumors etc. The function indicates whether the gland produces an appropriate quantity of thyroid hormones. Therefore, overfunction (Hyperthyroidism), normal function (Euthyroidism), and underfunction (Hypothyroidism) are distinguished. For over- and underfunction a more refined diagnosis containing additional information about the degree of malfunction, i.e. overt or latent, is often required. Furthermore, physicians also consider the possible cause of underfunction by classifying it into primary or secondary underfunction.

The production of hormones by the thyroid gland is controlled by a negative feedback that is regulated by the Hypothalamus in the brain. Depending on the thyroid hormone concentration in the blood, the Hypothalamus releases the so-called TSH (thyroidea stimulating hormone) in order to stimulate the hormone production of the thyroid gland. An underfunction of the thyroid gland may have one of two causes. The first one (primary underfunction) is that the thyroid gland itself has a defect and doesn't produce enough hormones. The second cause (secondary underfunction) is a too low presence of TSH (i.e. caused by the Hypothalamus).

A precise diagnosis of thyroid function is possible by measuring the concentration of the different hormones which are produced by or stimulating the thyroid gland. The different hormones are measured by applying various tests. The evaluation of these tests can be represented by means of the decision tree in Fig. 1. This decision tree reflects the strategic knowledge of a medical expert. According to this strategy, the TSH (spontaneous) test is interpreted first. If Euthyroidism can be clearly concluded, no other test is needed. An application of TRH and repeated TSH concentration measurement can be used for further differentiation. The RIA-$T_3$ and RIA-$T_4$ tests are needed for a refined diagnosis (degree of malfunction, detection of secondary Hypothyroidism).
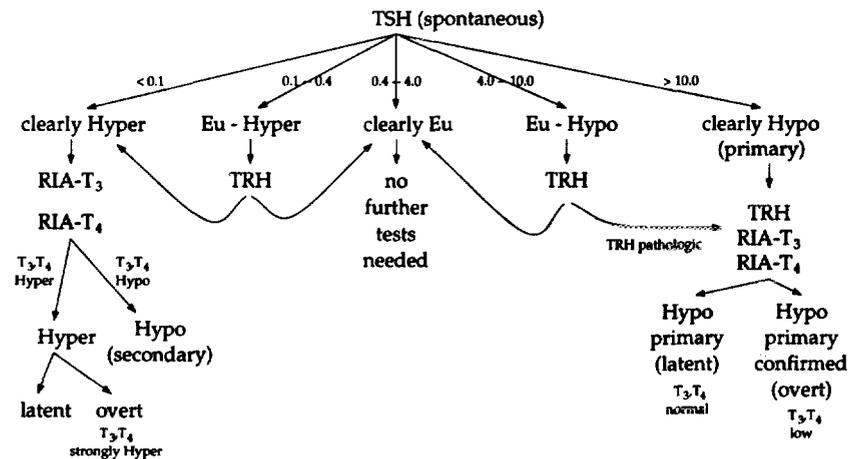


Figure 1: Decision tree

# 3 Fuzzy Expert System

An expert system for the diagnosis of thyroid gland function has been developed based on the medical knowledge presented in the previous section. This system uses laboratory test data as main input and classifies the thyroid gland's function as Eu-, Hyper-, or Hypothyroidism. Furthermore, a refined diagnosis (latent or overt, type of Hypothyroidism) is also determined.

## 3.1 First approach

The range of the possible values of a lab test for the determination of thyroid function can be subdivided into three subranges (normal, decreased, and increased concentration). The range 0.3 - 4.0 $\mu U/ml$ is usually considered as normal TSH concentration. A decreased TSH value most likely points to Hyperthyroidism and an increased one to Hypothyroidism. However, a strict classification based on the

given ranges alone leads to a fundamental problem. E.g. a TSH value of 0.31 would be interpreted as Euthyroidism and one of *0.29* as Hyperthyroidism. On the other hand, a human expert would consider Eu- and Hyperthyroidism as more or less equally possible in such cases and arrange additional tests. In order to overcome this problem, the evaluation of the lab tests is modeled by means of fuzzy set theory. Each lab test is described by a so-called *linguistic variable* consisting of several *fuzzy sets*. In the case of the TSH test, a linguistic variable *TSH* with the three fuzzy sets *decreased*, *normal*, and *increased* is used. Each fuzzy set is defined by means of a membership function as shown in Fig. 2. In this way, a TSH value of 0.31 (e.g.) is interpreted to belong to a degree of 0.45 to *decreased* and to a degree of 0.55 to *normal*. The determination of the degree of membership of a given observed value is called *fuzzification*.
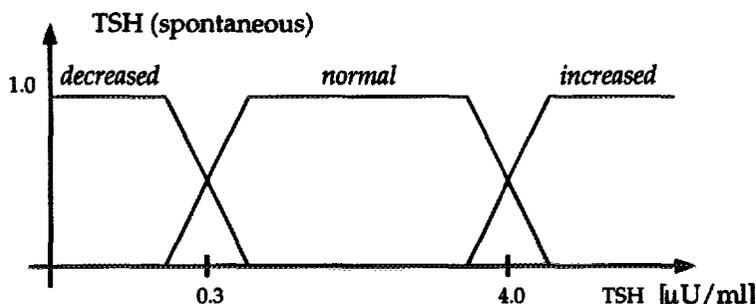


Figure 2: Membership functions for the TSH test

Rules including linguistic variables and fuzzy sets are used in order to interpret the observed test results and assign evidence to one or more diagnoses; e.g.:

**IF** *TSH = decreased* **THEN** *Hyper = probable* **AND** *Eu = possible* **AND** *Hypo = possible*

**IF** *TSH = normal* **THEN** *Eu = certain* **AND** *Hyper = improbable* **AND** *Hypo = improbable*

**IF** *TSH = increased* **THEN** *Hypo = probable*

In the above rules, *Hyper*, *Eu*, and *Hypo* denote linguistic variables describing the resulting main diagnosis. Each one consists of the five fuzzy sets *impossible*, *improbable*, *possible*, *probable*, and *certain* (see also Fig. 3). The first two rules obviously support more than one diagnosis, however to different degrees.

The evaluation of the rules is done by means of *max-min-inference* [7]. In our example (TSH=0.31), the premise of the first rule is fulfilled to a degree of 0.45. The degree of 0.45 is propagated to the right hand side of the rule, thus leading to the conclusions *Hyper = probable* to a degree of 0.45, *Eu = possible* (degree 0.45), and *Hypo = possible* (degree 0.45). Analogously, *Eu = certain* (0.55), *Hyper = improbable* (0.55), and *Hypo = improbable* (0.55) is concluded by the second rule whereas the third rule cannot be applied as its premise is assigned a zero evidence. If more than one rule derived the same fuzzy set (which is not the case in our example), the maximum evidence obtained would be assigned. Furthermore, the fuzzy set of each conclusion is combined with its degree by taking pointwise the minimum value of both, e.g.:

$$\mu_{Hyper.probable}(x) = min\{\mu_{probable}(x), 0.45\}$$

In oder to compute a measure of belief for a possible diagnosis, the center of gravitiy ($x$-coordinate) of the area defined by the concluded fuzzy sets of the respective linguistic variable is determined. This process is called *defuzzification*. In this way, *Eu* is assigned an evidence of 84% in our example, *Hyper* 47%, and *Hypo* 24%.

Similar linguistic variables, fuzzy sets, and rules have been defined for all lab tests supported by the actual version of our system (TSH spontaneous, TSH after application of TRH, TRH test, RIA-T$_3$, RIA-T$_4$) and all possible diagnoses and subdiagnoses. Each test is evaluated individually in analogy to the TSH test described above. In order to combine the results of different tests, the maximum degree is used if more than one test computes a membership degree for the same fuzzy set. After evaluation of all test results and corresponding rules, the most evident diagnoses are determined by means of defuzzification as explained above.

114
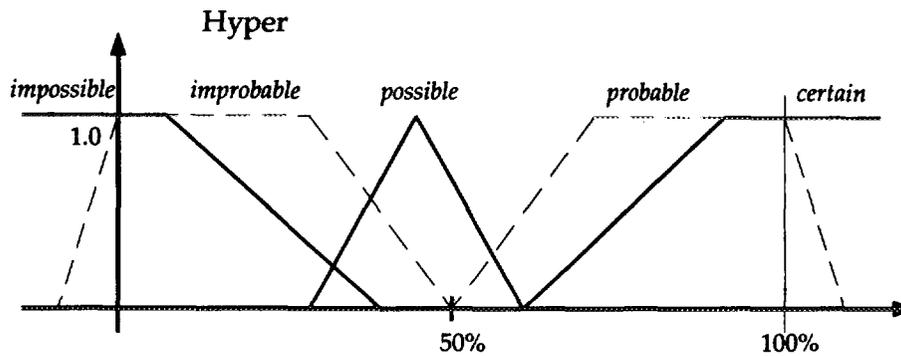
Figure 3: Fuzzy sets of the diagnosis *Hyper*

## 3.2 Second approach

The application of the first approach has shown useful diagnostic results. It has been observed, however, that the individual evaluation of the tests doesn't correspond exactly to the way physicians are making their decision. They rather make use of the decision tree of Fig. 1 in which the principal diagnosis (Eu, Hyper, Hypo) is exclusively determined by the TSH and TRH tests whereas the $T_3$ and $T_4$ tests are only used for further refinement of the diagnosis (even if they can also be used for the principal diagnosis if considered individually). The individual evaluation of each test gives the same weight to each test which obviously can lead to other results than the decision tree. For this reason, our first approach has been extended in order to reflect also the strategy expressed by the decision tree of Fig. 1. For that purpose, new conditions have been added to the premises of several rules to be fired only if it is appropriate according to the actual position in the decision tree. If membership degrees other than 0 and 1 are obtained, more than one branch of the tree may be traversed. The fuzzy logic and decision tree approach could be successfully combined in this way. As expected, the diagnostic results of our expert system have been improved substantially by means of this modification (see also Section 4).

## 4 Experimental Results

Neuron Data's expert system shell Nexpert Object has been used for the implementation of our expert system. This shell, however, doesn't support fuzzy logic. For this reason, a fuzzy logic interpreter has been developed in C++. A special integration module provides an interface between rules and objects of Nexpert Object and the external fuzzy logic interpreter.

The fuzzy expert system has been tested by comparing the system's results of 66 real patient cases with the respective diagnoses of the medical experts. Disregarding the influence of inter- and intra-observer variability of the interpretation of the medical experts, the comparative tests led to the results as reported in Table 1.

| Diagnosis of the function | *Medical Expert* | | | |
|---|---|---|---|---|
| of the thyroid gland | Hypo | Eu | Hyper | Total |
| *Fuzzy Expert System*  Hypo | 8 | 0 | 0 | 8 |
| Eu | 1 | 27 | 2 | 30 |
| Hyper | 0 | 2 | 26 | 28 |
| Total | 9 | 29 | 28 | 66 |

Table 1: Summary of the test results

The table shows that the system (second approach) reaches the same conclusions as the experts in about 92% of the cases. In absolutely no case, a wrong type of malfunction (Hyper- instead of Hypothyroidism or vice versa) has been reported. The only divergent results provided by the experts and our system,

respectively, concern patient cases where a latent malfunction has been concluded instead of a normal function (or vice versa), i.e. a difference which anyhow is rather subtle. In comparison with the above results, the first approach which didn't include the decision tree knowledge of Fig. 1 had a performance of about 81% correct diagnoses on the same test cases as above. However, there have been 11 cases where the same evidence has been computed for Hyper- and Euthyroidism. I.e., no clear decision could be made. Theses cases have been counted as 50% correct in our statistics.

# 5 Conclusions

This paper has presented an expert system for the evaluation of laboratory tests in order to determine thyroid function. The test interpretation knowledge has been modeled by means of fuzzy logic methods. Fuzzy logic itself includes methods for combining results of various tests which, however, are only appropriate as long as all tests are to be considered with the same importance. In our system, a combination of pure fuzzy logic and traditional decision tree technique has proved to be most advantageous. Altogether, this approach has turned out to be well adapted to problems like the one considered and shown good diagnostic performance.

# References

[1] E. H. Shortliffe. *Computer-Based Medical Consultations: MYCIN.* American Elsevier Publishing, 1976.

[2] H. Pople. The Formation of Composite Hypotheses in Diagnostic Problem Solving. *Proceedings 5th IJCAI, Boston*, pages 1030–1037, 1977.

[3] S. Weiss, C. Kulikowski, and A. Safir. Glaucoma Consultation by Computer. *Comp. Biol. Med.*, 8:24–40, 1978.

[4] Glenn Shafer. *A Mathematical Theory of Evidence.* Princeton University Press, New Jersey, 1976.

[5] A. P. Dempster and A. Kong. Uncertain Evidence And Artificial Analysis. In Glenn Shafer and Judea Pearl, editors, *Readings in Uncertain Reasoning*, pages 522–528. Morgan Kaufmann Publishers Inc., San Mateo, 1990.

[6] L.A. Zadeh. The concept of a linguistic variable and its application to approximate reasoning. *Memorandum ERL-M 411 Berkeley*, October 1973.

[7] H.-J. Zimmermann. *Fuzzy Set Theory and its Applications.* Kluwer Academic Publishers, Boston, Dordrecht, London, second, revised edition, 1991.

[8] Abraham Kandel, editor. *Fuzzy Expert Systems.* CRC Press, Boca Raton, 1992.

[9] F. Grimm, X. Fábregas. An expert system for the diagnosis of thyroid gland diseases. In *Proceedings of the SHARE Europe Spring Meeting, Hamburg, April 19-22*, pages 275–292, 1993.

# Probabilistic Rule Induction Method from Clinical Database based on Rough Sets and Resampling Methods

Shusaku Tsumoto and Hiroshi Tanaka
Department of Informational Medicine
Medical Research Institute,Tokyo Medical and Dental University
1-5-45 Yushima, Bunkyo-ku Tokyo 113 Japan *

**Abstract**
One of the most important problems in rule induction methods is how to estimate the reliability of the induced results, which is a semantic part of knowledge to be induced from finite training samples. Since ordinary inductive learning methods, such as induction of decision trees, AQ method and rough sets, are mainly oriented to discover the syntactic structure of training data, they assume the deterministic nature of those samples for efficiency, and in most of the cases, the reliability of acquired knowledge is not evaluated statistically. Hence they are not well applicable in their present form to probabilistic domain, such as medical domain. In order to reduce these disadvantages further, we introduce repeated cross-validation method and the bootstrap method, both of which are studied as nonparametric error estimation methods or statistical model estimation ones in the community of statistics. The results show that this combination estimates the accuracy of the induced results correctly.

## 1 Probabilistic Rules in RHINOS2

Our approach is firstly motivated by automatic rule generation for RHINOS [5].RHINOS is an expert system which diagnoses the causes of headache or facial pain from manifestations. For the limitation of the space, in the following, we only discuss about the acquisition of inclusive rules,which are used for differential diagnosis. For further information,refer to [5].

Inclusive rule consists of several rules,which we call positive rules. The premises of positive rules are composed of a set of manifestations specific to a disease to be included for the candidates of disease diagnoses. If a patient satisfy one set of the manifestation of a inclusive rule, we suspect the corresponding disease with some probability. These rules are derived by asking the following questions in relation to each disease to the medical experts:*1.a set of manifestations by which we strongly suspect a corresponding disease. 2.the probability that a patient has the disease with this set of manifestations:SI(Satisfactory Index) 3.the ratio of the number the patients who satisfy the set of manifestations to that of all the patients having this disease:CI(Covering Index) 4.If sum of the derived CI(tCI) is equal to 1.0 then end. If not, goto 5. 5.For the patients suffering from this disease who do not satisfy all the collected set of manifestations, goto 1.* An inclusive rule is described by the set of manifestations, and its satisfactory index. Note that SI and CI are given experimentally by medical experts.

Formally, we can represent each positive rule as a tuple: $\langle d, R_i, SI_i(, CI_i) \rangle$, where $d$ denotes its conclusion, and $R_i$ denotes its premise. The inclusive rule is described as: $\langle \{\langle d, R_1, SI_1(, CI_1) \rangle, \cdots, \langle d, R_k, SI_k(, CI_k) \rangle\}, tCI \rangle$. where total CI(tCI) is defined as the sum of CI of each rule with the same conclusion:$\sum_i CI_i$.

## 2 Probabilistic Extension of Rough Sets

Rough set theory is developed and rigorously formulated by Pawlak[7]. This theory can be used to acquire certain sets of attributes which would contribute to class classification and can also evaluate how precisely these attributes are able to classify data.

*Correspondence to: Shusaku Tsumoto, Department of Informational Medicine, Medical Research Institute,Tokyo Medical and Dental University, 1-5-45 Yushima, Bunkyo-ku Tokyo 113 Japan, TEL: +81-03-3813-6111 (6159), FAX: +81-03-5684-3618 email:tsumoto@tmd.ac.jp

For the limitation of space, we mention only how to extend the original rough set model to probabilistic domain, which we call PRIMEROSE( Probabilistic Rule Induction Method based on ROugh Sets ). And we denote a set which supports an equivalence relation $R_i$ by $IND(R_i)$ and we call it an *indiscernible set*. For example, if an equivalence relation $R$ is supported by a set $\{1,2,3\}$, then $IND(R)$ is equal to $\{1,2,3\}$ ( $IND(R) = \{1,2,3\}$ ).

## 2.1 Definition of Probabilistic Rules

We extend the definition of consistent rules to probabilistic domain. For this purpose, we use the definition of inclusive rules which Matsumura et.al [5] introduce for the development of a medical expert system, RHINOS(Rule-based Headache and facial pain INformation Organizing System). This inclusive rule is formulated in terms of rough set theory as follows:

**Definition 1 (Definition of Probabilistic Rules)** *Let $R_i$ be an equivalence relation and $X$ denotes one class, which is the subset of U. A probabilistic rule of $X$ is defined as a tuple, $< X, R_i, SI(R_i, X), CI(R_i, X) >$ where $R_i$, SI, and CI are defined as follows. $R_i$ is a conditional part of a class $X$ and defined as:*

$$R_i \quad s.t. \quad IND(R_i) \bigcap IND(X) \neq \phi$$

*SI and CI are defined as:*

$$SI(R_i, X) = \frac{card\ \{(IND(R_i) \bigcup IND^c(R_i)) \bigcap (IND(X) \bigcup IND^c(X))\}}{card\ \{IND(R_i) \bigcup IND^c(R_i)\}}$$

$$CI(R_i, X) = \frac{card\ \{(IND(R_i) \bigcup IND^c(R_i)) \bigcap (IND(X) \bigcup IND^c(X))\}}{card\ \{IND(X) \bigcup IND^c(X)\}}$$

*where $IND^c(X)$ or $IND^c(R_i)$ consists of unobserved future cases of a class $X$ or those which satisfies $R_i$, respectively.* $\qquad\qquad\square$

A total rule of $X$ is given by $R = \bigvee_i R_i$, and then total CI(tCI) is defined as: $tCI(R,X) = CI(\bigvee_i R_i, X)$. Since the above formulae include unobserved cases, we are forced to estimate these measures from the training samples. For this purpose, we introduction cross-validation and the Bootstrap method to generate "pseudo-unobserved" cases from these samples as shown in the next subsection.

## 2.2 Cross-Validation and the Bootstrap Method

Cross-validation method for error estimation is performed as following: first, the whole training samples $\mathcal{L}$ are split into $V$ blocks: $\{\mathcal{L}_1, \mathcal{L}_2, \cdots, \mathcal{L}_V\}$. Second, repeat for V times the procedure in which we induce rules from the training samples $\mathcal{L} - \mathcal{L}_i (i = 1, \cdots, V)$ and examine the error rate $err_i$ of the rules using $\mathcal{L}_i$ as test samples. Finally, we derive the whole error rate $err$ by averaging $err_i$ over $i$, that is, $err = \sum_{i=1}^{V} err_i / V$ (this method is called $V$-fold cross-validation). Therefore we can use this method for estimation of $CI$ and $SI$ by replacing the calculation of $err$ by that of $CI$ and $SI$, and by regarding test samples as unobserved cases.

On the other hand, the Bootstrap methods is executed as follows: first, we create empirical probabilistic distribution($F_n$) from the original training samples. Second, we use the Monte-Carlo methods and randomly take the training samples by using $F_n$. Third, rules are induced by using new training samples. Finally, these results are tested by the original training samples and statistical measures, such as error rate are calculated. We iterate these four steps for finite times. Empirically, it is shown that about 200 times repetition is sufficient for estimation.

Interestingly, Efron[3, 4] shows that estimators by 2-fold cross-validation are asymptotically equal to predictive estimators for completely new pattern of data, and that Bootstrap estimators are asymptotically equal to maximum likelihood estimators[3, 4] and are a little overfitted to training samples. Hence, we can use the former estimators as the lower bound of SI and CI, and the latter as the upper bound of SI and CI.

Furthermore, in order to reduce the high variance of estimators by cross-validation, we introduce repeated cross-validation method, which is firstly introduced by Walker[9]. In this method, cross-validation methods are executed repeatedly(safely, 100 times), and estimates are averaged over all the trials. In summary, since our strategy is to avoid the overestimation and the high variabilities, we adopt combination of repeated 2-fold cross-validation and the Bootstrap method in this paper.

118

Table 1: Information of Database

| Domain | Samples | Classes | Attributes |
|--------|---------|---------|------------|
| headache | 121 | 10 | 20 |
| meningitis | 99 | 3 | 25 |
| CVD | 137 | 6 | 27 |

## 2.3   Cluster-based Reduction of Knowledge

Reduction technique removes dependent variables from rules. This dependence is originated from algebraic dependence, that is , if $f(a_1, a_2, \cdots, a_n, a_{n+1}) = f(a_1, a_2, \cdots, a_n) = 0$ then $a_{n+1}$ is dependent on $a_1, a_2, \cdots, a_n$. Hence, intuitionally, if the removal of one variable does not change the former consistent classification, we can remove this variable. In PRIMEROSE, we extend the concept of reduction to probabilistic domain: we delete an attribute when the deletion does not make apparent SI change. For example, if one rule support one class with some probability and other classes with some probabilities, we minimize its conditionals by the cluster-based reduction: that is, if the removal of one attribute does not change the above probabilities, we can remove this attribute.

This process means that we fix the probabilistic nature of the induced rules and is very effective when databases include inconsistent samples. This method is very similar to VPRS model introduced by Ziarko[10].

On the other hand, in the original Pawlak's models inconsistent parts is ignored and only reduction of the consistent parts is executed. For precise information, please refer to [7, 10, 8].

## 2.4   Algorithm for PRIMEROSE

Algorithms for rule induction can be derived by embedding rough set theory concept into the algorithms discussed in Section 2. An algorithm for induction of inclusive rules is described as follows:

1)Using all attributes,calculate all equivalent relation $\{R_i\}$ which covers all of the training samples, that is, calculate $\{R_i | \bigcup IND(R_i) = U\}$.
2)For each class $d_j$, collect all the equivalent relation $R_i$ such that $IND(R_i) \cap IND(d_j) \neq \phi$. For each combination, calculate its possible region.
3)Calculate $SI(R_i, d_j)$.
4)Apply probabilistic reduction of knowledge to each relation $R_i$ until SI is changed ( Minimize the components of each relation ). If several candidates of minimization are derived, connect each with disjunction.
5)Collect all the rules,perform the cross-validation method to estimate utCI for each $d_j$.

# 3   Experimental Results

We apply PRIMEROSE to headache(RHINOS's domain), meningitis, and cerebrovascular diseases, whose precise information are given in Table 2. These data are incomplete, and include many inconsistencies.

The experiments are performed by the following three procedures. First, we randomly splits these samples into pseudo-training samples and pseudo-test samples. Second, by using the pseudo-training samples, PRIMEROSE induces rules and the statistical measures. Third, the induced results are tested by the pseudo-test samples. We perform these procedures for 100 times and average each accuracy and the estimators for accuracy over 100 trials. We compare PRIMEROSE with AQ15[6] and CART[1].

Experimental results are shown in Table 2. These results suggest that PRIMEROSE performs a little better than the other two methods and that the estimation of accuracy performs very well.

119

Table 2: Experimental Results

| Domain | Method | Accuracy | 2-fold CV Estimator | the Bootstrap Estimator |
|--------|--------|----------|---------------------|-------------------------|
| headache | CART | 62.8% | | |
| | AQ15 | 61.2% | | |
| | PRIMEROSE | 74.4% | 58.7% | 91.6% |
| meningitis | CART | 60.6% | | |
| | AQ15 | 67.7% | | |
| | PRIMEROSE | 74.7% | 59.6% | 88.3% |
| CVD | CART | 65.7% | | |
| | AQ15 | 73.0% | | |
| | PRIMEROSE | 81.7% | 70.1% | 87.5% |

# 4 Related Works

## 4.1 Comparison with AQ15

AQ is an inductive learning method based on incremental STAR algorithm developed by Michalski [6]. This algorithm selects one seed from positive examples and starts from one "selector"(attribute). It adds selectors incrementally until the "complexes" (conjunction of attributes) explain only positive examples. Since many complexes can satisfy these positive examples, according to a flexible extra-logical criterion,AQ finds the most preferred one.

It would be surprising that the complexes supported only by positive examples corresponds to the positive region. That is, the rules induced by AQ is equivalent to consistent rules introduced by Pawlak[7]. However, as shown in [7], the ordinary rule induction by rough set theory is different from AQ in strategy;Pawlak's method starts from description by total attributes, and then performs reduction to get minimal reducts,that is, rules are derived in a top-down manner. On the contrary, AQ induces in a bottom-up manner. While these approaches are different in strategies, they are often equivalent because of logical consistency, and this difference suggests that when we need the large number of attributes to describe rules, induction based on rough set theory is faster.

One of the important problem of the AQ method is that it does not work well in probabilistic domain [6]. This problem is also explained by matroid theory: inconsistent data do not satisfy the condition of independence, so we cannot derive a basis of matroid in probabilistic domain using the proposed definition, which is the same problem as the Pawlak's method,as discussed in Section 4. Hence it is necessary to change the definition of independence to solve those problems.

As discussed earlier,in PRIMEROSE, we adopt cluster membership as the condition of independence, instead of using class membership. Restricting the probabilistic nature, we can use almost the same algorithm as class-consistency based reduction. Then we estimate the probabilistic nature of the derived rules using some resampling plans, such as cross-validation method in this paper. This is one kind of solution to the above problems, and the similar approach can also solve the disadvantage of AQ.

## 4.2 Variable Precision Rough Set Model(VPRS)

Variable Precision Rough Set Model(VPRS) is a probabilistic generalization of rough set model which Ziarko independently introduce[10].

This model focuses on boundary region which cannot be classified by original rough set model. Ziarko introduces precision $\beta$ and allow for some degree of misclassfication in the largely correct classification. This means that the elements of boundary region nearer to the positive region are included in the $\beta$-lower approximation $\underline{R}^{\beta} X$. Hence VPRS also uses our possible region for rule induction, although our model is based on primitive cluster, which appears in training samples.

Interestingly, although defined region is probabilistic, this generalized model inherits all basic mathematical properties of the original model. Hence VPRS complements the rough set model as a methodological tool for rule induction in probabilistic domain.

However, this model left one problem unsolved. Although precision is predefined, the induced results

may not reflect "real" precision level. For example, we induce rules from training samples by precision $\beta = 0.1$. Then the accuracy of rules is also derived by this precision level. But when we get some new samples, the accuracy will be higher than this level. So it is desirable to estimate the bias of precision. For this purpose, our repeated cross-validation method may be used. Such kind of estimation is also important problems in probabilistic rule induction.

Since we have not implemented a system based on VPRS yet, we do not compare the performance of this model with ours. In near future, we will report the comparison of each method.

# 5 Conclusion

We introduce a new approach to knowledge acquisition, PRIMEROSE, and develop an program based on this method to extract rules for an expert system from clinical database. It is applied to three medical domains. The results show that the derived rules performs a little better than CART and AQ15 and that the estimation of statistical measures performs well.

# References

[1] Breiman,L.,Freidman,J.,Olshen,R.,and Stone,C. *Classification And Regression Trees.* Belmont,CA:Wadsworth International Group, 1984.

[2] Efron B. *The Jackknife, the Bootstrap and Other Resampling Plans.* Pennsylvania:CBMS-NSF,1982.

[3] Efron B. Estimating the error rate of a prediction rule: improvement on cross validation. *J.Amer.Statist.Assoc.* **78**,316-331,1983.

[4] Efron B. How biased is the apparent error rate of a prediction rule ? *J. Amer. Statist. Assoc.* **82**,171-200,1986.

[5] Matsumura,Y, et al. Consultation system for diagnoses of headache and facial pain: RHINOS,*Medical Informatics*,**11**,145-157,1986.

[6] Michalski,R.S.,et al. The Multi-Purpose Incremental Learning System AQ15 and its Testing Application to Three Medical Domains, *Proc. of AAAI-86*, 1041-1045,Morgan Kaufmann,1986.

[7] Pawlak,Z *Rough Sets*,Kluwer Academic Publishers, 1991.

[8] Tsumoto,S.and Tanaka,H. PRIMEROSE:Probabilistic Rule Induction based on Rough Sets and Resampling MEthods, *Proc. of RSKD'93*, 1993.

[9] Walker,M.G. and Olshen,R.A. Probability Estimation for Biomedical Classification Problems. *Proc. of SCAMC-92*,McGrawHill,1992.

[10] Ziarko,W. Variable Precision Rough Set Model, *Journal of Computer and System Sciences*,**46**,39-59,1993.

# Multiple models and flexible reasoning for a breast cancer knowledge-based system

Eleni Christodoulou
Department of Computer Science
University of Cyprus
P.O.Box 537, Nicosia, CYPRUS
e-mail:cseleni@jupiter.cca.ucy.cy

February, 1994

**Abstract**

This paper reports on the development of a breast cancer knowledge-based system, which uses multiple models and multiple reasoning operators to provide prognostic and diagnostic conclusions on malignant masses of suspected breast areas. The program uses three different knowledge representation models for capturing the domain knowledge: a rule-based model, a "class-oriented" model consisting of a class taxonomy describing the different breast cancer histological types and a case-based model, each providing an alternative form of diagnostic reasoning. Different reasoning operators are used to generate a diagnosis from these knowledge representation models. A separate, flexible, metareasoning module dynamically decides among the different representation models and the phase of reasoning to be performed. This structure of many models and multiple reasoning strategies gives the diagnostic system the flexibility to determine at every choice point, during a problem-solving activity, the "best" strategy for the particular case, and to change strategy when the currently pursued one ceases to be the best, thus going out of an unpromising route. The result is a flexible system that supports efficiently the decision-making procedure of histopathologists.

## 1 Introduction and design philosophy

Breast cancer is the commonest of all malignant diseases in women in recent years. It is the leading cause of death from cancer in western women. There are many different histological breast cancer types, each associated with a different prognosis and recurrence rate. Such information is vital in determining immediate treatment of the patient and conducting long term follow up [1]. Presently several subjective schemes are available for the diagnostic classification of carcinoma of the breast. Histopathologists base their conclusion on a number of diagnostic factors, extracted from a manual microscopic and macroscopic analysis of breast biopsies and on some other clinical data of the patient.

We are developing, in a collaborative project with the Department of Histopathology of the Nicosia General Hospital and the Cyprus Institute of Neurology and Genetics, a knowledge-based computer system to assist histopathologists in an objective diagnostic and prognostic decision making of carcinoma of the breast. Towards this goal, conventional and advanced knowledge representation and reasoning techniques have been used to reach high diagnostic accuracy.

In general, a human domain expert, for example a histopathologist possesses a large amount of specialized factual knowledge and a number of reasoning methods that he flexibly applies on the factual domain knowledge to achieve efficient diagnostic conclusions. In a knowledge-based

system this specialized knowledge is normally expressed in different representation formalisms and kept separate from the code that performs the reasoning. Histopathologists are highly skilled at quickly identifying which component of their knowledge to use for a specific problem case. Particular sets of features prompt the histopathologist to use the appropriate knowledge component. By modelling the diagnostic system behaviour on that of human experts, including their knowledge structure and inference methods, the user acceptance of the medical system is enhanced.

## 2 System design

It has been argued that the use of many knowledge representation models is better than the use of a single model and that there is not a unique model able to meet all the requirements needed to develop a knowledge-based sytem [4], [5], [6], [10]. In the development of the breast cancer diagnostic system we emulated the decision-making behaviour of a histopathologist, by incorporating three different knowledge representation models, a number of reasoning operators and a higher level metareasoner to guide the reasoning between the different models. The organization of the different models and operators is sketched in figure 1. The operators are identified by rectangles and the data/knowledge bases by ovals. The diagram shows only the main interactions between the system components. In the next part of this section the structure of the knowledge base and the different reasoning operators are described.

**Knowledge base**

Three different models make up the knowledge base of the breast cancer system. Two of these form a static physiologic knowledge base and the third is built dynamically from cases. The three models are respectively a rule-based model, a deep model, and a case-based model. The latter is built through a case-based reasoner and it is organized in a way to facilitate the matching with new cases.

The rule-based representation model consists of a set of rules combining the different nuclear and cytoplasmic features with resulting conclusions (diagnoses). The nuclear and cytoplasmic features include for example: *cytoplasm stain, clumb thickness, uniformity of cell shape, marginal adhesion, single epithelial cell size, bare nuclei, bland chromatin, normal nucleoli, cut surface texture, cut surface colour, existence of extracellular mucin, etc.*

An example of a rule is given below:

```
Rule: IF the < lesion contain pools of extracellular mucin >    and
         the < mucin pools contain islands of malignant cells >    and
         the < entire malignant part of the lesion consist of
              islands of malignant cells in mucin pools >
         THEN  diagnosis = mucinous carcinoma
```

The different nuclear and cytoplasmic features are defined in the knowledge base of the system in the form of variables taking different qualitative or quantitative values (e.g. yes/no, low/intermediate/high or 0-100)

The "deep" knowledge model is based on a *class taxonomy*. In general, a class represents an object, a concept or an event about which data are to be stored. Simple examples are patient(an object) and diagnosis(a concept). Normaly, a class is described by a collection of attributes (characteristics of the class) and their associated values. A class taxonomy is a hierarchy of different classes, where attributes of classes are inherited by classes lower in the hierarchy. In the breast cancer model a class is a cluster of a number of important items of diagnostic knowledge that are utilized by all pathologists in the identification of histological breast cancer

types. These include information about the possible clinical symptoms of a disease, its macroscopic and microscopic appearances. Macroscopic features are features observed macroscopically such as colour or consistency. Microscopic features are features observed microscopically such as distended ducts, tubule formation or mitoses.
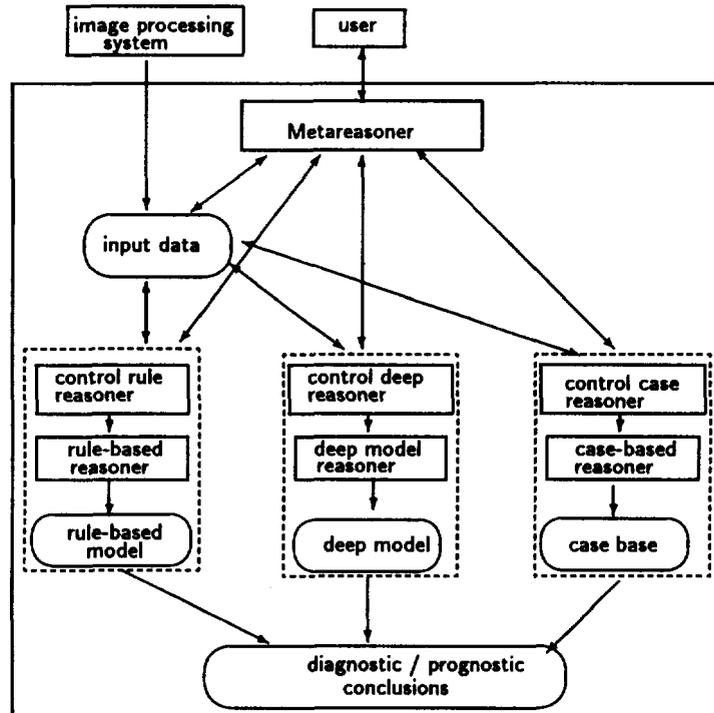


figure1: System Architecture illustrating the different Models and Reasoners

A class represents for example one of the following breast cancer types: *ductal carcinoma, lobular carcinoma, medullary carcinoma, colloidal carcinoma, tubular carcinoma, invasive cribriform carcinoma and mucinous carcinoma.*

The use of deep modelling in medical diagnosis results in higher problem-solving flexibility and facilitates the extensibility of the system [2], [3] .

The case-based model is a discrimination network of completed breast cancer diagnosed cases. Whenever the correct diagnosis is determined for a case, possibly through the static knowledge representation models, the new case is then added to the discrimination network. The purpose of the third model is to allow a *case-based reasoner* to rapidly find the cases that are similar to a new case as a starting point for a diagnosis. The first step in case-based diagnosis is using the input parameters to find similar cases in the case network. Once a partial match of the parameters has been found, the next step is to try to adapt the stored case to the new parameters. If the new parameters can not be adapted to an existing case then the search is abandoned and the matching is rejected. Case-based reasoning constitutes a promising approach to modelling and reasoning with medical knowledge [8], [12]. Diagnoses are made by recognizing similarities between previous cases encountered and by adapting previous solutions to new problem cases.

Learning is intrinsically related with the case-based model [8]. The system learns from its experience in problem solving. Experience plays two important roles in problem solving:

- It contributes to the refinement and modification of reasoning processes and knowledge.

- Individual experiences act as exemplars upon which to base later decisions.

A system that can make use of previous experience can cut down on its work by using previous cases to suggest a solution to a new but similar case [12]. A reasoning system that uses experience

124

is able to take advantage of the similarities between novel cases. Since most cases are not classic ones, and since exceptions seem to recur in medicine, such a system has many advantages over one that must go through the same long and complicated reasoning process to analyze two similar, complex, cases. In the breast cancer diagnostic system, experience is gained by sets of already diagnosed cases. After a diagnosis has taken place, this is added in a case-based structure implemented in the form of a network. These cases will then be used in evaluating new and relatively novel situations.

The capability of using experience in reasoning, enables us to build a more adaptable diagnostic system. In many situations the histopathologists cannot accurately tell the system all the information the system needs. This problem, could at least be partially alleviated through the development of a means of learning from experience. Our principle is to build a system based on the best knowledge available and to gradually train this system with actual cases.

**Reasoning operators**

The reasoning operators are the active elements of the system. They take the input, turn it into a patient specific model, and generate the diagnostic or prognostic information. Every model of the knowledge base is associated with a separate reasoner that performs inferencing on the given knowledge formalism. Each model reasoner uses its own local control reasoner. The aim is to provide flexible reasoning that enables each time the selection of the relevant subset of knowledge to be activated, thus reducing the search space for the particular problem and enhancing the system performance.

A metareasoner is used to drive the behaviour of the overall system. It interacts with the user for gathering the input data and decides which is the best knowledge model to be used for a particular problem case. The decision is taken by identifying from the input data a set of features that can be characteristic for the appropriate model to be used. The control then passes to the control reasoner of the relevant model. The need to use a metareasoner for building powerful systems is pointed out in [7], [9].

In a future stage the diagnostic system will be expanded to cover prognostic assessment of carcinoma of the breast. The input data of this prognostic system would be partly extracted by a computer-based image processing system applied to biopsy slides of the breast masses. This entails the development of an integrated decision-support system for the diagnosis and prognosis of carcinoma of the breast.

# 3 Conclusion

There is a relatively high incidence of breast cancer in women in the last years. Different research groups have been working towards an accurate and objective computer-based diagnostic system for carcinoma of the breast. It has been proven that knowledge-based systems provide the best solution in problems where the decision-making behaviour of an expert in a specialized, knowledge-intensive field has to be emulated. Systems already in existence for this domain [13], [14] base their solution on knowledge-based mechanisms using only one representation model of the domain knowledge and one reasoning method. In [14] a rule-based model is used and in [13] a class-model is used. Using only one represenation model and one reasoning stategy reduces the performance of the system as well as the user acceptance of the system, since the user, in this case the histopathologist, has to deal with a number of questions concerning knowledge that is not relevant for every case. Histopathologists identify quickly which part of their knowledge must be used for a specific case, avoiding an elicitation of unnecessary data. An accurate emulation of the diagnostic behaviour of a human expert must be based on multiple representation models and flexible reasoning operators, effectively controlling, for every problem case, the use of the relevant subset of knowledge. A system making use of three different models, a rule-based, a class-model and a case-based model, a set of different inference operators and a

metareasoner driving the behaviour of the overall system, provides the user with an efficient and powerful solution to the breast cancer diagnostic problem. This architecture is not restricted to the medical task of diagnosis alone but it can be applied to all the other medical tasks [5] including prognosis, therapy planing and monitoring.

**Acknowledgment**

I would like to thank Associate Professor Elpida Keravnou (University of Cyprus) for her constructive comments and fruitful interaction.

# References

[1] Dixon, J. M., Page, D. L., Anderson, T. J. *et al Long-term survivors after breast cancer*, British Journal of Surgery, 1985, 72, 445-448

[2] E. Keravnou, J. Washbrook *Deep and shallow models in medical expert systems*, Artificial Intelligence in Medicine 1 (1989) 11-28

[3] E. Keravnou, J. Washbrook *What is a deep expert system? An analysis of the architectural requirements of second-generation expert systems*, The Knowledge Engineering Review, No. 4: 3, 1989, 205-233

[4] G. Lanzola, M. Stefanelli *A specialized Framework for Medical Diagnostic Knowledge-Based Systems*, Computers and Biomedical Research 25, 351-365 (1992)

[5] M. Stefanelli, R. Bellazi, C. Berzuini, L. Ironi, S.Quaglini *Towards a General Architecture for Medical Expert Systems*, working notes AAAI spring symposium series, Artificial Intelligence in Medicine, Stanford University, March 25, 1992

[6] G. Lanzola, M. Stefanelli *Computational Model 1.0*, AIM Project report: A General Architecture for Medical Knowledge-Based Systems, June 27, 1992, University of Pavia

[7] G. Lanzola, M. Stefanelli *Computational Model 3.0*, AIM Project report: A General Architecture for Medical Knowledge-Based Systems, December 30, 1993, University of Pavia

[8] J.L. Kolondner, R.M. Kolondner *Using Experience in Clinical Problem Solving: Introduction and Framework*, IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-17, No. 3, May/June 1987

[9] S. Russel, E. Wefald *Principles of Metareasoning*, Artificial Intelligence 49 (1991) 361-395

[10] M. DeJongh and J. Smith *Integrating models of a domain for problem solving*, Expert Systems and their applications, General Conference, Second Generation expert systems, Avignon 91, volume 2, May 27-31, 1991

[11] B. Chandrasekaran, T. Johnson *Generic Tasks and Task Structures: History, Critique and New Directions*, in *Second Generation Expert Systems*, Jean-Marc David, Jean-Paul Krivine and Reid Simmons, editors, Springer Verlag

[12] J.A. Campbell, J. Wolstencroft *Cases and the elucidation of deep knowledge*, in *Deep Models for Medical Knowledge-Engineering*, E. Keravnou (editor), 1992 Elsevier Science Publishers B.V.

[13] H. Heathfield, N. Kirkham *A cooperative approach to decision support in the differential diagnosis of breast disease*, Med. Inform. (1992), vol. 17, No. 1, 21-33

[14] Van Diest PJ, Belien J.A., Baak J.P. *An expert system for histological typing and grading of invasive breast cancer*, Pathol. Res. Pract. 1992 Jun; 188(4-5): 405-9

# Designing for Decision Support
# in a Clinical Monitoring Environment

## Enrico Coiera[1]

**Hewlett Packard Laboratories**
**Filton Rd., Stoke Gifford, Bristol, BS12 6QZ, UK**
**email:ewc@hplb.hpl.hp.com**

**Abstract** - The conceptualisation of user needs that drives much of the technological development in decision support for patient monitoring is inadequate of flawed. In an attempt to develop a clearer picture of such user needs, preliminary results of a study into the behaviour of clinicians performing real-time monitoring are discussed. The study informally suggests that a critical reassessment needs to be made of the type of decision support that should be provided in monitoring situations. More generally, the study results are used to argue that such clinician studies are generally lacking in the design of all clinical decision support systems.

## 1 Introduction

It seems almost a truism to say that the role of a clinical decision support system is to assist health care professionals in carrying out their clinical duties. Our purpose in building them is also apparently straightforward - we wish either to provide assistance for tasks that are subject to human error, or to provide tools that augment human performance. The measured results of our efforts should either be reduced patient morbidity and mortality, or a reduction in the costs of providing care.

For example, decision support has long been advocated in patient monitoring, based upon the perception that monitoring constitutes a complex set of tasks that are prone to human error. Further, it is believed that such enhancements to monitors can lead to the earlier detection of significant clinical events, and that this will reduce morbidity, mortality and costs.

Unfortunately, while there have been many attempts at developing computer systems for decision support with patient monitors, they have failed to make any

significant impact upon clinical practice. The current state of the art in monitoring reflects improvements in measurement fidelity, and the development of new measurements, rather than any advance in the way users interact with the monitoring system. In this paper, the motivations that underlie much of the technological development for decision support in patient monitoring are critiqued. Difficulties with the present conceptualisation of user needs are presented that may in part explain the current state of affairs. In response to these, preliminary results of clinician studies are presented that suggest a critical reassessment needs to be made of the type of decision support that should be provided in monitoring situations. More generally, it will also be argued that such clinician studies are generally lacking and that much clinical information system design needs to be reappraised in this light.

## 2 Problems with the State of the Art

In many ways, work to date in the design of decision support systems for patient monitoring has been fragmented. On the one hand, there has been a steady effort concentrated on the technical aspects of diagnosis, monitoring and control of physiological signals [2]. On the other hand, some work has taken place exploring clinical errors in patient monitoring, for example in the delivery of anaesthesia [5] [15]. There is also some literature on human factors in the clinical workplace, identifying causes of suboptimal performance and suggesting areas in which decision support might be appropriate (e.g. [8]). Further removed from clinical medicine, there is a large literature on the cognitive aspects of process control [16]. These latter bodies of work help to characterise clinical tasks, and should provide the motivational impetus and focus for the technological development of intelligent patient monitoring systems.

Unfortunately, there seems to have been little connection between the technical developments of decision support systems, and the work on the cognitive aspects of clinical decision making that should motivate these developments. Yet it is the cognitive motivations that should identify which aspects of clinical tasks should be supported by a decision support system.

Much of the work done in monitoring decision support has paid lip service to the notion of task understanding, and has operated on a set of largely untested assumptions about clinical need. For example, a strong paradigm for the design of decision support systems has been the provision of diagnostic information in real time situations. This view was possibly motivated by early research in decision analysis which suggested that human judgements were prone to decision biases [12], and the corollary which stated that such flawed judgements could be normalised by relying on sounder, more formally based, diagnostic systems. There is now growing evidence that questions such an assessment. Based upon studies of individuals in the field rather than in controlled laboratory situations, evidence now suggests that the primary effort for decision makers is not at the moment of choice, but rather in *situation assessment* [13]. In other words, it may well be the case that the majority of clinicians do not have difficulty in making a diagnosis, but rather in establishing a clear picture of the state of the world, and clarifying their goals and assumptions, prior to attempting to make a diagnosis. If it is the case

that situation assessment is the key bottleneck in decision making, this has implications for the design of decision support tools for monitoring and control. Systems that assist clinicians in making an assessment of monitored data may be of more utility than systems that attempt to manufacture a diagnosis.

With this background, we undertook to re-examine our understanding of clinical behaviour during patient monitoring situations, in the hope that the insights such a study might provide would translate into requirement specifications for a clinical decision support system.

## 3 Studying Clinical Decision Making

How does one gain an understanding of the needs of clinical practice? There are a number of alternatives, which can be roughly grouped into the following four categories:

*Anecdotal or individual experience:* At the most superficial level one can draw upon individual experience, or rely on anecdotal evidence from other colleagues, to help pinpoint the tasks clinicians need help with. This method usually operates within a framework of the existing assumptions about user needs.

*Asking clinicians:* If one wants to be a little more methodical, one might conduct a survey, and ask clinicians about their needs. While this may seem attractive, it has its problems. The biggest is that it is unlikely that clinicians really know what they want or need - "Although users are expert at what they do, they have difficulty predicting what they would like" [8].

More specifically, both these approaches suffer from a number of methodological difficulties. The problem of attitudinal bias has long been understood in the social psychology literature - people's actions differ from their verbalised responses [18]. We don't necessarily do as we say. Secondly, these methods require a degree of self-reporting. When asked why they made a certain judgement, or how they solved a particular problem, people are capable of providing apparently plausible reports of their mental events. However, asking people to introspect about their behaviour is a contentious investigative method in psychology because verbal reports of influences on behaviour may not be valid. Since there is no method of independently checking the validity of self reports, the details of cognition remain private to the individual [15].

*Non-participatory observation:* The next stage is not to ask clinicians what they want, but to study the way they work in the field. There are many approaches that one could take, including techniques derived from ethnography [8], as well as the more traditional software design methods of data-flow and task analysis. Here researchers try to understand the demands made upon individuals through detailed observations of them engaged in routine tasks. By making such observations, one is in a position to identify needs that are not apparent to the individuals being studied.

*Formal psychological experiments:* Finally, at the most detailed and scientifically rigorous level, one might have a set of hypotheses about specific aspects of clinical behaviour, and set up controlled studies to test them out. Some of the classic work in clinical decision making has been of this type [6].

While it may seem that the ideal method of proceeding is to invest in formal psychological studies of clinical decision making, this is not necessarily the case. Laboratory experiments run the risk of treating decision making as an abstract process, independent of the vagaries and interrupts of working life. By their nature, such studies are conducted in laboratory conditions which factor out the interruptions and pressures of the real clinical workplace. In the real world, time pressures mean that short cuts are taken, and that people and events interrupt clinical activities.

What we need to do is to characterise medical decision processes as much as possible in the way they really occur, as opposed to the way they *should* occur. This will give us the best chance of designing systems for use in such environments, as opposed to designing what would be needed if everything else was ideal. It may well be the case for example, that in an ideal consultation a doctor may want to consult a computer to help with the diagnostic process. However, with the pressures of the real workplace, the bottlenecks to increased performance may be much more mundane and less glamorous, and may be more closely associated with easy and timely access to information, or better communications between professionals, or even simply reducing the amount of paperwork in the clinical workplace.

This argument strongly favours pursuing non-participatory methods in the field, such as the ethnographic approach. These are sufficiently formal to allow robust statements to be made about results, and have the advantage of being grounded in the realities of the clinical workplace. Formal experiments will only answer narrow questions and are relatively expensive to conduct - one must therefore have a broad understanding of the domain to make a value judgement about the most likely avenues for investigation prior to making the investment in experimentation.

## 3.1 Real-time decision making

Most analyses of monitoring practice have sought to focus on the causes of errors that have an impact on patient care [4][5]. This approach is limited in several ways. Firstly, categorising and tabulating errors says nothing about the behavioural context within which they occurred, specifically the sequence of behaviours that led up to them. Secondly, while errors are clinically important, they account for a very small portion of the total behaviour of a clinician [14]. Thirdly, the emphasis on removing causes of errors is one sided. We are unlikely to ever be in a position where all causes of error have been eliminated. We should thus acknowledge that errors and mishaps are inevitable and rather than solely focusing on the design of practices and support systems to eliminate error, we should aim to design ones that are explicitly intended to cope with error [1].

An alternate strategy is to examine the routine behaviour which occupies the bulk of clinical practice. This approach has two advantages. Firstly, by focusing on the elements of normal behaviour that are responsible for sub-optimal clinical performance, we may be in a better position to suggest useful forms of behavioural support. The emphasis here is on optimising behaviour which clinicians spend the majority of their time engaged in, rather than supporting critical but rare failures. Secondly, major errors usually arise in the context of a sequence of minor errors [10][4]. By analysing the dynamic context within which minor errors occur, as

opposed to simply identifying and categorising them, we may obtain a clearer idea of how to prevent them. An important consequence may be that potentially critical errors can be prevented by optimising "normal" clinical practice. An analysis of normal behaviour should also assist in exploring the ways in which clinicians routinely successfully recover from incidents, since in the main, most significant events are successfully dealt with and do not result in negative outcome for the patient [9]. Equally, we are interested in identifying the circumstances under which such strategies fail.

## 4 The Study

In an attempt to understand the types of decision problem facing clinicians in a real-time monitoring environment, a study of anaesthetists during cardiac bypass surgery was undertaken at two large teaching hospitals in the United Kingdom. While the full results of this study will be reported elsewhere, a brief description of the type of analysis performed, and the type of information that it supplies will be presented here.

The study aimed to collect qualitative data about the types of behaviour anaesthetists exhibit when monitoring patient data, and to generate a cognitive model of the monitoring strategies that they employed. The eventual aim of this work is to design monitoring decision support systems that naturally blend in with the anaesthetist's work practices, and support the tasks with which they experience difficulty.

The study involved 12 clinicians with a range of expertise from junior registrar to senior consultant anaesthetist. Each of the subjects was followed though one surgical procedure of between 2 and 4 hours in duration. The data collection commenced with a preoperative interview with the subject. The interview tried to capture the subject's understanding of the patient, the intended surgical procedure, and the intended anaesthetic. During the operation a variety of data were recorded, including video and audio records of the anaesthetist interacting with the monitoring and anaesthetic machines. A second shadow clinician who was not responsible for patient care acted as an observer, and recorded all therapeutic interventions and clinically significant events. This event record served as an annotation to the real-time data that was gathered from the monitoring systems attached by the subject clinician to the patient. Subsequent to the operation, a short interview obtained the clinician's reactions to the operation, attempting to gather data on how the clinicians's preoperative expectations were matched by the actual outcome of the operation.

Within a week of the operation, each subject was re-interviewed for about 2 hours, during which their performance was explored with reference to two or more clinically interesting portions of the operation that had been recorded. The data from the operation were available to act as prompts for the subjects' memory.

Once the entire data gathering exercise was completed, the data were extensively analysed, and episodes that demonstrated errors or potential errors were examined, along with periods of good performance. The result of the analysis was a description language of anaesthetic behaviour, and a number of hypotheses about

the way anaesthetists work in the monitoring environment. These hypotheses form a cognitive model which will allow us to further explore anaesthetic behaviour, and develop advanced decision support systems for monitoring tasks.

# 5 An Example Event - Missed Hypoxia

Several events were culled from each of the operations studied, and used as the subject matter for each of the 2 hour post-operative interviews. These events ranged in duration from a few minutes to periods that were 30 to 40 minutes long. The subject was interviewed in an attempt to gain an understanding of the intention behind observed clinical actions, and to obtain as complete a description as possible of the data used by the subject in framing each clinical decision.

In the following example, the behaviour of an anaesthetist is presented as he deals with events during an operation involving a coronary artery bypass graft. The event is reconstructed both from the data gathered during the surgery, as well as from the post-operative interviews.

## 5.1 The clinician

The subject anaesthetist in this example was an experienced clinician who had been practising for several years. He had substantial experience with the type of cardiovascular surgery for which he was providing anaesthesia.

## 5.2 The patient

The patient was a 47 year old gentleman who had a fairly large myocardial infarction several months prior to surgery. Following the myocardial infarction he had had recurrent episodes of left ventricular failure and subsequently developed a left ventricular aneurism. The patient was admitted to have the aneurism excised.

The problems noted by the subject anaesthetist prior to surgery were recurrent bouts of ventricular failure with episodes of pulmonary oedema that had not responded to conventional medical therapy. The patient's relatively poor left ventricular contractility with dyskinesia from the aneurism was considered a problem by the clinician, and he suggested that the patient could require fairly intensive haemodynamic support when coming off cardiac bypass.

## 5.3 The event

The event demonstrates a gradually declining $SaO_2$ level post bypass, probably caused by pulmonary oedema (Figure 1.). The downward trend started off with an $SaO_2$ at 100 immediately after bypass, and dropped down to a low of about 87. The clinician did not detect the trend, despite it reaching clinically significant levels, until he took a routine blood gas sample about 30 minutes after the drop had commenced. As usual with most clinicians we observed, the alarm capabilities of the monitoring device had been suspended - many clinicians perceive that patient monitors have a high rate of false alarms. Consequently, no alarm sounded. In some sense, however the presence or absence of alarms was irrelevant, since the downward trend would have been visible for quite some time before the SaO2 dropped into the alarm region.
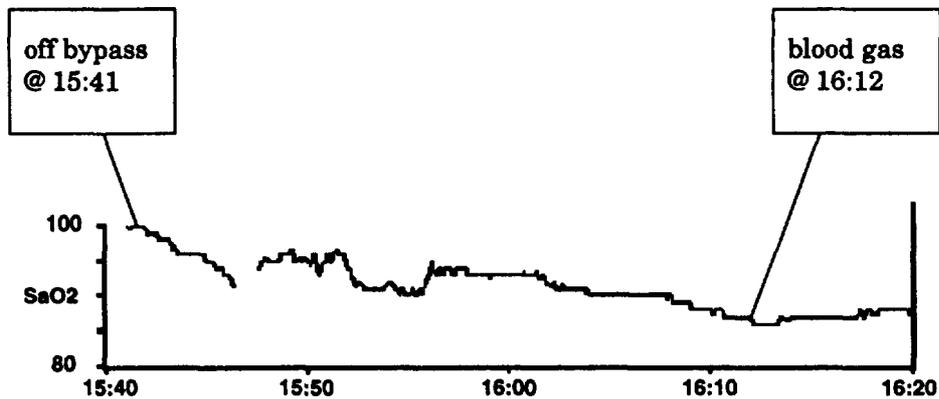
**Figure 1. A decline in patient's SaO2 level in the immediate post-bypass phase remained undetected for over 40 minutes, despite the clinician repeatedly examining the patient monitor screen.**

### 5.3.1 Chronology of clinical actions

The goal of event analysis is to uncover what the clinician was doing during this period, understand what his objectives were, and what factors may have contributed to his failure to notice the clinically significant fall in oxygen saturation.

The chronology is as follows. The subject normally had his patients coming off cardiac bypass with an infusion of an inotropic agent running to support cardiac contractility. In this particular case, the provision of inotropic support was not the surgeon's normal practice. The surgeon preferred instead to have the patient on an infusion of glyceryl trinitrate (GTN). Thus, counter to the subject's normal plan, an intravenous infusion of GTN was running at the end of the cardiac bypass.

The subject was concerned about the patient's cardiac performance, as demonstrated in the preoperative interview. In the absence of his preferred strategy, he gave the patient repeated doses of Calcium Chloride (CaCl) (Figure 2.) This was given for its inotropic action - rather than giving a more potent inotrope which would run counter to the surgeon's expressed practice. The clinician had a rule that by 20 minutes after bypass the various pressures should have returned to normal range. Since this had not happened, he finally felt justified to depart from the surgeon's practice and commence treatment with standard inotropes, and so administered adrenaline at 15:56.

At 16:12, he started to take a routine blood gas. In so doing noted that the $SaO_2$ value was lower than acceptable, and immediately reacted by increasing the patient's inspired oxygen ($FiO_2$). He then commenced hand bagging, and increasing the end-expiratory pressure on the ventilator (PEEP), aiming to treat pulmonary oedema. The routine blood gas value taken at 16:12 showed an arterial $PO_2$ of 7.2

There were several possible explanations for the prolonged drop in saturation. Firstly, the lungs may have been partially collapsed after bypass and choosing to hand bag the patient with a bit of positive pressure to try and re-expand them

136
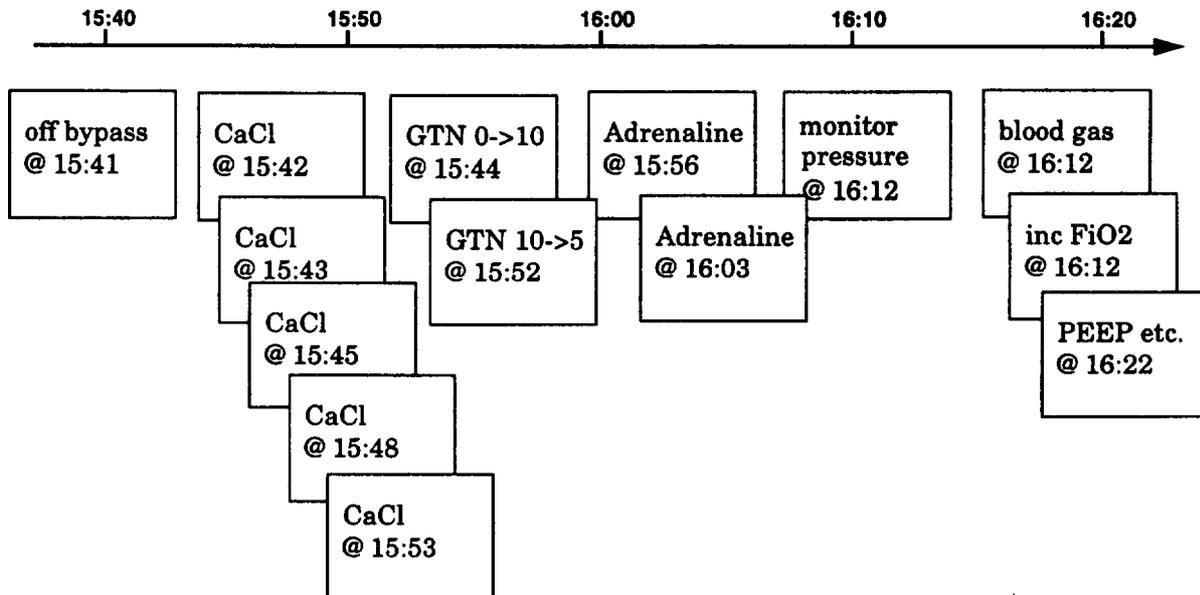
**Figure 2. Clinical actions taken during the period of evolving hypoxia**

previous history, the tightening of the left ventricle following surgery, and the high filling pressures, pulmonary oedema was a more likely explanation. The addition of PEEP and increase in FiO2 were commenced in response to this possibility.

### 5.4 An analysis

The next stage in our study involved analysing each event, and for this we needed to develop a descriptive language with which to bring some order to the data set. To this end we developed a classification system for clinical actions. Actions were classified as either *strategic* or *reactive*, based upon whether the subject had intended to carry the action out in a preplanned way, or whether the action was in response to a clinical event that had not been anticipated. Events were also classified according to whether they were associated with *monitoring*, *diagnosis* or *control*. This simple structure defines a vocabulary of 6 possible classifications for clinical events. (Figure 3.).

Using this vocabulary, the sequence of the subject clinician's actions was labelled for our example event (see Figure 4.). Much of the clinical activity throughout the drop in saturation was seen to be reactive. In other words, the subject was forced to deal with a combination of events for which he had not developed a preplanned response. In particular, he was faced with a poorly performing heart, and pulmonary oedema. However, as a skilled and experienced clinician, these factors on their own would not be expected to trouble him. The third confounding element of the scenario was that the subject was forced to adopt the therapeutic plan favoured by the surgeon. This was not the subject anaesthetist's normal way of

|            | strategic | reactive |
|------------|-----------|----------|
| diagnosis  | SD        | RD       |
| monitoring | SM        | RM       |
| control    | SC        | RC       |

**Figure 3. Clinical actions were classified according to whether they were based on a predefined strategy or were in reaction to unexpected clinical events**

working (having to place the patient on a regime of GTN for preload reduction, when he would rather give direct inotropic support).

It appears that this combination of circumstances led to an period of intense concentration, when the subject was preoccupied with the management of the patient's cardiovascular performance. One consequence of this preoccupation was that, despite looking repeatedly at the monitor, the subject did not detect the saturation trend.

The video and physiological data provide sufficient evidence to substantiate the hypothesis that the subject was preoccupied with issues of cardiac performance. From the physiological data, we can see he switches the CVP line to measure LA pressure repeatedly during the period between 16:05 and 16:10 (Figure 5.). The intraoperative video demonstrates the subject repeatedly spending long periods in front of the monitor, dispelling any doubts that the SaO2 value was missed because there had been any failure to use the monitoring equipment.

## 5.5 A possible explanation

With this single example, one would not want to draw anything more than preliminary conclusions. A common wisdom states that one of the major problems facing clinicians is *data overload* (in other words, they have difficulty taking in the large amount of data available to them on their patient monitoring systems). It may at first glance appear that this is a reasonable explanation for this event - that the clinician simply had too much to look at, and so missed a vital piece of clinical information. The evidence however, does not support this. There was ample time available to look at the monitor (in excess of an half hour).

Based upon evidence both from the actual operation and subsequent interviews, it seems reasonable to suggest that he was trying to develop a mental picture of cardiac function over the period in question. This involved the integration of several pieces of indirect evidence (the cardiac pressures) and therefore involved

RC -> SC -> RC -> RM -> SM -> RC

15:40    15:50    16:00    16:10    16:20

| off bypass @ 15:41 | CaCl @ 15:42 | GTN 0->10 @ 15:44 | Adrenaline @ 15:56 | monitor pressure @ 16:12 | blood gas @ 16:12 |

CaCl @ 15:43

GTN 10->5 @ 15:52

Adrenaline @ 16:03

inc FiO2 @ 16:12

CaCl @ 15:45

CaCl @ 15:48

CaCl @ 15:53

PEEP etc. @ 16:22
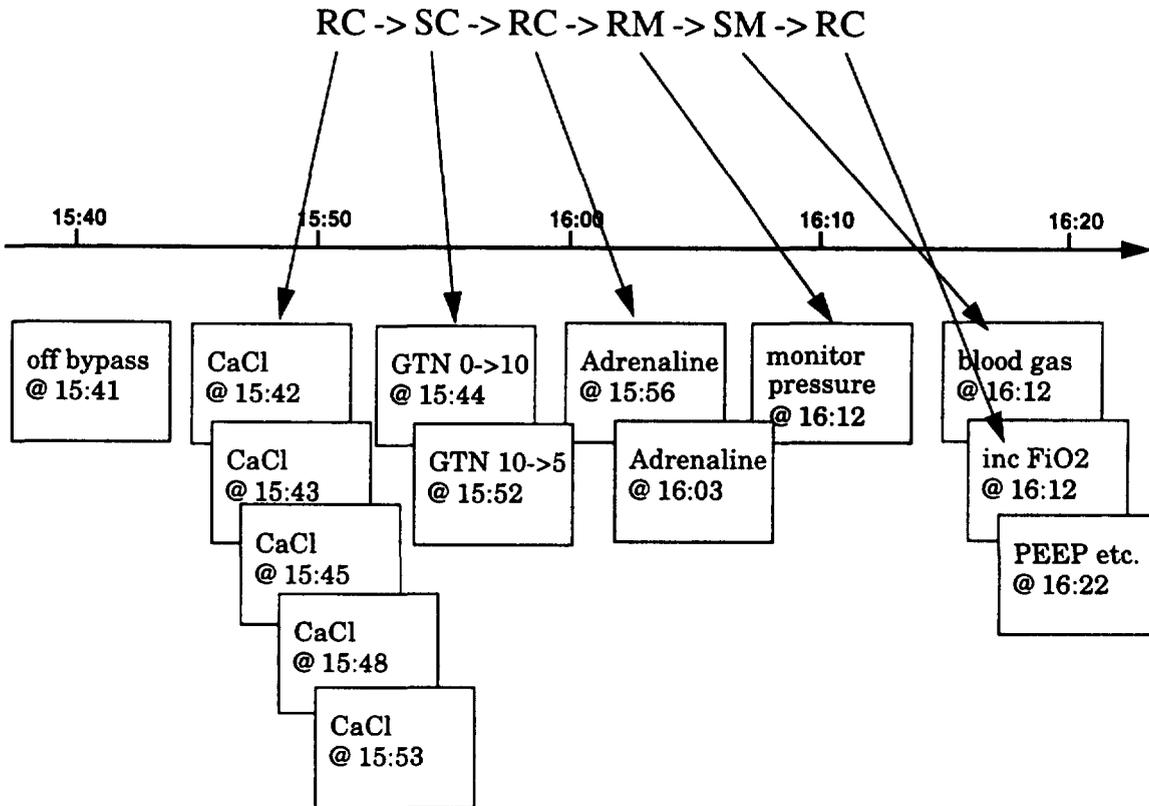
**Figure 4. Strategic(S) and Reactive (R) monitoring and control actions taken during the period of evolving hypoxia**

the constructing of some kind of mental model of the evolving cardiac haemodynamic status.

At particular points, there is clear evidence of focus on this activity - especially the switching behaviour around 16:05-16:10. This period of activity reflects the intensity with which the subject was pursuing the goal of assessing myocardial performance.

Cognitive psychological theories of attentional resource allocation suggests that the subject's difficulties in detecting the $SaO_2$ trend may have been due to this reasoning process swamping his available mental resources. So intense was the activity, that there were no mental resources available to attend to even simple monitoring activities like looking at parameters other than the pressures. Fundamental cognitive principles indicate that unless a clinician specifically allocates attentional resource to the perception of something like the $SaO_2$ numeric, that under such a heavy load the only thing the clinician will be able to reliably report is the numeric's colour [11].
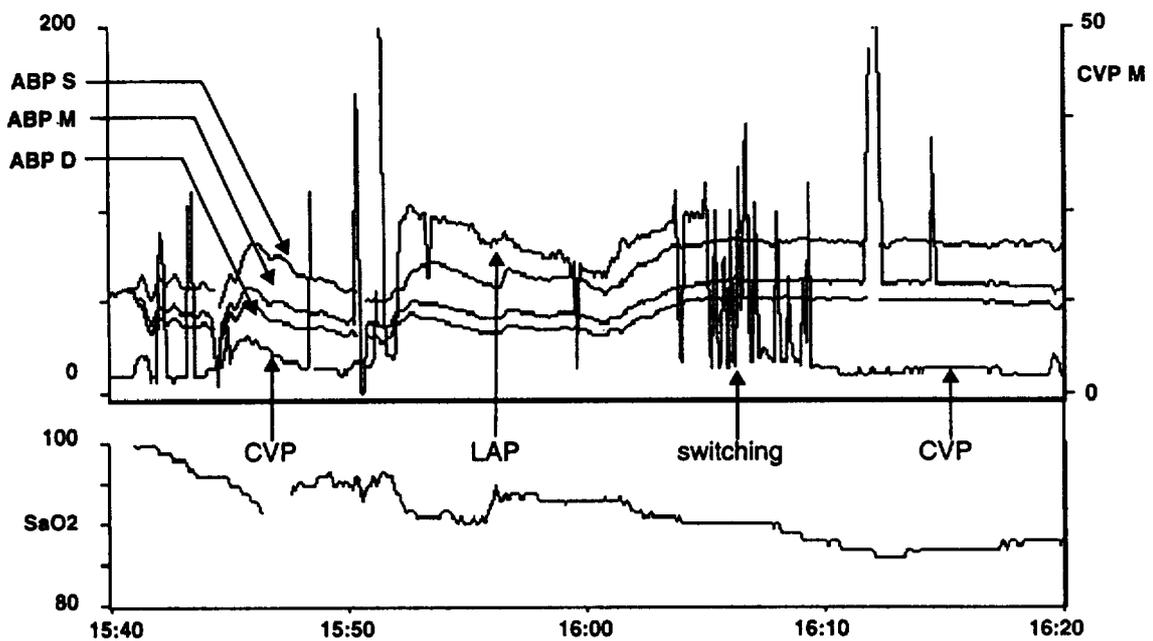
Figure 5. Blood pressure changes during the time of the dropping trend in SaO₂ revealed relatively poor cardiac performance immediately post-bypass. By 16:00 the ABP systolic value is below 80 mmHG. The lower curve in the blood pressure set shows central venous pressure (CVP). For the middle third of this curve, the clinician uses the CVP line to measure left atrial pressure (LAP), and towards the end of this middle third, he switches repeatedly between the two (as evidenced by the spikes in the trace). The switching behaviour indicates a period of intense concentration on these values.

We hypothesise that the clinician was not data overloaded; in fact he was suffering from *processing overload*. In other words, in the absence of direct data, the clinician's attention is being devoted to interpreting the meaning of several pieces of indirect information. This correlates well with the observations made at the beginning of this paper that situational assessment represents the primary effort for a decision maker; that most effort is devoted not to individual decisions such as diagnosis, but to establishing an understanding of the state of the world upon which such decisions can be based.

## 5.6 Implications from the example

The notion of data overload has long driven monitor designers to consider that the simplification and abstraction of monitored signals is an ideal to be aimed for. The idea is to reduce the burden placed on clinicians swamped with data by presenting them with information that has been condensed in some useful way.

The notion of processing overload developed from the previous example suggests that monitor displays should be constructed with a quite different aim - to assist clinicians with the cognitive tasks that tax them. This may indeed mean presenting

even more information than we expect, if that information assists them in completing the tasks before them.

In our example, the clear presentation of information about cardiac performance may have helped our subject clinician. For example, there may be clear ways of assisting with the visualisation of clinical data that makes estimation of cardiac contractility an easier cognitive task. Even simple grouping of data may be of help.

The concept of attentional load becomes a central one in this discussion, as it becomes clear that a clinician's attention is a finite resource whose use should be optimised. Thus, rather than presenting clinicians with alarms that immediately demand chunks of attention (and indeed may distract attention from important tasks at hand), we should aim to unload the clinician and free up attentional resources. If we recognise that the clinician is the one best able to handle the complexities of decision making in the domain, then we can begin to design systems that get out of the way, and instead help free up the clinician's inherent problem solving abilities.

# 6 Conclusion

For applications like patient monitoring, the complexity of the tasks for which the systems are used make it difficult to specify how information should be displayed. Yet it is unusual to find clinical support systems that have their designs based upon a detailed understanding of the way clinicians handle information. Some of this information does exist in the literature, but more specific information about the behaviours and tasks of clinicians is lacking.

In the clinical example presented here, the complexities of the processes involved in understanding how a clinician might come to miss a simple but critical piece of information are evident. If we are to design systems that support such complexity, then we need to invest in studies that make clinical decision processes explicit.

More generally, the development of advanced decision support systems in health care settings is a challenging task that requires a sensitivity to the unique characteristics of that environment. To date, most developments in this area have been motivated by the perceived need for decision support without any clear data pointing to its particular nature. What is now needed is a concerted effort to understand the nature of clinical practice, and in particular the nature of the user tasks that we hope to support in the future. Such studies need not be of an extravagant nature, but they do need to be methodical and focused in their approach. To assist this endeavour, there is a significant body of work in cognitive psychology and human computer interaction that can be harnessed. Our goal, as always, is to develop clinical information systems that have a genuine and positive impact on the delivery of health care, and with the right tools and approaches, this will be a singularly achievable goal.

# References

[1]J. S. Brown, Research that reinvents the corporation, *Harvard Business Review*, Jan-Feb, (1991), 102-111.

[2]E. Coiera, Editorial: Intelligent Monitoring and Control of Dynamic Physiological Systems, *Artificial Intelligence in Medicine*, 5, (1993), 1-8.

[3]E. W. Coiera, Incorporating user and dialogue models into the interface design of an intelligent patient monitor, *Medical Informatics*, 16, 4, (1991), 331-346.

[4] J. B. Cooper, C. D. Long, R. S. Newbower, Human Error in Anaesthesia Management, in B. I. Grundy, J. S. Gravenstein et al. (eds), Quality of Care in Anesthesia, (1982) 114-130.

[5]J.B. Cooper, R. S. Newbower, R. J. Kitz, An analysis of major errors and equipment failures in anaesthesia management: considerations for prevention and detection, *Anesthesiology*, 60,1, (1984), 34-42.

[6]A. Elstein, L. Shulman, A. Sprafka, *Medical Problem Solving - An Analysis of Clinical Reasoning*, Harvard University Press, Cambridge, Ma., (1979).

[7]D. Evans, V. Patel (eds), Cognitive Science in Medicine: Biomedical Modelling, MIT Press, (1989).

[8]D. Fafchamps, C. Young, P. Tang, Modelling Work Practices: Input to the design of a Physician's Workstation, *Proceedings of 15th Symposium on Computer Applications in Medical Care*, (1991).

[9]D. M. Gaba, M. Maxwell, A. DeAnda, Anesthetic Mishaps: Breaking the Chain of Accident Evolution, *Anesthesiology*, 66, 5, (1987), 670-676.

[10]D. M. Gaba, Dynamic Decision-making in Anesthesiology: Cognitive Models and Training Approaches, in D. A. Evans, V. L. Patel (eds), *Advanced Models of Cognition for Medical Training and Practice*, Berlin, Springer Verlag, (to appear).

[11]G. Higgins, The Cognitive Psychology of the Monitoring Environment, *Proceedings of the 4th Annual Meeting, European Society for Computing and Technology in Anaesthesia and Intensive Care*, Salzburg, 6-9 October, (1993)

[12]D. Kahneman, P. Slovic, A. Tversky (eds), Judgment under uncertainty: heuristics and biases, Cambridge New York, Cambridge University Press, (1982).

[13]G. A. Klein, R. Calderwood, Decision models: some lessons from the field, *IEEE Trans. Sys. Man Cyb*, 21, 5, (1991), 1018-1026.

[14]J. S. McDonald, R.R. Dzwonczyk, A time and motion study of the Anaesthetist's intraoperative time, *British Journal of Anaesth*esia, 61, (1988), 738-742.

[15]R.E. Nisbett and T. D. Wilson. Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84:231–259, 1977.

[16] M.K. Sykes, Essential Monitoring, *Br. J. Anaesth.*, (1987), 59, 901-912.

[17]C. D. Wickens, *Engineering psychology and human performance*, Harper Collins, New York, (1992).

[18]A. W. Wicker, Attitudes v. Actions: the Relationship of Verbal and Overt Responses to Attitude Objects, in *Attitudes*, N. Warren, M. Jahoda (eds), 2nd. Ed. Penguin Books Ltd., Middlesex, (1976).

# A PC-BASED NEURAL NETWORK MEDICAL DIAGNOSTICS SYSTEM

C. C. Neocleous [1] and C. N. Schizas [2]

## ABSTRACT

A general purpose neural network program that may be used as a powerful PC tool in medical (and other) data classification / diagnostics is presented. The system has been developed in the Microsoft EXCEL (V.4) Macro Language. It uses the standard backpropagation algorithm, suitably modified for a spreadsheet. The program is versatile in the sense that it enables the user to specify different operating parameters. More specifically, the system provides options for selecting: i) either one or two hidden layers, ii) the number of neurodes in each layer, iii) whether normalization is to be used and how it will be implemented, iv) the gain, momentum, biases, and upper and lower values of the initial random weights, and v) the desired accuracy. The effects of changes in above, can easily be observed and demonstrated through the use of a system of user friendly interfaces, and output presentation schemes such as animated graphs. A number of case studies in medical diagnostics are presented. These are based on a reanalysis of previously published data in the fields of EMG signal classification and on Clinical and Clinical/Laboratory data for neuromuscular disorder classification. Investigation of the system sensitivity to parametric variation of the gain, initial random weights, and the number of neurodes in the hidden layer has been done and presented.

## INTRODUCTION

This work is a further extension of previous work done, aiming at providing a powerful Artificial Neural Network (ANN) design and analysis tool. [Neocleous and Schizas (1993)]. In recent years there has been a steady growth in the availability of ANN programs on PC platforms, such as the Brainmaker, Autonet, Neuroshell, Winbrain, Neuralyst, and Braincell. The last two have been written in the EXCEL macro language. These programs have different degrees of complexity, versatility, and cost. Their primary target is business analysts, medical practitioners, and engineers. In the field of medical diagnosis, where the user is not usually well versed on either PC's or ANN, it is important that such programs are simple, effective and versatile. Work has been done recently, in an effort to develop PC based ANN medical diagnostic systems. More specifically, in the area of classification, work has been done by [Poli, et. al. (1991)], [Bounds et al. (1988)], [Schizas et. al. (1990)], [Schizas et. al. (1992)], [Schizas et. al. (1993].

[1] Department of Mechanical Engineering, Higher Technical Institute, P.O.Box 2423, Nicosia, Cyprus.
Email: costas@jupiter.cca.ucy.cy
Fax +357-2-494953

[2] Department of Computer Science, University of Cyprus, 75 Kallipoleos, Nicosia, Cyprus.
Email: schizas@jupiter.cca.ucy.cy
Fax +357-2-360881

## SYSTEM DESCRIPTION

The system is composed of two parts (see Figure 1) . The first part (main program) is a macro file that has the overall supervision and control of the system, i.e. its auto-generation, data exchange, and the proper implementation of the backpropagation algorithm. The second part is a group of interactive files that incorporate the network architecture, the forward path, the feedback path, the various operating parameters and the learned weights. The system has also a number of support utilities such as statistical information tables and graphs, animated graphs that continuously indicate and monitor the learning progress, and means for automatic evaluation of the correct classification yield.

```
                    ┌─────────────────────┐
                    │    MAIN PROGRAM     │
                    └─────────────────────┘
                              │
                              ▼
 ┌────────────────────────────────────────────────────────┐
 │  ┌──────────────┐  ┌──────────────┐  ┌──────────────┐  │
 │  │ CONTROLS FILE│  │  DATA FILE   │  │AUXILIARIES FILE│ │
 │  └──────────────┘  └──────────────┘  └──────────────┘  │
 │                                                        │
 │      ┌──────────────┐      ┌──────────────┐            │
 │      │ FORWARD FILE │      │ FEEDBACK FILE│            │
 │      └──────────────┘      └──────────────┘            │
 └────────────────────────────────────────────────────────┘
```
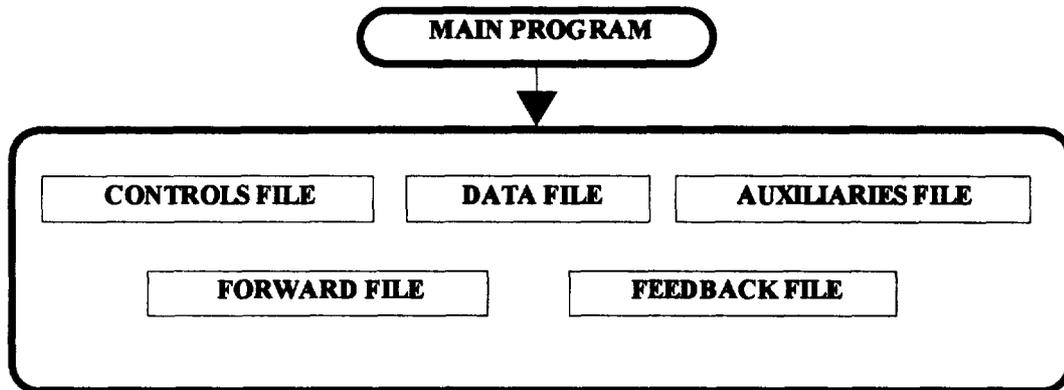
Figure 1. Block Diagram of the System

Figures 2 and 3 show snapshots of a typical Pattern Learning Curve and a typical TSS Vs Epoch curve, while Table 1 shows a window from the main input/output control file. It is reminded that an epoch is reached when all training patterns have been applied (and processed) once, while TSS is the Total Sum of Squares of the error between the desired output and the actual output, calculated for all training patterns during an epoch.
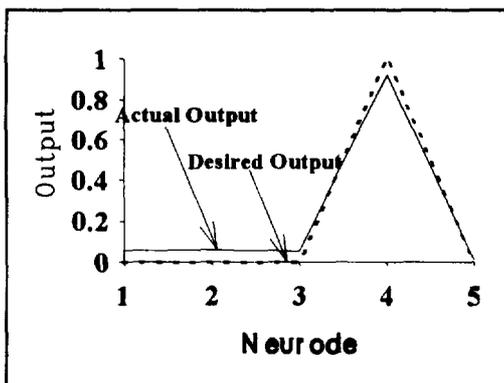


Figure 2. Typical Pattern Learning Curve
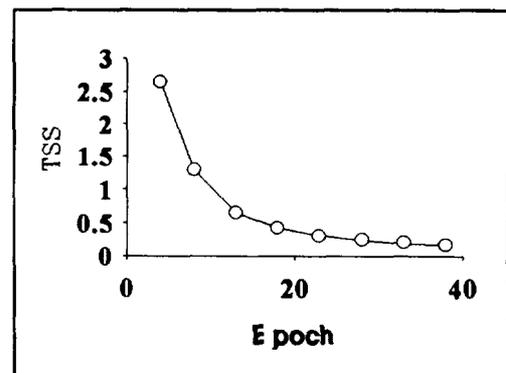
Figure 3. Typical TSS vs Epoch Curve

```
                    PROGRAM DEVELOPED BY
            COSTAS NEOCLEOUS AND CHRISTOS SCHIZAS
          UNIVERSITY of CYPRUS  -  HIGHER TECHNICAL INSTITUTE


TOPOLOGICAL PARAMETERS              DESIRED PERFORMANCE PARAMETERS      ACTUAL   DESIRED
No. OF HIDDEN LAYERS=           1   MAX No OF ITERATIONS/PATTERN=   3   OUTPUT   OUTPUT
No. OF NEURODES IN INPUT LAYER=  16  MAX NUMBER OF EPOCHS=        1000   0.029
No. OF NEURODES IN HIDDEN LAYER= 40  MAXIMUM PSS=                  0.1   0.027
No. OF NEURODES IN OUTPUT LAYER= 5   MAXIMUM TSS=                  0.1   0.011
No. OF TRAINING PATTERNS=        41  DO YOU WANT STATISTICS (Y/N) ?  Y   0.009
No. OF TESTING PATTERNS=         30  EPOCH LOG STEP=                5    1.012      1


OPERATING PARAMETERS               ACTUAL TRAINING PERFORMANCE
LEARNING RATE=                0.5  PATTERN UNDER TRAINING=       P35
MOMENTUM FACTOR=              0.1  ACTUAL PSS=
UPPER OF RANDOM No=           0.3  ACTUAL No. OF ITERATIONS=      1
LOWER OF RANDOM No=          -0.3  ACTUAL EPOCHS DONE=           39
BIAS OF FIRST LAYER=          -1   ACTUAL TSS=                  0.11
BIAS OF SECOND LAYER=         -1
BIAS OF THIRD LAYER=
BIAS OF FOURTH LAYER=


OTHER PARAMETERS
TYPE OF ACTIVATION FUNCTION=   SIGMOIDAL
TYPE OF NORMALIZATION=         NO NORMALIZATION
```

Table 1. The main control window

A user may temporarily freeze the learning process, modify any of the parameters in the window shown in Table 1, and either rebuild the complete structure or continue the learning process with the updated parameters. This feature enables the user to have complete control on the workings of the program, and facilitates elaborate experimentation on Artificial Neural Networks.

## APPLICATIONS IN MEDICAL DIAGNOSTICS

### 1. EMG SIGNATURE CLASSIFICATION

Two sets of data, a training set and an evaluation set have been used. These are the same data used by Pattichis (1992). A network structure of 14 input, 40 hidden and 3 output neurodes has been used. The number of training cases were 24 and the evaluation 10. The TSS versus Epoch learning progress curve for this model is shown in Figure 4.
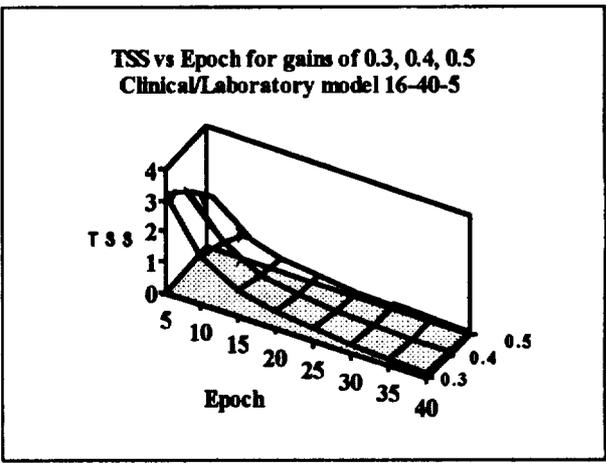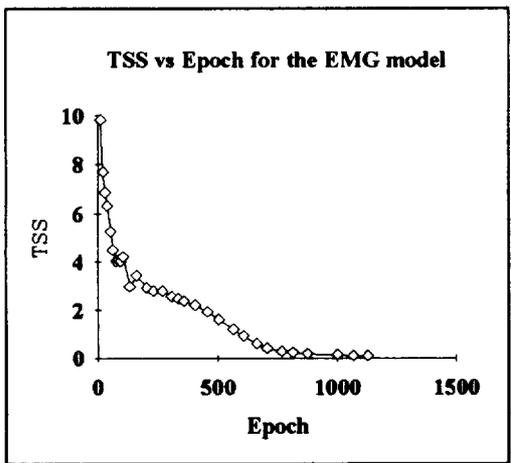


Figure 4. TSS vs Epoch for the EMG 14-40-3 model.    Figure 5. TSS vs Epoch for gains 0.3, 0.4, 0.5 Clinical/Laboratory - Model 16-40-5

The network's performance with this model is shown in Table 2.

| % CORRECT CLASSIFICATIONS ON THE TRAINING SET | % CORRECT CLASSIFICATIONS ON THE TESTING SET |
|---|---|
| 100 | 80 |

Table 2. EMG 14-40-3 model Network Yield.

This result is of some improvement on Pattichis's work, since the network employed in our model is simpler. It employs only one hidden layer of fewer neurodes (a 14-40-3 model) in comparison to a two hidden layer equivalent model of 14-100-10-3. This difference may be attributed to the fact that we have used a slightly different learning strategy and different values of the initial random weights. We have forced the system to attempt to learn in few iterations, any given applied pattern, prior to moving to the next pattern.

## 2. CLINICAL NEUROMUSCULAR DISORDER CLASSIFICATION

The training and evaluation set data developed by Schizas and Pattichis (1994) have been used. These involve a training set of 41 and an evaluation set of 30 subjects. We have used a systematic series of different models having 8 input, 5 output, 40 or 50 hidden neurodes, and various values of gains (0.3, 0.4, 0.5). Furthermore we have preprocessed the age field so that the input vectors have a range from 0 to 10. We have also experimented on models using an adaptive gain in order to speed up the learning process. In these models the gain is adapted on-line depending on the requirements of the learning progress. The gain is slightly increased or decreased depending on the value of TSS at the completion of each epoch. The diagnostic yield for different models is shown in Table 3.

| CLINICAL MODEL | % CORRECT CLASSIFICATIONS ON THE TRAINING SET | % CORRECT CLASSIFICATIONS ON THE TESTING SET |
|---|---|---|
| 8-40-5 | 100 | 93.3 |
| 8-50-5 | 100 | 93.3 |

Table 3. Diagnostic yield of the clinical neuromuscular disorder classification models.

| CLINICAL LABORATORY MODEL | % CORRECT CLASSIFICATIONS ON THE TRAINING SET | % CORRECT CLASSIFICATIONS ON THE TESTING SET |
|---|---|---|
| 16-40-5 | 100 | 96.7 |
| 16-50-5 | 100 | 100 |

Table 4. Diagnostic yield of the clinical/laboratory neuromuscular disorder classification models.

## 3. CLINICAL/LABORATORY NEUROMUSCULAR DISORDER CLASSIFICATION

Similar approaches as for the clinical models have been used. The number of input neurodes is 16 instead of 8 in order to accommodate the laboratory data. The diagnostic yield is shown in Table 4. It is noted that 50 hidden layer neurodes are needed so that the system will be able to generalize well (100% on the unknown evaluation set). What is though of interest is the fact that the system, which is using an on-line learning procedure, as opposed to the more familiar batch proc-

essing, has achieved this performance in only few epochs ( about 50 epochs). Figure 5 shows typical learning curves for some models.

## CONCLUSIONS

This work has shown that a PC based ANN tool for medical diagnostics applications is feasible. It may prove a powerful element in the hands of a medical doctor, to be used not only for diagnosis, but also for prognosis. The size of the system as well as the network structure are crucial factors that govern the overall effectiveness of the job. Our system has demonstrated the capability to be used directly as an application tool, as well as a framework for explorations in Artificial Neural Networks. The fact that the system has been embedded in a powerful spreadsheet, enables the user to quickly input data, analyze the results and conduct all sorts of statistical manipulations as they may be required.

## REFERENCES

[1] **Bounds G., Lloyds J., Mathew B., Wadell G.** (1988), "A Multilayer Perceptron for the Diagnosis of Low Back Pain", Proc. IEEE Int. Conf. On Neural Networks, pp. II-481-489, San Diego.

[2] **Neocleous C., Schizas C.** (1993), "Computer Aided Instruction of Artificial Neural Networks", 2nd International Conference on Computer Applications in Engineering, Nicosia.

[3] **Pattichis C.** (1992), "Artificial Neural Networks in Clinical Electromyography". Phd Thesis.

[4] **Poli R., Cagoni s., Livi R., Coppini G., Valli G.** (1991), "A Neural Network Expert System for Diagnosing and treating Hypertension", Computer, pp. 64-71, March 1991.

[5] **Schizas Chr., Pattichis C., Schofield I., Fawcett P., and Middletton L.** (1990), "Artificial Neural Nets in Computer Aided Macro Motor Unit Potential Classification", IEEE Engineering in Medicine and Biology Magazine, Vol.9, No.3, 31-38

# An Intelligent Medical System for Diagnosis of Bone Diseases

I. Hatzilygeroudis[1] [2], P. J. Vassilakos[3], A. Tsakalidis[1] [2]

(1) University of Patras, School of Engineering, Dept of Computer Engin. &
Informatics, 26500 Patras, Greece.
(2) Computer Technology Institute, P.O. Box 1122, 26110 Patras, Greece.
(3) Regional University Hospital of Patras, Dept of Nuclear Medicine,
Patras, Greece.

## Abstract

In this paper, aspects of the design of an intelligent medical system for diagnosis of bone diseases that can be detected by scintigraphic images are presented. The system comprises three major parts: a user interface (UI), a database management system (DBMS), and an expert system (ES). The DBMS is used for manipulation of various patient data. A number of patient cases are selected as prototype and stored in a separate database. Diagnosis is performed via the ES, called XBONE, based on patient data. Knowledge is represented via an integrated formalism that combines production rules and a neural network. This results in better representation, and facilitates knowledge acquisition and maintenance.

## 1. Introduction

Computer-based methods are increasingly used to improve the quality of medical services. Those methods include both conventional techniques, such as database management systems (DBMSs), and artificial intelligence (AI) techniques, such as knowledge-based systems (KBSs) or expert systems (ESs) [1, 6].

Medical diagnosis is a very active field as far as introduction of the above techniques is concerned. In medical diagnosis, DBMSs are used for storing, retrieving and generally manipulating patient data, whereas ESs are mainly used for performing diagnoses based on patient data, since they can naturally represent the way experts reason. Diagnosis of bone diseases is greatly facilitated by the use of nuclear medicine methods, more specifically by the use of scintigraphic images (or scintigrams or scans)[1], and a number of relevant expert systems have been developed [2, 7].

In this paper, an intelligent medical system for diagnosis of bone diseases that uses the above methods is presented. The structure of the paper is as follows. In section 2 the medical knowledge involved is outlined. In section 3 the architecture of the system is discussed. Section 4 deals with the knowledge representation formalism used by the ES. Finally, section 5 concludes.

## 2. Medical Knowledge

### 2.1 Patient Data

Patient data concerns information related to the patient and can be distinguished in three types: demograhic, clinical and nuclear medicine (NMI) data. *Demographic data* concerns information such as patient's age, sex etc. *Clinical data* is further distinguished in physical

---

[1]Terms interchangeably used.

findings and laboratory results. *Physical findings* are those detected by a physical examination of the patient, like e.g. the existence and the kind of a pain etc. *Laboratory results* are those detected via laboratory tests, like e.g. blood tests etc.

Finally, *NMI data* is that extracted from scintigraphic images. The images are acquired by a γ-camera and depict the concentration of an administered radio-pharmaceutical (99m Tc-MDP) on patient's osseous tissue. NMI data concerns description of the concentration patterns depicted in the scintigrams. Description consists in an account of two types of observed features, qualitative and quantitative.

## 2.2 Heuristic knowledge

Heuristic knowledge represents experience accumulated through years and concerns the way an expert uses the above knowledge to make diagnoses. A diagnosis basically consists in relating patient data with corresponding diseases. Especially in the case of the NMI data, a physician should recognise the *characteristic patterns* of the radio-pharmaceutical concentration in the scintigrams. Each pattern gives an indication of the category of the suspected bone disease. Diagnosis of bone diseases is two-fold. An initial diagnosis is made based on the demographic and clinical data of the patient. This is then used either to specify the kind of the scan needed (simple or 3-phase) or to be compared with a later diagnosis based on the NMI data. This later diagnosis may or may not coincide with the initial. It is then up to the doctor to decide which is the right one.

## 3. System Description

The architecture of the system is illustrated in fig.1. It consists of three major parts: a User Interface (UI), a Data Base Management System (DBMS) and an Expert System (ES), called XBONE.

The DBMS is used to manipulate patient data. Scintigrams acquired by a γ-camera are transferred to the system. Although a γ-camera is typically accompanied by a dedicated computer, we developed an interface with a PC [5] so that acquired images can be directly introduced into the system and stored into the Patients Database (PDB) via the Data & Image Manager (DIM). The rest of the patient data is also introduced in the PDB by the user-physician.

A number of patient cases are selected from the PDB as *prototype cases* by the user-physician, after final diagnoses have been made, and stored in the Prototype Cases Database (PCDB). The selection is based on a standard protocol [10]. Prototype cases are also classified according to the disease category they concern. They are then used to identify characteristic scintigraphic patterns for each category and subcategory of the diseases.

The ES is used to perform diagnosis. The Knowledge Base (KB) contains the heuristic knowledge for diagnosing bone diseases. An integrated knowledge representation (KR) formalism based on production rules, the most widely employed KR formalism by ESs [3], is used (see next section). There are two rule sets in the KB, one dealing with the initial diagnosis and the other with the later one. The Working Database (WDB) contains the case-specific data, that is the (initial) patient data, partial conclusions, data given by the user, and any other information relevant to the case under consideration. Patient data are transferred from the PDB. The Inference Engine (IE) uses the available knowledge to draw conclusions and make diagnoses. The Explanation Mechanism (EM) creates explanations when asked by the user to do so. The Training Mechanism (TM) is used for rule training (see next section).

Finally, the UI performs a number of functions. It allows interaction with the PDB and the PCDB through the DIM, so that the user can manipulate patient data and the prototype cases. Also, the user can interact with the IE to start a diagnosis process as well as with the EM to ask "why" and "how" questions during the process. Finally, the user can use the TM to perform rule training.
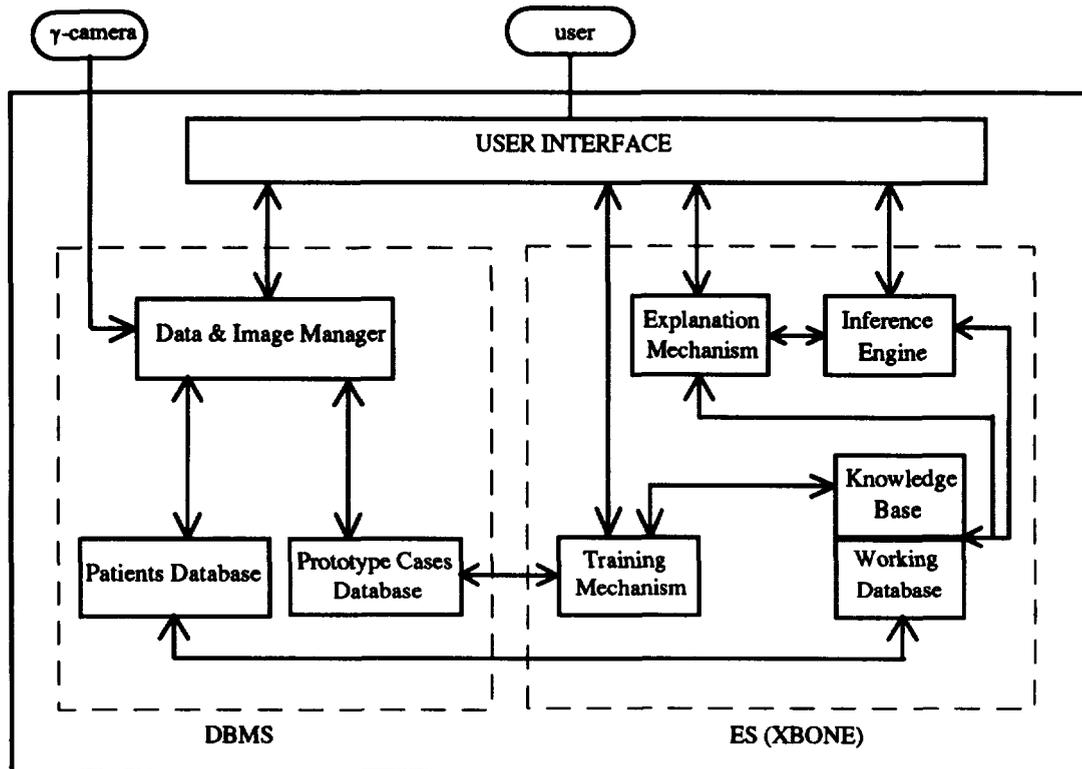


Fig.1 The Architecture of the System

## 4. Knowledge Representation in XBONE

### 4.1 Motivation

A situation met in medical diagnosis is the following. There are a number of symptoms (i.e. patient data) that all contribute in diagnosing a disease. However, not all of the symptoms have the same significance. For example, a symptom S1 may give much stronger evidence for diagnosing a disease D than a symptom S2. Also, some combinations of symptoms may give stronger evidence than others in diagnosing the disease. To be able to represent this situation in a production rules formalism, we introduce a factor assigned to each condition of a rule, representing its significance in drawing the conclusion. To determine the values of and manipulate those factors, the adaline unit (see e.g. [4, 8]) is employed.

### 4.2 The Integrated Formalism

Each rule is considered as an adaline unit (fig.2). The inputs $c_j$, $i=1,...,n$ of the unit are the conditions of the rule. Each condition $c_j$ is assigned a number $sf_j$, called a *significance factor*,

150

corresponding to the weight of the input of the adaline unit. Moreover, each rule itself is assigned a number $sf_0$, called the *bias factor*, corresponding to the weight of the bias input ($c_0$ = 1) of the unit. Each input takes a value from the following set of discrete values:
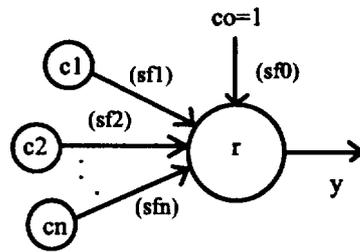


**Fig.2.** A Rule as an Adaline Unit

$$c_i = \begin{cases} 1 & \text{if condition is true} \\ -1 & \text{if condition is false} \\ 0 & \text{if value is unknown} \end{cases}$$

This gives the opportunity to distinguish between the falsity and the absence of a condition, in contrast to conventional rule-based systems. The output y is calculated as the weighted sum of the inputs filtered by a threshold function [4, 8]. Thus, the output can take one of two values, '0' and '1', representing failure and success of the rule respectively.

The general syntax of a rule is the following[2]:

<rule> ::= [(<bias-factor>)] **if** <conditions> **then** <conclusions>

<conditions> ::= <condition> {**and** <condition>}

<conclusions> ::= <conclusion> {**and** <conclusion>}

<condition> ::= <object> <l-operator> <value> [(<significance-factor>)]

<conclusion> ::= <object> <r-operator> <value>.

<object> acts as a variable and represents a concept in the domain. <l-operator> and <r-operator> denote a symbolic (e.g. is, isnot) or a numeric operator (e.g. <, >=). <value> denotes a value of the <object>, numeric or symbolic. Finally, <bias-factor> and <significance-factor> are real numbers. Since significance factors are optional, conventional production rules can be used.

The factors assigned to the rules are determined by the TM. Each rule is individually trained. To this end, a number of prototype training patterns, called the *training set*, are supplied for each rule. For example, [0 1 1] could be a training pattern for the rule **if** S1 **and** S2 **then** D. The standard LMS learning algorithm [4, 8] is used.

## 5. Conclusions

In this paper, an intelligent medical system for diagnosis of bone diseases that can be detected by scintigraphic images is presented. The system is still under development by the collaboration of the Depts of Computer Engineering & Informatics and Nuclear Medicine of

---

[2]A BNF notation is used here, where '[]' denotes optional occurrence and '{}' zero, one or more occurrences of the enclosed expression. Also, '<>' denotes a nonterminal symbol. All other symbols are terminal.

the University of Patras (e.g. [9]), and it is intended for use by nuclear medicine physicians. It is currently implemented in the C language on a PC.

A DBMS is used for patient data manipulation. Scintigrams are directly acquired by the system. However, NMI data are extracted by the user-physician. Although there are systems using computer-based methods for NMI data extraction [7], image processing techniques are not very reliable, due to the inherent noise and the very large number of possible normal and abnormal situations, so that extraction of the NMI data by the physicians is preferred [2]. The prototype cases can be used for characteristic patterns standardisation for each category of diseases.

An expert system is used for performing diagnosis. Knowledge is represented via a formalism integrating production rules and the adaline neural network. This results in better representation, since one can represent more complex relations between conditions (symptoms). Because determination of the significance factors is automated, knowledge acquisition is also partially automated, thus helping the experts to express their knowledge. All that an expert has to do is to determine the symptoms involved in diagnosing a disease and the training set. Also, KB maintenance is easier, since rules can be periodically trained using the updated PCDB.

## References

[1]  Clancey W. J. and Shortliffe E. H., (Eds), *Readings in Medical AI: The First Decade*, Addison-Wesley, Reading, MA, 1984.

[2]  Ellam S.V. and Maisen M.N., "A Knowledge-based System to Assist in the Diagnosis of Thyroid Disease from a Radioisotope Scan", in Pretschner D.P. and Urrutia B., (Eds), *Knowledge-based systems to aid medical image analysis*, vol.1, Commission of the European Community, 1990.

[3]  Hayes-Roth F., "Rule-Based Systems", *CACM*, vol. 28, 1985, 921-932.

[4]  Hecht-Nielsen R., *Neurocomputing*, Addison-Wesley, Reading, MA, 1990.

[5]  Metallidis S. and Gorila K., *Acquisition and Processing of Images from a* γ *-camera to a PC*, Diploma Thesis (in Greek), Dept of Computer Engineering & Informatics, University of Patras, 1990.

[6]  Miller P. L., (Ed), *Selected Topics in Medical AI*, Springer-Verlag, NY, 1988.

[7]  Reiber J.H.C., Bloom G., Gerbrands S.S., Backer E., van de Herik H.J., Reijs A.E.M., van der Feltz I. and Fioretti P., "An Expert System Approach for the Objective Interpretation of Thallium-201 Scintigrams", in the same as [2].

[8]  Simpson P. K., *Artificial Neural Systems: foundations, paradigms, applications, and implementations*, Pergamon Press, NY, 1990.

[9]  Siozos A. and I. Tzanetis, *XBONE: An Expert System for Diagnosing Bone Diseases*, Diploma Thesis (in Greek), Dept of Computer Engineering & Informatics, University of Patras, 1992.

[10] Vassilakos P.J., "Quality Assurance of Bone Phantom Data Base", in Britton K.E., *COST B2 - 4th Annual Report*, Commission of the European Community, 1992.

# BIOLOGICAL APPLICATIONS OF THE SENSIVITY ALGORITHM FOR THE ARTIFICIAL NEURAL NETWORK

M. Giacomini, T. Parisini, C. Ruggiero, R. Sacile

Dept. Communication, Computer and System Science
University of Genova, Italy.

## Abstract

The present paper focuses on the determination of the most appropriate size of artificial neural networks, taking into account that small nets are more efficient both in forward computation and learning and that neural nets which are too large tend to have poor generalization ability. Specifically, an application of a pruning algorithim is presented for the determination of the main classification features in three biological fields.

## Introduction

Artificial neural networks have proved to be a reliable tool for classification in many fields, above all in those cases where it's impossible, a priori, to define features that can lead to a correct classification. This is the case of biological applications, where often the high number of features present in the data may make it difficult to assess the most significant ones. In this case the use of artificial neural networks can be very useful, since this tool can cope with a great number of inputs independently of a preliminary assessment of the importance of each one.
One of the greatest problems in using an artificial neural network is the determination of its most appropriate size. In theory, if a problem is solvable with a network of a given size, it can also be solved by a larger net which imbeds the smaller one with all the redundant connections having a zero weight. However, the learning algorithm will typically produce a different structure, with nonvanishing synaptic weights on all connections, hiding therefore the existence of a smaller size neural network.
The advantages of the use of the smallest reliable network can be summarized as follows:
-a small net is more efficient both in forward computations and in learning;
-neural nets which are too large tend to have poor generalization ability (specially if they are trained with a small set of patterns);
It may be possible that a smaller net will exibit a behaviour that can be described by a simple set of rules.
The last feature above is very important in biology where artificial neural networks can give the opportunity of determining the most important features for classification in the considered field.
In this paper we present an application of the pruning algorithm by Karnin [1] for the determination of the main classification features in three biological fields: the determination of the protein secondary structure starting from NMR data, the determination of the same secondary structure starting from the primary structure, the classification of marine bacteria starting from their gaschromatographic profile.

## Description of the sensitivity algorithm used.

To determine the most appropriate size of a neural network for a given case, we have chosen to estimate the sensitivity of the global error cost function to the inclusion/exclusion of each synapse in the considered network.
We use additional arrays that keep track of the incremental changes to the synaptic weights during a single successful step of the SuperSAB learning [2]. The synapses are then ordered by decreasing sensitivity values so that the network can be efficiently pruned by discarding the last

items of the list. This procedure does not require any modification of the cost function, it does not interfere with the learning process and its computational weight is negligible.
Defining the global error of the net as:

$$E = \sum_p \sum_i \left(o_{pi} - t_{pi}\right)^2$$

(where p is one particular pattern, $o_{pi}$ is the desidered output neuron i in response to pattern p, $t_{pi}$ is the present value of the output neuron i), we can define the sensitivity as

$$s = - \frac{E(w^f) - E(w^i)}{w^f - w^i} w^f$$

However this formula can be applied if only one weight (w) changes whereas all other weights are constant; but this is not the case of normal learning. As shown in [1] the sensitivity for the weight $w_{ij}$ can be estimated as:

$$\hat{s}_{ij} = - \sum_{n=0}^{N-1} \left[ \frac{\partial E}{\partial w_{ij}}(n) \, \Delta w_{ij}(n) \right] \frac{w^f_{ij}}{w^f_{ij} - w^i_{ij}}$$

where all the terms are always calculated in each learning process, so the only extra computational demand is the summation.
As we use the SuperSAB algorithm for learning, the calculations above are extended only to the successful steps in which the changes to the weights are actually applied.


## Determination of the secondary structure of protein from NMR data.

At present the recent - and not yet completely assessed - techniques of NMR are being used to study proteins in their natural state, that is in solution. The raw information from NMR describes the interactions among spatially close atomic nucleic-specially hydrogen, due to their abundance and strong response to NMR-and the interactions between nuclei displaced by only a few chemical bonds. The interpretation of this information is ultimately important.
The data from the first type of interactions can be used to determine the type of residue bound to a particular amino-acid; the data from the second type of interactions can be used to determine the distances from particular pair of atoms and these distances collect all information about the secondary structure of an amino-acid. It is possible to obtain from each amino-acid four different distances between hydrogen atoms. These distances are measured between: the two hydrogen atoms bound to the azote atoms at the beginning of the side chain of two following amino-acids ($d_{NN}$); the hydrogen atom bound to the central carbonium atom and the hydrogen atom bound to the azote atom at the beginnning of the side chain of the following amino-acid ($d_{aN}$); the hydrogen atom bound to the second carbonium atom and the same azote atom ($d_{bN}$); the hydrogen atom bound to the azote atom at the beginning of the side chain and the hidrogen atom bound to the central carbonium atom of the same aminoacid ($d_{Na}$).
The knowledge used in examinig NMR raw details often largely heuristic, reflecting both the incompleteness of the data available and the incertainty as to the best method to interpret it. In this situation of absence of a well assessed method for the analysis of these data, we thought it would be worthwile to use a neural network which learns from a certain number of examples the way in which it is possible to determine in which state an amminoacid can be, given the four characterizing distances.
We have already demonstrated the possibility of such a net to determine the correct state of an amminoacid [3]. Here we want to present the first results of the sensitivity check to determine if all these distances have the same importance or not. Our first results show that the variation in

sensitivity of the four distances is not significant. This indicates that their importance for the determination of the secondary state is about the same.

## Prediction of secondary structure of protein from their primary structure.

A great number of computer based strategies have been so far employed to predict the secondary structure of a protein starting from the amino acid sequence, but no method has been shown to be more efficient than others [4,5]. In the last few years, methods based on neural nets for secondary structure prediction have been introduced [6,7].

In protein secondary structure prediction the goal is to find a general mapping from the amminoacid sequence to the secondary structure. These two structures are represented in vectorial form in order to obtain a match between a sliding window $2n+1$ aminoacids in the primary structure and the correspondent secondary structure state of central amino-acid in $(n+1)$ the position of the window [8].

The application of the sensitivity check to this problem aims to determine the best width of the input window to avoid the huge number of calculations due to an overestimated value of n and not to miss important information about the cooperation of nearby aminoacids. After some trials we can affirm that sensitivity is quite constant for the four aminoacids nearer to the central one (for both sides) and decreases considerably for farther aminoacids, becoming almost negligible for amminoacid farther than 6 places than the central one. So we conclude that the best width for the input window is of 13 aminoacids.

## Marine bacteria

Gas liquid chromatography has been found to be very useful for a correct and speedy classification of the bacterial genus; in fact all the bacterial genus do not contain the same quantities of the different fatty acids. So we have trained a neural network with gas-chromatographic data (that are percentage of presence of a fatty acid in each sample well known bacteria. In order to classify bacteria according to 10 different genera we used 113 indexes of different fatty acids; so our network is composed by 113 input neurons 150 hidden neurons and 10 output neurons [9].

In this case the application of the sensitivity check aims to determine the real importance of such a high number of fatty acids for the correct classification.

Using this method we have found that the 46% of the input fatty acids are almost ininfluent for the classification. In fact, after pruning the corrections of these uninfluent input neurons we can classify correctly the 90,6% of an ensamble of bacteria not considered in the learning set; the percentage of correctly predicted with the original connections is only of 91,2%. So it is clear that all these input neurons whose average sensitivity are lower than $10^{-4}$ are quite unuseful.

Moreover, this sensitivity analisys can evidentiate the importance of particular fatty acids for specified genera, an information of great biological value.

## References

[1] Karnin, E. D., 1990, A simple procedure for pruning back-propagation tranied neural networks, *IEEE Transactions on Neural Networks* 2: pp. 239-242.

[2] Tollenaere T., 1990, "SuperSAB: Fast Adaptive Back Propagation with Good Scaling Properties", *Neural Networks* 3, 561 - 573.

[3] Giacomini, M., Parisini, T., and Ruggiero, C., 1992, Secondary structure of proteins from NMR data by neural nets, *Artificial Neural Networks* 2 (I. Aleksander and J. Taylor eds) Elsevier.

[4] Sternberg, M. J., and Thrton, M. J., 1978, Prediction of protein structure from amino acid sequence, *Nature* 271, pp. 15-20.

[5] Kabsh, W., and Sander, C., 1983, How good are prediction of protein secondary structure?, *FBES Lett.* 155, pp. 179-182.

[6] Quian, N, and Sejnowski, T., 1988, Predicting the secondary structure of globular proteins using neural network models, *J. Mol. Biol.* **202**, pp. 865-884.

[7] Mc Gregor, M. J., Flores T. P. and Sternberg, M. J., 1989, Prediction of b-turn in proteins usign neural networks, *Prot. Eng.* **2**; pp. 521-526.

[8] Ruggiero, C., Sacile, R., and Rauch, G. 1993, Peptides secondary structure prediction with neural networks: a criterion fro building appropriate learning sets, *IEEE Transaction on Biomedical Engineering* **40**, pp. 1114-1121.

[9] Giacomini, M., Ruggiero, C., Calegari, L., Bertone, S., Casareto, L., and Reina, S., 1993, Marine bacteria classification by neural nets, *MIE93*, pp. 453-455.

# REDUCTION OF PAIN IN TRANSCRANIAL ELECTRICAL STIMULATION

Eskola H, Suihko V(*), Zentner J(**), Häkkinen V, Malmivuo J(*)

Department of Clinical Neurophysiology, Tampere University Hospital, P.O Box
2000, FIN-33521 Tampere, Finland

(*) Ragnar Granit Institute, Tampere University of Technology, P.O.Box 692, FIN-
33101 Tampere, Finland

(**) Department of Neurosurgery, University of Bonn, 5300 Bonn 1, Germany

*Abstract.* As the magnetic transcranial stimulator was introduced by Barker in 1985, the method superseded the painful electrical transcranial stimulation in examining awake subjects. However, there are also several advantages in electrical stimulation compared to the magnetic one. The stimulation field can be better localized and the stimulation current can be controlled. The stimulator itself is smaller, lighter and cheaper to produce. We have studied various mathematical, technical and experimental methods for decreasing the pain associated with electrical stimulation

## 1. INTRODUCTION

The examination of central motor nervous system proceeded a long step after the development of electrical transcranial stimulation by Merton et al. in 1980 /1/. However, the painfulness or discomfort of the stimulation was prominent. This drawback was removed by Barker in 1985 /2/, who transmitted the stimulation current inside the skull inductively by using a circular coil. This electromagnetic method was named magnetic stimulation.

A magnetic stimulator produces a relatively short, smooth-shaped pulse into a wide area of cortex. The recent technical development has enabled the focusing of the field by changing the size or geometry of the coil (for instance "small", "butterfly" or "figure-of-eight", "cap" and "flat" coils) /3/, /4/, /5/. Also magnetic stimulators producing repetitive stimulation have been constructed. However, the modification of the stimulation field or the stimulation pulse is much easier when electrical stimulation is applied. Further, electrical stimulators are small, light and relatively inexpensive.

The painfulness of electrical stimulation is based on the fact that the current from a

surface electrode must pass through the scalp and the resistive skull bone to reach the cortex. Although some muscle contraction can be obtained during the stimulation, the discomfort seems to be generated in the pain receptors of the scalp /6/. On the other hand, short pulses probably produce minimal discomfort /7/, /8/. From these starting-points we have developed various methods in order to decrease the discomfort associated with the electrical stimulation method.

## 2. METHODS

### Mathematical methods

To optimize the current distribution inside and outside the skull we chose an analytical mathematical approach, based on the three-concentric-spheres model /9/. Resistivities of 222 Ωcm, 17,760 Ωcm and 222 Ωcm were given for brain, skull and scalp tissues, respectively.

### Technical methods

We have constructed an electrical stimulators which produce short constant voltage pulses. For testing purposes the pulse from a commercial stimulator (Device A, Digitimer) was shortened to 50 µs by using a customized box (Device B) developed by Zentner /8/, or a special stimulator (Device C) was used, which produces rectangular pulse shape.

### Medical methods

We have used topical anaesthesia (EMLA cream) and blockade of occipital nerves as potential methods for pain reduction.

### Experiments

The mathematical, technical and medical methods were tested by series of studies. In each study, 15 - 54 normal subjects were stimulated several times and after each stimulation the subjects assessed the pain by using visual analog scale (VAS). The effects of electrode type, electrode separation, anaesthesia, and type of stimulator (Device A, device B or device C with pulse length of 50 µs) were studied.

## 3. RESULTS

According to the mathematical model the extracranial current density is high between adjacent electrodes and decreases if inter-electrode distance is increased (Fig.1). Correspondingly, the intracranial portion of current increases when increasing the distance, as well as the amount of activated cortical neurons. However, although the extracranial current decreases with increasing inter-electrode distance, the pain was increasing according to our experimental results. From Fig.1 it can be evaluated that if the intracortical current density is high enough to activate cortical neurons, most extracranial pain receptors situating between the electrodes are activated in both cases (distances 3cm and 9cm). Thus the amount of activated pain receptors is probably higher when the distance is 9 cm.
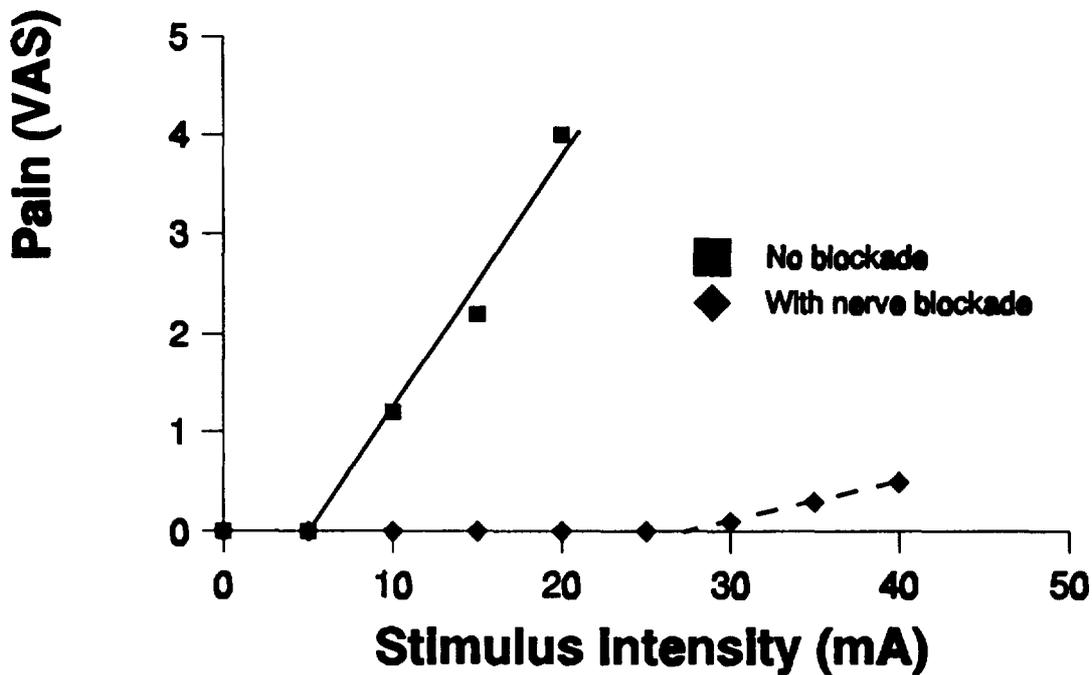
**Fig.2.** Effect of stimulus intensity and nerve blockade on pain sensation (VAS) in occipital scalp area.

## 4. DISCUSSION AND CONCLUSIONS

This study shows that the pain sensation in electrical transcranial stimulation can be altered by several methods. One can choose the electrode type, adjust the distance of electrodes or use of several electrodes. The blockade of sensory nerves having branches over motor cortex would practically remove the pain, but this procedure is impractical. A short stimulus pulse (50 µs) is probably less painful than a longer one.

We have applied a spherical head model for calculations, which does not include anatomical details of the head, while numerical models take them into account. However, it has been shown that the overall distribution is similar in both models /10/. According to recent experimental studies the anisotropy of white matter seems to be very important in modelling of stimulation /11/. This far, no such models have been presented.

An appropriate application of technical and experimental knowledge combined with results from theoretical calculations can be expected to noticeably reduce painfulness of transcranial electrical stimulation, thus making this method acceptable for the awake patient.

159
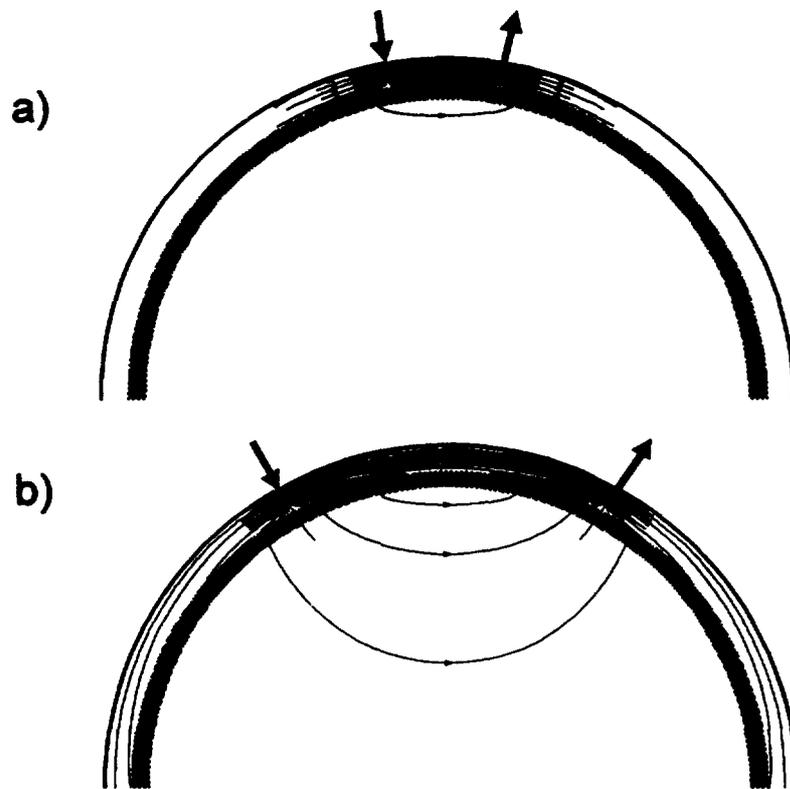
Fig. 1. The effect of inter-electrode distance ( a) 3cm, b) 9cm) on calculated intracranial and extracranial current density.

The mathematical calculations showed that motor cortex stimulation produces minimal discomfort when inter-electrode distance is more than 5 cm. Similar results were also obtained on test subjects. Both methods also showed that disk electrodes produce slightly less pain than needle electrodes.

According to our anaesthesia experiments the topical anaesthetics did not affect the pain sensation at all. As can be obtained from Fig.2, the blockade of occipital nerves practically removed all discomfort in occipital areas.

The preliminary results from the comparison of stimulators showed that similar voltage setting affected statistically similar level of painfulness (according to VAS gradings). However, with a low voltage setting of 300 V distinct motor responses were only obtained with devices B and C.

160

# REFERENCES

1. Merton PA., Morton HB., Stimulation of the cerebral cortex in the intact human subject. Nature 1980; 285: 227.

2. Barker AT., Jalinous, R., Freeston IL., Non-invasive magnetic stimulation of the human motor cortex. Lancet 1985; 1:1106-1107.

3. Ueno S, Tashiro T, Harada K, Localized stimulation of neural tissues in the brain by means of paired configuration of time-varying magnetic fields. J. Appl.Phys 1988; 64, 5862-5864.

4. Cohen LG, Roth BJ, Nilsson J, Dang N, Panizza M, Sandinelli S, Friauf W, Hallett M, Effects of coil design on delivery of focal magnetic stimulation. Technical consideration. Electroenceph Clin Neurophysiol 1990; 75:350-357.

5. Kraus KH, Cugino LD, Levy WJ, Cadwell J, Roth BJ, Use of cap-shaped coil for transcranial magnetic stimulation of the motor cortex. J. Clin.Neurophysiol. 1993; 3: 353-362.

6. Häkkinen V, Eskola H, Yli-Hankala A, Nurmikko T, Kolehmainen S, Which structures are sensitive to painful transcranial electric stimulation. Accepted to: Electromyogr Clin Neurophysiol.

7. Geddes LA, Optimal stimulus duration for extracranial cortical stimulation. Neurosurgery 1987; 20: 94-99.

8. Zentner J, Modified impulse diminishes discomfort of transcranial electrical stimulation of the motor cortex. Electromyogr Clin Neurophysiol 1989; 29: 93-97. '

9. Driscoll D, An investigation of a theoretical model of the human head with application to current flow calculations and EEG interpretation. Ph.D.thesis, The University of Vermont, 1970; 179p.

10. Hämäläinen M, Sarvas J, Realistic conductivity geometry model of the human head for interpretation of neuromagnetic data. IEEE Trans Biomed Eng 1989;36: 165-171.

11. Amassian VE, Eberle L, Maccabee PJ, Cracco RQ, Modeling magnetic coil excitation of human cerebral cortex with a peripheral nerve immersed in a brain-shaped volume conductor: the significance of fibre bending in excitation. Electromyogr Clin Neurophysiol 1992; 85: 291-301.

# Artificial Neuronal Network in Medical Microbiology.

J.Schindler, P.Parýzek, P.Urbášková

3. Medical Faculty, Charles University Prague, and National Institute of Public Health, Prague, Šrobarova 48, Prague 10, CZ 100'42, Czech Republic

## Abstract

Summary of experience with artificial neuronal network in identification of bacteria and in clinical microbiology is presented. Back propagation type of the neuronal network was used with 47 input neurons, 70 neurons in the hidden layer and 97 output neurons. 78 taxa contained more than 10 strains each . If the training set was selected out from strains with the reference identification confirmed by probabilistic identificationt the working file contained 4634 strains belonging to 78 taxa and 98,6 % strains were identified in agreement with the reference at the genus level and 97.4% at the species level. The Hopfield network applied to antibiotic resistance prediction failed due to overlapping antibiotic resistance profiles in the strains isolated from a hospital.

## Introduction

Neuronal network appears to be a plausible method for the identification of bacteria. Bungay and Mary Bungay (1991) reported on the software NeuroShell and showed the way of its application in a network using eight input neurons and seven output neurons. At the same time Rataj and Schindler (1991) showed the applicability of the neuronal network in the identification of bacteria on a set of 29 most probable character patterns of strains. They used a vector of 17 phenotypic characters coding for nine selected taxa. Both these papers used a desktop computer both for training and for testing. Accordingly only small sets of strains could be used and those papers can be considered as pilot trials.

The present paper reports on experience with the identification of strains characterized and kept in CDC database. The aim of it is to find out if neuronal identification of bacteria is applicable and to what extent.

## Methods

Briefly, the basic element of the network is described by the equation:

$$y = f (\sum_{i}^{N} i=1 \ ( Wi \ . Xi )+ \textit{theta} ),$$

wherein $y$ is the neuron output, $W_i$ weight of input $i$, $X_i$ neuron inputs, *theta* is the threshold level of the neuron, f(t) is a non linear transfer function defined:

$$f (t) = 1/ (1+ \exp (- t)).$$

Strains. 6945 strains of 111 taxa were selected from the database of CDC, Atlanta compiled by J.Farmer III were used for training and testing. The database contains data on strains which, for any reason, are interesting or remarkable. Thus the collection contains also atypical strains.

Strains with uncertain identification given or with more than 5 missing tests were eliminated. 78 taxa contained more than 10 strains each. Back propagation type of the neuronal network was used with 47 input neurons, 70 neurons in the hidden layer and 97 output neurons.Test results recorded in the database were expressed 0 for negative; numbers 1 - 9 expressing the day positive results appeared; blank for not done. The highest activity value of an output neuron was used as a criterion of identification.

Neuronal network operation.The result of identification is expressed by the neuron output activity $a$, $0<a <=1$. After activation of the network different values for each taxon are obtained. Two different criteria were used for interpretation of identification: 1. the highest value, 2. the difference between the first and the second highest value. Eventually the simple highest activity value was used. The taxon with the highest one was taken as a result of identification.

Results

First, the training set randomly selected from the working file contained 10% strains of each taxon, but at least 10 strains. All strains from the working file including those from the training set were used in the testing set.

Results were evaluated by comparison with CDC reference identification. Only species with more than 9 strains in the database were evaluated. Frequency distribution of activities shows that the majority of results lies in the range 0.98 - 1.00. On the species level all strains of 30 species out of 78 were identified in complete agreement with the reference identification. In 32 species 90 to 99% strains were in agreement . (Figure)
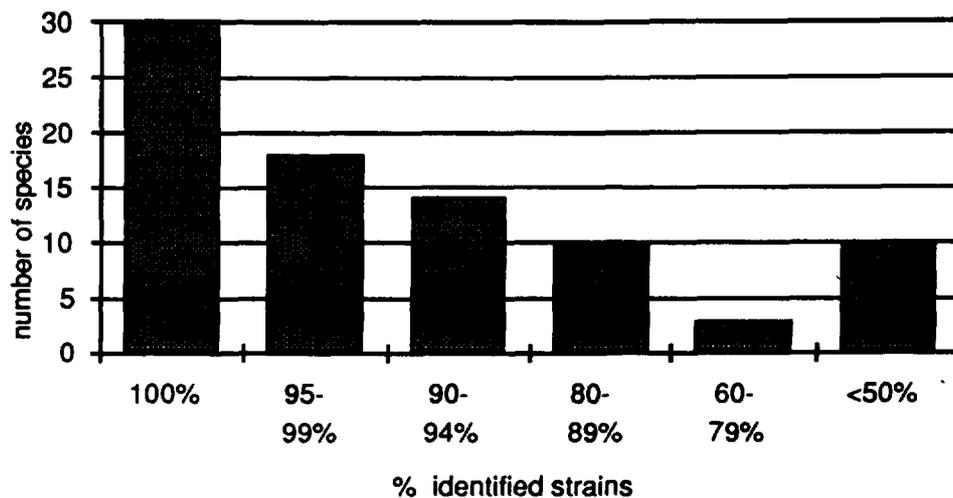


Figure . Percentage of strains in particular taxa identified in agreement with original reference identification. Total number of taxa is 78.

At the species level 92.3 %, at the genus level 95.7% strains were identified in agreement with the reference. Due to the phenotypic similarity some species may be grouped into so called *convenience groups* . Thus e.g. *E.coli* and shigellae belong to one convenience group, salmonella - arizona, kluyvera - butiauxella to others. Using these groups as the criterion, identification rate rose to 97.7% strains . (Table 1).

| Level of evaluation | Identified in agreement<br>% strains |
| --- | --- |
| Species | 92.3 |
| Genus | 95.7 |
| Convenience group | 97.7 |

*Table 1 Result of identification evaluated by agreement on species, genus and convenience group level. Traing set contained 10% strain of the whole set.*

Misidentifications were due to phenetic similarity of the taxa tested. Strains which were incorrectly identified as *E.coli.* were shigellae phenotypically very heterogeneous .

To get more information on strains which were not identified in agreement with the reference, probabilistic identification was computed using Farmer' s identification matrix (1985). In some instances.probabilistic identification was in disagreement with the reference identification. The results showed that some strains in the database cannot be identified unambiguously.

Then a training set was used selected out from the strains with the reference identification confirmed by probabilistic identification. Thus a working file was obtained containing 4634 strains belonging to 78 taxa. Strains up to 10% of the whole set were selected at random and included into the training set. The results are given in Table 2.

| Agreement on | |
| --- | --- |
| genus level | species level |
| 98.6% | 97.4% |

Table 2. *Result of identification of strains of gramnegative ferenting rods.. The network was trained by 10% strains of the whole set. The identity of strains in the training set was confirmed by probabilistic identification.*

It was tempting to apply artificial neuronal network to solve some problems encountered in clinical microbiology especially in connection with nosocomial infections and resistance of bacteria to antibiotics. For this purpose a Hopfield network was trained with antibiotic resistance patterns. Each strain was represented by a vector of 25 resistance markers, identification- and ward data. The attempt to use a pattern for determining the species or ward of origin was not successful. However, the network can be used to predict resistance to one or two antibiotics following input of resistance pattern reduced to certain resistance markers antibiotics linked to a group determining the most frequent multiresistance pattern.

Discussion

It is a well established fact that identification reflects the classification. The effectivity of the neuronal network in identification depends on learning. The main problem is the the selection of patterns in the training set for the particular taxa to be identified. They need not be rather typical, however they should be well classified. In contrast to probabilistic identification, neuronal network identification depends more on individual strains.

Misidentifications were mostly due to phenetic similarity of taxa tested and to atypical characters of strains which have a very low modal likelihood fraction in probabilistic identification.

The artificial neuronal network is applicable and effective in bacterial diagniostics. The method will be further developed. The main problem appears to be selection of the training set with respect to missing characters. It is assumed that coding by two-state code (+, -, ND) only, instead of 1 to 9 code for positivity could bring better performance of the network. The training set shouldn't contain strains with missing characters. On the one hand, this might lead to incorrect emphasizing of typical and "safe" patterns and to less discrimination between late and early fermenters and acidifiers. On the other hand, errors due to too fuzzy input patterns can be eliminated.

At present the type of neuronal network used can predict resistance to one or two antibiotics. However, the input data are limited to those markers which form a cluster of multiresistance. The neuronal network failed to predict the sensitivity/resistance to several antibiotics to a useful extent, or to detect the species or the ward of origin on the basis of a known pattern. This is explained by the inherent character of resistance patterns of strains in the hospital. The patterns of multiresistant strains are overlapping to a high extent.

References

Bungay, H., and M.L.Bungay. 1991. Identifying Microorganisms with a Neuronal Network. Binary,3: 51-52.

Rataj, T., and J.Schindler. 1991. Identification of Bacteria by Multilayer Neuronal Network. Binary, 3:159-164.

Farmer, J.J., III, B.R.Davis, F.W.Hickman-Brenner, A.McWhorter, G.P.Huntley-Carter,

M.A.Asbury, C.Riddle, H.G.Wathen-Grady, C.Elias, G.R.Fanning, A.G.Steigerwalt, C.M.O'Hara, G.K.Morris, P.B.Smith, and D.J.Brenner.1985. Biochemical identification of new species and biogroups of Enterobacteriaceae isolated from clinical specimens. J. Clin. Microbiol. 21: 46-76.

# PHYSICAL AND MATHEMATICAL GROUNDS OF IMAGE RESOLUTION IMPROVEMENT AT ULTRASOUND DIAGNOSTICS

T.V.Yakovleva, Z.M.Benenson

Scientific Council on Cybernetics of the Russian Academy of Science, Moscow , Russia

The present paper deals with the theoretical investigation and computer simulation of the process of ultrasound pulse propagation through the inhomogeneous biological medium with strongly dispersive absorption coefficient. The analytical solution of the wave equation for ultrasound irradiation has been obtained by means of Green function. The properties of ultrasound pulse signal have appeared to be dependent upon some average carrier frequency which changes with the depth of the medium. The formulas have been derived for the calculation of current carrier frequency of ultrasound pulse in dependence upon the absorptive properties of a medium and the depth of observation. It has been analytically shown that time delays of the pressure signals at receiving aperture change from one point of phased array to another, what distorts the image of investigated object at ultrasound diagnostics. The possible mathematical techniques of the correction of these distortions have been proved theoretically and by computer simulation. These techniques are based upon the calculation of mutual coherence function at nearby transducers. The accurate theoretical consideration of the algorithms of the phase distortions correction has proved the possibility of image quality improvement in the process of ultrasound visualisation.

## 1. Introduction

As it is well known, in ultrasound imaging devices applicable for medicine purposes the pulse signal is used. The range of frequency band of this signal is commensurable with the carried frequency. Besides, the inhomogeneous medium in which the signal propagates, is strongly dispersive by its absorption coefficient. These peculiarities cause the necessity of the special theoretical consideration of the problem of phase aberration correction for pulse ultrasound systems. In present paper the theoretical and computer investigations of the process of ultrasound propagation in inhomogeneous biological medium have been carried out with taking into account of strong dispersion of the medium by its absorption coefficient.

The solution of wave equation for pressure function is analytically obtained for inhomogeneous and essentially absorptive medium for pulse signal when the frequency band is commensurable with the carried frequency. The formulas describing the dependence of carrier frequency upon the medium's depth have been determined. The analytical formula for pressure signal has been obtained. This has made it possible to find the expressions for the function of mutual coherence. On the basis of obtained theoretical results the problem of adaptation of ultrasonic irradiation in the inhomogeneous biological medium is considered and two algorithms are constructed for the calculation the correction of phase aberrations: the so-called correlation algorithm, based upon the calculation of the sample function of mutual coherence for nearby subapertures of phase array and the phase algorithm, which consists in the calculation of the phase difference of pressure signals on nearby subaperture.

The computer simulation of these algorithms has been carried out. The accuracy of the correction of phase aberrations has been investigated. The results of mathematical computer simulation of the process of ultrasound propagation in the inhomogeneous medium have proved the possibility of effective adaptive processing of acoustic signals and ultrasound device resolution improvement due to adaptive focusing.

## 2.The basic dependencies

Let us consider the ultrasound propagation in the strongly absorptive inhomogeneous medium. The wave equation for $P(\bar{R}, \omega)$ - the Fourier transform of the spatially-temporal function of pressure $P(\bar{R}, t)$ is as follows:

$$\nabla^2 P + \frac{\omega^2}{c^2} P + 2f(\omega, \bar{R}) i P = \frac{\omega^2_i}{c^2} \hat{\beta}(\bar{R}) P + \nabla(\hat{\rho}(\bar{R})) \nabla P \tag{1}$$

where $\omega$ is the angular frequency, $c = (\rho_0 \beta_0)^{-1/2}$ - the sound velocity,

$\hat{\beta}(\bar{R}) = \beta(\bar{R}) / \beta_0$ - the respective variation of elasticity,

$\hat{\rho}(\bar{R}) = \rho(\bar{R}) / \rho_0$ - the respective variation of density,

$f(\omega, \bar{R})$ - frequency-dependent function of sound absorption in a medium.

Using Born approximation we can represent the equation (1) in the form:

$$\nabla^2 \hat{P} + \frac{\omega^2}{c^2} (1 - \gamma_c(\bar{R})) \hat{P} = \frac{\omega^2}{c^2} \gamma_{in} P_I \tag{1a}$$

$$\hat{P}(\hat{\bar{R}}, \omega) = P(\hat{\bar{R}}, \omega) \exp\left\{ c \int_0^z (f(\omega, z') / \omega) dz' \right\},$$

$P_I(\bar{R}, \omega)$ is the Fourier transform of the incident ultrasound wave, $P(\hat{\bar{R}}, \omega)$ is a pressure function in the plane of aperture,

$\gamma_c(\bar{R})$ - smooth or large scale function causing beam defocusing,

$\gamma_{in}(\bar{R})$ - small-scale function causing the scattering,

$\hat{\bar{R}}$ is the coordinate in the plane of aperture (Fig.1).

Having solved the equation (1a) and performing the inverse Fourier-transform of $P(\bar{r}, z, \omega)$ we obtain the analytical solution of wave equation for pressure $P(\bar{r}, z, t)$ in the following form in space-time coordinates:



Fig.1

$$P(\bar{r}, z, \hat{t}) = \frac{\omega^2_g}{c^2 z^2} A_0 \exp\left\{ -i\omega_g (\hat{t} - \tau(\bar{r}, z)) \right\} \cdot \exp[-\alpha(z) \cdot z] \cdot$$

$$\int_{V_1} L_{tr}(\bar{r}', z) \cdot \exp\left[ -i\omega_g \frac{(\bar{r} \cdot \bar{r}')}{cz} \right] \cdot \chi(\bar{r}', \bar{r}, \hat{t}) \cdot d^2 \bar{r}' \tag{2}$$
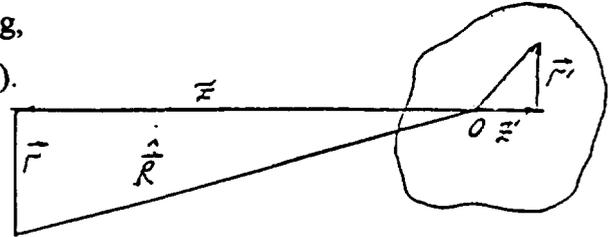
where $L_{tr}(\bar{r}', z)$ is the radiation pattern of transmitting aperture, $\bar{r}, z$ - the coordinates in the aperture plane, $\bar{r}', z'$ - the coordinates in the scattering volume, $\hat{t} = t - 2z/c$, $\tau(\bar{r}, z)$ - time delay, determined by function $\gamma_c(\bar{R})$, $\alpha(z) = 2 \cdot c \cdot f(\omega) / \omega$ - the absorption coefficient of the medium, $\omega_g(z)$ is the carrier frequency which is defined by the absorption function $f(\omega, z)$ and depends on the depth of observation. We suppose transmitted beam to be Gaussian

171

$$A(\omega) = A_o \cdot \exp\left[ -\frac{(\omega - \omega_0)^2}{8} \cdot \tau_I^2 \right],$$

which is common in practise. Here $\tau_I$ is the pulse duration.

The scattering function $\chi(\vec{r}', \vec{r}, \hat{t})$ is described by the following formula:

$$\chi(\vec{r}', \vec{r}, \hat{t}) = \int\limits_{V_{II}} \exp\left[ -2(\hat{t} - 2\frac{z'}{c} + \Delta(\vec{r}, \vec{r}'))^2 / \tau_I^2 \right] \exp\left[ 2i\frac{\omega_g}{z}(z' + \frac{|r'|^2}{2z}) \right] \gamma_{in}(\vec{r}', z') dz'$$

where $V_\perp$ - the transverse area of the scattering volume,

$V_{II}$ - the longitudinal interval of the scattering volume.

The important conclusion from formula (2) is as follows: the properties of the received scattered signal depend upon the carrier frequency $\omega_g$. So instead of the whole spectrum of frequencies we can deal now with one average frequency $\omega_g$, which can be calculated as dependent upon the absorptive properties of a medium and the depth of observation $z$.

From formula (2) one can also see that in different points $\vec{r}$ of the aperture the signal comes with delay $\tau(\vec{r}, z)$ which cause the defocusing phase shifts $\omega_g(z) \cdot \tau(\vec{r}, z)$. These phase aberrations distort the image of investigated object. The purpose of the adaptive focusing is the correction of such distortions by means of calculation and compensation of phase aberrations. .

## 3. The function of mutual coherence

The construction of the algorithms of adaptive focusing in ultrasound device is based on the calculations of sample values of the mutual coherence function of the signals received by different transducers:

$$f_{m,m-n} = P(\vec{r}_m, z, 0) \cdot P^*(\vec{r}_{m-n}, z, 0) \tag{3}$$

and on averaging it by different depths of observation.

Supposing that the ultrasound is scattered by independent point reflectors we obtain the following formula for the sample function of mutual coherence in the points $\vec{r}_m$ and $\vec{r}_{m-n}$ of aperture for the same moment of time:

$$f_{m,m-n} = \exp\left[ i\omega_g \cdot (\tau_m - \tau_{m-n}) \right] \cdot B_{m,m-n}$$

$$B_{m,m-n}(z) = B_{m,m-n}^{(c)}(z) + B_{m,m-n}^{(n)}(z) \tag{4}$$

where functions $B_{m,m-n}^{(c)}$ and $B_{m,m-n}^{(n)}$ describe the contribution to the mutual coherence function of the coherent and noncoherent (or noise) parts correspondingly. The mathematical analysis of these functions shows that the average value of $B_{m,m-n}^{(c)}$ is positive (at the absence of bright points in a medium). As for the value of $B_{m,m-n}^{(n)}$, it is complex with zero average values of its real and imaginary parts. As it follows from (4) the phase difference $\omega_g \cdot (\tau_m - \tau_{m-n})$ can be determined if the value $B_{m,m-n}$ is real and positive. The error in definition of this phase shift will depend upon the value of noise part $B_{m,m-n}^{(n)}$ and upon the presence of bright point.

## 4. The algorithms of the calculation of the defocusing delays

Let us consider now two algorithms of calculation of the defocusing delays $\tau(\vec{r}_m) = \tau_m$ for the case of linear phased array. As the characteristic radius of correlation of random functions of time delays $\tau$ is significantly larger than the distance between the nearby receivers

of an array we will perform the adaptive processing of ultrasound signals for so-called subapertures, each subaperture containing some number $n$ of receivers. The size of one subaperture is approximately equal to the correlation radius of random phase shifts.

## 4.1 The correlation algorithm

This algorithm is based on the calculation of the mutual coherence function for two nearby subapertures at the same moment of time. Taking into account that the time delays $\tau_m$ change slightly with changing of depth it is reasonable to average the values $f_{m,m-1}(z_j,0)$, calculated for different depths $z_j$ in order to decrease the influence of random changes of incoherent part of mutual correlation function:

$$\tilde{f}_{m,m-1} = \frac{1}{h+1} \sum_{j=L-h/2}^{L=h/2} f_{m,m-1}(z_j)$$

where $L$ is the index of the mean value of depth $z_L$, $h$ is the reasonable number of steps $\delta z$. We determine the value in question $\omega_g \tau_m$ by means of calculating the main value of the argument of averaged mutual correlation function:

$$Arg_{m,m-1} = arctg \left[ \frac{Im(\tilde{f}_{m,m-1})}{Re(\tilde{f}_{m,m-1})} \right]$$

and then implementing some mathematical processing of it in order to escape the errors which could be caused by the presence of bright point.

## 4.2 The phase algorithm

The so-called phase algorithm consists in the calculation of the phase differences of the analytical signals $P(\vec{r}_m, z_j, 0)$ at nearby subapertures. The formulas for the signals received by nearby subapertures can be presented as follows:

$$P(\vec{r}_m, z_j, 0) = \exp(i\omega_g \tau_m) \cdot D_m^{(j)} \cdot \exp(i\varphi_m^{(j)})$$

$$P(\vec{r}_{m-1}, z_j, 0) = \exp(i\omega_g \tau_{m-1}) \cdot D_m^{(j)} \cdot \exp(i\varphi_m^{(j)}) \cdot (1 + O(|\vec{r}_m - \vec{r}_{m-1}| / d)$$

where $d$ is the size of the aperture. As it follows from these formulas, the phase difference of the signals being considered is approximately equal to the value in question, namely:

$$Arg_m^{(j)} - Arg_{m-1}^{(j)} \approx \omega_g (\tau_m - \tau_{m-1})$$

where $Arg_m^{(j)} = arctg \left[ \frac{Im(P(\vec{r}_m, z_j, 0))}{Re(P(\vec{r}_m, z_j, 0))} \right]$. The calculation of this value in the phase method is

followed by some mathematical transformations which are similar to ones used in correlation method in order to improve the accuracy of the definition of phase aberrations.

## 4.3 The results of computer simulation of ultrasound adaptive focusing

The computer simulation of the process of ultrasonic irradiation propagation and its adaptive focusing in the inhomogeneous biological medium has been carried out for various values of parameters of the irradiation and of the medium. Fig.2 illustrated the possibility of the calculation and compensation of the phase distortions at the adaptive processing of the ultrasound signals. The two curves correspond to the initial and calculated values of phase delays $\omega_g \tau_m$ on subapertures. Fig.2a and 2b illustrate the accuracy of the calculation of phase

aberrations by the phase and by the correlation algorithms correspondingly. As one can see, the initial and mathematically calculated values of phase distortions $\omega_g \tau_m$ are very close to each other. This means that the adaptation algorithms being considered, give the opportunity to correct effectively the phase aberrations of the ultrasonic irradiation at the diagnostics process.
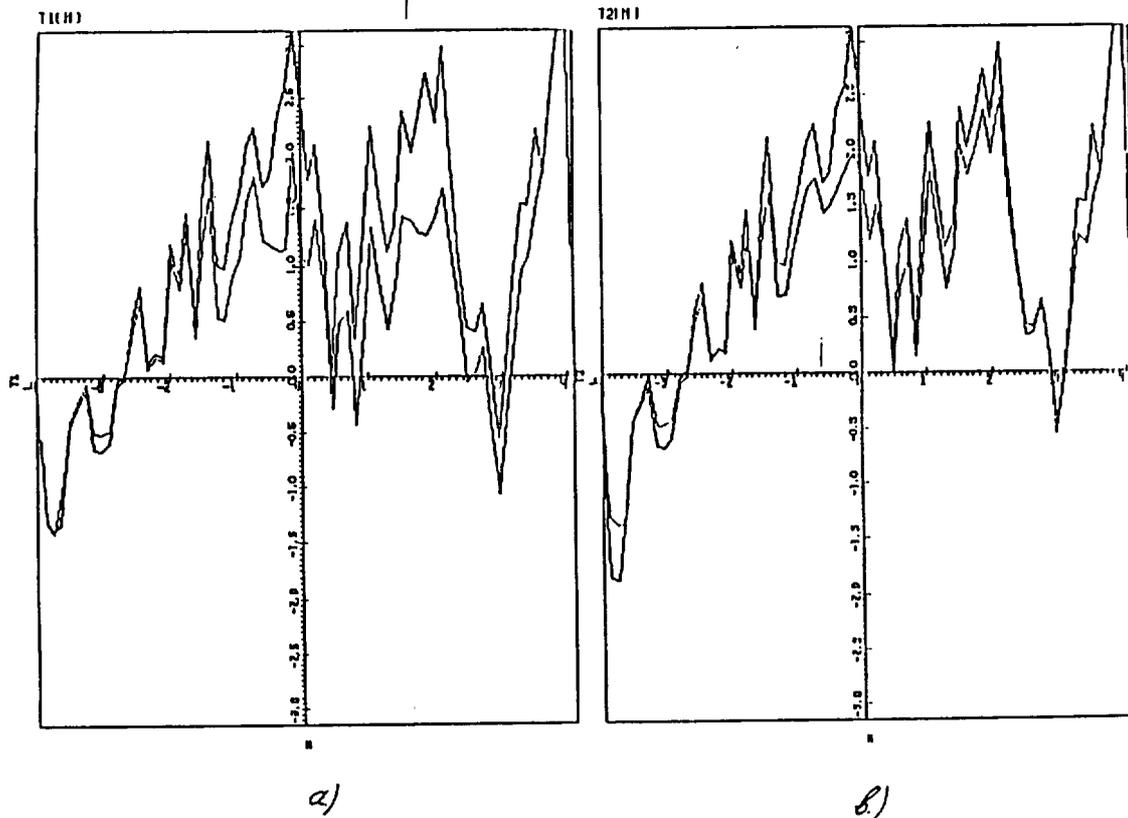


Fig.2

It has been also shown that the adaptive processing of ultrasound irradiation by means of the phase distortions correction leads to the significant optimisation, or narrowing of the receiving radiation pattern what makes it possible to reconstruct the medium's structure more precisely.

## 5. Conclusion

The principle results of the paper are the following:
-the wave equation is analytically solved for broad band ultrasound pulse in the inhomogeneous biological medium with the dispersive absorption coefficient.
- the formulas have been found for the current frequency and analytical expression for the mutual coherence function of signals in different points of receiving aperture.
- the algorithms of the calculation of ultrasonic signals time delays (or phase aberrations) due to inhomogeneity of the medium have been constructed.
-the computer simulation of ultrasound adaptive focusing in the biological medium has proved the possibility of the calculation and effective correction of the phase aberrations.
- both the theoretical analysis and the computer simulation of the correlation and the phase methods of adaptation have shown that the accuracy of determination of defocusing delays of signals is approximately the same for both these algorithms. But the phase method is more simple in its apparatus realisation and so is more preferable to be used in practice for adaptive processing of signals in ultrasound imaging devices.

# Simulated Annealing Image Reconstruction for Positron Emission Tomography

E. Sundermann *, I. Lemahieu †, P. Desmedt ‡
Department of Electronics and Information Systems,
University of Ghent, St. Pietersnieuwstraat 41, B-9000 Ghent,Belgium
email: erik@petdec.rug.ac.be

### Abstract

In Positron Emission Tomography (PET) images have to be reconstructed from noisy projection data. The noise on the PET data can be modeled by a Poisson distribution.

In this paper, we present the results of using the simulated annealing technique to reconstruct PET images. Various parameter settings of the simulated annealing algorithm are discussed and optimized. The reconstructed images are of good quality and high contrast, in comparison to other reconstruction techniques.

## 1  Introduction

Positron Emission Tomography (PET) is a tomographic method to display metabolic activity in a slice through a patient's body. The particular construction of the PET scanner and the use of a radioactive tracer entail the modeling of the data by a Poisson distribution.

The reconstruction method most commonly used today in PET is the Filtered Backprojection (FB) algorithm [1]. This reconstruction technique is based on a Fourier Transform algorithm and is extremely fast. However, since FB is a deterministic algorithm and thus does not account for statistical fluctuations in the measurements, the obtained reconstructed images can suffer from very annoying stripe-like artifacts.

To address the problem of noise, the study of statistical (iterative) reconstruction techniques has received much attention in the past few years. One of the most popular methods is the Maximum Likelihood Expectation Maximization (ML-EM) algorithm [2, 3], which searches the image that maximizes the likelihood of the data. However, when the algorithm is iterated too long, the reconstructed image starts to degrade [4].

In this paper we present the results of introducing the simulated annealing technique in the statistical reconstruction algorithms for PET. Maximization algorithms (e. g. conjugate gradient, EM) achieve convergence by monotonically increasing some metric (e. g. a likelihood) at every iteration. Simulated annealing is a Monte Carlo technique which allows occasional negative increments of the cost function so that it can avoid getting trapped in local minima. The technique has been proposed earlier by Kearfott et. al. [5] and Webb [6] for SPECT, but it has never been applied to PET before.

## 2  Methods

Simulated annealing was introduced by Metropolis et. al. [7] and is used to approximate the solution of very large combinatorial optimization problems (e. g. NP-hard problems). The technique originates

---

from the theory of statistical mechanics and is based upon the analogy between the annealing of solids and solving optimization problems.

Let us assume we are looking for the configuration that minimizes a certain cost function $E$. The algorithm can then be formulated as follows [8]. Starting off at a given configuration, a sequence of iterations is generated. Each iteration consists of the random selection of a configuration from the neighbourhood of the current configuration and the calculation of the corresponding change in cost function $\Delta E$. This neighbourhood is defined by the choice of a *generation mechanism*, i.e. a prescription to generate a transition from one configuration into another by a small perturbation. If the change in cost function is negative, the transition is unconditionally accepted; if the cost function increases the transition is accepted with a probability based upon the Boltzmann distribution

$$\mathbf{P}_{acc}(\Delta E) \sim \exp(-\frac{\Delta E}{kT})$$

where $k$ is a constant and the temperature $T$ is simply a control parameter. This temperature is gradually lowered throughout the algorithm, from a sufficiently high starting value (i.e. a temperature where almost every proposed transition, both positive and negative, is accepted) to a "freezing" temperature, where no further changes occur. In practice, the temperature is decreased in stages, and at each stage the temperature is kept constant until thermal quasi-equilibrium is reached. The whole of parameters determining the temperature decrement (initial temperature, stop criterion, temperature decrement between successive stages, number of transitions for each temperature value) is called the *cooling schedule*.

It follows that the implementation of simulated annealing presupposes the following four key "ingredients":

- the definition of configurations

- the choice of a cost-function

- a generation mechanism, i.e. the definition of a neighbourhood on the configuration space

- a cooling schedule

We will now discuss the particularities of the annealing algorithm in the case of reconstruction of PET images.

# 3 Results and discussion

## 3.1 Simulated annealing for PET

The simulated annealing algorithm for PET can be formulated as follows. We will iteratively reconstruct an image that fits best the measured data $\mathbf{P}^m$. To do so, we will calculate at each iteration step the pseudo data $\mathbf{P}^p$ that correspond to the present state of the reconstructed image. We assume that, by minimizing the difference between the measured data and the pseudo data, the reconstructed image will converge towards the sought-after original image. Therefore we will choose as cost function a function that expresses the difference between both data sets. At each iteration, the intensities of one or a few pixels are altered. We will then calculate the corresponding change in cost function, and decide upon this difference in cost function whether the proposed transition is accepted. As the cost function decreases throughout the algorithm, the reconstucted image "crystallizes" as can be seen in Fig. 1.

## 3.2 Configurations

We have performed studies on software generated phantom images. The pixel intensities are integer values ranging between 0 and 10000. To reduce reconstruction times, we have experimented with relatively small (40 x 40) images, since there is no reason why the obtained results should not be valid for larger dimensions. We have furthermore restricted our experiments to the reconstruction of noisefree data.
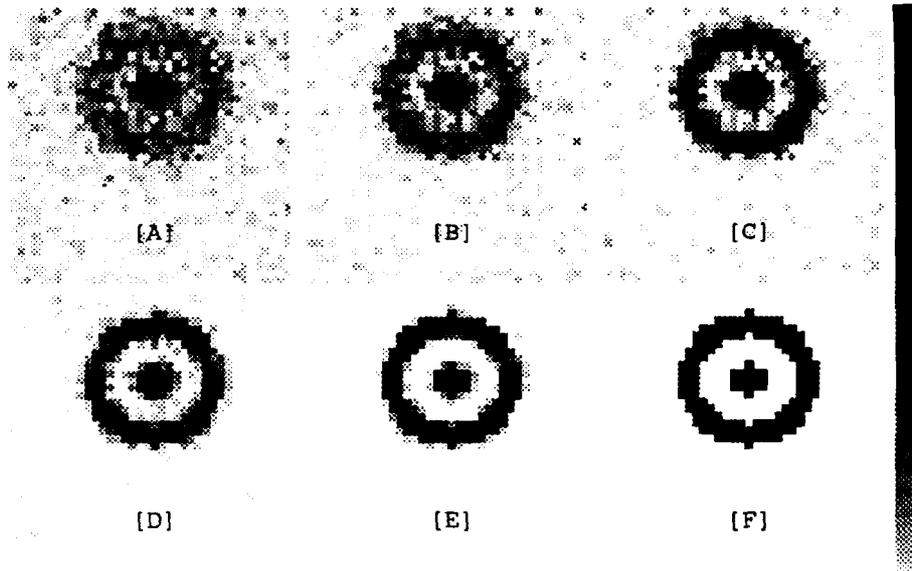
Figure 1: Reconstructed image after [A] 500.000 [B] 1 million [C] 1.5 million [D] 2 million [E] 3 million [F] 5 million iterations

## 3.3 Cost function

We first performed experiments using the classical cost function

$$E = \sum_{i,j} (P_{ij}^m - P_{ij}^p)^2 \qquad \forall i,j = 1,\ldots,40$$

which is the mean squared distance between $\mathbf{P}^m$ and $\mathbf{P}^p$. This approach is suggested by most authors, including [5] and [6], and produces good results for a number of different testphantoms. The reconstructed images show that simulated annealing, in comparison to other reconstruction algorithms, is very good at reconstructing edges between image-areas with a constant intensity, but has more difficulties in reconstructing these constant areas themselves. This problem can be addressed by adding an appropriate smoothness factor to the cost function [9].

We later derived a new cost function, based upon the maximum likelihood principle. Maximization of $P(\mathbf{I}|\mathbf{P}^m)$, with $\mathbf{I}$ the reconstructed image, implies maximization of $\prod_{i,j} P(P_{ij}^m|\mathbf{I})$. Taking into account the Poisson-character of the detection process, we find that

$$P(P_{ij}^m|\mathbf{I}) = e^{-P_{ij}^p} \frac{(P_{ij}^p)^{P_{ij}^m}}{P_{ij}^m!}$$

After taking the logarithm of the likelihood and omitting the constant term, we find for the cost function

$$E = \sum_{i,j} (P_{ij}^m \ln P_{ij}^p - P_{ij}^p) \qquad \forall i,j = 1,\ldots,40$$

This new cost function achieves reconstructed images of the same quality as the classical cost function, but clearly shows faster convergence in the early stages of the reconstruction. Since this new cost function is based upon the statistical character of the detection process, we expect it to outperform the classical one when applied to noisy data.

## 3.4 Generation mechanism

As explained above, the generation mechanism provides a means to move from one image to a neighbouring one. As this means changing the intensity of one or more pixels, we have to decide on both the number of pixels to be changed simultaneously and the amount of intensity to be changed (i.e. the *grainsize*).

177
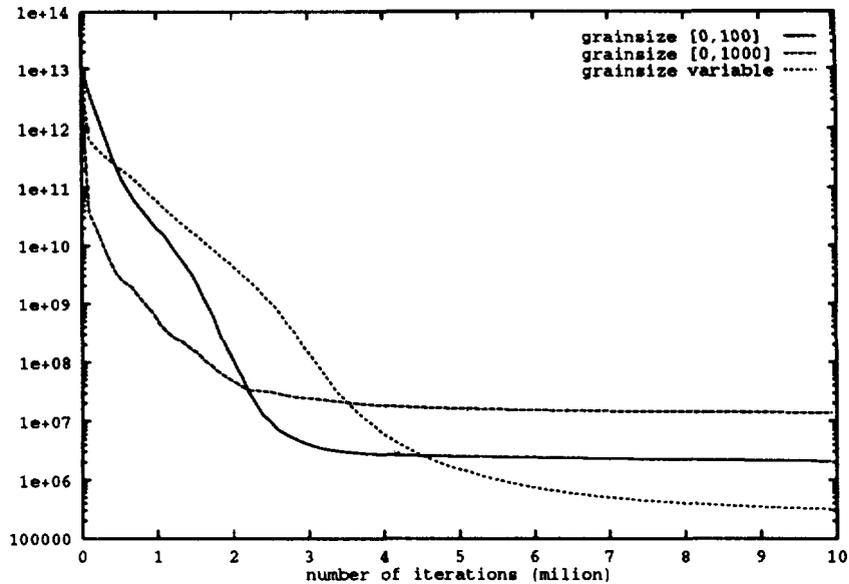
Figure 2: Evolution of the cost function for different grainsize-settings

**Pixel updating.** We considered two possible approaches to the first problem, based upon [10]. There is the *Grain Allocation Method* (GAM), which allocates a grain to one pixel, and there is the *Grain Transfer Method* (GTM), which transfers a grain from one pixel to another. Since GTM conserves the total image intensity (which is initially unknown), we constructed a third method which uses both GAM and GTM at the same time. Our experiments showed that GAM performs marginally better than the other two methods.

**Grainsize.** First, it should be clear that the algorithm must allow both positive and negative grainsizes, but must forbid negative image intensities. Most authors suggest choosing grainsizes from a uniform distribution in the interval $[-\beta, +\beta]$. As can be seen in Fig. 2, our experiments showed that better results can be obtained by not keeping the boundaries constant, but letting the algorithm itself decide upon the boundary values throughout the process. This can be achieved by keeping track of the mean accepted grainsize, and choosing the boundaries hereupon.

## 3.5 Cooling schedule

**Starting temperature.** The initial temperature should be chosen so that about 80 percent of all positive transitions (i.e. transitions which increase the cost function) are accepted [8]. This temperature can be determined by running the algorithm shortly for a few temperature values, and calculating the corresponding acceptance rate.

**Temperature decrement between successive stages.** The temperature is decreased by multiplication with a factor $\alpha$. There is actually a trade-off between temperature decrement between stages and the number of iterations per stage. For example when the temperature is kept constant long enough for the equilibrium to be reached, greater decrements are allowed. Best values for $\alpha$ are between 0,8 and 0,98.

**Number of iterations per stage.** Frequently used criteria are a constant number of iterations, or iterating until a constant number of moves is accepted. Experiments show that better results are achieved by considering the physical background of simulated annealing and the concept of *thermal quasi-equilibrium*. This means keeping the temperature constant until the cost function has reached a constant value (or is oscillating around this constant value).

**Stop criterion.** Since the emphasis of our research was on the quality of the reconstruced images, we have not devoted much attention to the development of a workable stop criterion. It is clear however that when the equilibrium values of the cost function for succesive stages are constant themselves, the iteration process can be stopped.

# 4   Conclusion

We have found the simulated annealing algorithm to reconstruct images of good quality in the case of noiseless reconstruction. As already mentioned, it is notable that the reconstructed images show a high contrast, in comparison to other reconstruction techniques. These results certainly justify the continuation of our research for the reconstruction of noisy images. The main problem with simulated annealing are the large reconstruction times. Since it has been proven that simulated annealing is extremely suited for parallellization [11], we will also try to reduce the reconstruction time by parallellizing the algorithm.

# References

[1] A.C. Kak and M. Slaney. *Principles of Computerized Tomographic Imaging.* IEEE Press, New York, 1988.

[2] L.A. Shepp and Y. Vardi. Maximum Likelihood Reconstruction for Emission Tomography. *IEEE Trans. Medical Imaging*, 1:113–122, 1982.

[3] P. Desmedt, K. Thielemans, I. Lemahieu, F. Vermeulen, D. Vogelaers, and F. Colardyn. Measured Attenuation Correction Using the Maximum Likelihood Algorithm. *Medical Progress Through Technology*, 17:199–204, 1991.

[4] D.L. Snyder, M.I. Miller, L.J. Thomas, and D.G. Politte. Noise and Edge Artifacts in Maximum Likelihood Reconstructions for Emission Tomography. *IEEE Trans. Medical Imaging*, 6:228–238, 1987.

[5] K.J. Kearfott and S.E. Hill. Simulated Annealing Image Reconstruction Method for a Pinhole Aperture Single Photon Emission Computed Tomograph (SPECT). *IEEE Trans. Medical Imaging*, 9(2):128–143, 1990.

[6] S. Webb. SPECT Reconstruction by Simulated Annealing. *Phys. Med. Biol.*, 34(3):259–281, 1989.

[7] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller. Equation of State Calculations by Fast Computing Machines. *J. of Chem. Phys.*, 21(6):1087–1092, 1953.

[8] P.J.M. van Laarhoven and E.H.L. Aarts. *Simulated Annealing: Theory and Applications.* Kluwer Ac. Publ., Dordrecht, 1987.

[9] S. Geman and D. Geman. Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 6:721–741, 1984.

[10] B.R. Frieden. Restoration of Pictures by Monte-Carlo Allocation of Pseudograins. In *Proc. 2nd Int. Joint Conf. on Pattern Recognition (IEEE, New York)*, pages 141–142, Copenhagen, 1974.

[11] K.A. Girodias, H.H. Barrett, and R.L. Shoemaker. Parallel Simulated Annealing for Emission Tomography. *Phys. Med. Biol.*, 36(7):921–938, 1991.

# On the use of successive data in the ML-EM algorithm in Positron Emission Tomography

P. Desmedt,* I.Lemahieu†

University of Ghent - ELIS Department

Sint-Pietersnieuwstraat 41, B-9000 Gent, Belgium

## Abstract

The Maximum Likelihood-Expectation Maximization (ML-EM) algorithm is the most popular statistical reconstruction technique for Positron Emission Tomography (PET). The ML-EM algorithm is however also renowned for its long reconstruction times. An acceleration technique for this algorithm is studied in this paper.

The proposed technique starts the ML-EM algorithm before the measurement process is completed. The ML-EM algorithm thus begins the reconstruction with temporary data. This temporary data is updated, yielding successive data, as the measurement process proceeds. Since the reconstruction is initiated during the scan of the patient, the time elapsed before a reconstruction becomes available is reduced.

Experiments with software phantoms indicate that the quality of the reconstructed image using successive data is comparable to the quality of the reconstruction with the normal ML-EM algorithm.

## 1 Introduction

A PET scanner is a tomographic device to investigate the metabolic functioning of a patient [Phelps86]. A PET scan produces a number of noisy projection measurements. The noise on the data can result in severe noise artifacts in the reconstructed image obtained by standard techniques, i.e., Filtered Backprojection [Herman79]. To avoid noise artefacts statistical reconstruction methods are examined. The most popular statistical reconstruction method is the Maximum Likelihood-Expectation Maximization method [Shepp82, Lange84, Vermeulen83]. This method succeeds in retrieving a reasonable good image. But the ML-EM method entails a substantial enlargement of the reconstruction time.

The proposed acceleration method starts the reconstruction before the measurement process is completed. Thus only part of the final data is available. This temporary data is extended as the measurements proceed. Two different decompositions of the final data are studied: angular data and additive data (see next section).

Of major interest is the number of partial data sets the total data can be decomposed in. Using more partial data sets will result in a larger reduction of the waiting time. But breaking up the data into too many sets can also result in poorer reconstructions. Indeed, the information (signal to noise) provided by the first temporary data sets decreases as the partioning of the final data set increases.

Experiments to determine the trade-off between reduction of reconstruction time and image quality are performed on software phantoms.

## 2 The use of successive data

The ML-EM algorithm implies the following iterative scheme:

$$\mathbf{f}^{(k)} = \frac{\mathbf{1}_N}{\mathbf{A}^t \mathbf{1}_L} \cdot \mathbf{A}^t \left( \frac{\mathbf{d}}{\mathbf{A} \mathbf{f}^{(k-1)}} \right) \cdot \mathbf{f}^{(k-1)}, \tag{1}$$

where $\mathbf{f}^{(k)}$ (dimension $N$) is the estimated image at the $k$-th iteration and $\mathbf{d}$ (dimension $L$) is the measured data. The transfer matrix $\mathbf{A}$ (dimension $L \times N$) models the measurement process of the scanner. The "·" and the divisions are element-wise operations.

The ML-EM algorithm searches the image that maximizes the likelihood of the data. In contrast with Maximum A Posteriori methods, e.g., [Hanson87], the ML-EM algorithm does not use any particular prior distributions for the image. The data $\mathbf{d}$ of this iterative scheme usually consists of the measurements after the scan is finished. In this paper, the iterative procedure is started before the final data is available. According to different scanning modalities two different types of partioning the final data are considered: the angular partial data and the additive partial data. The temporary data used in the ML-EM algorithm is obtained by combining a number of partial data sets of the same type.

The angular partial data consists of measurements for only a restricted number of view angles. This simulates a camera where the data for different view angles are successively collected. The view angles are equally spread over 180 degrees. The new temporary data contains data for twice the number of angles of the previous data. The updated data corresponds to equally spread angles and includes all the angles of the previous set.

The additive partial data procedure handles data collected for all angles simultaneously. In this case the number of counts in each projection bin increases with the measurement time. The new temporary data equals the previous temporary data with another partial data set added to it. All partial data sets have approximately the same signal to noise ratio. Since the data has a Poisson statistics, the signal to noise ratio of the temporary data increases as more and more partial data sets are added.

## 3 Experiments

The experiments are performed on a software phantom with intensities based on real scanner measurements. The phantom is digitized on a 64 by 64 grid. The phantom is represented in fig. 1. A range of activity levels of the measurements is used. The data (with 64×64 measurement points) is generated from the software phantom by application of a projection procedure (without considering attenuation, scatter, etc.). Poisson noise is added on all the partial data sets. The figures presented here only show the most typical results. Other experiments confirm these results.

The transitions between the different temporary data are triggered by the change of log likelihood of the data (this is inspired on [Ranganath88]). If the increase of log likelihood is too small between two successive iteration steps (less than 1%), the temporary data is updated.

As a measure of the image quality, the scaled quadratic distance is taken between the estimated image $\mathbf{f}^{(k)}$ and the reference image $\mathbf{f}^{(ref)}$. The scaled quadratic distance is given by

$$sc\_dist = \frac{1}{N} \sum_{i=0}^{N} \left( \frac{1}{(f_i^{(ref)})^2} (f_i^{(k)} - f_i^{(ref)})^2 \right), \quad f_i^{(ref)} \neq 0. \tag{2}$$

Fig. 2.1 shows the evolution of the scaled distance as a function of the iteration number for the additive data. The mean intensity of the final data is 800. The different curves correspond
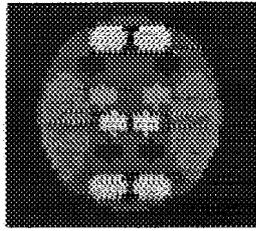
Figure 1: The PET phantom

to a partioning of the final data set into 2, 4 and 8 parts and the normal ML-EM algorithm. The figure reveals that the lowest value for the sc_dist criterion is obtained by the normal ML-EM algorithm. As the partitioning number increases, the ML-EM reconstructions using the partial data get worse. However (considering the large scale of the y-axis) the penalty for the sc_dist introduced by the successive data procedure is always small.

Fig. 2.2 shows the sc_dist evolution as a function of the iteration number for the angular data. The curves on this figure correspond to the reconstructions for mean intensity in the final data set of 100. The figure illustrates the importance of the incorporation of a convolution step at the updating of temporary data. Without this convolution the image qualify suffers from the successive data procedure. This deterioration of the image quality can entirely be avoided by the convolution. This convolution step also becomes necessary in the additive data procedure whenever number of counts becomes too low.
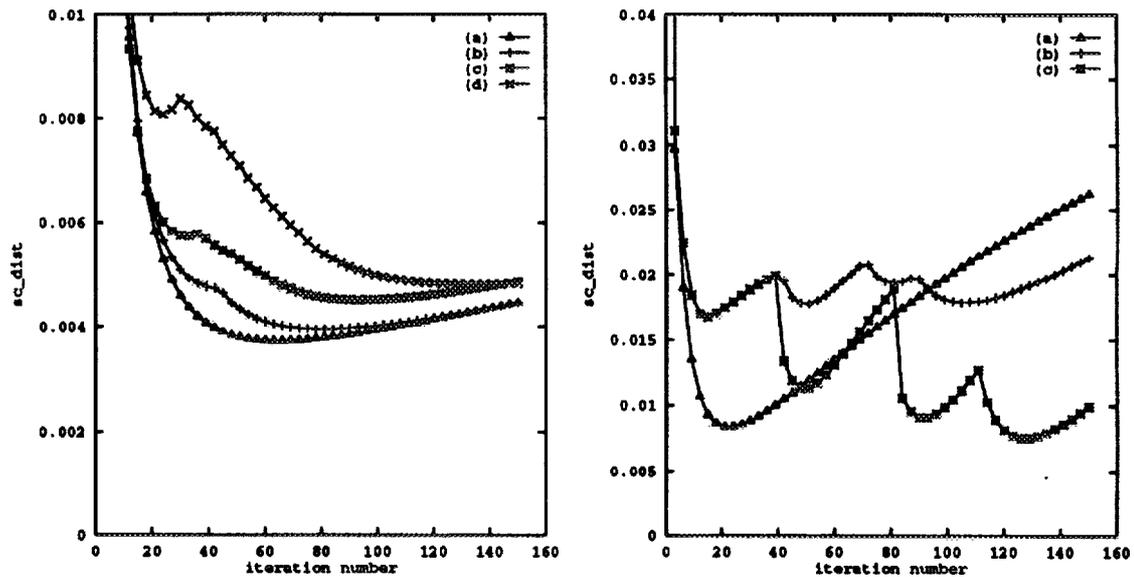


Figure 2: The sc_dist criterion as a function of the iteration number for (1) additive data and (2) angular data. (1) Additive data (mean intensity level 800): number of successive data: (a) standard ML-EM, (b) 2, (c) 4, (d) 8. (2) Angular data (mean intensity level 100): The starting number of angles is 8: (a) standard ML-EM, (b) successive data without interpolation (c) successive data with interpolation

182

Fig. 3 shows reconstructions for additive data (with mean intensity 100 and decomposition in 8 sets), compared to the images obtained with the standard ML-EM algorithm. It is clear that at iteration 150 the image produced by the standard ML-EM algorithm (fig. 3.2) is comparable to the image obtained at the same iteration by the successive data image (fig. 3.3). The image at iteration 20 (fig. 3.1) of the standard ML-EM algorithm is similar to the last image of the successive data with interpolation (fig. 3.4). This indicates that even for such poor data, the image quality does not suffer from the successive data procedure.

(1)        (2)



(3)        (4)
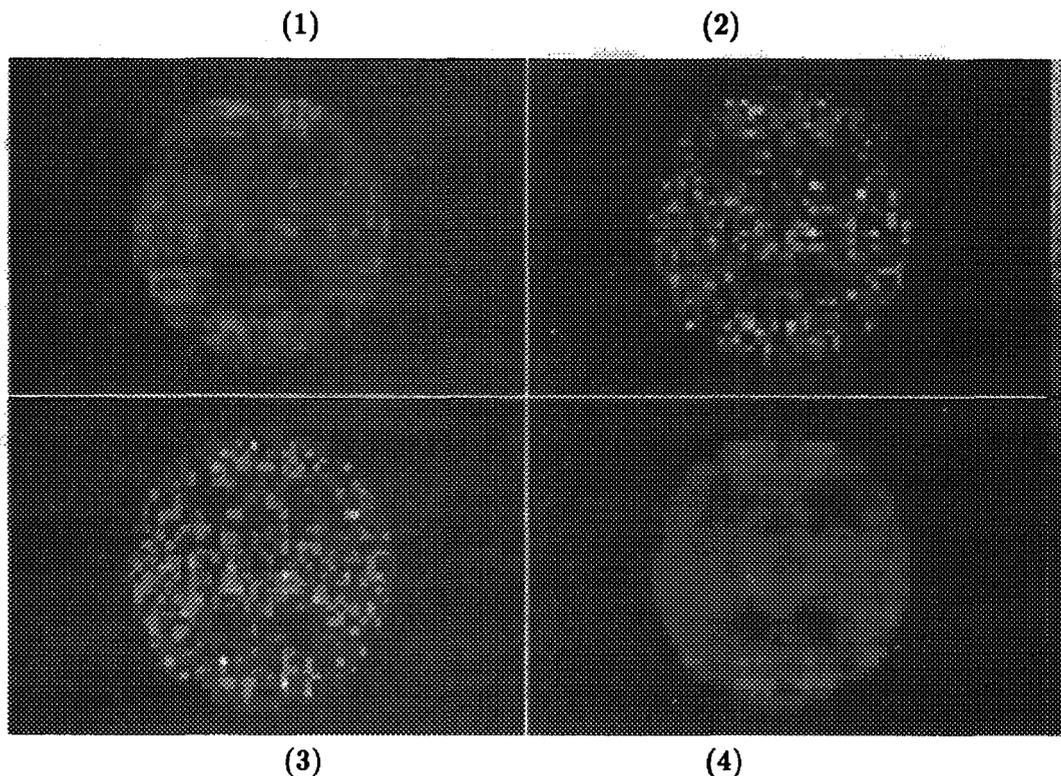
Figure 3: Reconstructed images: (1) standard ML-EM reconstruction at iteration 20, (2) standard ML-EM reconstruction at iteration 150, (3) partial additive data reconstruction without interpolation at iteration 150, (4) partial additive data reconstruction with interpolation at iteration 150

## 4  Conclusions

The experiments with software phantoms indicate that the use of successive data does not result in a loss of reconstructed image quality. For final data sets with a low number of counts a convolution step at the transitions of temporary data are required. These conclusions are valid for both examined types of successive data: additive and angular partial data.

The actual reduction of the time before the scan is available to the physician will depend on the length of the scan time and the power of the computer on which the ML-EM algorithm runs. It is conceivable that within a few years, when CPU speed will be increased by another factor 10, the ML-EM reconstructed images will be at hand when the patient is removed from the scanner.

# References

[Phelps86]    Positron Emission Tomography and Autoradiography: Principles and Applications for the brain and heart, eds. M. Phelps, J.C. Mazziotta, H.R. Schelbert, Raven Press, New York, 1986

[Herman79]    Herman G.T., *Topics in applied physics Vol 32: Image reconstruction from projections*, 1979, Ed. G.T. Herman, Springer-Verlag

[Shepp82]     Shepp L.A., Vardi Y., *Maximum likelihood reconstruction for emission tomography*, IEEE Trans. Med. Imaging, Vol. MI-1, 1982, pp. 113-122

[Lange84]     Lange K., Carson R., *EM reconstruction algorithms for emission and transmission tomography*, J. Comput. Assist. Tomogr., 1984, pp. 306-316

[Vermeulen83] Vermeulen F.L., *On the performance of maximum-likelihood reconstructions in positron emission tomography*, Proceedings of the 1983 SPIE International Technical Conference/Europe, Vol. 397, Applications of digital image processing, 1983, pp. 274-279

[Hanson87]    Hanson K.M., *Bayesian and related methods in image reconstruction from incomplete data*, in Image recovery: theory and application, Ed. H. Stark, Academic Press, 1987, pp. 79-125

[Ranganath88] Ranganath M.V., Dhawan A.P., Mullani N., *A multigrid expectation maximization reconstruction algorithm for positron emission tomography*, IEEE Trans. Med. Imaging, Vol. 7, 1988, pp. 273-278

# A SPECIAL DESIGNED LIBRARY FOR MEDICAL IMAGING APPLICATIONS

D. LYMBEROPOULOS*, C. SPYROPOULOS** S. KOTSOPOULOS*, V. ZOUPAS* and N. YOLDASSIS*

*Dept. of Electrical Engineering University of Patras, Patras 26 110 GREECE
**School of Medicine, Regional University Hospital University of Patras, Patras 26 110 GREECE

Abstract: The present paper deals with a sophisticated and flexible library of medical purpose image processing routines. It contains modules for simple as well as advanced gray or color image processing.

## I. INTRODUCTION

Medical Electronics manufacturing and research industry in all over the world has created many diagnostic imaging systems [1] which have been in use for more than a decade. A wide range of the most sophisticated modalities [2] can be seen in any modern hospital. Physisians and radiologists in cooperation with engineers have carried out medical examinations by operating the high perfomance commercial Image Diagnostic Systems.

The communication between radiologists and physicians is under the direction of a new scenario based upon a dynamic interactive diagnosis and consultancy procedure [3, 4]. For this reason, image manipulations during diagnostic procedures have to be studied and analyzed in order to create sets of appropriate image processing functions. Medical Workstations supporting these funcions are capable to manipulate and display the acquired images from various modalities in a way similar to the modality's monitor.

Medical Modalities produce a lot of digital images of different size, depth and resolution. Doctors have to analyze and estimate these medical images very carefully in order to accomplish their diagnostic procedures. Recognizing the need and the importance for the physician to select the appropriate image processing function suitable for each medical case [5], and the potential developer to write and develop image processing applications using code portable and easy to maintain, this medical purpose library designed very carefully providing:

- Object orientation
- High degree of security. All of the input and output parameters are validated in each routine protecting users from errors
- No need of complicated parameters declaration
- Ability of handling gray and color images produced by different modalities, thus having different attributes
- Ability of handling ROIs (Regions Of Interest) of any shape

The developed routines can be used by experts in the area of image processing. They can use the well known algorithms included in this library, or develop their own modules by using and combining the routines.

In order to help users compile, link and execute their programs, a simple parser has also been implemented. Using this parser the routines can be treated as language commands instead of simple routine calls.

## II. DESCRIPTION OF ROUTINES

The routines of the developed library [6] are written in ANSI C language on a HP Apollo computer running HP - UX operating system. The use of ANSI C and the modular structure of the routines make the library compatible to any machine that runs this language, regardless of its

operation system or processing architecture. Thus, the potential developer is able to develop sophisticated applications working with X - Windows, MS - Windows, APPLE Macintosh System 7, or any other GUI, having no limits in the combination of the routines, structures and the format of the image data.

All the modules are function calls. The return value of each function is an error code. These errors occur when the library can not proceed with an operation probably due to invalid parameters, or when the operation is finished but some minor problems have been detected thus depending on the user to decide for any possible action or not.
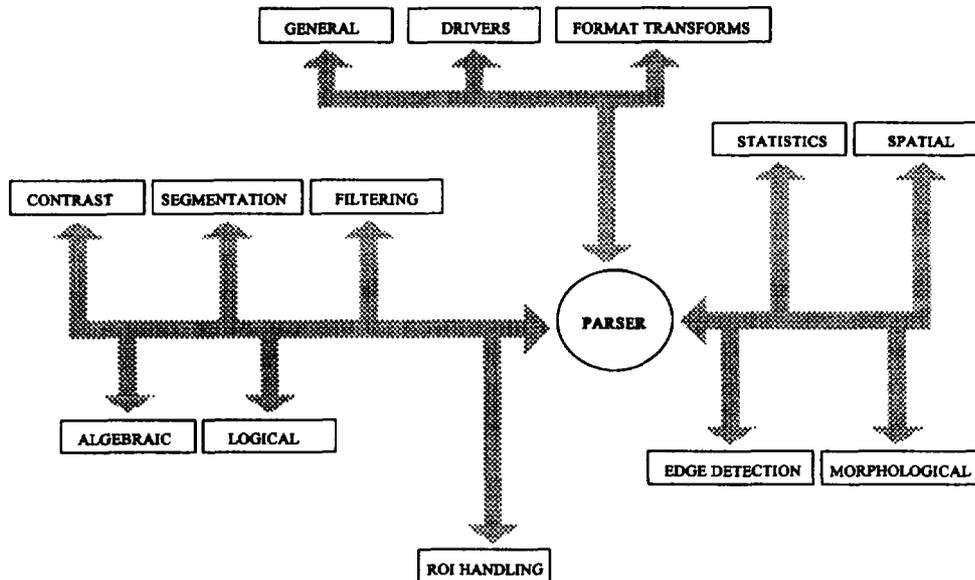


Fig. 1

The routines have been associated into thirteen different groups depending on their functionality ( Fig. 1).

Routines for general purpose image manipulation such as object creation, object destruction, image copy etc., which are essential in any image processing application.

The library offers drivers for the communication with high quality graphics monitors, hardcopiers and various peripherals used in a hospital.

Routines for image format transformations are also included in the package. All major image formats are supported, providing the ability to handle TIFF, BMP, XWD, PICT, GIF etc.

Routines that perform contrast modification is an other powerful feature of the library, dealing with images that do not properly utilize the allowable grayscale or colorscale.

Other features of the implemented library include extraction of statistical information using data reduction routines which are of great importance in image analysis.

Routines for spatial(operations are also included in the library. There are routines for zooming, rotation and other transformations, very useful for comparison of pairs of organs in a single individual, or for improving the detail within an image.

Special routines are offered for image filtering, which can be used for noise removal (low pass filters), as well as for image enhancement (high pass filters).

Algebraic operations is an other class of algorithms implemented in the software package. Single-operand and Double-operands as well as multiple-operands operators are supported, very useful in removing undesired patterns or objects from an image, superimposing images, simulating the use of a medical light box with variable intensity, and even noise reduction.

Routines for bitwise logical operations are also included, providing a fast way of handling images for a variety of purposes.

Other features of the software package include segmentation, which is a key step in medical image analysis. It entails the separation of the image into regions which correspond to objects or parts of objects. Such algorithms are thresholding, suppression, window slicing, adaptive

186

thresholding, contouring etc.

Sophisticated routines for edge detection are supported too, the fomer being a problem of fundamental importance in medical image analysis. Many different algorithms have been implemented, thus providing a way of processing a variety of medical images.

Morphological processing is an other class of routines supported by the library, dealing with the modification of the spatial form or structure of objects within a grayscale or colorscale image. This is of great importance in medical applications.

Routines for handling ROIs is an other powerful feature of the library. These are very useful when only a small part of the original image is of interest, or when high speed of processing has to be achieved. Rectangle and irregular ROIs are supported, the latter describing ROIs of any shape.

## III. CONCLUSIONS

In this paper a special purpose library has been presented. It is a sophisticated modular medical purpose library of image processing routines.

This library offers powerful features for medical image processing and analysis applications, thus providing the physicians with a means of analyzing and estimating medical images in order to accomplish their diagnostic procedures.

The software has been developed on a HP Apollo machine under HP - UX operating system. It is written in ANSI C offering full compatibility with other systems running C. The overall design helps users in the development of any gray or color image processing application.

## IV. REFERENCES

[1] G. Kossoff, "Diagnostic Imaging Systems", J. of Electrical and Electronics Eng., Australia, Vol. 4, No 2, June 1984, pp 93 - 100.

[2] D. Myers, "Developments in Digital Image Processing Equipment", J. of Electrical and Electronics Eng., Aust., Vol. 6, No 3, Sept 1986, pp. 206 - 211.

[3] Dvering A. J., Dwyer S. J., Prewitt J. M., "Digital Picture Archiving and Communication System in Medicine, Computer Mag.16, No. 8, 1983 pp.14.

[4] Mun S. K., Freedman M, Kapur R., "Image Management And Communications For Radiology, IEEE Eng. in Medicine and Biology Mag, Vol. 12, No 1, March 1993, pp. 70.

[5] D. Lymberopoulos, S. Kotsopoulos and G. Kokkinakis, "A Prototype Medical Workstation with Unified Image Viewing Facilities", J. of Electrical and Electronic Eng. Aust., ;Vol 12, No 1, March 1992.

[6] N. Yoldassis, S. Kotsopoulos, V. Zoupas, D. Lymberopoulos, "HIPPOCRATES: A New Software Package for Medical Imaging Applications", Proceedings of the 15th Annual International Conference of the IEEE EMBS, October 1993.

# Two Dimensional Nonlinear Spectral Estimation Techniques for Breast Cancer Localization

*P.T. Stathaki   A.G. Constantinides*

Signal Processing Section, Department of Electrical and Electronic Engineering
Imperial College,  Exhibition Road,  London SW7 2BT,  UK
Telephone: (071)-225 8520    Fax: (071)-823 9919    email: a.constantinides@ee.ic.ac.uk

## Abstract

*In this paper the problem of image texture analysis in the presence of noise is examined from a higher-order statistical perspective. The approach taken involves the use of two dimensional second order Volterra filters where the filter weights are derived from third order cumulants of the two dimensional signal. The specific application contained in this contribution is in mammography, an area in which it is difficult to discern the appropriate features. The paper describes the fundamental issues of the various components of the approach. The results of the entire texture modelling, classification and segmentation scheme contained in this paper are very encouraging.*

## 1   Introduction

In many problems of digital signal processing it is necessary to introduce nonlinear systems as for example, in the  well known  detection and estimation problems where nonlinear filters arise when the Gaussian assumptions are not valid. In the search for efficient nonlinear signal processing systems, the task is often concerned with the determination of general characterisation procedures for such systems that retain at least a part of the simplicity that the impulse response method affords for linear filters. A possible way to describe the input-output relationship in a nonlinear filter that is amenable to characterisation, analysis and synthesis is the use of discrete Volterra filter representations. Although the class of nonlinear systems that can be represented by the Volterra series expansion is not extensive, nevertheless, the Volterra series is one of a few representations that are easily amenable to characterisation and analysis. In many cases, a nonlinear system can be represented by a truncated version of the Volterra series, which results in a simpler representation and requires a limited amount of knowledge of higher order statistics[2].

If the input-output relation is restricted to the third term of a Volterra processor, the system becomes a quadratic filter. It is very common to reduce a nonlinear system to the quadratic part. In image processing applications, linear filters tend to blur the edges, do not remove impulsive noise effectively and do not perform well in the presence of signal dependent noise. It is also known that, although the exact characteristics of our visual system are not well understood, experimental results indicate that the first processing levels of our visual system possess nonlinear characteristics. For such reasons, nonlinear systems are a very powerful tool in image processing[2].

In this paper we use specific types of nonlinear systems for modelling and segmentation of mammograms. Mammograms are amongst the most difficult of radiological images to interpret because (i) the contrast of different regions in the mammogram is often very low (ii) features in mammograms indicative of breast disease are often very small[6] and (iii) the noise component may be too high. The specific objective of our study makes the above conditions even more difficult in that we desire to localise the areas of breast cancer at the very early stage of formation. The size, therefore, of such lesions is very small and hence the discriminatory ability of any technique capable of resolving their presence must be very powerful.

## 2   Mathematical Representation of the Image

We represent the image with a 2-D random field $y[m,n]$, where $y[m,n]$ denotes the value of   the random field at the point $[m,n]$ defined theoretically over the integers $-\infty < m,n < \infty$ in the 2-D plane. A typical finite extent for a realistic sequence is the measured data array, which is always assumed to have a region of support $0 \le m \le M-1, 0 \le n \le N-1$[1]. Consider an image $y[m,n]$ that has one or more types of different textural primitives. For the purposes of spectral estimation and modelling we represent $y[m,n]$ as the output of a two dimensional non linear time invariant system driven by  noise $x[m,n]$

$$y[m,n] = x_o + \sum_{i_1}\sum_{j_1} a[i_1,j_1]x[m-i_1,n-j_1] + \sum_{i_1}\sum_{j_1}\sum_{i_2}\sum_{j_2} b[(i_1,j_1),(i_2,j_2)]x[m-i_1,n-j_1]x[m-i_2,n-j_2]$$

$$+ \sum_{i_1}\sum_{j_1}\sum_{i_2}\sum_{j_2}\sum_{i_3}\sum_{j_3} c[(i_1,j_1),(i_2,j_2),(i_3,j_3)]x[m-i_1,n-j_1]x[m-i_2,n-j_2]x[m-i_3,n-j_3]+...$$

where $a[i_1,j_1],b[(i_1,j_1),(i_2,j_2)],c[(i_1,j_1),(i_2,j_2),(i_3,j_3)],...,$ are called the linear, quadratic, cubic etc. filter weights respectively, $x_o$ is a constant term whose value depends on the input $x[m,n]$ and $0 \le i_1,i_2,j_1,j_2 \le N-1$ where $N$ denotes the filter length[2]. The two dimensional second order Volterra filter is given by the following relationship:

$$y[m,n] = \sum_{i_1}\sum_{j_1} a[i_1,j_1]x[m-i_1,n-j_1] + \sum_{i_1}\sum_{j_1}\sum_{i_2}\sum_{j_2} b[(i_1,j_1),(i_2,j_2)]x[m-i_1,n-j_1]x[m-i_2,n-j_2]$$

where we assumed that $x_o = 0$ without loss of generality[7]. By assuming that the input signal $x[m,n]$ is a discrete, stationary, white Gaussian process, the output process is also discrete, stationary, non-Gaussian process[2]. It is observed that we do not require the mean of the output to be zero.

In this research the only information about the texture image that assumed to be known a priori is the number of texture classes. In all other senses the image is completely unknown. For this reason our method consists of representing each pixel in the image by a nonlinear model, parameters of which are estimated by using an appropriate neighbourhood for that pixel, for example a rectangular window. Then these parameters are used as features for classification. We make the assumption, that all pixels in the small window belong to the same class. The results of the segmentation are influenced slightly by the size of the window over which filter parameters are extracted[5].

## 3    Second Order Statistical Analysis

The autocorrelation function of the real process $y[m,n]$ is given by:
$$R[k,l] = E\{y[m,n]y[m+k,n+l]\}$$
and in view of the second order Volterra model can be written as:
$$R[k,l] = \sum_{i_1}\sum_{j_1} a[i_1,j_1]E\{x[m-i_1,n-j_1]y[m+k,n+l]\}$$

$$+\sum_{i_1}\sum_{j_1}\sum_{i_2}\sum_{j_2} b[(i_1,j_1),(i_2,j_2)]E\{x[m-i_1,n-j_1]x[m-i_2,n-j_2]y[m+k,n+l]\} \qquad (4.1)$$

It is observed here that $b$ is a symmetric kernel, that is to say $b[(i_1,j_1),(i_2,j_2)] = b[(i_2,j_2),(i_1,j_1)]$.

The terms of equation (4.1) involve averaging over the product of one, two, three and four Gaussian random variables. It is well known in statistics that the average of the product of an odd number of zero-mean jointly Gaussian random variables is identically zero irrespective of their correlation. Moreover, the average of the product of an even number of zero-mean jointly Gaussian random variables is equal to the summation over all distinct ways of partitioning the random variables into products of averages of pairs[2]. For example, if $x_1,x_2,x_3,x_4$ are zero-mean jointly Gaussian random variables, then:

$$E\{x_1x_2x_3\} = 0 \qquad (4.2)$$
$$E\{x_1x_2x_3x_4\} = E\{x_1x_2\}E\{x_3x_4\} + E\{x_1x_3\}E\{x_2x_4\} + E\{x_1x_4\}E\{x_2x_3\} \qquad (4.3)$$

With (4.2) and (4.3), $R[k,l]$ reduces to the form:

$$R[k,l] = \beta\sum_{i_1}\sum_{j_1} a[i_1,j_1]a[i_1+k,j_1+l] + \beta^2\sum_{i_1}\sum_{j_1}\sum_{i_2}\sum_{j_2} b[(i_1,j_1),(i_2,j_2)]b[(i_1+k,j_1+l),(i_2+k,j_2+l)]$$

$$+\beta^2\sum_{i_1}\sum_{j_1}\sum_{i_2}\sum_{j_2} b[(i_1,j_1),(i_2,j_2)]b[(i_2+k,j_2+l),(i_1+k,j_1+l)]$$

$$+\beta^2\sum_{i_1}\sum_{j_1}\sum_{i_2}\sum_{j_2} b[(i_1,j_1),(i_1,j_1)]b[(i_2,j_2),(i_2,j_2)] +2\beta^2\sum_{i_1}\sum_{j_1} b[(i_1,j_1),(i_1,j_1)]b[(i_1+k,j_1+l),(i_1+k,j_1+l)] \qquad (4.4)$$

where $\beta$ is the variance of the input driving noise $x[m,n]$.

The autocorrelation function given by (4.4) is not sufficient to solve the problem because the number of unknowns present in (4.4) is much greater than the number of useful samples of $R[k,l]$. However, additional information can be provided by examining higher order statistics[2].

## 4    Third Order Statistical Analysis

If we define $M[(k_1,l_1),(k_2,l_2)]$ to be the third order moment sequence of $y[m,n]$, then[4]:

$$M[(k_1,l_1),(k_2,l_2)] = E\{y[m,n]y[m+k_1,n+l_1]y[m+k_2,n+l_2]\} \qquad (5.1)$$

In the following, the third-order moment sequence of the second order Volterra filter is derived. First we use the following symbols:

$$G_1[k,l] = \sum_{i_1}\sum_{j_1} a[i_1,j_1]x[m-i_1+k,n-j_1+l] \qquad (5.2)$$

$$G_2[k,l] = \sum_{i_1}\sum_{j_1}\sum_{i_2}\sum_{j_2} b[(i_1,j_1),(i_2,j_2)]x[m-i_1+k,n-j_1+l]x[m-i_2+k,n-j_2+l] \qquad (5.3)$$

Based on (5.2), (5.3) one can easily expand (5.1) in the following compact form[2]:

$$\begin{aligned}
M[(k_1,l_1),(k_2,l_2)] = E\{ &G_1[0,0]\cdot G_1[k_1,l_1]\cdot G_1[k_2,l_2] + G_1[0,0]\cdot G_1[k_1,l_1]\cdot G_2[k_2,l_2] \\
&+ G_1[0,0]\cdot G_2[k_1,l_1]\cdot G_1[k_2,l_2] + G_1[0,0]\cdot G_2[k_1,l_1]\cdot G_2[k_2,l_2] \\
&+ G_2[0,0]\cdot G_1[k_1,l_1]\cdot G_1[k_2,l_2] + G_2[0,0]\cdot G_1[k_1,l_1]\cdot G_2[k_2,l_2] \\
&+ G_2[0,0]\cdot G_2[k_1,l_1]\cdot G_1[k_2,l_2] + G_2[0,0]\cdot G_2[k_1,l_1]\cdot G_2[k_2,l_2] \} \qquad (5.4)
\end{aligned}$$

The first, fourth, sixth and seventh terms of (5.4) involve averaging over an odd number of zero-mean jointly Gaussian random variables. Therefore are identically zero. Equation (5.4) then becomes:

$$\begin{aligned}
M[(k_1,l_1),(k_2,l_2)] = E\{ &G_1[0,0]\cdot G_1[k_1,l_1]\cdot G_2[k_2,l_2] + G_1[0,0]\cdot G_2[k_1,l_1]\cdot G_1[k_2,l_2] \\
&+ G_2[0,0]\cdot G_1[k_1,l_1]\cdot G_1[k_2,l_2] + G_2[0,0]\cdot G_2[k_1,l_1]\cdot G_2[k_2,l_2] \} \qquad (5.5)
\end{aligned}$$

Each term of (5.5) involves averaging over an even number of zero-mean jointly Gaussian random variables. Keeping in mind the procedure we described in the previous paragraph, one can decompose the average of the product of an even number of jointly Gaussian random variables into a summation of products of averages of pairs. The first term of (5.5) (not using the fact the fact that $b[(i_1,j_1),(i_2,j_2)]$ is a symmetric kernel), can then be written as follows:

$$E\{G_1[0,0]\cdot G_1[k_1,l_1]\cdot G_2[k_2,l_2]\} = E\{(\sum_{i_1}\sum_{j_1} a[i_1,j_1]x[m-i_1,n-j_1])\cdot(\sum_{i_1}\sum_{j_1} a[i_1,j_1]x[m-i_1+k_1,n-j_1+l_1])$$

$$\cdot(\sum_{i_1}\sum_{j_1}\sum_{i_2}\sum_{j_2} b[(i_1,j_1),(i_2,j_2)]x[(m-i_1+k_2,n-j_1+l_2)x(m-i_2+k_2,n-j_2+l_2)])\} =$$

$$\beta^2 \sum_{i_1}\sum_{j_1}\sum_{i_2}\sum_{j_2} a[i_1,j_1]a[i_1+k_1,j_1+l_1]b[(i_2,j_2),(i_2,j_2)] +$$

$$+2\beta^2 \sum_{i_1}\sum_{j_1}\sum_{i_2}\sum_{j_2} a[i_1+k_2,j_1+l_2]a[i_2+k_2-k_1,j_2+l_2-l_1]b[(i_1,j_1),(i_2,j_2)]$$

Now we define $\phi_1[k_1,l_1]$ and $\phi_2[(k_1,l_1),(k_2,l_2)]$ to be as follows:

$$\phi_1[k_1,l_1] = \sum_{i_1}\sum_{j_1}\sum_{i_2}\sum_{j_2} a[i_1,j_1]a[i_1+k_1,j_1+l_1]b[(i_2,j_2),(i_2,j_2)]$$

$$\phi_2[(k_1,l_1),(k_2,l_2)] = \sum_{i_1}\sum_{j_1}\sum_{i_2}\sum_{j_2} a[i_1+k_1,j_1+l_1]a[i_2+k_2,j_2+l_2]b[(i_1,j_1),(i_2,j_2)] \quad \text{so that}$$

$$E\{G_1[0,0]\cdot G_1[k_1,l_1]\cdot G_2[k_2,l_2]\} = \beta^2\phi_1[k_1,l_1] + 2\beta^2\phi_2[(k_2,l_2),(k_2,l_2)-(k_1,l_1)] \qquad (5.6)$$

Similarly, one can show that:

$$E\{G_1[0,0]\cdot G_2[k_1,l_1]\cdot G_1[k_2,l_2]\} = \beta^2\phi_1[k_2,l_2] + 2\beta^2\phi_2[(k_1,l_1),(k_1,l_1)-(k_2,l_2)] \qquad (5.7)$$

$$E\{G_2[0,0]\cdot G_1[k_1,l_1]\cdot G_1[k_2,l_2]\} = \beta^2\phi_1[(k_2,l_2)-(k_1,l_1)] + 2\beta^2\phi_2[-(k_1,l_1),-(k_2,l_2)] \qquad (5.8)$$

The fourth term of (5.5) is quite different from the first three terms. It involves averaging over the product of four Gaussian random variables as well as averaging over the product of six Gaussian random variables. The later can be broken into the sum of fifteen terms, where each term involves a product of three averages of distinct pairs of random variables. By doing so and defining:

$$\phi_3[(x_1,y_1),(x_2,y_2),(x_3,y_3)] =$$

$$= \sum_{i_1}\sum_{j_1}\sum_{i_2}\sum_{j_2}\sum_{i_3}\sum_{j_3} b[(i_1,j_1),(i_2,j_2)]b[(i_1,j_1)+(x_1,y_1),(i_3,j_3)+(x_2,y_2)]b[(i_2,j_2)+(x_3,y_3),(i_3,j_3)]$$

$$k_o = \beta\sum_{i_1}\sum_{j_1} b[(i_1,j_1),(i_1,j_1)] \quad \text{we obtain:}$$

$$E\{G_2[0,0]G_2[k_1,l_1]G_2[k_2,l_2]\} =$$

$$= 8\beta^3\phi_3[(k_2,l_2),(k_2,l_2)-(k_1,l_1),(k_1,l_1)] + k_o(R[k_1,l_1]+R[k_2,l_2]+R[(k_2,l_2)-(k_1,l_1)]) - 2k_o^3 \qquad (5.9)$$

We replace (5.6), (5.7), (5.8), (5.9) in (5.5) and we obtain $M[(k_1,l_1),(k_2,l_2)]$. It is now possible to use (4.4) in conjunction with (5.5) to provide a sufficient number of equations required to solve for the unknown system parameters $a$ and $b$ and the variance of the white Gaussian process $\beta$. The equations however, are nonlinear, and hence a least-squares method and other similar methods can be used to achieve the solution. In this paper we will use specific types of two dimensional second order Volterra systems that provide a simplified solution to these equations[2].

# 5  Simulations

For the mammograms we use a quadratic model with simple linear term:

$$y[m,n] = x[m,n] + \sum_{i_1}\sum_{j_1}\sum_{i_2}\sum_{j_2} b[(i_1,j_1),(i_2,j_2)]x[m-i_1,n-j_1]x[m-i_2,n-j_2], \qquad 0 \le i_1,j_1,i_2,j_2 \le N \quad \text{and}$$

$(i_1,j_1) \ne (i_2,j_2)$. It is assumed that the noise $x[m,n]$ is white, Gaussian, with zero mean and unknown variance $\beta$. We determine the coefficients $b[(i_1,j_1),(i_2,j_2)]$ from third order cumulants up to the scale factor $\beta^2$ i.e.

$$b[(i_1,j_1),(i_2,j_2)] = \frac{1}{\beta^2} M[(k_1,l_1),(k_2,l_2)]$$

where $M[(k_1,l_1),(k_2,l_2)]$ are the third order cumulant samples of the output and $m$ is the mean of the output. Some indicative results shown in figures below. In these pictures we have used windows of size $32 \times 32$.
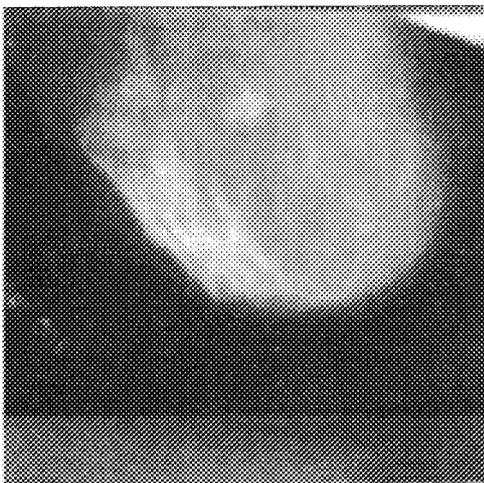


**Figure 1a:** Mammogram 1
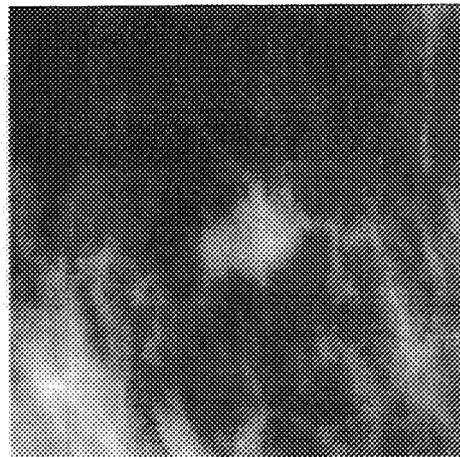


**Figure 1b:** Part of the mammogram around the tumour
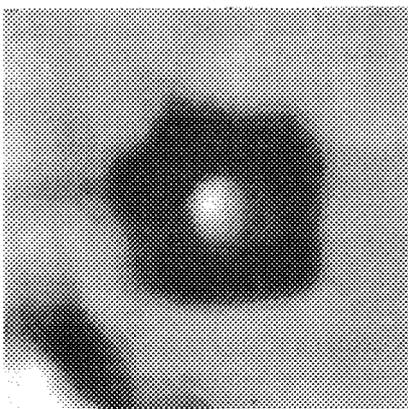


**Figure 1c:** $b[(0,1),(1,0)]$
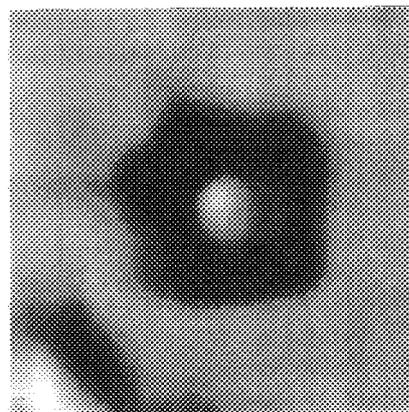


**Figure 1d:** $b[(0,1),(1,1)]$
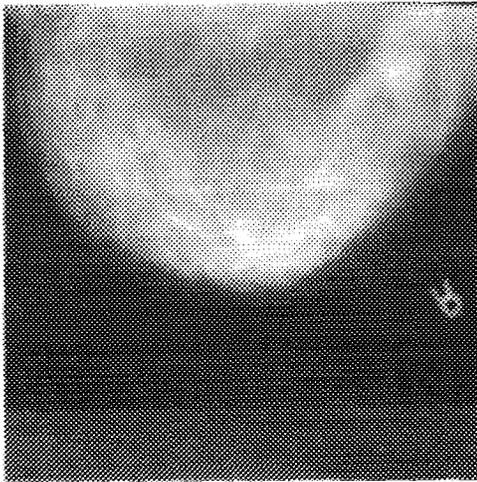
**Figure 2a:** Mammogram 2
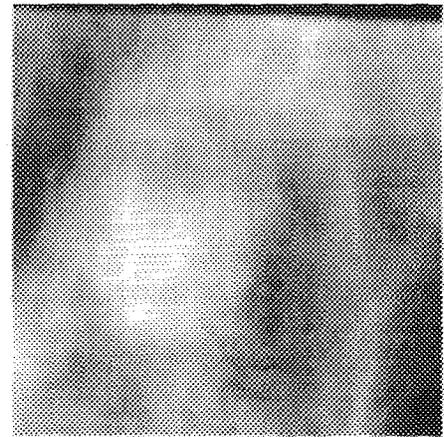


**Figure 2b:** Part of the mammogram around the tumour
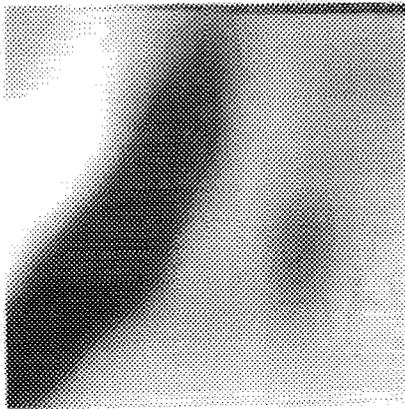


**Figure 2c:** $b[(0,1),(1,0)]$
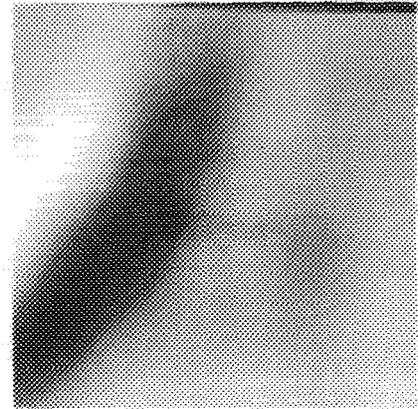


**Figure 2d:** $b[(0,1),(1,1)]$

## Remarks

The segmentation shown in figures 1,2 clearly indicates the region containing the malignant lesion. The discriminatory properties of the approach are evident in view of the small size of the lesion localised. Further research is under way to explore additional and useful properties of the method.

## REFERENCES

[1] Steven M Kay, *Modern Spectral Estimation: Theory and Application*, Prentice Hall 1988

[2] I.Pitas and A.N. Venetsanopoulos, *Nonlinear Digital Filters, Principles and Applications*, Kluwer Academic Publishers 1990

[3] Bart Kosko, *Neural Networks for Signal Processing*, Prentice Hall 1992

[4] C L Nikias and M Raghuveer, "Bispectrum Estimation: A digital signal processing framework", *Proc IEEE*, vol 75, pp 869-891, 1987

[5] P T Stathaki and A G Constantinides, "Noisy Texture Analysis based on Higher Order Statistics and Neural Network Classifiers", *Trans. IEEE Int. Conference on Neural Network Applications to Signal Processing*, Singapore, August 1993

[6] P T Stathaki and A G Constantinides, "Higher Order Spectral Estimation Techniques in Mammography", *Trans. Int. Conference on DSP*, Cyprus, July 1993

[7] T Koh and E J Powers, "Second-Order Volterra Filtering and Its Application to Nonlinear System Identification", *IEEE Trans. ASSP*, Vol. 33, No. 6. Dec. 1985

# PERFORMANCE EVALUATION OF BREAST IMAGE COMPRESSION TECHNIQUES

G. Anastassopoulos[*], G. Panayiotakis[**], D. Lymberopoulos[*]
and A. Bezerianos[**]

\* Wire Communications Laboratory, Electrical Engineering Dept.
\*\* Medical Physics Dept., School of Medicine
University of Patras

## ABSTRACT

Novel diagnosis orienting teleworking systems manipulate, store and process medical data through real time communication - conferencing schemes. One of the most important factors affecting the performance of these systems is image handling. Compression algorithms can be applied to the medical images, in order to minimize: a) the volume of data to be stored in the database, b) the demanded bandwidth from the network, c) the transmission costs, and to maximize the speed of the transmitted data. In this paper an estimation of all the factors of the process that affect the presentation of breast images is made, from the time the images are produced from a modality, till the compressed images are stored, or transmitted in a Broadband network (e.g. B-ISDN). The images used were scanned images of the TOR(MAX) Leeds breast phantom, as well as typical breast images. A comparison of seven compression techniques has been done, based on objective criteria such as Mean Square Error (MSE), resolution, contrast, etc. The user can choose the appropriate compression ratio in order to achieve the desired image quality.

## 1. INTRODUCTION

The use of digital techniques in various diagnostic modalities has increased. Even in conventional screen film radiography, digital techniques show great promise.

From the communication point of view, there are two major factors affecting the performance of a conferencing system handling multimedia data (especially images) and they are:

a)  *the system's response time*, defined as the time interval between the moment when a request arises by a conference participating entity (workstation, database, image acquisition device) and the moment when the image data is loaded in the entity's memory.

b)  *the quantity of exchanged data* during the telework, which depends on the running processes, the communication networks and protocols.

The increasing number of digital images requires additional storage capacity; image compression by error free or irreversible methods [1] not only enhances storage capacity but also increases the rate of image transmission. Most image coding systems provide a fixed compression ratio which allows reconstructed image quality to vary with the inherent "compressibility" of the original image. In order to achieve an acceptable quality, from the diagnostic point of view, an appropriate compression techniquehas to be applied at each case.

In the present paper we have applied a selected number of compression algorithms in both phantom and real breast images. The major target of this work was to compare the quality of the reconstructed images in both cases. Combination of basic compression algorithms have been used in order to improve the resulted quality and objective and subjective evaluation criteria have been employed. Performance evaluation of compression algorithms was assessed with the

aid of the Leeds TOR(MAX) breast phantom. It is one of the mammographic phantom more widely used in quality control and in comparative evaluation of mammographic systems [2]. Its physical and geometrical characteristics mimic microcalcifications and low contrast lesions, which are important clinical indications.

The implementation has been developed using a UMAX UC1200S Professional Colour Scanner, two SUN SPARC stations, a PC-486 equipped with extra hardware for image processing (MATROX MVP-AT real time imaging card), interconnected in a LAN.

Data from breast as well as phantom images were digitized, compressed and transmitted from one Medical Workstation (MWS) [3] to another, where they were decompressed and displayed in MWS's monitors and interpreted by radiologists. The original images were interpreted independently and compared with the compressed images, to determine the performance of some compression techniques in mammography.

The structure of this paper is as follows: In section 2, the implemented image compression techniques, are presented. In section 3, the applied compression techniques and the results are presented. Finally, some conclusions are given in section 4.

## 2. IMAGE COMPRESSION TECHNIQUES

General waveform coding can be classified into seven major categories: Statistical, PCM, predictive, transformation, hybrid, interpolative and extrapolative coding and the eighth class which consists of miscellaneous schemes that do not fall into any of the other seven classes [4]. Each of these classes can be further divided based on whether the parameters of the coder are fixed or whether they change as a function of the type of data that is being coded (adaptive). There are many image coding systems providing a fixed or adaptive compression ratio of the original image [5,6]. There are two major categories of image compression algorithms being used by the above systems:

a) The lossless ones where the reconstructed image is identical with the original image. The compression ratio in this case varies with the statistics (characteristics) of each image and is about 1:2 up to 1:5 [1].

b) The lossy compression algorithms for images that do not necessarily need exact reconstruction of the original achieve a high compression ratio. With the lossy techniques, once the original image is compressed, it can never be fully recovered. However, its use results in a much higher compression ratio than with the lossless methods. Lossy techniques are based on quantization of data (pixel) values [4,7]. Only a small number of bits are used to represent the data value. The minimum number of bits needed for an acceptable image quality varies greatly depending on the application. In medical image compression, the reconstructed image should be as close to the original as is technically possible. This requirement leads to a rather low compression ratio. On the other hand, a high-compression ratio is the top priority in teleconferencing applications, even to the extent of sacrificing image quality [6].

The compression algorithms that have been used in this paper are:

*Huffman coding* is a statistical data compression technique [1]; it's employment will reduce the average code length used to represent the pixel values of an image. The Huffman code is an optimum code since it results in the shortest average code length of all statistical encoding techniques. In addition, Huffman codes have a prefix property which means that no short code group is duplicated as the beginning of a longer group.

*Lempel-Ziv compression* is generally much better than that achieved by Huffman coding, or adaptive Huffman coding and takes less time to compute. The amount of compression obtained depends on the size of the input, the number of bits per code and the distribution of common substrings.

In *predictive coding* the sample to be encoded is predicted from the encoded values of the previously transmitted samples. The error resulting from the subtraction of the prediction from the actual value of the sample is quantized into a set of discrete amplitude levels. These levels are represented as binary words of fixed word length and sent to the channel coder for transmission [4,7,8]. Whenever these codewords are not equiprobable, further lossless compression can be achieved by using variable-length codewords. The first step of the predictive coding is the predictor. The predictor implemented in this paper, estimates the value of the present pixel as the average of three previously transmitted pixels (the previous one in the same line and the two previous ones respectively in the previous line), subtracts it from the original value and transmits the predicted error to the second step. The second step of the predictive coding consists of the quantizer, which is designed so that the quantization error is at the threshold of visibility, while minimizing the number of quantizer levels or the entropy of the quantizer output. The quantizer implemented in this paper is a uniform quantizer.

In *transform coding* an alternative representation of the signal is made first by taking linear combinations of samples in a block of data (called coefficients) and then quantizing the selected coefficients for transmission. In the present paper the entire image is transformed using a 2D-DCT. Then, a bit allocation table designating the number of bits required to store the approximate value of each pixel in this transformed image is produced and a quantized image is formed. The bit allocation table is used to encode the quantized image, forming a one dimensional, sequentially compact data file of the compressed image file. The compression ratio of this file depends on the information in the bit allocation table and the amount of quantization on the transformed image. The compressed image file and bit allocation table are used to reconstruct the image [4,7,9,10,11,12].

In this paper a performance evaluation of the compression techniques has been done.

## 3. RESULTS AND ANALYSIS

The above algorithms have been applied separately to the phantom and to the breast images. Also, combinations of the above algorithms have been applied. The hybrid coding techniques used consist of:

1. Huffman coding
2. Lempel-Ziv coding
3. Predictor - Huffman coding
4. Predictor - Quantizer - Huffman coding
5. Predictor - Lempel-Ziv coding
6. Predictor - Quantizer - Lempel-Ziv coding
7. Discrete Cosine Transform

The parameters that had been taken into account during the interpretation of the images were both objective (Mean Square Error - MSE) and subjective criteria (contrast, resolution, histogram matching, the subtracted image between the original and the reconstructed image, etc.).

The MSE is calculated using the following formula:

$$MSE = \frac{1}{N^2} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} [f(x,y) - f'(x,y)]^2$$

where f(x,y) and f'(x,y) denote pixel values of the original and the reconstructed image at the position (x,y).

The subject assigns an overall quality for the reconstructed images and they are rated

into several categories. In the rating scale method, the radiologists view the reconstructed images and assign each image to one of the several given categories (1: bad, 2: poor , 3: fair, 4: good and 5: excellent). h Figure 1 the compression ratio versus the MSE for the phantom image is presented. In Figure 2 the compression ratio versus the subjective ratio for the phantom image is presented. Employing a DCT with a compression . ratio of up to 62%, the reconstructed image is characterized as excellent (category 5), between 62% and 83% as good, 83% and 93% as fair and after that as not acceptable. Among the rest of the



Figure 1

MSE vesus compression ratio (Phantom Image)

techniques, Lempel Ziv yields the best results with 70.2% compression ratio and 0 MSE, followed by the Predictor-Lempel Ziv with a 62% ratio and 0 MSE. It must be noted, that although the MSE in Predictor-Quantizer-Lempel Ziv is 1245 the reconstructed image can be characterized as fair.
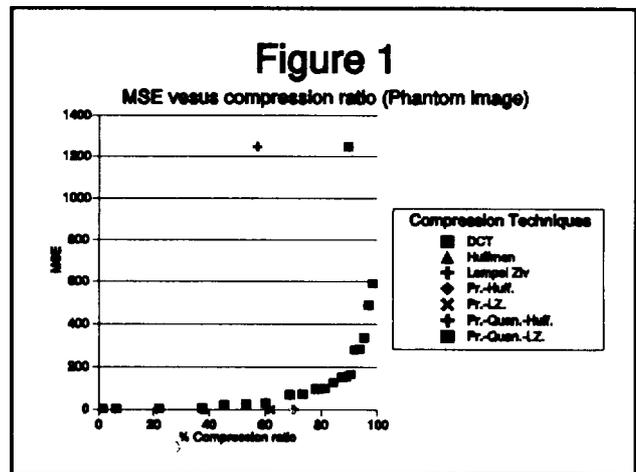
In Figure 3 the compression ratio versus the MSE for the breast image is presented. In Figure 4 the compression ratio versus the subjective ratio for the breast image is presented. In this case, the DCT with a compression ratio of up to 87.5%, the reconstructed image is ·characterized as excellent (category 5), between 87.5% and 92% as good, 92% and 95.5% as fair and after that as not acceptable. Among the rest of the techniques, Lempel Ziv yields the best results with a 64.6% compression ratio and 0 MSE, followed by the Predictor-Lempel Ziv with a 63.4% ratio and 0 MSE. The Predictor-



Figure 2

Subjective versus compression ratio (Phantom Image)

Quantizer-Lempel Ziv has an MSE of 98 and the reconstructed image can be characterized as good.
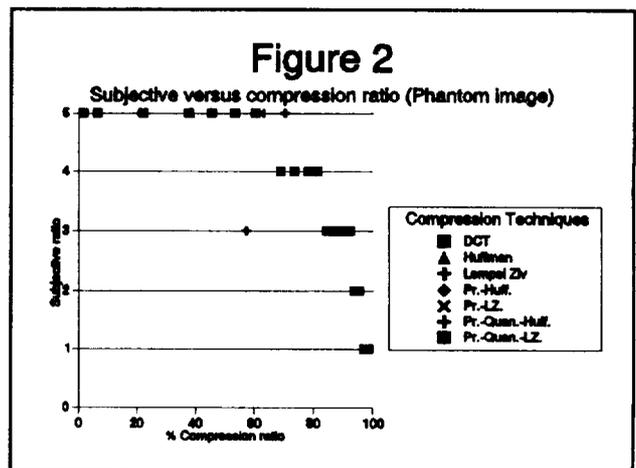
Comparing the above results, it is obvious that in breast images a better compression ratio of 15-20% than phantom images can be achieved and keeping the same quality level.

## 4. CONCLUSIONS

In this paper a diversity of individual or combined coding methods have been applied in both phantom and real breast medical images and the reconstructed images have been compared. The problem of information loss is directly coupled with the diagnostic value of the image. However, the use of different compression techniques, resulting reconstructed medical images with similar quality, could facilitate the data bank storage and transmission processes of the radiographic images, because of the different execution times of each technique. This work will be continued through the analysis of an extended gama of compression techniques and the recording the processing times of each one in order to create an efficient compression

service communication element.

## 5. REFERENCES

[1] Gilb. Held, "Data Compression", chapter 2, Wiley, Chichester, New York, 1989.

[2] G. Panayiotakis, E. Giakoumakis, A. Flioni, S. Xenofos and B. Proimos, "Quality control of mammographic units in Greece", Proc. VI Mediterranean Conference on Medical and Biological Engineering, pp. 623-626, 5-10 July, Capri, Italy, 1992.

[3] D. Lymberopoulos, S. Kotsopoulos, and G. Kokkinakis, "A Prototype Medical Workstation with Unified Image Viewing Facilities", Journal of Electrical and Electronics Engineering, Australia - IE Aust. & IREEE Aust., Vol. 12, No. 1, March 1992.

[4] A. Netravali, and J. Limb, "Picture Coding: A Review", Proc. IEEE, vol. 68, no. 3, pp. 366-406, Mar. 1980.

[5] G. Anastassopoulos, D. Lymberopoulos, S. Kotsopoulos and G. Kokkinakis, "A new image compression scheme for digital angiocardiographyapplication",SPIE's Medical Imaging 1993, vol. 1897, pp. 461-468, Feb. 14-18, California, USA, 1993.

[6] G. Anastassopoulos, D. Lymberopoulos and G. Kokkinakis, "Image Processing on Compressed Data for Medical Purpose", SPIE's Medical Imaging 1994, Feb. 13-18, California, USA, 1994.

[7] A. Jain, "Image Data Compression: A Review", Proc. IEEE, vol. 69, no. 3, pp. 349-389, Mar. 1981.

[8] K. Rose and A. Heiman, "Enhancement of One - Dimensional Variable-Length DPCM Images Corrupted by Transmission Errors", IEEE Transactions on Communications, vol. 37, no. 4, April 1989.

[9] N. Ahmed, T. Natarajan and K. R. Rao, "Discrete Cosine Transform", IEEE Transactions of Computers, January 1974.

[10] J. Makhoul, "A Fast Cosine Transform in one and two dimensions", IEEE Transactions on Acoustics, Speech and Signal Processing, vol. ASSP-28, no. 1, February 1980.

[11] L. Chung and HK. Huang, "Radiological Image Compression: Full-Frame Bit-Allocation Technique", Radiology, Vol. 133, No.3, pp. 811-817, Nov. 1985.

[12] C. Christopoulos, W. Philips, A. Skodras and J. Cornelis, "Discrete Cosine Transform Coding of Images", Proceedings of the International Conference on Digital Signal Processing and II International Conference on Computer Application to Engineering System, pp. 164-169, July 14-16, Nicosia, Cyprus, 1993.

Figure 3
MSE versus Compression ratio (breast image)



Figure 4
Subjective versus Compression ratio (breast image)
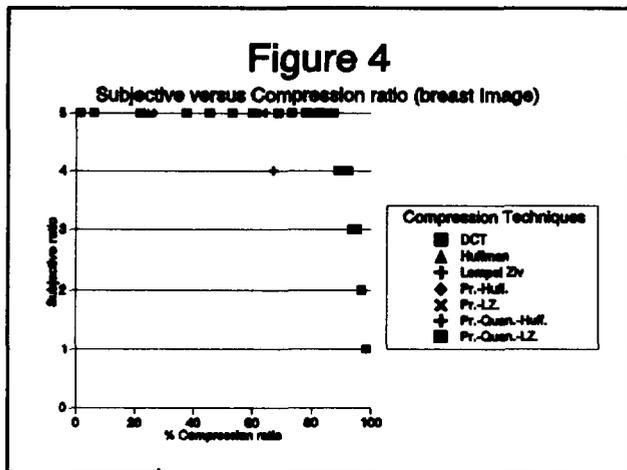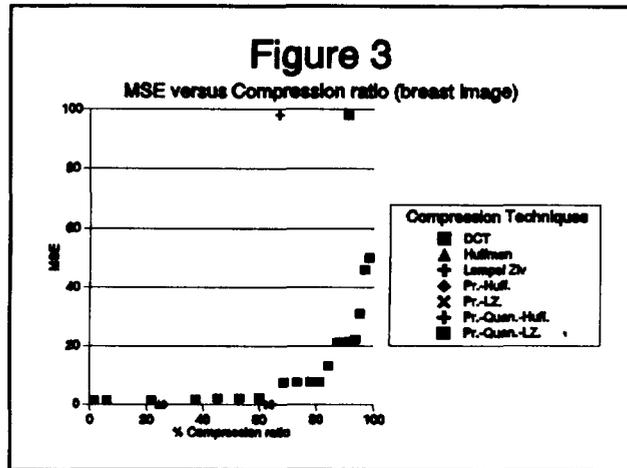
# Nonlinear Time Series Analysis of
## Resting, Moderate and High Exercise Electrocardiograms

A. Bezerianos[1], T. Bountis[2], G. Papaioannou[2], P. Polydoropoulos[1]

1 Department of Medicine , Section of Medical Physics
2 Department of Mathematics, Section of Applied Analysis
Univerity of Patras, 261 10 Patras, Greece

## ABSTRACT

In recent years, there has been an increasing number of papers in the literature, applying the methods and techniques of Nonlinear Dynamics to the ECG of various human subjects. Most of them are based primarily on correlation dimension calculation and conclude that the dynamics of ECG signal is deterministic and occurs on a chaotic attractor, whose dimension can distinguish between healthy and severely malfunctioning cases. In this paper we demonstrate that correlation dimension calculations , while qualitatively useful in revealing the presence of a deterministic attractor, must be used with care, as they do not always yield reliable estimates of the attractor's dimension. We suggest that further tests are needed to provide a more complete description of the dynamics. We conclude that a detailed investigation, using as many of the available techniques of nonlinear time series analysis is needed to investigate more deeply into dynamical properties of ECG signals and hopefully provide a more sensitive diagnostic tool, than has been achieved with other more traditional approaches.

## I. INTRODUCTION

As it is well-known the usual ECG records the electrical activity of a large mass of atrial and ventricular cells. Since cardiac depolarization and repolarizasion normally occur in synchronized fashion, the ECG is able to record these electrical activities in the form of certain well-defined waves (Fig. 1).

The human ECG, even in the case of resting subjects, displays a considerable amount of fluctuations in time (the so-called heart rate variability), as well as in amplitude.It has been shown that conventional statistical measures of these variabilities have significant prognostic implications in cases of myocardial infractions and are useful for detecting arrythmias .

In recent years, considerable attention has been devoted to unifying various aspects of cardiac pathophysiology using nonlinear dynamics, particularly through applications of chaos theory and concepts of fractal geometry. The main efforts, so far, have concentrated on the analysis of QRS complex and heart rate variability and not so much on the full ECG time series .

In this paper, we have used the methods and techniques of nonlinear time series analysis to compute the correlation dimension of the ECG time series of 5 male subjects, 3 of whom are healthy (ages 22,23 and 30) and 2 heavy smokers (35 and 38 years old). It suggests that the dynamics of the full ECG signal has a strong deterministic component and occurs on a chaotic attractor of dimension as low as 2-3.5. The time-derivative of the signal howover showed that the signals contain a significant amount of "noise" (or randomness) in the high frequency part of the power spectrum. In all cases studied, this

time differentiation series were found to have a significantly higher dimension than the original ECG, filled out all the embedding space it is placed in.The latter was confirmed also by Theiler' s test, the results of which showed that low dimensionality of ECG is most likely due to temporal correlations. Finally Singular Value Decomposition (SVD) method was applied for noise elimination and the results were in good agreement with the low correlation dimension estimates of the attractor.

In this paper, we shall also examine the relation of various parameters (autocorrelation factor, embedding dimension, length of time series analyzed ) with the correlation dimension and the underlying physiological process in the phases of ECG studied.
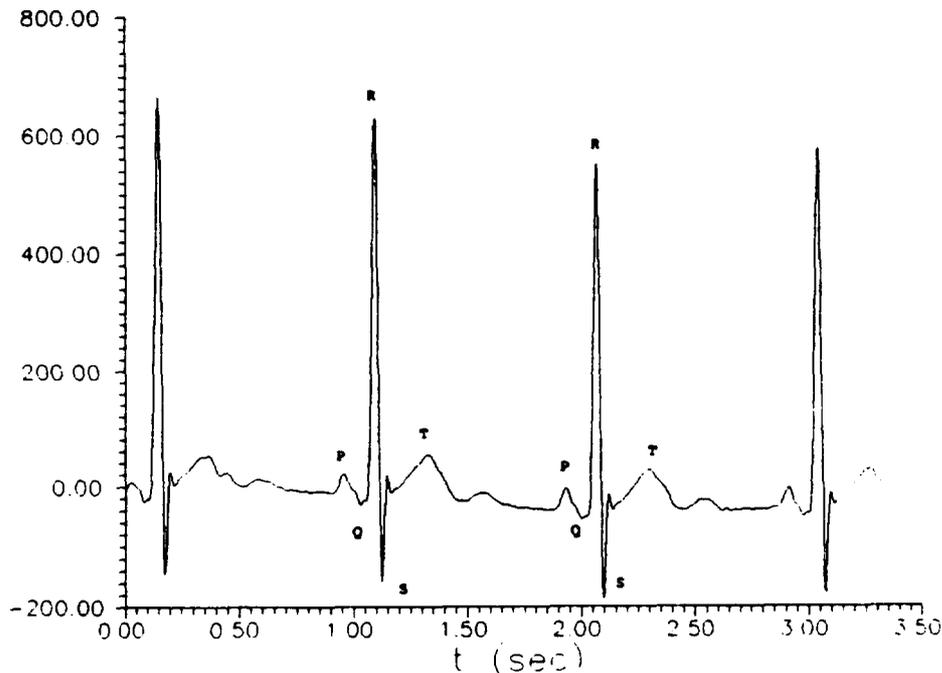


**Fig. 1   ECG of a healthy person (rkollman)**

## II. MATERIALS AND METHODS

The cases studied in this research consisted of 3 healthy young men of ages 22, 23 and 30 and of 2 heavy smokers, 35 and 38 years old. Their ECG signals were recorded at the clinical ECG laboratory of the University of Patras Hospital, under M.D. monitoring. No evidence of cardiovascular disease was observed in any of our 5 cases, as assessed by history, clinical examination and resting and exercise tests conducted by 12-lead ECG recordings.One ECG lead, with clearly visible P and T waves and QRS complexes of large amplitude (corresponding to lead V ) was recorded continuously for 10 to 15 min, in supine (resting) and upright (exercise) positions. The exercise test was carried out with a treadmill speed of 25 km/hr and 10% inclination and the high exercise test speed was carried out with a treadmill speed of 7.5 km/hr and 20% inclination, while the recording period started after the heart rate and blood pressure of the subject had been stabilized.The ECGs were amplified using a CASE 15 exercise system (Marquette, OH) with standard amplification and filters set between 0.04 Hz and 80 Hz. The signals were digitized on line (12 bit Metrabyte DAS 16F, ADC board) at a rate of 300 Hz and stored on hard disc for further processing. Each resulting file of 800-1000 heart beats finally consisted of 10 blocks of 22000 data points.For our calculations we used the first block of these files.

The procedure in order to calculate the correlation dimension is as follow:

From the delayed copies matrix $a_k$=[ $X(t_k)$, $X(t_k+\tau)$, $X(t_k+2\tau)$,....,$X(t_k+(n-1)\tau)$] of the input signal $X(t)$, where $\tau$ is the delay time and n is the embedding dimension, the correlation integral

$$c(r) = \frac{1}{N^2} \sum_{k,j=1}^{N} H(r - |a_k - a_j|), \quad k \neq j$$

is computed. r is the radius of the located hyperspheres, $|a_k-a_j|$ is the Euclidean distance between the vectors $a_k$ and $a_j$, N is the total number of elements of the signal and H is the Heaviside function. The correlation dimension value is the slope of the line logC(r) versus log(r) ( d [ log C(r)]/ d (log r) ).[1]

The time differentiation series $\Delta y(t_i)=x(t_i)-x(t_{i-1})$ where X is the input signal and the randomized phase time series G(t) which is the input signal with random phase, were studied. The G(t) time series was calculated as follows :

The FFT of the input signal X(t) was found and the phases were changed in random way. The reconstructed signal is G(t). [6]

The Singular Value Decomposition method (SVD) showed that the ECG signal is strongly deterministic and the results were in good agreement with the correlation dimension of the attractor.[7]

## III. RESULTS AND DISCUSSION

Fig 2. shows a typical power spectrum [2,3] and the autocorrelation [4,5] of a healthy subject . The spectrum shows a negative slope in log log scale and the autocorrelation shows that the ECG is a strongly periodical signal.
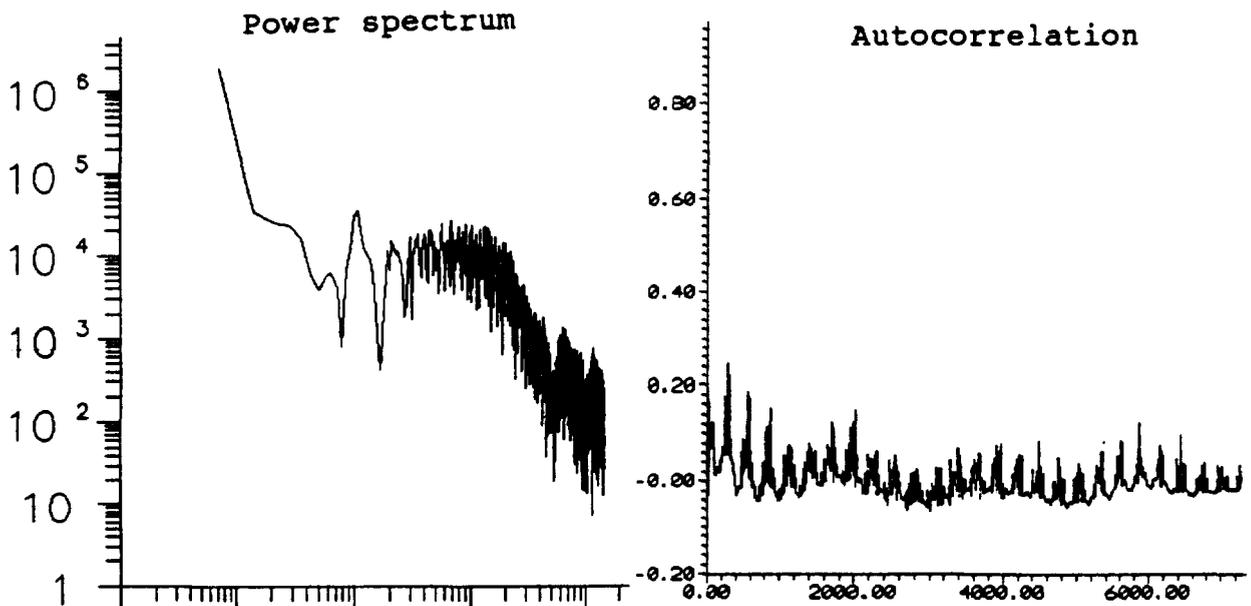


Fig. 2 Power spectrum and autocorrelation of a healthy pesron (rkollman)

In Fig 3 two pictures are showed. The first is the correlation dimension versus embedding dimension of a healthy ECG and the second is the slope of the line, log C(r) versus log r , versus log r for several embedding dimensions. It can be observed that when the embedding dimension is greater than six the correlation dimension remains the same, near 3. Seven was chosen as the best embedding dimension for our signals.In

order to calculate the embedding dimension we used delay time 100 and 22000 points of our signal.
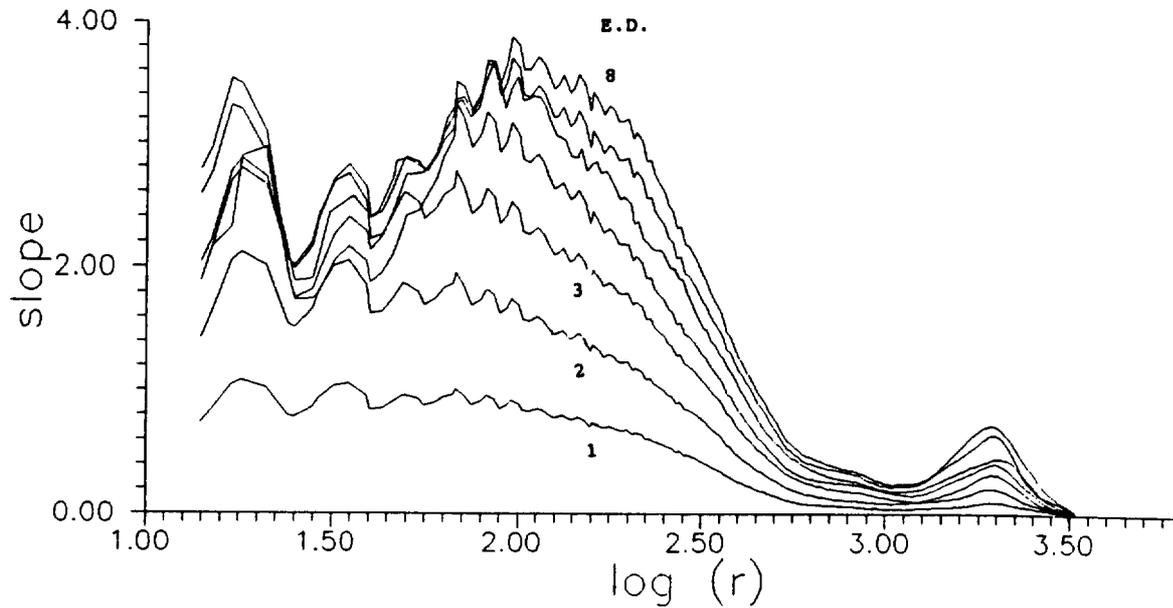


**Fig. 3.** *Correlation dimension versus embedding dimension*

Table 1 shows the correlation dimension for several subjects. In order to calculate correlation dimension we used 22000 points and delay time 100 for all the cases.

**Table 1.** *Correlation dimension for several subjects*

| Case | Files | Correlation Dimension of Input Signal | Corr. Dimens. of Time Differentiation Signal | Correlation Dimension of Random. Phase Signal |
|---|---|---|---|---|
| Resting,Healthy | Rkollman | 3.04 ± 0.626 | 3.79 ± 0.078 | 6.56 ± 0.905 |
| Exers.,Healthy | Ekollman | 3.58 ± 0.846 | 5.68 ± 0.518 | 6.50 ± 0.432 |
| Resting,Healthy | Rspyrosa | 2.08 ± 0.704 | 4.21 ± 0.404 | 6.58 ± 0.305 |
| Exers.,Healthy | Espyrosa | 3.01 ± 0.877 | 5.21 ± 0.776 | 6.65 ± 0.662 |
| Resting,Healthy | Rpol | 2.63 ± 0.640 | 4.23 ± 0.549 | 6.75 ± 0.525 |
| Exers.,Healthy | Epol | 2.79 ± 0.396 | 6.14 ± 0.743 | 6.22 ± 0.635 |
| Resting,Smoker | Rgiana | 2.74 ± 0.292 | 5.31 ± 0.348 | 6.01 ± 0.834 |
| Exers.,Smoker | Egiana | 2.71 ± 0.725 | 4.81 ± 0.484 | 6.58 ± 0.924 |
| Resting,Smoker | Rmara | 2.21 ± 0.503 | 4.71 ± 0.435 | 6.58 ± 0.505 |
| Exers.,Smoker | Emara | 3.49 ± 0.452 | 5.91 ± 0.478 | 6.81 ± 0.642 |
| High Exersise Healthy | Fkollman | 2.16 ± 0.572 | 3.44 ± 0.466 | 6.68 ± 0.472 |

It can be seen the original ECG signal was found to have a correlation dimension between 2-3.5. If we look at the power spectrum we can see that the signals have a significant amount of "noise" in high frequency. The time differentiation signal contains the high frequency fluctuations of the data and it can be seen that time differentiation series have a significantly higher dimension than the original signal between 3-7,"filling out " practically all of the embedding dimension space it is placed in,just as ordinary noise

does. The randomize phase time series was found to have a correlation dimension between 6-7.(Table 1)

More over , Theiler's test of eliminating dynamically neighboring points in dimension calculation showed that the apparent low-dimensionality of ECGs at small scales, is most likely due to temporal rather than spatial correlations between the points (Fig. 4).

Fig. 4 shows the result of the Singular Value Decomposition (SVD) method on signal Rkollman. It can been seen that this file was found to have 3 basic eigenvalues which is in good agreement with the low correlation dimension of the attractor which is 3.04. The SVD method was applied for all our signals and all the results supported the low estimates found for the correlation dimension.
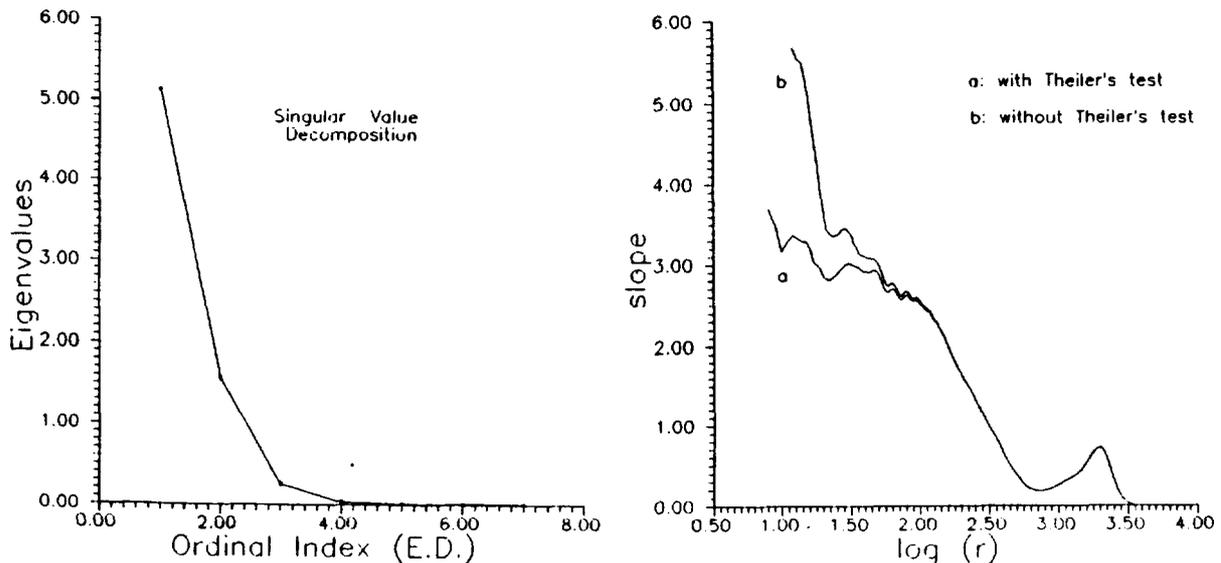


**Fig. 4  SVD method and Teiler's test**

## REFERENCES

[1] J.D. Farmer,  E Ott, J. A. Yorke: The dimension of chaotic attractors,  Physica D7, p. 153, 1983

[2] A. L. Goldberger, M. D. Bruse, J. West: Fractals in Physiology and Medicine,  The Yale    Journal of Biology and Medicine, Vol. 60, p.p. 421-435,1987

[3] W.H. Press, S. A. Teukkolsky, B.P. Flannery, Wi T. Vetterling: Recipes in C Cambridge University Press, C Edition,1992

[4] A. Babloyantz, A. Destexhe: Is the Normal Heart a Periodic Oscillation ? Biological Cybernetic, Vol 58, p.p. 203-211, 1988.

[5] A.L. Goldberger, B.J. West : Applications of nonlinear dynamics of clinical cardiology, Annals New York Academy of Sciences 1986.

[6] Broomhead D.S. and G.P. King, Extracting qualitative dynamics from    experimental data, Physica 20D, 217-236,1986

[7]Theiler J., Spurious dimension from correlation algorithms applied to limited    time-series data,Phys. Rev. A34,2427,1986

# Digital Signal Processor based on the Field Programmable Gate Array (FPGA) for the detection QRS complex on ECG signal in real-time environment

## Jalal SABOR, Miloud HASSANI, Michel HUBIN

CRITT-GBM, LCIA-La3i, INSA de Rouen, BP 08 F-76131 Mont Saint Aignan Cedex
France

## Abstract

Detectors of the *QRS* complex actually used in the electrocardiogram monitoring propose pure numeric structures based on microprocessors. The recognition of the *QRS* by these methods is far from being a real-time process.
In fact, we decided to develop a pure logic Digital Signal Processor based on the Field Programmable Gate Array (FPGA). This recent technology of ASICs permits an infinite reprogramming of DSP and give a high speed operation going till 250 Mhz. This DSP detects the *QRS* complex of the ECG signal precisely and calculates the average cardiac frequency for such recognition of bradycardia or tachycardia.
This hardware approach gives us the possibility to conceive very fast miniaturized systems for rigorous analysis of electrocardiogram signal.

Keywords : Processor, Detector, *QRS* complex, Electrocadiogram, FPGA, real-time.

## 1.Introduction

Actually, the recognition of the *QRS* complex in the ambulatory monitors is made either by analog methods [1], [2], [3], either by numeric methods [4], [5], [6] based on microprocessor structures in real-time environment.
In fact, the detection of *QRS* is difficult, not only because of the physiological variability of the complex *QRS*, but also because of the various types of noise that can be present in the ECG signal.
Noise sources include muscle noise, artifacts due to electrode motion, power-line interference and baseline wander (see figure 2).
In our approach, the band pass analog filter reduces the influence of these noise sources, and thereby improves the signal-to-noise ratio. Of the many *QRS* detectors proposed in the literature, few give serious enough attention to noise reduction.
Our DSP complex *QRS* detector is developped by the pure logic method based on the Field Programmable Gate Array ( FPGA of XILINX ). With this recent technology of ASICs, the DSP is operational in real-time environment.
The DSP is based on two important modules during the processing (see Figure 1): the ECG signal acquisition module and *QRS* detection module.This last module calculates the average cardiac frequency for such recognition of bradycardia or tachydycardia.

Fig.1. DSP organisation

## 2. ECG signal acquisition

The database of DSP is provided by one ECG signal channel. This signal is recorded in our laboratory using the EINTHOVEN D1 bipolar derivation.

Before the numeric conversion, the ECG is filtered with an analog band pass filter between 2 Hz and 15 Hz [3] to attenuate the components included in the respiratory spectral domain. This filter lets only pass the components of the ECG spectral domain (see figure2).

In the first step, The DSP activates the numeric analog conversion the ECG signal, in generatin all necessary commands signals including the sample rate (40 Hz). In addition, with the interruption controller unit, the DSP can manage simultaneously the acquisition of 256 signals at different sample rates.

In the second step, the DSP recuperates the data, to stock it in external memory, and also to transmit it to the *QRS* complex detection module.
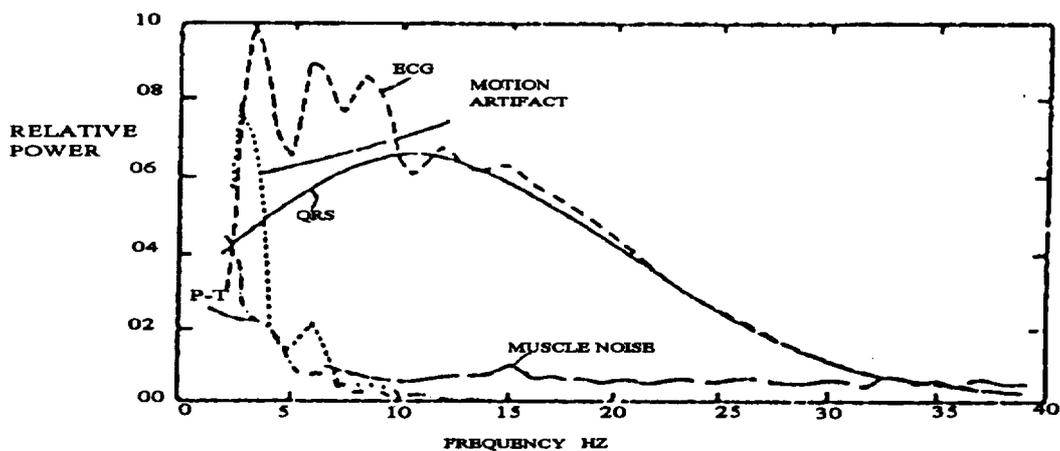


Fig.2. Typical spectra power of the ECG

208

# 3. QRS complex detection

Our algorithm is based on the analysis of differences [6]. Moreover, the analysis of the first and second derivation of the ECG permits to recognize the *QRS* complex. The detection threshold is autoadjustable, it is calculated regularly after the detection of four succesive *QRS*. With this technique, the algorithm is always adaptative with respect to the variations of the *QRS* amplitude.

The synoptic of the algorithm is given in figure 3
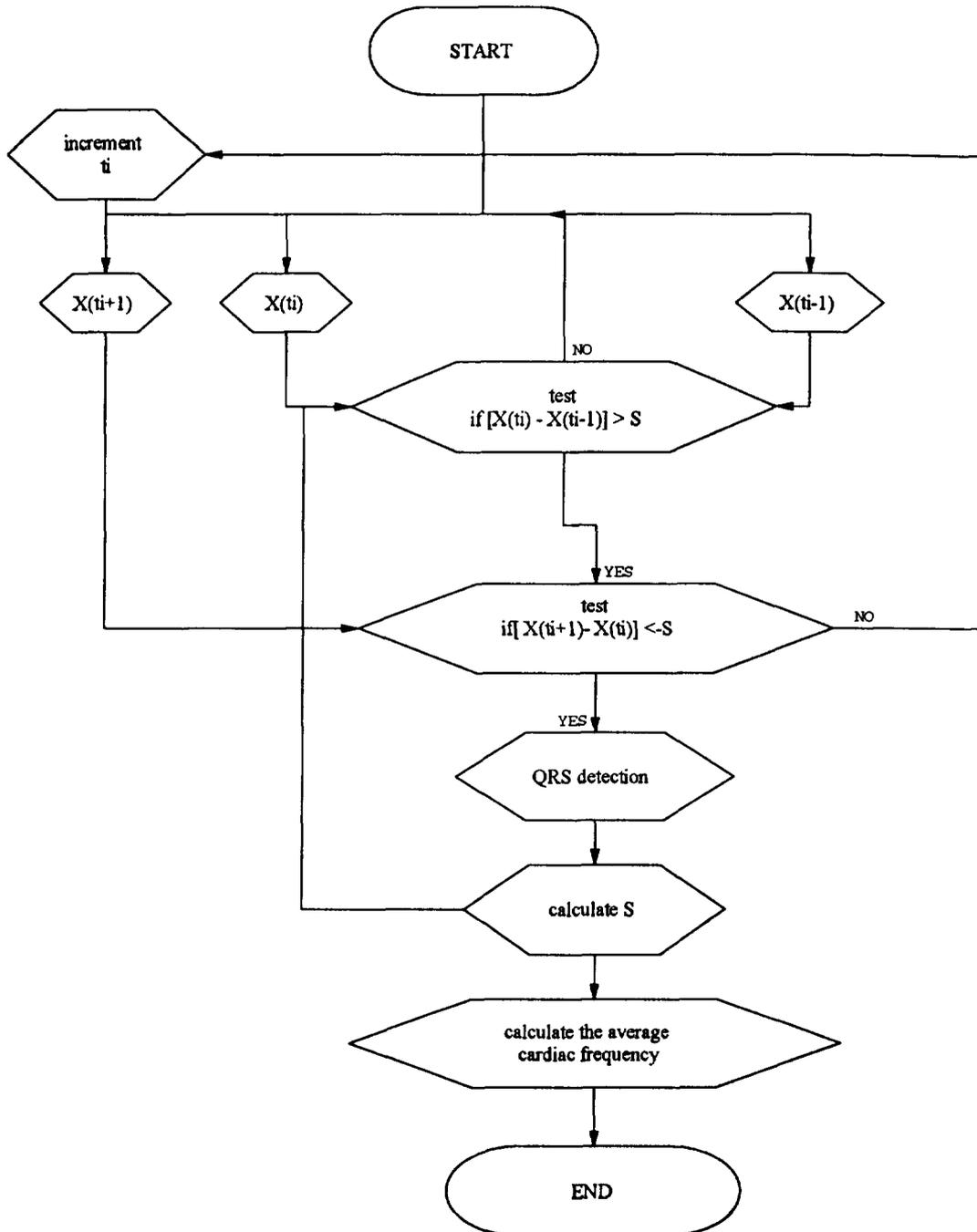


Fig.3. QRS detection algorithm

x(ti): sample of ECG signal at instant ti.

After the detection of N *QRS* complexes in a period of time T, the DSP can calculate the average cardiac frequency in beatings per minute by the following formula:

$$FC = 60*N/T.$$

This algorithm has been tested firstly by the MATLAB software with same database, and after having found satisfying results, we have decided to implement it in an ASIC based on FPGA (xc4020-200Mhz XILINX family). We have used only 60% of the Combinatory Logic Blocs (CLBs) of the FPGA, this gives us a large possibility to improve our algorithm in the same DSP, by means of upgraded theories. The logic simulation has been made by the VIEWSIM software.

## 4. Results

Figure 3 shows the different steps of the *QRS* complex detection.



Fig. 4.
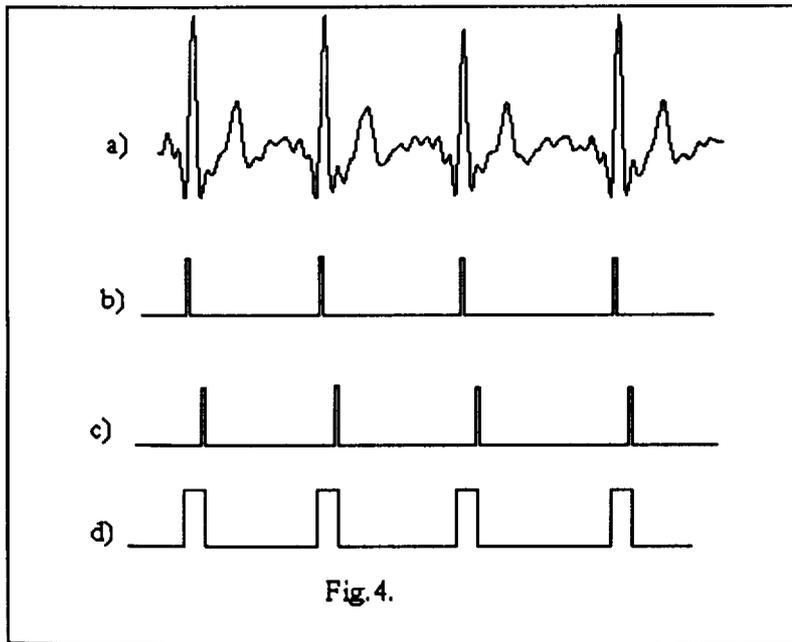
a) the ECG signal prefiltered analogically between 2 Hz and 15 Hz.
b) detection of the positive slope of the *QRS* complex.
c) detection of the negative slope of the *QRS* complex.
d) detection of the *QRS* complex.

## 5. Conclusion

With the recent technology of Field Programmable Gate Array (FPGA), we have implemented a logic DSP for *QRS* complex detection on ECG signal in real-time. In fact, this DSP is designed in order to be used in miniaturized ambulatory monitors, since it is the best and more precise and reliable way to identify in real-time bradycardies and tachycardies. Particularly, we will integrate this DSP in a SID (Sudden Infant Death) helping research monitor [7].

# References

[1]  J. Framden, M. R Neuman, *"QRS Wave detection"*, Med. Biol. Eng. Comput.,May, 1980, pp 125-132.

[2]  N. V Thakor, J Webster, W Tompkins, *"Optimal QRS Detector"*, Med. Biol. Eng. Comput., May 1983, pp 343-350.

[3]  O.Y. Vel, *"R-Wave Detection in presence of Muscle Artifacts"*, IEEE Trans. Biomed. Eng., vol. 31, No.11, November1984, pp 715-717.

[4]  Pan Jiapu, W. Tompkins, *"A Real Time QRS Detection Algorithm"*, IEEE. Trans. Biomed. Eng., vol. 32, March 1985, pp 230-236.

[5]  D. Forbres, H. Jimson, *"A QRS Detection Algorithm"*, Journal of Clinical Monitoring, Vol. 3-1, Janvier 1987, pp 54-63.

[6]  R. Armington, F. Geheb, *"Optimizing The Utiliziation of Multi-Lead ECG Data for Accurate QRS Detection"*, IEEE Annual Conf. of the Eng. in Medecine, 1987, pp 259- 262.

[7]  M. Hubin, J. Sabor, M. Hassani, *"Multisensor Intelligent System for Infant Monitoring"*, Proc. Int. Conf. on Sensor, Nuremberg, Vol. 3, November 1993, pp 179 - 185.

[8]  N .V Thakor, Yi Sheng zhu, *"Application of adaptive filtering to ECG analysis: Noise cancellation and arrythmia detection"*, IEEE Trans. Biomed. Eng., Vol. 38, No. 8, August 1991, pp 785-794.

[9]  R. A Frankel, Erik. W Pottala, R. W Bower, James. J Baily, *"A filter to suppress ECG baseline wander and preserve st-segemnt accuracy in a real-time environment"*, Journal of Electorcardiology, Vol. 24, No. 4, October 1991, pp 315-323.

# Pattern recognition of ECG waves forms using wavelet transform: experimental results in the sudden infant death syndrome monitoring.

M. HASSANI, I. FACHRUDIN, M. HUBIN

CRITT-GBM, LCIA-La3i, INSA de Rouen, BP 08 F-76131 Mont Saint Aignan Cedex, France

**Abstract** - Until now, many studies taken about physiological signal analysis had not lead to any explanation about sudden infant death syndrome. According to previous statistics studies, the fundamental hypothesis formulated by the practicians concerns the cardiac rhythm disorder, as long QT syndrome and large or fine QRS wave (supra ventricular tachycardia) .

For that reason, different studies about ECG signal have been done. Nevertheless, it's to be proved that classical signal analysis is not sufficient because the most algorithm developed earlier detect neither QRS duration nor T wave.

Time-frequency analysis is proposed to establish the time evolution of all the ECG signal details. This paper describes a method to detect different waves as QRS and T of ECG signal observed during infant monitoring by using wavelet transform theory. A qualitative analysis of different waves as QRS and T extracted is made, and a correlative study is done after the use of different levels of signal decomposition for pattern detection, in order to recognize different tachycardia and long QT syndrome occurred to the infant in SIDS.

**Key Words** - Wavelet, Filters Banks, SIDS, ECG, QRS-T Detection.

## I. INTRODUCTION:

The widespread interest in ambulatory monitoring of the ECG , especially in the case of infant risk of sudden death, has prompted the development of software QRS, RR, and QT detection. Some authors [7],[9],[10] had developed algorithms to detect the QRS waves.

The demands of the ECG waves detector performance in ambulatory home monitoring are high [1]. Since a recording typically lasts 24h, it is crucial to design computationally feasible algorithm in order to arrive at reasonable overall processing time of the analysis system.
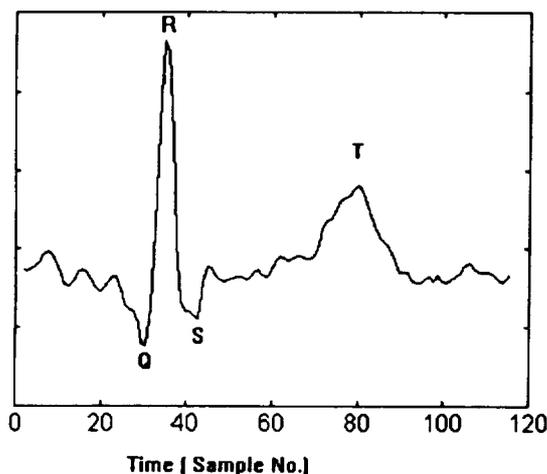


**Fig.1. Original ECG signal with QRS and T complexs**

To watch the risk infants, we propose an optimal QRS and T waves detection algorithm based on the wavelet transform. Its analysis leads to a recognition of pathologic events linked with cardiac disfonctionment ( supra ventricular tachycardia TSV, catecholergic ventricular tachycardia and long QT syndrome).

A new detection method of QRS and T complex waves based on the wavelet transform is presented. We know not only that the wavelet transform can be interpreted as the band pass filter action for the signal [8], but also that in the power spectral density of the QRS and T in the ECG complex, the QRS and T wave are localised in different frequency bands. So we can built the local optimal band pass filter with some filters banks which are obtained using scale parameter variation of the wavelet, for each frequency band.

In [7] the authors proposed a real time algorithm QRS detection, based on the filtering, derivation, squaring and moving window integration. The new idea is to improve this algorithm in order to determine the width of the QRS complex. Instead of using the cascaded high pass and low pass filter , we use local optimal band pass filter built with wavelet transform.

In section II , we will first recall the theory of wavelet transform. We will give in section III the principle of the algorithm proposed. In section VI, we will show the results of QRS and T waves detection, and the results of width QRS wave calculations. Lastly, we will give the prospect of an implementation possibility of this method in a "STAR Corp." DSP which will become integrated

into a real time signal processing chain.

## II. WAVELET TRANSFORM PRESENTATION:

Let give a time varying signal s(t) and ψ(t) with finite energy. The wavelet transform consists in computing coefficient that are inner products of signal and a family of "wavelet". In continuous domain, the wavelet corresponding to scale $a$ and time location $b$ is

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi(\frac{t-b}{a})$$

$$(1)$$

where ψ(t) is the wavelet "prototype", which can be considerate as a band pass function. The factor $|a|^{-1/2}$ is used to ensure energy preservation.

The Formalism of the continuous wavelet transform was first introduced by [2]. Time $b$ and scale $a$ parameters vary continuously:

$$S_{a,b}(t) = \int_{0}^{\infty} s(t).\psi^*_{a,b}(t)dt$$

$$(2)$$

(the * stands for complex conjugate)

If s(t) and ψ(t) were discretized with one period T, the (2) equation becomes:

$$S(a,kT) = \frac{T}{\sqrt{a}} \sum_{n} s(nT).\psi^* (\frac{(n-k)T}{a})$$

$$(3)$$
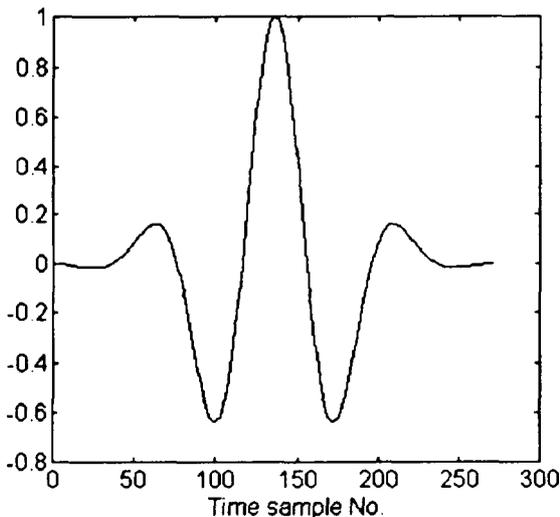


Fig. 2. Wavelet Morlet with a=2,α =2 and c=2π

In our case we use the Morlet wavelet prototype:

$$\psi_1(t) = e^{ict} e^{(\frac{-\alpha^2 t^2}{2})}$$

$$(4)$$

using only real-value of (4) (Fig. 2)

$$\psi_2(t) = \cos(ct) e^{(\frac{-\alpha^2 t^2}{2})}$$

$$(5)$$

How the wavelet transform is assimilated as a zooming function of the signal s(t) around the time location b, with increase in size of a,

- if s(t) is practically constant at the scale a , around the time location b, the s and ψ2(t) product (with null surface) are too weak, then S(a,b) is also weak.
- when around the time location b, s(t) presents variations which coincident with the one of ψ2(t/a), then the s(t) by ψ2(t) product becomes important, and also the wavelet coefficient.

The wavelet used must satisfy some conditions :

- times continuous, absolutely integrable and square-integrable (finite energy)
- analytic: the Fourier transform ψ(ω) must be equal to 0 for ω <0,
- admissible: it means that

$$\int_{0}^{+\infty} \frac{|\psi(\omega)|^2}{\omega} d\omega = \int_{-\infty}^{+\infty} \frac{|\psi(t)|^2}{|\omega|} d\omega = C_\psi \langle +\infty$$

$$(6)$$

this condition implies that

$$\int_{-\infty}^{+\infty} \psi(u)du = 0$$

In case of the Morlet Wavelet :

$$\psi_1(\omega) = \frac{1}{\alpha} \sqrt{2\pi}.\exp\left[\frac{-(\omega-c)^2}{2\alpha^2}\right]$$

$$(7)$$

we obtain :

$$\int_{-\infty}^{+\infty} \psi_1(t)dt = \frac{1}{\alpha} \sqrt{2\pi}.\exp\left[\frac{-c^2}{2\alpha^2}\right]$$

$$(8)$$

The value obtained for c and α used ( c=2π and α=2) is $0.9x10^{-3}$. The wavelet ψ1(t) does not give satisfaction about admissibility, it would introduce difficulties if we are interested in the signal reconstruction, but it is not our case. Nevertheless, it permits a good analysis of the ECG signal.

## III. PRESENTATION OF THE ALGORITHM

### III.1 QRS AND T WAVES DETECTION ALGORITHM

As the spectral analysis of the ECG signal shows that there are several harmonics (Fig. 3), we

remark that the QRS wave spectral varies from 0 to 30Hz, with the maximum around f=15Hz whereas the T wave spectrum varies from 0 to 10Hz.

We also show that only in the frequency band 0-5Hz, the T wave spectrum density is higher than the corresponding QRS wave.
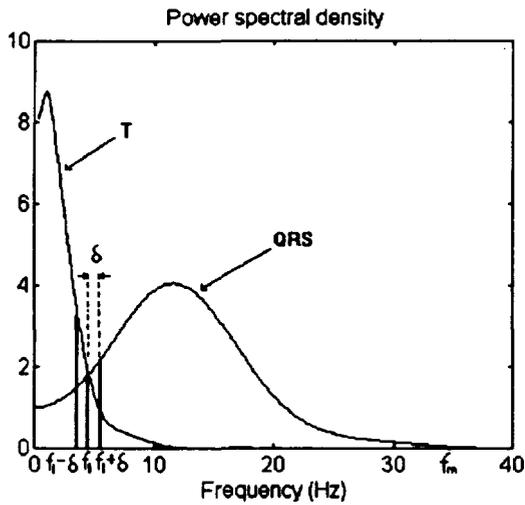


**Fig. 3. Typical power spectral density of QRS and T complexs**

We know that the wavelet transform acts as a band pass filter. The Power spectral density of Morlet wavelet in the fig.4 illustrates the characteristics of this filter. In this case, for a given value of a, $S(a,b)$ will be maximum when the signal contains one harmonic at the frequency $f=c/2\pi a$ in time location value b.
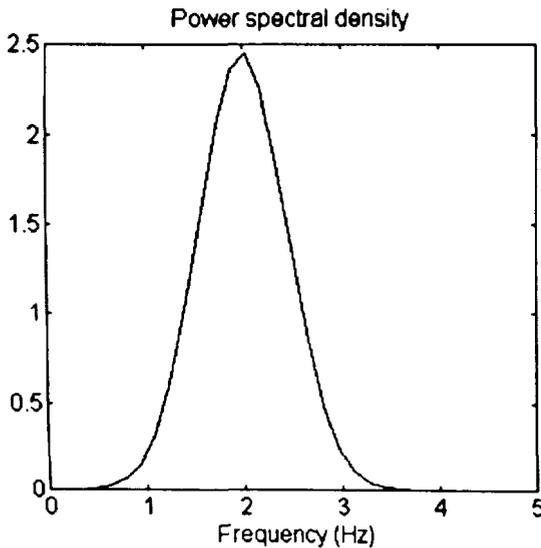


**Fig. 4. Power spectral density of Wavelet in Fig. 2.**

Basing on these establishment, we built the local optimal band pass filter with some filters banks

which are obtained using wavelet scale parameter variation .

For each frequency band belonging to $D=\{D_1=(0,f1-\delta),D_2=(f1-\delta,f1+\delta),D_3=(f1+\delta,fm)\}$ we choose several value of a such that with the given relation $f=c/2\pi a$, $f \in D_i$. And then we compute the wavelet transform for these values of a.
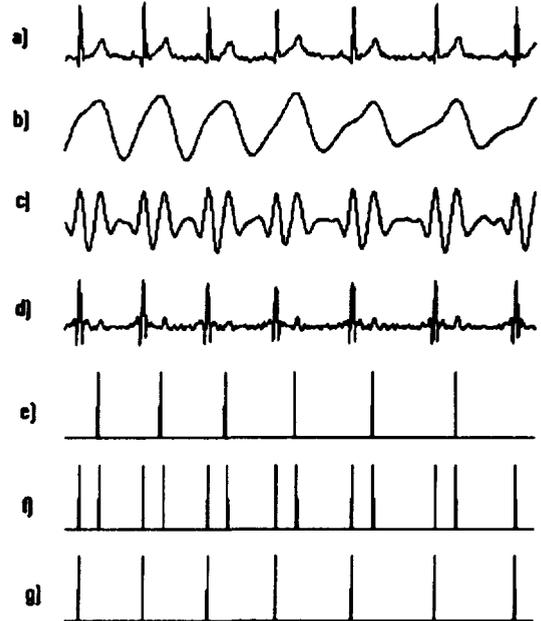


**Fig. 5. QRS and T detection algorithm. a) Original ECG signal. b) ST. c) SQT. d) SQ. e) Output pulse of ST maximum local detection. f) Output pulse of SQT maximum local detection. g) Output pulse of SQ maximum local detection.**

The various steps of algorithm are shown in fig. We can resume this algorithm as follows:

1) determinate the $f_1$ (limit frequency) and $\delta$ (Fig. 3).
2) compute the wavelet transform $S(a,b)$ for different values of a such as $f=c/2\pi a \in D_i$.
3) compute:
   i) the meaning of $S(a_i,b)$ for $f=c/2\pi a_i \in D1$ for each value of b.
      $ST=mean(S(a_i,b))$
   ii) the maximum of $S(a_i,b)$ for $f=c/2\pi a_i \in D2$, for each value of b
      $SQT=max(S(a_i,b))$
   iii) the meaning of $S(a_i,b)$ for $f=c/2\pi a_i \in D3$ for each value of b
      $SQ=mean (S(a_i,b))$.
4) search the peaks of ST, SQT and SQ.
The peaks are such:
- ST correspond to the T wave peak
- SQT correspond to the QRS and T waves peak
- SQ correspond to the QRS wave peak.

To search the peaks, we begin to detect the

extremums. And then using the threshold we determine the peaks corresponding to the QRS and T waves peaks.

First we compute the derivative of the signal:

$$dx(nT) = x(nT-T) - x(nT).$$

The extremums are obtained by

$$dx(nT)*dx(nT-T) <=0.$$

The extremums corresponding to QRS or T peaks is obtained when x(nT) is superior to a threshold. The threshold is function of the previous QRS and T peaks.

## III.2 QRS WIDTH DETECTION ALGORITHM

Fig. 6. shows the signal at various steps in algorithm process. First, in order to reduce the influence of noises (baseline wander and T wave interference), we compute SQ which not other than the pass band filtered signal. The next process is to make signal differentiation to provide the QRS wave slope information, followed by squaring in order to obtain positive value and to intensify the slope wave. Then to obtain QRS wave feature, the computation with the moving window integration is considered.
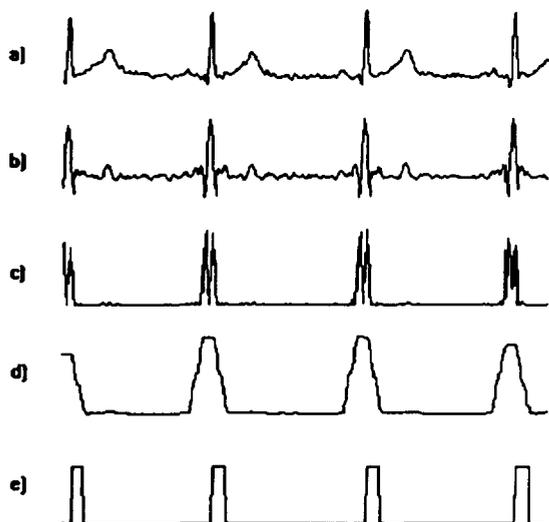


FiG.6. QRS width algorithm determination. a) Original ECG signal. b) SQ. c) Output of differentiation and squaring. d) Moving-window integration output. e) QRS width output pulse.

We can resume this algorithm as follows:

1). compute SQ (see the first algorithm).

$$y(nT)= SQ(nT)$$

2). compute the differentiation

$$y(nT)=x(nT-T)-x(nT)$$

3). compute the squaring

$$y(nT)=[x(nT)]^2$$

4). compute the moving window integration

$$y(nT)=(1/N)[x(nT-(N-1)T)+x(nT-(N-2)T)+...+x(nT)]$$

where N is the number of samples in the width of window integration .

The relationship between the QRS width and the window moving integration is shown in Fig. 7. The width of the window N should be approximately the same as the widest possible QRS wave[7].

5). compute the width of the Y slope to determine the QRS width.
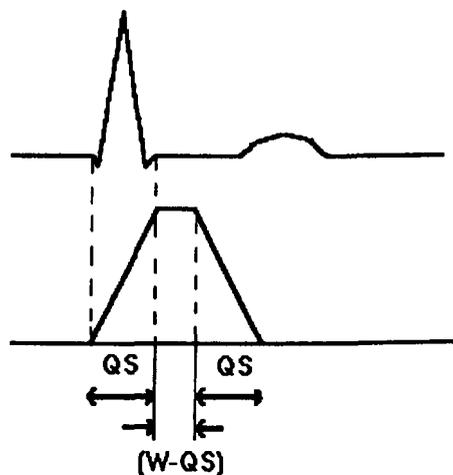


Fig. 7. ECG signal and output of moving-window integrator. QS: QRS width. W: Width of the integrator window.

## IV. EVALUATION

The algorithm will be evaluated later to a standard arrhythmia database in the paediatric service at Rouen Hospital .

Our algorithm is simulated using Matlab process. The ECG complex is prefiltred analogically with a band pass filter [0,5-40Hz] and sampled with f=160Hz .

The result of the algorithm evaluation is shown in fig.5 and fig.6. To prevent tachycardia occurred during infant monitoring, practician can evaluate the QT and QRS duration between two consecutive QRS complex which is done in our case in the paediatric service.

If the T detection give an unwedging mean of 15ms, the QRS one is 0. The QRS width is been detected with incertitude which not exceed 5%. The results obtained shows clearly the reliably of the detection algorithm.

215

## V. CONCLUSION

The aim of this paper was not only to show an efficient QRS wave forms detection algorithm, but rather to consider other possibilities offer by the wavelet theory of such an approach in the pattern recognition of biological signals. Also, the correlation will be made between different case of tachycardia and RR, QT, QRS result detection in infant monitoring. Future development may include an implementation possibility of this method in a "STAR Corp." DSP which will become integrated into a real time signal processing chain, and the definition of automatic alarms occurred in ECG waves forms disorder.

### References:

[1] M. Hubin, J. Sabor, M. Hassani, *"Multisensor Intelligent System For Infant Monitoring"*, In Proc.Int.Conf.on Sensors , Nürnberg, pp179-185, 1993.

[2] A.Grossmann and J.Morlet, *"Decomposition of Hardy Functions into Square Integrable wavelets of Consatnt Shape,"* SIAMJ.,vol.15,pp.723-736,1984

[3] S. Mallat, W.J. Hwang, *"Singularity Detection and Processing With Wavelets"*, IEEE.Trans.on Informat.Theory, Vol.32.march 1992.

[4] A. Cohen, L. Daubechies and J.C. Feauveau, *"Biorthogonal Bases of Compactly Supported Wavelets"*, AT&T Bell Laboratories, Technical Report, n°TM 11217-900529-07.

[5] J.C. Feauveau, *"Analyse multiresolution par ondelettes non orthogonales et Bancs de filtres numériques"*, Thèse de Doctorat, Université de Paris Sud, France 1990.

[6] M. Barrat, O. Lepetit, *"Fast processing of the wavelet transform"*, Traitement du signal, Vol.8, No. 1, 1990.

[7] J. Pan, WJ. Tompkins, *"A real time QRS detection Algorithm"*, IEEE Trans. on Biomedical Eng. Vol. 32, No. 3, March 1985.

[8] R. Alcantara, R. Gutierrez, *"Une nouvelle methode pour la detection de singularites utilisant la transformee en ondelette"*, XVI colloque GRETSI, pp.141-144, 1993.

[9] Y.V.Olivier, *"R-wave detecttion in the presence of muscle Artifacts"*, IEEE Trans.on Biomedical Eng., Vol. 31, No. 11, November 1984.

[10] A.D. Forbes, H.B. Jimison, *"A QRS detection algorithm"*, Journal of Clinical Monitoring, Vol. 3, No. 1, January 1987.

# Optimization of preprocessing during the learning process of Neural Networks. Application in pattern recognition

H. Witte, M. Galicki, J.Dörschel, A. Doering , G. Grießbach, M. Eiselt*

Institute of Medical Statistics, Computer Sciences and Documentation; *Institute of Pathophysiology; Medical Faculty; Friedrich Schiller University Jena; D-07740 Jena

## Abstract

The final aim of the strategy given in this study is to train the Neural Network (NN) classifier and the adaptive preprocessing units (APU) simultaneously, i.e. that properties of preprocessing will be chosen automatically during the training phase. The strategy was developed to obtain optimized recognition units which can be efficiently integrated in strategies for monitoring the cerebral status of neonates [1]. Therefore, applications e.g. in neonatal EEG pattern recognition and QRS (ECG) detection (preprocessing for heart rate analysis used in brain stem monitoring) will be demonstrated.

## 1 Introduction

Artificial Neural Networks (NN) have been used as classifiers in pattern recognition system for many applications, due to their ability to extract the essential features from a training set of patterns (set of feature vectors) and to identify unknown patterns. Additionally, by means of a previous preprocessor and feature extractor, an amplification of the main features of the patterns can be performed supporting the performance of NN pattern classification. A separate optimization of preprocessing (and feature extractor) and NN can only result in a suboptimized whole recognition system. In contrast to this, our aim was to find an optimization strategy which optimizes preprocessor and NN simultaneously [1].

In our applications, real-time preprocessing is needed and therefore different recursive adaptive estimation algorithms as preprocessing units were introduced [2]. The properties of these estimations can be controlled by adaptation factors. This is an important fact in realizing the optimization strategy of the preprocessing units in connection to and during the training (learning phase) of the NN. In this way, the preprocessing units can be seen as an "off-shore" layer of the NN.

The optimized recognition systems can be used for the main components within the concept of monitoring the cerebral status of neonates. Main application is related to EEG pattern recognition. Functional methods like EEG may determine disturbances before morphological damage can be verified. The so-called **discontinuous EEG** in neonates is characterized by two sequential patterns of EEG activity, named burst and interburst. The quantification of the discontinuous EEG may be a first step in the direction of automatic EEG monitoring in neonates [1]. However, both false-positive and -negative results with respect to the outcome were obtained when only the EEG was considered. This might be due to the fact that brain-stem damage may occur without EEG abnormalities. Heart rate fluctuations (HRF) are able to describe the vegetative brain-stem function [3].

Based on the ECG recording, the detection of the QRS wave in the ECG has to be carried out to find time points which designate identical moments within subsequent ECG events. These time points can be designated by impulses. Accordingly, a series of impulses results, which forms the basis for representation and analysis of HRF. Therefore, a recognition unit for transient signal components (QRS waves of ECG; spike events in EEG) have to be available. Recognition of spike activity must be taken into consideration to prevent false positive detections of burst activity.

Furthermore, other recognition units can be integrated into such a monitoring strategy, e.g. recognition of surface EMG patterns (e.g. classification of body movements of the neonate).

All these efforts result in optimized recognition subsystems consisting of preprocessor, feature extractor and NN classifiers.

## 2 Description of the adaptive algorithms (preprocessors)

The adaptive basic algorithms used are:
* **mean value**

$$M_{n+1} = M_n + c^*(x_{n+1} - M_n) \qquad \Rightarrow (\text{operator } \mathbf{Y} = \mathbf{M^c(X)}), \qquad (1)$$

* **second statistical moment**

$$E_{n+1} = E_n + c^*(x^2_{n+1} - E_n) \qquad \Rightarrow (\text{operator } \mathbf{Y} = \mathbf{E^c(X)}), \qquad (2)$$

* **momentary power within a frequency band**

$$E^{c1}\{M^{c2}[X - M^{c3}(X)]\} \qquad\qquad (3)$$

($c_1$ determines the time resolution, $c_2$ the upper, and $c_3$ the lower cut-off frequency of the pass band)

* **alpha-quantile**

$$Q_{n+1} = \begin{cases} Q_n + c\,\alpha & \text{if } x_{n+1} \geq Q_n \\ Q_n - c\,(1-\alpha) & \text{if } x_{n+1} < Q_n \end{cases} \qquad (4)$$

Such adaptive estimations can be viewed as processing units with one signal input (sequence X), one or more control input(s) to vary the adaptation factor(s) and one output (sequence Y).

Through the variation of the adaptation factors, essential properties of the estimation quality can be influenced. **Such properties include** estimation accuracy and time resolution, the quantity of signal power above an adaptive threshold as well as filter properties of adaptive estimations.

## 3 Neural Networks

The Neural Networks (NN) used were multi-layer networks [4] (error backpropagation; 2 (1) hidden layers; logistic activation function; learning rate = 0.2; momentum term = 0.2). The network structure was created empirically. Additionally, the common learning algorithm (APU and NN) uses a modification of the backpropagation algorithm for simultaneous optimization of adaptive preprocessing. The modifications differ between the applications 5.1 and 5.2. As an alternative, a global learning algorithm based on a random search technique was used.

## 4 Instantaneous optimization of the adaptation parameter

An error function $e(c)$ was introduced which depends on the scalar c and it presents a value obtained after the one-step backpropagation algorithm. Variations of the parameter c result in different values of this error function $e(c)$. The high complexity and nonlinearity of the error function $e(c)$ prevents an analytical solution. Simple numerical methods will be introduced which find an optimal c-value for the minimization of this modified error function.

Firstly, a gradient technique was used, when the error function $e(c)$ was differentiable by c. This method for the optimization of the adaptation parameter c is based on the following formula

$$c' = c - \gamma\, \text{grad}(e(c)) = c + \Delta c$$
$$\text{grad}(e(c)) = (\partial e/\partial c)_c,$$

where c' is the actual value, c the previous value and $\gamma$ the step length (**applications 5.1**).

Secondly, a gradient-free method (simplified version of Hooke-Jeeve's algorithm) can be used for the computation of $\Delta c$ if a differentiation of $e(c)$ by c is not possible (**applications 5.2**).

These numerical optimization methods were embedded into the whole optimization of the error function, which can be described in general as follows [1]:

**A) Starting stage:** The correction term for the weight matrix $\Delta W'(W_0, Y(X,c_0))$ was obtained after performing one backpropagation step, where $W_0$ is the starting matrix of the weights (stochastical choice) and $c_0$ is the starting value of c.

**B) Modification stage** for the adaptation parameter $c' = c + \Delta c$ where c' is the actual value, c the previous value and $\Delta c$ a correction term (computed as shown above).

**C) Learning stage of modification:** In this stage the usual backpropagation learning step is performed, using the modified value of the adaptation parameter $c' = c + \Delta c$.

A principle representation of the whole optimization scheme is shown in Fig. 1.
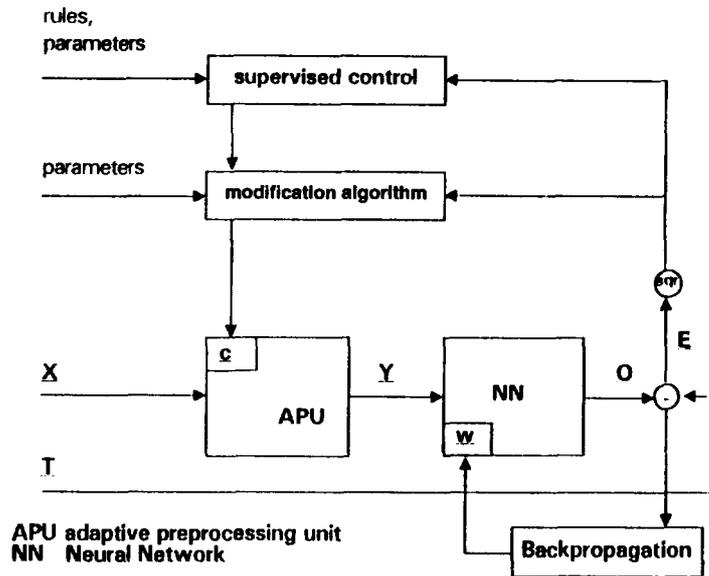
Fig. 1: The principal representation of the optimization scheme used.

## 5 Results

### 5.1 Strategy in EEG (and EMG) pattern recognition

The pattern recognition unit consists of the instantaneous computation of the spectral band power (momentary power) via equ. 3 followed by an averager (feature extractor) and NN. As NN input parameters (features), the sequence of mean power values in 4(8) disjunct time sub-intervals (averager as feature extractor) were used simultaneously (8 input neurons for each EEG channel and each frequency range). In this way, a power profile of the patterns (spatio-temporale structure) can be computed and used for the NN input parameter set (Fig. 2).
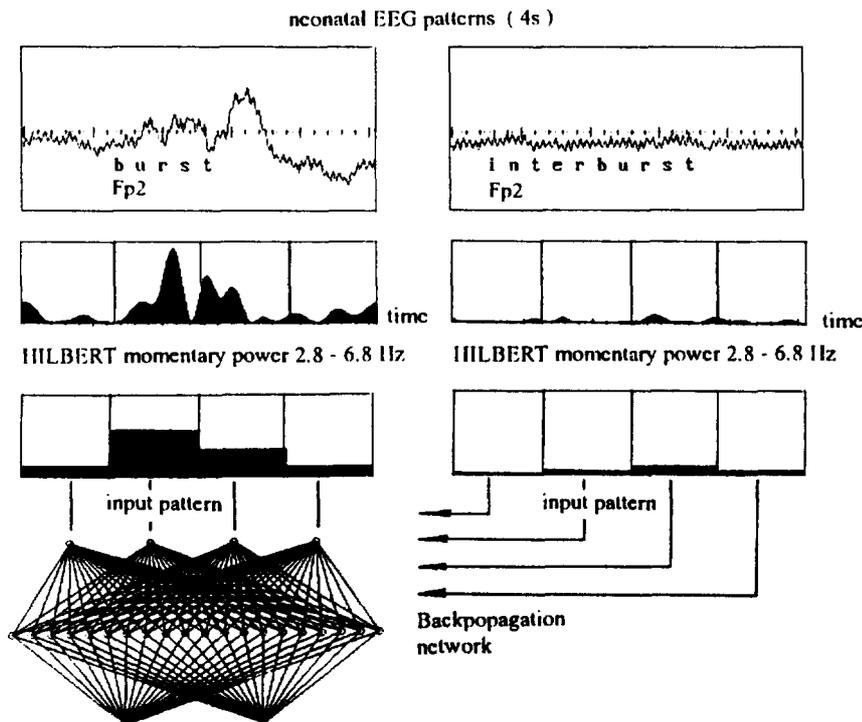


Fig. 2: Preprocessing and feature extraction for neonatal EEG pattern recognition (explanations, see text).

219

Selected burst and interburst patterns were used as a learning set (4 babies, 25 burst and 25 interburst patterns per child). During the learning phase the mid-frequency of a band pass characteristic (controlled by $c=c_2=c_3$) will be optimized. The aim is to find the most discriminating frequency band between both patterns.

For the segmentation of the discontinuous EEG into burst and interburst intervals, an optimal mid-frequency of 5 Hz was found ($c_2, c_3=0.2$ for sampling frequency of 128 Hz). The corresponding frequency range contains a characteristic initial wave of the burst pattern [5]. Therefore, this special feature of the burst pattern seems to be responsible for the good results of burst recognition.

The on-line classification can be performed via a moving window of a defined length (same length as training data), in which the power profile will be calculated from the time course of momentary power. This profile will be given (for each sampling point) at the inputs of the trained NN.

The same computing scheme can be used for pattern recognition in surface EMG recordings.

### 5.2 Learning optimal adaptive thresholds for detection of transients (QRS and spike detection)

During the occurrence of epileptic spikes (transient EEG components), higher amplitude values (power values) and low variance of momentary frequency estimation can be observed, in comparison to the surrounding background activity [6]. The joining of the characteristics of these parameters during spike and surrounding EEG activity can be performed through adaptive threshold calculation. The adaptive estimation of the $\alpha$-quantile ($Q_n$) can be used as such a threshold. An $\alpha$-quantile ($Q_n$) estimation was used to detect high power values (course of momentary power), in comparison to the surrounding activity. Fig. 3 shows the results of the stepwise computing of an EEG recording with (periodic) spike activity. The preprocessing will be realized through the calculation of the momentary spectral power (A) and a subsequent calculation of an adaptive threshold (B). The quantity of overcrossings (C) were used as one important parameter for spike detection.
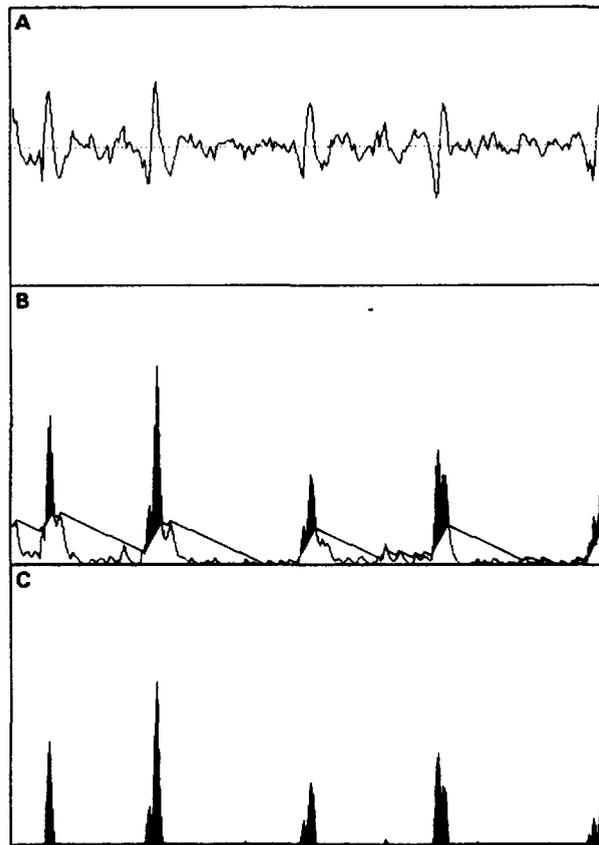


Fig. 3.: Preprocessing and feature extraction for spike as well as QRS detection. For this example spike activity was used.

For the optimization of the threshold the parameters c and α were controlled (equ. 4). Additionally, both band pass and threshold can be integrated simultaneously into the training of the NN for the optimization of spike detection. The same strategy was used for QRS detection in ECG recordings. Results of QRS detection are represented in Fig. 4.
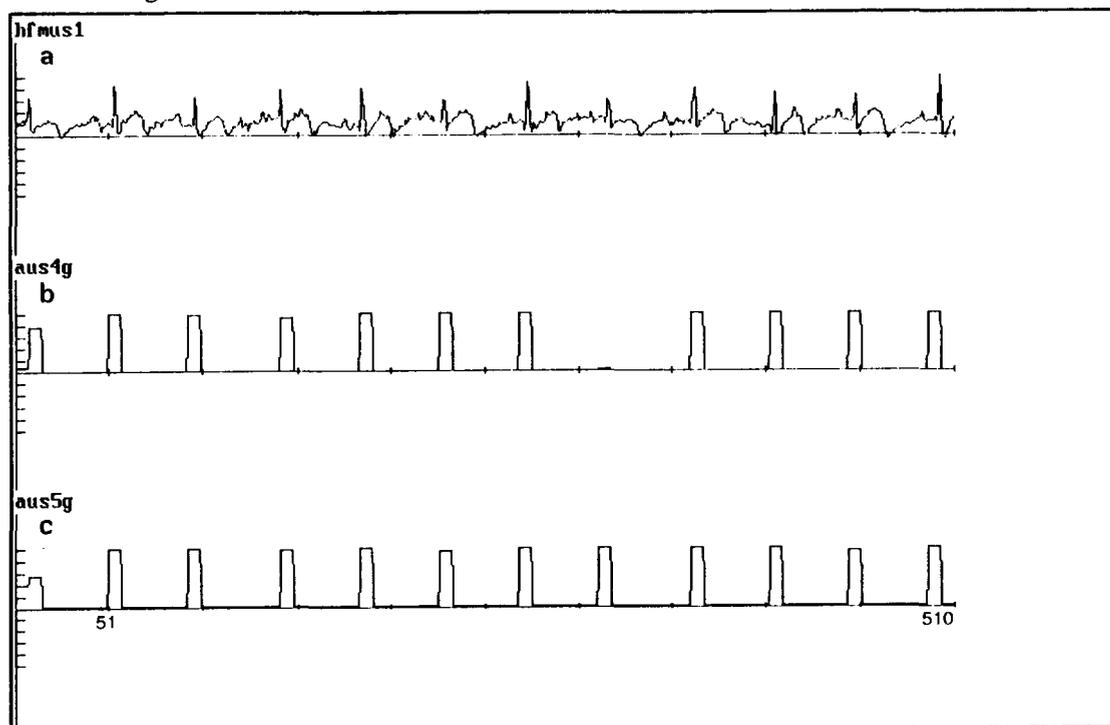


Fig. 4: Results in QRS detection for two c-values (equ. 4) which were found as local minima. The original ECG recording (a) was preprocessed (see Fig..3) and the QRS wave was detected by NN (1-5-1).
The parameters are (b) α=0.5, c=0.326 (c) α=0.5, c=0.49

## 6 Conclusions

A strategy was introduced for the common optimization of preprocessing units and for the classifier during the training process. Furthermore, applications were demonstrated in the field of monotoring the cerebral function in neonates. In this context, the Neural Network classifier (supervised learning) was used, because

* experiences exist in using dynamic spectral parameters as input parameters of NN [1,5,6]

* expert knowledge can be utilized via the preparation of the training set by a physician, and

* the learning algorithm of NN can be modified to simultaneously optimize characteristics of preprocessing during the training phase.

## 7 References

[1] Witte,H., Galicki,M., Dörschel,J., Griessbach,G., Eiselt,M., Doering,A., Hoyer,H.: Optimization of adaptive preprocessing units during the learning process of Neural Networks. Application in EEG pattern recognition. Informatik aktuell (1993), Springer-Verlag, 71-80.

[2] Grießbach,G., Schack,B.: Adaptive quantile estimation and its application in analysis of biological signals. Biom.J. 35 (1993), 165-179.

[3] Witte,H., Rother,M.: High frequency and low frequency heart rate fluctuation analysis in newborns - A review of possibilities and limitations. Basic Research in Cardiology 87 (1992), 193-204.

[4] Rumelhart,D.E., Winton,G.E., Williams,R.J.: Learning internal representation by error backpropagation. In: Rumelhart,D.E., McClelland,J.L. (eds.): Parallel Distributed Processing. Explorations in the Microstructure of Cognition. The MIT Press, Cambridge (Mass.), 1986.

[5] Witte,H., Dörschel,J., Grießbach,G., Eiselt,M., Galicki,M., Arnold,M.: The combination of on-line preprocessing with Neural Network classification for EEG monitoring in neonates. Beitr.Anaesth.Intensivmed., 1994 (in press).

[6] Witte,H., Eiselt,M., Patakova,I., Petranek,S., Griessbach,G., Krajca,V., Rother,M.: Use of Discrete Hilbert Transformation for automatic spike mapping. A methodical investigation. Med.&Biol.Eng.&Comput. 29, 242 (1991).

# BIOMEDICAL ENGINEERING EDUCATION IN POLAND

Grzegorz Pawlicki & Tadeusz Pałko
Warsaw University of Technology
Institute of Precision and Biomedical Engineering
Chodkiewicza 8, 02-525 Warsaw, Poland

Keywords: biomedical engineering, clinical engineering, education.

## INTRODUCTION

The history of Biomedical Engineering (BME) in Poland has begun from Physics Department of Radium Institute in Warsaw organized by Nobel Prize winner Maria Sklodowska-Curie.

Education in BME started in poland in 1946 at the Electrical Faculty of the Warsaw Uniwesity of Technology (WUT) as specialization termed Electromedical Technology. It was probably the first european - and perhaps in the world - system of training in the BME at the university level. Students with background in electrical engineering were trained in radiology and radiometry, X-ray arrangements, electromedical equipment: ECG, EMG, EEG, electrostimulators, diathermy, light therapy lamps etc. From the 1951 BME education was transferred from the Electrical to the Electronics Faculty. In the 1970 the BME Division transferred from Electronics Faculty to the interdisciplinary Faculty of Fine Mechanics. At the Electronics Faculty lasted only Nuclear Medicine Laboratory, that changed later into Nuclear and Medical Electronics Division, which exist till now.

BME at Fine Mechanics Faculty extended the training range on some mechanical aspecte of biomedical instrumentation like transducers, optical systems and artificial organs.

In the 1970's the training in BME started at the Silesian, Gdansk and Wroclaw Universities of Technology as well as in Physics Faculty of Warsaw University. Till now in Poland were graduated more than 2,5 thousand of BME engineers. Among them about 1500 at the WUT. Some of them achieved Ph.D. degree - in Poland about 150 persons.

## CURRICULUM OF BME AT THE FINE MECHANICS FACULTY

Graduate 5 years training with the only terminal degree of M.S. and postgraduate studies of Ph.D., as well as continuous education courses are offered. The education programs of BME for M.S. degree comprises about 3500 hrs (table 1.) and is divided as follow: 2625 hrs for background and 855 hrs specialization subjects. It does not include hours for dissertation preparation and practicals (14 weeks, comprising diploma practical).

Background (table 2.) includes: general subjects - mathematics, physics, material engineering and foundations of: mechanical and electrical engineering, electronics, computer science (software), metrology, automation, designing of mechanical and micromechanical structure and technological processes of precision and electronic devices manufacturing.

Table 1

# Distribution of the teaching hours for M. S. study in BME

| Subjects | Hours | The year of study | | |
|---|---|---|---|---|
| | | I and II | III and IV | V |
| General background | 2625 | 1455 | 990 | 180 |
| Specialization | 855 | - | 525 | 330 |
| Total | 3480 | 1455 | 1515 | 510 |

Table 2

## LIST OF BACKGROUND SUBJECTS IN BME

| Nr | The name of subject group | Total number of hours | The year of study | | |
|---|---|---|---|---|---|
| | | | I and II | III and IV | V |
| 1 | Mathematics | 270 | 240 | 30 | |
| 2 | Physics | 210 | 150 | 60 | |
| 3 | Mechanics | 210 | 150 | 60 | |
| 4 | Electrotechnics (Including electronics) | 390 | 180 | 210 | |
| 5 | Material Engineering | 60 | 45 | 15 | |
| 6 | Design of precision devices | 270 | 210 | 60 | |
| 7 | Manufacturing technol. of precision devices | 210 | 60 | 150 | |
| 8 | Informatics (computer science) | 120 | 90 | 30 | |
| 9 | Metrology | 210 | 105 | 105 | |
| 10 | Automation | 120 | 60 | 60 | |
| 11 | Projects | 180 | | 90 | 90 |
| 12 | Diploma seminar | 60 | | | 60 |
| 13 | Economics | 90 | 30 | 30 | 30 |
| 14 | Humanistic subjects | 210 | 120 | 90 | |
| 15 | Propedeutics of precision technology | 15 | 15 | | |
| | Total | 2625 | 1455 | 990 | 180 |

Specialized education (table 3.) includes biology, outline of anatomy and medical foundation of BME, health services organizations, biomedical informatics, biological processing, electromedical instruments, ultrasonography, radiological equipment, biomechanics, biomaterials, biomedical optics, artificial organs, intensive care systems, assurance quality and testing of devices.

Table 3

# LIST OF SPECIALIZATION SUBJECTS WITH THE TEACHING HOURS

1. Outline of Anatomy (study of human body structure) . . 30
2. Basic Physiology for Biomedical Engineers (study
   of chosen physiological processes) . . . . . . . . . . . . 90
3. Propedeutics of Medicine (principles of practical
   medical procedure in diagnostic and therapy) . . . . . . . 30
4. Outline of Biology (general cell physiology
   and genetics) . . . . . . . . . . . . . . . . . . . . . . . . . 30
5. Electromedical Devices (equipment for electrical
   signals of human body detection and analysis:
   electrocardiograph, electromyograph, impedance
   rheograph, etc.) . . . . . . . . . . . . . . . . . . . . . . 120
6. Radiological Devices (equipment for internal
   structures of human body visualization using x-rays;
   Computer Assisted Tomograph . . . . . . . . . . . . . . . 105
7. Biomedical Informatics (data basis for computer
   assisted diagnostic, expert system,
   hospital administration) . . . . . . . . . . . . . . . . . . 105
8. Biomechanics (mathematical description
   and analysis of human body mechanical functions) . . . . 45
9. Specific Requirements for Biomedical Equipment
   (patient safety during medical treatment, equipment
   reliability, compatibility, devices testing) . . . . . . . . . 45

It is offered additionaly facultative (selected) subjects such as haemodynamic measurement devices, biomedical optics, rehabilitation equipment, special technologics, sterilization devices, biosystems etc. In our teaching process we cooperate closely with hospitals and Medical Academies. We also offer a 3 years training period for preparation of a doctoral dissertation. Programs are structured according to the individual interets of Ph.D. students. We propose as well in the framework of continuous education, one year long training in BME containing about 200 hrs of teaching.

## SKILL AND PERSPECTIVE OF EMPLOYMENT

Our graduates are prepared first of all to take the charge of clinical engineers and clinics. They can also be employed in scientific institutes and industry of biomedical technology and medical arrangement. They have got knowledge on compatibility, legalization, quality assurance related to medical devices systems. Our BME graduates are also prepared to design, in collaboration with physcicians and engineers of other specialties.

## CONCLUSIONS

Presented above system of education provides good fundamentals in BME. The system is flexible and allows to develop individual interests of the students. Specialization in the specific areas of BME is possible through professional training in hospitals, producers and post graduate courses. Fundamental and extensive education allows graduate to change his area of activity.

## REFERENCES

1. Jaron D. (1992) Biomedical Engineering Education - A perspective Proceedings of the 1992 International Biomedical Engineering Days. pp. 18-21. Bogazici University. Turkey.
2. Plonsey L. (1973) New directions for bioimedical engineering education. Eng.Educ.64:177-179.
3. Potvin A.R., Long F.M., Webster J.G., Jendrucko R.J. (1981) Biomedical engineering education: enrollment, courses, degrees and employment. IEEE Transaction on Biomed.Eng. BME 28 1:22-28.
4. Schwartz M.D. (1984) The emerging field of clinical engineering and its accomplishments. IEEE Tranations on Biomed.Eng. BME. 31 12:743-748.

# Designing the PELICAN-System

Avgeridis K.[1], Panagopoulou G.[2][3], Sirmakessis P.S.[2][3], Tsakalidis A.[2][3]

[1]Sociologist,
University Hospital of Patras
26500 Patras, Greece

[2]Computer Technology Institute,
P.O. Box 1122, 26110 Patras,Greece

[3]Department of Computer Engineering and Informatics,
University of Patras, 26500 Patras, Greece

## Abstract

In this paper we present the design of a complete blood-transfusion information system, named the PELICAN-System. This system is now under development in the University of Patras in co-operation with the Blood Centre of the University Hospital. In this work we describe the main aspects of its architecture and the basic functions that the system should offer.

## 1. Introduction

In recent years a lot of work in computer science has focused in the usage of computer systems in hospitals and other health centres. This attempt aims to introduce new technologies in medical science, in order to improve the quality of the services provided. Blood Centres are one of the most active parts of a hospital with significant contribution to hospital's operation. Especially nowadays we have a significant increase in the needs of blood and its products in Greece, due to the increase in car accidents and the needs of patients suffering from Thalassemia. Moreover now it is more significant than before to have a computerised way to test and control blood tests for diseases such as AIDS, Hepatitis B and C, and sex related diseases.

Moving in this direction we have started a project that tries to combine computer technology with blood transfusion processes. In this project we firstly studied the way a blood centre in Greece works. We tried to collect any available information from local hospitals and blood centres and to study any previous work in this area in Greece and abroad. At the end of this process we arrived at the specifications and the description of a complete blood-transfusion information system named PELICAN-System. The design of PELICAN follows the standards described by ISO quality control elements ([2], [8], [9], [10]) and the requirements stated in international literature ([1], [6], [14], [16], [17], [18], [19]).

As far as we know, various European countries have already implemented different blood bank information systems. More precisely the Hesse Central Blood Bank of the German Red Cross is using an automated routine for blood donation testing and blood bank organisation based in a computerised system ([3]). This system is using the Groupamaticsystem and the ABBOTT Commander System to test blood products. These instruments are connected with a data management system installed on a HP 1000 computer in the lab area. Similar strategies have been adopted by other systems ([7], [15]), like the system currently established in the South East Scotland Transfusion Service ([13], [14]) or in the prototype of the Long Island Blood Distribution System ([5]).

The PELICAN-System is a complete blood-transfusion information system, which is under development in the University of Patras in co-operation with the Blood Centre of the University Hospital. We named the system using the name of the bird pelican, as this bird uses its own blood as an antidote to save the life of its children in case of diseases or snake bites.

In this work we present the basic architecture of the PELICAN. A more detailed description can be found in ([4], [12]). The paper is organised as follows: section 2 describes the architecture of the system by presenting the different databases and utilities sections. The work that has already been done and other design issues are the subject of section 3.

## 2. The Architecture of PELICAN-System

The PELICAN-System follows the international standards in the design of its architecture. Different independent modules are integrated in a communication protocol and operate as a whole system. A schematic representation of the system can be seen in figure 1. The independence of each module guarantees the ease of expansion and the way that new approaches can be considered. In the following sections the modules of PELICAN are described.
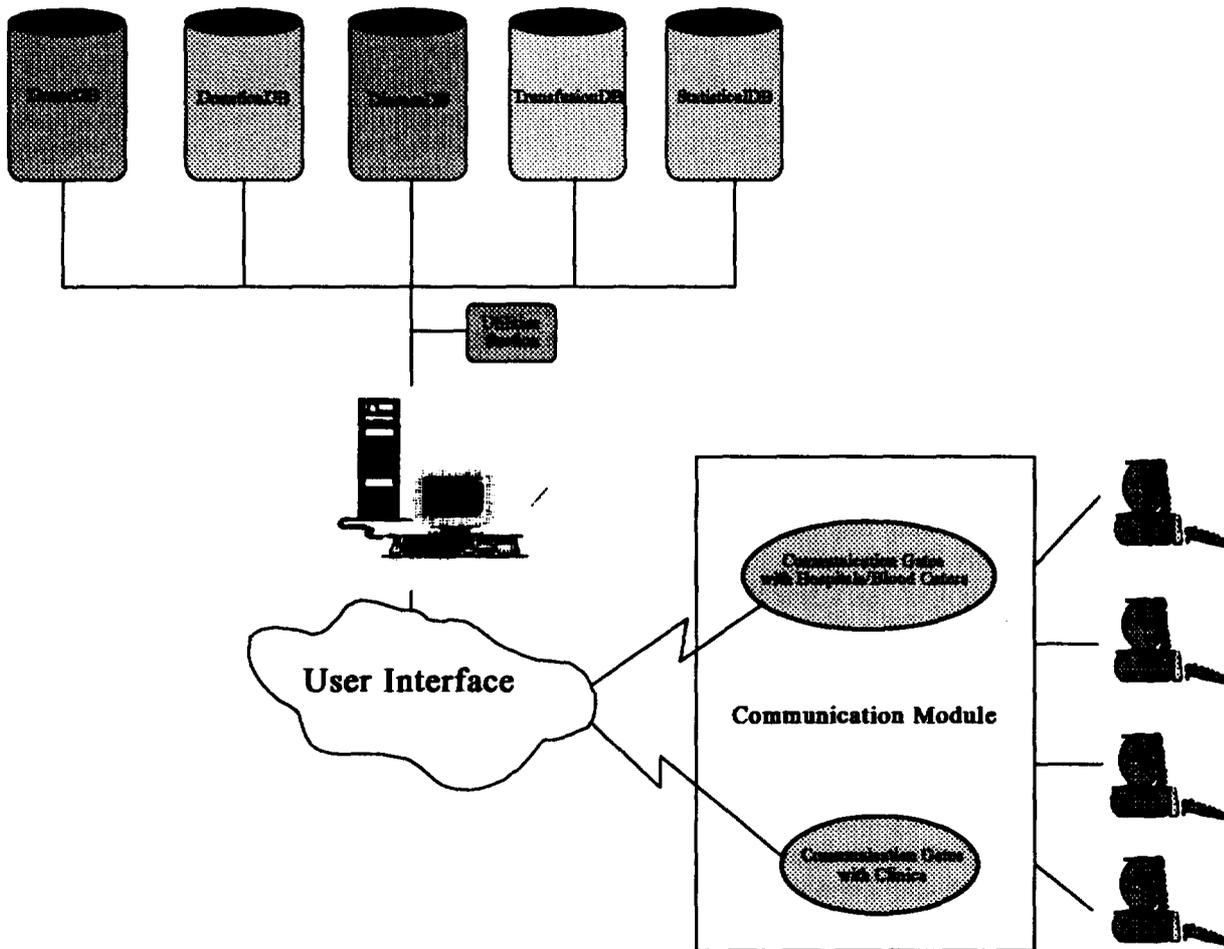


**Figure 1.** The PELICAN architecture and its connections

## 2.1. The User Interface Module

The user interface of PELICAN has been designed in a way which guarantees that it will be accepted by the personnel working in blood centres. Following the international way of designing user interfaces, a graphical interface has been adopted. We have taken under consideration that our design should be able to follow every characteristic of the hospital's or blood centre's environment. So two hardware platforms will be used. On the one side we have a DOS platform of a local network based on personal computers and on the other side a UNIX

platform controlled by a mainframe server with various inexpensive terminals connected to the server. Microsoft Windows for the personal computer version and X Windows for the mainframes have been adopted as the proper interface for PELICAN.

On line help should be always available via brief messages presenting in the last two lines of the screen and an extended help support can also be accessed via a key press (F1 is the windows standard).

An educational program will be available as a tutorial or a self teaching method in the system's use. All screens have been designed to look like the forms that personnel is now using. This means that we tried to make the screen forms look the same as the forms currently used in practice in order to achieve a faster way passing from the hand typing to the computer usage.

Different input devices such as scanners or ABBOTT's control devices with digital output have been studied and most of them will be connected to the system in order to reduce the case of typing errors during the input process.

## 2.2. The Database's Module

The database module of PELICAN consists of five databases:
- Donor DB
- Donation DB
- Disease DB
- Transfusion DB
- Statistical DB

The existence of five different databases guarantee that any new hospital and blood centre can ask for the contents of any one of these and receive a copy as soon as possible.

*Donor Database* contains information relative to donors. Basic demographic elements and contact information can be stored there. Secondary information such as the preferable time for transfusion, donation period etc. can also be stored. The database can handle the case of voluntary donation of individual persons or donations organised by different groups. Queries like "find the donor with number xxx", or "find all the donors that can give blood in the local police department" can be performed in the DonorDB. Every person gets a personal unique number as long as he becomes donor. This number follows a predefined format that describes his blood type, rhesus and the preferred way for donation (e.g. with other members of a group, or when a relative needs blood). Using this number the system can search for every information relative to the donor very quickly. Moreover in case of power failure or system crash, the personnel of the blood centre can easily find out a few things about the donor by just looking at his number.

The *Donation Database* contains information about each person's donations, that is blood pressure measurements, temperature, hematocrit, etc. The dates of donations can also be found here. In this database a user can execute queries that present different measurements during donations. More time consuming queries can be also asked. For example a user can ask for a list of all donors with a predefined blood type, rhesus and antigens that should give blood (meaning that the necessary interval from the previous donation has passed). Another important query, called emergency list, can present the donors that have more than three months to give blood. This query can help to cover emergency needs for blood products.

Disease and other important laboratory tests can be inserted into the *Disease Database*. The contents of this database are actually "sensitive" data and can be retrieved only by authorised personnel. For this reason special care is taken to ensure system's security. First of all, every tuple in the database uses the serial number of the blood bag to identify the input. This guarantees that all tests are not connected to the donor himself, but to a bag. Moreover every positive test result is stored as negative. The correct result is stored in encrypted form using different encryption algorithms from time to time. The DES and RSA encryption algorithms are now being tested for accuracy.

The fourth database, called *Transfusion Database*, contains information about blood transfusion. The whole history of a blood bag can be found here. Information for every blood unit that has been transfused into any patient at the local hospital or at any hospital, results of compatibility tests and all the possible blood reactions are also stored here. In case of reactions all the actions taken are stored here. Quality control tests for the blood, for all its products and for blood bags can be inserted in the proper tuples of the database. The contents of the blood

refrigerators are organised in the TransfusionDB. The system will inform the personnel for the items that are near the expiration date every morning during the boot process.

The last database, the *Statistical Database*, is a database storing statistical data. More precisely reports of the way donors are distributed according to their sex, age, job, family state (married, single, divorced), education, frequency of donation and many other criteria are provided by the Statistical Database. The database is working, making the reports, during non working hours, possibly nights and Sundays, in order to have all results ready on request. A typical statistical form is presented in Table 1.

| Sex | Male | | | | | | Female | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| *Age* <br> *Blood Type* <br> *Rhesus* | *18-22* | *23-27* | *28-32* | *...* | *63-65* | **Total** | *18-22* | *...* | *63-65* | **Total** |
| *AB-* | | | | | | | | | | |
| *AB+* | | | | | | | | | | |
| *O-* | | | | | | | | | | |
| *O+* | | | | | | | | | | |
| *A-* | | | | | | | | | | |
| *A+* | | | | | | | | | | |
| *B-* | | | | | | | | | | |
| *B+* | | | | | | | | | | |
| **Total** | | | | | | | | | | |

**Table 1.** A statistical table for donors

## 2.3. The Utilities' Module

The utilities' section of the PELICAN-System contains a few utilities necessary for the centre's functionality. An electronic-mail option will be provided for the exchange of messages between the personnel. An easy-to-use word processor for preparing simple forms can also be used. A storage management program, that gives the ability to manage the contents (reagents and other disposables) of the centre's store, will work in the laboratory area. New orders would be transmitted to the technical department to get official approvement. This system should be connected with all the different network configurations of the hospital, in order to have connections with the administrative and technical departments.

## 2.4. The Communication Module

The communication module deals with the connection of the PELICAN-System with the clinics of the local hospital, as well with other hospitals and blood centres. Local communication is done via NOVELL or other LAN protocol, while communication to other hospitals is established via X.25 protocol. Personnel from the clinics can ask the system for available blood products and the system chooses the appropriate one, informing the blood centre for this action.

## 3. Conclusion

PELICAN-System is now developed in both DOS and UNIX platforms. At the time being only functions concerning the Donation, Donor and Disease Database have been implemented. A lot of work should be done in the communication module, due to the fact that there is neither a common platform in the hospitals' computer centres, nor a common organisational structure in the blood centres of Greece.

For the future our research work is focused in the enhancement of the PELICAN-System with a neural network, in order to support a "clever" mechanism for the checks of blood transfusions. We believe that this attempt will get the government's interest in order to produce a computerised platform for a complete regional blood bank organisation in Greece

# References

[1] Allen F., Brodheim E., Hirsch R., Steel D., Ying W., "Regional Blood Center Automation: Computer Surveillance of Donor Blood Processing", Transfusion, Vol. 18, No. 6, 1978.

[2] "An American National Standard IEEE Guide to Software Requirements Specification"

[3] "Automation of Routine Blood Donation Testing, Blood Product, Manufacturing and Testing and Product Distribution and Shipping Operations in a Blood Bank Organization" Poster of The Hess Central Blood Bank of the German Red Cross, Frankfurt am Main, West Germany

[4] Balatos K., Panagopoulou G., Pantelis A., Sirmakessis S., "Using Computers in the Blood Center of the University Hospital of Patras", Diploma Thesis in Department of Computer Engineering and Informatics (In Greek), (1992).

[5] Brodheim E., Prastacos G., "The Long Island Blood Distribution System as a Prototype for Regional Blood Management", Interfaces, vol 9, pp 3-20, 1979

[6] Brodheim E., Hitch R., Prastacos G., "Setting Inventory Levels for Hospital Blood Banks", Transfusion, vol 16, pp 63-70, 1976

[7] Frankfurter G. et. al., "A Regional Computer-based Blood Inventory Information System", Transfusion, Vol. 10. No. 4, 1970.

[8] "ISO 9000 Quality Management and Quality Assurance Standards"

[9] "ISO 9001 Quality Systems"

[10] "ISO 9004 Quality Management and Quality Systems Elements"

[11] Jennings J., "Blood Bank Inventory Control", Management Science", Vol. 19, 1973.

[12] Maniatis A., Balatos K., Panagopoulou G., Pantelis A., Sirmakessis S., Tsakalidis A., "BLOOD BANK DATABASE (BB-DB) Blood Centers in a Computerised World", in Proceedings of the 4th Greek Computer Conference (in Greek), vol 2, pp 563-574, 1993

[13] Mohr J.R, Kluge A., "The Computer and Blood Banking", Springer-Verlag, 1981

[14] Moores M. "Progress in Transfusion Medicine" Chapter 3: Computer Applications to Transfusion Practice, Principles and Practice

[15] Prastacos G., Brodheim E., "PBDS: A Decision-support System for Regional Blood Management", Management Science, Vol. 26, No. 5, 1980

[16] Prastacos G.P., Brodheim E., "Computer-Based Regional Blood Distribution", Computers and Oper. Research, vol 6, 1979

[17] Prastakos G., Maniatis A., "The Greek Blood Banking System: An Overview" EEC Report supported by OMAS, October 1992

[18] Rabinowitz M., Valinsky D., "Hospital Blood Banking: An Evaluation of Inventory Control Policies", Technical report, 1980.

[19] "Responsibilities in Implementing and Using a Blood Bank Computer System", American Association of Blood Banks Information Systems Committee [1989]

# DESIGN OF NEW IMPLANTABLE NEURAL CUFF ELECTRODES

M. Sawan, H. Barada, F. Duval[†], M. Hassouna[††], and M. Elhilali[††]

Dept. of Elect. and Comp. Eng., École Polytechnique de Montréal
Montréal, Qc, Canada, H3C 3A7
Email sawan@vlsi.polymtl.ca
[†]Dept. of Elect. Eng., Université de Sherbrooke, Sherbrooke, Canada
[††]Royal Victoria Hospital, Montréal, Canada

**Abstract:** New types of spiral cuff tantalum and platinum electrodes, based on a new design technology, have been developed. These electrodes form the bio-electronic interface between the tissues (nerves) and a multichannel urinary prostheses system. They are especially designed for bladder stimulation purposes. *In vivo* tests have been performed and a series of results are reported.

## INTRODUCTION

Available advanced implantable neural stimulation systems lack flexible electrodes. Current techniques and materials to produce biocompatible electrodes present important deficiencies, especially the lack of a flexible interface to tissue. Major progress are accomplished but still need improvement and development of new horizons [1,2]. On the biomedical side, the use of such electrodes is growing rapidly especially for neuroprosthetic applications such as the treatment of chronic pain, urinary incontinence and retention, auditory deficiencies, extremities control and others.
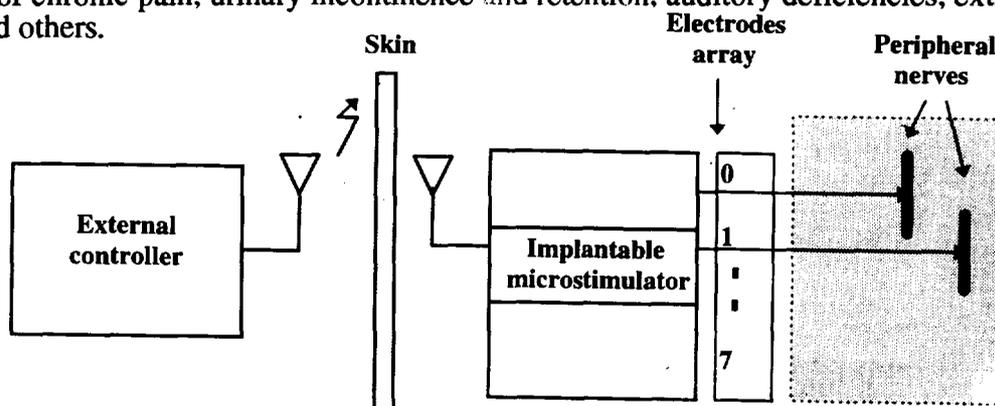


*Figure 1. General representation of the system.*

Implantable electrodes are divided into two main categories, muscular electrodes (intramuscular or epimysial) and neural electrodes (intraneural or nerve-cuff). The nerve-cuff is a promosing method due to the continual increase in investigations and developments of a such type of electrodes [1,2]. In fact cuff electrodes overcome several disadvantages of intramuscular and epimysial electrodes [2]. On the mechanical side, the failure of electrode/electrode leads is a major problem[1,5] so cuff electrodes are expected to provide a stable interface between the electrode and tissues if they are placed in areas of relatively low stress [2]; therfore they should improve the mechanical properties of the electrodes or leads. On the electrical side, they have lower excitation threshold currents which will permit the use of smaller electrode surface to increase selectivity without the risk of corrosion. They also exhibit low power requirements which make them an attractive choice for a fully implanted neural prosthesis system; furthermore, on the medical side, implantation of cuff electrodes should place minimal demand on both the patient and the surgeon [5]. In the present paper, we propose new techniques to produce implantable electrodes using tantalum and platinum materials. These electrodes serve as the interface between the tissues (nerves) and a multichannel implantable microstimulator

dedicated for bladder control. This system is composed of two main parts. The first part, an internal prosthesis, constitutes the implant [1] using CMOS gate array chip to control a wide variety of waveforms via eight monopolar stimulation channels which are connected to the electrodes. The second, is an external controller incorporates the necessary circuitry to transmit data transdermally via a radio frequency-coupled technique. Fig.1, shows a general block diagram of the system. The following section describes the techniques of designing and fabricating the tantalum electrodes.

## STIMULATION ENVIRONMENT AND BIOCOMPATIBILITY

Neural stimulation electrodes are intended for a hostile body environment, they have to be corrosion resistant and unconditionally biocompatible. In addition they should have a flexible electrode-nerve interface. This interface must be assured by a noble metal (platinum, iridium, etc...) for their inner characteristics, and in order to attenuate tissue damage and hydrolysis, a large electrode-nerve contact reducing the current density is necessary [1,4]. To avoid the friction between the electrode and the body environment, a thin layer of medical silicon for final insulation of the electrode will reduce the irritation and decrease the current leakage. Moreover, applying biphasic stimuli can optimize reversible reactions. On the other hand tantalum, as a conducting metal, is a particularly interesting choice. *In vitro* and *in vivo* tests have demonstrated that tantalum produces less effects on tissue than do silver and other materials, and it is easier to insulate it electrically from saline fluid than gold and platinum [1,3].

## ELECTRODES DESIGN AND REALIZATION

The development of the tantalum insulation technology and modeling of its insulating properties permit us to use efficiently this material during the design and the fabrication process. In fact, the use of tantalum as the conductor material has resolved the problem of insulation integrity, also the self passivating nature of the $Ta_2O_5$ provides excellent electrical stability. Moreover, the use of noble metals, such as platinum alloyed with irridium, rhodium, or rhenium were found to be the most suitable for long term excitation in biological medium. The chief factors upon which this decision is based are the minimal loss of metal due to electrolysis, and the biocompatibility of the platinum with biological surrounding [3]. The noble metals have proved better resistance to electrolysis than other commonly used metals such as Gold, Palladium, or stainless steel. In addition, studies have shown that there is no evidence of neural damage induced by long-term electrode implantation, both with or without active electrical stimulation. As result of these studies, pure platinum become the best choice to realize the electrode-tissue stimulation interface.

Fig. 2 shows the different steps that we used to make electrodes. Tantalum pentoxide ($Ta_2O_5$), an insulating and highly corrosion-resistant material, is formed on a tantalum surface by an anodization process, whereby a tantalum sample, the anode, is immersed in a dilute electrolytic solution with a current allowed to flow between the tantalum anode and the cathode, typically a platinum reference plate. The thickness of the oxide formed increases nonlinearly with the magnitude of the voltage applied and with the time of application [3].

The design steps are the following: tantalum foils are cleaned, and a solution of citric acid is used to help grow a 1000 Å of a tantalum pentoxyde ($Ta_2O_5$). At one end of the electrode tip (1x2 mm) is prepared to make the electrode-nerve contact. On the other end of the electrode (1x10 mm), a gold-plated connector is mechanically fixed, together with a gold plated tube to avoid breaking of the tantalum electrode at the level of the junction. The final step of the process consists of encapsulating the electrode for biological insulation. This step is done by surrounding the electrode with silastic tubing and moulding. With this process many type of electrodes were fabricated, monopolar, bipolar, and one double bipolar electrode configurations. Finally, the electrode-nerve contact is made by welding a foil of platinum on the foil of tantalum using the electrical discharge technique. This technique provides a good mechanical and electrical configuration of electrode-nerve area.
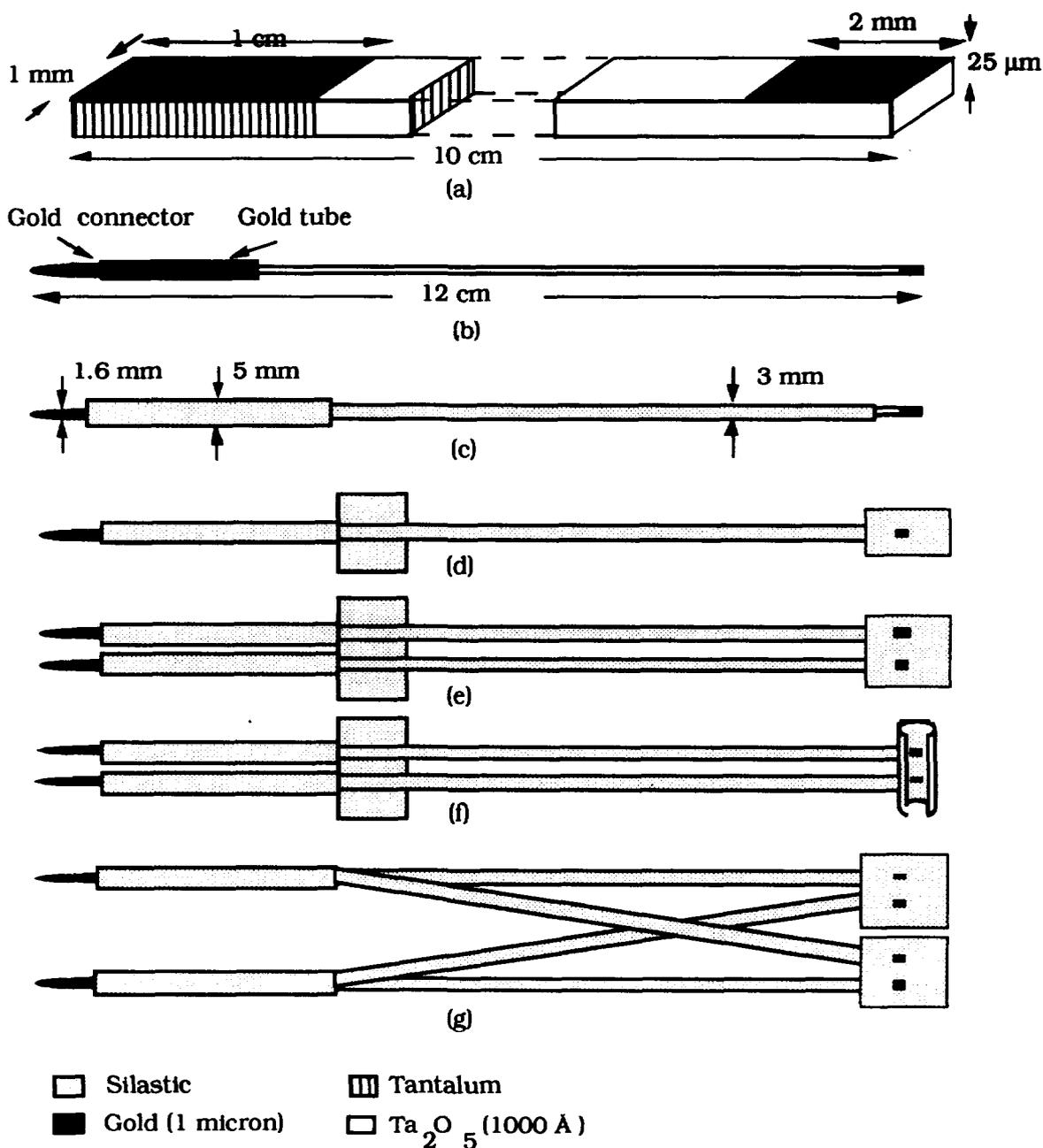
Figure 2. Electrodes design and categories: (a) Tantalum foil ready to make extremities contacts, (b) Mechanical connector and tube fixed on the electrode tip, (c) Completed sample monopolar electrode, (d) Monopolar type, (e) Bipolar with Silastic foil as interface, (f) Bipolar with Silastic tube as interface, (g) Bipolar with two parallel outputs with Silastic foil.

## RESULTS

*In vivo* tests on animals (dogs) have been performed. The tantalum electrodes were implanted in an appropriate positions to meet the requirements for bladder stimulation protocoles. Using the stimulation system described above parameters were determined to be a bipolar train waves of a high-frequency signal of 1-600 Hz and a low-frequency signal of 1-100 Hz. The pulse amplitude was found to be 0.5-2.0 mA, and the stimulation technique has been applied

either intermittently or continuosly during 10-40 Sec. In order to protect the stimulating tissues, we limited the maximum charge quantity. This limitation is obtained by coupling capacitors that allow the stimulus to be dc voltage free and to generate completely balanced waveforms. Each waveform period is composed of an active phase and a charge-recovery phase. The charge quantity of the active phase is maintained at a value lower than 900 mC. The electrical stimulation was carried out during eight months, and at least twice a day. The electrode-tissue interfaces showed a steady increase in impedance during the first two months, after which the impedance remained stable; furthermore, the implanted electrodes have been proved good functionality over the time. In most cases, failure of electrodes, happened at the begining of experiments, was due to mechanical problem (breakage of the leads and male connector failure of tantalum foil). Moreover, the electrode-tissue interface proved a good contact during the implantation period. On the medical side, better voiding of bladder is obtained by a high-frequency signal ranging between 20 and 200 Hz, modulated by a low-frequency signal of 1-2 Hz.

In designing the second group of electrodes, we have focussed on the multi-stranded stainless steel wires (316L, Cooner Wire Co.) due to their flexibility, the hardness of these wire materials, the passive compatibility with tissue, and the ability to pass biphasic electrical current without corrosion [5] . All of theses properties make the multi-stranded stainless steel wires to have long working life. We have adopted the method described by Naples and Mortimer [2] in realising theses electrodes. We are working on miniaturizing and making more dedicated electrodes to our applications. Specifically, electrodes that are suitable for the stimulation of sacral-roots and pudendal nerves. A variety of electrodes sizes, diameters, and lead length with multiple configurations will be reported.

Another more advanced technique has been investigated in our laboratory to realize the electrode-tissue contact. This technique is based on the thin-film technology. The fabrication process is similar to that employed in the integrated circuits technology. A platinum material is deposited on a polymeric flexible silicone substrate. The last two groups are under tests, also new stimulation algorithms are now under investigation to better understand the behavioral bladder response to electrical stimulation. Results of this work will be published soon.

## CONCLUSION

We have described new design techniques for implantable electrodes. Preliminary *in vivo* results are also presented. Tantalum electrodes showed good functionality and stability over time. Tissues and electrodes reactions appear to be acceptable. We also developed other electrodes based on two new techniques. We are presently investigating the SMA (Shape Memory Alloy) technology and their remarkable characteristics to develop neural spiral cuff electrodes.

## ACKNOWLEDGMENTS

## REFERENCES

[1] M. Sawan, F. Duval, M. Hassouna, et al. Computerized transcutaneous control of multichannel implantable urinary prosthesis. IEEE Trans. Biomed. Eng., 39: 600-609, 1992.
[2] G. Naples, J.T. Mortimer, et al. A spiral nerve cuff e'ectrode for peripheral nerve stimulation. IEEE Trans. Biomed. Eng., 35: 905-916, 1988.
[3] G.A. May, S.A. Shamma, and R.L. White. A tantalum-on-saphire microelectrode array. IEEE Trans. Electron Dev., 26: 1932-1939, 1979.
[4] A. M. Dymond. Characteristics of the metal-tissue interface of stimulation electrodes. IEEE Trans. Biomed. Eng., 23: 274-280, 1976.
[5] A. Scheiner, J.T. Mortimer, T.P. Kicher. A study of the fatigue properties of small diameter wires used in intramuscular electrodes. Journal of biomedical materials research, 25, 589-608, 1991.

# HEATSTIM:
## An integrated hard-/software system for feedback controlled radiant heat stimulation in somatosensory research.

Jörg Breuer & Kay H. Steen
Department of Dermatology, Dermatophysiology,
University of Bonn,
D-53105 Bonn, Germany.

The development of extracellular single nerve fiber recording techniques *in vitro* in pain research which are sensitive enough to investigate even the unmyelinated C-fibers provided a direct approach in studying the excitability of nociceptive nerve endings in mammals. The hard-/software packet "HEATSTIM", which initiates a focused halogen bulb to produce a temperature slope, predicts the rise in local temperature of the examined receptive field of a single nerve fiber from a certain starting temperature $t_0$ to an end value $t_{end}$ (e.g. 32°C - 46°C) within exactly 20 seconds. The temperature range corresponds to the range of heat sensitivity of nociceptive nerve endings. The temperature rise is feedback controlled by an IBM-compatible PC with an AD/DA-Converter, which is connected to a local thermocouple via an integrated thermo-amplifier. In conclusion we present a highly accurate method to perform ramp-shaped heat stimulation for biological receptive systems in neuroscience and cell-biology research.

## 0. Introduction
Recently, *in vitro* techniques for the study of afferent nerve endings in mammals have been introduced, thereby allowing the application of test substances to the nociceptive terminals in well-controlled concentrations. We have investigated identified primary afferents in a rat skin-saphenous nerve preparation. Previous to chemical stimulation, identification and characterization of the single primary afferents were achieved using established criteria of sensory properties and of conduction velocities through electrical stimulation in the receptive fields of the unit (Lynn & Carpenter, 1982[1]; Fleischer et al., 1983[2]). The determination of sensory properties included measuring the responsiveness of the C-units to punctuate mechanical stimulation (tested by "von Frey-hairs", calibrated in millinewtons) and the sensitivity to heat and cold. Heat responsiveness and mechanical sensitivity are major criteria in characterizing a single unit and identifying the so-called "polymodal" nociceptors, the heat- and mechanosensitive and, most importantly, chemosensitive nociceptors (Perl, 1968[3]).

The role of different chemical substances in pain mediation were evaluated by increasing the concentration of these substances locally at the receptive fields of afferent C-fibers and recording the single nerve fiber-activity via a low-noise AC-coupled amplifier. The recordings were digitized (12-bit A-D-Converter, 4 kHz sampling rate) and processed in an 80386-type computer using a DAP 1200 interface card with built-in digital signal processor (Microstar labs., Redmond, WA, USA).

## 1. Components of the Hard- and Softwaresystem HEATSTIM
To identify heat sensitive fibers we developed the hard-/software packet "HEATSTIM", which initiates a focused 150-Watt halogen bulb to produce a temperature slope. This predicts the rise in local temperature of the examined receptive field from a starting temperature $t_0$ (32°C) to a final value $t_{end}$ (46°C) within exactly 20 seconds. The temperature range corresponds to the range of heat sensitivity of nociceptive nerve endings. The temperature rise is feedback-controlled via a local thermocouple and is passed through an ice point controlled integrated thermo-amplifier with an analogous connection to an input port of the AD/DA-Converter inside our IBM-compatible PC.

A cybernetical system consisting of an analogue temperature-sensor, a voltage-controlled power mate and a PC AT 286 with a built-in 12 bit AD/DA-converter provides a minimum deviation from the ideal value. Dynamic nominal and actual value comparison is implemented by programmed specifications. The input is provided by digitized incoming temperature values. The power mate output voltage is controlled by the analogue output port of the D-A-Converter; its output level is dynamically specified, in the form of 12 bit digits causing a defined adjusting voltage, by calculations dependent on proportional, differential, and integral parts of deviation. Depending on system timer and user-defined constant values, a function makes the slope-type nominal values available for immediate access in the output control routine. All values are available by a global data record, which is used for graphical presentation and data storage. An off-line variance analysis is integrated to ensure minimal adjustment deviation.

The **stimulator** is realised by an 150 watts OSRAM-type halogen bulb, which bundles the emitted light within a range of 40 mm. The maximum power emission is achieved when a voltage of 15 volts with a current of 10 amperes is led on the lamp. For our purposes a voltage reduction to 10 volts $U_{max}$ was necessary to avoid heat oversaturation of the *in vitro* preparation. In our experience the resulting power of 100 watts was efficient enough for this purpose: only about 60 per cent of the maximum capacity was needed; beyond it the life expectancy of the lamp increases considerably. To avoid mechanical damage of the bulb we had to put it into a special manufactured box with an aperture for the reflector in order to maintain free light emission in one direction. An isolated two-wire lead with a diameter of 1.5 mm² is sufficient for the connection with the power supply up to a distance of two meters (-> a).

The **power supply** consists of 220/18 V transformator (330 VA) with a regulated AC/DC-Converter. In detail, AC/DC conversion is performed by a 25A-rectifier supplying a N-channel Darlington transistor and three standard bipolar transistors with current; a 20000 µF capacity is used for noise filtering and buffering load dependent peaks. Output controlling takes place in adjusting the high impedance input of the Darlington transistor, whose emitter is connected with the base of the three parallel switched bipolar transistors, first to avoid damaging of the D-A-Converter and second to prevent causing undesirable load-dependent feedbacks on the control unit (-> b).

A temperature sensor in form of a K-type-(chromel-alumel)-**thermocouple** is applied to the *in vitro* preparation to get information about the actual local temperature. In order to obtain a high sensitivity for temperature changes we have chosen the thinnest standard couple (0.5 mm diameter) available by Philips/Germany. A K-type thermocouple produces a mean voltage of 40.44 µV/K, and the resulting voltage curve has to be linearised before analysis (-> c).

Adaptation of the thermocouple to the analogue input of the A-D-Converter requires an amplification and a linearisation of the thermo-voltage in a suitable range, dependent on the temperature range the measurement must be performed, and the input range of the A-D-Converter, where maximum resolution can be expected. In addition, the resulting thermo-voltage has to be compared with a reference temperature, since only a certain relative deviation to environmental temperature can be measured. For this purpose we used a monolithic **thermocouple amplifier** with cold junction compensation AD595, distributed by Analog Devices (Norwood, U.S.A.), which provides a low impedance output of 10 mV/°C, i.e. 25°C will cause an output voltage of 250 mV; ice-point compensation and noise suppression are realised internally. A following standard operation amplifier adjusts the resulting (AD595-)voltage to the maximum input range of the A-D-Converter input. It should be mentioned that the AD595 provides an internal set point controller, which can be used to switch a connected circuit (e.g. a heater driver) on or off. Depending on an adjacent set point

voltage and an incoming thermo-voltage, the chip toggles the temperature comparator output, whereby a simple hysteresis can be constructed easily. We did not make any use of it, because our demand of uniformity and accuracy is not reachable with this feature. Beyond it, the internal controller would be too slow and cause unacceptable noise peaks in the very sensitive isolation amplifier used for our practical application (-> d).

The hardware system used is an IBM-compatible AT 286-12 configuration with a monochrome display adapter, harddisk, 1 MB RAM and additional 12 bit AD/DA-Converter; this is the minimal hardware requirements for the integrated system described here. A resolution of 12 bit allows 4096 conditions to be distinguished; in our case a temperature range from 0°C - 50°C can be measured in 4096/50°C steps, i.e. the accuracy is about 1/82°C. Accuracy can be increased by decreasing the temperature range, but in practice it would be overlayed by incoming noise. The AD/DA-Converter used here provides 16 analog input channels (only one of them is needed) and one analog output channel. In practice the computer represents first the user interface for interaction and supervision before, after and during the experiment, controlling the connected additional hardware; second, i.e. a high grade of flexibility can be maintained, since all hardware independent features of the application - representation, feedback-controlling, supervision - are realised by programmed specification(-> e).
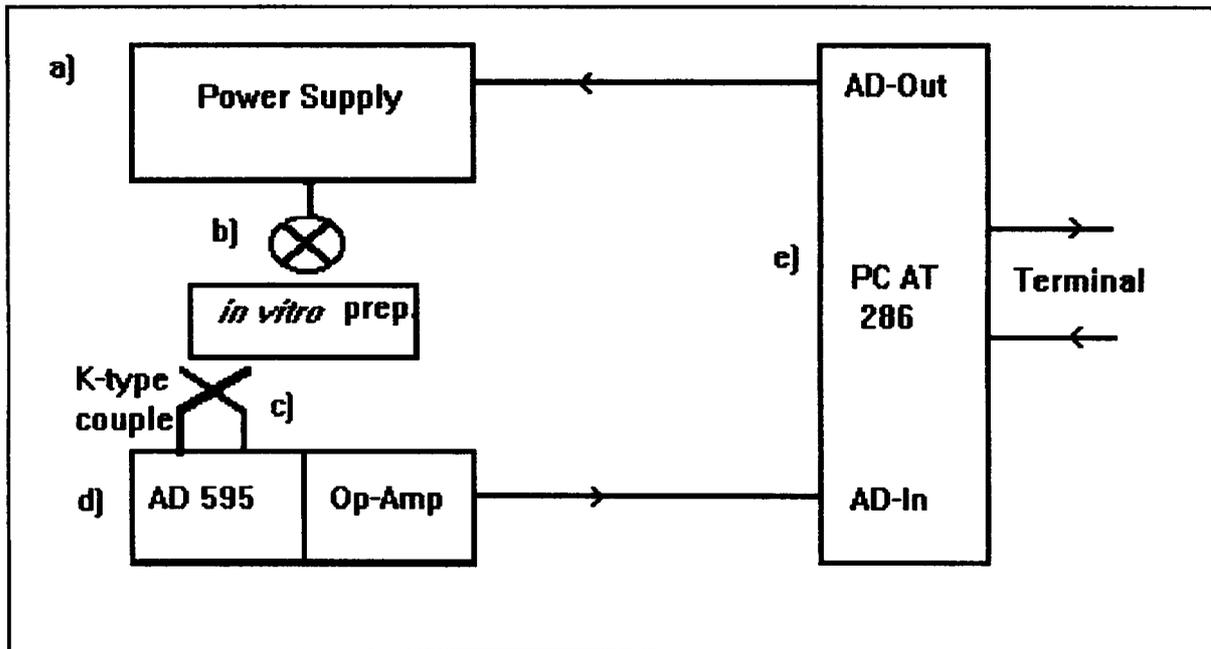


Figure 1 : Blockdiagram of stimulus device HEATSTIM.

The program is completely written in Turbo Pascal. It covers routines for user interaction, time period measurement, D/A-conversion, A/D-conversion, nominal value calculation, PID-controlling, deviation analysis, and recording of experiment data, furthermore a routine for preadjusting the start temperature of 32° C.

## 2. Discussion

The heat threshold of mechano-heat sensitive C-fibers (CMH) in primates is greater than 38°C, but less than 50°C (Meyer et al. for review [4]). For that reason HEATSTIM is arranged to maximum accuracy between 25°C and 50°C. An example of the heat-response-characteristic of a CMH is shown in figure 2. Considering other temperature-sensitive nerve fibers also with

different temperature change response-characteristics (AMH, warm fibers, cold fibers, etc.) there is a comprehensive need for heat response investigation which could be covered by an adoption of HEATSTIM. In spite of that, no ready-made stimulus device exists which allows the study of heat-response-characteristics of these fibers in variable experimental situations. In psychophysiological research in humans the availability of a highly accurate and flexible heat-stimulation- and measurement equipment for the investigation of heat-sensitivity, heat-thresholds and hyperalgesia to termal stimulation in various experimental situations (e.g. inflamed tissue) might open new research fields. Also, in cell-biology studies a heat stimulation is needed to create heat shock responses of cell cultures for the investigation of second messenger responses (e.g. $IP_3$) and "heat shock proteins".
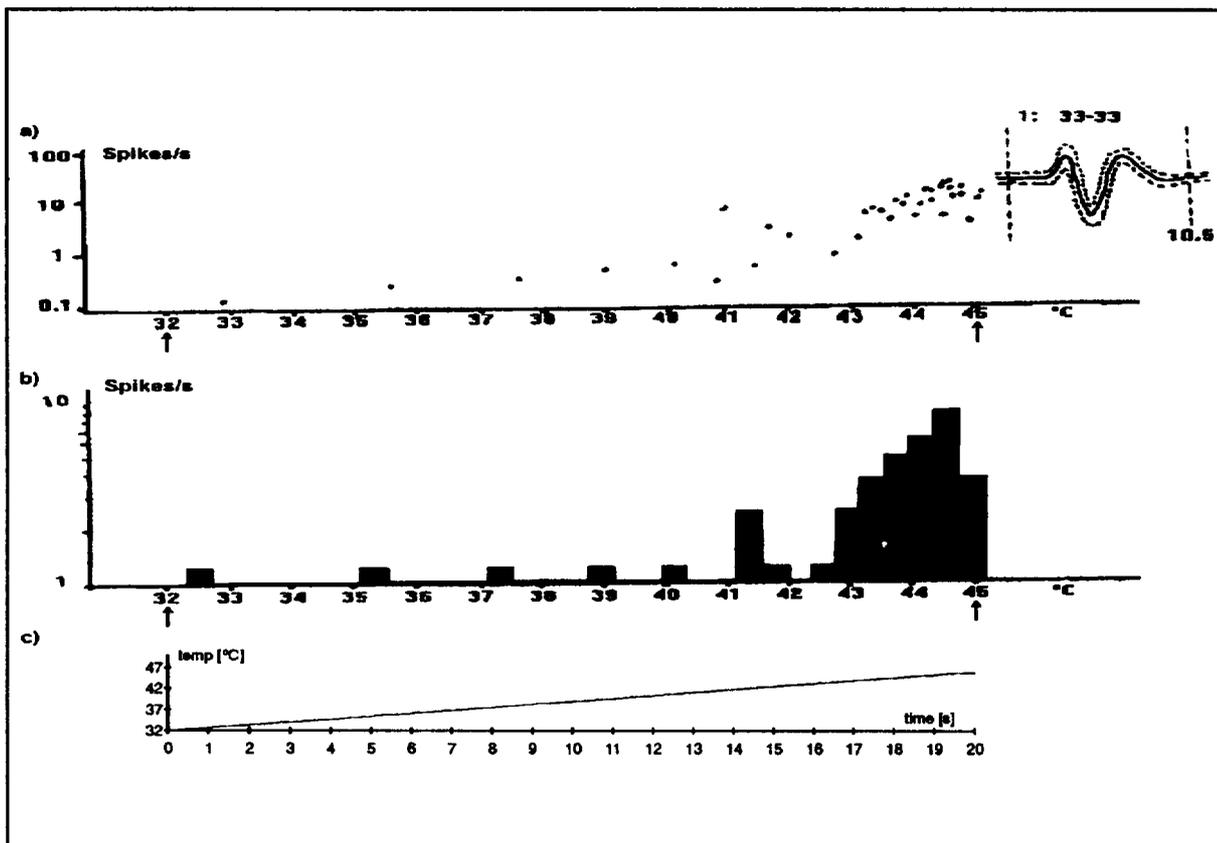


Figure 2 : a) Interval histogram
b) Poststimulus time histogram
c) Temperature slope

Figure 2 shows the evaluation of a heat response of a mechano-heat sensitive C-fiber (CMH), which was characterized by a von Frey threshold of 64 mN and a conduction velocity of 0.41 m/s. Beginning and ending of heat stimulation are marked by arrows. The first spike has been measured at 32.5 °C; a burst has been recorded at 43 °C. The interpretations have been performed by the analysis software SPIDI, a program designed for the evaluation of spike-sequences recorded by the acquisition software SPIKE; both SPIKE and SPIDI have been developed by Clemens Forster, University of Nürnberg/Erlangen, Germany.

## 3. Conclusion

In conclusion we present a highly accurate method to perform ramp-shaped heat stimulation for biological receptive systems, based on a digital feedback-controlled emitter for steady adjusted power application. This hard/software system is flexible for different laboratory needs of neuroscience and cell-biology research.

## References

[1] Lynn B, Carpenter S E : Primary afferent units from the hairy skin of the rat hind limb. Brain Res 238, 29-43, 1982

[2] Fleischer E, Handwerker H O, Joukhadar S : Unmyelinated nociceptive units in two skin areas of the rat. Brain Res 267, 81-92, 1983

[3] Perl E R : Myelinated afferent fibers innervating the primate skin and their response to noxious stimuli. J Physiol (Lond) 197, 593-615, 1968

[4] Meyer R A , Campbell J N, Raja S N : Peripheral neural mechanisms of nociception In : Textbook of Pain (3. Ed.) Wall P D, Melzac R (Eds.), Churchill Livingstone, Edinburgh, London, Madrid, Melbourne, New York and Tokyo, 1994

# PAINVAS:
## An integrated hard- and software system for pain- and simultaneous tissue pH-recording.

Osman Altindal, Kay H. Steen
Department of Dermatology, Dermatophysiology,
University of Bonn,
D-53105 Bonn, Germany.

Investigators of chronic pain in humans meet with the problem that the experimentally applied painful stimuli do not match that of a painful disease, and that pain estimated in the laboratory situation is not similar to the undulating pain experiences of patients during the day under different physical efforts. In vascular diseases (e.g. arterial occlusion in working heart or leg), pain becomes evident after physical labour and is part of diagnostic manoeuvres. To study the chronic pain of a patient in his natural daily environment, different pain recording-techniques have been introduced, which provide estimates over a longer time period, including well-established pain questionnaires and experimental electronic pain diaries. In ischemic and inflamed regions of the human body pH-levels as low as 5.4 have been measured. With the same concentration of hydrogen ions, nociceptors of rat skin were shown to be tonically-excited and sensitized. In human skin it is possible to induce constant on-going pain with continuous infusion of acid buffers. "PAINVAS", an integrated hard-/software system, works as a pain-diary with simultaneous tissue pH-recording to investigate the relationship between pain and pH changes in variable experimental situations and in chronic pain patients.

## 0. Introduction

Psychophysiological pain studies involve the investigation of the relationship between physiological data and pain reports from human beings, preferably in the same experiments at the same time. A well-known problem in the interpretation of pain research-results, especially research of clinical pain, is caused by the gap between the laboratory and the "real-life situation", that of undulating pain experiences of patients under different physical loads during the day. In patients with vascular occlusion pain becomes evident after physical labor. High hydrogen ion concentrations were found in inflammation (down to pH 5.4), in fracture-related hematomas (down to pH 4.7) and in cardiac ischemia (pH 5.7). To investigate a correlation between pain and local acidity in different pathological states, a 24-hour recording technique to measure pH via a thin needle electrode and simultaneous pain-ratings via a visual analogue scale (VAS) had to be developed. Patients should be able to behave and to work as usual. Therefore the recording equipment must be easy to carry and an impediment should be avoided.

The integrated hard-and software system "PAINVAS" makes it possible for the patient to record ambulant data. The investigator only has to transfer the recorded data to the personal computer for further graphic and statistical analysis. There are three categories of information which must be registered. The first category concerns the patient's physical status and medications. The second concerns the intensity of the pain. And the last one is the data regarding pH-measurement. We required "PAINVAS" to be able to record all three categories of information without any interruption during e.g. 24 hours, to be extendable for other technical requirements, and to be portable.

## 1. Hard-/Software components of "PAINVAS"

The "PAINVAS" consists of a Pascal program named "PVAS.EXE" which manages the recording procedures, stores all of the patient's recorded data in a file and also controls the six hardware units.

1) The pocket-personal computer Atari-Portfolio provides almost the same capability of a personal computer with an Intel 8088 compatible processor and an Operating System compatible to Microsoft Disk Operating System (MS-DOS Version 2.11). With the 128 KiloByte Random Access Memory (RAM) and 128 KiloByte exchangeable memory-card, which has its own power-supply, that of a small lithium battery, the Atari-Portfolio is the kernel of "PAINVAS". It can be connected to 220 V AC using an AC-DC-adapter (6 V) or power supplied by three mignon-type batteries (each 1.5 V).

2) A parallel interface together with a special function embedded in the operating system is used for data transfer from the pocket-computer to a printer or to other computers.

3) An A/D-converter makes it possible to access an analogue 8 bit data block which is available from a connected gauge. The data access rate is programmable using an application program on the pocket-computer up to 10,000 data blocks per hundredth second. For the pH-registration it may be adjusted

to 100 data blocks per hundredth second. The 8 bit data block can describe a value in the range of 0 to 255.

4) A portable gauge called Portamess 654 by Knick Berlin, Germany is used as the pH meter. It transfers the analogue data from an electrode that is connected to the gauge's BNC input port. With this gauge pH-values in the range of pH 0.0 to pH 14.0 can be measured to an exactness of pH 0.01 or mV-values in range of -1999 mV to 1999 mV or temprature in range of -50.0 °C to +150.0 °C. At the analogue output port data is encoded at 1 mV per digit. This gauge is power-supplied by 9 V batteries or a NiCd accumulator type IEC 6 F 22. This pH meter with a low impedance scriber output port is necessary for impedance conversion between the electrode and A/D-converter.

5) A symmetrical +/- 9 V power-supplied amplifier converts the output of the Portamess 654 to a range 0.0 V to 2.5 V, so that the digitalized mV-value is in range of 0 to 255 and fits in an 8 bit data block. The input port of this hardware unit is connected to the pH meter Portamess 654 and the output port is connected to the A/D-converter.

6) A specially-made combination pH electrode in 20 G needle [Figure 1] with an impedance of 300 MegaOhms is used to measure the concentration of hydrogen ions in tissues. It is the smallest combination sensor that requires no external reference and causes minimal tissue damage. This unit is connected to the input port of the pH meter Portamess 654 by a coax cable. The response time of this microelectrode is about 10 seconds and its stability is within pH 0.05 a day.
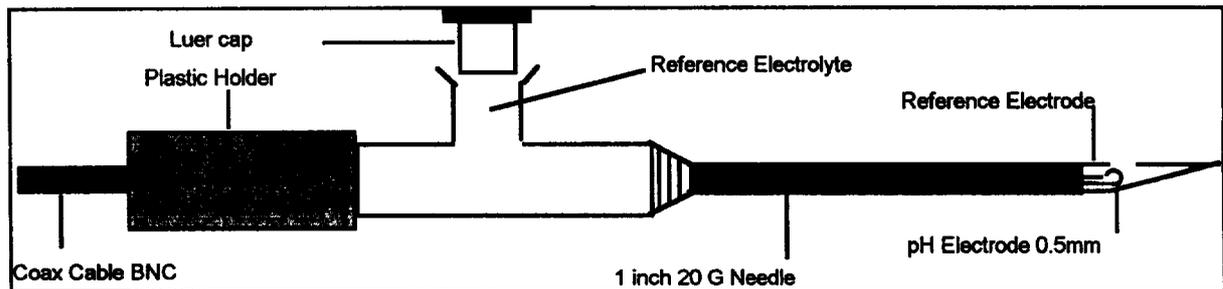


Figure 1: Combination pH electrode in 20 G needle with an internal reference electrolyte.

## 2. The program functions and their usage

The software system has three main functions: The comments or remarks describing the changes in the physical and psychological status and the medication of the patient are registered permanently, independent of the program configuration. The second function manages the recording of data concerning the current intensity of pain (pain-rating). The third function manages the recording of the pH-values in the pain-causing tissue. All data is recorded separately with a corresponding time-stamp. The behavior of the last two functions are dependent on the configuration, i.e. they can be separately activated or deactivated in the "setup"-phase of the program. We developed this system with independent units in order to provide flexibility for investigators planning to practice new methods. In this way it is possible to combine and activate the functions of interest for each session individually. Furthermore, new units can be added to this system in order to record some other kind of data. The investigator is asked to enter some personal data concerning the patient, the name of the session and the size of the interval X for the timer-event in seconds (X >= 30s) for the interactive pH-measuring and pain-rating. He can access the option to set or reset the pain-rating as well as the pH-measuring using the TABULATOR-key. Furthermore, the investigator must check the entries and acknowledge their correctness. Before the function for pH-measurement is started, the pH-electrode should be calibrated together with the pH meter. The pH-value is calculated by the program "PVAS.EXE" dependent on the constant $\triangle$ pH = pH 3,0 = $pH_{max}$ 7.0 - $pH_{min}$ 4.0 and the corresponding $\triangle$ mV which is defined exactly after the calibration. To enable the program to function, the calculation of $\triangle$ mV the $mV_{max}$ and $mV_{min}$ must be entered manually or by repeating the calibration when all the hardware components are connected together and the program asks for the corresponding mV-values. When the system is ready to start the session, the pH-needle must be introduced into the tissue that causes the pain, and the pokket computer must be fastened to the patient. Pressing the ENTER-key starts the session. The patient can always enter a remark in between the timer-events, i.e. a string of up to 20 characters, which is written with a time-stamp into a file on the pocket computer. The number of remarks is unlimited. When activated, the pH-data is recorded every X seconds interactively. The pH-measurement function reads out the mV-values of the A/D-converter buffer and calculates the corresponding pH-values every

X seconds simultaneously to the VAS-ratings and stores them with a time-stamp into the file together with the rated data. Every X seconds an acoustic signal draws the patient's attention to the pain-rating procedure. The patient is asked to rate the current value as a percentage, which describes the subjective intensity of pain, via a Visual Analogue Scale (VAS) that is monitored as a horizontal bar, ranging from pain threshold (0%) to maximum "unbearable pain" (100%) [Figure 2]. The bar moves back to zero after each rating to assure that authentic magnitude estimates were given each time. The VAS values are immediately written into an ASCII-file with a corresponding time-stamp. When the intensity does not differ from the last rated value, the last value is stored with the current time-stamp automatically after half of X seconds are passed.

When data becomes available, it is immediately written into an ASCII-file on the memory-card; therefore the risk of data-loss due to power-supply failure is diminished.
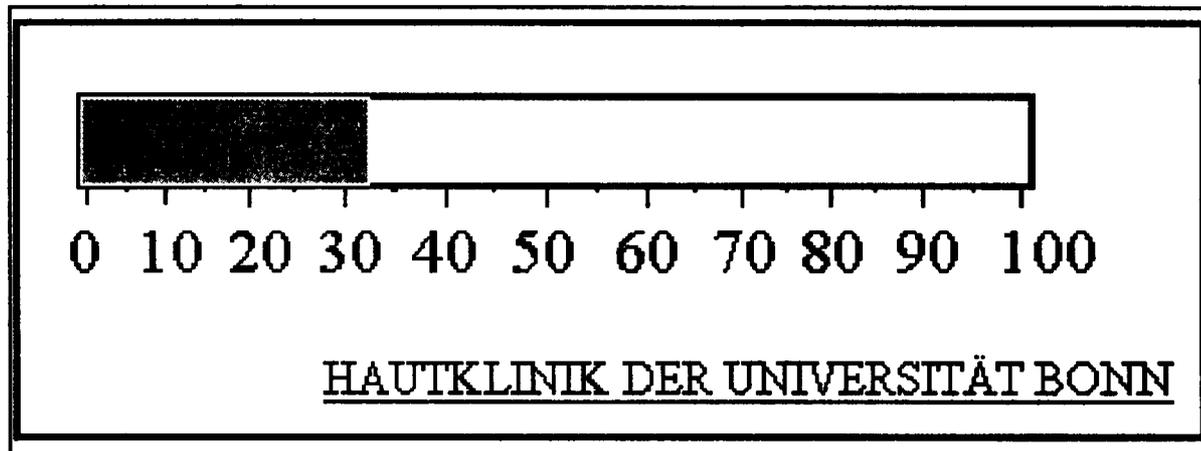


Figure 2: The patient can enter the intensity of the pain by positioning the bar on the estimated pain threshold within the range of 0% (no pain) to 100% (unbearable pain).

## 3. Calibration of the pH-electrode

At first the pH-electrode is connected to the BNC input of the Portamess 654. 1) Then the tip of the needle must be placed in pH 7.00 buffer and the pH meter must be adjusted to pH 7.00. 2) The needle must be placed in pH 4.00 buffer and the pH meter set to pH 4.00. Now the first and the second procedure must be repeated until the values are achieved without any manipulations of the pH meter. This calibration is necessary because voltage range at the output of the pH meter must correspond to the pH range from pH 4,0 to pH 7,0 , in order to increase the accuracy of the pH-value, which is calculated by a special software unit on the pocket computer.

## 4. Format of the recorded data

All the recorded data is stored in one file on external memory-card. The name of that file (*YYMMDDNN*.DAT) is automatically constructed by the software system itself using the current date, the first letter of the patient's name and the first letter of the patient's surname followed by the extension ".DAT". Before the software system creates a new file using that name, it checks whether a file with that newly calculated name already exists. In this way it recognizes when a session with the same patient has been interrupted before and appends all the data from the new session to that file, in order to maintain the continuation of the data. The first six lines contain the data, which has been entered by the investigator when he initialized the session. That data block is n·eded for the management of the sessions. Lines which contain remarks begin with a time-stamp followed by the string "#oa#" to distinguish the remark lines from the other lines. For the rest of the file each non-remark line begins with a time-stamp followed by the corresponding pain-rating data and ends with the corresponding pH-data. Blanks are used to separate data [Figure 3]. The information in that file is totally encoded in ASCII on a Disk Operating System (DOS) to enable the data to be easily transfered to a stationary computer, where a lot of standard or special programs can be used to analyse this accumulated data statistically or even to make graphical presentations very quickly and very easily. It is not necessary for the investigator to learn the usage of more programs than needed. To make our system

as effective as possible for usage in laboratories we have developed a hard-/software system where the accumulation of data and its analysis are strictly separated from each other [Figure 3, 4].

```
28- 1-1994    5:24:10
NAME
SURNAME
Birthday
Born in
5:24:27   #oa#1. REMARK TEXT
5:24:46   0.00     7.4
5:25:16   5.50     7.39
5:25:33   #oa#2. REMARK TEXT
5:25:46   30.00    7.37
5:26:16   42.00    7.38
```

Figure 3: These are the first 11 lines of the file that is named "942801NS.DAT" automatically. All the text is encoded in ASCII-code. The program's start date and time is in the 1st line; it differs from the session's start time in the 7th line. The comments are to identify with the string "#oa" at the beginning of each remark. The pH-values are periodically registered together with the pain-rating value. Any kind of a standard analysis program can be used to work out this datafile [Figure 4].  <---

Figure 4: This graphical presentation of the pH-values and the pain-ratings at a time is made using a standard program (MS-Excel 4.0). A motorized syringe pump was used to infuse a phosphate buffered solution of pH 5.2 into the palmar forearm skin of human subjects [Figure 5], during a period of about 40 minutes. A constant ongoing pain was induced immediately and it took 5 minutes to measure values down to pH 5.2. After the infusion was stopped, the skin-tissue pH-values began to increase again.    --->
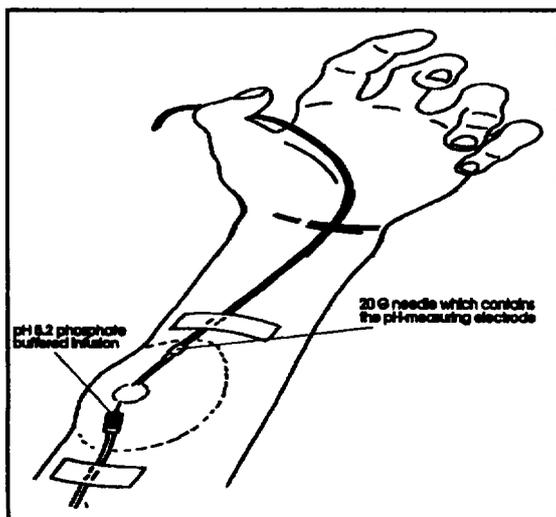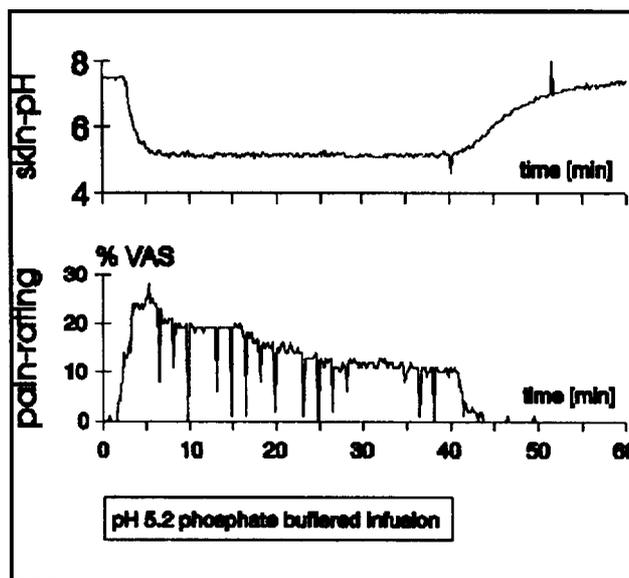


Figure 5:   The 20 G needle sends the pH- status in the skin tissue to the pH-meter, while the smaller needle, which is driven by a motorized syringe, infuses continuously a phosphate buffered solution of pH 5.2 into the skin of a human subject.
<---



## 5. File transfer from the Atari-Portfolio to a stationary computer

There are some transfer procedures already embedded in the Basic Input Output System of the Atari-Portfolio, therefore the investigator must only have an equivalent program on his stationary computer. Before beginning any transfer the AC/DC-converter on the pocket computer must be exchanged for the parallel interface unit and the computers have to be linked by a parallel cable from port to port.

When the transfer programs have been started on both sides, the transfer procedures begin to communicate using the so called hardware -handshake protocol. When the stationary computer is chosen as server, the investigator must enter the filename to be transferred.

# 6. Discussion

## 6.1 Visual Analogue Scale

The electronic pain diary presented here is based on a Visual Analogue Scale (VAS) [1]. This scale has proven to be used effectively in hospital clinics and provides valuable information about pain and analgesia (e.g. effect of analgesic drugs). The VAS consist of a horizontal line with the points 0% (=no pain) up to 100% (=unbearable pain). The patient is required to place a mark at a point which corresponds to the level of pain intensity presently felt [Figure 2]. The advantage of the VAS compared to other pain measurement tools is the implied equality of ratios, its ease and brevity of administration and scoring, minimal intrusiveness and its conceptual simplicity [2]. These advantages are even more important in a "non-laboratory situation" where the concentration of the patient cannot be fixed solely on the experimental procedures. The easy handling of the technical equipment and the scale is essential for the older patients. The main disadvantage of a pure VAS is that the endless variety of qualities of pain are not considered. To specify the qualities different pain questionnaires have been developed - the most established one is the "McGill Pain Questionnaire" [2]. In the presented "PAINVAS"-Software a free recording of pain quality and description of momentary activity is possible including the documentation of time of the record.

## 6.2 pH-measurement

The cause of clinical pain is still not completely understood. Elevated hydrogen ion concentrations have been repeatedly suggested as an important factor of inflammatory pain [3,4]. Indeed, acid pH was found in inflamed and ischemic tissues (see introduction). Psychophysiologists evaluated the role of pH in experimental pain but still were not able to demonstrate low tissue pH-values during painful sensations of a chronic disease or during increasing pain in exercise situation. The collection of physiological data is needed to put findings of *in vitro*-investigations together [3] and to evaluate significance of these findings.

The simultaneous recording of pH-values and pain-rating in order to correlate them both is one possibility to use the Hard-/Software "PAINVAS" demonstrated here. The investigations of other correlations between physiological data (e.g. $pCO_2$-, $pO_2$-concentration) and sensory data (e.g. pain, itch, loudness, light intensity) are also conceivable.

# 7. Conclusion

We conclude that, with "PainVAS", it is possible to record data on a Visual Analogue Scale of ambulant patients for a long time period and any kind of analogue data signals simultaneously. With the use of an interface link it is possible to transfer the created file to any personal computer easily.

# References

[1] Huskisson E.C., Visual Analogue Scales. In: Melzack R. (ed) Pain measurement and assessment. Raven Press, New York, page 33-37, 1983.

[2] Melzack R. & Katz J., Pain measurement in persons in pain. In: Textbook of Pain (3. Ed) Wall PD, Melzack R. (Eds.), Churchill Livingstone, Edinburgh, London, Madrid, Melbourne, New York and Tokyo, 1994.

[3] Steen K.H., Reeh P.W., Anton F., Handwerker H.O.. Protons selectively induce long lasting excitation and sensitization to mechanical stimulation of nociceptors in rat skin, *in vitro*. J. Neurosci 12, 86-95, 1992.

[4] Steen K.H., Reeh P.W.. Sustained graded pain and hyperalgesia from harmless experimental tissue acidosis in human skin. Neurosci Lett 154, 113-16, 1993.

# Luminescence Studies on Phosphor Screens

G. Panayiotakis, C. Nomikos, A. Bakas and B. Proimos
Medical Physics Department, University of Patras
265 00 Patras, Greece

## Abstract

We report our results on x-ray phosphor screens prepared of some new materials focusing attention on their efficiency under fluoroscopy conditions, on optimization conditions and on comparisons among the various materials. All data are presented in absolute values. A theoretical model is presented,that takes into account the granular structure of the screens, permitting the explanation and prediction of the luminescence properties of the screens.

## Introduction

Phosphor screens are used in a vast number of x-ray medical applications and their properties are closely related to the dose received by the patient as well as to the quality of the image produced. During the last years rare earth oxysulphide screens became of great interest [1,2,3]. In this paper we present the results of a study concerning the luminescence of screens prepared of some rare earth oxysulphide and other phosphor materials. Apart from the detailed experimental data, comparisons are given among the various materials and theoretical calculations supporting the data and permitting the estimation of the efficiency parameters are presented.

## Materials and Methods

The phosphor materials studied were $YVO_4:Eu$, $La_2O_2S:Tb$ and $CdPO_3Cl:Mn$. The phosphor screens necessary for the experiments were prepared by sedimentation on fused silica substrates.

Phosphor screens were excited to luminescence both in transmission and in reflection mode observation. In the transmission mode, the x-ray beam incidents on the one side of the screen and the light emitted is detected from the other side, while in the reflection mode the x-ray beam incidents on the one side of the screen and the light emitted is detected from the same side.

The x-ray beam for exciting the screens was supplied by a classical radiotherapy unit with a tungsten anode tube (tube voltage ranging from 50kV to 250kV peak). Since almost in all cases the measurements have been carried out under fluoroscopy conditions [4], the x-ray beam has been filtered by an aluminum filter 20mm thick, which is considered to simulate the patient's body.

The exposure rate at the screens position was measured by repalacing the screen with the measuring chamber of a dosemeter and the light emitted from the screen was measured by a photomultiplier (having an S-20 photocathode). In order from the photomultiplier's output current to calculate the total light flux emitted from the screen, we had to take into account the respective position, shape and size of the screen and the photocathode as well as the spatial distribution of the light emitted from the screen [5]. We have also taken into account the matching factor between the screen's emission and the detectors sensitivity [6] which compensates for the color difference among the various phosphors.

From these exposure rate and photomultiplier current values we have calculated the efficiency of the various phosphor screens in absolute units $[W/m^2]/[mR/s]$. The various errors arising during these efficiency measurements and calculations are in most cases of the order of 5% or even less.

### Results and Discussion

We present now briefly the most important results of our study. The efficiency of the screens proved to depend on the x-ray beam properties as well as on the screen properties and ofcourse on the phosphor material. Considering the dependence of the luminescence on the beam and screen properties we remark the following:

a) The efficiency depends on the tube voltage increasing initially with it, reaching a maximum and decreasing slowly thereafter for all phosphor materials, all screens and both modes of observation. Typical examples are shown in figure 1.

b) The efficiency does not appear to depend on the tube current at least up to an exposure rate of about 800R/s (absence of saturation effects).

c) The efficiency strongly depends on the screen's thickness. This dependence is different in the case of transmission and in the case of reflection mode observation. As shown in figures 2 and 3, in transmission mode observation for all kinds of phosphors the efficiency shows an initial increase with the screen's thickness, it reaches a maximum and decreases slowly therafter, while in reflection mode observation the efficiency increases continuously with the screens thickness towards a saturation value.

The voltage value at which the efficiency peak appears in the n=f(V) curves depends also on the screen's thickness, increasing linearly with it; the effect is more pronounced in the case of transmission mode observation. The thickness value at which the efficiency peak appears in the n=f(w) curves depends strongly on the phosphor material, been in general higher for high mean atomic number materials.

A detailed explanation of all the above experimental data is given in the following chapter.

### Theoretical Calculations

According to the theoretical model introduced by Hamaker [7] and improved by Ludwig [8] the absolute efficiency of a fluorescent screen (homogeneous fluorescent layer) supported by a partially transparent and reflecting substrate is given, in the case monochromatic x-ray beam excitation and transmission mode observation, by the relation:
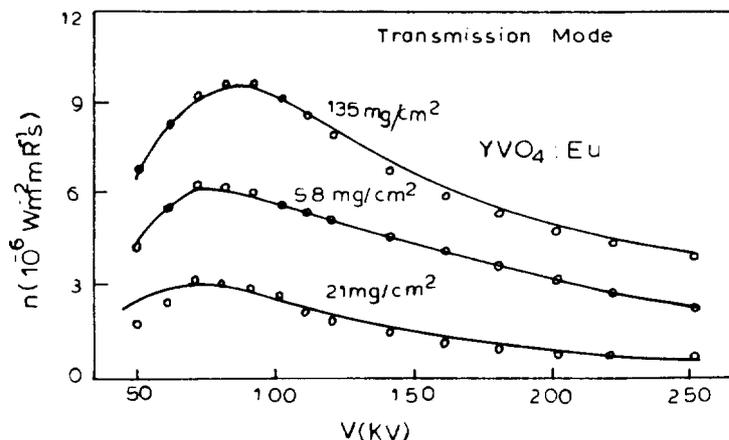


Figure 1:    Absolute efficiency of x-ray screens versus the x-ray tube voltage.
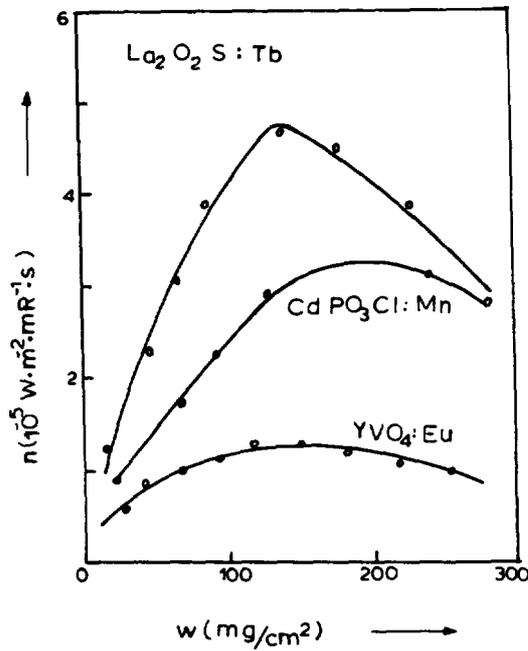
Figure 2: Absolute efficiency of x-ray screens versus the screen's thickness in transmission mode observation.
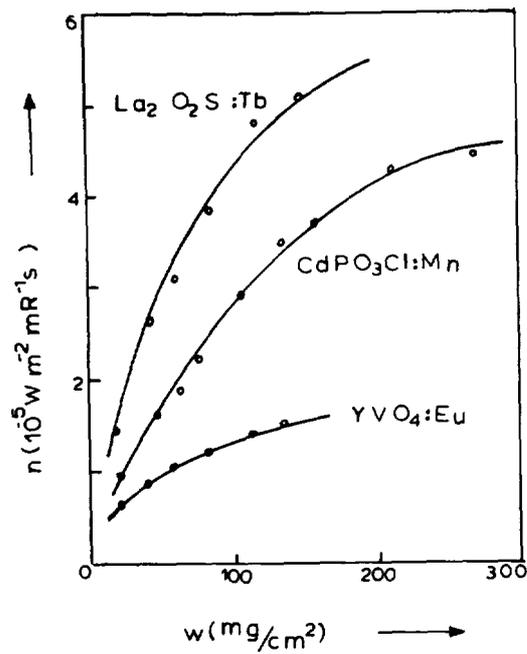


Figure 3: Absolute efficiency of x-ray screens versus the screen's thickness reflection mode observation.

254

$$\eta \;=\; \frac{n_c \gamma t \mu (1+\varrho) \exp(-\mu w)}{2(\mu^2 - \sigma^2)}$$

$$\frac{(\mu-\sigma)(1-\beta)\exp(-\sigma w)+2(\sigma+\mu\beta)\exp(\mu w)-(\mu+\sigma)(1+\beta)\exp(\sigma w)}{(1+\beta)(\varrho+\beta)\exp(\sigma w)-(1-\beta)(\varrho-\beta)\exp(-\sigma w)}$$

(1)

where
$\varrho = (1-r)/(1+r)$

$\sigma = \sqrt{a(a+2s)}$ and $\beta = \sqrt{a/(a+2s)}$

r    reflectivity of the screen's substrate

nc    intrinsic efficiency coefficient of the phosphor

t    transparency of the screen's substrate

w    screen's weight (thickness)

γ    convertion factor converting the exposure rate to energy fluence

s    light scattering coefficient of the phosphor material

a    light absorption coefficient of the phosphor material

μ    mass absorption coefficient of the phosphor material

A similar relation holds for the reflection mode observation [9]. Since the x-rays in our case were not monochromatic we have used for the spectral distribution function of the x-ray beam after the filtration necessary for simulating fluoroscopy conditions the relation [10]:

$$f(E) = (1-E/E_0)\exp(-\mu_{Al} d) \qquad (2)$$

where d is the total Al filter thickness (22mm) and Eo is the peak energy of the x-ray photons determined by the tube anode voltage.

As a result the final expression for the efficiency under a continuous spectrum x-ray beam excitation has been calculated [11] equal to:

$$n = \int_0^{E_o} n(E)\, f(E)\, dE \Big/ \int_c^{E_o} f(E)\, dE \qquad (3)$$

where n(E) is the relation for transmission or for reflection mode observation respectively.

The values of $\mu$ for the various phosphor materials have been evaluated from the values given by Saloman et al [12] for the various elements. The values of $t$ and $\beta$ have been calculated from transmissivity and reflectivity measurements according to Ludwig [8]. The coefficient $\sigma$ and the intrinsic efficiency $n_c$ have been estimated by fitting relation (3) to our experimental data.

Using the above values for these parameters we have calculated the absolute efficiency of the screens according to relation (3). The corresponding theoretical results are presented in the various figures in the form of continuous lines (experimental data are presented as dots); the agreement between theory and experiment is obviously very good and in most cases within the magnitude of the experimental error.

## References

1   Stevels A L N 1975 Med. Mundi 20 12.

2   Giakoumakis G E, Nomicos C D and Sandilos P X 1989 Phys. Med. Biol. 34   673.

3   Schwarz L, Gurvich S, Golovkova S I, Starick D and Herzog G 1993 Nucl. Tracks Radiat. Meas. 21 35.

4   Holm T and Moseley R 1964 Radiology 82 898.

5   Giakoumakis G E and Nomicos C D 1985 Phys. Med. Biol. 30 993.

6   Giakoumakis G E 1993 Appl. Phys. A52 7.

7   Hamaker H 1947 Philips Res. Rep. 2 55.

8   Ludwig G W 1971 J. Electrochem. Soc. 118 1152.

9   Giakoumakis G E, Nomicos C D, Yiakoumakis E N and Evangelou E K 1990 Phys. Med. Biol. 35 1017.

10  Storm E 1972 Calculated bremmstrahlung spectra from thick tungsten targets Phys. Rev. A5 2328.

11  Giakoumakis G E and Nomicos C D 1985 J. Appl. Phys. 58 2742.

12  Saloman E B, Hubbell J H and Scofield J H 1988 Atomic Data and Nuclear Data Tables 38 1.

# THE DECADE OF THE BRAIN : CREATIVITY IN SCIENCE UNDERSTANDING

By: Paolo Manzelli, President of ECO-CREA International Institute, c/o Laboratory of Educational Research, Department of Chemistry , University of Florence Italy

## Abstract:

The "Decade of the Brain" characterise the need to improve a very deep change in science understanding.

In truth the brain capability is not inscribed into the traditional paradigm of science, because, from the Rene' Descartes times (XVII century), science believe in the splitting between " res extensa" and "res cogitans; i.e. the encoding strategies of the brain are not related with the evolution of knowledge. But at this time we need to think up that the modalities of the appearance of reality that we perceive and know, is not only a consequence of the nature of external objects and processes.

For traditional science understanding the internalisation of information, by means the construction of brain frames, is a not scientific issue, owing the separation between two cultures ; consequently the objectivity in science is merely considered as a naive mirror image of that we perceive, without any attention about the complex action of recognition and information processing of the brain.

Nowadays the International Institute ECO-CREA will develop a new paradigm of science inclusive of the brain relationships among Energy, Information and Matter, with the aim to improve creativity in contemporary science understanding.

# THE DECADE OF THE BRAIN : CREATIVITY IN SCIENCE UNDERSTANDING

By: Paolo Manzelli, President of ECO-CREA International Institute, c/o Laboratory of Educational Research, Department of Chemistry , University of Florence Italy

The "Decade of the Brain" characterise the need to improve a very deep change in science understanding.

In truth the brain capability is not inscribed into the traditional paradigm of science, because, from the Rene' Descartes times (XVII century), science believe in the splitting between " res extensa" and "res cogitans; i.e. the encoding strategies of the brain are not related with the evolution of knowledge. But at this time we need to think up that the modalities of the appearance of reality that we perceive and know, Is not only a consequence of the nature of external objects and processes.

For traditional science understanding the internalisation of information, by means the construction of brain frames, is a not scientific issue, owing the separation between two cultures ; consequently the objectivity in science is merely considered as a naive mirror image of that we perceive, without any attention about the complex action of recognition and information processing of the brain.

Hence a big challenge is open in science inside of the "Decade of the Brain" (1990-2000), because depending from our traditional mechanic approximation the thinking brain is to be considered as a black-box.

The cutting operated by the mechanic viewpoint of science is that the information is not defined as a function of the energy / matter interactions, because any explanation can be exclusively calculated as a result of forces and motion of particles; hence within this traditional approach the information processing of the brain need to be correlate to mechanical interaction of particles. The above general approach yield a collection of machinery analogies for approaching the thinking process in terms of molecular recognition ( like the key -lock opening mechanism) that are clearly inadequate to express any one idea about the sensible and logic functions of brain communication and its evolution in nature.

Moreover the analogy originally proposed by Shannon between "Information and Entropy" in thermodynamics, meets a deep conceptual restriction, because information is only related to the disorder generated by the entropy of the system and this approach can be only useful to find the better conditions to fit a diminution of the noise of information's transmission. In addition the concept of negative entropy (neg-entropy) associated to the concept of order that is coming out some considerations about the emergence of order from within chaos during a molecular transformation, can be only defined as a qualitative statement.

Really considering that information cannot derived from mechanical point of view , and referring to the need to understand the brain thinking function, we need to develop a conceptual context in which information can be clearly defined as a function of energy / matter interactions. The above means that we need to assume constructive creative approach to build up a new paradigm of science understanding.

Starting with the idea to construct a new paradigm, we can consider the following reasoning pattern: free-energy and matter can be seen as two forms of energy with different categories of codex of information. Consequently the interactions between energy and matter can be seen and understood as a process of energy dissipation, working for de-codify and codify new forms of energy / matter yielding a transformation.
Hence in the context of the above interpretation, Information process is equivalent to the energy dissipative activity that any energy / matter interaction requires to spend giving rise to a transformation.
Because any process of change in the universe can be thought in the same manner we are allowed to introduce a new general variable that is associated with the last fraction of total energy . In such a way we define the new paradigm of science referring to three fundamental variables i.e. Energy, Information and Matter. (< E / I / M > paradigm)

For advancing in this approach, we apply the following procedure to construct the basic conception of <E/I/M> paradigm, in which information is now proposed as a fundamental variable linked to a fraction of the total energy involved in each energy / matter interaction.

In the beginning, we start selecting the most essential postulate of science ; i.e. Energy cannot be created or destroyed.

A logical consequent of this postulate , the total Energy must be a constant, equal to say "One" at any time.
As a result of the previous assertion, we note that if we considered the information as a parameter of the general description of the energy-matter transformation, the global variation (d) of the different aspect of energy , obtained from the sum of the free-Energy (Ef), and the codified-Energy like Matter - (Em), and also (Ei), then also Energy dissipated into the information process, (including that we spend for our brain understanding) , must be equal to zero. Certainly a variation of a constant is zero for definition.

In synthesis considering :

$$E \text{ total} = (Ef) + (Em) + (Ei) = 1$$

$$\{ \text{ i. e. } \} / \text{ at any time } / \ d\,[\,(Ef) + (Em) + (Ei)\,] = 0$$

$$\text{Therefore :} \quad + d\,(Ei) = - d(Ef) - d(Em) \quad \dashrightarrow \ ( \ I° \ )$$

The last conclusion is very important to understand the role of information in nature.

In fact from formula ( I° ) now we know that the energy linked to information processing is growing at any time in which we observe a lowering of the forms of energy codified as free-energy and matter.
Henceforth the evolution of each process of energy / matter interaction, can be seen and understood as a program in which nature progressively transforms Energy and Matter to develop an increase of Energy linked to the process of information (Ei).

Such increase of +d(Ei), we can indicate as the principle of "fertile evolution", that correspond to dissipate ( - Ef) and / or decrease ( - Em); the above means that the principle that regulating the increase of Energy linked to the information's procedures of natural evolution, is complementary to the equivalent principle, of the "minimum of action", settled by J.B. d'Alembert, that governs the motion of particles.

Brain intelligent functions born from the "fertile evolution" principle.

This line of reasoning permits to us to argue that, for a "complete" theory of physics, that would include the information processing that is evolving in nature with the aim to think about to the brain functioning; the required condition is to renew the mechanic paradigm of traditional science understanding in a way that each element and every event of physical reality would be described within a new strategy of thought, alike to the proposed theoretical context that we indicate as <E/I/M> paradigm.

Therefore our choice to search for a conceptual procedure inserting the brain as an object of the complete reality need to be developed to find an exhaustive interpretation of modern science understanding inclusive of the brain functioning.

The above short approach is proposed to put in evidence the importance of the challenge of the Decade of The Brain. The ECO-CREA International Institute with the above starting exposition of a general constructive creative methodology of science, call for help to scientists from all over the world for an important collaboration finalised to understand the brain thinking functions and its creative natural evolution.

Finally the ECO-CREA Institute will emphasise that the challenge of the Decade of the Brain is very significant for developing creativity in science. Really science until now is only able to conceive and thus to produce machinery , without understanding life and its clever evolution. But now is so easy to forecast that going forward with the mechanic model of traditional science thinking evidently we risk to destroy life on our planet.
The pure science is polluting the world and now we gain the urgency to research for an ecology of mind.

# The Design of a Text Database for Biomedical Information (FINE DataBase)*

*Krotopoulou A.,Lafazanis M., Panagopoulou G., Sirmakessis S., Spirakis P., Tambakas V., Terpou D., Tsakalidis A.*
*Department of Computer Engineering and Informatics*
*University of Patras*
*26500 Patras, Greece*

*Vlahopoulos P.*
*Institute of Biomedical Technology*
*Ellinos Stratiotou 50A*
*26441 Patras, Greece*

## Abstract

FINE DataBase (FDB) is a document database that keeps, in a structural way, information suitable for effective communication and co-operation among Clinical Engineers. The majority of this information is free texts with biomedical interest. An electronic magazine is also provided to the users, while additional data concerning medical devices, the user themselves and Clinical Engineers all around the Europe can be found in FDB. The goal of organising the above information is to assure the dissemination and the effective use of the gained experience as well as to strengthen the collaboration among Clinical Engineers. This paper gives a short description of the FDB design, which may serve as a framework for other specialised databases which deal with free texts.

## 1. Introduction

Nowadays a great amount of information relative to Clinical Engineers is produced by different organisations and institutes. Information concerning medical devices such as failures, malfunctions, solutions to technical problems, reports from European or other international organisations, quality control standards etc. is provided by many resources. Furthermore, announcements for new products, seminar calls for participation and bibliographic notes are distributed from different means. Every Clinical Engineer interested in these topics can be informed by reading relative magazines, conference proceedings or by communicating by means of every source of the above information.

As it can easily be seen, a lot of valuable time is spent in search of this kind of information. FINE DataBase (FDB) can reduce the time spent, because it is a tool that contains all this kind of information and supports the exchange of ideas in the Clinical Engineering area.

FINE DataBase is a tool that is being developed under the Biomedical Equipment Assessment and Management (BEAM) AIM project. It is part of the Facility for INformation Exchange (FINE) tool. FINE is a BEAM sub-system aiming to develop a prototype conferring tool to be used for information exchange in the area of Biomedical Technology. The basic system's principles is based on the need for communication's support among Clinical Engineers and other professionals. Its objective is to provide to the Clinical Engineering Community a central source of services, information and news over the existing Public Data Networks or Public Switched Telephone Networks. Therefore, Clinical Engineers even though they are physically isolated, they can form a distributed working environment.

This paper is an overview of the FINE DataBase design. In section 2, there is a brief presentation of the information stored into FDB. Section 3 contains a description of the design and architecture of the database. The final section summarises our conclusions for the work done till now and gives in short our future plans.

---

# 2.    An Overview of the FINE DataBase

Fine DataBase is a document database that keeps in a structural way free texts with Biomedical interest. According to the needs of Clinical Engineers, various kinds of free texts need to be supported. Consequently, texts have to be distinguished into categories and this has to be made in a way that guarantees the effective communication and co-operation among them ([9]).

According to a widely accepted consideration, free texts can be divided into *reports* and *announcements*. Reports, in their turn, can be distinguished into *official* and *general content* reports. Official reports are well constructed and approved documents provided by organisations or any authorised person. These reports are addressing issues related to accident alerts, quality control protocols and maintenance for any kind of medical devices included in FINE DataBase. Moreover EC Directives and International Standards can be found. On the other hand, general content reports are free texts provided by the FINE DataBase users, as a result of the interaction among them, reflecting their experience related to the use of Biomedical equipment. These reports can be relative to any kind of information that Clinical Engineers consider that is of interest to their community. General content reports can be distinguished in a way similar to the official ones.

As far as announcements concerns, it has to be pointed out that they are actually working as real life announcements, while they are provided in order to encourage the communication and co-operation among FINE DataBase users. Following the idea of effective division of texts into categories, seven kinds of announcements have been specified: announcements for bibliography, new products, meetings, seminars, expositions, job openings and collaborations. Any new book or product can be presented in the corresponding announcement while meetings, seminars or expositions are announced. Job openings and themes for collaboration can also be propound ([9], [10]).

The existence of an *electronic magazine* was found to be essential for the promotion of information exchange, which is the basic goal of FINE system. Electronic magazine actually mimics the structure of "real" magazines. In other words, it is edited in issues, while each issue is formed by articles being free texts. Each issue can be distributed to every subscriber who can be any FINE DataBase user.

During the specification of FDB user requirements, the need of having additional information which is not structured as texts was detected. This information mainly refers to FINE DataBase users and Biomedical devices. As far as FDB users concerns, information like their names, addresses and other contact information, or more specific one, like their specialities, is kept.

# 3.    The FDB Design

In order to achieve efficient storing and manipulation of FDB data -especially of free texts- special techniques ([3], [6]) have been considered. The main FDB design task was the achievement of a proper organisation of free texts. This has been firstly succeeded by the efficient use of basic data as indices for the documents. In other words, special basic data have been selected to represent the *basic keywords* of the texts, according to which the retrieval of the texts is performed. Basic keywords are relative to the content of the texts and have been carefully detected, while different basic keywords are defined for different text kinds. The number of basic keywords has such a value that both efficient representation of the content of the text is achieved and information redundancy is avoided. To have an idea about how the basic keywords are formally defined for every text kind, we give the representative example of the definition of the basic keywords of accident alert reports. They are indicated by the following entity set:

*Define entity set accident_alert (text_id, model_id, failure, event_date, insertion_date)*

The attribute "text_id" indicates the code of the report, while the rest attributes define the basic keywords which indicate the medical device ("model_id") on which the certain failure

("failure") took place on the defined date ("event date"). Finally, "insertion_date" indicates the date that the report was inserted in the FINE DataBase.

The values of the basic keywords for every free text can be defined manually (the user actually types them when he inserts the text).

In a second step, in order to achieve efficient organisation of the free texts we gave to the user the ability of specifying some additional keywords that he assumes that indicate the basic concepts touched on the text. They are called *free keywords*. Furthermore, the subject of the text is kept. The formal definition of free keywords and subject of accident alert texts is indicated by the following expression:

*Define entity set acal_text (text_id, (keywords)\*, subject, file_name)*

The attribute "file_name" indicates the full path of the file containing the text. This is a special technique which was adopted in order to override the limitations of the traditional RDBMS's concerning free texts. More precisely, texts would not be stored as tuples of the FINE DataBase. Instead, only the filename of the file containing the text is saved. In this way we can handle documents of any size, containing any kind of information (images, graphics, pure text) ([1], [7], [15]).

Additionally, information retrieval techniques similar to ([12]), can be integrated with the FDB design for the efficient indexing and retrieval of free texts. The use of Vector Space Model ([13]), enriched with statistical methods for automatic indexing, and the Global Thesaurus, is expected to increase both the recall and the precision of the database. Especially for the Thesaurus we are going to support the system with a Global Thesaurus with automatic construction based on Document Clustering due to the connected components evaluation method ([4], [5]).

The way we confront the structural organisation of free texts makes the representation of electronic magazine an easy task, since every issue is a collection of articles (free texts). The entity-set which indicates the organisation of electronic magazine is the following one:

*Define entity set electronic_magazine (issue_id, issue_date, (subject, file_name)\*)*

As far as non text-structured information concerns, no representation anomalies occur. A representative entity-set expression of this category of FDB information is given below:

*Define entity set device_model (model_id, model_name, manufacturer, model_description)*

The only point that has to be mentioned here, is that this information is closely related to the content of free texts, as it can be easily seen in the "accident_alert" entity set. All the entity sets and relationships that organise FDB information, are presented in the FDB entity-relationship diagram (Figure 1).

Another issue that has to be mentioned has to do with security aspects in the design of FDB. Due to the fact that texts can contain anything and users are allowed to insert any new text, the database should have mechanisms to protect itself from having garbage or irrelevant information. Every text is marked with the "id" of the user who inserted it. This is indicated by the following relationship:

*Define relationship user_texts (user_id, text_id)*

‘ If something is wrong or irrelevant, using the above relationship, any available information about the author can be retrieved. Furthermore, any text can be retrieved by any user having this right, but it can be modified or be deleted only by the author. This effort is controlled by the same relationship.

Access rights to database functions are defined by the assignment of a user into one or more of four security groups. Different groups are granted with different access rights. Crucial operations like deletions are provided only to authorised users. Every user is assigned into one or more groups by inserting his unique "user_id" into the tuples of the "group" entity set.

266
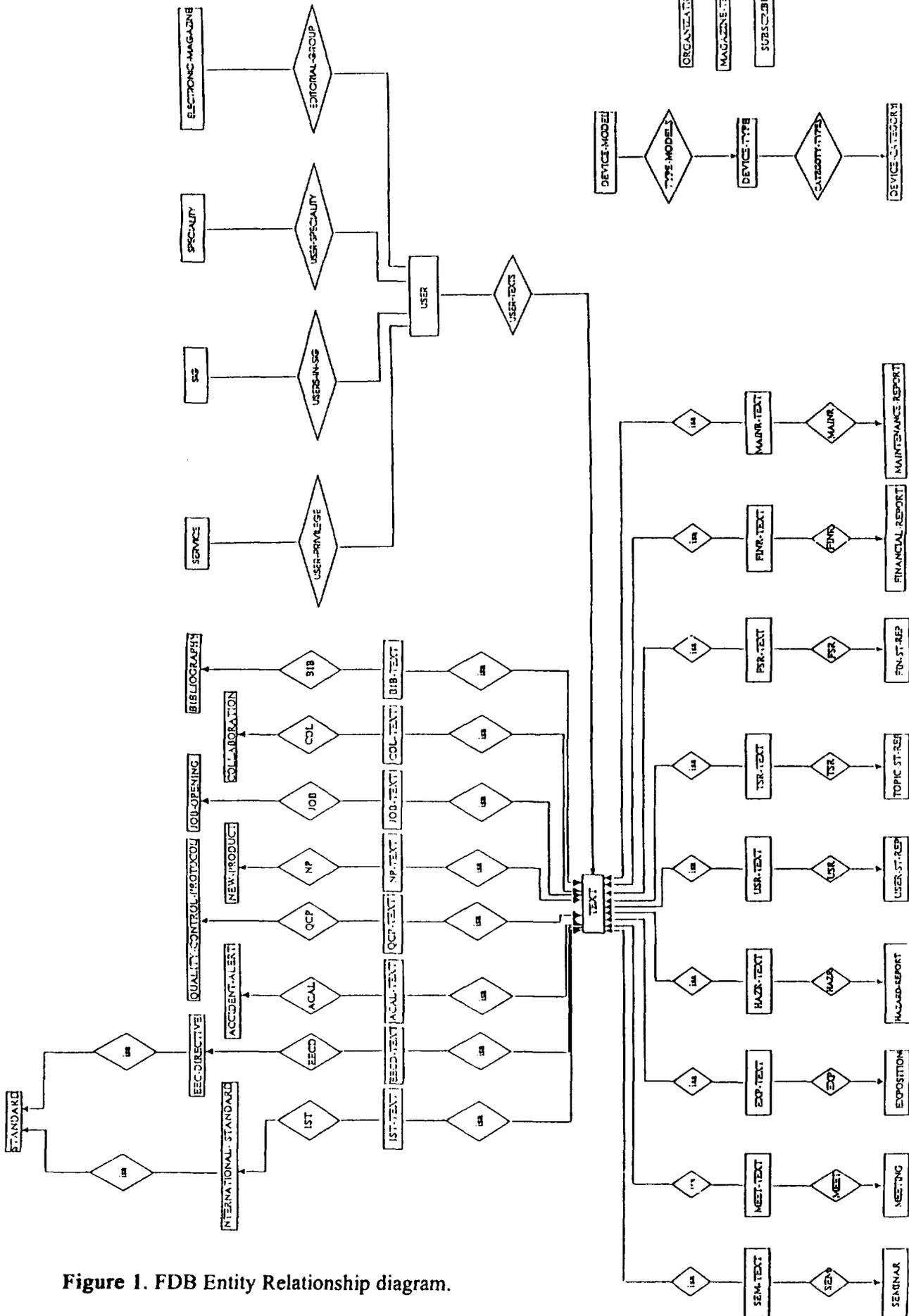
**Figure 1**. FDB Entity Relationship diagram.

## 3.  Conclusion

FINE Database has been designed in a modular way to ensure that future extensibility can be done. For this reason only a few modifications are needed to design a distributed architecture of the system in order to implement a flexible tool for retrieving up-to-date information from any remote database, which is actually some of our future work.

Additionally we are going to investigate the use of document standards for the architecture (e.g. ODA, CDA), interchange (e.g. ODIF, DDIF) and text mark-up (e.g. SGML) in order to support, the maintenance of the revisable form of the FDB documents, the exchange of documents among the nodes of the systems and the visualisation of documents in its true form.

Although the design of a document database has been conformed by other projects, individual requirements, such as efficiency, multi-tiered security, integrity and data independence, can make different approaches considerable. Because of the high level of interest with which database research has focused on representation of free texts, we hope that the proposed modelling techniques may find broader applicability in various relative cases.

## References

[1]  Amur R., Mohan A., Ramaswami M., '"TMS: A Free Form Text Management System", Software Practice and Experience, 1990, Vol. 20, No. 3, pp. 321-324.

[2]  Anick P.G., Flynn R.A., "Integrating a Dynamic Lexicon with Dynamic Full-Text Retrieval System", SIGIR 1993, pp 136-145.

[3]  Clifton C., "The design of a document database", ACM Conference on Document Processing Systems, 1988, pp. 125-134.

[4]  Fowler H.R., Fowler A.L.W., Wilson A.B, "Integrating Query, Thesaurus, and Documents through a Common Visual Representation", in proceedings of SIGIR 1991, pp. 142-151.

[5]  Fuhr N., "Integration of Probabilistic fact and text retrieval", in proceedings of SIGIR 1992 pp. 211-221.

[6]  Fuller M., Mackie E., Sacks-Davis R., Wilkinson R., "Coherent Answers for a Large Structured Document Collection", in proceedings of SIGIR 1993, pp 204-213.

[7]  Jasinski P., Meinzer H., Sandblad B., "An Approach to Integrate Collections of Images and Relational Databases", Methods of Information in Medicine, 1988, Vol. 27, No. 4, pp. 177-183.

[8]  Korth E., and Silberschatz A., "Database system concepts", McGraw-Hill, University of Atlas at Austin, 1986.

[9]  Krotopoulou A., Lafazanis M., Panagopoulou G., Sirmakessis S., Spirakis P., Tambakas V., Terpou D., Tsakalidis A., "Design of the S/W components of the FINE system and system configuration", Chapter 5, Deliverable #13, AIM/BEAM, FINE no A2001, December 1992.

[10] Krotopoulou A., Lafazanis M., Panagopoulou G., Sirmakessis S., Spirakis P., Tambakas V., Terpou D., Tsakalidis A., "FINE: The Pilot Phase Demonstrator", Chapter 5, Deliverable #21, AIM/BEAM, FINE no A2001, October 1993.

[11] Krotopoulou A., Panagopoulou G., Sirmakessis S., Spirakis P., Tambakas V., Terpou D., Tsakalidis A., Pallikarakis N., "FINE DataBase: A Database for Information Exchange Among Clinical Engineers", to appear in Twelfth International Congress on Medical Informatics, MIE 1994, Lisbon.

[12] Lafazanis M., Mamalis V., Spirakis P., Tambakas V., Tsakalidis A., "FIRE: An efficient Information Retrieval System for Large Text Collections", T.R. 94.01.5, Computer Technology Institute, Patras, Greece.

[13] Salton G., "Automatic text indexing using complex identifiers", ACM Conference on Document Processing Systems, 1988, pp. 135-144.

[14] Soren V., "SQL and Relational Databases", Microtrend publication, 1991.

[15] Stonebraker M., Stettner H., Lynn N., Kalash J., Guttman A., "Document Processing in a Relational Database System". The INGRES Papers, 1983, Wiley.

[16] Walker J. H., "The role of modularity in document authoring systems", ACM Conference on Document Processing Systems, 1988, pp. 117-124.

# A query system in an integrated medical application environment

Demetrios Paparoditis[1], Eckart Fleck[1], Bernd Mahr[2]

[1]Department of Cardiology, German Heart Institute Berlin, FR Germany
[2] Department of Computer Science, Technical Universtity Berlin, FR Germany

email: dimi@dhzb.de

## Abstract

This paper presents our current efforts towards the development of a medical query system to be applied in our clinic and whose main principles are based on the exploitation of medical conceptual knowledge. Considering the heterogeneity and distribution of the information sources available inhouse we developed a specific system architecture suited to these conditions. Physicians can, using their terminology, access in a transparent and uniform manner existing information stored in different medical documentation systems. The "query model" is based on empirical results showing that physicians' questions in a clinical problem context investigating patient care are instances of predefined question types. Building upon this experience we introduced the concept of Medical Query Pattern (MQP). Analysing the medical terms and their interrelationships by extracting medical concept information and domain knowledge from existing medical knowledge sources, the system selects the most appropriate MQP. Associated with each MQP is a set of specific frames providing information to support the query formulation, the query interpretation, the selection of the appropriate information source and finally the synthesis and the presentation of the query results.

## 1. Introduction

The issue of developing intelligent user interfaces for query purposes is subject of current and future research in medical informatics towards the efforts to fulfill the physician's information needs[1]. If physicians are going to obtain information (patient and non-patient related) that is important to them for effective patient care, they must be able to query different medical information sources and expect to receive informative answers. Their working context is overloaded with a huge amount of digital information, both in-house and external, stored in different dedicated medical databases. Accordingly they are often faced with the problem of localizing the appropriate information source. Even if this problem is overcome, the next obstacle is to find the "right" terms to describe the inquiries in order to extract the desired piece of information.

Different attempts have been made towards the development of "intelligent" systems in order to overcome querying difficulties appearing in computerized medical environments e.g. due to the lack of comfort of user interfaces[2], the geographical distribution of the information sources[3], the heterogeneity in the structure of the databases[4], resulting in more or less sophisticated systems. Techniques applied in those systems have their roots in the artificial intelligence, the data base and the information retrieval area.

The proposed approach to overcome the above difficulties is based on the idea that physicians' inquiries in a clinical environment are instances of question types [5][6]. Expanding this idea we introduce the notion of the medical query pattern, which is a central concept in developing our integrated medical query system called IMQS. Medical domain knowledge and consideration of medical terminology is supported through the incorporation of two knowledge sources, the semantic network[7] and the Metathesaurus[8] of the Unified Medical Language System UMLS[1].

The paper is organized as follows: In section 2 we elaborate on the idea of question types by examining the elements of physicians' queries. Section 3 introduces the concept of MQP and manifests its representation. In section 4 we present the architecture of the IMQS, discussing

its functional components. Finally in section 5 we give some concluding remarks and outline the direction of our future activities.

## 2. Elements of physician's queries and derivation of query types

It is obvious that the development of a query system designed to support physicians in our clinic to access the available medical information sources, presumes, besides the analysis of the physician's problem context and the physician's role in this context, additional knowledge acquisition with regard to how physicians articulate their inquiries[9][10].

We have examined fifty physicians' questions after interviewing five physicians of the cardiological department and "on-line" observations of physicians' work investigating patient care in the cardiological out-patients department. We found that inquiries concerning medical information (patient related and non-patient related) are formulated combining natural language with medical language elements (e.g. show the hemodynamics registered by the heart catheterization). The "backbone" of the inquiries, the medical expression, is a composition of medical concepts describing entities like "hemodynamics", "heart catheterisation" and relationships indicating links between them like "registered by", "prescribed". This "backbone" is clothed with lexical items of natural language as verbs (e.g. is, show), question nominal words (e.g. where, what), conjunctions, articles, quantifiers, etc.. Additional "cosmetic" elements (e.g. please) were identified, whose elimination does not affect the query interpretation.

We avoided categorizing the selected material according to different points of view such as grammatical (e.g. what-questions, who-questions)[11], logical-semantic (e.g. goal orientation questions, quantification questions)[12][13] or according to the type of the information needed[14] (e.g. medical knowledge, factual information from the patient record etc.). Instead of that we gave emphasis to a "deeper" examination of the "backbone" of physicians' inquiries, seeking question "types", representing "similar" questions, from which query patterns could then be derived. The observation of the selected material has shown that for physicians investigating patient care in the cardiological out-patients department it was significant that most of the questions posed were directed to the patient's medical record, retrieving information only from those portions of the record that were relevant to a specific current clinical problem or relevant to support future decision-making. Fourteen question types were derived e.g. "optimal timing for invasive/non-invasive therapeutic procedure", "trend of parameter(s) since invasive/non-invasive therapeutic procedure", "trend of parameter(s) since invasive/non-invasive diagnostic procedure", "parameter(s) registered by invasive/non-invasive diagnostic procedure", "effectiveness of invasive/non-invasive therapeutic procedure", "report of invasive/non invasive diagnostic procedure".

Building upon our empirical results and on results of other researchergroups examining physicians inquiries in other problem contexts[5][15], we are convinced that by extending the corpus of question types, we will be able to cover the demand of needed information in relevant clinic contexts, which can be answered from the available digital information sources in the clinic.

## 3. Medical Query Patterns

Motivated by the fact that query types form an abstraction of physicians' queries and from the hypothesis that if a query matched with a particular query type then this should tell us something about how to process this query, we introduce the concept of Medical Query Pattern (MQP). Each query type ist associated with a MQP. A MQP is described using a frame structure consisting of the following four components:

(a) representation (b) specification (c) result synthesis (d) presentation

The aim of the structure is to accommodate a clear conceptual representation of the query type and to organise all the essential knowledge required to guide both the physician during query formulation/reformulation and the system during the different query processing phases.

The "*representation*" component encapsulates the semantic representation of a query type. For its representation, medical terminological knowledge must be considered, encompassing knowledge on the medical conceptual model and knowledge on single medical concepts. The medical conceptual model describes the domain of discourse at an abstract level using medical concept classes and binary hierarchical and non hierarchical relationships between them. Modelling entities such as concept classes and relationships can be thought of as elements of a uniform "meta language" which can be used to represent a query type. According to this representation a query type forms exactly one consistent view on the medical conceptual model (see Fig. 1). For the representation of query types we use the semantic network of the UMLS which purports a conceptual taxonomy of the biomedical domain in medical concept categories and a set of permissible relationships. Using the network formalism i.e. the semantic types and the semantic relationships as elements of the "meta language" we apply them to represent query types. For example the query type "parameter(s) registered by invasive/non-invasive diagnostic procedure" using the semantic types <Organ or Tissue Function> and <Diagnostic Procedure> and the semantic relationship <measured_by> can be represented as <Organ or Tissue Function> <measured_by> <Diagnostic Procedure>. The instantiation of this representation using the medical concepts "Hemodynamics", "Heart Catheterization" from the UMLS Metathesaurus corresponds to the permissible query "show the hemodynamics measured by the heart catheterization".
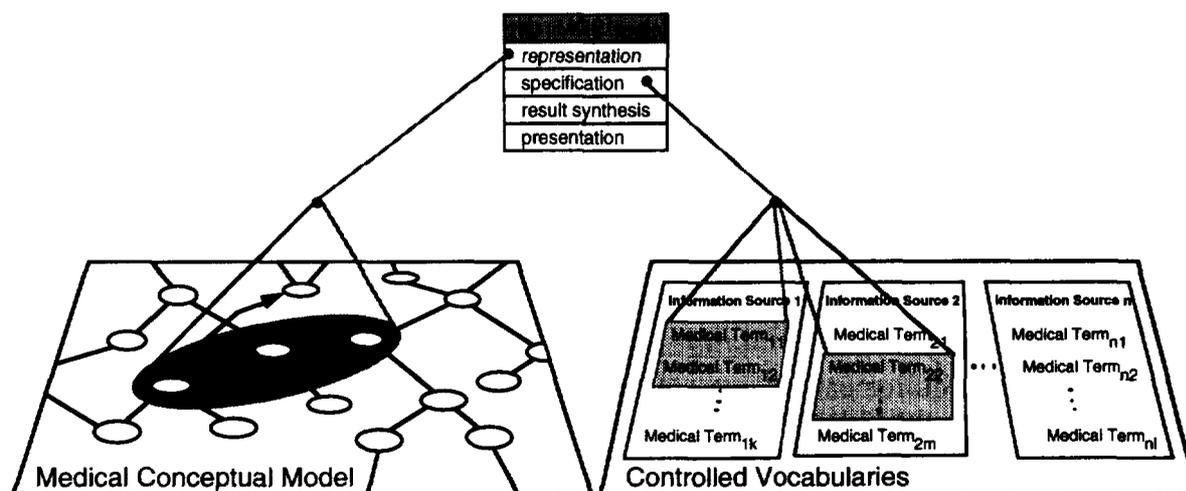


**Fig. 1:** Structure of a medical query pattern

The "*specification*" component indicates whenever the system requires more accurate specification of the physician's initial query before its submission for execution.

Consider an example where a physician using the query system poses the query "outcome of revascularisation treatment on patient X" which can be answered from a dedicated knowledge base system. In order to execute the query, the knowledge base system requires a more accurate specification of the query in terms of parameters describing the patient status e.g. weight of patient, length of stenosis etc. If the query system just "passes through" the initial query statement to the knowledge base system for execution then the execution will fail.

The purpose of the "specification" component is therefore to prevent the query system from submitting syntactically and semantically correct queries which are going to fail due to insufficient parametrisation of the initial query.

In cases where further query parametrisation is necessary, the component provides a formal specification of a parameter list holding the missing information. Elements of the parameter list are expressions of the form <parameter name> <operator> <parameter value>. We distinguish further between single-value and multi-value parameters. Logical connections between expressions and values of multi-value parameters can be specified. Parameter

instantiation can be done in two modes: the "interactive" and the "interconnectivity" modus. The "interactive" mode indicates that after initial query formulation and selection of a corresponding MQP the physician must be requested to instantiate the parameters listed in the "specification" component. The "interconnectivity" mode indicates that the query system itself will have to select the required information, by first accessing the information sources for data selection and then directing the "complete" query i.e. the initial physician's query statement extended with the selected parameters, to the dedicated information source for execution. Which information sources have to be queried is specified in the parameter list in form of "references" pointing to entries of the controlled vocabularies of the information sources (see Fig. 1). The incorporation of standardized controlled medical vocabularies for both parameter names and parameter values reduces problems by mapping operations during query processing.

The often distributed localisation of the information sources, which is a typical situation in a computerized medical environment, implies that to answer the physician's inquiry the system has to query different information sources and merge the results. The "*result synthesis*" component that encapsulates knowledge about how an answer to a user query has to be combined after initial query decomposition and subsequent execution. In order to combine the intermediate results and produce the final answer the query system has to apply this knowledge. Knowledge assigned to this component can further be applied in cases where already during description of a MQP it is known that none of the existing information source(s) can "furnish" a complete answer to the physician's query. They are, however, able to deliver, at most, intermediate results, from which an answer can then be derived. Instead of expanding the functionality of the external information source(s), procedural knowledge (e.g. inferences) can be assigned to the component in order to expand the functionality of the query system for answer completion.

The "*presentation*" component describes how end-results have to be presented to the physician. Query results as responses to physicians' inquiries can have different presentation types e.g. simple string (type of stenosis), table (results of labor examinations), graphic (trend of cardiac output), image (X-ray) etc.. Furthermore, a combination of different result types as answers to a physicians' queries are possible.

## 4. System Architecture

Considering the requirements of querying information in a medical application environment and having in mind the concepts mentioned above we are currently developing a query system based on the architecture shown in Fig. 2. In order to minimize the figure's complexity, we eliminate the explicit "use" relation between two functional components belonging to different groups (shaded shapes) and place it on the level of the groups.

The *query editor* component is responsible for the user interaction during the query formulation/reformulation, directs the user input to the analyser and is further responsible for result presentation.

For each medical knowledge base there is an associated server, handling the access to the medical knowledge base to which it is assigned, providing basic services e.g. browsing functions and functions for the selection of the best matching MQP(s) to a specific user query. The *query manager* handles the interaction between the components of the query processor group. The *analyser* is responsible for scanning the user query, to localise the medical concepts and the concept relationships, and to perform term standardisation. Furthermore it calls services from the *MQP-server* to obtain the best matching MQP(s) to the user's query. In order to overcome ambiguity problems when the number of the selected MQPs are greater than one, matched MQPs are instantiated using the standardized medical terms from the initial user query and presented to the user for selection. The selected MQP is then called the active MQP and its description is decisive for the further query processing. The analyser checks the specification component of the active MQP to determine if further query parametrisation is necessary.

The task of the *selector-decomposer* is to apply convergence mechanisms in order to narrow down the space of the accessible information sources to only those which are able to provide

272

an answer to the query. The selection process is supported from knowledge extracted from the active MQP (from its representation and specification component), from the medical concepts
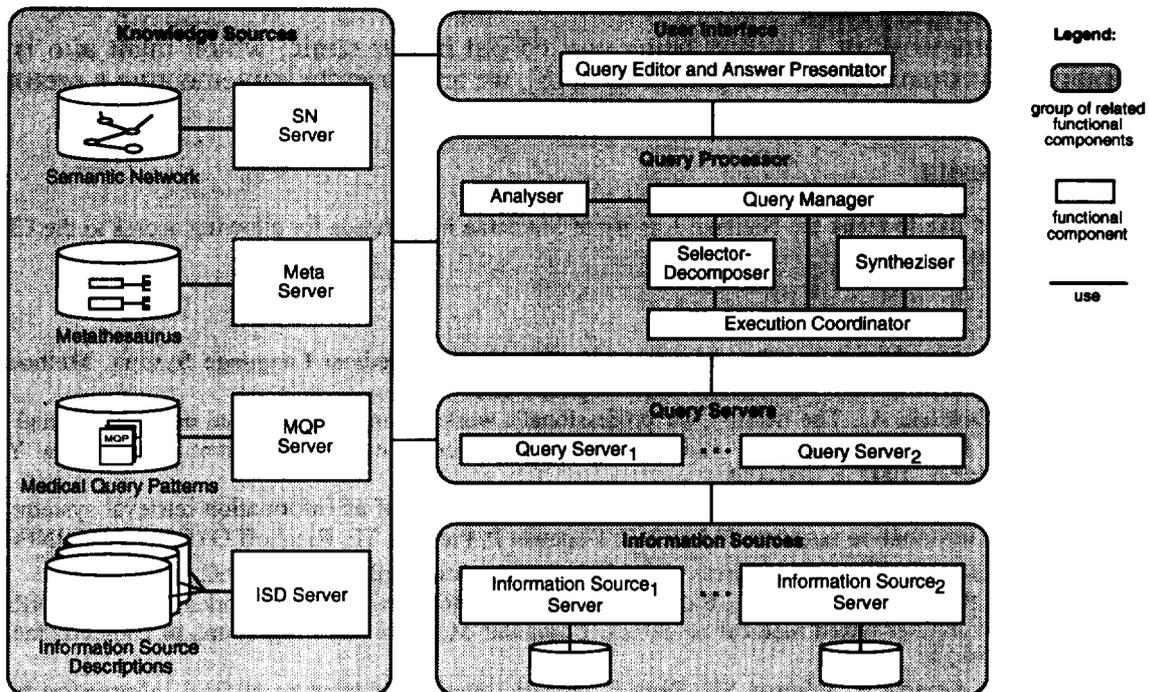


**Fig. 2:** Functional components of the IMQS

and relations of the user query and from the descriptions of the information sources. Remaining ambiguities can then be resolved by user consultation. When more than one information source must be accessed in order to answer the query, the component performs query decomposition. The results of the query decomposition annotated with information about information source names are then passed to the *execution coordinator* who dispatches the query(s), monitors and synchronises its/their execution. Based on the intermediate result(s) collected from the *execution coordinator* and extracting knowledge from the active MQP's components, the *synthesiser* builds the final answer and converts it for presentation.

With each information source a *query server* is associated. It is responsible for mapping operations, for translation of a source query in statement(s) of the target query language, for query execution, and finally for result standardisation. In order to complete its tasks a query server exploits knowledge from the descriptions of the information sources and the Metathesaurus.

## 5. Conclusion

In this paper we highlighted a "top-down" approach for the development of an integrated medical query system. Beginning with the medical terms in the physician's query and supported by both the knowledge concentrated in the associated MQP and medical terminological knowledge the system provides answers to a large number of questions arising in the context of the care setting of the cardiological out-patients department.

Our future investigations are going to concern the clinical evaluation of the semantic network for the representation of query types, covering physicians' inquiries in the cardiological out-patients department.

The description of MQP's, the construction of "complex" MQP's based on "atomic" MQP's and the possibility of chaining MQP's in order to model some physicians' tasks is going to be a further area of study.

Finally for the user interface we plan to support three possibilities of query entry (a) by allowing restricted natural language query statements (b) by typing in single medical terms or by selecting them from the current physician's problem context e.g. during "reading" the computerized patient record and (c) by allowing the choice of a query type and then driving

the physician to construct incrementally a query instance by filling the "gaps" with legitimate medical terms. Alternatives (a) and (b) can also be exploited for evaluation and expansion of the corpus of the query types.

Within the framework of a project being carried out in our clinic, which main aim is the integration and communication of patient data[16], we are currently implementing a prototype of the IMQS.

## Acknowledgments

## References

[1]   Lindberg DAB, Humphreys BL, McCray AT. The Unified Medical Language System, Methods of Information in Medicine, Vol. 32, No. 4, 1993, p. 281-291.

[2]   Silva JS, Zawilski AJ. The health Care professional's workstation: its functional components and user impact. In Ball MJ, Collen MF (Eds.): Aspects of the computer-based patient record, New York, Springer-Verlag, 1992, p. 102-124.

[3]   Nelson SJ, Sherertz DD, Tuttle MS. Issues in the development of an information retrieval system: The physician's information assistant. In Lun KC, Degoulet P, Piemme TE, Rienhoff O (Eds.): MEDINFO 92, proceedings of the 7th world congress on medical informatics, North-Holland, 1992, p. 371-375.

[4]   Cimino C, Barnett GO, Hassan L, Blewett DR, Piggins J. Interactive Query Workstation: Standardizing access to computer-based medical resources. Computer Methods and Programms in Biomedicine, 35, 1991, p. 293-299.

[5]   Cimino JJ, Aguirre A, Johnson SB, Peng P. Generic queries for meeting clinical information needs. Bulletin of the Medical Library Association, 1993, 81(2), p. 195-206.

[6]   Lang KL, Graesser AG, Dumais ST, Kilman D. Question asking in human-computer interfaces. In Lauer TW, Peacock E, Graesser AC (Eds.): Questions and Information Systems, Lawrence Erlbaum Associates, Publishers, New Jersey, 1992, p. 131-165.

[7]   McCray AT, Hole WT. The scope and the structure of the first version of the UMLS Semantic Network. In Miller RA (Ed.): Proceedings of the 14th annual symposium on computer applications in medical care, IEEE Computer Society Press, Washington, DC, 1990, p. 126-130.

[8]   Tuttle MS, Nelson SJ, Fuller LF, Sherertz DD, Erlbaum MS, Sperzel WD, Olson NE, Suarez-Munist ON. The semantic foundations of the UMLS Metathesaurus. In Lun KC, Degoulet P, Piemme TE, Rienhoff O (Eds.): MEDINFO 92, proceedings of the 7th world congress on medical informatics, North-Holland, 1992, p. 1506-1511.

[9]   Sager N. Computer Processing of Narrative Information. In Sager N, Friedman C, Lyman MS (Eds.): Medical Language Processing-Computer management of narrative data, Reading (Mass.), Addison-Wesley, 1987.

[10]  Cimino C, Barnett GO. Analysis of physician questions in an ambulatory care setting. In Kingsland LC (Eds.): Proceedings of the 15th annual symposium on computer applications in medical care, IEEE Computer Society Press, Washington, DC, 1991, p. 995-999.

[11]  Ulrike Reiner. Anfragesprachen für Informationssysteme, Deutsche Gesellschaft für Dokumentation, Germany, 1991.

[12]  Lehnert WG, The process of question answering, Lawrence Erlbaum Associates, Publishers, New Jersey, 1978.

[13]  Graesser AC, Person N, Huber J. Mechanisms that generate Questions. In: Lauer TW, Peacock E, Graesser AC (Eds.): Questions and Information Systems, Lawrence Erlbaum Associates, Publishers, New Jersey, 1992, p. 167-187.

[14]  Osheroff JA, Forsythe DE, Buchanan BG, Bankowitz RA, Blumenfeld BH, Miller RA. Physicians' Information Needs: Analysis of questions posed during clinical teaching, Ann. Intern. Med., 1991, 114(7), p. 576-81.

[15]  Rada R, Barlow J, Zanstra P, de Vries Robbe P, Bijstra D, Potharst J. Expertext for medical care and literature retrieval. Artificial Intelligence in Medicine 2, Elsevier Science Publishers, 1990, p. 341-355.

[16]  Fleck E, Hansen H, Mahr B, Oswald H. Integration von Techniken und Information: Diagnoseunterstützung zum Nutzen des Patienten. In Fleck E, Oswald H (Eds.): Neue Techniken und Konzepte der Diagnoseunterstützung bei Herz-Kreislauferkrankungen, Blackwell-Wissenschaft Verlag, Berlin, 1992.

# Design and Development of a Biomedical
## Equipment Management Database

by Dr Andreas Mallouppas* and
Constantinos Mardakoftas*
Biomedical Research Foundation
P.O. Box 913, Nicosia, Cyprus.

## ABSTRACT

The effective and efficient management of a hospital's equipment inventory (medical and plant) is of paramount importance if a high quality health care delivery system is to be achieved. The need for an overall policy and technical infrastructure at hospital as well as national levels is discussed. Particular focus is given to the characteristics of an equipment management database, which is designed to be applicable to the level of technical and other expertise available in developing country health systems.

## INTRODUCTION

It has long been recognised [2],[3] by many, including the World Health Organisation, that for developing countries to reach a level of expertise by which they can fully utilise and obtain the benefits of technology in health delivery they will need to strengthen their existing health care technical service capabilities and expertise. Otherwise the great influx of modern technology in all types of medical and hospital equipment will not bring the benefits and efficiency that is experienced by developed nations. On the contrary it will impose further problems on an already weak technical and economic situation. Technology has solved problems and has made faster and cheaper a lot of procedures in all sectors, however it has also brought with it new difficulties that need to be addressed and solved. Problems of safety, equipment standards, good operating practices, maintenance and repair of equipment, are facing many ministries of health, traditionally weak in technical expertise. Many developing countries are finding that even simple equipment such as ECG's (Electro Cardiographs) are sometimes not functioning or are under utilised due to inexperienced operators or inadequate maintenance and repair. When it comes to very large and expensive items of inventory such as MRI's, CT's, specialised intensive care equipment and a lot of essential life support systems, then the need for efficient management of equipment becomes very obvious. In many cases it affects life or death situations.

The data base developed and discussed here is of absolutely
no value if it is not to be utilised by experienced,
trained, qualified, technical personnel who will have the
capability to apply safety, good operating practices and
equipment standards and other pertinent issues. This has to
be emphasised since at many decision making levels the idea
still prevails that if the equipment and tools are bought
(and the data base is a tool) then the problems are solved.
What is forgotten or ignored is that the human factor is the
most difficult resource in any organisation to secure and
the most expensive, which takes up the biggest effort, time
and training in order to develop it. Therefore adequate
manpower training, career prospects and service
infrastructure are a prerequisite to effective management of
equipment, if any of the above measures are to succeed.


A national health delivery system cannot offer the desired
level of health care unless all its medical facilities (both
public and private) are managed properly and have the
required technical support. Thus having one hospital,
usually a major hospital in the capital, at some acceptable
level of equipment and technical management, does not imply
that the national health system as a whole is functioning
correctly. Hence the need for a management system to be
easily accessible to all parts of the health delivery
system.


## EQUIPMENT MANAGEMENT DATA BASE

A manual inventory and stock control system is difficult to
use and time consuming to access and update. Moreover access
by other remote facilities is impossible and due to its
manual nature it is space and time consuming. A computerised
data base system, however, based on a PC makes the
management of the inventory fast and easily accessible, even
by remote facilities. Its updating is also very simple.

The aim of the data base is to provide the necessary
information , red by the engineers and technicians in
order to effectively plan routine preventive maintenance
procedures, stock control and cataloguing of the equipment
on "card" form. The search by equipment type, by location,
department etc is available immediately and hence a lot of
details and statistics concerning the equipment, its spares,

repairs and maintenance is quickly available. In addition information on manufacturers, equipment codes, standards etc can be incorporated into the data base. Data files in ASCII form from other data bases can also be used.

The data base language used is Clarion, version 2.1 for DOS. The integrated environment of Clarion provides a set of utilities and facilities which permit the easy development of the data base. The language itself is a blend of functions, procedures and control structures integrated with some very powerful database management features in the form of functions. It is a well structured language like Pascal. There is a screen generator that is very developer friendly. Due to the fact that there is in the language the notion of procedures and functions and thus the idea of passing parameters, a program can be very modular and contained. The global variables can be kept to a minimum. The development of a data base system using Clarion is only controlled by the user's ability to use the language. Also it should be noted that the data base engine is a proprietary one. The latest version of Clarion reads during execution time and manipulates other popular data base files, such DBase. However import and export of ASCII files is very easy. Moreover this feature can be used to produce ASCII files that can be read by packages such as Excel or LOTUS spreadsheets and thus produce graphical representations of reports and results. There also exists a writer facility that completely frees the user from the need to go to a developer to produce a report codewise in the source code of the program. Also due to the fact that the language produces an executable file, Overlays must be used where the program is vast.

## MANOMEQ (MAnagement Of Medical EQuipment)

The data base system has been given the name MANOMEQ and aims to assist the Biomedical or Clinical Engineer or Health Care Technical Service Department to effectively manage the inventory of medical and hospital plant equipment.

The data base is modular and contains information, in the form of drop-down menus representing the familiar manual "card" system. It has information concerning equipment records, equipment maintenance and repair "cards", costs, manufacturer details, retrieval of information by various fields, such as location, department, type of equipment etc,

a library of medical equipment and accessories with codes (provided by the US/ECRI[3] database), as well as a print facility. Figures 1-3 show examples of the various screens that are available in the data base.

The data base has the following features:

(i) **Equipment "Card"**: Here all the relevant details concerning the equipment are entered or can be edited, including details of accessories and manuals available with the equipment.

(ii) **Maintenance and Repair "Cards"**: As a pull-down menu these two "cards" show all the details concerning the equipment's M&R record, the technician who performed it and on another menu card the relevant costs in materials and labour.

(iii) **Stock Control**: A stock control entry and reporting system is available to enable the generation of stock inventory and reports. A warning facility exists for reordering of spares. This allows for a smooth availability of spares at optimum cost.

(iv) **Fault Records**: It is important to know for reasons of planning, equipment appraisal, stock procurement, the type of faults that occur in equipment.

(v) **Manufacturer Information**: Gives details of manufacturers for easy ordering and correspondence.

(vi) **Importing or Updating Data**: Data from other data bases may be imported, such as equipment codes, equipment standards, manufacturer details etc which are useful for more effective management of equipment.

(vii) **System Entry**: The present security system allows entry via a **User ID** and **Password**.

**Future Developments**:
System access via different levels of authority will need to be implemented, giving different degrees of access to personnel according to their security status.
Multi-user capability can easily be provided, allowing access for vewing and/or change of records according to user security level.

# REFERENCES

(1) "Interregional Meeting on the Maintenance and Repair of Health Care Equipment", Nicosia, 1986, WHO/SHS/NHP/87.5
(2) "Global Action Plan on Management, Maintenance and Repair of Health Care Equipment", 1987,WHO/SHS/NHP/87.8
(3) "Universal Medical Device Nomenclature System", by ECRI/WHO Collaborating Center, Pensylvania, USA, 1991.

```
                  ┌──────────────────────┐
                  │  M A I N   M E N U   │         April 11, 1994
                  └──────────────────────┘

                    MEDICAL EQUIPMENT INVENTORY PROGRAM


                      1. Equipment Records
                      2. Search Records
                      3. U.M.D.N.S.
                      4. Stock Control
                      5. Standard Faults Inventory
                      6. System Maintenance
                      7. System Reports

                      0. System Close Down
```
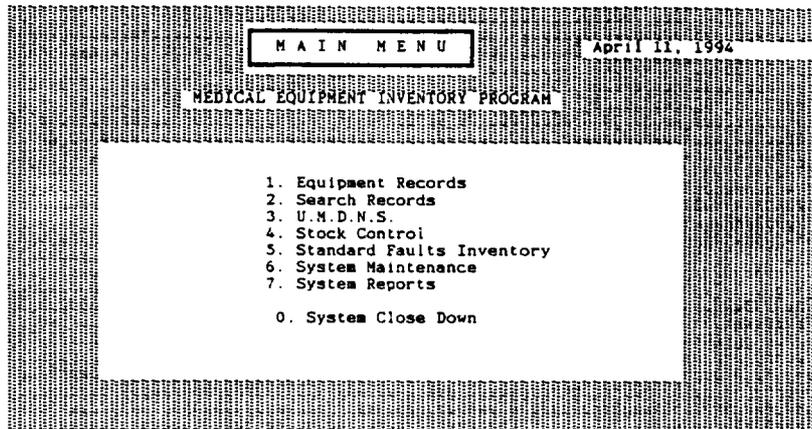
Figure 1 : Main Database Modules



```
        M E D I C A L   E Q U I P M E N T   R E G I S T R Y   1994/04/11

                                               Record Number :          1
Movable Item (Y/N)  : Y                        Inventory No. :        555
Department          : Nicosia General          Under Repair (Y/N) :   N
Location            : Casualty          -
Classification      : Electro Medical
Description of Item : Portable ECG

      Manufacturer :                         Equipment Specifications

   SIEMENS                               Mfr. Serial No. : 222345
   Model/Type : EC/II                        Voltage   :   240 V
     P.O. No. : 2156                       Frequency :   50 Hz
     P.O Date : 1989/03/21                     Phase : 1 -ph 1 2 3
   Comm. Date : 1989/04/15        Additional Spec. : none
                                  Accessories / Remarks: ECG leads/PS
   Available Manuals (X) :        Planned Maintenance Every   20   Weeks

   Operating (X)  Service (X)  Installation (X)  Recommended Spare Parts (X)

 Action : F   Edit  Find  eXit        Pg-Up - Previous   Press F1 for more
              Ins - Insert new Record Pg-Dn - Next       Function Keys
```

Figure 2 : Equipment "Card"



```
                                                              94/04/11
                        Record Maintenance Data                    1
 Movabl                 ---,---,-----------                       555
 Depart                                                         : N
 Locati      Down Date  : 1990/10/10
             Up Date    : 1990/10/21           Cost
                                      Labour   :       30
             Man Hours Spend :   1    Material:       55
                                      Total    :      105
  Plann      Company / Department                               27
             Performing Maintenace : NGH Workshop
  Date
             Description of Work Done :
 1990/0      Recharging Battery Unit
 1990/1
             Signed By :  A Antony
                                Technician
```

Figure 3: Equipment Maintenance Record