CONF-900290--2

HYBRID INTELLIGENT PERCEPTION SYSTEM: INTELLIGENT PERCEPTION THROUGH COMBINING ARTIFICIAL NEURAL NETWORKS AND AN EXPERT SYSTEM*

CONF-900290--2

DE90 007483

C. W. Glover
P. F. Spelt

Center for Engineering Systems Advanced Research
Oak Ridge National Laboratory
Post Office Box 2008
Oak Ridge, Tennessee 37831-6364

Paper To Be Presented At:

First Workshop on Neural Networks:
Academic/Industrial/NASA/Defense

Auburn University
Auburn, Alabama

February 5-6, 1990

*Research sponsored by Oak Ridge National Laboratory, operated by Martin Marietta Energy Systems, Inc. under Contract No. DE-AC05-84OR21400 with the U.S. Department of Energy.

Hybrid Intelligent Perception System: Intelligent Perception through Combining Artificial Neural Networks and an Expert System

C. W. Glover and Philip F.Spelt
Center for Engineering Systems Advanced Research
Building 6025, Mail Stop 6364
Oak Ridge National Laboratory
Post Office Box 2008
Oak Ridge, Tennessee 37831-6364

This paper presents a report of work-in-progress on a project to combine Artificial Neural Networks (ANNs) and Expert Systems (ESs) into a hybrid, self-improving pattern recognition system. The purpose of this project is to explore methods of combining multiple classifiers into a Hybrid Intelligent Perception (HIP) System. The central research issue to be addressed for a multiclassifier hybrid system is whether such a system can perform better than the two classifiers taken by themselves. ANNs and ESs have different strengths and weaknesses, which are being exploited in this project in such a way that they are complementary to each other: Strengths in one system make up for weaknesses in the other, and vice versa. There is presently considerable interest in the AI community in ways to exploit the strengths of these methodologies to produce an intelligent system which is more robust and flexible than one using either technology alone. Perception, which involves both data-driven (bottom-up) and concept-driven (top-down) processing, is a process which seems especially well-suited to displaying the capabilities of such a hybrid system. This work has been funded for the past six months by an Oak Ridge National Laboratory seed grant, and most of the system components are operating in both the PC and the hypercube computer environments. Here we report on the efforts to develop the low-level ANNs and a graphic representation of their knowledge, and discuss ways of using an ES to integrate and supervise the entire system.

INTRODUCTION

This paper describes research on a hybrid neural network and rule-based pattern recognition system which is capable of self-modification or learning. Our hybrid system exploits the complementary strengths and weaknesses of artificial neural networks (ANNs) and expert systems (ESs), or rule-based (RB) systems. This combination is expected to produce a pattern classification system which performs better than either classifier by itself. In addition, the hybrid system is capable of self-modification through a feedback loop between the ANNs and the ES. This feedback loop allows an ANN to be automatically trained by the ES, while the ES can also modify the models in its knowledge base from information supplied by the ANNs.

One strength of an ES is that general problem solving strategies can be built into the system (Michalski, Carbonell and Mitchell, 1983). This allows for complex control structures, which are often based on heuristics, to be easily incorporated. In contrast, it is very difficult to build general problem solving strategies based on complex control structures, especially those utilizing heuristics, from ANNs. A second strength of an ES is its ability to provide an explanation (a decision trace) of the steps leading to its conclusions, a capacity clearly not possible with ANNs. On the other hand, an important weakness of an ES is its inability to convert sensor data into symbolic information -- the transformation between the two representations must be known _a priori_ and be explicitly incorporated into the system. An ANN learns the transformation through the training process.

Perhaps the most important complementarity between ANN and RB systems is constraint satisfaction. RB systems employ what Smolensky called "hard constraints", whereas ANNs use "soft constraints" (Smolensky, 1988). The ES's hard constraints are provided by a set of rules which describe an aspect of the environment, the conditions of which must be met exactly in order for the rule base to recognize a match between the environmental events and its _a priori_ model(s). Thus, in a traditional ES, a control panel as shown in the middle of Figure 1a might be described as a "large rectangular black blob", with a set of white blobs contained within and arranged according to a certain specified configuration. Failure to meet these specifications could occur for two reasons: the vision system fails to detect all the small white blobs, thus one hard constraint is not met and the MATCH fails; on the other hand, noise in the vision system could create extra white blobs within the black one, again causing the MATCH to fail. Specifically, an ES which recognizes the central panel in Figure 1a would fail to recognize the same panel in Figures 1b and 1c because the lower white circles are missing from the description. This is a strength if the two examples really belong to different categories, but a weakness should the extra white blobs be due to vision system noise. Thus, an ES's hard constraint satisfaction is not very fault tolerant, which is both a strength and a weakness.

On the other hand, an ANN's fault tolerant, soft constraint satisfaction produces a different outcome to the same panel identification problem. In an ANN trained to recognize the control panel on the left in Figure 1a, the representation of the control panel would consist of a fixed length feature vector, with each vector component representing a different geometrical moment distribution of black and white pixels around a suitably chosen set of axes. The training set then is a set of feature vectors representing both positive and negative panel examples. After suitable training, a feature vector for another panel presentation to the vision system is sent to the ANN. The output from the ANN is a number representing a similarity measure between the extracted feature vector and the feature vector of the panel the ANN was trained to identify. If the ANN's output value is above some threshold value, then a MATCH condition is returned. Generally, a threshold value is chosen such that a MATCH condition is returned in the presence of noisy input vectors. With a suitably chosen threshold value, the ANN MATCHes the control panel on the left in Figure 1b and 1c with the control panel it was trained to identify from Figure 1a. This soft constraint satisfaction makes the ANN fault tolerant. Just as with the ES's hard constraint satisfaction, the ANN soft constraint satisfaction can be perceived as both a strength and weakness: the ANN may classify these two panel examples as the same, when actually the additional white blobs (buttons) at the bottom require it to be placed in a different category.

It is clear that the strengths and weaknesses of the ES's hard constraint satisfaction are complementary to those of the soft constraint satisfaction of an ANN. By exploiting the complementary strengths and weaknesses of ESs and ANNs, the hybrid system is an attempt to build a more robust pattern classifier than either an ES or ANNs acting alone. The next section is an overview of the salient features of the hybrid system. There follows a detailed discussion of the current implementation of the hybrid system as it applies to a machine vision problem. Finally, a report on the current status of the hybrid system is presented, along with an indication of future work.

HYBRID SYSTEM OVERVIEW

In general, the hybrid perception system architecture is a hierarchical layering of pattern classifiers. Figure 2 presents a generic hybrid system architecture. Information from a sensor or some other system enters from the left, where it is processed by a hierarchical layering of low-level pattern classifiers. These classifiers extract features or attributes from the data for use in the higher level classifiers. Any suitable pattern classification algorithm can be used for the low-level pattern classifiers. Our hybrid system uses ANNs for these classifiers.

For the system to function in its hybrid mode, the ES is given an a priori set of models of certain expected environmental states, conditions, or objects, together with any heuristics needed to cope with the task. These models can be stored in the knowledge base in a variety of representations. One of the goals of the ES is to determine whether certain expected environmental events are present in the current data stream. The first job of the RB classifier system is to assimilate the feature set from the low-level classifiers into the same representation that is used for the models. The ES then uses a suitably chosen metric of the similarity between assimilated and model representations to decide whether a match exists. If a match is found, then the ES invokes a routine to automatically create and train a new high-level ANN pattern classifier from the appropriate low-level feature data. The purpose of the new high-level pattern classifier is to recognize this model state when it is present in future data streams. A different high-level classifier is created for each model state found by the RB system.

Both the high-level classifiers and the RB system receive all subsequent low-level feature vector data streams, and each provides an estimate of whether a certain model state exists in the data stream. If the two classifiers (the high-level ANN and the ES) agree -- the event either is or is not there -- the hybrid system needs only to report the outcome. It is when the high-level ANN and the ES disagree that the strengths of the hybrid system are exploited. Because the RB system must compile fragments of information to determine whether a specific environmental state exists, one can be reasonably sure that one of the model states is present in the data when an ES yields a MATCH condition. On the other hand, the high-level classifier (ANN) will yield a MATCH when only some fraction of the input data overlaps with training data. Thus, the decision outcome of the two high-level classifiers may be different. The tradeoff between the hard constraint satisfaction of the ES and soft constraint satisfaction of the high-level ANN classifier provides the basis for exploiting the strengths of both types of classifiers and minimizing the weaknesses. The resolution of the disagreement is the unique contribution of the hybrid system, as will be described in the section on our implementation. Clearly, the disagreement must be resolved, and the methodology for doing so is both the primary contribution of a hybrid system and the focus of our research on this system.

HYBRID SYSTEM IMPLEMENTATION

The first, proof-of-principle, version of the hybrid system is being built to process vision data from a mobile robot. This section illustrates the current hybrid system's information flow and decisions process using a specific example. There are many other problems, in other research domains, where the current hybrid system also can be used. The problem addressed by this particular implementation of the hybrid system can be explained from Figure 1 as follows: A mobile robot is to find and identify specific types of control panels located anywhere in a building (see Figure 1a). The hybrid system's RB contains _a priori_ knowledge of each control panel type, in the form of a qualitative symbolic model. Pre-processed visual data from the robot's CCD camera are presented to both the low-level ANN classifiers and the RB system. Once a panel is identified, the RB system automatically creates and trains a high-level ANN to recognize the control panel from the current data stream. As the robot moves closer to the control panels (see Figures 1b and 1c) new features may emerge which were not known _a priori_. In this event, the hybrid system must learn the new features and update its models and categorization scheme.

Our implementation of a hybrid system to solve the problem just described is shown in Figure 3. For the first image processing, the robot camera's gray-level image is passed to a set of preprocessing routines for feature extraction: the gray-level image is converted to a binary image, and a component labelling operation is performed in which all connected black or white regions ("blobs") are detected and assigned unique integer values as labels.* Then certain blob features are extracted: the centroid and area, the size and coordinates of a bounding box, and the Zernike moments. Zernike moments are used because they are scale, translation, and rotation invariant representations of the standard moments (Teague, 1987). The routines that perform these feature extraction operations are completely parallel algorithms running on a hypercube computer (Jones, Mann and Simpson, 1988).

Ruck (1987) has used Zernike moments extracted from Doppler-Shift RADAR data as input to previously trained nearest-neighbor (Duda and Hart, 1973) and backpropagation ANNs (Rumelhart and McClelland, 1986). Using these Zernike moments, Ruck showed that the backpropagation algorithm provides better discrimination among the examples to classify tanks, jeeps, and trucks. Based on this result, a backpropagation ANN algorithm is used as a low-level classifier to classify the shape of each blob in an image as a circle, ellipse, square, rectangle, triangle, or unknown. Each blob's Zernike moments, as

---

* Simply put, connected regions are blobs in an image, and each blob in the image is counted. A blob's label is its number in the count.

components in an ANN input vector, are passed to low-level ANNs for shape classification. The resulting blob shape classification is concatenated with the blob's other features (color, centroid, area, bounding box size) into a blob attribute vector. (Details of this procedure can be found in Glover, Silliman, and Walker, in preparation).

The attribute vector of every blob in the initial scene image is passed to the ES, where the vectors are integrated into the same type of representation as the a priori models of the control panels. In this implementation, directed graphs represent both the control panel models and the assimilated blob information from the image. A directed graph is constructed for each blob whose bounding box encloses other blobs. The directed graph structure was chosen because it is invariant under scale changes and translation. Construction of a directed graph is a multistep process: For each blob located in the interior of a bounding blob, a graph node is located at the blob's centroid coordinates. A line is then drawn through all nodes that are collinear to within some tolerance, which defines, e.g., the jth level in the graph structure with d_j nodes on that level. Once all the levels of a graph have been established, the connections between nodes on the jth level and those on the j - 1 level are obtained by dropping a perpendicular ray from each node on the jth level and rotating the ray counterclockwise until it intersects a node on the j - 1 level. Connections between nodes on the jth and the j + 1 levels are obtained by erecting a perpendicular ray from the node on the jth level and rotating the ray clockwise until it intersects a node on the j + 1 level. The model graphs have been constructed in a similar manner and stored in the ES's knowledge base.

An entire image may generate a set of directed graphs, $G = \{G_i \mid i \in N\}$, where N is the number of graphs extracted from an image. Each image graph G_i must then be compared with each model graph G^*_k in the set of M model graphs, $G^* = \{G^*_k \mid k \in M\}$, in order to determine if a MATCH condition exists. The ES must extract a similarity measure between each G_i and G^*_k, for all i, k pairs. The hybrid system's RB uses a similarity metric that is equal to the number of transformations needed to transform Gi into G^*_k. This number is found by counting the number of levels and the number of nodes per level that must be added to G_i in order to transform it into G^*_k. Rao and Glover (1989) have developed a nontrivial extension of Hirschberg's string matching algorithm (Hirschberg, 1975) for 2-dimensional planar graphs. The Rao and Glover algorithm finds the number of levels l and the maximum number of nodes per level d by which G_i and G^*_k differ; i.e. $G_i - G^*_k = (l,d)$.

A MATCH condition is returned by the ES if $G_i - G^*_k = (l,d) \leq (l',d')$ and the similarity metric $\leq l' + d'$, where (l',d') are thresholds initially determined by the user. A MISMATCH condition will be returned if these criteria are not satisfied, and the RB system will consider the next pair of graphs from sets G and G^*. An exact match between G_i and G^*_k would yield $(l,d) = (0,0)$. An exact match between G_i and G^*_k also requires that each G_i node's attribute vector match the corresponding G^*_k node's attribute vector.

For each MATCH condition returned from the initial image, the ES automatically creates and trains a high-level ANN to recognize the matched control panel using the Zernike moments extracted from the image. Parametric studies (Glover and Walker, in preparation) have shown that the following set of heuristics facilitates the automatic training of an ANN backpropagation algorithm: First, the ES creates a backpropagation ANN of a predetermined size. Then the matched blob's Zernike moment vector is added, as a positive example, to a predetermined list of Zernike moment vectors which serve as negative examples. The ES monitors the convergence rate during ANN training. If the convergence rate drops below some threshold value while the ANN's global error is still above an acceptable level, then the ES adjusts the step size and/or the momentum parameters in the backpropagation learning rule. If this still does not lead to convergence, then it is assumed that the ANN is stuck in a local minimum. In this case, the ES resets the learning parameters, changes the ANN initial weight values, and restarts the ANN training procedure just described. If this fails to yield an acceptable convergence rate and global error value, the ES then adds more hidden layer nodes to reduce the number of local minima and restarts the entire ANN training procedure. Thus far, these heuristic training procedures have always produced a convergent ANN.

Once a high-level ANN has been trained for a recognized environmental event, the hybrid system is ready to accept the next image, and the image processing proceeds as before. This processing of additional images is now synchronized with the high-level ANN processing. Simultaneously with the low-level ANN processing just described, the high-level ANN receives as input each image blob's Zernike moment vector and produces a similarity measure as its output. This value is sent to the decision module in the ES along with the output information from the low-level ANN processing.

Three decision outcomes are possible once any trained high-level ANNs exist: MATCH, RB-MATCH, and ANN-MATCH. The decision module in the RB returns a MATCH condition if the ES similarity measure satisfies its matching criteria, and the high-level ANN's output is above some predetermined threshold. If this is the case, the ES adds the Zernike moment vector extracted from the current image to the training set of the high-level ANN, as a positive example; but no high-level ANN training occurs. The decision module in the ES returns a RB-MATCH condition if the ES similarity measure satisfies its matching criteria, while the high-level ANN's output is below some predetermined threshold. This situation produces the first example of self-modification in the hybrid system: In this situation, the ES adds, as a positive example, the Zernike moment vector extracted from the current image to the training set of the high-level ANN, and trains the high-level ANN with the modified training set. The ES does not initialize the ANN weights, but uses the old weights as starting values. It has been found that convergence is always obtained with only a few iterations through the training set (Glover and Walker, in preparation).

The ES's decision algorithm returns a ANN-MATCH condition, if the ES similarity measure does not satisfy its matching criteria, but the high-level ANN's output is above some predetermined threshold. In this case, the ES adds the Zernike moment vector extracted from the current image to the training set of the high-level ANN as a positive example. However, no ANN training occurs. Since the RB system did not produce a match and the high-level ANN did, indications are that new features have emerged which are not present in the model. In this case, the ES creates a temporary model graph from the graph extracted from the image and adds it to the ES's knowledge base. If the temporary model graph is verified in subsequent images it will replace the old model graph; otherwise it will be deleted.

Again, one of the main research issues with the hybrid system is whether it is capable of unsupervised learning without destroying its knowledge base. Initially, this research will focus on the choices of the threshold values used for the matching conditions, and the strategy used to update the knowledge base. Investigation of this research issue is in progress.

SUMMARY AND FUTURE DIRECTIONS

This paper describes a hybrid neural network and rule-based pattern recognition system capable of self-modification or learning. The central research issue to be addressed for a self-modifying multiclassifier hybrid system is whether such a system can perform better than either of the two classifiers taken by themselves. The complementary strengths and weaknesses of ES and ANN classifiers served as the motivation for building the hybrid system. Self-modification of the hybrid system is achieved through a novel feedback loop between the ES and the high-level ANN. This feedback loop allows the hybrid system to exploit the ES's hard constraint and the ANN's soft constraint matching capabilities. New information about objects in a scene can be gleaned from the various combinations of RB and ANN matching outcomes. Both the ES and ANN can learn form this new information, which could not be inferred by a system employing only one type of classifier.

The major thrust of research is to measure the hybrid system's performance on a variety of sensor driven problems, where each problem is designed to systematically explore the hybrid system's operating limits. A secondary research issue focuses on the implementation of the hybrid system in parallel form on a hypercube computer architecture. The hypercube architecture will be decomposed into many smaller sub-hypercubes where each component of the hybrid system will be executed on the sub-hypercubes. Synchronized communications between the sub-hypercubes will insure that the information needed by each hybrid system component will arrive at the appropriate time. Each hybrid system component currently runs in modular form on the nodes of a hypercube computer. The implementation of a general control structure to insure synchronized communications is being developed.

Future hybrid system research will proceed in two areas. The first initiative is directed toward multiple sensor integration. The hybrid system will be interfaced with multiple sensors with the data from each sensor being processed by a suitably chosen hierarchy of low-level classifiers. The integration of the sensor information then will be done in symbolic form by the ES. Subsequently, high-level classifiers will be created to process multiple sensor data. The second initiative is directed toward incorporating more disparate types of high-level classifiers into the hybrid system architecture. All types of classifier have their own sets of strengths and weaknesses. By having several different types of classifiers trained to recognize the same pattern sets, the ES decision module will be provided with several different similarity measures. The ES can exploit the different classifiers' strengths and weaknesses to accomplish more robust and precise classification.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the outstanding work of M. Silliman and M. Walker in developing the ANN components of the system described here. Dr. N. S. V. Rao, of Old Dominion University, is responsible for much of the work on the directed graph portions of the hybrid system. We also thank Dr. Reinhold Mann, Dr. Ed Oblow, and Dr. Judd Jones, all of Oak Ridge National Laboratory, for enlightening discussions, support, and help. This research was funded internally by an Oak Ridge National Laboratory exploratory research grant.

BIOGRAPHY

Philip F. Spelt, Ph. D. - a cognitive scientist in Oak Ridge National Laboratory's Engineering Physics and Mathematics Division - developed the Learning and Inferencing system components for an autonomous mobile robot at the Center for Engineering Systems Advanced Research (CESAR) laboratory. He was a co-recipient of a Laboratory R & D grant to develop the hybrid perception system described in this paper. He is currently developing an adaptive planning expert system for two cooperating autonomous robots, as well as doing cognitive engineering work for remotely controlled robotic systems. He received his BA in psychology from Grinnell College, Iowa, and his MSc and PhD degrees from the University of Pittsburgh in experimental psychology. He has done work on animal learning and memory, and on computer simulation during his 20-year career. Dr. Spelt is a member of SCS, AAAI, and IEEE.

Address: Dr. Philip F. Spelt
CESAR, Bldg 6025, ms 6364
Oak Ridge National Laboratory
P. O. Box 2008
Oak Ridge, Tennessee 37831-6364
(615) 574-7472

REFERENCES

Duda R. O. and P. E. Hart. "Pattern Classification and Scene Analysis", John Wiley & Sons, New York, 1973.

Glover. C. W. and M. Walker. "A Study on Automatic Artificial Neural Network Training Procedures", Oak Ridge National Laboratory Technical Memorandum, in preparation.

Glover, C. W., M. Silliman, and M. Walker. "Shape Classification of Objects in Images using Zernike Moments and Artificial Neural Networks", Oak Ridge National Laboratory Technical Memorandum, in preparation.

Hirschberg, D. S. "A Linear Space Algorithm for Computing Maximal Common Sequences", Comm. ACM, Vol. 18, no. 6, pp. 341-343, 1975.

Jones, J. P., R. C. Mann, and E. M. Simpson. "A Computer Vision System for a Hypercube Concurrent Ensemble", Oak Ridge National Laboratory TR-89-037, 1988.

Michalski, R. S., J. G. Carbonell, and T. M. Mitchell. Machine Learning, Vol. I, Tioga Publishing Co., Palo Alto, 1983.

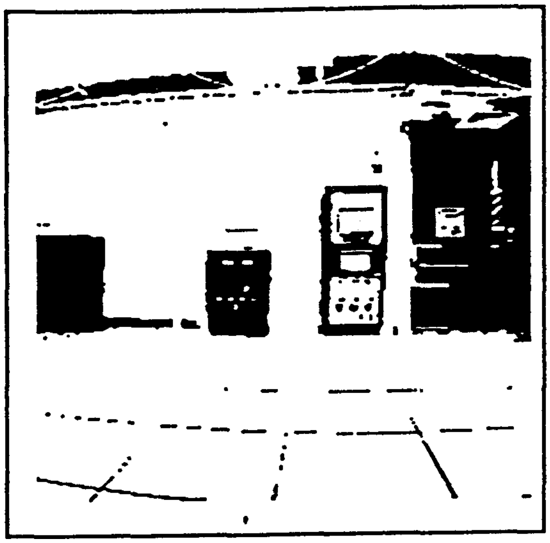
Rao, N. S. V. and C. W. Glover. "Symbolic Pattern Processing System for Planar Spatial Structures", Dept. of Comp. Sci., Old Dominion Univ., TR-89-037, 1989 (Oak Ridge National Laboratory Technical Memorandum in preparation.)

Ruck, D. W. "Multisensor Target Detection and Classification", Masters Thesis, Air Force Institute of Technology, AFIT/GE/ENG/87D-56, Dec. 1987.

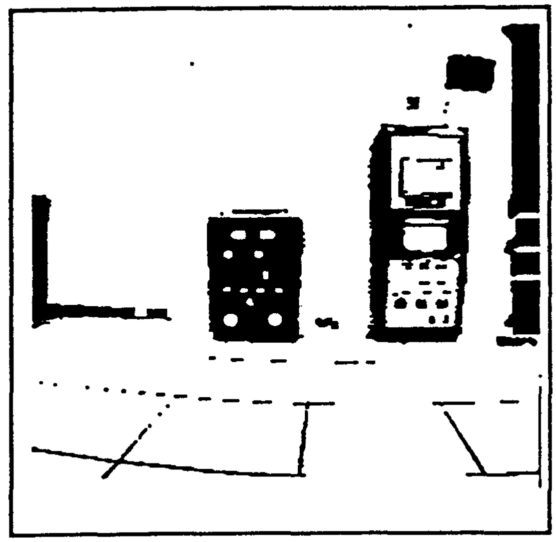
Rumelhart D. E. and J. L. McClelland. Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1, MIT Press, Cambridge, 1986.

Smolensky, P. "On the Proper Treatment of Connectionism", Behavioral and Brain Sciences, Vol. 11, pp. 1-74, 1988.

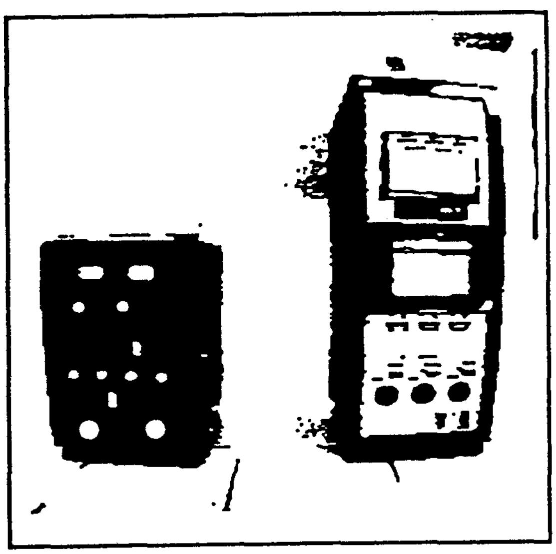
Teague, M. R. "Image Analysis via the General Theory of Moments", J. Opt. Soc. Am., Vol. 70, pp. 920-930, Aug. 1987.



(1a)



(1b)



(1c)

Fig. 1. Binary image of two control panels taken by a mobile robot from three progressively closer positions.

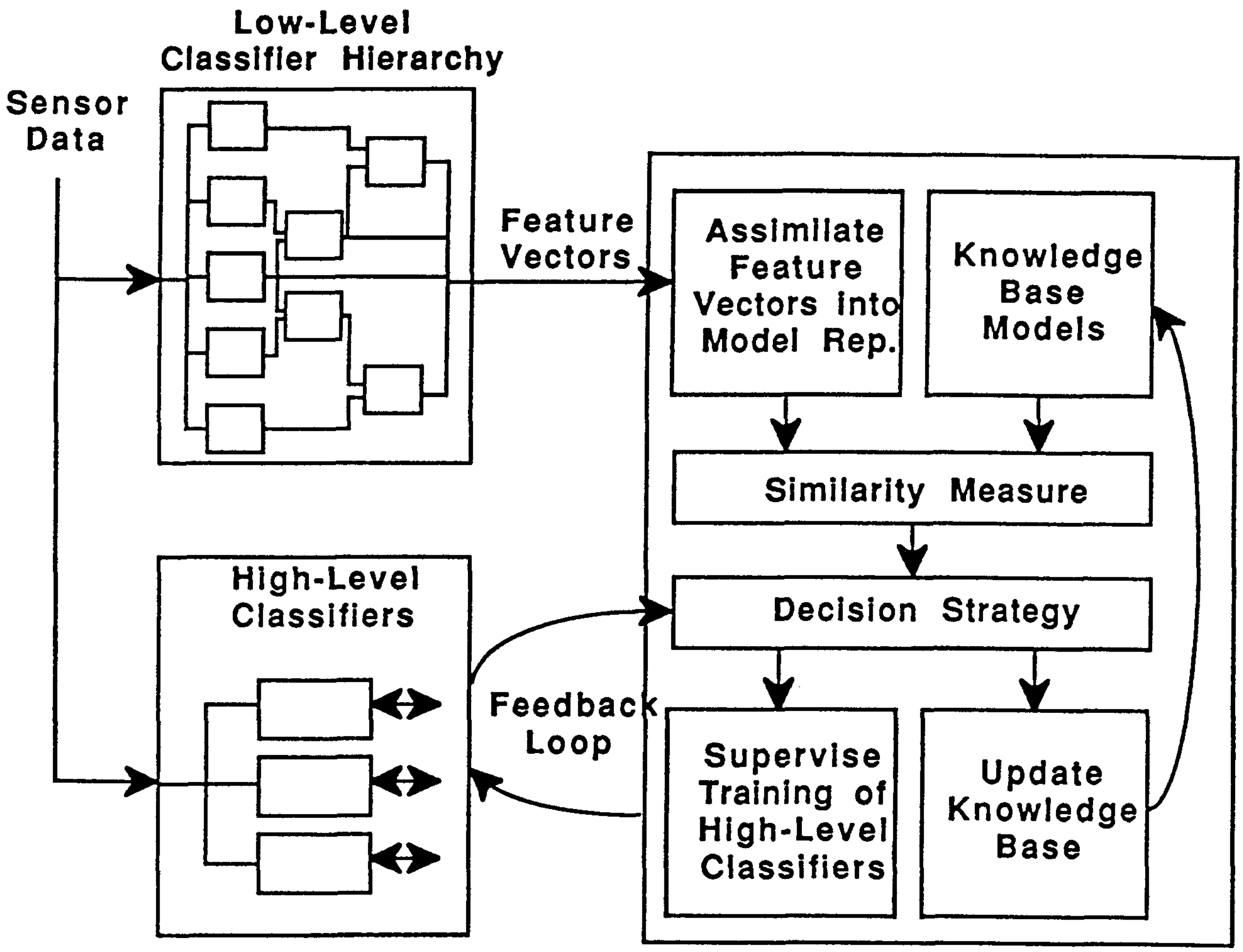


Fig. 2. Schematic diagram of a generic hybrid perception system.

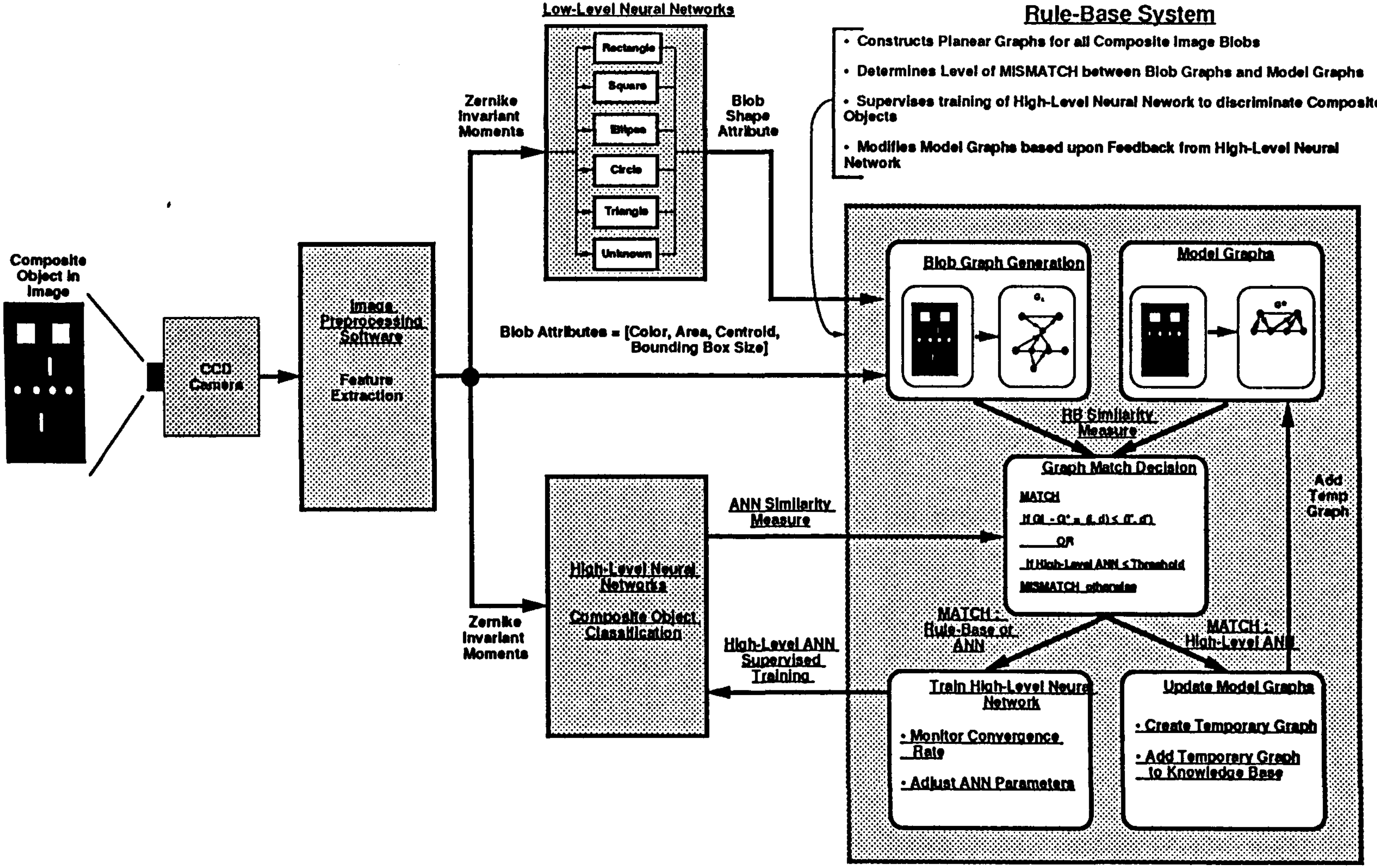


Figure 3. Detailed diagram of the hybrid perception system.