REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
		FINAL 30 SEP 92 TO 15 SEP 95

4. TITLE AND SUBTITLE
MACHINE VISION THROUGH MACHINE LEARNING

5. FUNDING NUMBERS
F49620-92-J-0549
62301E 8855/00

6. AUTHOR(S)
RYSZARD MICHALSKI

AFOSR-TR-96

0161

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
GEORGE MASON UNIVERSITY
4400 UNIVERSITY DRIVE
FAIRFAX VA 22030

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFOSR/NM
110 DUNCAN AVE SUITE B115
BOLLING AFB DC 20332-0001

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
F49620-92-J-0549

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE:
DISTRIBUTION UNLIMITIED.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)
SEE REPORT FOR ABSTRACT

19960502 031

14. SUBJECT TERMS

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAR

FINAL TECHNICAL REPORT ON THE PROJECT

Machine Vision Through Machine Learning

Grant No F49620-92-J-0549
30 September, 1992 - 15 September 1995

(a) PROJECT OBJECTIVES

This research has been concerned with the development of initial methodologies and vision systems capable of learning descriptions of visual objects or scenes, and the application of the learned descriptions to the efficient recognition of objects in a scene. The underlying motivation for this project is that learning capabilities will make computer vision systems adaptable to a wider range of practical problems than current vision systems that in most cases lack learning capabilities.

In this project, we concentrated on the following topics:

1) Development of the MLT ("multilevel logical templates") methodology for learning image transformations that characterize classes of visual objects.

2) Implementation of the MLT methodology and its application to the acquisition of texture descriptions by learning them from object samples presented in a scene under varied perceptual conditions and noise.

3) Development of methods that use a simple form of analogy for learning visual concepts (the PRAX project).

5) Application of the developed methods and systems to selected practical problems in the area natural object recognition, object detection in a scene, and target recognition.

(b) STATUS OF THE RESEARCH EFFORTS

Below is a brief description of the results obtained.

1) DEVELOPMENT OF THE MULTILEVEL LOGICAL TEMPLATES METHODOLOGY FOR LEARNING VISUAL CONCEPT DESCRIPTION

The *multilevel logical templates* (MLT) methodology has been developed for training a vision system to perform a given set of vision tasks. The methodology, developed by Michalski and

implemented by Bala, consists of three phases: 1) image marking, 2) automated model development, and 3) model testing (Figure 1).

In **Phase 1**, an operator selects and classifies samples from a training image that represent visual concepts to be learned (e.g., specific objects, parts of a scene, etc.)

In **Phase 2,** the system iteratively executes the following sequence of modules: Training Input Formulation, Model Learning and Refinement and Model Testing.

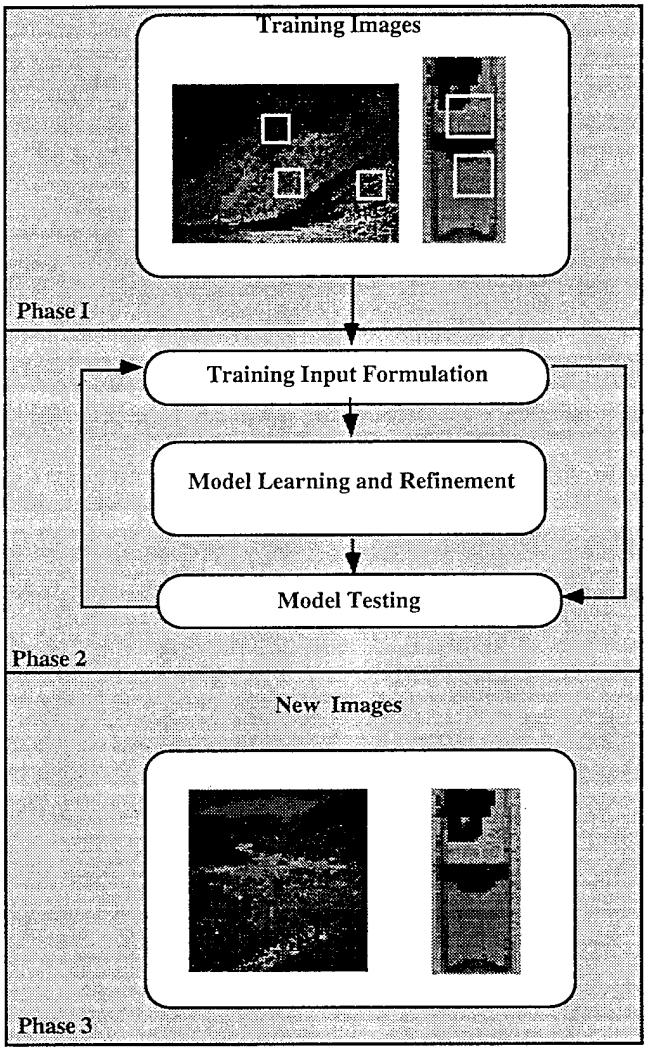

Figure 1: MLT Methodology.

The **Training Input Formulation** module performs two basic steps: 1) optimizing the image volume (by adjusting the resolution and the number of gray levels accordingly to the given vision task), 2) computing high-level features from the training image samples, and 3) creating "training events," which constitute input to the learning process. The **Model Learning and Refinement** module executes a learning system to determine general descriptions of indicated visual concepts from the given samples (and background knowledge).

At each iteration, the generated descriptions are applied to the whole training area of the image and a "symbolic" image is created, in which the "pixels" denote numerical labels of the visual concepts being learned. The descriptions are called "logical templates," because in the original implementation of the methodology they were logic-style decision rules that will be applied to the image in parallel.

The **Model Evaluation** module evaluates the quality of the descriptions obtained at a given iteration by relating the symbolic images they produce to the target image. If the descriptions need further improvement, the process is repeated as the current symbolic image is input. The process ends when the obtained symbolic image is sufficiently close to the target image labeling (indicating the "correct" labeling of the image). Complete object descriptions are sequences of image transformation operators (rule sets) that produce the output image, and serve as symbolic object models.

Phase 3 involves an application of the learned models to new images, to compute confidence scores for recognition.

To recognize an unknown surface sample, the system matches it with candidate surface descriptions. This is done by applying decision rules to the events in the sample. For each event, the class membership is determined. To increase the confidence of recognition, the majority class of the events in a window is taken as the decision.

Advantages of this approach are that the recognition process can be very fast, as it is amenable to parallel execution, and that the recognition accuracy for new images is very high.

The MLT methodology has been initially applied to learning multilevel rules characterizing given surface classes from surface samples [Michalski et. al., 1993]. The rules were determined using the inductive learning program AQ-15 [Michalski et al., 1986] and represented in the VL_1 logic-style language (Variable-Valued Logic System 1) [Michalski, 1972]. These rules serve as "logical templates" that can be matched in parallel or sequentially against window-size samples of surface to classify the image.

2) DEVELOPMENT OF THE P-MLT METHODOLOGY FOR LEARNING VISUAL CONCEPT DESCRIPTIONS

The aim of the *Parallel Multiple-level Logical Templates* (P-MLT) methodology is to extend the original MLT methodology by combining rule-based and neural nets learning in order to increase the speed of image processing and recognition.

A preliminary system implementating the P-MLT methodology, called AQ-ANN/1, works in two stages. In the first stage, a set of decision rules in the VL_1 (Variable-valued Logic System 1) which approximately characterize objects of interest are induced from examples. In the second stage, the rules are transformed into an equivalent neural one layer neural net, and the resulting neural net is further trained to improve its recognition performance.

The AQ algorithm generates decision rules in a "greedy" fashion, at each step determining one rule that covers a maximal portion of the "uncovered" training data, and so on until all positive training examples are covered, and all negative examples are excluded. To create rules from examples, it employs "inductive generalization operators" that make the decision rules as general as possible without becoming inconsistent [Michalski, 1972; Michalski et al., 1986]. When noise is present in the training data, the rules are allowed to be partially inconsistent and/or incomplete with regard to the input data.

The learning process is executed in two phases:

1. Rule learning using the AQ algorithm.
   This phase generates rules that describe the training examples (those that cover only a few examples are truncated from class description).

2. Backpropagation network learning.
   Each node in a one-layer network corresponds to a single rule. The degree of match of an example to the node rule represents node activation. This activation value is input to the sigmoid transfer function associated with each node. Weight values for the connections between nodes and outputs are obtained using the backpropagation learning mechanism.

The node rules in the network are a form of receptive field transfer function. The network architecture is similar to the Radial Basis Function network (RBF network). The RBF network models data by a Gaussian distribution function associated with each node. The network generated by the AQ algorithm is constructed based on rules that represent generalization of the initial examples. Our approach overcomes two important drawbacks of RBF learning algorithms, namely, choosing the right number of nodes (clusters to be modeled by the Gaussian distribution) and the measure of the spread of the data associated with each cluster.

Deliverables:

An initial method for learning surface descriptions using the P-MLT approach.

3) DEVELOPMENT OF THE DYNAMIC RECOGNITION METHODOLOGY

Currently, there are two major approaches to object recognition: model-based recognition, and feature-based classification. In model-based recognition an instance of one or several known objects are located in an image, using their geometric models. A big drawback of this approach is that it is only feasible for a small number of possible objects. A bigger drawback is that it only works for objects whose geometry is precisely known and thus excluding many objects such as most natural objects which don't have a precise and well defined geometry.

In feature-based classification, object instances are assigned to classes based on vectors of feature values. Usually in a system using this approach, the feature extraction and classification are two isolated processes. The feature extraction module first extracts all the relevant features of the image which are necessary for achieving correct classifications of all objects which the system is trying to recognize. The classifier will then classify the image by comparing these extracted features to those from the models stored in the database. The disadvantage of such a system is that in order to recognize an object, it needs to always measure the same properties of it, namely all its relevant features. This is, however, not desirable since extracting all the relevant features can be computationally very expensive and is not always possible. This is specially true for a system which recognizes a large number of classes since this usually requires extraction of a large number of features in order to discriminate between those classes. The other major drawback of feature-based classification is its limited descriptive power because in such an approach there is no explicit means for describing structural properties of objects. Therefore this approach is not suitable for recognizing complex objects for which structural properties are critical for the classification process.

We have developed an alternative approach to recognition. Our approach combines the descriptive powers of both model-based and feature-based approaches for building *characteristic* descriptions of objects. A characteristic description of an object is a collection of all information known about it. This information includes all the known object features such as color, size, and texture as well as

structural properties of the object. However, in order to classify instances of objects by using these characteristic descriptions, we use the Dynamic Recognition methodology which is fundamentally different from both model and feature vector matching. The main idea behind Dynamic Recognition which was originally introduced by Michalski in 1986 is that the system determines "key" attributes from characteristic description of objects. These attributes are determined by conducting inductive inference on candidate object descriptions.

The proposed Dynamic Recognition methodology involves three steps:

1- REDUCE
2- INDUCE
3- INQUIRE

In the REDUCE step, some "striking features" of objects in the image are used to reduce existing characteristic descriptions and determine candidates. In other words, all the descriptions which are not satisfied by the values of these features are removed from the set of candidate descriptions. The "striking features" are somewhat domain dependent, but they are usually those features which are easily detectable by the system such as color and size. In addition some background knowledge can be used to further reduce the set of possible candidates. For instance if we know that the image was taken from the sky, then the set of possible candidates will include those objects which are expected to be found in the sky.

In the INDUCE step a learning program is applied to the reduced set of characteristic descriptions to determine the simplest *admissible discriminant* descriptions. These discriminant descriptions will usually contain only the discriminant features, i.e., fewer features than the original characteristic descriptions. These discriminant descriptions should contain only measurable features in the given context. For example color can not be considered as a discriminant feature in a gray scale image or area of the object should not be considered as a discriminant feature if the object is known to be occluded.

In the INQUIRE step, an evaluation function is applied to each remaining feature in the discriminant descriptions and the value of the feature with the highest score is extracted from the image of the object to be recognized. An important parameter of this evaluation function is the *cost* of the feature, which measures the difficulty of extracting it from the image. Rules not satisfied by the value of the extracted feature are removed from the set of candidate descriptions. The INQUIRE step is repeated until we are left with one candidate description, namely the description of the object in the image.

Thus, in the Dynamic Recognition methodology, recognition is considered as an inductive inference process that determines the discrminant features of the objects in a given context, and not as a matching process.

Deliverables:

1. An initial methodology for dynamic recognition.
2. Results of Learning system: DR (partial implementation).

Papers:

Michalski, R., Bala J., Pachowicz P. "GMU RESEARCH ON LEARNING IN VISION: Initial Results," *Proceedings of the 1993 DARPA Image Understanding Workshop,* Washington DC, 1993.

Bala J., Michalski, R., Pachowicz P. "GMU Research on Learning in Vision," *Proceedings of the 1994 ARPA Image Understanding Workshop,* Monterey, November, 1994.

4) THE PRAX METHOD FOR DETERMINING DESCRIPTIONS FOR A LARGE NUMBER OF VISUAL CONCEPTS

Most research on concept learning from examples concentrates on algorithms for generating concept descriptions of a relatively small number of classes. In conventional methods, when the number of classes if growing, their descriptions become increasingly complex. In some computer vision applications, the number of classes may be very large, and they may not be known entirely in advance. Therefore, in such situations, the learning method must be able to learn incrementally new classes. Such a *class-incremental* mode is different from the conventional *event-incremental* mode, in which examples of classes are supplied incrementally, but the set of classes remains unchanged.

The PRAX method is specifically oriented toward learning descriptions of a large number of classes in a class-incremental mode. The learning process consists of two phases. In Phase 1, symbolic descriptions of a selected subset of classes, called *p*rincipal ***axes*** (briefly, *praxes*) are learned from concept examples (here, samples of textures). The descriptions are expressed as a set of rules. In Phase 2, the system incrementally learns descriptions of other classes (*non-prax* classes). These descriptions are expressed in terms of the similarities to praxes, and thus the second phase represents a form of analogical learning. To utilize a uniform representation, the prax descriptions are also transformed into a set of similarities to the original symbolic descriptions.

Deliverables:

1. A methodology for learning PRAX-based concept descriptions.
2. Learning systems: PRAX-1, PRAX-2.

Papers:

Bala, J., Michalski R., and Wnek J., "Learning a Representation for Efficient Recognition of a Large Number of Visual Concepts," *1993 AAAI Fall Symposium on Machine Learning in Computer Vision*, (AAAI Technical Report FS-93-04), Raleigh, NC, Oct. 22-24, 1993.

Bala, J., Michalski R., and Wnek J., "The Principal Axes Method for Constructive Induction," *Proceedings of the Ninth International Conference on Machine Learning,* Aberdeen, Scotland, July 1992.

5) NOISE-TOLERANT LEARNING OF OBJECT MODELS FROM COMPLEX SENSORY DATA

This project aims at the development of new techniques for learning from very complex and noisy sensory attributional data. The guiding premise of this research is that erroneous data can be detected more effectively on the model level — where relationships between data clusters and between classes to be learned is expressed better than in raw training data. These techniques are dedicated for symbolic learning programs, however, they can also be adapted to the other classifiers.

Model acquisition from noisy data sets is a difficult problem for symbolic learning programs when applied to image analysis domain. Inductive learning systems perform a generalization of the input data in order to anticipate unseen examples. In a standard mode, when all the input examples can be assumed to be correct, a concept description generated by an inductive learning system should be complete (cover all training examples) and consistent (cover no examples of other concepts). In the case of noisy data, the system does not seek such complete and consistent descriptions. There are two basic approaches to symbolic learning from noisy data. The first approach, tree pruning (elimination of some subtrees from the learned decision tree), taken by the ID family of algorithms, allows a certain degree of inconsistent classification of training examples so that the descriptions will be general enough to describe the basic characteristics of a concept. The second approach, taken by the AQ family of programs, is to remove some of the unimportant rules (or conditions) from a set of rules, and retain only those covering the largest number of examples. Traditional learning methods based on pruning/truncation try to handle noise in one step. Therefore, they share a common problem: the final concept descriptions are based on the initial noisy training data.

A new approach has been developed which extends the traditional one-step method of noise handling to a closed-loop two- or multiple-step process. The learning loop can be run once or multiple times with changing learning and/or truncation/pruning criteria. This learning loop includes:

1) Concept acquisition by a concept learning system such as AQ or ID;
2) Evaluation of learned class descriptions, detection of less significant disjuncts/subtrees, which are not likely to represent patterns in the training data, and removal of detected rules/subtrees; and
3) Filtration of training data through optimized rules/trees (i.e., removal of all examples not covered by truncated or pruned concept descriptions).

In this approach, pruned/truncated concept descriptions are used as a filter to improve the training data set. Then, the concept acquisition phase is repeated from the improved training data. Consequently, those training examples which caused the generation of pruned/truncated concept components are no longer taken into account when concept learning is repeated. Since the detection of erroneous examples is executed on the concept description level rather than on the input data level, data filtration reflects attribute combination in the construction of concept descriptions and inter-class distribution over the attribute space.

Initially, we prototyped a version of a rule learning program and showed basic results for a texture recognition problem involving six texture classes. We reported that the recognition rate increased and the complexity of object models decreased substantially.

Next, we implemented the above approach to rule learning and decision tree learning programs and tested them on several vision problems. The new version of the learning program AQ-NT uses the AQ14 learning program. The decision tree version, the ID-NT program, uses the C4.5 learning program. Both programs were tested on the acquisition of attributional descriptions of twelve similar texture classes from texture energy measures. Different image sections were used for training and for testing. We notice improvement in the recognition error and the stability of the recognition system over increasing pruning/truncation levels. For higher truncation levels the maximum error rate stabilized and the recognition of the worst recognizable class improved substantially. The results were compared to other learning programs.

Recently, the developed noise-tolerant learning method was tested on real images of natural outdoor scenes composed of the "Grass" area, "Tree" area, and "Rocks" area. All the images were taken in different places but in the same mountain area. The images were characterized by (i) the lack of a clear border area between the "Grass" area and the "Tree" area, (ii) many isolated large

rocks, (iii) overlap of the "Grass" area and the "Rocks" area, and (iv) difficulty in the interpretation of some small image region. There were difficulties with the precise segmentation of the test image when other learning programs were used. However, using the developed approach we achieved two major improvements. First, the distinction between the "Tree" area and the "Grass" area was improved substantially. Second, the false classification of large grass sections was eliminated. Moreover, the segmented images better highlighted surface details corresponding to large rocks and small bushes.

Deliverables:

1. A methodology for learning object descriptions from noisy sensory data using symbolic learning programs.
2. Two prototype learning systems: the AQ-NT rule-based learning program (demo) and the ID-NT decision-tree learning program.

Papers:

J. Bala and P.W. Pachowicz, "Recognizing Noisy Patterns via Iterative Optimization and Matching of Their Descriptions," *International Journal on Pattern Recognition and Machine Intelligence*, Vol.6, No.4, pp.513-538, 1992.

P.W. Pachowicz, J. Bala and J. Zhang, "Methodology for Iterative Noise-Tolerant Learning and Its Application to Object Recognition in Computer Vision," *Proceedings of the Int. Conf. on Systems Research, Informatics and Cybernetics 92*, Baden-Baden, August 1992.

P.W. Pachowicz, J.Bala and J. Zhang, "Iterative Rule Simplification for Noise Tolerant Inductive Learning," *Proceedings of the IEEE Conference on Tools with AI*, Arlington, VA, pp.452-453, 1992.

J. Bala and P.W. Pachowicz, "Issues on Noise Tolerant Learning from Sensory Data," *Proceedings of the AAAI Symposium on Machine Learning and Vision*, Raleigh NC, pp.135-138, 1993.

P.W. Pachowicz and J. Bala, "A Noise-Tolerant Approach to Symbolic Learning from Sensory Data," *Journal of Intelligent and Fuzzy Systems*, Vol.2, No.4, pp.347-361, 1994.

6) MODEL EVOLUTION PARADIGM TO OBJECT RECOGNITION IN DYNAMIC ENVIRONMENTS

This project aims at object recognition under the gradual change in perceptual conditions and/or under varying object appearances; the development of a new paradigm (related to *active vision*) for object recognition in dynamic environments.

Most past research on object recognition has been focused on learning to recognize objects under a given subset of stationary perceptual conditions (such as lighting, resolution and positioning) and for known object appearances (e.g., subsets of IR or SAR object signatures; subsets of object silhouettes). Object recognition in dynamic environments, however, has to deal with changes in perceptual conditions and object appearances not known to the system beforehand. Frequently, models learned under given perceptual conditions are not effective in recognizing objects under other conditions. This problem is particularly severe for object recognition in outdoor environments where the variability of perceptual conditions and object appearances can be

extremely high.

Most approaches to object recognition do not adapt an object recognition system directly to changing perceptual conditions and object appearances. These methods use stationary models acquired during the off-line training phase. Such an approach requires that each condition influencing the change of object characteristics is represented in the model, a conclusion which is hard to satisfy for realistic environments.

We have developed a model evolution paradigm (called, CHAMELEON) for object recognition under variable perceptual conditions and changing object appearances. The paradigm relies on the on-line dynamic modification of object models according to perceived changes in object characteristics. This paradigm was tested for a scene segmentation problem based on texture characteristics of surfaces. It assumes that a change in, for example, texture characteristics is gradual and is reflected in the images of a sequence. Given texture descriptions (models) learned from the first image of a sequence, the system applies these descriptions to the next image to recognize the objects. Then, the system computes a recognition confidence for each object and compares the results with those obtained when working with the previous images. Dynamic characteristics of the confidence change are modeled. If the recognition confidence deteriorates, so that the system will have more problems in recognizing the object in the next image, the system indicates which descriptions must be modified and activates data selection and learning processes. New training examples, which represent the change in object characteristics, are selected and provided to an incremental learning program. The modified models are verified to insure the soundness of the evolution process.

Using the model evolution paradigm, a vision system adapts to the changes in the environment by adapting the object models on-line and autonomously. This allows for capturing any variability in object characteristics without knowledge about object properties and without building complex, dedicated modules to deal with changes in a given perceptual condition. Thus, an object model can be adapted to any combination of perceptual conditions. Moreover, the system can adapt to a change in the internal state of an object (e.g., to a change in a target's IR signature). Model evolution is an active agent process actively working on its internal knowledge and models of the environment and the objects. Model evolution includes (but is not limited to) and integrates: vision processes, model evaluation, reasoning about the models, guidance for model modification, data selection, and control processes. A kernel of the model evolution system is an incremental learning program.

We have developed the CHAMELEON-1 (semi-autonomous evolution) and CHAMELEON-2 (fully autonomous evolution) systems for the recognition of textures and for texture-based scene segmentation under gradual changes in resolution and lighting conditions. The CHAMELEON-1 and -2 systems were intensively tested. We used different image sequences and different control strategies for the selection of new training data for model evolution. We investigated the soundness of the model evolution in critical situations — i.e., situations where the system mistakenly selects incorrect data or the dynamics of model evolution is too slow when compared to the dynamics of the change in object characteristics.

Conclusions from the development and testing of the CHAMELEON-1 and -2 systems have been used in the design of a new framework for a new CHAMELEON-3 system. This new system will apply a Bayes classifier and later a radial basis function classifier (RBF) to serve as the incremental learning kernel of a model evolution system. The new kernels will be capable of modifying the models more effectively using:

(i) statistical information and/or selected new training data,
(ii) gradient information about the direction and the dynamics of model change within the attribute space, and

(iii) prediction of model change beyond the image sequences already seen.

We also investigated (1) architectures for the integration of vision and learning processes of model evolution particularly for automatic model evolution guidance, (2) problems with instability in model evolution, and (3) different strategies for the selection of new training examples for model modification in the incremental mode. We also developed a synthetic environment for the CHAMELEON-3 model evolution system.

Deliverables:
1. A methodology to object recognition in dynamic environments.
2. CHAMELEON-1 system for semi-autonomous evolution of object models.
3. CHAMELEON-2 system: for autonomous evolution of object models for scene segmentation.

Papers:

P.W. Pachowicz, M. Hieb and P.Mohta, "A Learning-Based Incremental Model Evolution for Invariant Object Recognition," *Proceedings of the Int. Conf. on Systems Research, Informatics and Cybernetics 92*, Baden-Baden, August 1992.

P.W. Pachowicz, "A Learning-Based Evolution of Concept Descriptions for an Adaptive Object Recognition," *Proceedings of the IEEE Conference on Tools with AI*, Arlington, VA, pp.316-323, 1992.

P.W. Pachowicz, "Invariant Object Recognition: A Model Evolution Approach", *Proceedings of the DARPA Image Understanding Workshop*, Washington DC, pp. 715-724, 1993.

P.W. Pachowicz, "Semi-Autonomous Evolution of Object Models for Adaptive Object Recognition," *IEEE Trans. on Systems, Man and Cybernetics*, Vol. 24, No. 8, pp.1191-1207, 1994.

7) AUTONOMOUS VISION AGENTS: LEARNING, EVOLVING AND SELF-GOVERNING

This research project aims at the design and development of adaptability mechanisms for a vision module which is already prestructured for application-specific data gathering and/or image analysis/understanding. These mechanisms will allow a vision module to undergo on-line modification of its internal knowledge/models, structure and/or processes in an active manner.

This research focuses on how an autonomous vision agent can manage itself while working in dynamic environments, under varying task parameters, and employing dynamic links with associated subsystems. This is due to the following basic issues that the agent has to deal with on-line:

(i) change in scene complexity influencing the time, quality and complexity of processes needed for image analysis/understanding,
(ii) change in object appearances, influencing the change of object/scene models,
(iii) occurrence of unexpected situations the system has barely been trained to deal with,
(iv) on-line change in task parameters, and
(v) interruptions/requests coming from processes that the agent communicates with (sensor hardware, host task processes, and application processes).

Sensory systems working in realistic dynamic environments may have to deal with one or more of

these issues in order to become autonomous and no longer rely on an engineer to reconfigure the system. An autonomous vision agent must be able to minimize the impact of these issues on its perceptual skills.

The way we have chosen to realize this goal is to develop an active vision agent (AVA) which will be capable of modifying its internal resources over a sequence of images affected by situations which differ from those the system was prestructured for. We have designed a framework for an AVA. This framework includes the following three elements which will insure the system's adaptability to changes in environments, parameters of perceptual tasks, and interactions with the other processes of the application system:

1) introduction of different learning functions into the agent's data processing/analysis algorithms,
2) introduction of model evolution processes into the agent's model/knowledge base, and
3) introduction of self-governing processes into the agent.

The first element of an AVA, learning functions for data analysis algorithms, allows the agent to optimize itself to operate better and faster for repetitive tasks/conditions. Using these functions, the system constantly looks for better data analysis solutions through a network of prestructured/available image analysis procedures. This recently initiated research has shown how the introduction of learning functions within the traditional *train-recognize* paradigm can transform this paradigm into an active agent paradigm.

The second element of an AVA, model evolution, insures system adaptability to changing object appearances and perceptual conditions not reflected in the initial models. We have developed and tested model evolution systems operating in semi-autonomous and fully autonomous modes for scene segmentation and recognition tasks.

The third element of an AVA, the self-governing aspect, supports automatic reconfiguration of agent processes due to changes in scene complexity, time restrictions, task parameters, external requests, and dynamics of the environment. This research has roots in our previous work where we showed how a vision system can restructure itself on-line using simple image measures and a feedback control loop. Our recently developed framework for an AVA includes self-governing functions for the agent through the use of the following tools:

(i) *Focus-of-Attention*: allowing for selective analysis of local image data and/or time events,
(ii) *Resolution-on-Demand*: allowing for accessing data at appropriate levels of detail,
(iii) *Abstraction-on-Demand*: allowing for accessing models/knowledge on appropriate levels of competence, and
(iv) *Event-on-Pipeline*: allowing for incremental analysis of scene objects and events over image sequences.

This research is continued and reflects our belief that by introducing this paradigm into machine perception, autonomous vision agents will gain enough degrees of freedom to adapt to changing external influences.

Deliverables:
An approach to design autonomous vision agent; a system which will achieve adaptability to environments through control and learning mechanisms.

Papers:

P.W. Pachowicz, "Invariant Object Recognition: A Model Evolution Approach", *Proceedings of the DARPA Image Understanding Workshop*, Washington DC, pp. 715-724, 1993.

P. W. Pachowicz, "Integration of Machine Learning and Vision into an Active Agent Paradigm," *Proceedings of the AAAI Symposium on Machine Learning and Vision*, Raleigh NC, pp.110-114, 1993.

8) DEVELOPMENT OF A MULTISTRATEGY MODEL ACQUISITION METHOD FOR MODEL ACQUISITION ACCORDING TO MULTIPLE RECOGNITION OBJECTIVES.

This method integrates two forms of learning, inductive generalization and genetic algorithms, in a closed-loop fashion in order to achieve robust concept learning capabilities. The learning process cycles between two phases (Figure 2): an inductive learning phase and a genetic algorithm phase. In the inductive learning phase cognitively-oriented concept descriptions are produced in standard disjunctive normal form (DNF). In the GA phase the performance of these concepts is improved using a set of tuning data. After the concepts are modified, they are refined again by the AQ algorithm resulting in somewhat simpler descriptions. In this way, the learning loop is closed and two learning modules are able to exchange concept descriptions while improving them according to different criteria.

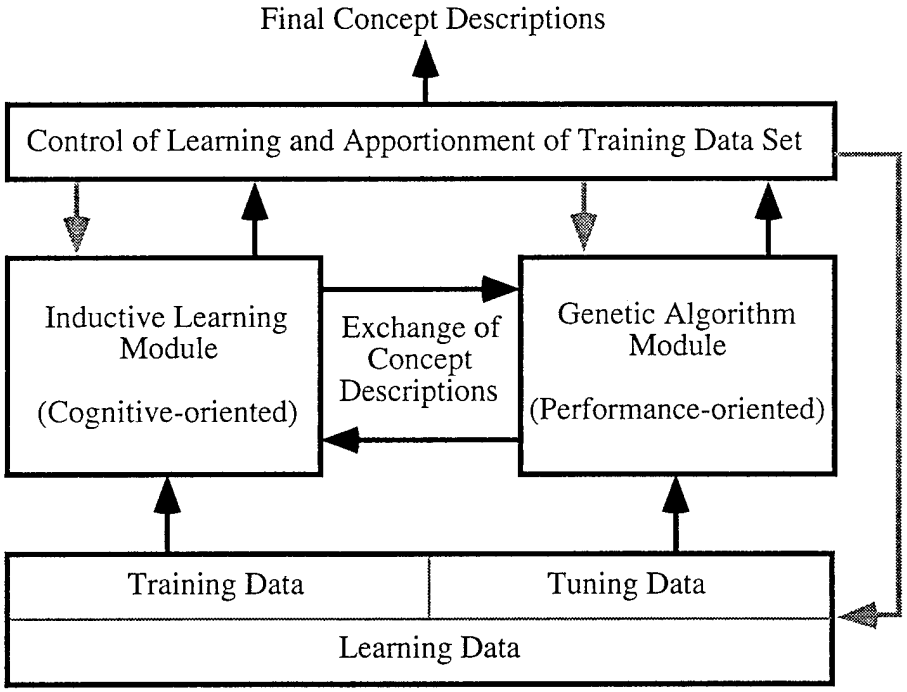

Figure 2: An architecture for the AQ-GA system.

By combining inductive learning with genetic algorithms, the learned concept descriptions are no longer required to be complete and consistent with respect to the initial training data, which reduces overfitting problems and leads to better predictive performance. Also, the use of GAs reduces the effects of noise on the learned concept descriptions. The fact that the GA starts its search with plausible AQ-generated concept descriptions results in much shorter search times. Also, the performance-oriented GA search provides the ability to escape from some of the local minima traps resulting from AQ biases. The method was successfully applied to texture recognition problem.

Deliverables:

Learning system: AQ-GA

Papers:

Bala, J., DeJong K., and Pachowicz P., "Multistrategy Learning from Engineering Data by Integrating Inductive Generalization and Genetic Algorithms," in *Machine Learning: A Multistrategy Approach*, Vol. IV, R.S. Michalski and G. Tecuci, (Eds.), Morgan Kaufman, San Mateo CA., 1994.

Bala, J., K. DeJong and P. Pachowicz, "Using Genetic Algorithms to Improve the Performance of Classification Rules Produced by Symbolic Inductive Methods," vol. 542 Springer-Verlag Lecture Notes in Computer Science, Oct. 16-19, 1991.

Bala, J., K. DeJong, P. Pachowicz, "Integrated Inductive Learning And Genetic Algorithms For Texture Recognition," *Proceedings of the ML92 Workshop on Integrated Learning in Real-World Domains, Aberdeen,* Scotland, July 1992.

Bala, J., DeJong K., and Pachowicz P., "Integration of Inductive Learning and Genetic Algorithms to Learn Optimal Descriptions from Engineering Data," *International Workshops on Mulistrategy Learning*, Harpers Ferry, West Virginia, Nov., 1991.

9) LEARNING TO RECOGNIZE 2D SHAPES IN X-RAY IMAGES.

The goal of this research is to develop a methodology for applying symbolic inductive learning techniques to the machine vision problem of object recognition. Presently, the methodology consists of 5 steps, which are pictured in Figure 3. The first step is Region of Interest (ROI) determination, in which image objects that are potentially of interest are determined by image processing and low-level vision manipulations. Step 2, Event Extraction, is designed to extract features using regions of interest and to produce classified training examples expressed in a representation space suitable for inductive learning. If using a symbolic inductive learning system, such as AQ15c, a discretization step is necessary (Step 3) to abstract the training examples into a discrete representation space. To properly validate a learning system, training events are split into training and testing sets, according to a specific validation methodology that is determined based on the number of training events. The training set is used for learning (Step 4), which produces generalized concepts of the original training examples. These concepts can then be used to recognize unknown objects or shapes coming from the original representation space. Following learning, the testing set is used to validate inductively learned concepts in a recognition phase (Step 5), which produces a classification for each of the classified examples in the testing set. Because examples in the testing set are classified, the learner's performance can be determined by calculating the percentage of testing events that were correctly classified during the recognition phase.

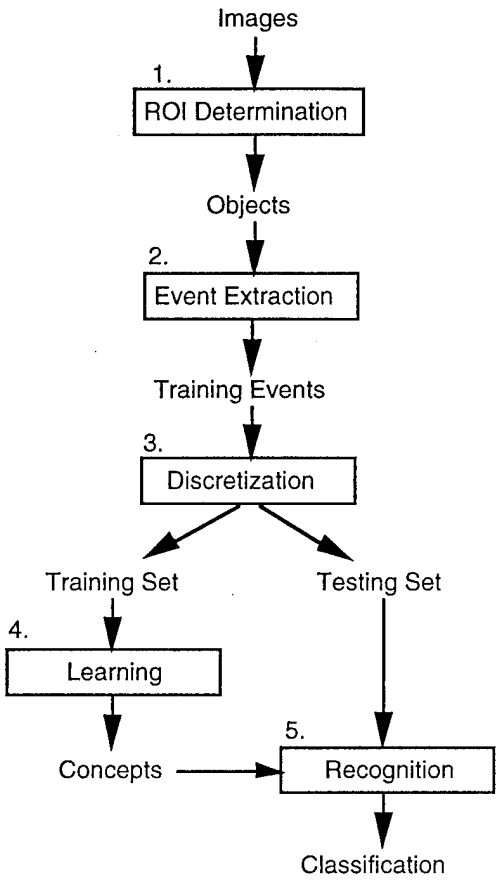

Figure 3: Five step learning and recognition methodology.

To demonstrate the viability of this methodology, it is being applied to a variety of image sets that contain a shapes under varying perceptual conditions. Initial results using this methodology were reported by Maloof and Michalski (1994). The results reported here are for an image set of x-rayed airport luggage containing blasting caps (see Figure 4). One potential application of this research is for an intelligent system to assist airport security personnel with luggage screening.

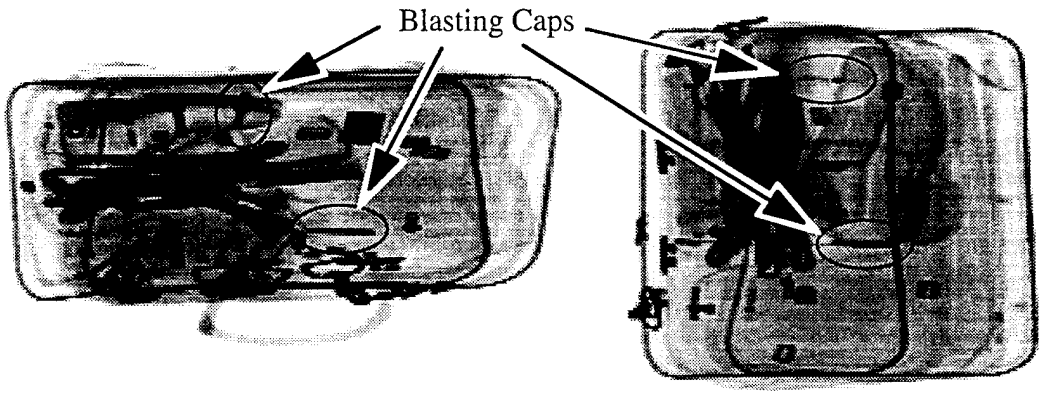

Figure 4: Sample x-ray images of luggage containing blasting caps.

Experimental comparisons were made between three learning methods: AQ15c, a symbolic inductive learning system, a feed-forward artificial neural network (ANN), a non-symbolic inductive learning method, and *k*-nn, a statistical pattern recognition technique. These learning methods were compared using average predictive accuracy, best predictive accuracy, and average learning and recognition times. The validation methodology used for each learning method consisted of 500 learning and recognition runs. For each run, training events were divided evenly into disjoint training and testing sets. The training set was used for learning, while the testing set was used for recognition. A predictive accuracy was computed for the run, which was the percentage of correctly classified testing events. For a 500 run experiment, the average predictive accuracy is average of the predictive accuracies computed for each run. The best predictive accuracy for a 500 run experiment is the highest predictive accuracy the learner achieved on any single run. Finally, the average learning and recognition time for a 500 run experiment is the average amount of CPU time the system spent learning and recognizing. These results are presented in Table 1 and were also reported by Maloof and Michalski (1995d).

Learning Method	Average Predictive Accuracy (%)	Best Predictive Accuracy (%)	Average Learning and Recognition Time (seconds)
AQ15c	95	100	0.1
ANN	79	95	8.1
k-nn	69	88	0.03

Table 1: Performance summary for classification technique.

Deliverables:

1. A methodology for learning to recognize shapes.
2. A shape-learning system.

Papers:

Maloof, M. A., and Michalski, R. S. (1994) Learning descriptions of 2D blob-like shapes for object recognition in x-ray images: an initial study. *Reports of the Machine Learning and Inference*

Laboratory, MLI 94–4. Center for Machine Learning and Inference, George Mason University, Fairfax, VA.

Maloof, M. A., and Michalski, R. S. (1995d) Learning symbolic descriptions of 2D shapes for object recognition in x-ray images. 8th International Symposium on Artificial Intelligence. Monterrey, Mexico, October 16–20, 1995 (submitted).

10) INCREMENTAL LEARNING USING A PARTIAL MEMORY APPROACH

The goal of this research is to develop an incremental learning methodology to support active vision applications. The methodology, pictured in Figure 5, consists of a development phase, in which traditional concept learning is used to provide the system with its initial concepts, and a deployment phase, in which the system receives criticism and reinforcement from its environment and user and learns incrementally. The methodology maintains a partial memory consisting of representative examples that provide the learner with a historical context and decrease learning time. Mechanisms for determining representative examples and for aging and maintaining these examples permit the learner to incrementally acquire changing concepts over time, which is necessary not only for active vision applications, but also for intelligent agents and dynamic knowledge-based systems.

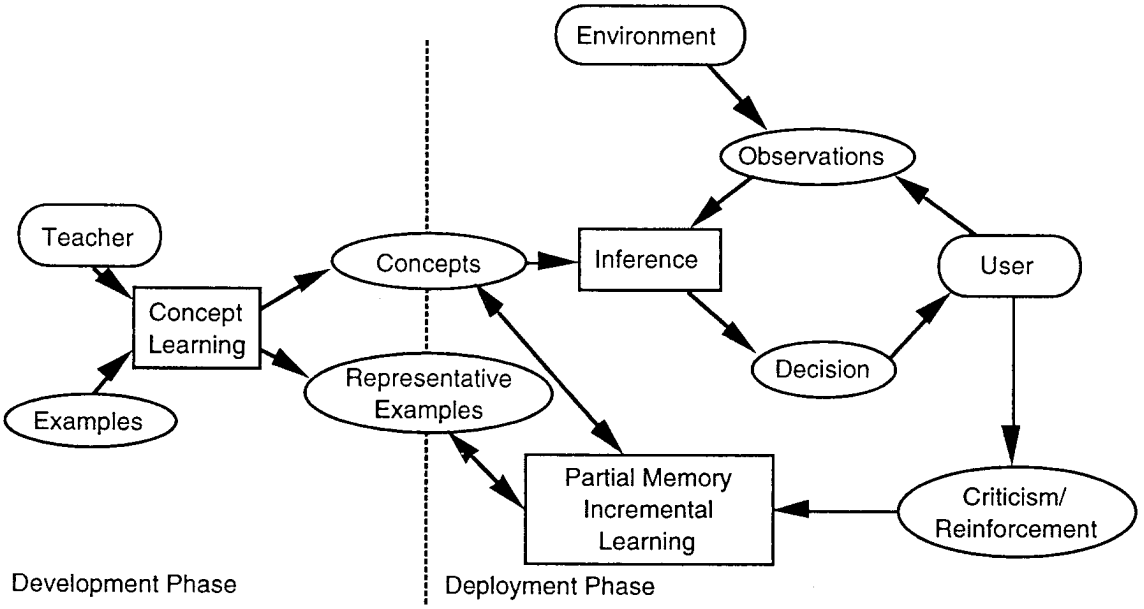

Figure 5: Partial memory incremental learning architecture and methodology.

Initial experiments have been conducted using the dynamic knowledge-based application of computer intrusion detection, in which use patterns are learned for computer users for anomaly detection. Experimental comparisons were made between three learning methods: AQ15c, a symbolic inductive learning system, a feed-forward artificial neural network (ANN), a non-symbolic inductive learning method, and *k*-nn, a statistical pattern recognition technique. These learning methods were compared using average predictive accuracy, best predictive accuracy, average learning time, and average recognition time. These learning methods were validated using 100 2-fold cross validation methodology, as follows. For each learner, 100 learning and

recognition runs were made in which classified training events were split evenly into disjoint training and testing sets. Learning was conducted on the training set, while the induced concepts were tested using the testing set. Since the examples in the testing set were classified, the learner's predictive accuracy can be computed as a percentage of the testing examples correctly classified. The amount of CPU time the system spent learning and recognizing can also be computed. For a 100 run experiment, the average predictive accuracy is simply the average predictive accuracy of the 100 learning and recognition runs. The best predictive accuracy for a 100 run experiment is the highest predictive accuracy achieved during any of the 100 learning and recognition runs. The average learning time is the average CPU time the system spent learning over the 100 run experiment. Finally, the average recognition time is the average CPU time the system spent classifying testing events during the 100 run experiment. These results are summarized in Table 2. These results are also reported by Maloof and Michalski (1995a, 1995b), while the methodology is discussed in general by Maloof and Michalski (1995c).

Learning Method	Average Predictive Accuracy (%)	Best Predictive Accuracy (%)	Average Learning Time (seconds)	Average Recognition Time (seconds)
k-nn	83	89	0.7	1.43
ANN	85	94	580.0	0.17
AQ15c	88	96	68.8	0.3

Table 2: Comparative summary of results for batch learning experiments.

Deliverables

A methodology for incremental.

Papers:

Maloof, M. A., and Michalski, R. S. (1995a) A partial memory incremental learning methodology and its application to intrusion detection. *Reports of the Machine Learning and Inference Laboratory*, MLI 95–2. Center for Machine Learning and Inference, George Mason University, Fairfax, VA.

Maloof, M. A., and Michalski, R. S. (1995b) A partial memory incremental learning methodology and its application to intrusion detection. 7th IEEE International Conference on Tools with Artificial Intelligence. Washington, DC, November 5–8, 1995 (submitted).

Maloof, M. A., and Michalski, R. S. (1995c) Learning incrementally changing concepts using a partial memory approach. AAAI 1995 Fall Symposium on Active Learning. Cambridge, MA, November 10–12, 1995 (submitted).

11) LEARNING HYBRID MODELS FOR ROBUST ATR FROM PARTIAL OR DISTORTED SAR DATA

This project aims at the development of an approach to target modeling from sequences of SAR signatures through the transformation of target data from raw pictorial representation into *feature*

token representation. (A feature token is a hybrid data object collecting significant invariant local characteristics of a target/object.) The transformation involves the detection, extraction, fusion and organization of local dominant information about the target into local *tokens,* and *tokens* organization into a hybrid model. The goal for developing a representation transformation is to represent target data in a form which can be manipulated by tools of reasoning and learning in the training and recognition phases. Such transformation should also be simple enough and executable on parallel hardware (traditional or NN computers).

The representation transformation is the initial step in the design of an alternative approach to the ATR problem which will allow for recognizing targets of reduced signature (under overlap with other objects, partial visibility, background urban structures, heavy camouflage) through an AI-related reasoning process. This approach should allow for recognizing an unknown object of reduced signature from a larger number of candidate targets (over 20 different targets of varying pose signature).

This research is related to the most recent works of Waxman and the MIT Lincoln Lab team, ARPA research and systems engineering team and other researchers working with SAR datasets. Primary distinctions of this approach are:

- the detection of scatter blobs (rather than single pixels or regions) on different hierarchical levels,
- organization of adjacent blobs into local structural elements (tokens) which provide a base for further modeling of a target,
- fusion of structural, morphological, spectral and tactical information into a token,
- automatic selection of the most characteristic tokens for a given target, and
- tokens organization into a target model.

Transformation of target pictorial data into feature-token data is performed in the following six steps:

Step_1: Detection of scatter blobs,
Step_2: Construction of a graph representing target dominant blobs,
Step_3: Extraction of structural, morphological and spectral target data for detected blobs,
Step_4: Extraction of token frames to represent local target information,
Step_5: Fusion of feature data for each token frame, and
Step_6: Learning target model from a sets of feature-tokens taken for a range of target poses.

The detection of scatter blobs is executed by the DOG operator for gradually changing deviation (from a fine to relatively wide). In the second step, the DOG filtered signature is processed to extract blob pikes and three regions of interest (ROI). Blob pikes indicate centers of local blobs and ROIs indicate signature areas for maxima filtration. Blob pikes are extracted by a local 3x3 maximum operation. Base ROI is extracted by thresholding the DOG filtered signature at the level 0. Adjusted ROI is extracted by thresholding the DOG filtered signature at a level automatically adapted to provide dominant blob data for natural background of a target (or detailed blob data for urban structural background of a target). Rank ROI is extracted (for the training phase only) from raw signature, and it is used to eliminate blobs of the background. Detected blob pikes are then filtered by ROIs to form a corresponding base graph and an adjusted graph of a target. The base graph is more detailed, while the adjusted graph indicates the most significant blobs only. The adjusted graph is used as a starting graph in feature-token construction both in the training and recognition.

The graph extraction process is run for three different deviations of the DOG operator. In the third step, each candidate graph node is evaluated through these three levels to determine its structure. Additional structural, morphological and spectral features are computed and associated with each

blob. This process is run for a set of target signatures influenced by the change in target pose.

In the forth step, each target graph is transformed into a set of separate token frames. A token frame arranges two adjacent blobs into a structural abstract object. In the fifth step, a token frame is complemented by morphological and spectral features, and additional structural features. In the sixth step, redundant tokens are merged and the relationship between tokens is arranged for each target pose. Given a sequence of token sets for changing target pose, a representation and target model is learned (generalized).

Majority of operations implemented are parallel operations performed on target signatures or target graphs. Some sequential operations are involved in the training process. The project is continued on developing a learning system to acquire a set of feature graphs from feature tokens.

We have shown that the traditional approach to target pose estimation and feature extraction fails when applied to the target recognition task. A methodology for learning and recognizing targets from reduced SAR signatures has been developed to overcome problems with the traditional approach. Separate programs for segmentation and decomposition of SAR signatures have successfully been developed to extract shape and spectral data for the training phase. Developed approaches and programs have been tested for sequences of target data (target pose change). Training database of target graph data has been developed for the next phase of this research.

Deliverables:
1. A feature token paradigm (called OCTOPUS) for representing target data using structural, spectral and morphological features.
2. A set of programs for segmentation and decomposition of SAR signatures into token data.

Papers:

Pachowicz, P.W., "Representation Transformation for Target Recognition from Reduced SAR Signatures," submitted to International Conference on Image Processing, 1995.

13) RECOGNIZING NOISY PATTERNS IN SENSORY DATA

This project aims at the development of an approach to the recognition of noisy patterns in sensory data. The approach is based on the acquisition and analysis of dynamic characteristics of the matching process (a recognition curve - a sequence of confidence values) rather than on a static confidence measure received from a single match.

Most approaches to the problem of object recognition are based on the traditional architecture. This architecture emphasizes a separation of the training and recognition systems, so there is no cooperation between both systems during the recognition phase. In such architecture, the choice of the optimization degree needed to form concept desctiptions has to be determined by a teacher. This optimal degree can be found through the acquisition of recognition characteristics and the search for the most optimal optimization value. Such training, however, assumes that a teacher is well prepared and is able to interpret the data properly. Even if a teacher is able to find the optimization degree correctly, this degree can differ, for example, with changes in perceptual conditions. Then, the pike of recognition characteristics can be shifted out of the range of initially selected optimization degree.

To mitigate this problem, the recognition process has been redesigned and arranged into the iterative optimization loop of concept descriptions. This loop consists of three modules: concept optimization, inductive assertion, and a module of control and decision making. The loop is

controlled by an optimization parameter and it activates inductive assertion. Inductive assertion processes are performed each time for optimized concept descriptions. The system increases the optimization degree for each iteration loop. The decision making module completes believe values of partial recognition results computed for each optimization loop. In this way, a recognition curve is created versus the optimization degree.

The classification decision is made based on the evaluation of obtained recognition curves for each object class. The recognition algorithm that incorporates iterative optimization of concept description and flexible matching with test data performs as follows:

Step 1: Label the first section of each recognition curve as uptrend or downtrend recognition pattern,
Step 2: Select these recognition curves that have the uptrend recognition pattern only, and
Step 3: Make the final classification decision indicating this class for which the uptrend pattern runs through the highest recognition rates.

Such a classification decision is made based on the characteristics of the recognition curve through a fusion of recognition trend (over increasing optimization levels) with the recognition level. It means, the classification decision is made based on a sequence of matches rather than on a single match. We demonstrated that this approach is capable of recognizing very noisy concepts while traditional techniques based on a single confidence level fail to do so.

Deliverables:
A recognition method capable to classify very noisy patterns in sensory data.

Papers:

J. Bala and P.W. Pachowicz, "Recognizing Noisy Patterns via Iterative Optimization and Matching of Their Descriptions," *International Journal on Pattern Recognition and Machine Intelligence*, Vol.6, No.4, pp.513-538, 1992.

SUBCONTRACTOR
Center for Automation Research, University of Maryland, College Park

The Computer Vision Laboratory at the University of Maryland, College Park has been conducting research on various aspects of the relationship between machine learning and machine vision under Subagreement GMU-5-25010-1 to AFOSR grant F49620-92-J-0549. The research performed under this subagreement over the past two years has dealt with three applications of learning to the design of sensor-based agents:

(i) Development of specifications for agents that are capable of performing given tasks in a given environment. This is being done in the context of a formal framework for agent and task specification.

(ii) Development of exploratory and computational strategies that can be used by an active agent to "learn to navigate", i.e. to discover and organize information about the structure of its environment. This too is being done within a task-dependent formal framework. A paper [1] describing initial work in this area appeared in the Proceedings of the ARPA Image Understanding Workshop in November 1994. A technical report describing further work is close to completion [2].

(iii) Definition of methods of sensor-based manipulator control based on perceptual-kinematic maps, which relate properties of the sensory data (e.g., positions of features in an image) to properties of the kinematic chain that drives the manipulator (e.g., joint angles). In this framework, learning how to control the manipulator can be regarded as a problem of planning paths on a perceptual-kinematic surface. A paper [3] describing initial work in this area appeared in the Proceedings of the ARPA Image Understanding Workshop in November 1994; several technical reports on the perceptual-kinematic surface have been published [4] or are in preparation.

The Maryland principal investigators also contributed to the preparation of a report on research issues involved in the application of learning techniques to machine vision [5].

REFERENCES

1. S. Khuller, E. Rivlin, and A. Rosenfeld, "Learning to Navigate on a Graph," Proc. ARPA Image Understanding Workshop, Monterey, CA, November 1994, pp. 789-795.

2. P. Cucka, N. Netanyahu, and A. Rosenfeld, "Learning in Navigation: Goal Finding in Discrete Spaces," University of Maryland Technical Report 3433, April 1995.

3. J.Y. Herve, "Learning Hand-eye Coordination by an Active Observer," Proc. ARPA Image Understanding Workshop, Monterey, CA, November 1994, pp. 743-751.

4. J.Y. Herve, "Learning Hand-eye Coordination by an Active Observer Part I: Organizing Centers," University of Maryland Technical Report 3319, July 1994.

5. R.S. Michalski, A. Rosenfeld and Y. Aloimonos, "Machine Vision and Learning: Research Issues and Directions," University of Maryland Technical Report 3358, October 1994.

(c) TECHNICAL JOURNAL PUBLICATIONS

to be completed

Wnek, J. and Michalski, R.S., "Experimental Comparison of Symbolic and Subsymbolic Learning," *HEURISTICS, The Journal of Knowledge Engineering,* Special Issue on Knowledge Acquisition and Machine Learning, Vol. 5, No. 4, pp. 1-21, 1992.

Bala J. W. and Pachowicz P., "Recognizing Noisy Pattern Via Iterative Optimization and Matching of Their Rule Description," *International Journal on Pattern Recognition and Artificial Intelligence*, Vol. 6, No. 4, 1992.

Michalski, R.S., "Toward a Unified Theory of Learning: Multistrategy Task-adaptive Learning," in *Readings in Knowledge Acquisition and Learning: Automating the Construction and Improvement of Expert Systems*, B.G. Buchanan and D.C. Wilkins, Morgan Kaufmann, San Mateo, 1993.

Michalski, R.S., "A Theory and Methodology of Inductive Learning," in *Readings in Knowledge Acquisition and Learning: Automating the Construction and Improvement of Expert Systems*, B.G. Buchanan and D.C. Wilkins, Morgan Kaufmann, San Mateo, 1993.

Michalski, R.S., "Beyond Prototypes and Frames: The Two-tiered Concept Representation," in *Categories and Concepts: Theoretical Views and Inductive Data Analysis*, I. Van Mechelen, J. Hampton, R.S. Michalski and P. Theuns (Eds.), Academic Press, New York, 1993.

Michalski, R.S., "Learning = Inferencing + Memorizing: Introduction to Inferential Theory of Learning," in *Foundations of Knowledge Acquisition, Vol. 2: Machine Learning*, S. Chipman and A. Meyrowitz (Eds.), 1993.

Michalski, R.S., Bergadano, F., Matwin, S. and Zhang, J., "Learning Flexible Concepts Using a Two-tiered Representation," in *Foundations of Knowledge Acquisition, Vol. 2: Machine Learning*, S. Chipman and A. Meyrowitz (Eds.), 1993.

Michalski, R.S., "Inferential Theory of Learning as a Conceptual Basis for Multistrategy Learning," *Machine Learning*, Special Issue on Multistrategy Learning, Vol. 11, pp. 111-151, 1993.

Pachowicz, P. and Bala J., "A Noise-Tolerant Approach To Symbolic Learning From Sensory Data," *International Journal of Intelligent and Fuzzy Systems*, vol. 2, no. 4, pp. 347-362, 1994.

Michalski, R.S. and Tecuci, G. (Eds.), *Machine Learning: A Multistrategy Approach*, Vol. IV, Morgan Kaufmann, San Mateo, CA., 1994.

Bala, J.W., De Jong, K.A. and Pachowicz, P., "Multistrategy Learning from Engineering Data by Integrating Inductive Generalization and Genetic Algorithms," In Michalski, R.S. and Tecuci, G. (Eds.), *Machine Learning: A Multistrategy Approach*, Vol. IV, Morgan Kaufmann, San Mateo, CA, 1994.

Pachowicz, P. and Bala J., "A Noise-Tolerant Approach To Symbolic Learning From Sensory Data," *International Journal of Intelligent and Fuzzy Systems*, vol. 2, no. 4, pp. 347-362, 1994

Submitted or in preparation:

Bala, J. and Michalski R., "Rule-Structured Networks: A Multistrategy Approach for Learning from Continuous Data," *IEEE Transactions on Neural Networks.*

Bala, J., Michalski, R.S. and Wnek, J., "The PRAX Approach for a Similarity-based Recognition of a Large Number of Visual Concepts," *Pattern Recognition..*

(d) LIST OF PROFESSIONAL PERSONNEL ASSOCIATED WITH THE RESEARCH EFFORT

George Mason University

Principal/Co-Principal Investigators:
Ryszard S. Michalski
Peter W. Pachowicz

Postdoctoral fellows:
Jerzy Bala

Postdoctoral fellows:
Janusz Wnek

Supported Graduate Assistants:
Ali Hadjarian
Mark Maloof

Not-supported Graduate Assistants:
Eric Bloedorn

Awarded degrees:
Ph.D. in Information Technology. Date: May 1993. Recipient: Jerzy W. Bala. Title: Learning to Recognize Visual Concepts: Development And Implementation of a Method For Texture Concept Acquisition Through Inductive Learning.

Subcontract to University of Maryland

Principal/Co-Principal Investigators:
Azriel Rosenfeld
Yiannis Aloimonos
Larry Davis

Postdocs:
Jean-Yves Herve
Ehud Rivlin

Supported Graduate Assistants:
Tomas Brodsky
Brad Stuart

(e) INTERACTIONS (COUPLING ACTIVITIES)

Regular meetings and research discussions with members of Computer Vision Laboratory, University of Maryland, MD.

(i) Papers presented at meetings, conferences, seminars, etc.

Bala, J., K. DeJong, P. Pachowicz, "Integrated Inductive Learning And Genetic Algorithms For Texture Recognition," *Proceedings of the ML92 Workshop on Integrated Learning in Real-World Domains, Aberdeen,* Scotland, July 1992.

Bala, J., Michalski, R.S. and Wnek, J., "The Principal Axes Method for Constructive Induction," *Proceedings of the 9th International Conference on Machine Learning*, D. Sleeman and P. Edwards (Eds.), Aberdeen, Scotland, July 1992.

Pachowicz, P.W., Hieb, M.R. and Mohta, P., "A Learning-Based Incremental Model Evolution for Invariant Object Recognition," *Proceedings of the 6th International Conference on Systems Research, Informatics and Cybernetics*, Baden-Baden, Germany, August 1992.

Michalski, R.S., "Inferential Theory of Learning: Developing Foundations for Multistrategy Learning," *Reports of Machine Learning and Inference Laboratory*, Center for Artificial Intelligence, George Mason University, MLI 92-03, September 1992.

Wnek, J. and Michalski, R.S., "Comparing Symbolic and Subsymbolic Learning: Three Studies," *Reports of Machine Learning and Inference Laboratory*, Center for Artificial Intelligence, George Mason University, MLI 92-04, September 1992.

Pachowicz, P.W., "A Learning-Based Evolution of Concept Descriptions for an Adaptive Object Recognition," *Proceedings of the 4th International Conference on Tools with Artificial Intelligence*, pp. 316-323, Arlington, VA, November 1992.

Pachowicz, P.W., Bala, J. and Zhang, J., "Iterative Rule Simplification for Noise Tolerant Inductive Learning," *Proceedings of the 4th International Conference on Tools with Artificial Intelligence*, pp. 452-453, Arlington, VA, November 1992.

Michalski, R.S., Pachowicz, P.W., Rosenfeld, A. and Aloimonos, Y., "Machine Learning and Vision: Research Issues and Promising Directions," *NSF/DARPA Workshop on Machine Learning and Vision* (MLV-92), Harpers Ferry, WV, October 15-17, 1992. *Reports of Machine Learning and Inference Laboratory,* MLI 93-1, Center for Artificial Intelligence, George Mason University, February 1993.

Michalski, R.S., Bala, J.W. and Pachowicz, P.W., "GMU Research on Learning in Vision: Initial Results," *Proceedings of the DARPA Image Understanding Workshop,* Washington D.C., April 18-21, 1993.

Bloedorn, E., Wnek, J. and Michalski, R.S., "Multistrategy Constructive Induction," *Proceedings of the Second International Workshop on Multistrategy Learning (MSL93)*, Harpers Ferry, WV, Morgan Kaufmann, May 27-29, 1993.

Hieb, M. and Michalski, R.S., "Knowledge Representation for Multistrategy Task-adaptive Learning: Dynamic Interlaced Hierarchies," *Proceedings of the Second International Workshop on Multistrategy Learning (MSL93)*, Harpers Ferry, WV, Morgan Kaufmann, May 27-29, 1993.

Michalski, R.S. and Tecuci, G. (Eds.), *Proceedings of the Second International Workshop on Multistrategy Learning (MSL93)*, Harpers Ferry, WV, Morgan Kaufmann, May 27-29, 1993.

Michalski, R.S. and Tecuci, G., "Multistrategy Learning," *Tutorial at the National Conference on Artificial Intelligence, AAAI-93*, Washington D.C., July 11-12, 1993.

Bala, J., Michalski, R.S. and Wnek, J., "The PRAX Approach to Learning a Large Number of Texture Concepts," *Technical Report FS-93-04 Machine Learning in Computer Vision: What, Why and How? AAAI Fall Symposium on Machine Learning in Computer Vision*, AAAI Press, Menlo Park, CA, October 1993.

Bala, J. and Pachowicz, P.W., "Issues on Noise Tolerant Learning from Sensory Data," *Technical Report FS-93-04 Machine Learning in Computer Vision: What, Why and How? AAAI Fall Symposium on Machine Learning in Computer Vision*, AAAI Press, Menlo Park, CA, October 1993.

Pachowicz, P.W., "Integration of Machine Learning and Vision into an Active Agent Paradigm on the Example of Face Recognition Problem," *Technical Report FS-93-04, Machine Learning in Computer Vision: What, Why and How? AAAI Fall Symposium on Machine Learning in Computer Vision*, AAAI Press, Menlo Park, CA, October 1993.

Bala J., Michalski, R., Pachowicz P. "GMU Research on Learning in Vision," *Proceedings of the 1994 ARPA Image Understanding Workshop,* Monterey, November, 1994.

(ii) Consultative and advisory functions to other laboratories and agencies

TASC, Spatial Data Engineering Center, Linda Tischer, November 19, 1993.

NIH, Image Processing Group, Dr. Ronald Levin, Bonnie Douglas, March 30, 1994

(f) NEW DISCOVERIES, INVENTIONS, OR PATENT DISCLOSURES AND SPECIFIC APPLICATIONS STEMMING FROM THE RESEARCH EFFORT

1) Ideas and methodology for applying symbolic learning for determining visual object descriptions (MTL).

2) Method for learning new concepts descriptions in terms of the previously learned concept descriptions (PRAX method).

3) Method for determining "key" object features through inductive inference ("dynamic recognition")

4) Method for multistrategy learning that combines learning decision rules with neural net learning.

(g) ADDITIONAL INSIGHT AND INFORMATION

NSF/DARPA Workshop on Machine Learning and Vision, organized by George Mason University in collaboration with the University of Maryland, October 15-17, 1991 in Harpers Ferry, WV.

The Workshop brought together leading researchers in vision and learning to discuss the possibilities of cross-fertilizing the two fields, and implementing learning capabilities in vision systems. It was attended by about 40 participants representing universities, industrial and governmental laboratories, and several sponsoring agencies.

Partially based on this Workshop and partially based on further investigations and collaborations of the authors, a report " Machine Vision and Learning" has been prepared by R.S. Michalski, A. Rosenfeld and Y. Alomonois, 1994.

Publications

Bala, J. and Michalski R., "Recognition of Textural Concepts Through Multilevel Symbolic Transformations," *Proceedings of the 3rd International Conference on Tools for Artificial Intelligence*, San Jose C.A., Nov. 1991.

Bala, J., "Learning To Recognize Visual Concepts," Ph.D. Dissertation, George Mason University, University Microfilms Int., Ann Arbor, MI, 1993.

Michalski, R., Bala J., Pachowicz P. "GMU RESEARCH ON LEARNING IN VISION: Initial Results," *Proceedings of the 1993 DARPA Image Understanding Workshop,* Washington DC, 1993.

Bala J., Michalski, R., Pachowicz P. "GMU Research on Learning in Vision," *Proceedings of the 1994 ARPA Image Understanding Workshop,* Monterey, November, 1994.

Bala J., Michalski, R., Pachowicz P. "GMU Research on Learning in Vision," *Proceedings of the 1994 ARPA Image Understanding Workshop,* Monterey, November, 1994.

APPENDIX A

References

Bala, J., "Learning To Recognize Visual Concepts: Development And Implementation of a Method For Texture Concept Acquisition Through Inductive Learning," Ph.D. dissertation, Graduate School of George Mason University, *Reports of the Machine Learning and Inference Laboratory*, Center for Artificial Intelligence, MLI 93-3, (also published by University Microfilms Int., Ann Arbor, MI).

Bala, J. and Michalski R., "Recognition of Textural Concepts Through Multilevel Symbolic Transformations," *Proceedings of the 3rd International Conference on Tools for Artificial Intelligence*, San Jose C.A., Nov. 1991.

Bala J. W. and Pachowicz P. W., "Recognizing Noisy Pattern Via Iterative Optimization and Matching of Their Rule Description," *International Journal on Pattern Recognition and Artificial Intelligence*, Vol. 6, No. 4, 1992.

Bala, J. and Pachowicz, P.W., "Issues on Noise Tolerant Learning from Sensory Data," *Technical Report FS-93-04 Machine Learning in Computer Vision: What, Why and How? AAAI Fall Symposium on Machine Learning in Computer Vision*, AAAI Press, Menlo Park, CA, October 1993.

Bala, J., Michalski, R.S. and Wnek, J., "The PRAX Approach to Learning a Large Number of Texture Concepts," *Technical Report FS-93-04 Machine Learning in Computer Vision: What,*

Why and How? AAAI Fall Symposium on Machine Learning in Computer Vision, AAAI Press, Menlo Park, CA, October 1993.

Bala, J., Michalski, R.S. and Wnek, J., "The Principal Axes Method for Constructive Induction," *Proceedings of the 9th International Conference on Machine Learning*, D. Sleeman and P. Edwards (Eds.), Aberdeen, Scotland, July 1992.

Bala, J.W., De Jong, K.A. and Pachowicz, P. W., "Multistrategy Learning from Engineering Data by Integrating Inductive Generalization and Genetic Algorithms," In Michalski, R.S. and Tecuci, G. (Eds.), *Machine Learning: A Multistrategy Approach*, Vol. IV, Morgan Kaufmann, San Mateo, CA, 1994.

Bloedorn, E., Wnek, J. and Michalski, R.S., "Multistrategy Constructive Induction," *Proceedings of the Second International Workshop on Multistrategy Learning (MSL93)*, Harpers Ferry, WV, Morgan Kaufmann, May 27-29, 1993.

Channic T., TEXPERT, An Application of Machine Learning to Texture Recognition, *MLI Reports* 89-27, George Mason University, Fairfax, VA, 1989.

Hieb, M. and Michalski, R.S., "Knowledge Representation for Multistrategy Task-adaptive Learning: Dynamic Interlaced Hierarchies," *Proceedings of the Second International Workshop on Multistrategy Learning (MSL93)*, Harpers Ferry, WV, Morgan Kaufmann, May 27-29, 1993.

Michalski, R.S. and Tecuci, G. (Eds.), *Machine Learning: A Multistrategy Approach*, Vol. IV, Morgan Kaufmann, San Mateo, CA., 1994.

Michalski, R.S. and Tecuci, G. (Eds.), *Proceedings of the Second International Workshop on Multistrategy Learning (MSL93)*, Harpers Ferry, WV, Morgan Kaufmann, May 27-29, 1993.

Michalski, R.S. and Tecuci, G., "Multistrategy Learning," *Tutorial at the National Conference on Artificial Intelligence, AAAI-93*, Washington D.C., July 11-12, 1993.

Michalski, R.S. and Wnek, J., "Constructive Induction: An Automated Design of Knowledge Representation Spaces for Machine Learning," *Reports of the Machine Learning and Inference Laboratory,* MLI 93-11, Center for Artificial Intelligence, George Mason University, Fairfax, VA, November 1993.

Michalski, R.S., "A Theory and Methodology of Inductive Learning," in *Readings in Knowledge Acquisition and Learning: Automating the Construction and Improvement of Expert Systems*, B.G. Buchanan and D.C. Wilkins, Morgan Kaufmann, San Mateo, 1993.

Michalski, R.S., Bala, J.W. and Pachowicz, P.W., "GMU Research on Learning in Vision: Initial Results," *Proceedings of the DARPA Image Understanding Workshop,* Washington D.C., April 18-21, 1993.

Michalski, R.S., Bergadano, F., Matwin, S. and Zhang, J., "Learning Flexible Concepts Using a Two-tiered Representation," in *Foundations of Knowledge Acquisition, Vol. 2: Machine Learning*, S. Chipman and A. Meyrowitz (Eds.), 1993.

Michalski, R.S., "Beyond Prototypes and Frames: The Two-tiered Concept Representation," in *Categories and Concepts: Theoretical Views and Inductive Data Analysis*, I. Van Mechelen, J. Hampton, R.S. Michalski and P. Theuns (Eds.), Academic Press, New York, 1993.