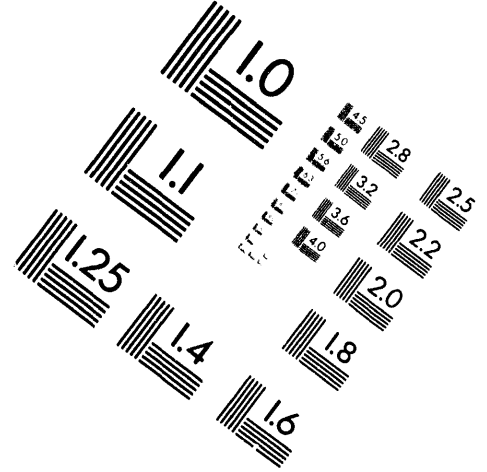
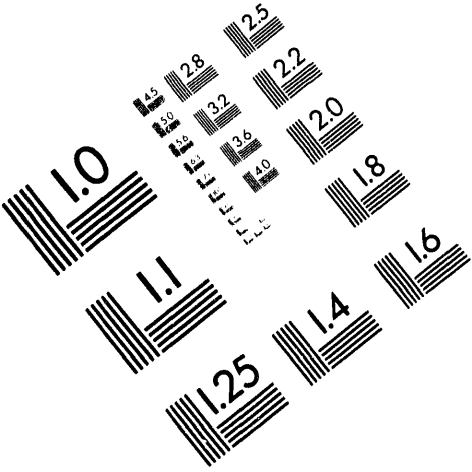




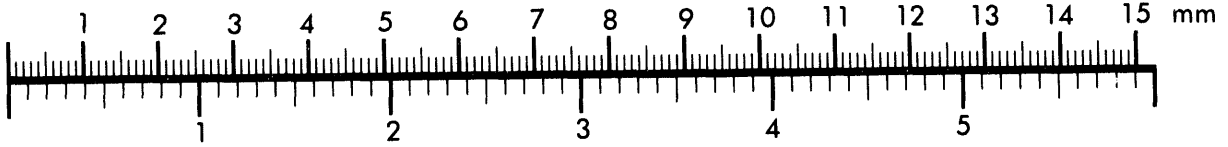
**AIM**

**Association for Information and Image Management**

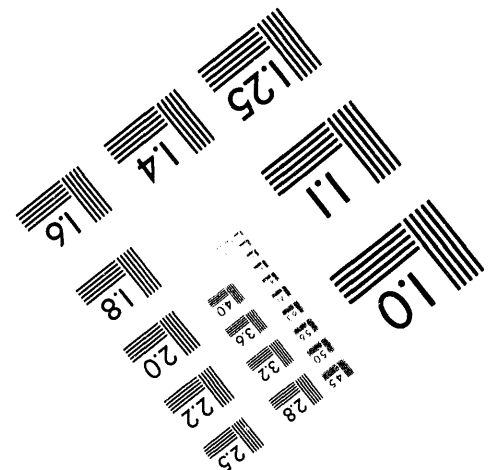
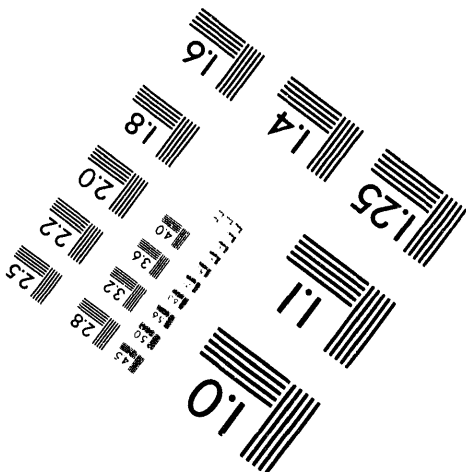
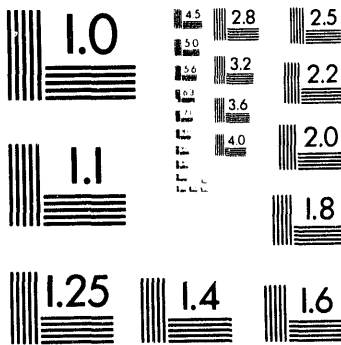
1100 Wayne Avenue, Suite 1100  
Silver Spring, Maryland 20910  
301/587-8202



Centimeter



Inches



MANUFACTURED TO AIM STANDARDS  
BY APPLIED IMAGE, INC.

---

**1 of 1**

---

SAND-94-1354C  
Conf-941219--1

## Fuzzy Clustering of Facial Form for Prototyping Environmental Protection Equipment

RECEIVED

JUL 06 1994

OSTI

David G. Robinson, Ph.D.  
GRAM, Inc.  
Albuquerque, NM

### Abstract

Emphasis on the human-to-aircraft interface has magnified in importance as the performance envelope of today's aircraft has continued to expand. A major problem is that there has been a corresponding increase in the need for better fitting protection equipment and unfortunately it has become increasingly difficult for aircrew members to find equipment that will provide this level of fit. While protection equipment has historically had poor fit characteristics, the issue has grown tremendously with the recent increase in the numbers of minorities and women. Fundamental to this problem are the archaic methods for sizing individual equipment and the methods for establishing a sizing system. This paper documents recent investigations by the author into developing new methods to overcome these problems. Research centered on the development of a new statistically based method for describing form and the application of fuzzy clustering using the new shape descriptors. A sizing system was developed from the application of the research, prototype masks were constructed and the hardware tested under flight conditions.

### Background

The US Army and Air Force have conducted numerous studies to collect anthropometric information to assist in determining standard sizes of clothing and protection equipment. In each study, attempts at developing a sizing system using classical statistical methods have consistently failed. Analysts and engineers typically ignore the statistical results and choose arbitrary linear dimensions for the design of protection equipment such as the oral-nasal oxygen mask. In this case, face length and width were chosen as primary design variables and five sizes successively redefined through extensive trial and error efforts.

As yet, no organized methodology exists to develop a sizing system for protection equipment, in particular oxygen masks and helmets. The remainder of this paper documents a recent attempt to develop and test such a methodology. The paradigm is based upon a new method to characterize facial form and the use of fuzzy clustering as an alternative to classical statistical methods in developing a sizing system.

### Problem

The form of an object will refer to both size and shape and is invariant under translation, rotation and reflection. Shape remains invariant under translation, reflection and scaling of an object [Lele]. Of interest in this study is the characterization of cranio-facial form to assist in protection equipment to be worn on the head and face.

The first aspect of the problem is how to characterize these forms in a logical and consistent manner. A number of popular methods were reviewed including procrustes [Lele], thin plate splines (TPS) [Bookstein], Euclidean distance matrix analysis (EDMA)

This work was supported by the United  
States Department of Energy under  
Contract DE-AC04-94AL85000.

MASTER 1

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED *875*

[Lele and Richtsmeier] and finite element scaling analysis [Richtsmeier, et.al.]. All methods currently in use depend on the existence of homologous anatomical landmarks on both the subject and a reference form. In addition, this comparison is based on measuring form difference as a function of deformation of the subject landmarks to the landmarks on some reference form. Several difficulties exist with the application of these methods to characterize facial form. First, normative models of the human face do not exist. Second, since it was required to characterize the form of the seal area for the protection equipment, and a unique set of anatomical landmarks could not be constructed to describe the seal line, it was necessary to develop an alternative method for describing facial shape and form.

The second aspect of the problem involves utilizing these form descriptors to develop a sizing system. Conflicting objectives exist: better fitting equipment with a minimum number of sizes. Landmark based characterization coupled with traditional statistical techniques have clearly fallen short on both accounts, forcing designers to bypass statistically based anthropometric methods and base design almost completely on the artistic abilities of the designer.

### **New Shape Descriptor**

The proposed form descriptor can philosophically be viewed as an extension of the methods proposed by Bookstein and those proposed by Lele and Richtsmeier. However, mathematically the new method is quite different. The basis for the method is the development of a functional descriptor for the surface being considered.

Let the surface of interest be described by the set of points  $(x_i, y_i, z_i), i = 1..n$ . These points are distinguished from landmarks in that there is no required association with a particular anatomical feature. Assume that the function  $z(x_i, y_i) = z(\mathbf{x})$  represents a regionalized variable [Cressie]. In particular, the function will be assumed to exhibit at least intrinsic stationarity; i.e. the first two moments are only required to be constant over a small region and this region may be infinitely small in size and the second moment is not necessarily bounded.

Let  $\mathbf{h}_{ij}$  define the vector between two points  $x_i, x_j$ . Note that  $\mathbf{h}_{ij}$  is a function not only of the distance between the two points, but also a function of the orientation between those points. Under intrinsic stationarity, the first two moments of the relationship between any point and another is given by:

$$E[z(\mathbf{x} + \mathbf{h}) - z(\mathbf{x})] = 0$$

$$V[z(\mathbf{x} + \mathbf{h}) - z(\mathbf{x})] = 2\gamma(\mathbf{h})$$

where  $\gamma(\mathbf{h})$  is referred to as the semi-variogram function and is a measure of the spatial independence of the two points. If we restrict the first two moments to be constant across the entire regionalized variable (thus only considering the function to exhibit weak stationarity), then there is a direct relationship between the variogram and the more common autocorrelation function. A number of functional forms for  $\gamma(\mathbf{h})$  have been verified. One of the most common is the spherical model:

$$\gamma(h) = \begin{cases} 0 & \text{when } h = 0 \\ C_0 + C \left( \frac{3h}{2a} - \frac{h^3}{2a^3} \right) & \text{when } 0 < h < a \\ C_0 + C & \text{when } h \geq a \end{cases}$$

Define the matrix:

$$\Gamma = \begin{bmatrix} 0 & \gamma(h_{12}) & \cdots & \gamma(h_{1n}) & 1 & x_1 & y_1 \\ \gamma(h_{21}) & 0 & \cdots & \gamma(h_{2n}) & 1 & x_2 & y_2 \\ \vdots & \vdots & \ddots & \vdots & 1 & \vdots & \vdots \\ \gamma(h_{n1}) & \gamma(h_{n2}) & \cdots & 0 & 1 & x_n & y_n \\ 1 & 1 & \cdots & 1 & 0 & 0 & 0 \\ x_1 & x_2 & \cdots & x_n & 0 & 0 & 0 \\ y_1 & y_2 & \cdots & y_n & 0 & 0 & 0 \end{bmatrix}$$

The best linear unbiased predictor of any point  $\mathbf{x}_o$  on the surface is then given by:  $\hat{z}(\mathbf{x}_o) = \gamma_o^T \Gamma^{-1} \mathbf{z}$ , where  $h_{io}$  is the variogram function between point  $i$  and the unknown point  $\mathbf{x}_o = (x_o, y_o)$  and:

$$\gamma_o = [\gamma(h_{1o}) \quad \gamma(h_{2o}) \quad \cdots \quad \gamma(h_{no}) \quad 1 \quad x_o \quad y_o]^T$$

$$\mathbf{z} = [z(\mathbf{x}_1) \quad \cdots \quad z(\mathbf{x}_n) \quad 0 \quad 0 \quad 0]^T$$

Let  $K = \Gamma_n^{-1}$ , where  $\Gamma_n^{-1}$  is the upper  $n \times n$  submatrix of  $\Gamma^{-1}$ . The matrix  $K$  is referred to as the energy matrix and uniquely characterizes the set of points used to describe the subject form. The matrix is invariant to translation, rotation, reflection but remains sensitive to scaling; required characteristics in the description of form. While unique, the matrix typically contains a large number of features (e.g. 57x57 for the current problem). A number of methods were evaluated to extract important information from the matrix. The most promising method investigated was based upon spectral decomposition of the matrix similar to the principle warps suggested by Bookstein. However, it was expected that future application would require an automated process and therefore a much simpler procedure was settled on.

If the matrix is treated as a two dimensional image of the surface, then classical image processing algorithms can be applied. In this study, an energy matrix was generated for each of the 99 subjects. A Hotelling transformation (a.k.a. Karhounen-Loeve transform (KLT) or principle components analysis ) is applied to the entire data set to generate an orthogonal, uncorrelated sample space. The result of the transformation is a set of ordered features which can be used for clustering.

### Fuzzy Clustering

As mentioned, the two major objectives underlying the development of a sizing system for protection equipment are to have the minimum number of sizes possible while at the same time maximizing the fit of the individual. Past difficulties with classical statistical clustering using landmark data highlight the need to consider alternative

methods to group similar facial forms. The natural use of linguistic descriptors when discussing the similarity between faces begs the application of fuzzy logic.

Since the sample size for this effort was limited to ninety-nine subjects, the curse of dimensionality was a significant consideration in both the selection of the number of clusters and the dimension of the feature space. A number of relationships have been suggested between sample size (N), number of clusters (c), and number of features (f). A typical requirement [Kanel, et.al., Raudys, et.al.] is that:  $cf\beta \leq N$ , where  $\beta$  is some safety factor.

A number of clustering algorithms were investigated ranging from the recent developments in adaptive resonance theory [Simpson] to the popular c-means [Bezdek]. The preliminary results indicated some minor benefits to using c-means, and therefore fuzzy clustering was performed to minimize the performance function:

$$J_m = \sum_{i=1}^c \sum_{j=1}^n (\mu_{ij})^m (\mathbf{x}_j - \mathbf{v}_i)^T \mathbf{A} (\mathbf{x}_j - \mathbf{v}_i)$$

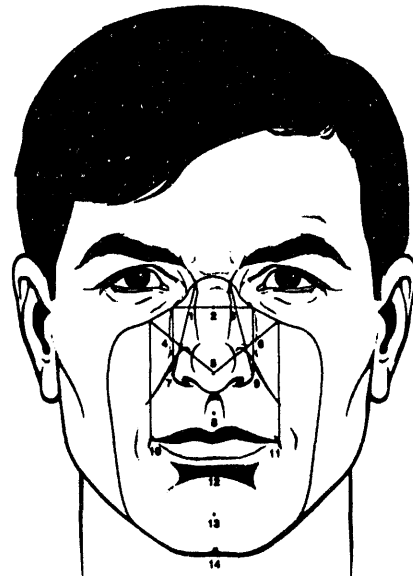
using the classical c-means algorithm with a subset of the ordered feature space generated by the KLT transform as input. It should be noted that the KLT/c-means algorithm was also successfully applied to both the EDMA and TPS shape descriptors.

## Application

### *Data Collection and Analysis*

At the core of the effort was the desire to incorporate new data collection methods such as highly accurate digital and scanning laser based systems. Figure 1 depicts the seal line and the additional points that were collected from each subject. Approximately thirty-nine points were collected from the seal line along with fourteen additional points. Data was collected from a total of ninety-nine subjects.

Energy matrices were calculated and a KLT analysis performed to identify a feature set. Fuzzy clustering was then performed using the c-means algorithm. Note that since the subjects were available as plaster casts, it was possible to investigate a number of clustering parameters and feature sets. Clusters were formed and the plaster molds were laid out on the floor in the resulting groups. A number of potential form descriptors were investigated and this immediate feedback was critical during the development. Using a sample size of ninety-nine, a maximum of six clusters from five features and a c-means coefficient of  $m=2$  was found to give consistent results. Figure 2 depicts front and side views of faces from three of the six clusters that resulted from the final analysis.



### ***Prototype Construction and Testing***

Those subjects whose faces had the highest degree of membership in each of the clusters were chosen as prototypes for each of the sizes. The data points from each of these faces was input to a computer-aided design program. The computer models were subsequently transferred to a rapid prototyping machine and molds for custom masks output. These molds were then used by the mask shop at Wright-Patterson AFB to construct a set of custom masks. To evaluate the fit, subjects from four different clusters were randomly chosen. However, as an extreme test of the methodology, subjects were chosen to be at the fringe of the clusters. The table below presents the membership of the prototype subjects and the subjects used for testing.

Those clusters which contained very few subjects were not evaluated in the fit test since it was obvious from the data that a good fit was going to be achieved for all members in those clusters.

Cluster	Center Subject	Membership	Test Subject	Membership
1	1986	0.92	0025	0.84
2	0027	0.96	-	-
3	0011	0.94	2024	0.89
4	0035	0.91	-	-
5	1980	0.93	1958	0.88
6	1918	0.99	1981	0.93

Table 1. Centroids and Survey Subjects

### ***Results***

The prototype masks were distributed to survey subjects and tested in either actual or simulated positive pressure flight conditions. The results of the survey from the fit testing were mixed but not unexpected. One survey subject (0025) refused to return her mask because the fit was so much better than her previous custom mask. Alternatively, one survey subject mask was unusable due to excessive leakage. Other survey results fell about the average. In reviewing the data associated with the original clusters it was concluded that due to the disparity in the number of subjects in each cluster, the degree of membership was not in itself an accurate measure of fit; again the curse of dimensionality becomes an issue. This problem is expected to resolve itself as more data becomes available.

### **Acknowledgments**

A portion of this work was accomplished while the author was an Associate Professor with the Air Force Institute of Technology. Funding for continued research in this area is being provided by Sandia National Laboratories and is gratefully acknowledged.

### **DISCLAIMER**

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.



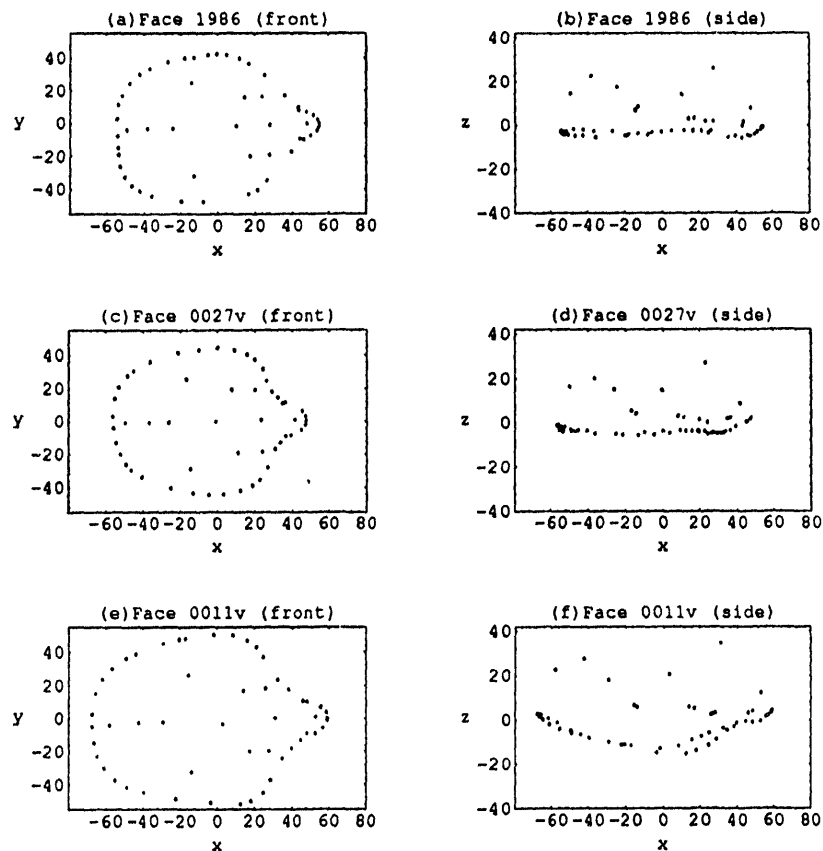


Figure 2. Three Typical Centroids (Front and Side Views)

## References

- Bezdek, J. *Pattern Recognition with Fuzzy Objective Function Algorithms*. New York: Plenum Press, 1981
- Bookstein, F. *Morphometric Tools for Landmark Data*. New York: Cambridge University Press, 1991.
- Cressie, N. *Statistics for Spatial Data*. New York: Wiley Interscience, 1991.
- Kanel, L., B. Chandrasekaran *On Dimensionality and Sample Size in Statistical Pattern Recognition*, *Pattern Recognition*, 3 (1971).
- Lele, S. *Some Comments on Coordinate-Free and Scale-Invariant Methods in Morphometrics*, *American Journal of Physical Anthropology*, 85:407-417, 1991.
- Lele, S. J. Richtsmeier. *Euclidean Distance Matrix Analysis: A Coordinate-Free Approach for Comparing Biological Shapes Using Landmark Data*, *American Journal of Physical Anthropology*, 87:49-65, 1992
- Raudys, S., A. K. Jain *Small Sample Size Effects in Statistical Pattern Recognition: Recommendations for Practitioners*, *IEEE Transactions on Pattern Analysis and Machine Intelligence* (March 1991).
- Richtsmeier, J. et. al. *Advances in Anthropological Morphometrics*, *Annual Review Anthropol.*, 21: 283-305 (1992).
- Simpson, P. *Fuzzy Min-Max Neural Networks -Part 2: Clustering*, *IEEE Trans Fuzzy Systems*, vol 1 no 1 (February 1993).

---

**DATE**

**FILMED**

**11 / 2 / 94**

**END**

---

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_