UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

DTIC FILE COPY

(2)

REPORT DOCUMENTATION PAGE

AD-A200 920

DTIC ELECTE ... 9 1988 D

1a. REPORT SECURITY CLASSIFICATION	1b. RESTRICTIVE MARKINGS
	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S) ARO 21474.10-MA
6a. NAME OF PERFORMING ORGANIZATION University of California, Irvine 6b. OFFICE SYMBOL (If applicable) Elec. Engrg.	7a. NAME OF MONITORING ORGANIZATION U. S. Army Research Office
6c. ADDRESS (City, State, and ZIP Code) Irvine, California 92717	7b. ADDRESS (City, State, and ZIP Code) P. O. Box 12211 Research Triangle Park, NC 27709-2211
8a. NAME OF FUNDING/SPONSORING ORGANIZATION U. S. Army Research Office 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER DAA29-84-K-0208
8c. ADDRESS (City, State, and ZIP Code) P. O. Box 12211 Research Triangle Park, NC 27709-2211	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)

THE DESIGN OF AUTOMATIC PATTERN CLASSIFIERS

12. PERSONAL AUTHOR(S) Dr. Jack Sklansky

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
FINAL	FROM 10-1-84 TO 6-30-88	10-6-88	

16. SUPPLEMENTARY NOTATION The view, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

ABSTRACT

Our research on "The Design of Automatic Pattern Classifiers" led to solutions of several fundamental problems associated with the use of automatic classifiers of potential military targets appearing in aerial and terrestial images. These problems fall in three areas: data analysis, classifier design, and classifier architectures.

-- continued on other side --

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Dr. Jack Sklansky	22b. TELEPHONE (Include Area Code) 714-856-6726 / 22c. OFFICE SYMBOL

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFIED

88 11 07 052

ARO 21474.10-MA

THE DESIGN OF AUTOMATIC PATTERN CLASSIFIERS-- FINAL REPORT

October 1, 1984 - June 30, 1988

1. STATEMENT OF PROBLEM

We will exploit relational graphs and "local" training of piecewise linear decision surfaces to achieve improved procedures and insights in the design of automatic pattern classifiers. We will use the orientations of the linear segments of these classifiers as guides in a branch-and-bound search for optimal subsets of features. We will develop procedures for determining near-optimal sequences of features for piecewise linear classifiers. We will develop new procedures for selecting small subsets of features from large initial sets for use in automatic pattern classifiers. We will extend our interactive relational graph design technique for the design of piecewise linear two-class classifiers to the design of multiple-class classifiers.

2. SUMMARY OF RESULTS

Our research on "The Design of Automatic Pattern Classifiers" led to solutions of several fundamental problems associated with the use of automatic classifiers of detected candidate military targets appearing in aerial and terrestial images. These problems fall in three areas: data analysis, classifier design, and classifier architectures.

2.1. Data analysis

We developed a highly effective means of discoverning clusters in multiple-dimensional data, using our CAMPA (Clustering and Mapping for Pattern Analysis) interactive software in conjunction with modern computer graphics.

We developed a new mapping technique -- the least squares mapping. In conjunction with this mapping we also invented a new clustering technique --

interative least squares (ILS) clustering. Experiments on a rich set of synthetic data demonstrated the superiority of ILS over other leading mapping and clustering techniques.

We carried out two computer-based experiments that compared the use of CAMPA against automatic techniques for the design of classifiers. One experiment evaluated cluster-finding processes; the other evaluated classifier design processes. We believe that the results of these experiments provide a reliable critique of mapping methods and a guide to their use.

We devised and tested two powerful techniques for selecting small subsets of features for automatic classifiers from an initially large set: one based on branch-and-bound and the other on genetic algorithms.

The first technique, based on a modified branch-and-bound, exploits our discovery of near monotonicity of the misclassification error with respect to subset inclusion among the features of piecewise linear classifiers. In tests carried out independently of us by the U.S. Army Missile Command on infrared imagery containing military targets, our modified branch-and -bound feature selector found a nine-feature subset from an initial set of 30 features extracted by the Perkins Global Segmenter. Our feature selector found this nine-feature subset after examining only 3000 subsets out of a possible maximum of 262143 subsets -- a result which demonstrated the superiority of our feature selector to all earlier feature selection methods tested by the Missile Command.

Recently we developed a new feature selection algorithm which seems especially effective for initial feature sets consisting of 20 or more features. This method is based on genetic algorithms -- i.e., algorithms that simulate natural evolution in the creation and selection of chromosomes. In our algoriothm each chromosome is a binary string that represents a feature subset. In this mechanism the chromosomes are allowed to <u>mate</u> or <u>crossover</u>; and to <u>mutate</u>. The mating of two chromosomes produces a pair of offspring chromosomes which are syntheses of their parents. A mutation of a chromosome produces a near-copy of the original chromosome with a few components altered.

Our preliminary experiments on the use of genetic algorithms indicate about two or more orders of magnitude reduction of search effort compared to our modified branch-and-bound when the number of features exceeds 20.

2.2. Classifier design

We devised three new techniques for the design of automatic classifiers: a) a method for designing trainable piecewise linear multiple-class classifiers that yields error rates that are close to Bayes-optimal, and mechanisms that are much simpler and faster than the k-nearest-neighbor classifier, which is the best known earlier classifier that provides near-Bayes-optimal error rates; b) a method for designing multiple-class tree classifiers with automatic minimization of the number of features -- thereby achieving high speeds of classification; and c) a promising new approach to the design of real-time classifiers, taking into account the accumulation of imperfect evidence in real time based on context and motion as well as shape.

Our method of designing multiple class tree classifiers builds on our new technique for designing multiple-class piecewise linear classifiers. This technique involves the cutting of arcs joining selected pairs of points in opposite classes in d-dimensional space. These pairs are selected on the basis of other data points in the neighborhood of the pairs. We refer to the arcs joining these pairs as links. We show how to nearly minimize the number of hyperplanes required to cut all or nearly all of these links, thereby yielding a near-Bayes-optimal decision surface regardless of the number of classes, and we describe the underlying theory. Experiments on multiple-class data obtained from ship images show that classifiers designed by this method yield approximately the same error rates as the best k-nearest neighbor rule, while enjoying computational efficiency of classification.

The error rate of our tree classifier is shown to be robust with respect to the design parameters. This makes the design, from a practical standpoint, free from parameters that have to be specified by users. The trade-off between speed and accuracy in pattern classification, where the speed is related inversely to the number of features, can be controlled by restricting the maximum allowable number of features at each node in the tree.

Our method of designing real-time classifiers is based on a new approach to the accumulation of "mass", representing degree of belief, in a hierarchical graph in which each link of the graph is associated with a level of abstraction in the description of an image. This technique takes into account the imperfect evidence of context and motion of candidate targets, as well as shape, in real time.

2.3. Classifier architectures

We devised a highly effective technique of dynamic task assignment to a network of processors analyzing long sequences of images in real time. This technique suppresses bottlenecks in the pipelining of tasks in image analysis and target classification.

Our implementation of dynamic task assignment is a constrained optimization of an objective function which estimates the system throughput. This objective function is dependent on the computation time for each task, the communication intensities among the tasks, and a mapping of the problem graph to a system graph. To find an efficient mapping, we devised a heuristic optimization procedure in which we first identified possible bottlenecks, and we gave them a high priority in the mapping procedure. Our method was tested using typical problem graphs and system graphs. Simulation results have shown that the proposed mapping method performs well enough to find a good mapping in a reasonably short time.

3. PUBLICATIONS AND TECHNICAL REPORTS

I. Foroutan, J. Sklansky, "Automatic Feature Selection for Non-Gaussian Data" Eighth International Joint Conference on Pattern Recognition, Paris, France, 1986, pp. 327-329.

I. Foroutan, J. Sklansky, "Feature Selection for Automatic Classification of Non-Gaussian Data," IEEE Trans. on Systems, Man, and Cybernetics, Vol. SMC-17, No. 2, March/April 1987, pp. 187-198.

W. Siedlecki, J. Sklansky, "Toward Optimal Feature Selection: Past, Present, and Future," Transactions of the Fourth Army Conference of Applied Mathematics and Computing, U.S. Department of Defense, pp. 721-729, 1986, ARO

W. Siedlecki, J. Sklansky, "On Automatic Feature Selection," International Journal of Pattern Recognition and Artificial Intelligence, Vol. 2, No. 2, June 1988, pp. 197-220.

W. Siedlecki, K. Siedlecka, J. Sklansky "An Overview of Mapping Techniques for Exploratory Pattern Analysis," Pattern Recognition, 1988, Vol 21, No. 5.

W. Siedlecki, K. Siedlecka, J. Sklansky, "Experiments on Mapping Techniques for Exploratory Pattern Analysis," Pattern Recognition, 1988, Vol. 21, No. 5.

W. Siedlecki, and J. Sklansky, "A Note on Genetic Algorithms for Large-Scale Feature Selection," submitted for publication.

Y. Park, J. Sklansky, "Automated Design of Piecewise Linear Classifiers of Multiple-Class Data," Ninth International Conference on Pattern Recognition Rome, Italy, November 1988.

Y. Park, J. Sklansky, "The Use of Tomek Links in the Design of Piecewise Linear Classifiers," Sixth Army Conference on Applied Mathematics and Computing, Boulder, Colorado, May-June 1988.

W. Siedlecki, K. Siedlecka, J. Sklansky, "Mapping Techniques for Exploratory Pattern Analysis," Proceedings of Conference on Pattern Recognition in Practice III, Amsterdam, May 18-20, 1988.

Y. Moon, N. Bagherzadeh, J. Sklansky, "Assignment of Tasks to Multiple Processors for Image Processing," Pattern Recognition for Advanced Missile Systems Conference, Huntsville, Alabama, November 1988.

Y. Moon, N. Bagherzadeh, J. Sklansky, "Macropipelined Multicomputer Systems for Image Analysis," SPIE Conference on Optics, Electro-Optics, and Laser Applications in Science and Engineering, Los Angeles, California, Jan. 1989.

Y. Park, J. Sklansky, "Automated Design of Multiple-Class Piecewise Linear Classifiers," submitted for publication.

4. SCIENTIFIC PERSONNEL

Patrick Durieux, M.S.

I. Foroutan, Ph.D.

W. Siedlecki, Programmer Analyst

K. Siedlecka, Programmer Analyst

Y. T. Park, Research Assistant

Y. S. Moon, Research Assistant

D. Gutfinger, Research Assistant

M. Silbermann, Postdoctoral Research Associate

Adam Krzyzak, Postdoctoral Research Associate

H. R. Patel, Research Assistant

J. Sklansky, Principal Investigator

5. APPENDIX: ABSTRACTS OF PAPERS AND REPORTS

1. I. Foroutan, J. Sklansky, "Feature Selection for Automatic Classification of Non-Gaussian Data"

2. W. Siedlecki, J. Sklansky, "Toward Optimal Feature Selection: Past, Present, and Future"

3. W. Siedlecki, J. Sklansky, "On Automatic Feature Selection"

4. W. Siedlecki, K. Siedlecka, J. Sklansky, "An Overview of Mapping Techniques for Exploratory Pattern Analysis"

5. W. Siedlecki, K. Siedlecka, J. Sklansky, "Experiments on Mapping Techniques for Exploratory Pattern Analysis"

6. W. Siedlecki, J. Sklansky, "A Note on Genetic Algorithms for Large-Scale Feature Selection"

7. Y. Park, J. Sklansky, "Automated Design of Piecewise Linear Classifiers of Multiple-Class Data"

8. Y. Park, J. Sklansky, "The Use of Tomek Links in the Design of Piecewise Linear Classifiers"

9. W. Siedlecki, K. Siedlecka, J. Sklansky, "Mapping Techniques for Exploratory Pattern Analysis"

10. Y. Moon, N. Bagherzadeh, J. Sklansky, "Assignment of Tasks to Multiple Processors for Image Processing"

11. Y. Moon, N. Bagherzadeh, J. Sklansky, "Macropipelined Multicomputer Systems for Image Analysis"

12. Y. Park, J. Sklansky, "Automated Design of Multiple-Class Piecewise Linear Classifiers"

FEATURE SELECTION FOR AUTOMATIC CLASSIFICATION OF NON-GAUSSIAN DATA

I. Foroutan and Jack Sklansky

Abstract

We present a computer-based technique for automatic selection of features for the classification of non-Gaussian data. The selection technique exploits interative cluster-finding and a modified branch and bound optimization of piecewise linear classifiers.

The technique first finds an efficient set of pairs of oppositely classified clusters to represent the data. Then a zero-one implicit enumeration implements a branch and bound search for a good subset of features.

A test of our feature selection technique on multidimensional synthetic and real data yielded close-to-optimum and, in many cases, optimum subsets of features. The real data consisted of a) 1284 12-dimensional feature vectors, representing normal and abnormal breast tissue, extracted from x-ray mammograms; and b) 1060 30-dimensional feature vectors representing tanks and clutter in infrared images.

TOWARD OPTIMAL FEATURE SELECTION: Past, Present, and Future

W. Siedlecki, Jack Sklansky

Abstract

We present an overview of relatively reliable and efficient methodologies of selecting features for automatic pattern classifiers. We also describe a family of suboptimal, robust methods developed under our ARO-supported research contract. These methods include:

1. methods utilizing the idea of approximate monotonicity,
2. methods estimating a trend of monotonicity, and
3. AI methods for graph searching.

Another promising method, currently under consideration, is based on the observation that the monotonicity property of classifier's error rate is highly related to the optimality of this classifier. This method does not require any search but evaluates all features at the same time in a fuzzy decision process that involves the assignment of a weight to each feature.

In the paper we discuss these four methods in detail.

ON AUTOMATIC FEATURE SELECTION

by

Wojciech Siedlecki and Jack Sklansky

Pattern Recognition Project
Department of Electrical Engineering
University of California, Irvine
Irvine, California 92717

ABSTRACT

We review recent research on methods for selecting features for multidimensional pattern classification. These methods include nonmonotoncity-tolerant branch-and-bound search, best-first search, and beam search. We described the potential benefits of Monte Carlo approaches such as simulated annealing and genetic algorithms. We compare these methods to facilitate the planning of future research on feature selection.

AN OVERVIEW OF MAPPING TECHNIQUES FOR EXPLORATORY PATTERN ANALYSIS

by

Wojciech Siedlecki, Kinga Siedlecka and Jack Sklansky

ABSTRACT

We describe a carefully designed collection of mapping methods for computer aided pattern analysis. This collection is so versatile that it is likely to be useful to every research facility engaged in multidimensional data analysis and pattern classification. The collection includes eight major groups of mappings. Among them is our own innovation -- the least squares mapping. We also introduce a method that accelerates and increases the precision of mappings based on the Fisher discriminant.

In a subsequent paper we describe two experiments that compare these mapping techniques against powerful automatic procedures: one for cluster-finding and the other for classifier design.

Experiments on Mapping Techniques for Exploratory Pattern Analysis

by

Wojciech Siedlecki, Kinga Siedlecka and Jack Sklansky

Department of Electrical Engineering
University of California, Irvine
Irvine, CA 92715

SUMMARY

We describe two computer-based experiments evaluating the effectiveness of several mapping techniques for exploratory pattern analysis. The first experiment compared various mappings and classical clustering techniques as aids to people whose objective is to find clusters in the data. The second experiment evaluated the effectiveness of two-dimensional displays produced by analytic mappings for people designing linear and piecewise linear classifiers. The performance of the classifiers designed by the people aided by these displays was compared with automatically trained classifiers.

Based on these experiments we selected three best mapping methods. Even the untrained users who took part in our experiments achieved very good results with the aid of these best mappings. In fact, these results were superior by a significant margin to to those obtained from renowned classical pattern recognition procedures.

Another valuable result of our experiments is that they allowed us to identify the sets of parameters most often used by the participants and, consequently, suggest guidelines for the best use of mapping techniques.

A Note on Genetic Algorithms for Large-Scale Feature Selection

by W. Siedlecki and J. Sklansky
University of California, Irvine
Irvine, California, 92717
U.S.A.

Abstract

We introduce the use of genetic algorithms (GA) for the selection of features in the design of automatic pattern classifiers. Our preliminary results suggest that GA is a powerful means of reducing the time for finding near-optimal subsets of features from large sets.

Automated Design of Piecewise Linear Classifiers of Multiple-Class Data

Youngtae Park and Jack Sklansky
Department of Electrical Engineering
Focused Research Program on Image Engineering
University of California, Irvine
Irvine, CA 92717
U.S.A.

Abstract - A new method for designing multiple-class piecewise linear classifiers is described. It involves the cutting of arcs joining pairs of opposed points in d-dimensional space. We refer to such arcs as *links*. We show how to nearly minimize the number of hyperplanes required to cut all of these links, thereby yielding a near-Bayes-optimal decision surface regardless of the number of classes, and we describe the underlying theory. This method does not require parameters to be specified by users. Experiments on multiple-class data obtained from ship images show that classifiers designed by this method yield approximately the same error rate as the best *k-nearest neighbor* rule, while enjoying computational efficiency of classification.

The Use of Tomek Links in the Design of Piecewise Linear Classifiers

Youngtae Park and Jack Sklansky
University of California, Irvine
Irvine, CA 92717, U.S.A.

Sponsoring agency: U.S. Army Research Office

Abstract

A new method for designing multiple-class piecewise linear classifiers is described. It involves the cutting of arcs joining pairs of opposed points in d-dimensional space, selected by a technique described by Tomek. If no data points exist inside the circle of influence between an opposed pair of data points, we refer to that pair as a *Tomek link*.

Given a training set $\mathcal{D}$ and a decision rule $\mathcal{R}$, we say that $\mathcal{D}$ is *training-set consistent* with respect to $\mathcal{R}$ if all data points in $\mathcal{D}$ are classified correctly by $\mathcal{R}$. Cutting all Tomek links by a hyperplane guarantees training-set consistency if the following two conditions are satisfied: a) the set of links which contribute to the 1-nearest neighbor decision boundary (we call these links *active links*) is the same as the set of Tomek links, and b) there exists a hyperplane that cuts all Tomek links.

It is easy to show that the set of Tomek links is a subset of the set of active links. In practical problems these two sets are almost identical, but the set of Tomek links is much easier to find than the set of active links. If at least two hyperplanes are required to cut all the Tomek links, almost all of the misclassified points will lie near the intersection of pairs of hyperplanes. Thus the Tomek links provide a powerful guide for the design of piecewise linear classifiers. In particular, we have found that an efficient way to design a piecewise linear classifier is to find a minimum set of hyperplanes that cuts all of the Tomek links.

We show how to nearly minimize the number of hyperplanes required to cut all of the Tomek links. This technique yields a near-Bayes-optimal decision surface regardless of the number of classes. It does not require parameters to be specified by the users. A nearly optimal set of hyperplanes is determined automatically by examining the error rates associated with various numbers of hyperplanes. We describe the mathematical theory underlying this design technique.

Experiments on real and synthetic data show that classifiers designed by this method yield approximately Bayes error rates while enjoying computational efficiency of classification.

Mapping Techniques for Exploratory Pattern Analysis

by

Wojciech Siedlecki, Kinga Siedlecka and Jack Sklansky

Department of Electrical Engineering
University of California, Irvine
Irvine, CA 92715

ABSTRACT

We describe a carefully designed collection of mapping methods for computer aided pattern analysis. This collection is so verstile that it is likely to be useful to every research facility engaged in multidimensional data analysis and pattern classification. The collection includes eight major groups of mappings. Among them is our own innovation – the *least squares mapping*. We also introduce a method that accelerates and increases the precision of mappings based on the Fisher discriminant.

We describe two computer-based experiments that compare this collection of mappings for computer-aided pattern analysis against powerful automatic techniques. One experiment evaluated cluster finding processes, and the other evaluated classifier design processes.

In both of these experiments naïve (i.e., untrained) humans aided by the generalized declustering mapping and our least squares mapping significantly outperformed the remaining computer-aided (i.e. semiautomatic) and automatic techniques – except that the classifier designed by humans aided by mappings performed about as well as the k-nearest-neighbor classifiers (the latter, it is known, are computationally inefficient).

Assignment of Tasks to Multiple Processors for Image Processing

by
Youngshik Moon, Nader Bagherzadeh and Jack Sklansky
Department of Electrical Engineering
Focused Research Program on Image Engineering
University of California, Irvine
Irvine, CA 92717
U.S.A.
September 7, 1988
Tel: 714-856-6726

Abstract

An efficient mapping scheme for maximizing the performance of real-time image processing system is presented. We first introduce a new objective function which can accurately estimate the throughput of a system. This objective function is shown to be dependent on the computation time for each subtask, the communication intensities among subtasks, and a particular mapping of problem graph to a system graph. Given a mapping, a method of evaluating the objective function is presented, where the link contention problem is carefully taken into account. To find an efficient mapping, a heuristic optimization procedure is proposed, where we first identify possible bottlenecks (we refer to these as *critical edges*) and then we give them higher priority in the mapping procedure. Our method is tested using some typical problem graphs and system graphs. Simulation results have shown that the proposed mapping method performs well enough to find a good mapping in reasonably short time.

Macropipelined multicomputer systems for image analysis

Youngshik Moon, Nader Bagherzadeh and Jack Sklansky
University of California, Irvine, Department of Electrical Engineering
Irvine, California 92717

ABSTRACT

We present a scheme for macropipelining in multicomputer systems to achieve high speeds in processing multiple images. Most image processing applications consist of a sequence of tasks - - e.g., preprocessing, detection, segmentation, feature extraction, and classification. This sequence lends itself to a pipelining strategy. To minimize the effects of bottlenecks in this pipeline, we introduce a performance model for data partitioning which includes both the computation and the communication aspects of parallel processing. With the help of this model, we assign the appropriate number of processors to each task so that the workloads are well-balanced. Then we generate a problem graph describing the relationships among tasks and subtasks. We use an estimator of the frame processing time of the image processing system as an objective function for choosing a mapping of the problem graph to a system graph. This estimator takes account of computation times and communication intensities among the subtasks in the problem graph, and accounts for link contentions. To find an efficient mapping, we use a heuristic optimization technique in which possible bottlenecks are given high priority in the mapping procedure. We tested our macropipelining scheme on a typical image processing application in a simulated hypercube computer system. The results support our belief that this scheme yields effective architectures for high-speed processing of long sequences of images.

Automated Design of Multiple-Class Piecewise Linear Classifiers

Youngtae Park and Jack Sklansky

Department of Electrical Engineering
Focused Research Program on Image Engineering
University of California, Irvine
Irvine, CA 92717
U.S.A.

Abstract - A new method for designing multiple-class piecewise linear classifiers is described. It involves the cutting of straight line segments joining pairs of opposed points (i.e., points from distinct classes) in d-dimensional space. We refer to such straight line segments as *links*. We show how nearly to minimize the number of hyperplanes required to cut all of these links, thereby yielding a near-Bayes-optimal decision surface regardless of the number of classes, and we describe the underlying theory. This method does not require parameters to be specified by users – an improvement over earlier methods. Experiments on multiple-class data obtained from ship images show that classifiers designed by this method yield approximately the same error rate as the best k-nearest neighbor rule, while enjoying computational efficiency of classification.