# Applications of Artificial Neural Networks in Image Processing IV
# Volume 3647

PROCEEDINGS OF SPIE

SPIE—The International Society for Optical Engineering

# *Applications of Artificial Neural Networks in Image Processing IV*

**Nasser M. Nasrabadi**
**Aggelos K. Katsaggelos**
*Chairs/Editors*

**28–29 January 1999**
**San Jose, California**

*Sponsored by*
IS&T—The Society for Imaging Science and Technology
SPIE—The International Society for Optical Engineering

*Published by*
SPIE—The International Society for Optical Engineering



**Volume 3647**

SPIE is an international technical society dedicated to advancing engineering and scientific applications of optical, photonic, imaging, electronic, and optoelectronic technologies.

The papers appearing in this book comprise the proceedings of the meeting mentioned on the cover and title page. They reflect the authors' opinions and are published as presented and without change, in the interests of timely dissemination. Their inclusion in this publication does not necessarily constitute endorsement by the editors or by SPIE.

Please use the following format to cite material from this book:
Author(s), "Title of paper," in *Applications of Artificial Neural Networks in Image Processing IV*, Nasser M. Nasrabadi, Aggelos K. Katsaggelos, Editors, Proceedings of SPIE Vol. 3647, page numbers (1999).

ISSN 0277-786X
ISBN 0-8194-3118-4

Published by
**SPIE—The International Society for Optical Engineering**
P.O. Box 10, Bellingham, Washington 98227-0010 USA
Telephone 360/676-3290 (Pacific Time) • Fax 360/647-1445



Printed in the United States of America.

# Conference Committee

*Conference Chairs*

**Nasser M. Nasrabadi,** U.S. Army Research Laboratory
**Aggelos K. Katsaggelos,** Northwestern University

*Program Committee*

**Pierre Baldi,** Jet Propulsion Laboratory
**Yoshua Bengio,** Université de Montréal (Canada)
**Bir Bhanu,** University of California/Riverside
**Terry M. Caelli,** Curtin University of Technology (Australia)
**Rama Chellappa,** University of Maryland/College Park
**Chang Y. Choo,** California State University/San Jose
**John Daugman,** University of Cambridge (UK)
**Sandor Z. Der,** U.S. Army Research Laboratory
**Edward R. Dougherty,** Texas A&M University
**Kunihiko Fukushima,** Osaka University (Japan)
**Erol Gelenbe,** Duke University
**David H. Haussler,** University of California/Santa Cruz
**Nicolaos B. Karayiannis,** University of Houston
**Christof Koch,** California Institute of Technology
**Bart Kosko,** University of Southern California
**Sun-Yuan Kung,** Princeton University
**Yann LeCun,** AT&T Bell Laboratories
**Richard P. Lippmann,** MIT Lincoln Laboratory
**Vincent Mirelli,** U.S. Army Research Laboratory
**Erkki Oja,** Helsinki University of Technology (Finland)
**Sankar K. Pal,** Indian Statistical Institute
**Tomaso A. Poggio,** MIT AI Laboratory
**Syed A. Rizvi,** CUNY/College of Staten Island
**David E. Rumelhart,** Stanford University
**Michael Seibert,** MIT Lincoln Laboratory
**Christoph Von Der Malsburg,** University of Southern California
**Jacek M. Zurada,** University of Louisville

*Session Chairs*

1 Use of Neural Networks for Automatic Target Recognition, Classification, and Characterization
**Syed A. Rizvi,** CUNY/College of Staten Island

2 Use of Neural Networks for Image Coding and Noise Reduction
**Anne Marie P. Marinelli,** U.S. Army Research Laboratory

3 Neural Network-based Feature Extraction, Reconstruction, and Fusion
**Syed A. Rizvi,** CUNY/College of Staten Island

# Introduction

Artificial neural network architectures, due to their highly parallel characteristics, offer unique solutions to problems encountered in image processing applications. This conference was intended to bring together academic and industrial researchers from all over the world to interact on their ideas and applications.

The conference contains 19 papers organized into 5 sessions that cover recent advances and applications of the artificial neural networks to image processing. Presentations included topics such as use of neural networks for automatic target recognition, classification, image coding, noise reduction, feature extraction, fusion, segmentation, three-dimensional reconstruction, texture classification, fuzzy neural networks, and feedback neural networks.

The chairs wish to thank all the contributors and session chairs for their contribution to a very successful meeting.

**Nasser M. Nasrabadi**
**Aggelos K. Katsaggelos**

# Contents

# Conference Committee

*Conference Chairs*

**Nasser M. Nasrabadi,** U.S. Army Research Laboratory
**Aggelos K. Katsaggelos,** Northwestern University

*Program Committee*

**Pierre Baldi,** Jet Propulsion Laboratory
**Yoshua Bengio,** Université de Montréal (Canada)
**Bir Bhanu,** University of California/Riverside
**Terry M. Caelli,** Curtin University of Technology (Australia)
**Rama Chellappa,** University of Maryland/College Park
**Chang Y. Choo,** California State University/San Jose
**John Daugman,** University of Cambridge (UK)
**Sandor Z. Der,** U.S. Army Research Laboratory
**Edward R. Dougherty,** Texas A&M University
**Kunihiko Fukushima,** Osaka University (Japan)
**Erol Gelenbe,** Duke University
**David H. Haussler,** University of California/Santa Cruz
**Nicolaos B. Karayiannis,** University of Houston
**Christof Koch,** California Institute of Technology
**Bart Kosko,** University of Southern California
**Sun-Yuan Kung,** Princeton University
**Yann LeCun,** AT&T Bell Laboratories
**Richard P. Lippmann,** MIT Lincoln Laboratory
**Vincent Mirelli,** U.S. Army Research Laboratory
**Erkki Oja,** Helsinki University of Technology (Finland)
**Sankar K. Pal,** Indian Statistical Institute
**Tomaso A. Poggio,** MIT AI Laboratory
**Syed A. Rizvi,** CUNY/College of Staten Island
**David E. Rumelhart,** Stanford University
**Michael Seibert,** MIT Lincoln Laboratory
**Christoph Von Der Malsburg,** University of Southern California
**Jacek M. Zurada,** University of Louisville

*Session Chairs*

1 Use of Neural Networks for Automatic Target Recognition, Classification, and Characterization
**Syed A. Rizvi,** CUNY/College of Staten Island

2 Use of Neural Networks for Image Coding and Noise Reduction
**Anne Marie P. Marinelli,** U.S. Army Research Laboratory

3 Neural Network-based Feature Extraction, Reconstruction, and Fusion
**Syed A. Rizvi,** CUNY/College of Staten Island

## SESSION 1

# Use of Neural Networks for Automatic Target Recognition, Classification, and Characterization

# EIGENSPACE TRANSFORMATION FOR AUTOMATIC TARGET DETECTION

Lipchen Alex Chan, Nasser M. Nasrabadi, and Don Torrieri

US Army Research Laboratory, Attn: AMSRL-SE-SE
2800 Powder Mill Road, Adelphi, MD 20783, USA

## ABSTRACT

In this paper, two eigenspace transformations are examined for feature extraction and dimensionality reduction in an automatic target detector. The transformations considered in this research are principal component analysis (PCA) and the eigenspace separation transform (EST). These transformations differ in their capabilities to enhance the class separability and to compact the information (energy) for a given training set. The transformed data, obtained by projection of the normalized input images onto a chosen set of eigentargets, are fed to a multilayer perceptron (MLP) that decides whether a given input image is a target or clutter. In order to search for the optimal performance, we use different sets of eigentargets and construct the matching MLPs. Although the number of hidden layers is fixed, the numbers of inputs and weights of these MLPs are proportional to the number of eigentargets selected. These MLPs are trained with a modified Qprop algorithm that maximizes the target-clutter class separation at a predefined false-alarm rate. Experimental results are presented on a huge and realistic data set of forward-looking infrared (FLIR) imagery.

**Keywords:** Principal component analysis, eigentargets, eigenspace separation, multilayer perceptron, automatic target detection, FLIR imagery.

## 1. INTRODUCTION

Human beings are usually very good at detecting and recognizing different targets even in relatively crowded and changing environments. However, human performance deteriorates drastically in a low-visibility environment or after an extended period of surveillance. Furthermore, certain working environments are either inaccessible or too hazardous for human beings. To compensate for such limitations of human operators, an accurate and versatile automatic target recognition (ATR) system is needed. For example, an ATR system in a battlefield might alert graveyard-shift watchmen with accurate information about any approaching vehicle, so that appropriate responses could be made in a timely fashion. Such a system might also reduce the workloads of pilots or tank commanders significantly by suggesting effective responses in real time.

Unfortunately, the development of such systems is hampered by large numbers of target classes and aspects, long viewing ranges, obscured targets, high-clutter background, different geographic and weather conditions, sensor noise, and variations caused by translation, rotation, and scaling of the targets. Furthermore, the recognition problem is made even more challenging[1,2] by the inconsistencies in the signatures of the targets, similarities between the signatures of different targets, limited training and testing data, camouflaged targets, the nonrepeatability of target signatures, and the difficulty of using contextual information (when it is available to the recognition system). To overcome these difficulties, Lampinen and Oja[3] subdivided the recognition task into two appropriate substages: feature extraction and classification. Using a combination of Gabor filters and multilayer self-organizing maps, Lampinen and Oja mapped the original images into a feature space of reduced dimensionality and complexity. A smaller, supervised subspace network classifier was then used to perform the classification in this feature space. The resulting system could handle a moderate number of classes for recognizing faces with relatively strong tolerance to distortions.

Further author information —
L.A.C.: Email: chan@netkonnect.net; Phone: 301-394-1677; Fax: 301-394-5357
N.M.N. (correspondence): Email: nnasraba@ragu.arl.mil; Phone: 301-394-0806; Fax: 301-394-5234
D.T.: E-mail: dtorr@arl.mil; Phone: 301-394-2484; Fax: 301-394-4797



Part of the IS&T/SPIE Conference on Applications of Artificial Neural Networks in Image Processing IV • San Jose, California • January 1999
SPIE Vol. 3647 • 0277-786X/99/$10.00

A complete ATR system may consist of several algorithmic components, such as preprocessing, detection, segmentation, feature extraction, classification, prioritization, tracking, and aimpoint selection.[1] The detection module is certainly one of the most important components, because the whole ATR system will not function properly without an excellent detector. Over the years, a number of detection algorithms have been proposed for ATR systems, such as the virtual agile retina target acquisition and classification (VARTAC) system proposed by Hecht-Nielsen et al.,[4] the fusion of morphological wavelet transform (MWT) algorithm and Gabor basis function (GBF) detection algorithm proposed by Casasent et al.,[5] and the ATR relational template matching (ARTM) algorithm proposed by Kramer et al.[6] A common problem for detection algorithms is false alarms, as shown in Figure 1, in which the boxes indicate the potential target areas that were detected by the ARTM algorithm. Techniques for reducing false-alarm rates are usually part of the detection algorithm; an example is fusing the output from different detection algorithms, a technique described by Casasent et al.[5]

In this paper, we present a clutter rejector that uses eigentargets obtained by two different methods for feature extraction and a multilayer perceptron (MLP) for clutter rejection. The inputs of this clutter rejector are the potential target areas (image chips) detected by the ARTM algorithm mentioned above, as well as other manually selected clutter chips from the same set of second-generation forward-looking infrared (FLIR) imagery. Section 2 discusses the two methods that we used to extract the eigentargets from the training images. Section 3 describes the neural clutter rejector, which uses the eigentargets as feature templates. Experimental results are presented in Section 4 and conclusions in Section 5.

## 2. EIGENTARGETS

In our experiments, we used two methods to obtain the eigentargets from a given set of training images. Principal component analysis is the most basic method, from which the more complicated eigenspace separation transform method is derived.

### 2.1. Principal Component Analysis

Also referred to as the Hotelling transform or the discrete Karhunen-Loève transform, principal component analysis (PCA) is based on statistical properties of vector representations. PCA is an important tool for image processing because it has several useful properties, such as decorrelation of data and compaction of information (energy).[7] We provide here a brief summary of the basic theory of PCA.

Assume a population of random vectors of the form

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} . \tag{1}$$

The *mean vector* and the *covariance matrix* of the vector population $\mathbf{x}$ are defined as

$$\mathbf{m_x} = E\{\mathbf{x}\} , \tag{2}$$

$$\mathbf{C_x} = E\{(\mathbf{x}-\mathbf{m_x})(\mathbf{x}-\mathbf{m_x})^T\} , \tag{3}$$

where $E\{\arg\}$ is the expected value of the argument, and $T$ indicates vector transposition. Because $\mathbf{x}$ is $n$-dimensional, $\mathbf{C_x}$ is a matrix of order $n \times n$. Element $c_{ii}$ of $\mathbf{C_x}$ is the variance of $x_i$ (the $i$th component of the $\mathbf{x}$ vectors in the population), and element $c_{ij}$ of $\mathbf{C_x}$ is the covariance between elements $x_i$ and $x_j$ of these vectors. The matrix $\mathbf{C_x}$ is real and symmetric. If elements $x_i$ and $x_j$ are uncorrelated, their covariance is zero, and therefore $c_{ij} = c_{ji} = 0$. For $N$ vector samples from a random population, the mean vector and covariance matrix can be approximated from the samples by

$$\mathbf{m_x} = \frac{1}{N}\sum_{p=1}^{N} \mathbf{x}_p , \tag{4}$$

$$\mathbf{C_x} = \frac{1}{N}\sum_{p=1}^{N} (\mathbf{x}_p\mathbf{x}_p^T - \mathbf{m_x}\mathbf{m_x}^T) . \tag{5}$$

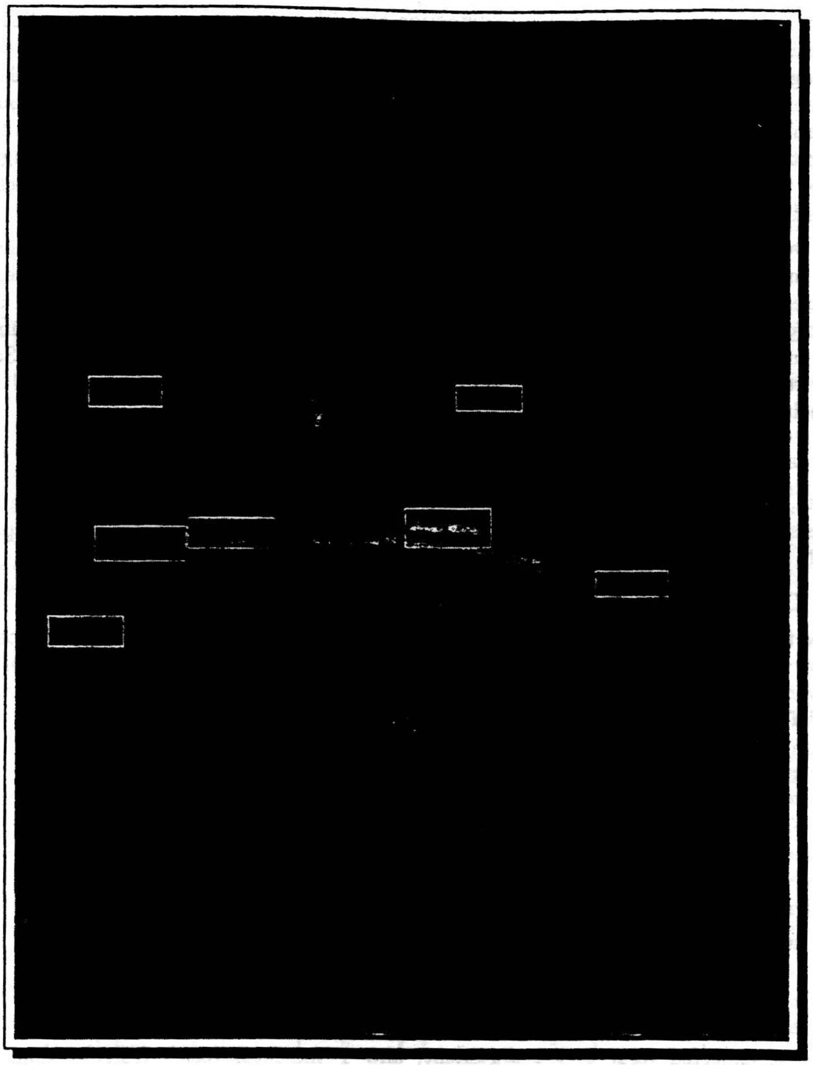

**Figure 1. Potential target areas detected by the ARTM algorithm.**

Because $C_x$ is real and symmetric, we can always find a set of $n$ orthonormal eigenvectors for this covariance matrix. A simple but foolproof algorithm to find these orthonormal eigenvectors for all real symmetric matrices is the Jacobi method.[8] The Jacobi algorithm consists of a sequence of orthogonal similarity transformations. Each transformation is just a plane rotation designed to annihilate one of the off-diagonal matrix elements. Successive transformations undo previously set zeros, but the off-diagonal elements get smaller and smaller, until the matrix is effectively diagonal (to the precision of the computer). We obtain the eigenvectors by accumulating the product of transformations during the process, while the main diagonal elements of the final diagonal matrix are the eigenvalues. Alternatively, a more complicated method based on the QR algorithm for real Hessenberg matrices can be used.[8] This is a more general method because it can extract eigenvectors from a nonsymmetric real matrix. Furthermore, it becomes increasingly more efficient than the Jacobi method as the size of the matrix increases. Given the considerable increase in efficiency for the size of our covariance matrix, we chose the QR method for our experiments described

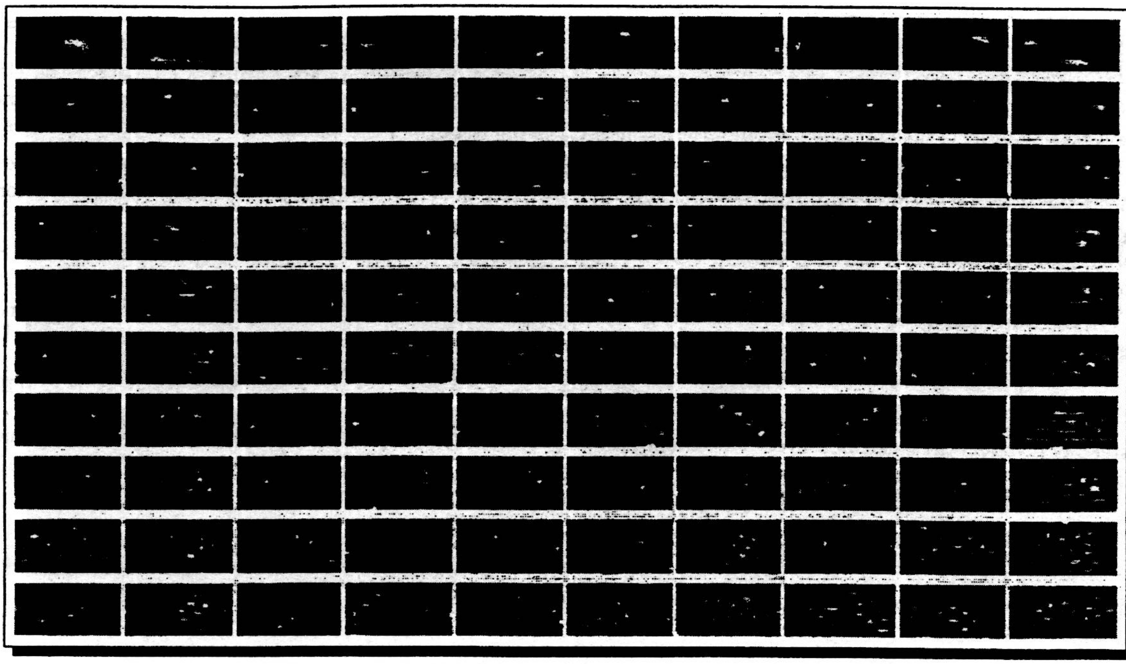

**Figure 2.** The 100 most dominant PCA eigenvectors (eigentargets) for the targets in the training set.

in this paper. Figure 2 shows the 100 most dominant eigenvectors representing the targets in the training set. Having the largest eigenvalues, these eigenvectors capture the greatest variance or energy among the training data. Therefore, their contrast level is also significantly higher than that of the remaining eigenvectors.

Let $\mathbf{e}_i$ and $\lambda_i, i = 1, 2, \ldots, n$, be the eigenvectors and the corresponding eigenvalues of $\mathbf{C_x}$, sorted in a descending order so that $\lambda_j \geq \lambda_{j+1}$ for $j = 1, 2, \ldots, n-1$. Let $\mathbf{A}$ be a matrix whose rows are formed from the eigenvectors of $\mathbf{C_x}$, such that

$$\mathbf{A} = \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \vdots \\ \mathbf{e}_n \end{bmatrix} . \tag{6}$$

This $\mathbf{A}$ matrix can be used as a transformation matrix that maps the $\mathbf{x}$'s into vectors denoted by $\mathbf{y}$'s, as follows:

$$\mathbf{y} = \mathbf{A}(\mathbf{x} - \mathbf{m_x}) . \tag{7}$$

The $\mathbf{y}$ vectors resulting from this transformation have a zero mean vector; that is, $\mathbf{m_y} = \mathbf{0}$. The covariance matrix of the $\mathbf{y}$'s can be computed from $\mathbf{A}$ and $\mathbf{C_x}$ by

$$\mathbf{C_y} = \mathbf{A}\mathbf{C_x}\mathbf{A}^T . \tag{8}$$

Furthermore, $\mathbf{C_y}$ is a diagonal matrix whose elements along the main diagonal are the eigenvalues of $\mathbf{C_x}$; that is,

$$\mathbf{C_y} = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix} . \tag{9}$$

Because the off-diagonal elements of $\mathbf{C_y}$ are zero, the elements of the $\mathbf{y}$ vectors are uncorrelated. Since the elements along the main diagonal of a diagonal matrix are its eigenvalues, $\mathbf{C_x}$ and $\mathbf{C_y}$ have the same eigenvalues and eigenvectors. In fact, the transformation of the $\mathbf{C_x}$ into $\mathbf{C_y}$ is the essence of the Jacobi algorithm described above.

Therefore, through the PCA transformation, a new coordinate system is established. The origin of this new coordinate system is at the centroid of the population, $\mathbf{m_x}$, with new axes in the direction specified by the eigenvectors $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$. The eigenvalue $\lambda_i$ becomes the variance of component $y_i$ along eigenvector $\mathbf{e}_i$. With its ability to realign unknown data into a new coordinate system based on the principal axes of the data, PCA is often used to achieve rotational invariance in image processing tasks.

On the other hand, we may want to reconstruct vector $\mathbf{x}$ from vector $\mathbf{y}$. Because the rows of $\mathbf{A}$ are orthonormal vectors, $\mathbf{A}^{-1} = \mathbf{A}^T$. Therefore, any vector $\mathbf{x}$ can be reconstructed from its corresponding $\mathbf{y}$ by the relation

$$\mathbf{x} = \mathbf{A}^T\mathbf{y} + \mathbf{m_x} . \tag{10}$$

Instead of using all the eigenvectors of $\mathbf{C_x}$, we may pick only $k$ eigenvectors corresponding to the $k$ largest eigenvalues and form a new transformation matrix $\mathbf{A}_k$ of order $k \times n$. In this case, the resulting $\mathbf{y}$ vectors would be $k$-dimensional, and the reconstruction given in Equation 10 would no longer be exact. The reconstructed vector using $\mathbf{A}_k$ is

$$\hat{\mathbf{x}} = \mathbf{A}_k^T\mathbf{y} + \mathbf{m_x} . \tag{11}$$

The mean square error between $\mathbf{x}$ and $\hat{\mathbf{x}}$ can be computed by the expression

$$\epsilon = \sum_{j=1}^{n} \lambda_j - \sum_{j=1}^{k} \lambda_j = \sum_{j=k+1}^{n} \lambda_j . \tag{12}$$

Because the $\lambda_j$'s decrease monotonically, Equation 12 shows that we can minimize the error by selecting the $k$ eigenvectors associated with the $k$ largest eigenvalues. Thus the PCA transform is optimal in the sense that it minimizes the mean square error (MSE) between the vectors $\mathbf{x}$ and their approximations $\hat{\mathbf{x}}$.

### 2.2. Eigenspace Separation Transform

The eigenspace separation transform (EST) has been proposed by Torrieri as a preprocessor to a neural binary classifier.[9] The goal of the EST is to transform the input patterns into a set of projection values such that the size of a neural classifier is reduced and its generalization capability is increased. The size of the neural network is reduced, because the EST projects an input pattern into an orthogonal subspace of smaller dimensionality. The EST also tends to produce projections with different average lengths for different classes of input, and hence improves the discriminability between the targets. In short, the EST preserves and enhances the classification information needed by the subsequent classifier. It has been used in a mine detection task with some success.[10]

The transformation matrix $\mathbf{S}$ of the EST can be obtained as follows.

1. Compute the $n \times n$ correlation difference matrix

$$\hat{\mathbf{M}} = \frac{1}{N_1}\sum_{p=1}^{N_1} \mathbf{x}_{1p}\mathbf{x}_{1p}^T - \frac{1}{N_2}\sum_{q=1}^{N_2} \mathbf{x}_{2q}\mathbf{x}_{2q}^T , \tag{13}$$

   where $N_1$ and $\mathbf{x}_{1p}$ are the number of patterns and the $p$th training pattern of Class 1, respectively. $N_2$ and $\mathbf{x}_{2q}$ are similarly related to Class 2 (which is the complement of Class 1).

2. Calculate the eigenvalues of $\hat{\mathbf{M}}$, $\{\lambda_i \mid i = 1, 2, ..., n\}$.

3. Calculate the sum of the positive eigenvalues

$$E_+ = \sum_{i=1}^{n} \lambda_i \qquad \text{if} \quad \lambda_i > 0 , \tag{14}$$

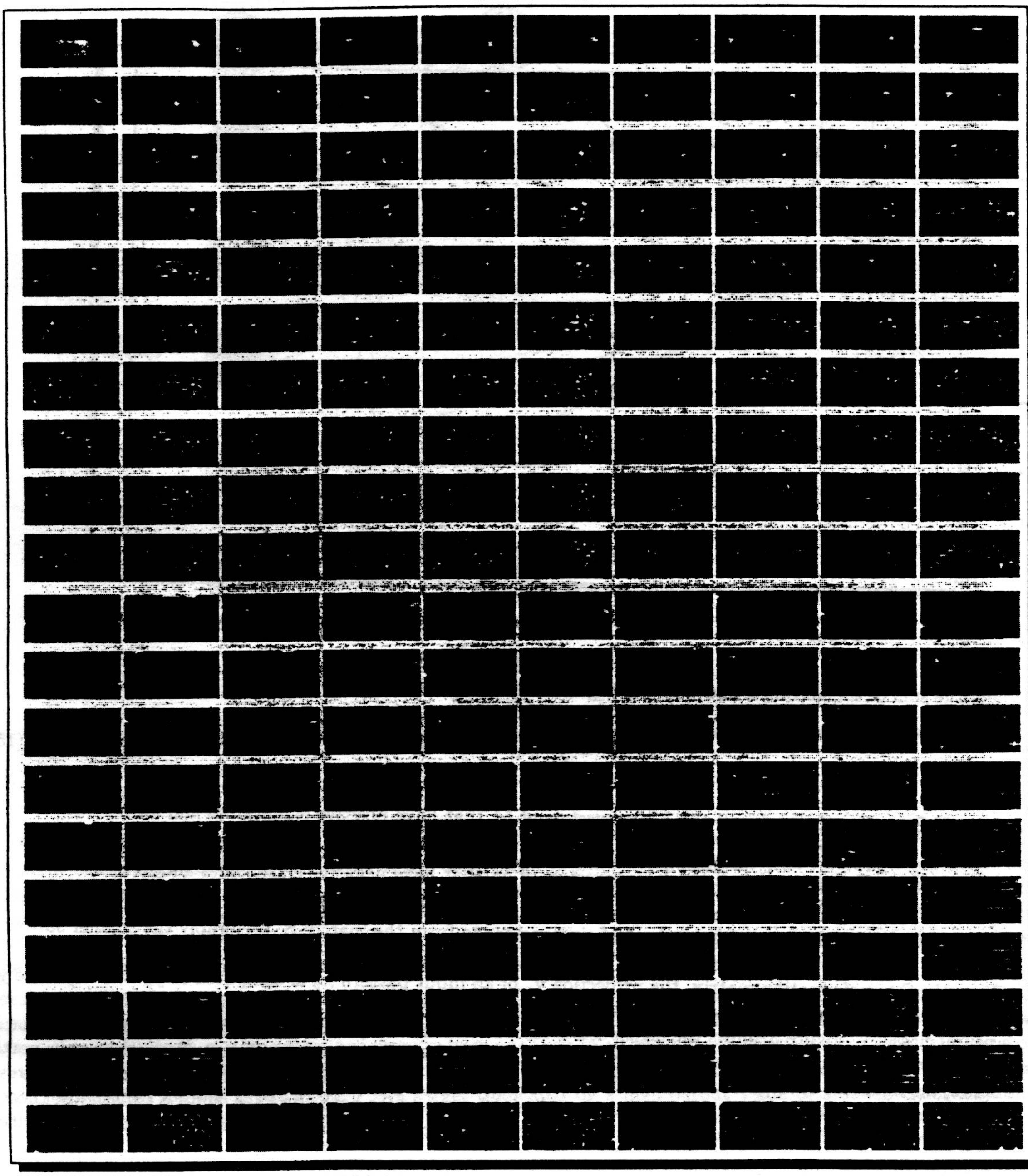

**Figure 3.** The 100 most dominant EST eigenvectors (eigentargets) associated with positive (top) and negative (bottom) eigenvalues for the training set.

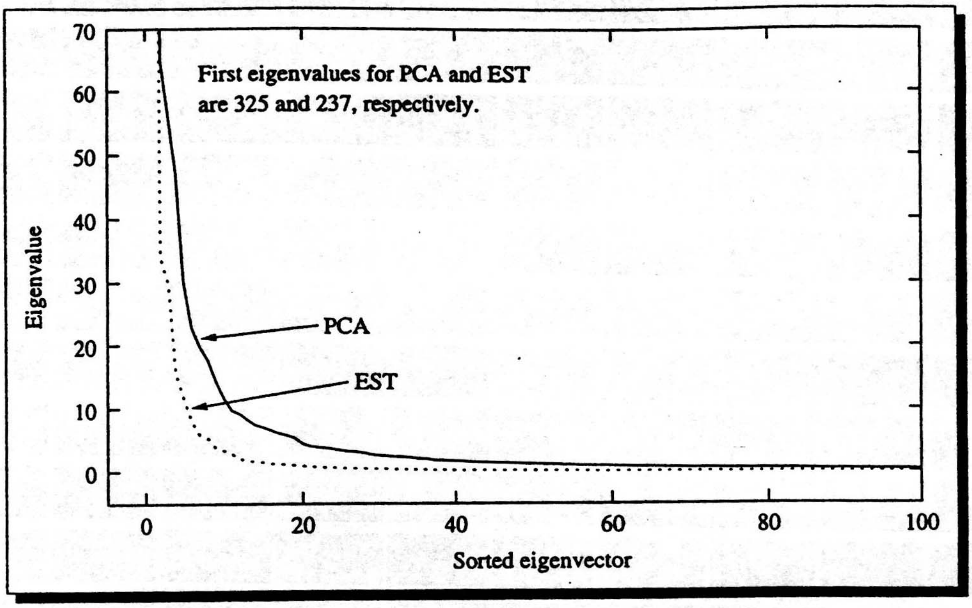


**Figure 4.** Rapid attenuation of eigenvalues in PCA and EST transforms.

and the sum of the absolute values of the negative eigenvalues

$$E_- = \sum_{i=1}^{n} |\lambda_i| \qquad \text{if} \quad \lambda_i < 0\,. \tag{15}$$

(a) If $E_+ > E_-$, then take all the $k$ eigenvectors of $\hat{\mathbf{M}}$ that have positive eigenvalues and form the $n \times k$ matrix $\mathbf{S}$.

(b) If $E_+ < E_-$, then take all the $k$ eigenvectors of $\hat{\mathbf{M}}$ that have negative eigenvalues and form the $n \times k$ matrix $\mathbf{S}$.

(c) If $E_+ = E_-$, then use either subset of eigenvectors to form the matrix $\mathbf{S}$, preferably the smaller subset.

Given the S transformation matrix, the projection $\mathbf{y}_p$ of an input pattern $\mathbf{x}_p$ is computed as $\mathbf{y}_p = \mathbf{S}^T\mathbf{x}_p$. The $\mathbf{y}_p$, with a smaller dimension (because $k \leq n$) and presumably larger separability between the classes, can then be sent to a neural classifier. Figure 3 shows the eigenvectors associated with the positive and negative eigenvalues of the $\hat{\mathbf{M}}$ matrix that was computed with the target chips as Class 1 and the clutter chips as Class 2. From the upper part of the figure, the signature of targets can be clearly seen. On the other hand, the lower part represents all the features of clutters. As shown in Figure 4, while the eigenvalues diminish rapidly for both the PCA and EST methods, those of the EST decrease even faster. In other words, the EST may produce a higher compaction in contextual information.

## 3. CLUTTER REJECTION

The inputs for our clutter rejection module are the image chips extracted from bigger scenes, as illustrated in Figure 1. The size of these image chips is fixed to a predefined dimension, which is common to both the targets and the clutters. To reduce the background information in target chips, we clip each image chip at a size that equals the dimension of the largest target in our training set. After the background removal, the input image is scaled to a preferred size based on a linear interpolation technique. This scaling is needed to achieve an image size that is efficient for feature extraction via the eigenspace transformation, while an effective amount of information is retained in the image.