AN ABSTRACT OF THE THESIS OF

Jin-Woo Eo for the degree of Doctor of Philosophy in Electrical and Computer Engineering presented on February 24, 1992.
Title : Multi-Variate Morphological Filtering with Applications to Color Image Processing

Redacted for Privacy

Abstract approved:_____
W. J. Kolodziej

Mathematical morphology, developed in the early 1960's for single-component signals, has been applied to a number of image processing applications. This investigation examines the systematic extension of mathematical morphology to multi-variate signals.

Two approaches are considered. The first approach, the extension of the theory of single-component morphological filters to multi-variate case, fails for reason of the lack of ordering within signal range space. Therefore, as a second approach, a two stage processing technique was proposed, consisting of the maximum separation of the object from its background feature and separate morphological filtering of each component. To separate the object from its background, a mapping technique, based upon the normalization and simultaneous diagonalization of sample covariance matrices (NAD-CVM), was applied. Sample variance morphological measure interpretation demonstrated that NAD-CVM mapping constitutes an excellent preprocessing tool for morphological filtering of multi-variate signals. An unsupervised NAD-CVM implementation and a morphological edge detector were tested experimentally to verify the properties of the theoretical algorithm. In addition, a method for the application of the proposed method to the analysis of color images was presented.

Multi-Variate Morphological Filtering
with Applications to Color Image Processing

by

Jin-Woo Eo

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Completed February 24, 1992

Commencement June 1992

APPROVED:

Redacted for Privacy

_____ _03-16-1992_
Associate Professor of department of Electrical and Computer
Engineering in charge of major

Redacted for Privacy

_ _ _ _   _____
Head of) department of Electrical and Computer Engineering

Redacted for Privacy

_ _ ＼_   _____
Dean of Graduate School

Date thesis is presented_____February 24, 1992_____

Typed by Jin-Woo Eo for_____Jin-Woo Eo_____

To my God

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

LIST OF FIGURES

LIST OF FIGURES (continued)

LIST OF FIGURES (continued)

LIST OF FIGURES (continued)

LIST OF TABLES

MULTI-VARIATE MORPHOLOGICAL FILTERING WITH APPLICATIONS TO COLOR IMAGE PROCESSING

1. INTRODUCTION

1.1 Motivations, Directions, and Contributions of this Study

Mathematical morphology, which was developed in early 1960's by Matheron and Serra, has been used for numerous image processing and analysis applications. Morphological filtering has proved to be an indispensable technique for the development of vision system algorithms. There are several reasons for the popularity of mathematical morphology, including the following:

(1) Morphological filters have the ability to filter out wide-band noise while preserving well-defined signal discontinuities or fast transitions; this can not be accomplished using a linear filter.

(2) Mathematical morphology allows to quantify the shape of a signal. This is an important property since the identification of objects and their features is related directly to shape matching; using the shape concept is a natural approach to pattern recognition.

(3) Morphological algorithms can easily be implemented for parallel computers (*e.g.* cellular logic computers), thus allowing very fast processing; references for parallel computers with the capability of performing morphological operations are listed in [9].

Mathematical morphology was originally developed for single-component signals. The current investigation is directed at the systematic extension of mathematical morphology to multi-variate signal analysis. Reported results have been based upon separate application of morphological operations to each component of a multi-

variate signal. From the current study, it may be observed that separate morphological operation for each signal component may yield poor results when the signal components are *correlated.* To address the latter problem, two approaches have been developed. The first is based upon the determination of a mapping method which provides *uncorrelated* signal components. The second one is to generalize the mathematical morphology for multi-variate signals. The basic research direction undertaken for these two approaches is summarized in Figure 1-1.
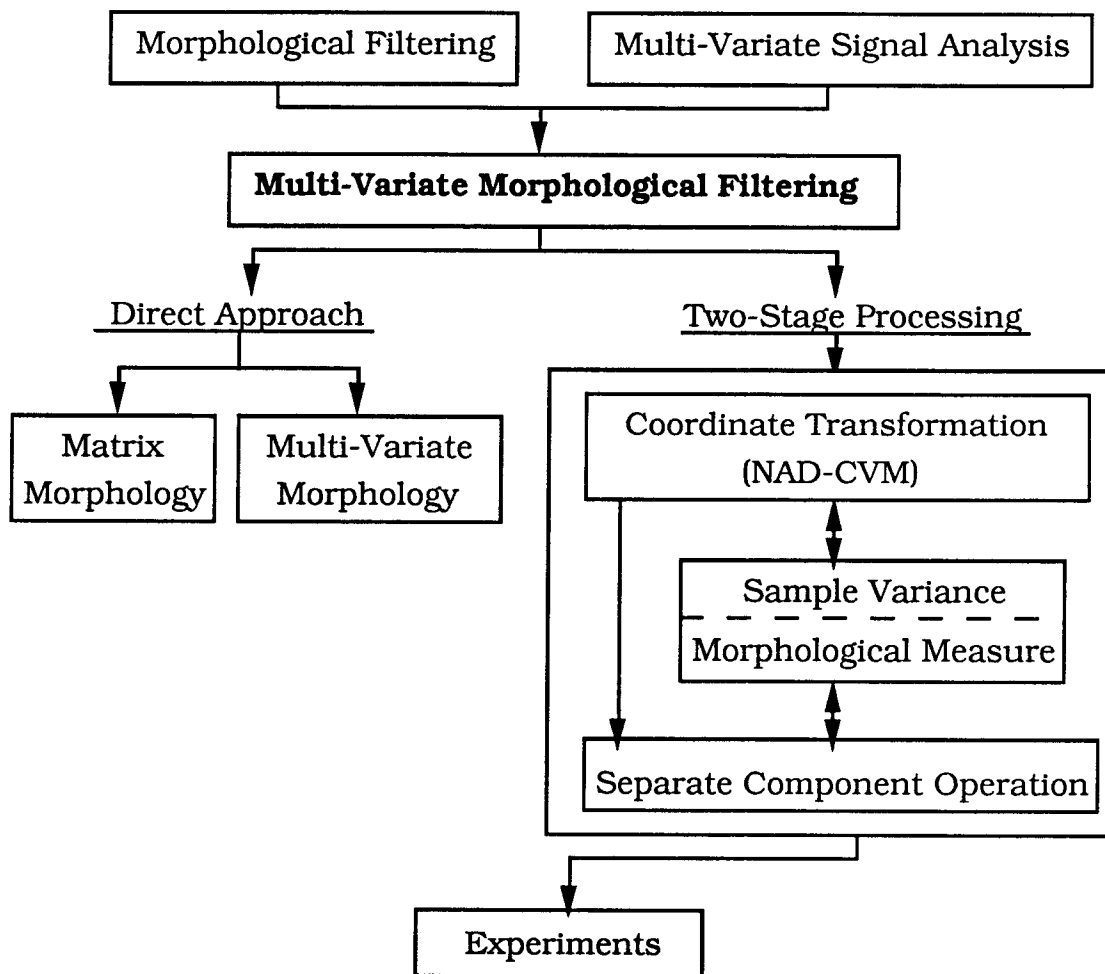
Figure 1-1 Basic research direction.

In [26], an attempt is made to extend the current concept of mathematical morphology (i.e., *scalar morphology*) to a *matrix*

morphology formalism. An image matrix was introduced as an array of separate images, as well as a structuring element for an image matrix. The umbra of a matrix was defined as an array of each component umbra. Thus *matrix morphology* is a method based upon separate morphological operations applied to each of the signal components. However, a principal difficulty in extending mathematical morphology to multi-variate signals is -- since basic morphological operations are based on maximum and minimum operations -- that the concept of ordering in signal range space must be preserved. This generalization of the ordering of real-line to a multi-dimensional space has not been fully studied, and only so-called *partial ordering* has been achieved. In this case, concepts of convex, cone-like set inclusions are potentially usable, but the complicated theory underlying this approach has obviated arrival at useful applications.

Therefore, for the current investigation attention was directed toward methods for processing signals in two stages. First, for convenient implementation of the morphological point of view, *decorrelation* of the signal components was performed, following which morphological operations were used to filter the *decorrelated* components separately. The principal component, or Karhunen-Loeve method, is a well known transformation which uncorrelates signal components. It should be noted that this method, although optimal with respect to representation of multi-variate signals in reduced dimensionality coordinates, is not necessarily optimal with respect to separation of pattern classes. The latter is important to the solution of the feature extraction problem. This drawback is incurred since a multi-variate signal is considered as single-class data.

Numerous mapping methods have been proposed for generalization of the Karhunen-Loeve method to cope with the problem of multi-class discrimination, particularly in the field of statistical pattern recognition. One approach, proposed by Fukunaga and Koontz [5] for the two-class problem, consisted of the use of a normalization process prior to application of principal components analysis. The fact that the so-called *best* feature for the analysis of one class, and the so-called *poorest* feature for the analysis of the second class, can be represented within the same component, has motivated

efforts to adapt the normalization process to morphological analysis. Then, following normalization, morphological operation by an appropriate structuring element can be used to extract information from each class selectively; that is, the essential shapes of one class are preserved while the irrelevant shapes of the second class are removed. The notion of *selective morphological operation* is studied here by relating sample variance and the performance of morphological information extraction to the *average height* of opening.

The mapping method, proposed in [5] is based on a normalization and simultaneous diagonalization of two autocorrelation matrices (NAD-CRM). Since the autocorrelation matrix is calculated by combining the covariance matrix with the mean vector, the NAD-CRM method provides maximum discrimination of variances and means of two classes. However, in the sense that for the best morphological separation, the maximum ratio of sample variances must be sought, this has not proved to be a desirable method. For this reason, the normalization and simultaneous diagonalization of sample covariance matrices (NAD-CVM) is proposed. In general, methods of this class have not received adequate attention as generalizations of the Karhunen-Loeve method.

Some of the useful properties of NAD-CVM include:

(1) The eigenvalues, obtained by the process of normalization followed by the simultaneous diagonalization of sample covariance matrices, provide an interpretation of sample variances.

(2) The eigenvectors for the normalized covariance matrix from one class are identical to those from the second class, and all eigenvalues are bounded by 0 and 1. Moreover, the sums of the corresponding eigenvalues are equal to one.

(3) There is an *unsupervised* NAD-CVM algorithm, which requires only one class sample covariance matrix to be known *a priori*.

When the properties (1) and (2) are combined, the NAD-CVM method can be used to produce an optimal signal component with maximum discrimination of sample variances (i.e., maximum ratio of sample variance from one class to the second class). Property (3) enables the

implementation of the NAD-CVM algorithm for a variety of two-class problems. The basic limitation of NAD-CVM is that it cannot be generalized to multi-class problems. Nonetheless, since an image can often be considered as a combination of object features and background features, respectively, Class 1 and 2, it is a useful mapping technique for multi-variate image analysis.

Morphological opening and closing operations provide selective filtering property, that is, they filter out shapes which do not fit structuring element while preserving those shapes which do resemble the structuring element. This property is attractive for the purposes of image processing and analysis. The effectiveness of the selective filtering by the use of the NAD-CVM method can be maximized, providing component with a maximum ratio of the sample variance of the object (which contains useful geometrical information) to the sample variance of the background (which contains irrelevant shapes). Furthermore the relationship between variance and *average height* can be used for the selection of a proper structuring element in a morphological filter.

In combination, the two-class separation property, morphological measure interpretation, and the ability to conduct an unsupervised implementation justify the effort of applying NAD-CVM mapping method to processing multi-variate signals. For two-class multi-variate signal analysis problems, there are numerous applications for which morphological filtering may be used. For example, to develop these concepts for the present study, the RGB color images of wooden boards were used, each with a defect, such as blue stain, an intergrown knot, or a spike knot. The performance of a morphological edge detection using open-close smoothing was experimentally examined. The quality of binary edge image was judged by the amount of detail provided in the object region (i.e., the wood defects) and lack of the detail in the background (i.e., among the wood grains). The problem of detecting blue stains provided the original motivation for the study of multi-variate (i.e., color) image analysis methods. Since the distribution of graylevels and texture of blue stain are similar to those of wood grains, blue stains cannot be easily detected using grayscale image. In case of blue stain detection,

separate operations for each RGB color image plane were shown by experimentation to be ineffective. However, a reasonable binary edge image were obtained using the NAD-CVM method, at the same time demonstrating the inadequacy of results obtained from the application of the NAD-CRM and Karhunen-Loeve methods. In case of knots detection, the results obtained by NAD-CVM method were similar to those obtained from the use of NAD-CRM, but at the same time were better than those obtained from the use of grayscale, RGB components, and Karhunen-Loeve method.

1.2 Thesis Organization

References to existing results on the mapping methods, mathematical morphology, and color image analysis are provided in the first sections of Chapter 2, 3, and 4, respectively. In Chapter 2, the properties of the principal components method are reviewed, the derivation and the properties of the NAD-CVM method are presented, and relevant mathematical results for an unsupervised NAD-CVM algorithm are discussed. Specifically, the mapping technique discussed in Section 2.2 are related to the mathematical morphology concepts for multi-variate signal analysis. Chapter 3 is initiated with a review of mathematical morphology, and an extension of grayscale morphology to multi-variate signal is proposed. The application of this technique produced different results than when separate morphological operations were conducted on each component. So-called *multi-variate morphology* is then introduced, using a generalization of the umbra concept, and sample variance is related to the properties of structuring elements, thus establishing a linkage between mapping techniques and mathematical morphology. In Chapter 4 color coordinate systems and the basic concepts of color image analysis are reviewed and the experimental results are presented. The conclusions derived from this investigation, as well as suggestions for future research, are presented in Chapter 5.

2. COORDINATES TRANSFORMATION TECHNIQUES FOR PREPROCESSING OF MULTI-VARIATE SIGNALS

Transformation of coordinates and reduction of dimensionality are two important issues in multi-variate signal analysis. There are numerous coordinate transformation (or mapping) techniques for pattern analysis [23]. Those mapping techniques can be divided into two categories, linear and nonlinear. Nonlinear mapping methods may be highly suitable in specific cases, however they lack generality in both, the theory and applications. Here the transformation techniques are used only as a vehicle to generalize morphological filtering. For this reason only linear transformation methods are investigated in this research. There are two basic linear transformation techniques which depend on a specific criteria of optimization :

(1) minimization of the mean-square reduced dimensionality representation error - this technique is called *principal component method* or Karhunen-Loeve expansion;

(2) generalized declustering method, which finds the projection direction to maximize a performance criterion; depending on the criterion used this technique is known as Fisher's discriminant method, optimal discriminant plane method, or declustering method [31], [3], [30].

Here, the main objective of coordinate transformation is to maximize the separation between the object of interest and the background, which represents no interest from the signal analysis point of view. For morphological analysis purposes, maximun separation of sample mean vectors is not important, because morphological edge and boundary detection is not related to sample means of the object and the background. Instead, the morphological filtering is related to the local data structure within support region of

the structuring element. For that reason, generalized declustering method is not optimal either, since its criterion includes a global between-class scatter matrix. Thus our approach to coordinate transformation for morphological analysis focuses on principal component method. This method provides an optimal local data structure separation, in the sense that sample variance depends on structuring element selection. This is why the normalization and diagonalization of sample covariance matrices (NAD-CVM) method is proposed here rather than normalization and diagonalization of sample autocorrelation matrices (NAD-CRM) method, (Fukunaga and Koontz [5]).

2.1 Principal Component Method

The basic idea of principal component method is to describe the dispersion of a set of n points in p-dimensional space by introducing a new set of orthogonal coordinates so that the sample variance of the given points with respect to the new coordinates are in descending order of magnitude. Thus the first principal component is such that the projection of the data points onto it has maximum variance among all possible coordinates, the second principal component has maximum variance subject to being orthogonal to the first principal component, and so on. The underlying concept of principal component method is based on the discrete Karhunen-Loeve expansion.

2.1.1 Review of Principal Component Method

Let $x(i) \in R^p$ be a p-dimensional observation vector, and let the number of observation be n, i.e., i=1,2,$\cdots$,n. Here we do not consider any randomness about observation. Define the *sample mean* and *sample covariance* as follows :

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x(i) \quad (sample\ mean\ ),$$ (2-1)

$$K = \frac{1}{n-1}\sum_{i=1}^{n} (x(i) - \bar{x})\,(x(i) - \bar{x})^{T} \quad (sample\ covariance\ ),$$ (2-2)

where T denotes transposition.

The principal component method consists of :

(1) determine a normalized transformation vector, ϕ_1, $\phi_1^{T}\phi_1=1$, such that the linear combination, $\phi_1^{T}x(i)$, $i=1,2,\cdots,n$, has maximum sample variance in the class of all linear combinations;

(2) determine the j-th normalized transformation vector, ϕ_j, $\phi_j^{T}\phi_j=1$ such that the linear combination $\phi_j^{T}x(i)$, $i=1,2,\cdots,n$; $j=2,3,\cdots$, has maximum sample variance in the class of all normalized linear combinations of $x(i)$ orthogonal to $\phi_1^{T}x(i)$, $\phi_2^{T}x(i)$, $\cdots$, $\phi_{j-1}^{T}x(i)$.

The last step is repeated until determining ϕ_p.

For a given arbitrary p-dimensional vector ϕ, sample variance of $\phi^{T}x(i)$, $i=1,2,\cdots,n$, can be written as

$$\frac{1}{n-1}\sum_{i=1}^{n} (\,\phi^{T}x(i) - \overline{\phi^{T}x(i)}\,)^{2}\ ,$$ (2-3)

where

$$\overline{\phi^{T}x(i)} = \frac{1}{n}\sum_{i=1}^{n} \phi^{T}x(i)\ .$$ (2-4)

Equation (2-3) can be rewritten as

$$\frac{1}{n-1}\sum_{i=1}^{n} (\phi^{T}x(i) - \phi^{T}\bar{x})^{2} = \phi^{T}K\phi\ .$$ (2-5)

Thus the problem of finding ϕ_1 is equivalent to determining a nonzero ϕ_1 such that the ratio (Rayleigh quotient) $\phi_1^T K \phi_1 / \phi_1^T \phi_1$ is maximized. It is well known that maximum value of this ratio is the largest eigenvalue λ_1 of the matrix K, and that ϕ_1 is a normalized eigenvector corresponding to λ_1. Similarly the solution of ϕ_i, $i=2,\cdots,p$ is the eigenvector corresponding to λ_i, $i=2,\cdots,p$, of K and $\phi_i^T K \phi_i / \phi_i^T \phi_i = \lambda_i$.

Since the matrix K is nonnegative definite and symmetric, diagonalization of K can be always achieved using the normalized eigenvector matrix Φ, such that $\Phi^T \Phi = I$ and

$$\Phi^T K \Phi = \Lambda = diag\{\lambda_1, \lambda_2, \cdots, \lambda_p\}, \qquad (2-6)$$

where $\Phi = [\phi_1 | \phi_2 | \cdots | \phi_p]$, ϕ_i is an eigenvector corresponding to the eigenvalue λ_i, and the matrix I denotes the identity matrix. Without loss of generality, the eigenvalues, λ_i's, are arranged in descending order i.e. $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p$. Note that the matrix $\Phi^T K \Phi$, is the sample covariance matrix of $y = \Phi^T x$. The transformation $\phi_k^T x(i)$, $i=1,2,\cdots,n$; $k=1,2,\cdots,p$, is called the k-th principal component. Thus the principal component method corresponds to the diagonalization of sample covariance matrix. An equivalent approach is known as the discrete Karhunen-Loeve expansion. The dimensionality reduction of principal component method, which is discussed in the next section, is mainly based on the properties of the discrete Karhunen-Loeve expansion.

2.1.2 Properties of Principal Component Method

(1) Dimensionality reduction property

Consider the transformation,

$$y(i) = P^T x(i), \quad i=1,2,\cdots,n, \qquad (2-7)$$

where the vectors x and y are p-dimensional vectors, and pxp matrix P is the transformation matrix, such that $P = [p_1 | p_2 | \cdots | p_p]$, and $P^T P = I$. Note that $y(i)$ is an orthogonal transformation of $x(i)$.

Suppose that we use only k<p components of vector $\mathbf{y}(i)$, to reconstruct $\mathbf{x}(i)$, $i=1,2,\cdots,n$, replacing those components of $\mathbf{y}(i)$ which we do not calculate with the preselected constants as follows :

$$\widehat{\mathbf{x}(i)} = \sum_{j=1}^{k} y_j(i)\, \mathbf{p}_j + \sum_{j=k+1}^{p} c_j\, \mathbf{p}_j . \qquad (2\text{-}8)$$

In (2-8) $y_j(i)$ denotes the j-th component of $\mathbf{y}(i)$. The corresponding reconstruction error can be defined in the mean-square sense as :

$$\varepsilon^2 = \frac{1}{n-1} \sum_{i=1}^{n} |\,\mathbf{x}(i) - \widehat{\mathbf{x}(i)}\,|^2$$

$$= \frac{1}{n-1} \sum_{i=1}^{n} \left(\sum_{j=k+1}^{p} \sum_{l=k+1}^{p} (y_j(i)-c_j)(y_l(i)-c_l)\mathbf{p}_j^{T}\mathbf{p}_l \right)$$

$$= \frac{1}{n-1} \sum_{i=1}^{n} \sum_{j=k+1}^{p} (y_j(i)-c_j)^2$$

$$= \frac{1}{n-1} \sum_{j=k+1}^{p} \sum_{i=1}^{n} (y_j(i)-c_j)^2 . \qquad (2\text{-}9)$$

The optimum choice of c_j to minimize ε^2 is given by setting $\partial\varepsilon^2/\partial c_j = 0$, which yields $c_j = (1/n)\,\Sigma_{i=1,n}\; y_j(i) \triangleq \bar{y}_j$, $j=k+1,\cdots,p$. Consequently we can rewrite ε^2 as

$$\varepsilon^2 = \frac{1}{n-1} \sum_{i=1}^{n} \sum_{j=k+1}^{p} (y_j(i)-\bar{y}_j)^2$$

$$= \frac{1}{n-1} \sum_{i=1}^{n} \sum_{j=k+1}^{p} \mathbf{p}_j^{T}(\mathbf{x}(i)-\bar{\mathbf{x}})(\mathbf{x}(i)-\bar{\mathbf{x}})^{T}\mathbf{p}_j$$

$$= \sum_{j=k+1}^{p} \mathbf{p}_j^{T}\mathbf{K}\,\mathbf{p}_j . \qquad (2\text{-}10)$$

Now the problem is to find the optimal $\mathbf{p}_j$'s which minimize mean-square reconstruction error ε^2. It is well known that the eigenvectors

of the matrix **K** minimize ε^2 over all choices of orthonormal basis vectors, i.e., $\mathbf{p}_j = \phi_j$. Using Equation (2-6), the minimum mean-square error is

$$\varepsilon^2 = \sum_{j=k+1}^{p} \lambda_j \ . \tag{2-11}$$

Note that if y_j is replaced by a constant, then the mean square error increases by λ_j. Therefore given dimensionality reduction k, the minimum mean-square reconstruction error can be achieved by taking the first k principal components.

(2) Geometric interpretation

The equation :

$$(x\text{-}\bar{x})^T K^{-1} (x\text{-}\bar{x}) = d \ , \tag{2-12}$$

for nonnegative values of d, defines a family of concentric ellipsoids in the p-dimensional space of **x**. The principal component transformation of the data is the projection of observation data onto the principal axes of this concentric ellipsoid family. The two-dimensional case illustration is shown in Figure 2-1. The original coordinates (x_1, x_2), are transformed by a shift of the origin to the sample mean $(\bar{x}_1, \bar{x}_2)$, followed by a rotation about this origin that yields the principal component coordinates (y_1, y_2).
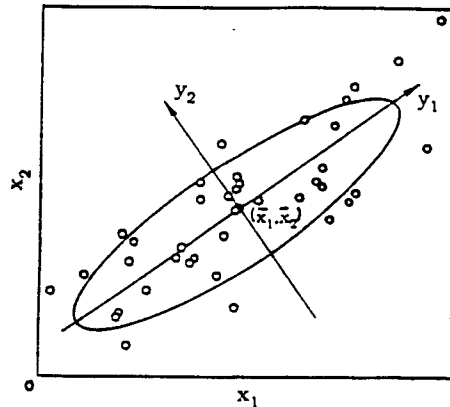
Figure.2-1 Geometric interpretation of principal component method.

2.2 Mapping Technique Using Normalization and Simultaneous Diagonalization of Sample Covariance Matrices (NAD-CVM)

2.2.1 Basic Properties of NAD-CVM Method

The mapping reviewed in Section 2.1 is also called the *total principal component mapping*, since the projection axes are computed from the total covariance matrix of a given data set. It is noted that although the total principal component mapping is optimal with respect to fitting data, it is not necessarily optimal with respect to separating pattern classes as far as feature extraction problem in pattern recognition is concerned. This suggests that we need another mapping technique which satisfies our objective, i.e., the maximum separation between the object and the background.

There are numerous separation techniques suggested in the literature [23],[6], [1]. Fukunaga and Koontz [5] proposed a normalization process, prior to applying principal component mapping, in order to extract the important features for separating two pattern classes. The normalization process provides eigenvectors which "best" fit one class and are the "poorest" for representing the other class. The fact that the best feature and the poorest one can be represented in one coordinate, motivates the effort to adopt this

normalization process for morphological data analysis. This property is convenient since we use the same structuring element for both classes in one stage of morphological operation. In [5] the transformation vectors are determined by the normalization and diagonalization process (described in the next section) of two autocorrelation matrices (NAD-CRM). Sample autocorrelation matrices are defined as follows :

$$R_j = \frac{1}{n_j-1}\sum_{i=1}^{n_j} x^{(j)}(i)\, x^{(j)}(i)^T = K_j + \overline{x}^{(j)}\overline{x}^{(j)^T} \quad (\,j = 1,2\,),\qquad (2\text{-}13)$$

where R_j, K_j, $x^{(j)}$, $\overline{x}^{(j)}$ and n_j denote sample autocorrelation matrix, sample covariance matrix, sample vector, sample mean, and number of samples in the j-th class respectively. It should be noted that the transformation vectors of the sample autocorrelation matrices are different from those of sample covariance matrices. The sample autocorrelation matrices are used to select features with respect to the discriminatory potential of both, the class sample means and the class sample variances. Accordingly this mapping will weaken the discriminatory power between class sample variances in exchange for adding the discriminatory power between class sample mean vectors. This makes the NAD-CRM mapping not optimal in the sense that the transformation vectors are not *principal axes*. Hence we propose NAD-CVM mapping whose transformation vectors are *principal axes*. The term *principal axes* refers here to the generalized eigenvalue problem (Equations (2-34) or (2-35)), while the term principal axes used in Section 2.1.2 refers to the standard eigenvalue problem. Eliminating the discriminating power of class sample means is justified in morphological analysis as explained in the introduction of this chapter.

2.2.2 Derivation of NAD-CVM Method

Let S_1 and S_2 be the weighted sample covariance matrices of Class 1 and Class 2, respectively :

$$S_i = \omega_i K_i \quad (i=1,2), \tag{2-14}$$

where

K_i is the (pxp) sample covariance matrix of Class i, and
ω_i is a weight of K_i, $\omega_i \geq 0$, $\omega_1 + \omega_2 = 1$.

Then the sample covariance matrix of the mixture of both classes S_0, can be written as :

$$S_0 = S_1 + S_2 . \tag{2-15}$$

The basic properties of normalization and simultaneous diagonalization are as follows :

Simultaneous diagonalization of the matrices S_0 and S_1 (or S_2) by the process of normalization of the matrix S_0 provides set of orthogonal eigenvectors which are also eigenvectors of S_2 (or S_1), and the sum of the eigenvalue matrices (as defined by Equation (2-6)) of the both classes is the identity matrix.

Sample covariance matrix is by definition symmetric. We assume that sample covariance matrix is positive definite without loss of generality, since the eigenvalues (which have interpretation of sample variances of principal components) equal to zero can be eliminated by eliminating the components which correspond to those eigenvalues. Thus S_0 is assumed to be positive definite and we can simultaneously diagonalize the matrices S_0 and S_1 by a congruence transformation matrix C, such that

$$C S_0 C^T = I \quad \text{and} \quad C S_1 C^T = \Lambda_1 , \tag{2-16}$$

where the matrices $\mathbf{I}$ and Λ_1 denote the identity matrix and a diagonal matrix respectively. Fukunaga and Koontz [5] showed that $\mathbf{S}_0$ and $\mathbf{S}_2$ can be also simultaneously diagonalized by the same matrix. $\mathbf{C}$, i.e. :

$$\mathbf{C} \, \mathbf{S}_2 \, \mathbf{C}^T = \Lambda_2 . \tag{2-17}$$

Furthermore, from Equations (2-15), (2-16), and (2-17),

$$\Lambda_1 + \Lambda_2 = \mathbf{I}. \tag{2-18}$$

The matrix $\mathbf{C}$, which satisfies Equations (2-16) and (2-17) can be determined by the following process:

Since $\mathbf{S}_0$ is positive definite, there exists a transformation matrix $\mathbf{P}$, such that

$$\mathbf{P} \, \mathbf{S}_0 \, \mathbf{P}^T = \mathbf{I} . \tag{2-19}$$

Let us define two matrices $\mathbf{Q}_1$ and $\mathbf{Q}_2$, as follows :

$$\mathbf{Q}_1 = \mathbf{P} \, \mathbf{S}_1 \, \mathbf{P}^T \text{ and } \mathbf{Q}_2 = \mathbf{P} \, \mathbf{S}_2 \, \mathbf{P}^T . \tag{2-20}$$

Then $\mathbf{Q}_1 + \mathbf{Q}_2 = \mathbf{I}$, and the eigenvalues and eigenvectors of $\mathbf{Q}_1$ are given by

$$\mathbf{Q}_1 \, \phi_j^{(1)} = \lambda_j^{(1)} \, \phi_j^{(1)} \quad (j=1,2,....,p), \tag{2-21}$$

where the superscript (1) indicates Class 1. Assuming that the eigenvalues are in the descending order, i.e.,

$$\lambda_1^{(1)} \geq \lambda_2^{(1)} \cdots \geq \lambda_p^{(1)} > 0. \tag{2-22}$$

Then the eigenvalues and eigenvectors for the Class 2 are

$$\mathbf{Q}_2 \, \phi_j^{(2)} = (\mathbf{I} - \mathbf{Q}_1) \, \phi_j^{(2)} = \lambda_j^{(2)} \, \phi_j^{(2)} . \tag{2-23}$$

From Equation (2-23) we have,

$$\mathcal{S}_1 \, \phi_j^{(2)} = (1 - \lambda_j^{(2)}) \, \phi_j^{(2)} \, . \tag{2-24}$$

Thus from Equations (2-21) and (2-24),

$$\phi_j^{(2)} = \phi_j^{(1)} \quad \text{for } j=1,2,\cdots,p, \tag{2-25}$$

and

$$\lambda_j^{(1)} = 1 - \lambda_j^{(2)} \quad \text{for } j=1,2,\cdots,p. \tag{2-26}$$

Equations (2-25) and (2-26) show that the eigenvector corresponding to the largest eigenvalue for Class 1 has the smallest eigenvalue for Class 2, and so on, i.e.,

$$1 > \lambda_1^{(1)} \geq \lambda_2^{(1)} \geq \cdots \geq \lambda_p^{(1)} > 0 \text{ for Class 1}, \tag{2-27}$$

and

$$1 > \lambda_p^{(2)} = 1 - \lambda_p^{(1)} \geq \cdots \geq \lambda_1^{(2)} = 1 - \lambda_1^{(1)} > 0 \text{ for Class 2}. \tag{2-28}$$

Let Φ be a (pxp) matrix composed of the eigenvectors ϕ_j ,i.e.,

$$\Phi = [\phi_1 | \phi_2 | \cdots | \phi_p].$$

Then the matrix

$$C = \Phi^T P \tag{2-29}$$

satisfies Equations (2-16), (2-17), and (2-18).

2.2.3 Useful Properties of NAD-CVM Method

(1) Sample variance interpretation of an eigenvalue

Let us consider the following transformation using matrix, $\mathbf{C}$ defined by Equation (2-29) :

$$\mathbf{y}^{(j)}(i) = \mathbf{C}^T \mathbf{x}^{(j)}(i), \quad j = 1,2 \; ; \; i=1,2,\cdots,n_1 \text{ (for Class 1) or } i=1,2,\cdots,n_2$$

(for Class 2)   (2-30)

where the vectors $\mathbf{x}$ and $\mathbf{y}$ are p-dimensional vectors, and $\mathbf{C} = [\mathbf{c}_1 | \mathbf{c}_2 | \cdots | \mathbf{c}_p]$. Let $\mathbf{y}^{(j)} = [y_1^{(j)} y_2^{(j)} \cdots y_p^{(j)}]^T$. Then the sample variance of $y_k^{(j)}$ for the Class j can be written as

$$\frac{1}{n_j-1}\sum_{i=1}^{n_j}(y_k^{(j)}(i)-\overline{y_k^{(j)}})^2 = \lambda_k^{(j)} \; , \; j=1,2. \qquad (2-31)$$

Since $\lambda_k^{(1)}$ satisfies $0 < \lambda_k^{(1)} < 1$, if we select the eigenvector $\mathbf{c}_1$ as the transformation vector, we get the maximum ratio of the sample variance of Class 1 to that of Class 2, i.e., $\lambda_1^{(1)}/1 - \lambda_1^{(1)}$. Consequently this transformation yields the best component for the Class 1 and the poorest for the Class 2 in the sense of the sample variances ratio. On the other hand, the transformation by the eigenvector $\mathbf{c}_p$ provides the maximum ratio of $\lambda_p^{(1)}/1 - \lambda_p^{(1)}$, and is the best component for the Class 2 and the poorest for the Class 1. Therefore the best selection of the transformation vectors is $\mathbf{c}_1$, $\mathbf{c}_2$, $\cdots$ for Class 1, $\mathbf{c}_p$, $\mathbf{c}_{p-1}$, $\cdots$ for Class 2, and $\mathbf{c}_1$, $\mathbf{c}_p$, $\mathbf{c}_2$, $\mathbf{c}_{p-1}$, $\cdots$ for both Classes 1 and 2.

(2) 2-dimensional geometric interpretation

The following four figures (Figure 2-2a - 2-2d) show the process of NAD-CVM mapping using the ellipsoid interpretation.

$K_1 = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$, $K_2 = \begin{bmatrix} 2.5 & -1.5 \\ -1.5 & 2.5 \end{bmatrix}$, and $\omega_1 = \omega_2 = 0.5$ are used. To show the interpretation of sample variances, sample means of each class are set at the origin.

(3) Unsupervised NAD-CVM algorithm

In general, we do not know *a priori* K_i and ω_i as used in Equation (2-14). Here we remove the assumption about knowing ω_i, and refer to such situation as an *unsupervised NAD-CVM algorithm*. The following lemmas provide characterization of the unsupervised NAD-CVM algorithm.

Lemma 2.1

Assume that K_2 is known, and K_0 (the sample covariance matrix of the whole image) can be calculated. Then the transformation matrix $\Phi^T P$ is given by

$$P K_0 P^T = I, \tag{2-32}$$

$$\Phi^T P K_2 P^T \Phi = \Lambda, \tag{2-33}$$

where $\Phi^T \Phi = I$, and Λ is a diagonal matrix whose diagonal elements are the eigenvalues of the matrix $P K_2 P^T$.

Proof.

Since $S_0 = K_0$, the matrix P in Equation (2-32) is the same as the matrix P in Equation (2-19). The matrices PK_2P^T and $P(\omega_2 K_2)P^T$ share the same normalized eigenvectors satisfying : $\Phi^T \Phi = I$. Therefore the transformation matrix $\Phi^T P$ is identical to the matrix defined by Equation (2-29). ∎
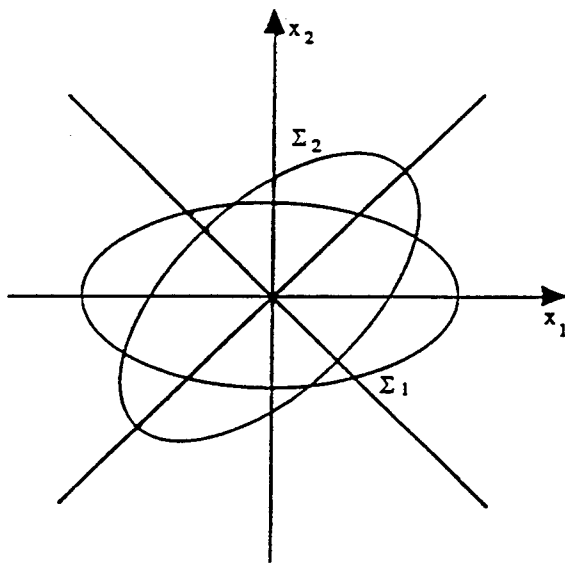
Figure 2-2a Ellipsoid interpretation of K_1 and K_2 .
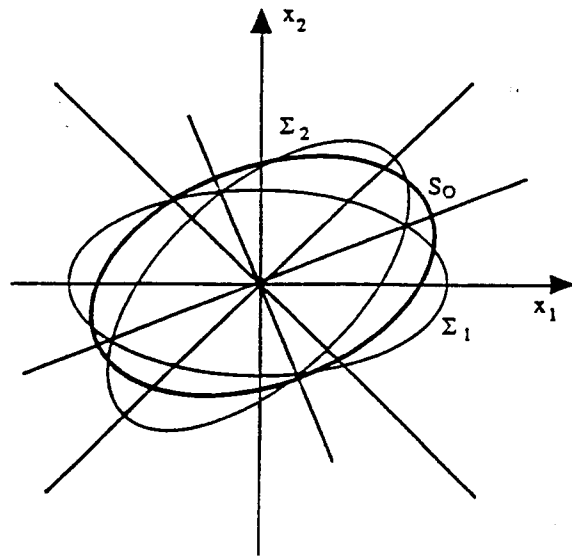
Figure 2-2b Generation of S_0 .
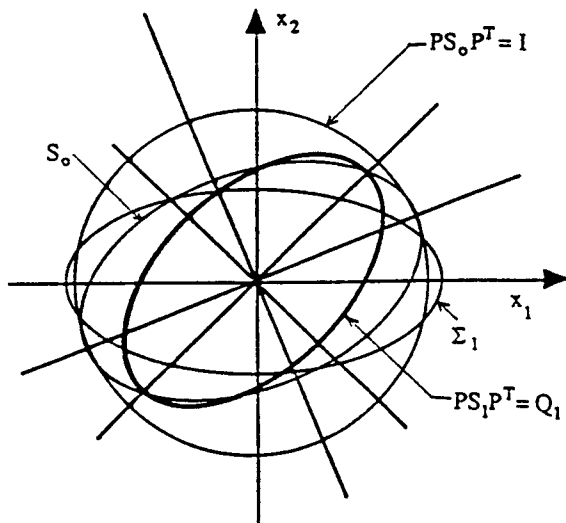
Figure 2-2c Generation of Q_1 .
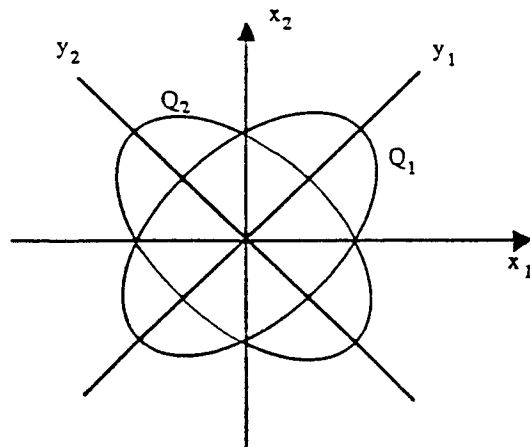
Figure 2-2d Relationship between Q_1 and Q_2 .

Lemma 2.2

Let K_1 and K_2 be symmetric and positive definite. Then the transformation matrix $\Phi^T P$ is invariant with respect to ω_1, ω_2, where $\omega_1 > 0$, $\omega_2 > 0$, and $\omega_1 + \omega_2 = 1$.

Proof.

Let $K_0 = \omega_1 K_1 + \omega_2 K_2$, $\omega_1, \omega_2 > 0$ and $\omega_1 + \omega_2 = 1$. Then the problem of finding the transformation matrix $C = \Phi^T P$ is equivalent to the solution of the following generalized eigenvalue problem :

$$(\omega_1 K_1) \, c_i = \lambda_i K_0 \, c_i, \tag{2-34}$$

or

$$(\omega_2 K_2) \, c_i = \lambda_i K_0 \, c_i \tag{2-35}$$

where the vector c_i is a generalized eigenvector corresponding to the eigenvalue λ_i, and forms the i-th column of the matrix C. We have proven in Section 2.2.2 that Equations (2-34) and (2-35) share the same eigenvectors. It remains to show that the two generalized eigenvalue problems, $(\omega_1 K_1) \, c_i = \lambda_i K_0 \, c_i$ and $(\omega_1' K_1) \, c_i' = \lambda_i' K_0 \, c_i'$, share the same eigenvectors, i.e., $c_i = c_i'$ for any i=1,2,···,p, $1 > \omega_1' > 0$, and $K_0' = \omega_1' K_1 + (1 - \omega_1') K_2$.

Since K_0 is positive definite, the generalized eigenvalue problem is equivalent to the standard eigenvalue problem ($K_0^{-1} \omega_1 K_1$)$c_i = \lambda \, c_i$. Let ω_1' be any positive real number such that $\omega_1' \neq \omega_1$, and $\omega_1' < 1$, and let $\omega_2' = 1 - \omega_1' > 0$. To show that the matrices $(K_0^{-1} \omega_1 K_1)$ and $(K_0'^{-1} \omega_1' K_1)$ have identical eigenvectors, it suffices to show that these two matrices commute [32, p.193]. To show that

$$(K_0^{-1} \omega_1 K_1)(K_0'^{-1} \omega_1' K_1) = (K_0'^{-1} \omega_1' K_1)(K_0^{-1} \omega_1 K_1),$$

we need to prove that $K_0^{-1} K_1 K_0'^{-1} = K_0'^{-1} K_1 K_0^{-1}$. Since K_1 is positive definite, ω_1 and ω_1' are scalars, and K_0, K_0', K_1 are invertible, it suffices to show that $K_0' K_1^{-1} K_0 = K_0 K_1^{-1} K_0'$. Replacing K_0 with $\omega_1 K_1 + \omega_2 K_2$ and K_0' with $\omega_1' K_1 + \omega_2' K_2$ we have :

$$\mathbf{K_0'K_1^{-1}K_0} = (\omega_1'\mathbf{K_1} +\omega_2'\mathbf{K_2})\mathbf{K_1}^{-1} (\omega_1\mathbf{K_1} +\omega_2\mathbf{K_2})$$
$$= \omega_1'\omega_1\mathbf{K_1} + \omega_2'\omega_1\mathbf{K_2} + \omega_1'\omega_2\mathbf{K_2} + \omega_2'\omega_2\mathbf{K_2K_1}^{-1}\mathbf{K_2} ,$$

$$(2\text{-}36)$$

and

$$\mathbf{K_0K_1^{-1}K_0} = (\omega_1\mathbf{K_1} +\omega_2\mathbf{K_2})\mathbf{K_1}^{-1}(\omega_1'\mathbf{K_1} +\omega_2'\mathbf{K_2})$$
$$= \omega_1'\omega_1\mathbf{K_1} + \omega_2'\omega_1\mathbf{K_2} + \omega_1'\omega_2\mathbf{K_2} + \omega_2'\omega_2\mathbf{K_2K_1}^{-1}\mathbf{K_2} ,$$

$$(2\text{-}37)$$

which proves the desired equality. ∎

Lemma 2.3

Let ω_1 and ω_1' be such that

$$\mathbf{K_0} = \omega_1\mathbf{K_1} +\omega_2\mathbf{K_2}$$

and

$$\mathbf{K_0'} = \omega_1'\mathbf{K_1} +\omega_2'\mathbf{K_2} .$$

Consider the generalized eigenvalue problems :

$$(\omega_1\mathbf{K_1}) \mathbf{c_i} = \lambda_i\mathbf{K_0} \mathbf{c_i} \text{ and } (\omega_1'\mathbf{K_1})\mathbf{c_i} = \lambda_i'\mathbf{K_0} \mathbf{c_i} \text{ ,i=1,2,} \cdots \text{,p.}$$

Let $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_p > 0$.

If $\omega_1 < \omega_1'$, then $\lambda_i < \lambda_i'$ for all i.

Proof.

For ω_1 we can rewrite the generalized eigenvalue problem as

$$(\mathbf{K_0}^{-1}\omega_1\mathbf{K_1})\mathbf{c_i} =\lambda_i \mathbf{c_i}.$$

Since $\mathbf{K_0}$ and $\mathbf{K_1}$ are both positive definite, we have :

$$(\mathbf{K_1}^{-1}\omega_1^{-1}\mathbf{K_0})\mathbf{c_i} = (1/\lambda_i) \mathbf{c_i}.$$

Replacing $\mathbf{K_0}$ with $\omega_1\mathbf{K_1} +\omega_2\mathbf{K_2}$ we have :

$$(\mathbf{I} + ((1- \omega_1)/ \omega_1)\mathbf{K_1}^{-1}\mathbf{K_2})\mathbf{c_i} = (1/\lambda_i) \mathbf{c_i} ,$$

which gives

$$(\mathbf{K_1}^{-1}\mathbf{K_2})\mathbf{c_i} = (\omega_1/1- \omega_1)((1/\lambda_i)-1) \mathbf{c_i} .$$

$$(2\text{-}38)$$

For ω_1' the generalized eigenvalue problem can be rewritten as

$$(\mathbf{I} + ((1- \omega_1')/ \omega_1')\mathbf{K_1}^{-1}\mathbf{K_2})\mathbf{c_i} = (1/\lambda_i') \mathbf{c_i} .$$

$$(2\text{-}39)$$

Eliminating $(\mathbf{K_1}^{-1}\mathbf{K_2})\mathbf{c_i}$ from Equation (2-38), (2-39) yields :

$$\mathbf{c_i} + (\frac{1-\omega_1'}{\omega_1'}) + (\frac{\omega_1}{1-\omega_1})(\frac{1}{\lambda_i} - 1)\mathbf{c_i} = [1+(\frac{1-\omega_1'}{\omega_1'})+(\frac{\omega_1}{1-\omega_1})(\frac{1-\lambda_i}{\lambda_i})]\mathbf{c_i} .$$

Thus $\lambda_i' = (1 + \dfrac{1-\omega_1'}{\omega_1'} \dfrac{\omega_1}{1-\omega_1} \dfrac{1-\lambda_i}{\lambda_i})^{-1} = \dfrac{\omega_1'(1-\omega_1)\lambda_i}{\omega_1'(1-\omega_1)\lambda_i + (1-\omega_1')\omega_1(1-\lambda_i)}$

and

$$\lambda_i' - \lambda_i = \frac{\omega_1'(1-\omega_1)\lambda_i - \omega_1'(1-\omega_1)\lambda_i^2 - (1-\omega_1')\omega_1(1-\lambda_i)\lambda_i}{\omega_1'(1-\omega_1)\lambda_i + (1-\omega_1')\omega_1(1-\lambda_i)}$$

$$= \frac{\lambda_i(1-\lambda_i)(\omega_1'-\omega_1)}{\omega_1'(1-\omega_1)\lambda_i + (1-\omega_1')\omega_1(1-\lambda_i)} .$$

Since $0 < \lambda_1 < 1$, $\omega_1 > 0$, and $\omega_1' > 0$, the denominator of the above equation is positive. The numerator is positive iff $\omega_1' > \omega_1$ which shows that $\lambda_i' > \lambda_i$. ∎

Lemma 2.1 can be used constructively when the sample covariance matrix of the background is known *a priori*. Lemma 2.2 implies that we do not need to calculate or estimate ω_1 and ω_2 as long as we know the sample covariance matrices K_1 and K_2 . However, according to Lemma 2.3 the eigenvalues depend on ω_i. Together, the above results show the feasibility of NAD-CVM implementation. An application of these properties to defect detection problem will be discussed in Chapter 4 .

2.2.4 NAD-CVM as a Preprocessing for Multi-variate Image Analysis

This section relates the mapping techniques discussed in this chapter to multi-variate morphology discussed in the sequel. The application of mapping methods to multi-variate morphology represents a novel approach. The mapping techniques for image analysis [21], [19] take advantage of the dimensionality reduction property of the eigenvector expansion. This approach may not be the best for image processing, since sample variance may not yield a good

measure of image quality. We have already introduced *the best and the poorest* NAD-CVM mappings in the sense that the resulting sample variance corresponds to a morphological measure. Such a measure can be directly related to the properties of morphological filtering. These concepts are further discussed in Section 3.3.

We note that NAD technique cannot be extended to a general multi-class problem. However, NAD-CVM mapping solves the segmentation problem , which is one of the three basic techniques in image analysis. Image segmentation is the most important in the sense that most image processing techniques, including morphological operations, are developed to provide a good segmentation results. Segmentation problems can be generally interpreted as the two class problems, i.e., segmentation of the object and the background. The term *segmentation* includes feature extraction techniques used for preprocessing.

The two class separation property of NAD-CVM, the morphological measure interpretation of sample variance, and the unsupervised implementation of NAD-CVM algorithm justify the effort of applying NAD-CVM mapping to preprocessing of multi-variate signals.

3. MATHEMATICAL MORPHOLOGY FOR MULTI-VARIATE SIGNALS

There are numerous applications of morphological filters in image processing, analysis, and related areas. The references for these applications can be found in [16]. Recently the multi-variate signal analysis and processing, mainly for color and multi-spectral images, have attracted a significant attention [21],[33]-[42]. Here we present a new approach to apply mathematical morphology to multi-variate signals.

First, we briefly review morphological operations using set-theoretical definitions. Next, an extension to grayscale morphology for multilevel signals using the concepts of *umbra* and *top surface* is discussed. Using Serra[22], Haralick *et al*[9], and Maragos[16] as the basic references, we review only essential definitions and properties of morphological operations, especially those which are used in the sequel. The morphological operations developed for single-component (uni-variate) signal (refered as *uni-variate morphology*) are extended to multi-variate signals using two different methods :

(1) separate morphological operations are applied to each signal component.

(2) *multi-variate morphology* concepts are developed using partial-ordering in finite dimensional spaces.

Since separate morphological operations are easy to define, but not effective when applied to original multi-variate signal components, and multi-variate morphology is meeting serious theoretical difficulties, we propose separate morphological operation for each transformed signal component. The transformed components are obtained from the original signal by a special mapping method. This approach is well founded by introducing a morphological measure related to the sample variance, and using it as a criterion to optimize the mapping technique.

3.1 Review of Mathematical Morphology

Mathematical morphology was developed in 1964 at the Paris School of Mines (France) by G. Matheron and J. Serra, who were asked to investigate the relationships between the geometry of porous media and their permeabilities and to quantify the petrography of iron ores in order to predict their milling properties [22]. By probing and transforming a geometric structure with different patterns of predefined shapes, or structuring elements, Matheron and Serra developed a number of different techniques. They called their geometric transformations "morphology", meaning the "study of forms." In order to avoid tying the results of morphological transformations to the specific application, they derived four mathematical quantification constraints with four corresponding principles [22, pp.6-15], or the method of "mathematical morphology."

Originally, Matheron and Serra represented image objects and structuring elements by sets in a Euclidean space. Hence, the morphological operations are actually set operations based on unions and intersections. The simplest morphological operations are erosion, dilation, opening, and closing, all of which grew out of Minkowski addition and subtraction. Since binary images can be represented by two-dimensional sets, Matheron [18] and Serra [22] applied the Minkowski set operations to image analysis using the concepts of mathematical morphology. Such an image analysis technique is called morphological image analysis or morphological filtering.

Binary signal analysis was extended to multilevel signals by Serra [22] and Sternberg [24,25]. Serra used signal cross sections to generalize the morphological operations of multilevel signals. Sternberg further generalized morphological transformations for multilevel signals by considering graytone images as surfaces of three-dimensional volumes and introducing the concept of *umbra*.

In this section a brief summary of basic morphological operators (erosion and dilation) and their secondary operators (opening and closing) is presented. Since mathematical morphology is based on the set-theoretical method, the discussion begins with binary signals

viewed as a set. Next this set-theoretical method is extended to multi-level signals.

There are several slightly different definitions of morphological operations found in the relevant literature, [22], [16], [9], and [8]. Here we follow Serra's definitions [22].

3.1.1 Set Processing Operators

An n-dimensional binary signal can be represented as a set in an n-dimensional Euclidean space ($\mathbf{R}^n$), or its discrete equivalent, the set of n-tuples of integers $\mathbf{Z}^n$, depending on whether the function is sampled or not. Serra [22] presents a complete analysis of discrete version of morphology. Here we do not distinguish between the continuous space ($\mathbf{R}^n$) and the discrete space ($\mathbf{Z}^n$) using common symbol $\mathbf{E}^n$ for both.

3.1.1.1 Erosion and Dilation

Let two sets A and B be subsets of $\mathbf{E}^n$. Minkowski set addition $A \oplus B$, of two sets A and B, consists of all points that can be expressed as a vector addition, a+b, where the vector a and b belong to the sets A and B respectively, i.e. :

$$A \oplus B = \{a+b : a \in A, b \in B\}.$$

Also

$$A \oplus B = \bigcup_{b \in B} A_b = \bigcup_{a \in A} B_a = \{z \in \mathbf{E}^n : A \cap (B^s)_z \neq \varnothing\}. \qquad (3\text{-}1)$$

where A_b denotes the translation of A by the vector b and is equivalent to $A \oplus \{b\}$. B^s denotes the symmetric set of B with respect to the origin, i.e., $B^s = \{-b : b \in B\}$.

<u>Dilation</u>

The dilation of A by B is the set $\{z \in \mathbf{E}^n : A \cap B_z \neq \varnothing\}$ containing the points z such that the translation B_z intersects A. Therefore the dilation of A by B is equal to $A \oplus B^s$.

Minkowski set subtraction of A by B, denoted by $A \ominus B$, is defined indirectly as the operational "dual" to Minkowski set addition with respect to the complement :

$$A \ominus B = (A^c \oplus B)^c = \underset{b \in B}{\cap} A_b = \{z \in E^n : (B^s)_z \subset A\}. \qquad (3\text{-}2)$$

Erosion

The erosion of A by B is defined by the set $\{z \in E^n : (B)_z \subset A\}$ containing the points z such that the translation B_z is included in A. This definition is equivalent to the operation "dual" to dilation, i.e., $(A^c \oplus B^s)^c$. Using the relationship (3-2), erosion is equal to $A \ominus B^s$.

3.1.1.2 Opening and Closing

By combining erosion and dilation, opening and closing operators are obtained. The opening $A \circ B$ of A by B and closing $A \bullet B$ of A by B are defined as follows :

$$A \circ B = (A \ominus B^s) \oplus B ,$$
$$A \bullet B = (A \oplus B^s) \ominus B . \qquad (3\text{-}3)$$

Since the opening of A by B is the union of the translations,(i.e., $A \circ B = \cup \{B_y : y \in E^n, B_y \subset A\}$, [18, p.19]), the output $A \circ B$ appears to be the result of fitting B around the inside perimeter of A. From this interpretation we can characterize opening as having a *fitting* property. Fig.3-1 shows an example of opening A by a non-symmetric structuring element B. On the other hand, the output of $A \bullet B$ appears to be the result of fitting B around the outside perimeter of A. This inside-outside duality between the opening and closing indicates that the complement of closing is equal to the opening of complement [8].

The morphological opening and closing defined by Equation (3-3) have the following three properties :

(1) Anti-extensivity for opening and extensivity for closing :

$$A \bigcirc B \subset A \subset A \bullet B .$$
<div align="right">(3-4)</div>

(2) Increasing Property :

$$A \subset A' \dashrightarrow A \bigcirc B \subset A' \bigcirc B ,$$
$$A \subset A' \dashrightarrow A \bullet B \subset A' \bullet B .$$
<div align="right">(3-5)</div>

(3) Idempotence :

$$(A \bigcirc B) \bigcirc B = A \bigcirc B ,$$
$$(A \bullet B) \bullet B = A \bullet B .$$
<div align="right">(3-6)</div>
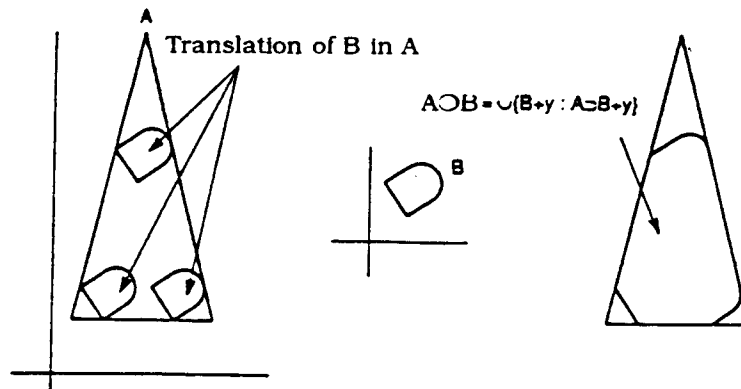
Figure 3-1 Opening by fitting. (Giardina and Dourgherty [8])

3.1.2 Relationship between Sets and Functions

To extend all the morphological set transformations to functions, certain relations between sets and functions must be established. Therefore, representing functions as sets is the main issue. There are two different, but equivalent, approaches to this problem: An n-dimensional (n-D) function can be represented either by an ensemble of n-D sets called its cross sections [22] or a single (n+1)-D set called its umbra. Due to its simple interpretation, the umbra representation of a function is generally used. We discuss here

the concept of umbra, since it can be used to generalize grayscale morphology to multi-variate signals.

In order to visualize the umbra concept, it is enough to observe that any function can be viewed as a 2-D image, without loss of generality. Umbra simply means shadow cast by the function "surface", and the umbra of a set B, in 3-D includes both B and the points of its shadow.
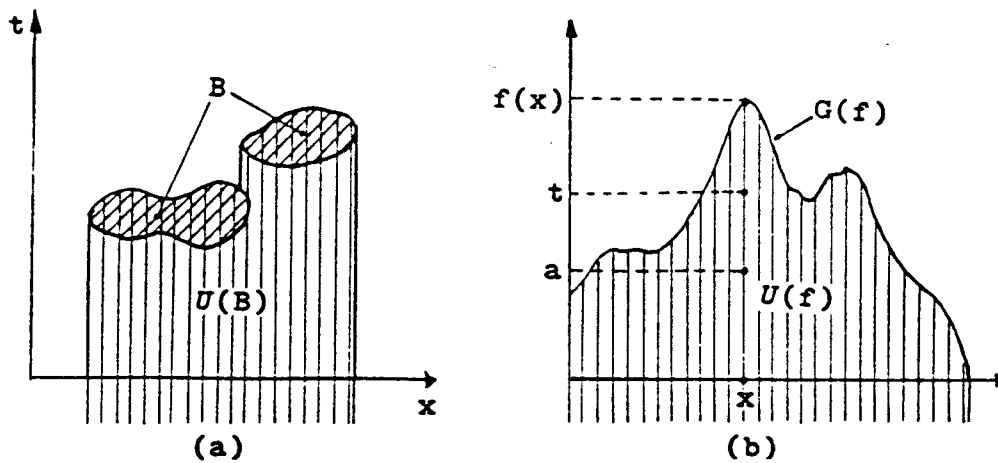
Figure 3-2 (a) Umbra of a set B, (b) umbra of a function.

The shadow is cast by a point light source at an infinite distance in the positive t-direction. Figure 3-2a illustrates this concept. Analytically, the umbra of set B can be expressed as a morphological transformation of B and a structuring element T, composed of the points belonging to the t-axis, including the origin. Using Minkowski sum $\oplus$ definition, the umbra of a closed set B equals to $B \oplus (-\infty, 0]$. This relation corresponds to $B \oplus [0, +\infty)^S$, i.e., dilation of B by the positive t-axis. Similarly, the umbra of a function f, is the Minkowski sum of the function and $(-\infty, 0]$. Thus the umbra of f, denoted by $U(f)$, is the set:

$$U(f) = \{(x, t) : f(x) \geq t\}. \tag{3-7}$$

Serra [22] maintained that mapping between an upper semi-continuous (u.s.c.) function f and $U(f)$ is one-to-one. Maragos [16] mathematically formalized this correspondence.

3.1.2.1 Theorem on One-to-One Correspondence between Signal and its Umbra

To any real valued u.s.c. function, $f(x)$, $x \in \mathbf{R}^n$, corresponds an unique umbra $U(f)$, which is a closed set in $\mathbf{R}^n \mathbf{x} \mathbf{R}$ such that

$$(x,t) \in U(f) \quad \leftrightarrow \quad t \leq f(x) . \qquad (3\text{-}8)$$

and

$$(x,t) \in U(f) \quad \dashrightarrow \quad (x,a) \in U , \text{ for all } a < t . \qquad (3\text{-}9)$$

Conversely, to any closed subset U of $\mathbf{R}^n \mathbf{x} \mathbf{R}$ satisfying relation (3-9), corresponds a unique u.s.c. function $f(x)$, which can be reconstructed from U as follows :

for $x \in \mathbf{R}^n$,

$$f(x) = \sup \{t \in \mathbf{R} : (x,t) \in U \} . \qquad (3\text{-}10)$$

Note that the case where $f(x)$ is defined on $\mathbf{Z}^n$ is trivial as far as the u.s.c. is concerned, because any sampled function is both u.s.c. and l.s.c. (lower semi-continuous). Hence Equations (3-8), (3-9), and (3-10) are also true for sampled functions, $f(x)$, $x \in \mathbf{Z}^n$. The supremum operation in Equation (3-10) is called *top surface* of U, and is denoted by

$$T[U] (x) = \sup \{t \in \mathbf{R} : (x,t) \in U \} . \qquad (3\text{-}11)$$

3.1.3 Function Processing Operators

Set-theoretic definitions of morphological operators can be extended to the case of multilevel signal and multilevel structuring element using the umbra.

Let us define function processing dilation and erosion using the umbra and the top surface.

<u>Definition 3-1</u> Let $f : F \subseteq \mathbf{E}^n \dashrightarrow \mathbf{E}$ and $g : G \subseteq \mathbf{E}^m \dashrightarrow \mathbf{E}$ ($m \leq n$). The dilation of f by g is denoted by $f \oplus g$, and is defined as

$$f \oplus g = T[U(f) \oplus U(g)]. \qquad (3\text{-}12)$$

The erosion of f by g is denoted by $f \ominus g$, and is defined as

$$f \ominus g = T[U(f) \ominus U^r(g)],$$ (3-13)

where $U^r(g)$ denotes the reflected set of $U(g)$,i.e.,

$$U^r = \{(x,t) : (x,-t) \in U\} .$$

Using Equations (3-1) and (3-2), Equations (3-12) and (3-13) can be rewritten as

$$(f \oplus g)(x) = \sup\{t \in \mathbf{R} : U^s(g) \cap U(f) \neq \varnothing \},$$ (3-14)

and

$$(f \ominus g)(x) = \sup\{t \in \mathbf{R} : [U^r(g)]^s \subset U(f)\},$$ (3-15)

where U^s denotes the transposed set of U, i.e.,

$$U^s = \{(x,t) : (-x,-t) \in U\},$$

and

$$[U^r]^s = \{(x,t) : (-x,t) \in U\}.$$

Alternatively, Equations (3-14) and (3-15) can be written in an algebraic form as follows:

$$(f \oplus g)(x) = \sup_{z \in \mathbf{E}^n} \{f(z)+g(x-z)\},$$ (3-16)

and

$$(f \ominus g)(x) = \inf_{z \in \mathbf{E}^n} \{f(z)-g(x-z)\},$$ (3-17)

where $f(x) = -\infty$, for $x \notin F$, and $g(x) = -\infty$, for $x \notin G$.

Function processing opening and closing are defined using function processing dilation and erosion as follows:

$$\text{opening} : f \circ g = (f \ominus g^s) \oplus g,$$ (3-18)

$$\text{closing} : f \bullet g = (f \oplus g^s) \ominus g,$$ (3-19)

where g^s denotes the transposition of function g such that $g^s(x) = g(-x)$.

3.2 Extension of Grayscale Morphology to Multi-variate Signals

In this section, we discuss a possible extension of the morphological concepts to analyze multi-variate signals. There are two possible approaches to such an extension. The first approach is to apply morphological operations to each signal component separately

and combine the results [2]. The second approach is to define an umbra of multi-variate function and to generalize definitions of set-theoretical operations. In [26] an attempt is made to extend the current concept of mathematical morphology (called scalar morphology) to a matrix morphology formalism. However the results of [26] are based on the application of separate morphological operations to each signal component. Here we discuss a possible extension of grayscale morphology to *multi-variate morphology* by modifying the definition of umbra.

3.2.1 Separate Component Morphological Operations

Let us define a multi-variate signals **f** and **g** , as follows:

$$\mathbf{f} : \mathbf{E}^n \text{ --> } \mathbf{E}^p \text{ , and } \mathbf{g} : \mathbf{E}^m \text{ --> } \mathbf{E}^p ,\qquad (3\text{-}20)$$

where **f**= $[f_1 \ f_2 \ ... \ f_p]^T$, **g**= $[g_1 \ g_2 \ ... \ g_p]^T$, $m \le n$. Separate morphological operations of each signal component of a multi-variate signal can be introduced using the following definitions of umbra and top surface.

<u>Definition 3-2</u> The vector umbra of **f**, denoted by $U_s(\mathbf{f})$, is defined by

$$U_s(\mathbf{f}) = [\ U(f_1) \ U(f_2) \ ... \ U(f_p) \ ]^T,\qquad (3\text{-}21)$$

where the notation U follows Equation (3-8).

<u>Definition 3-3</u> The top surface of $U_s(\mathbf{f})$, denoted by T_s, is defined by

$$T_s[U_s(\mathbf{f})] = [\ T(U(f_1)) \ T(U(f_2)) \ ... \ T(U(f_p)) \ ]^T,\qquad (3\text{-}22)$$

where the notation T follows Equation (3-11).

<u>Definition 3-4</u> Let $\star$ denote any morphological operation, i.e., dilation, erosion, opening, or closing. Then the separate component morphological operation is defined by

$$\mathbf{f}\star\mathbf{g} = T_s[U_s(\mathbf{f}) \star U_s(\mathbf{g})] \text{ , and}\qquad (3\text{-}23)$$

$$U_s(\mathbf{f}) \star U_s(\mathbf{g}) = [\ U(f_1)\star U(g_1) \ \ U(f_2)\star U(g_2) \ \cdots \ U(f_p)\star U(g_p) \ ]^T.\qquad (3\text{-}24)$$

Alternatively, Equations (3-23) and (3-24) can be rewritten as

$$\mathbf{f}\star\mathbf{g} = [\ f_1\star g_1 \ \ f_2\star g_2 \ ... \ .f_p\star g_p \ ]^T.\qquad (3\text{-}25)$$

We note that Equation (3-25) allows to define different morphological operations for each signal component. Definitions 3-2, 3-3,and 3-4 are

special cases which relate function processing operations to set processing operations.

Although the separate morphological operations are useful for analyzing multi-variate signal in certain circumstances [2], [26], they may extract less information than the grayscale morphological operations applied to a single component which is obtained by the Karhunen-Loeve, NAD-CVM, or NAD-CRM transformation.

3.2.2 Generalization of Umbra Concept

Definitions 3-2 and 3-3 define separate morphological operations. The following definition of umbra is proposed as a generalization of the umbra concept to multi-variate signals.

Definition 3-5 (Umbra of function $f : E^n \to E^p$)

$$U(f) = \{(x,t) : f(x) \succeq t\} \text{ such that } U(f) \subset E^{n+p}, \qquad (3\text{-}26)$$

where the vector inequality symbol $\succeq$, means that every component of $f(x)$ is greater or equal to the corresponding component of t. ($\succeq$ is a partial-ordering operator.)

The above definition of umbra is equivalent to the successive dilation of the graph $G(f)$ of f, by the positive axis $[0,+\infty]$ of the t_i axis, $i=1,2,\cdots,p$, i.e.,

$$U(f) = G(f) \oplus (-\infty,0]_{t_1} \cdots \oplus (-\infty,0]_{t_p} ,$$

where $(-\infty,0]_{t_i}$ denotes the negative axis of t_i .

Next we need to define the top surface to reconstruct a vector function from the umbra defined in 3-5. A suitable top surface can be defined by using the following *pseudo-supremum* operation (denoted by *psup*) on vector space.

Definition 3-6 (Top surface of the umbra of multi-variate function)

$$T[U(f)] (x) = psup\{t \in E^p : (x,t) \in U(f) \} , \qquad (3\text{-}27)$$

where *psup* operation is defined as follows :

<u>Definition 3-7</u> (Pseudo-supremum on vector space)
Let $B \subset \mathbf{E}^p$, then

$$psup(B) = \{t \in \mathbf{E}^p : t \not\succcurlyeq b \text{ for all } b \in B \text{ and } z \not\succcurlyeq t, z \in \mathbf{E}^p$$
$$\text{whenever } z \not\succcurlyeq b \text{ for all } b \in B\}. \qquad (3\text{-}28)$$

The following example shows an illustration of Definitions 3-5, 3-6, and 3-7 for p=2.

<u>Example 3-1</u> Consider a space $\mathbf{Z}x\mathbf{R}^2$. Define sets A and $B \subset \mathbf{Z}x\mathbf{R}^2$ as

$$A=\{(z_1,a_1,a_2) \in \mathbf{Z}x\mathbf{R}^2 : [2\ 3]^T \not\succcurlyeq [a_1,a_2]^T \text{ for a given } z_1 \in \mathbf{Z}\},$$

and

$$B=\{(z_2,a_1,a_2) \in \mathbf{Z}x\mathbf{R}^2 : [4\ 1]^T \not\succcurlyeq [a_1,a_2]^T \text{ for a given } z_2 \in \mathbf{Z}\},$$

where the sets A and B are sections of the umbra of a multi-variate function $f : \mathbf{Z} \to \mathbf{R}^2$ at $z = z_1$ and $z = z_2$, respectively, $f(z_1)=[2\ 3]^T$ and $f(z_2)=[4\ 1]^T$.

Let $\mathcal{A} = \{(a_1,a_2) : (x,a_1,a_2) \in A\}$, and $\mathcal{B} = \{(a_1,a_2) : (x,a_1,a_2) \in B\}$.
Then $psup\ (\mathcal{A} \cap \mathcal{B}) = (2,1)$, and $\qquad\qquad (3\text{-}29)$

$$psup\ (\mathcal{A} \cup \mathcal{B}) = (\max(2,4),\ \max(3,1)) = (4,3) \qquad (3\text{-}30)$$

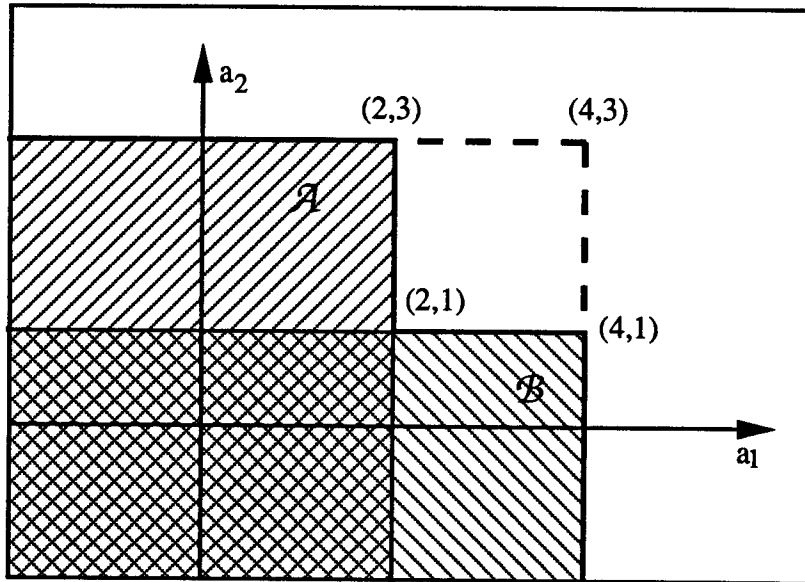Figure 3-3 shows a geometrical illustration of this example.

Figure 3-3 Umbra and top surface in vector space.

Using Definitions 3-5 to 3-7, we can define *multi-variate morphology* as follows.

<u>Definition 3-8</u> (*Multi-variate morphology*)

Let **f**, **g** and ★ be given by Equations (3-20) and (3-23). Then

$$\mathbf{f} ★ \mathbf{g} = T[U(\mathbf{f}) ★ U(\mathbf{g})] ,$$

(3-31)

where U and T are given by Equations (3-26) and (3-27) respectively.

Continuing with Example 3-1, let us assume that $\mathbf{f}(z_1)=[2\ 3]^T$ and $\mathbf{f}(z_1+1)=[4\ 1]^T$ for a given z_1. Then for a structuring element C={0 -1}, Equation (3-29) is equivalent to $(\mathbf{f} \ominus C)(z_1)$ and Equation (3-30) is equivalent to $(\mathbf{f} \oplus C)(z_1)$.

If we use Definitions 3-5, 3-6, and 3-7 for the umbra and top surface of multi-variate signal, Equation (3-31) results in separate morphological operations, and is equivalent to Equation (3-25). However the following problem may arise :

$$U(\mathbf{f}) ★ U(\mathbf{g}) \neq U[\ f_1★g_1 \quad f_2★g_2 \quad \cdots \quad f_p★g_p \ ]^T \text{ for some } x \in \mathbf{E}^n,$$

(3-32)

but

$$T[U(\mathbf{f}) ★ U(\mathbf{g})] = T[U[\ f_1★g_1 \quad f_2★g_2 \quad \cdots \quad f_p★g_p \ ]^T] \text{ for all } x \in \mathbf{E}^n.$$

(3-33)

Undesirable property (3-32), (3-33) is caused by the definition of supremum of $\mathbf{E}^p$, (3-28). Since we cannot relate every two elements using the partial-order $\succeq$, the vector space $\mathbf{E}^p$ is not a *well-ordered* vector space. Thus this extension of umbra applies only if the range of multi-variate signal is a *well-ordered* space under some ordering operation. The following Example 3-2 shows the difference between separate morphological operations and *multi-variate morphological* operations using a special ordering relationship.

<u>Example 3-2</u> Let $\mathbf{X} = [\ X_1\ X_2\ ]^T$,and $X_1\ :\ \mathbf{Z}^2 \to \{0,1\}$, $X_2 : \mathbf{Z}^2 \to \{0,1\}$. Figures 3-4-1, 3-4-2, and 3-4-3 show $\mathbf{X}$, X_1,and X_2, respectively. In Figure 3-4-1, symbols **a**, **b**, and **c** denote vectors $[1\ 0]^T$, $[0\ 1]^T$, and $[1\ 1]^T$, respectively.
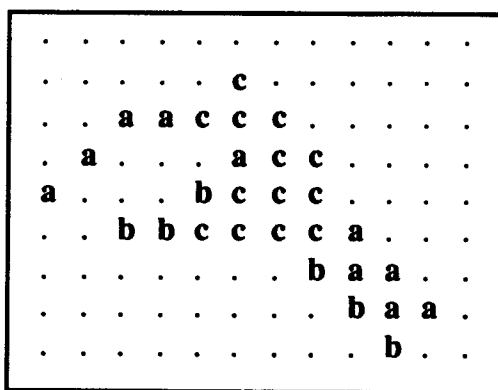
Figure 3-4-1 Multi-variate image **X**
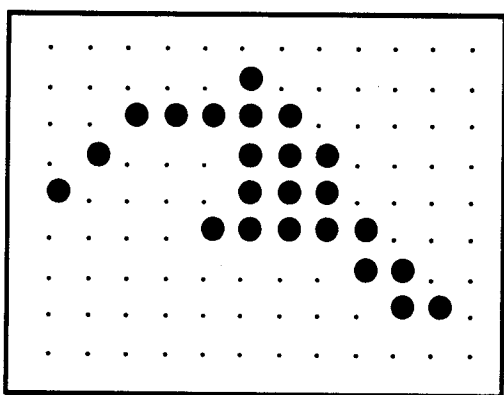
Figure 3-4-2 Image component X₁
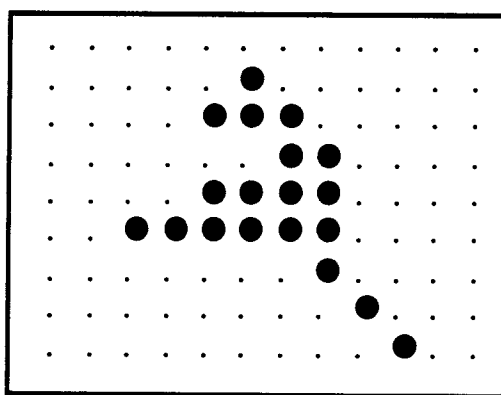
Figure 3-4-3 Image component X₂

Figures 3-4-4 and 3-4-5 show the results of separate opening by a structuring element (B), Rhombus ( ). Figures 3-4-6 and 3-4-7 show the results of separate closing.
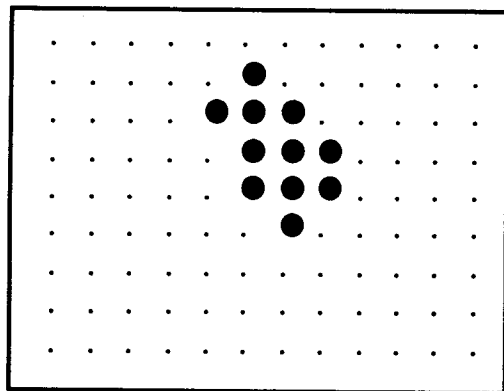
Figure 3-4-4 Opening X₁OB
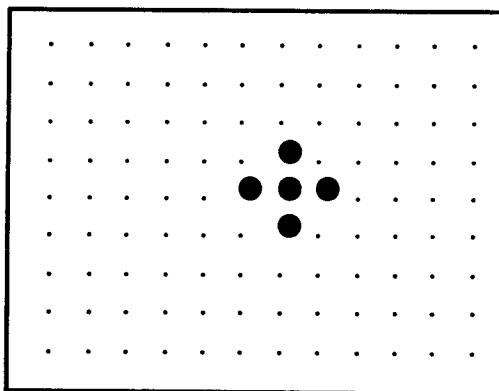
Figure 3-4-5 Opening X₂OB
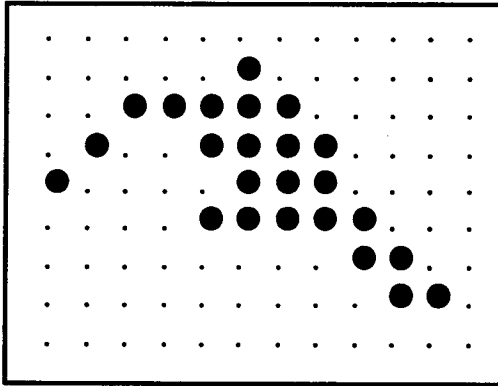
Figure 3-4-6 Closing $X_1 \bullet B$
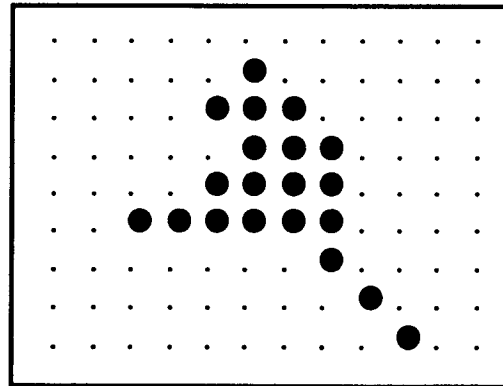
Figure 3-4-7 Closing $X_2 \bullet B$

```
          a
      a a a
        a c a
        c c c
          c
```

Figure 3-4-8 Opening $X O B$ by
separate operation

```
          c
    a a c c c
  a . a c c c
a . . b c c c
  . b b c c c c a
          b a a
            b a a
              b
```
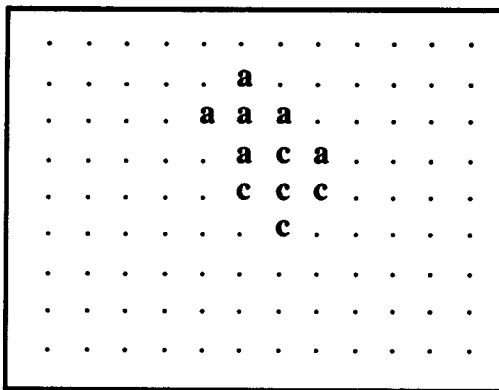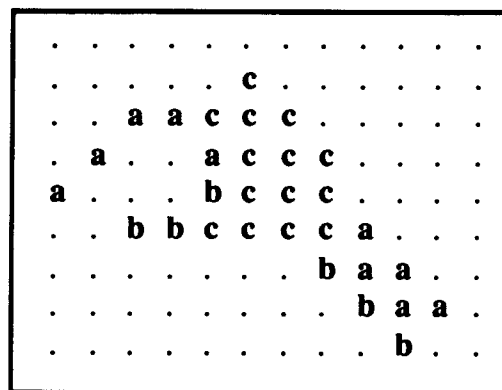
Figure 3-4-9 Closing $X \bullet B$ by
separate operation

Figures 3-4-10 and 3-4-11 show the results of *multi-variate* opening and closing when vector ordering a<b<c is assumed, and $(X \oplus B^s)(z) = \max\{X(y) : y \in B_z\}$, $(X \ominus B^s)(z) = \min\{X(y) : y \in B_z\}$, where $z, y \in Z^2$. Figures 3-4-12 and 3-4-13 show the results of opening and closing when b<a<c is assumed.

```
.  .  .  .  .  .  .  .  .  .  .  .  .  .
.  .  .  .  .  a  .  .  .  .  .  .  .
.  .  .  .  a  a  a  .  .  .  .  .
.  .  .  .  .  a  c  a  .  .  .  .
.  .  .  .  a  c  c  c  .  .  .  .
.  .  .  .  .  a  c  a  a  .  .  .
.  .  .  .  .  .  .  a  a  a  .  .
.  .  .  .  .  .  .  .  a  a  a  .
.  .  .  .  .  .  .  .  .  a  .  .
```

Figure 3-4-10 Opening XOB assuming a<b<c

```
.  .  .  .  .  .  .  .  .  .  .  .  .  .  .
.  .  .  .  .  .  c  .  .  .  .  .  .  .
.  .  a  a  c  c  c  .  .  .  .  .  .
.  a  a  a  a  c  c  c  .  .  .  .  .
a  a  a  a  b  c  c  c  .  .  .  .
.  .  b  b  c  c  c  c  a  .  .  .
.  .  .  .  .  .  .  b  a  a  .  .
.  .  .  .  .  .  .  .  b  a  a  .
.  .  .  .  .  .  .  .  .  b  .  .
```
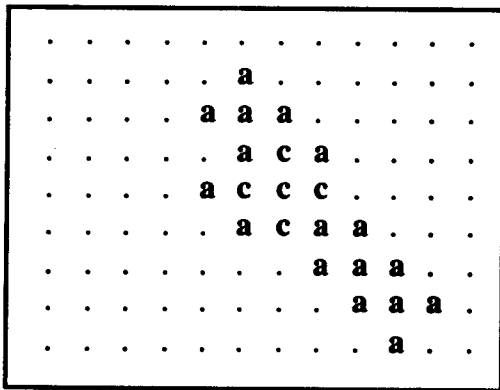
Figure 3-4-11 Closing X●B assuming a<b<c

```
.  .  .  .  .  .  .  .  .  .  .  .  .  .
.  .  .  .  .  a  .  .  .  .  .  .  .
.  .  .  .  a  a  a  .  .  .  .  .
.  .  .  .  .  a  c  a  .  .  .  .
.  .  .  .  b  c  c  c  .  .  .  .
.  .  .  .  .  b  c  b  b  .  .  .
.  .  .  .  .  .  .  b  b  b  .  .
.  .  .  .  .  .  .  .  b  b  b  .
.  .  .  .  .  .  .  .  .  b  .  .
```
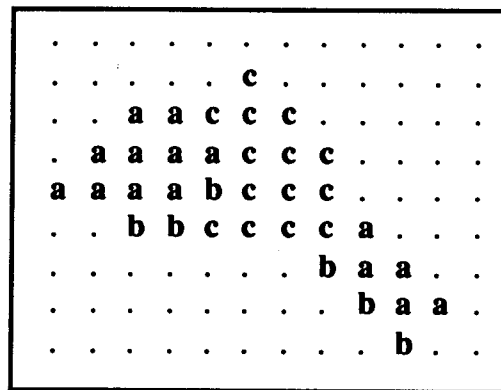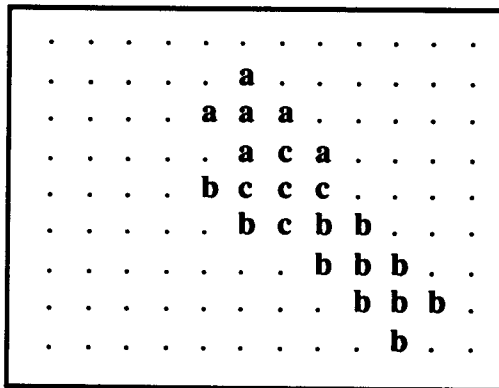
Figure 3-4-12 Opening XOB assuming b<a<c

```
.  .  .  .  .  .  .  .  .  .  .  .  .  .  .
.  .  .  .  .  .  c  .  .  .  .  .  .  .
.  .  a  a  c  c  c  .  .  .  .  .  .
.  a  b  b  a  c  c  c  .  .  .  .  .
a  b  b  b  b  c  c  c  .  .  .  .
.  .  b  b  c  c  c  c  a  .  .  .
.  .  .  .  .  .  .  b  a  a  .  .
.  .  .  .  .  .  .  .  b  a  a  .
.  .  .  .  .  .  .  .  .  b  .  .
```
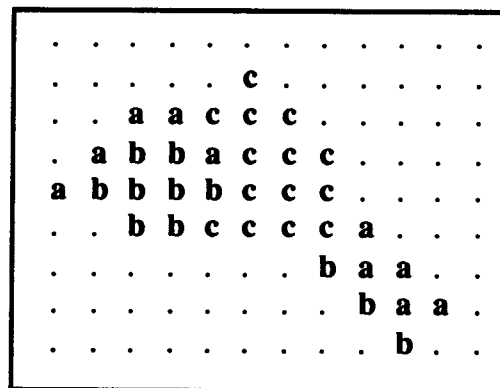
Figure 3-4-13 Closing X●B assuming b<a<c

As shown in Figure 3-4-8 and Figure 3-4-10 (or 3-4-12), separate opening eliminates a *tail* in the right-lower corner, while *multi-variate* opening preserves it. Figure 3-4-9 and Figure 3-4-11 (or 3-4-13) show that separate closing does not fill the hole on the left side, but *multi-variate* closing does. We observe that *multi-variate operations* yield more intuitive results, in the sense that closing fills small gulfs and opening suppresses sharp capes, where the reference point for "small" and "sharp" is the size and shape of the structuring element.

The advantages of using *multi-variate morphology* are overshadowed by the difficulties in dealing with partially-ordered sets. Convenient, partially-ordered sets are currently under investigation but no conclusive results are available at this point in time.

Consequently we devote our efforts to a two-stage processing, where a multi-variate signal undergoes preprocessing transformation and the morphological operations are applied to each component of the transformed signal separately. We propose an optimal transformation from the point of view of morphological filtering performance as discussed in Chapter 2.

3.3 Morphological Application of the Mapping Technique

The proposed mapping technique, NAD-CVM (normalization and diagonalization of sample covariance matrices), is based on the maximum separation between object and background in terms of the sample variances. The usage of sample variance in image processing and analysis is limited to the spatial feature extraction in the form of a histogram feature and to the dimensionality reduction. However we hypothesize that large sample variance, obtained by a specific mapping technique, gives better morphological information extraction. To justify this hypothetical statement we define a morphological measure which corresponds to the sample variance. We note here that it is a difficult task to find a morphological measure corresponding to the sample variance, the main difficulty being the difference in quantitative description of both concepts.

Therefore we propose to replace sample variance, with the first absolute central moment of observation :

$$\mu_j = \frac{1}{n} \sum_{i=1}^{n} |\, y_j(i) - m_j \,| \; , \tag{3-34}$$

where $y_j(i)$ is the observation for the j-th component, the number of observations is n, and m_j denotes sample mean of the j-th component. For simplicity we omit in the sequel the subscript j, unless otherwise noted. The first absolute central moment is not used in general as a measure of dispersion of observed data since the sample variance provides more computationally tractable algorithms using the well

developed theories in l_2 normed vector spaces. However there are several good reasons for using μ instead of σ^2 (sample variance) :

(1) it is difficult to find a morphological measure corresponding to σ^2 as noted above,

(2) μ can be reinterpreted by a morphological measure which is defined using the anti-extensivity property of opening,

(3) for the transformed components of some signals, ordering of variances can be preserved , i.e., if $\sigma_i < \sigma_j$, then $\mu_i < \mu_j$ for all i,j=1,2,...,p.

The last statement is not true in general, but holds for the transformation such that the "sample distribution" of transformed values approximates Normal distribution. For a normal distribution with variance σ^2 , the first absolute central moment is $(2/\pi)^{1/2}\sigma$. The term "sample distribution" denotes here the relative frequency distribution obtained from the observations. This provides a reasonable starting point for the analysis. (3) is also true for the sample distributions which approximate uniform, exponential, and gamma distributions. Thus (3) is approximately true for the transformed data unless there is a large number of outliers from the above distributions.

Next we need to define a morphological equivalent of μ. Let us consider a function f with bounded support D :

$$f : D \subset \mathbf{Z}^2 ---> \mathbf{R}. \qquad (3\text{-}35)$$

Let S_λ be a structuring element such that

$$S_\lambda = \underbrace{S \oplus S \oplus \cdots \oplus S}_{\lambda-1} \quad (\lambda\text{-1 times dilation operation by S}), \quad (3\text{-}36)$$

where $S \subset \mathbf{Z}^2$ is a convex set, and $\lambda=1,2,\cdots$.

Define the function f*: $D \subset \mathbf{Z}^2 ---> \mathbf{R}$ as f*(x) = | f(x) - m |, where m is the sample mean of f(x), i.e., $m=(1/M(D))\Sigma_{x\in D}f(x)$, and $M(D)$ denotes the number of elements of a set D. Now we can define morpholgical pattern distribution function r(λ) of f*, and morphological pattern density function p(λ), as follows :

$$r(\lambda) = Mes(\ f^* - (\ f^* \bigcirc S_\lambda\ )),$$ (3-37)

$$p(\lambda) = r(\lambda) - r(\lambda-1), \quad p(0)=0, \quad p(1)=r(1),$$ (3-38)

where $\bigcirc$ denotes opening operation and $Mes(g) = \Sigma_{x \in D}\ g(x)$. Let N be the minimum λ such that

$$Mes(\ f^* \bigcirc S_\lambda\ ) = Mes(\ f^* \bigcirc S_{\lambda+1}\ ) = \varepsilon,$$ (3-39)

for all $\lambda \geq N$. We note that λ such that $D \subseteq S_\lambda$, always satisfies the relationship (3-39), and ε can be determined by

$$\varepsilon = M(D) \cdot \min_{x \in D}\ (f^*(x)) \quad \text{for such a } \lambda.$$

Next we introduce another morphological measure $h(\lambda)$:

$$h(\lambda) = \sum_{i=\lambda}^{N} \frac{p(i)}{M(S_i)}, \quad \lambda=1,2,\dots,N.$$ (3-40)

Then the first absolute central moment of f can be represented using $h(\lambda)$ as follows :

From Equations (3-37) and (3-39), we have

$$
\begin{aligned}
r(N) &= Mes(f^* - (f^* \bigcirc S_N)) \\
&= Mes(f^*) - Mes(f^* \bigcirc S_N)) \\
&= Mes(f^*) - \varepsilon \\
&= \mu \cdot M(D) - \varepsilon.
\end{aligned}
$$ (3-41)

Equation (3-41) can be rewritten as

$$\mu = \frac{1}{M(D)}\ (r(N) + \varepsilon).$$ (3-42)

Using Equation (3-40), we have

$$\sum_{\lambda=1}^{N} (h(\lambda)(M(S_\lambda)-M(S_{\lambda-1})) = \sum_{\lambda=1}^{N} p(\lambda).$$ (3-43)

Using Equations (3-37) and (3-38), we get :

$$\sum_{\lambda=1}^{N} p(\lambda) = r(N).$$ (3-44)

Combining Equations (3-43) and (3-44), and substituting for r(N) in Equation (3-42) yields :

$$\mu = \frac{1}{M(D)}(\sum_{\lambda=1}^{N} (h(\lambda)(M(S_\lambda)-M(S_{\lambda-1})) + \varepsilon) .$$

(3-45)

Equation (3-45) can be interpreted geometrically using the stack of cylinders as shown in Figure 3-5.
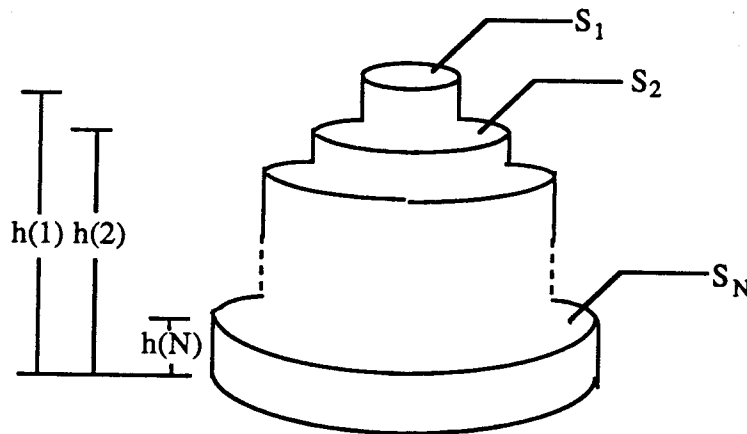

Figure 3-5 Interpretation of *average height* using the stack of cylinders.

The volume of the shape in Figure 3-5 is equal to $\Sigma\ h(\lambda)\cdot M(S_\lambda-S_{\lambda-1})$. Thus we can interpret $h(\lambda)$ as an *average height* of $(f^*\bigcirc S_\lambda)$, or alternatively $p(\lambda)/M(S_\lambda)$ as an *average height* of $(f^*-(f^*\bigcirc S_\lambda))$. The *height* is interpreted in the sense that $p(\lambda)/M(S_\lambda)$ is the value of volume divided by the area of S_λ.

Let us consider the NAD-CVM mapping. We expect that the locations of *discontinuities* in the original components will be maintained at least in one of the transformed components. On the other hand, the variance increase (or decrease) caused by the mapping is due to the facts that the mapping enlarged (or lessened) the amount of *jumps* and created (or eliminated) *discontinuities* by the normalization process and rotation of coordinates. This corresponds to increase (or decrease) of $p(\lambda)$ for a given structuring element, S_λ, i.e.,

increase (or decrease) of *average height*. This relationship between variance and the *average height* can be used to select a proper structuring element in a morphological algorithm. Such a proper selection eliminates small details in the background, but enhances discontinuities in the object.

To illustrate experimentally the above concepts, we use an RGB color image. Figure 3-6 shows the original R,G,B components at a horizontal section. Figure 3-7 is the result of NAD-CVM transformation designating x-axis region [40,110] as the object and the rest as the background. In Figure 3-7, NAD-CVM-1 is the coordinate which has the largest variance for the object and the smallest variance for the background. This is reversed for NAD-CVM-2. NAD-CVM-3 is the coordinate which represents no interest from the analysis point of view. Figures 3-8 and 3-9 show the plot of $p(\lambda)/M(S_\lambda)$ *versus* λ. We compare Figures 3-8 and 3-9 to identify the size of the largest structuring element λ^* which still provides a significant *average height*. In Figure 3-8 λ_1^* for the object is larger than the value for the background λ_2^*, for all three R,G,B coordinates. In Figure 3-9, as we expected, λ_1^* is larger than λ_2^* for the first coordinate (Figures 3-9a-1 and 3-9a-2), while λ_2^* is larger than λ_1^* for the second coordinate (Figures 3-9b-1 and 3-9b-2). Thus, if we use the structuring element with $\lambda=2$ for the first coordinate, we can eliminate all the details which have support smaller than S_2, and the remaining discontinuities with the support larger than S_2 can be used to extract information. Note that we have already eliminated nearly all the meaningless details of the background. The same approach with reversed order of object and background variances can be used for the second coordinate. Another important observation is that the first few structuring elements dominate the plot in the sense of significant *average height*. This allows to reduce the computational burden, and consequently makes the application of the proposed algorithm feasible by choosing smaller structuring elements.
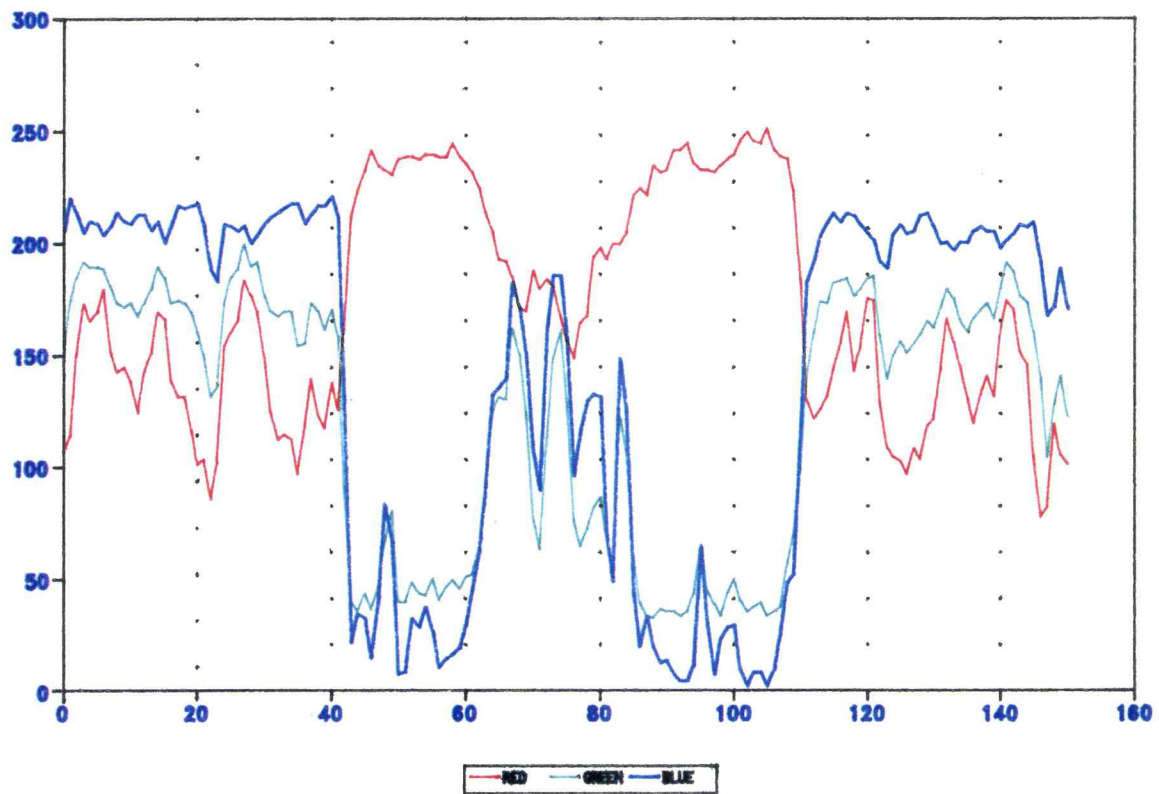
Figure 3-6 Original Data : R,G,B components at a section of color image.
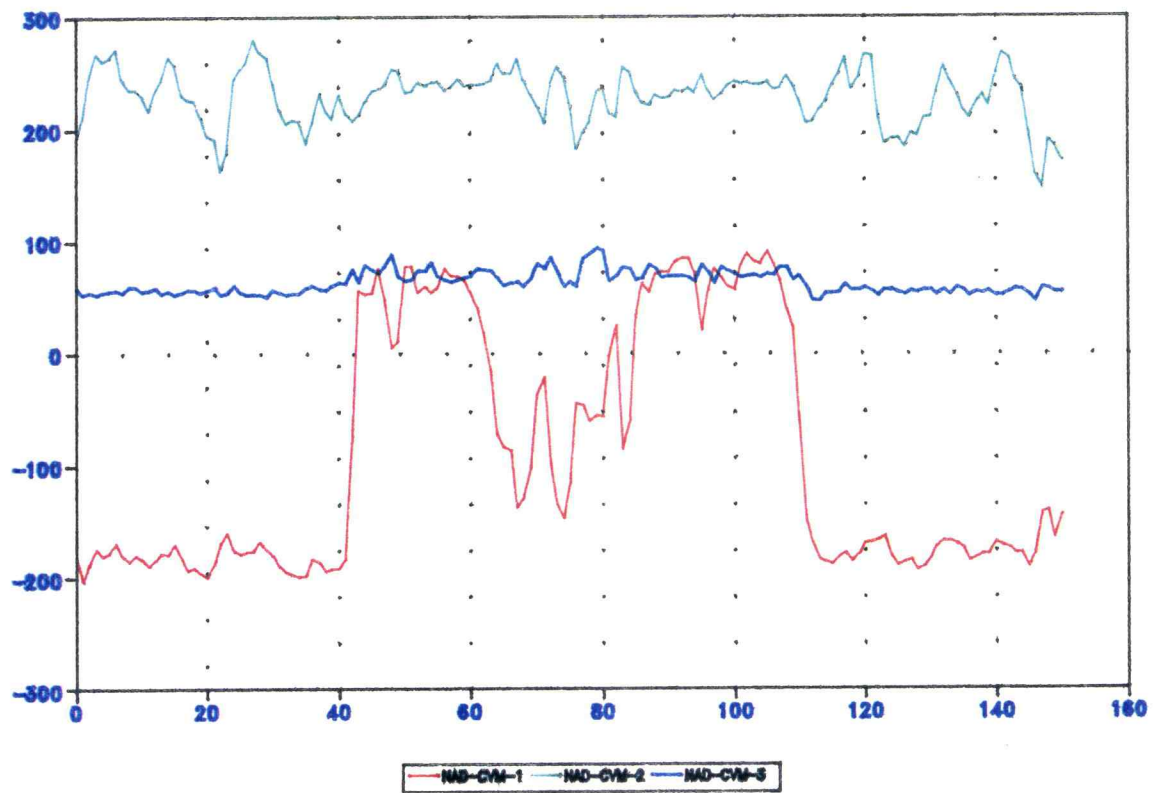
Figure 3-7 Results of NAD-CVM mapping of original data .

R-Component

3-8a-1

3-8a-2

G-Component

3-8b-1

3-8b-2

B-Component

3-8c-1
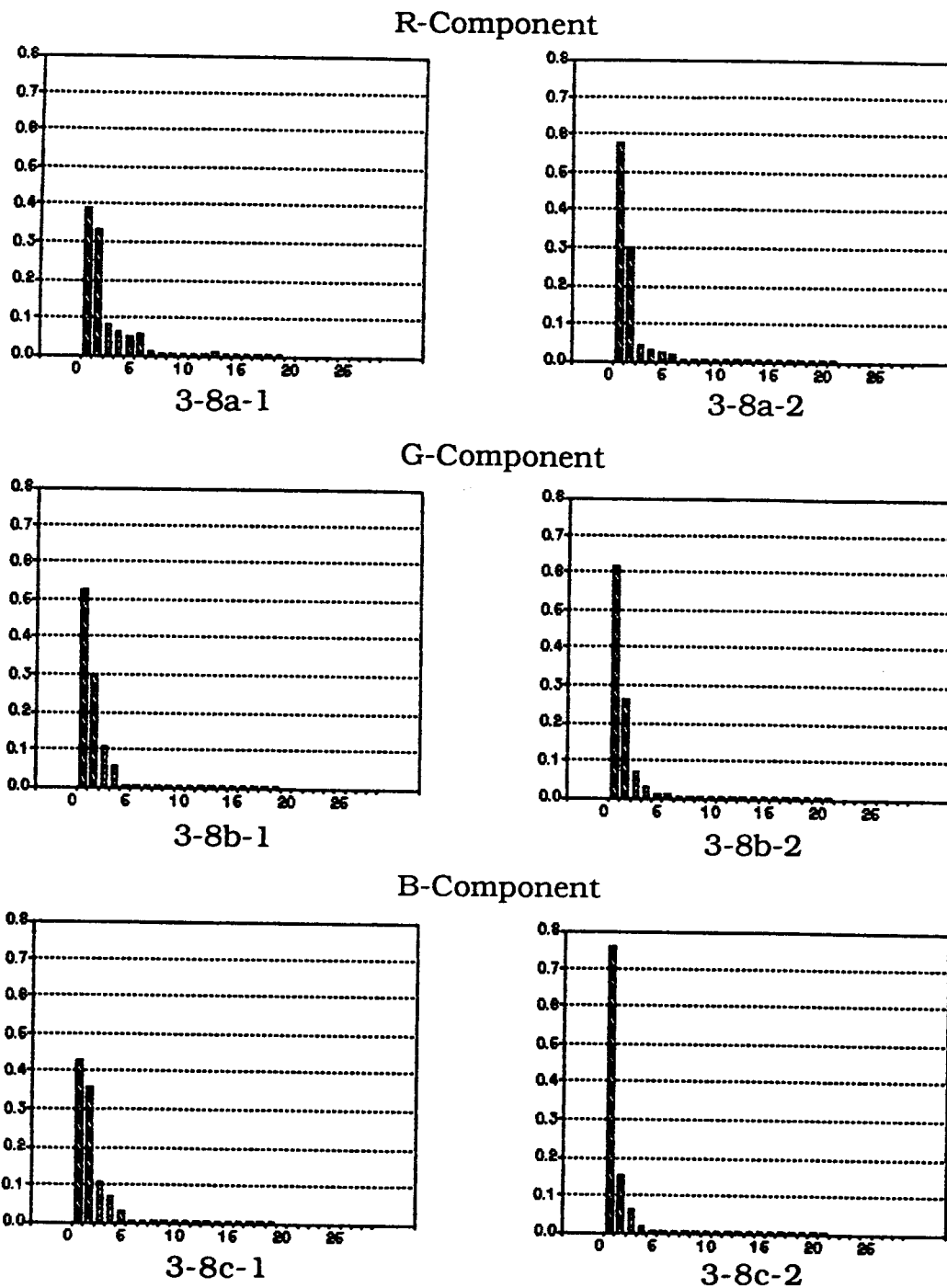

3-8c-2

Figure 3-8 Plots of normalized $p(\lambda)/M(S_\lambda)$ vs. λ for the original R,G,B components.

The left side applies to the object component, and the right side applies to the background component.

NAD-CVM-1

NAD-CVM-2

NAD-CVM-3

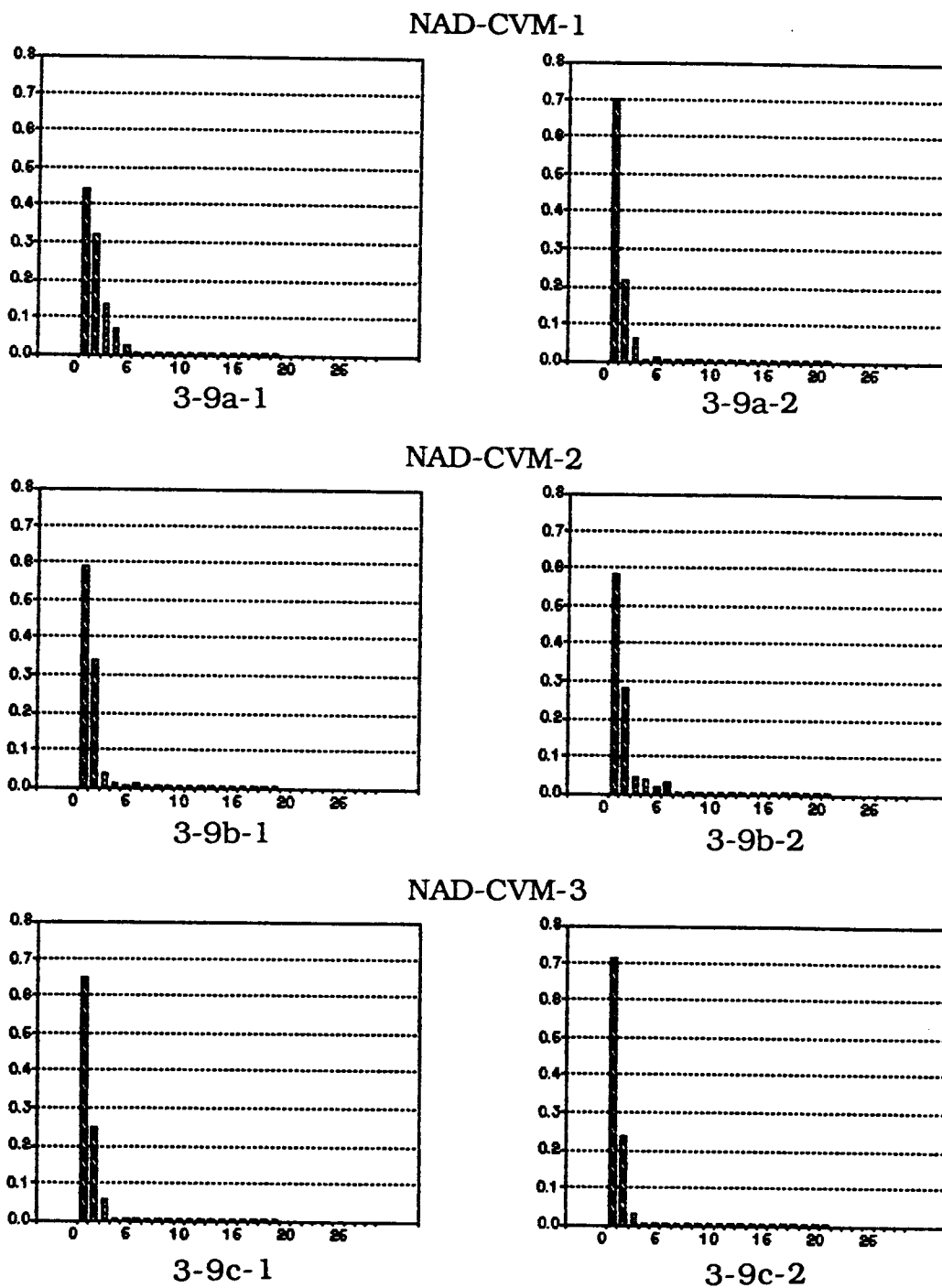3-9a-1

3-9a-2

3-9b-1

3-9b-2

3-9c-1

3-9c-2

Figure 3-9 Plots of normalized $p(\lambda)/M(S_\lambda)$ vs. λ for the components transformed by NAD-CVM method.

The left side applies to the object component, and the right side applies to the background component.

4. APPLICATIONS AND EXPERIMENTAL RESULTS

4.1 Applications

The combination of the NAD-CVM and mathematical morphology was applied to analyze color images. This combination seems to be promising since the NAD-CVM may be applied to any multi-variate two-class signal and the morphological filters have long proven their effectiveness in image processing. Thus any multi-variate signal analysis problem can benefit from the combined results of Chapter 2 and Chapter 3. In the following discussion the underlying assumption is that the analyzed multi-variate signal is a combination of an object and a background. The RGB color images and multi-spectral images are most frequently used as the examples of multi-variate signals.

In the first experiment, we use the RGB wooden board images. Various color coordinate systems have been developed to improve human visual perception, while mapping methods for color images are proposed to get a better image analysis performance in vision systems. We briefly review existing color coordinate systems in Section 4.1.1.

The color image with the RGB tristimulus components can be separated by NAD-CVM mapping method. We construct an unsupervised mapping algorithm using Lemma 2.1, under assumption that the background sample covariance matrix is known. Using the transformed components, a proper structuring element operates on the object and the background selectively. The examples of selective morphological operations are opening and closing, which remove and fill image features according to the sizes and shapes of the structuring elements. In Section 4.1.2, we recommend an edge detection method with an open-close smoothing which enhances edges in the object region while suppressing the edges in the background.

We further examine the effectiveness of the proposed edge-detector by processing remotely sensed ocean data and a complex color image.

4.1.1 An Overview of Color Image Analysis

Image analysis is concerned with the extraction of quantitative information from an image. Image analysis is distinguished here from other types of image processing, such as enhancement, restoration, and image data compression. The output of the image analysis system is usually numerical whereas the output of other image processing schemes is usually in a form of an image. Basically the image analysis includes feature extraction, segmentation, and classification techniques. Feature extraction is required for an identification of the object. Segmentation techniques are used to isolate the desired object from a scene so that the measurements can be made on the object. Quantitative measurements of object features allow classification and description of the image. Using the color images, we increase dramatically variety of features as compared to the features found in the monochrome images. Paradoxically this is both the main advantage and pitfall of color image analysis.

In general we may consider two different approaches to color image analysis : one uses colorimetry and color coordinate system, and the second concentrates on multi-spectral image analysis, treating tristimulus values of color as multi-variate signal. In this study we explore both approaches. However we will not use colorimetry or color constancy [4],[7],[14],[10], since these results are not directly related to the proposed here techniques.

The color can be specified by

(1) its tristimulus values for a given set of primaries,

(2) its chromaticity values and its luminance,

(3) some linear or nonlinear invertible function of its tristimulus or chromaticity values.

There are many different color coordinate systems employed for the specification of color. Pratt   [20, pp.72-83] provides a detail

discussions on those color coordinate systems together with their historical and analytic significance. C.I.E. spectral primary color coordinate system (R_C, G_C, B_C) and N.T.S.C. receiver primary color coordinate system (R_N, G_N, B_N) belong to category (1). C.I.E. X-Y-Z color coordinate system (X,Y,Z) and N.T.S.C. transmission color coordinate system (Y,I,Q) are directly transformed form of R_C-G_C-B_C and R_N-G_N-B_N coordinate system, respectively, such that Y tristimulus value is equivalent to the luminance of the color, and thus belong to category (2). C.I.E uniform chromacity scale color coordinate system (U,V,W) is a nonlinear transformation of the X-Y-Z coordinate system, which ensures that any unit change in the chromacity diagram of that system is perceived as an equivalently noticeable color shift to an observer. The U*-V*-W* coordinate system is an extension of the U-V-W coordinate system in an attempt to obtain a color solid for which unit shifts in luminance and chrominance are uniformly perceptible.

S-q-W* color coordinate system is the polar representation of the U*-V*-W* coordinate system, where W* represents the luminance (scaled and shifted), S and q components are related to saturation of color and its hue, respectively. It is known that the S-q-W* color coordinate system is the closest color coordinate system to human visual perception. The L-a-b coordinate system is a nonlinear transformation of X-Y-Z coordinate system to provide a relatively accurate measure of color in agreement with the Munsell color system. The above color coordinate systems can be considered linear or nonlinear transformations of the N.T.S.C. receiver primary coordinate system, R_N-G_N-B_N. The R_N-G_N-B_N tristimulus values of N.T.S.C. receiver primary system are highly correlated with one another. In image analysis and pattern recognition, feature selection problem is of primary interest. The importance of this step lies in the fact that if the features show significant differences from one class to another, an effective classifier can be easily designed. As mentioned earlier, tristimulus values of color can be the features of an object in image analysis system. Pratt [20] pointed out the fact that it is desirable to work with uncorrelated components, and Ohta [19] concluded that a large discriminant power of a feature can be achieved by large variance of that feature. These criteria (uncorrelation and large discriminant

power) can be fulfilled using Karhunen-Loeve transformation (or principal components method), which was reviewed in Chapter 2. The color coordinate system using Karhunen-Loeve transformation is called Karhunen-Loeve color coordinate system. (In the sequel Karhunen-Loeve will be abbreviated by K-L.) There are several modifications of the K-L transformation [5], [3], [13], [21], which were proposed to maximize the given criteria in multivariate-multiclass situation. It should be noted that those are not originally intended to obtain new color coordinate system, but to extract features used in pattern recognition of multivariate-multiclass data. NAD-CVM mapping is also included in that category. If we view the color image as composed of an object of interest and background which represents no interest from the analysis point of view, then we have a two-class color feature extraction problem. Thus the NAD-CVM mapping provides two features: one is the *best* for analyzing the object and the *poorest* for analyzing the background, the other one has complementary properties.

Next we survey image analysis techniques that can be applied to analyze color images. It is apparent that all monochrome image analysis techniques can be used for color images. There are three categories of image analysis techniques : feature extraction, segmentation, and classification. Since the classification techniques will not be changed by adding color information (i.e., color merely increases dimensionality), we exclude those techniques from our discussion. The segmentation techniques are basically based on feature extraction. Pratt [20] defines an image feature as a distinguishing primitive characteristic or attribute of an image field. Some features are natural in the sense that they are defined by the visual appearance of an image, while others, so-called artificial features, result from specific manipulations or measurements of an image. Natural features include the brightness of a region pixels (amplitude feature), edge outlines of objects, and grey scale textural region. Image amplitude histograms and spatial frequency spectra are examples of artificial features. Most applications of mathematical morphology involve natural features. Amplitude features are the simplest and perhaps the most

useful ones. To aid amplitude feature extraction, we propose a new color coordinate transformation, NAD-CVM. Amplitude features can form raw data for edge and texture feature extraction or can be used in amplitude thresholding. Variety of image analysis algorithms can be derived by using multi-variate amplitude features. The main problem is how to manipulate these multi-variate features. There are several known methods :

(1) Manipulate as if it is a uni-variate image using the weighted sum of features. This is equivalent to the case of a monochrome image analysis.

(2) Manipulate each feature separately, then combine the results in a single image.

(3) Manipulate the features in a vector space considering the features of a pixel as the elements of a vector.

As discussed in Section 3.2 and 3.3, methods (1) and (2), equipped with a proper mapping, can provide effective color image analysis algorithms.

4.1.2 Application of NAD-CVM Mapping and Mathematical Morphology to Color Image Analysis

In color image analysis, color of a pixel is usually given as three values corresponding to the N.T.S.C. receiver primary color coordinate system, (R_N, G_N, B_N). We consider the color image represented by (R_N, G_N, B_N) as an original color image, and denote it using the multi-variate function notation:

$$\mathbf{h} : D \subset \mathbf{Z}^2 \dashrightarrow \mathbf{R}^3, \qquad (4\text{-}1)$$

where $D = (1, 2, ..., N_1) \times (1, 2, ..., N_2)$.

Equation (4-1) can be rewritten as :

$$\mathbf{h}(x, y) = [h_1(x, y)\ h_2(x, y)\ h_3(x, y)]^T, \quad x = 1, 2, ..., N_1, \text{ and } y = 1, 2, ..., N_2 \qquad (4\text{-}2)$$

In many image analysis cases it is not unreasonable to assume that sample covariance matrix of background is known *a priori* . Denote this *a priori* known sample covariance matrix by K_2. Using Lemma 2.1, we can construct an unsupervised NAD-CVM mapping algorithm as

shown in Figure 4.1. Note that the eigenvalue matrix Λ, does not have property (2-18), since $\mathbf{K}_2$ is used in place of $\mathbf{S}_2 = \omega_2\mathbf{K}_2$. However the eigenvalues of $\mathbf{K}_2$ are equal to the eigenvalues of $\mathbf{S}_2$ multiplied by ω_2. Therefore transformation by the eigenvector which corresponds to the smallest eigenvalue of $\mathbf{K}_2$ yields maximum ratio of the object to the background sample variances. On the other hand, we get the minimum variance ratio using the eigenvector which corresponds to the largest eigenvalue of $\mathbf{K}_2$. Consequently, this algorithm can be used for a color image analysis, especially for algorithms that extract object information by separating background from the image, (*e.g.*, segmentation algorithm).

Next we consider morphological applications of NAD-CVM mapping. There are many desirable properties of morphological filters, which are used in numerous image processing and analysis applications. The following property of opening and closing is the most valuable as far as NAD-CVM mapping is concerned :

By properly selecting structuring element size and shape, the operation of opening and closing can selectively remove image features according to feature size and orientation. The smaller structuring element, the smaller the details which are filtered or modified. Transformations which apply opening and closing are less severe and introduce less distortion when a small structuring element is used before a larger one.

Combining the concepts of Section 3.3 and the above property of opening and closing , we can select a structuring element which operates on the object and the background selectively. An intuitive example is the open-close smoothing prior to applying an edge detection algorithm. The open-closing by a spherical structuring element is known as the rolling ball algorithm [25]. A simple morphological edge detector can be defined by considering the erosion residue as an edge strength. Let the image components, $f_1(x,y)$, $f_2(x,y)$ be the result of applying our algorithm (Figure 4.1), and represent the components with the maximum and minimum ratios of the object to the background variances, respectively.
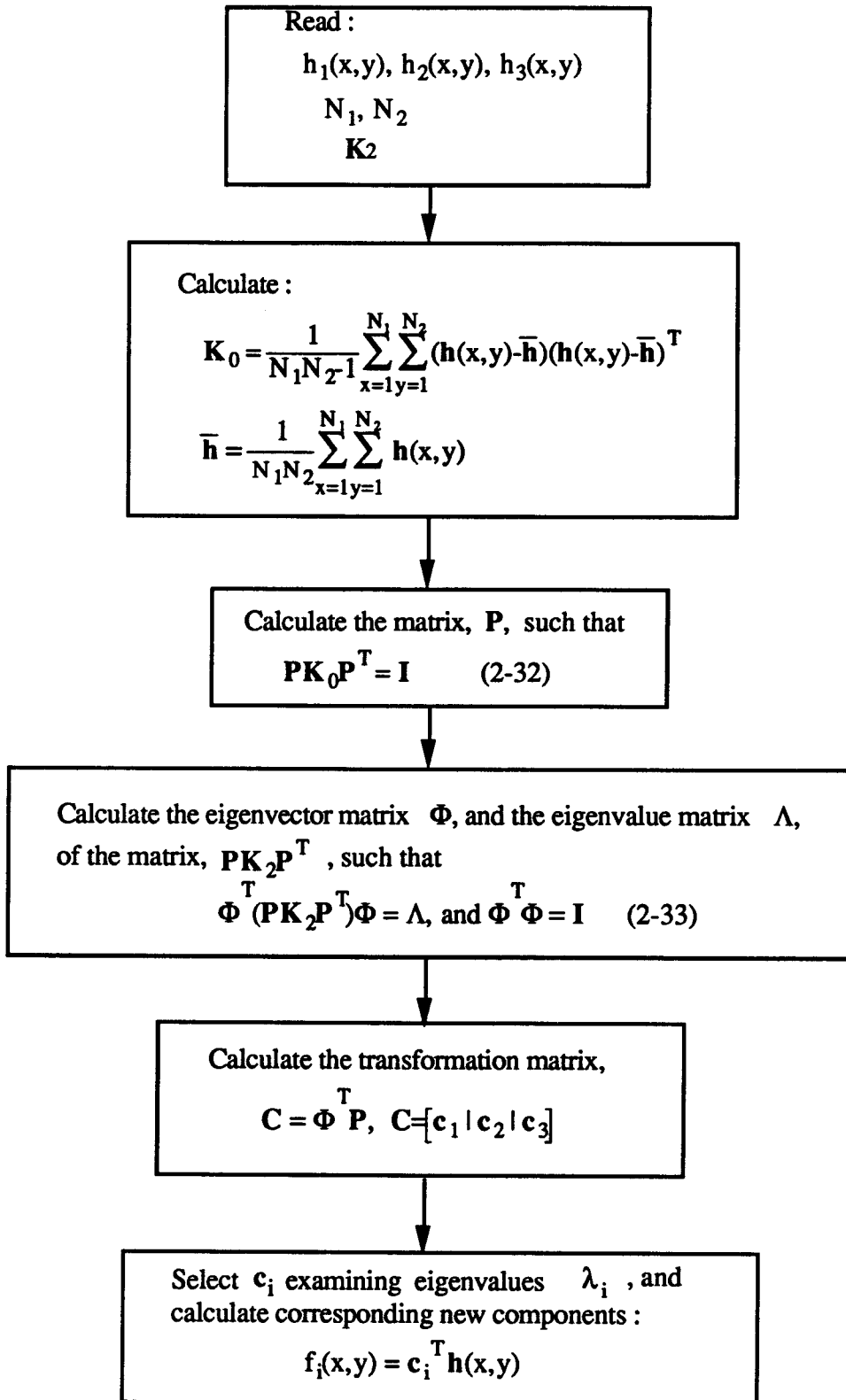
Read :

$h_1(x,y)$, $h_2(x,y)$, $h_3(x,y)$

N_1, N_2

K_2

↓

Calculate :

$$K_0 = \frac{1}{N_1 N_2 - 1} \sum_{x=1}^{N_1} \sum_{y=1}^{N_2} (h(x,y)-\bar{h})(h(x,y)-\bar{h})^T$$

$$\bar{h} = \frac{1}{N_1 N_2} \sum_{x=1}^{N_1} \sum_{y=1}^{N_2} h(x,y)$$

↓

Calculate the matrix, P, such that

$$PK_0 P^T = I \qquad (2\text{-}32)$$

↓

Calculate the eigenvector matrix Φ, and the eigenvalue matrix Λ, of the matrix, $PK_2 P^T$, such that

$$\Phi^T (PK_2 P^T)\Phi = \Lambda, \text{ and } \Phi^T \Phi = I \qquad (2\text{-}33)$$

↓

Calculate the transformation matrix,

$$C = \Phi^T P, \quad C = [c_1 | c_2 | c_3]$$

↓

Select c_i examining eigenvalues λ_i , and calculate corresponding new components :

$$f_i(x,y) = c_i^T h(x,y)$$

Figure 4-1 Unsupervised NAD-CVM algorithm.

Let the structuring element B, be a finite subset of $\mathbf{Z}^2$. Then an edge strength can be determined by

$$e_i(x,y) = f_i(x,y) - (f_i \ominus B)(x,y) , \quad i=1,2. \qquad (4\text{-}3)$$

A binary edge image can be produced by thresholding the edge strength $e_i(x,y)$. An appropriate threshold level should be determined by a design criterion. The operation of Equation (4-13) is a noise sensitive detector, and should not be applied directly to noisy images [15]. Thus a noise suppression algorithm is necessary before we apply the morphological edge detection operation. Open-closing is a good filter for noise suppression [25], [16], [27]. However the open-closing filter by a large structuring element can remove both noise and image features which do not fit the structuring element. Thus the structuring element of open-closing filter should be selected as small as possible, but large enough to remove noise. Here we consider image features in background as noise in the sense that those features represent no interest from the analysis point of view. Furthermore it is desirable to remove those features with minimal loss of object feature information. Thus the image component f_1 will have good object but not background edge features, provided that we use a structuring element selected according to the criterion discussed in Section 3.3.

The application of NAD-CVM and mathematical morphology to the edge detection is only one of many possible applications. The applications for NAD-CVM and mathematical morphology are our continuing effort.

4.2 Experiments on Color Wooden Board Images

The experiment algorithm was implemented in *FORTRAN* language and was tested on a *Stardent* computer running the *UNIX* operating system. Three 200x150 color images were used in this experiment:

(1) a wooden board with blue stains (Wood 1),
(2) a wooden board with an intergrown knot (Wood 2),
(3) a wooden board with a spike knot (Wood 3).

Figure 4-2 shows the original test images. All three test images were produced by scanning color pictures from [29] using the *HP* color scanner. All pictures shown in this chapter are displayed using *AVS* [28] and printed using *TEK-4693DX* color printer.

Experiments are designed to verify conjectures and proposed signal analysis technique described in Chapter 2 and 3. The algorithm performance measure is a visual inspection of binary edge images which are produced by morphological edge detector (Section 4.1.2). More detail in the binary image of the object, and stronger suppression of the background detail, better the algorithm performance.

The experiments are divided into three groups :

(1) test on selecting size and vertical shape of structuring element used in open-close smoothing,

(2) comparison of various mapping methods,

(3) test on the advantage of using color information and mapping method.

In calculation of NAD-CVM and NAD-CRM mappings, sample covariance matrix of the background K_2 and sample autocorrelation matrix of the background R_2 were calculated from a selected part of sample image. For convenience, we assign the names to the image components as shown in Table 4.1. The component marked by asterisk is selected as the *best* component for a given coordinate system. The *best* means either the maximum eigenvalue ratio of object to background for the NAD-CVM and the NAD-CRM mappings, or the maximum eigenvalue for the K-L mapping. The *best* for the RGB coordinate system is determined by a visual inspection. In group (2), we show only one binary edge image of a selected component per each mapping method.

Table 4.1 Names of component images

Coordinate System		Wood1	Wood2	Wood3
Original	R	* W1-R	W2-R	W3-R
	G	W1-G	* W2-G	W3-G
	B	W1-B	W2-B	* W3-B
NAD-CVM	1	W1-H	W2-H	* W3-H
	2	W1-I	* W2-I	W3-I
	3	* W1-J	W2-J	W3-J
NAD-CRM	1	W1-K	W2-K	* W3-K
	2	* W1-L	W2-L	W3-L
	3	W1-M	* W2-M	W3-M
K-L	1	W1-N	W2-N	W3-N
	2	W1-O	W2-O	W3-O
	3	* W1-P	* W2-P	* W3-P
Grayscale		* W1-Gray	* W2-Gray	* W3-Gray

4.2.1 Selecting Structuring Element

We use the three disc-like structuring elements to illustrate how the size of structuring element affects the performance of edge detection. Figure 4-3 shows all flat structuring elements used for experiments in this chapter. Figure 4-4 shows the results of processing W2-I using all three discs and the open-closing operation. Threshold level for each case is determined by the highest value which ensures the closure of knot edges. Without the transformation smaller structuring elements would yield more background as well as object edges. However Figure 4-4c shows that background edges are almost completely removed while object edges remain virtually independent of the structuring element size. This result verifies the

claims of Section 3.3, since W2-I is the *best* component for the NAD-CVM mapping.

Next we use structuring element rec(x,y), since the blue stain in Wood 1 appears in horizontally strapped fashion. Figure 4-5 shows the results of filtering W1-J using rec(9,5) and rec(13,5) in the process of open-closing. Comparing Figures 4-5a and 4-5b, we note that horizontally longer structuring element rec(13,5), gives better detection of horizontal edges. However neither structuring element provides an acceptable result.

The results of filtering W1-J using disc-like structuring elements are shown in Figure 4-6. Obviously the disc-like structuring elements provide better results than the rectangular structuring elements. This is due to the fact that the edges consist of not only horizontal lines, but various slopes as well. The discussion about selecting the disc size for Figure 4-4 holds also for this case, with Figure 4-6b showing better results than 4-6a.

As an auxiliary experiment, we tested W1-J using various dome-like structuring elements to improve the quality of edge detection. Figure 4-7 shows the results of processing W1-J using dome-like structuring elements with disc 4 supporting region. Due to the fitting property of opening and closing operations, a dome-like structuring element leaves more details which do not fit flat structuring element, disc 4. In general, the higher the dome is, the more details are preserved. Consequently, we can detect more edges in object region as well as background region using dome-shaped structuring elements. Again, using NAD-CVM mapping leads to selective morphological operations on object and background as shown in Figure 4-7. We note here that too high dome may worsen the results, since almost every detail fits a part of dome, thus canceling the effectiveness of open-close smoothing.
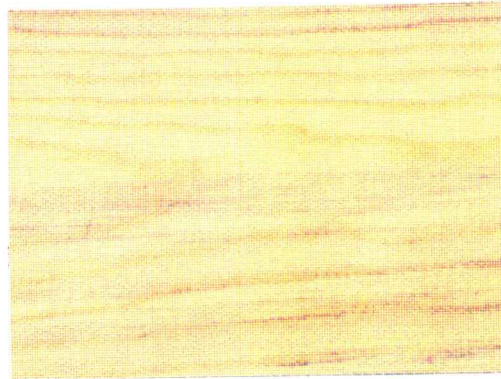
(a)
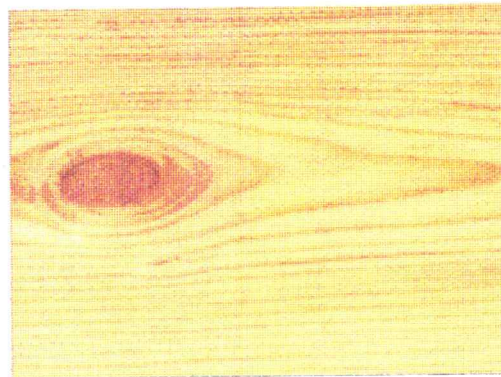
(b)
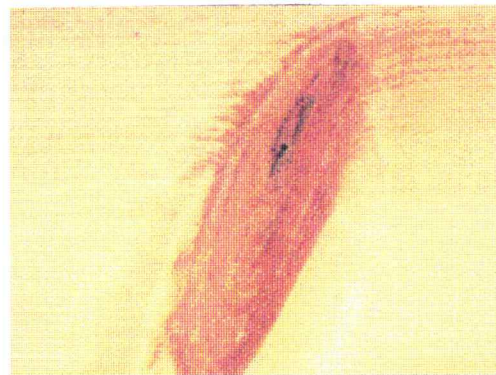
(c)

Figure 4-2 Original color images : wooden boards with (a) blue stains (Wood 1), (b) an intergrown knot (Wood 2), and (c) a spike knot (Wood 3).
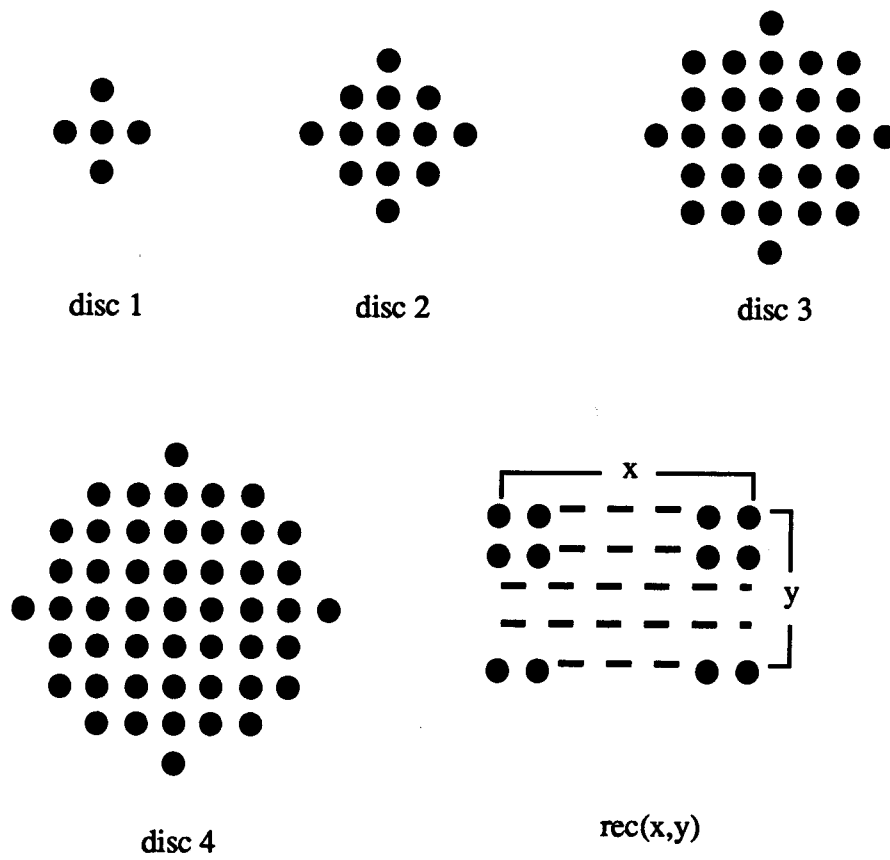
Figure 4-3 Structuring elements.
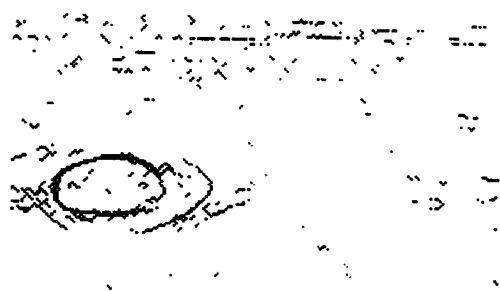
(a)
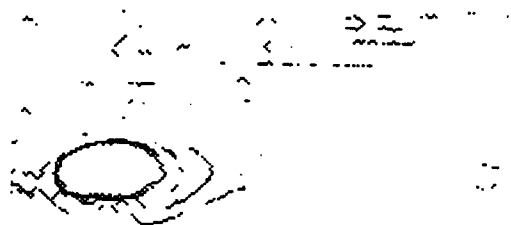
(b)

(c)

Figure 4-4 Results of processing W2-I using (a) disc 1, (b) disc 2, (c) disc 3.
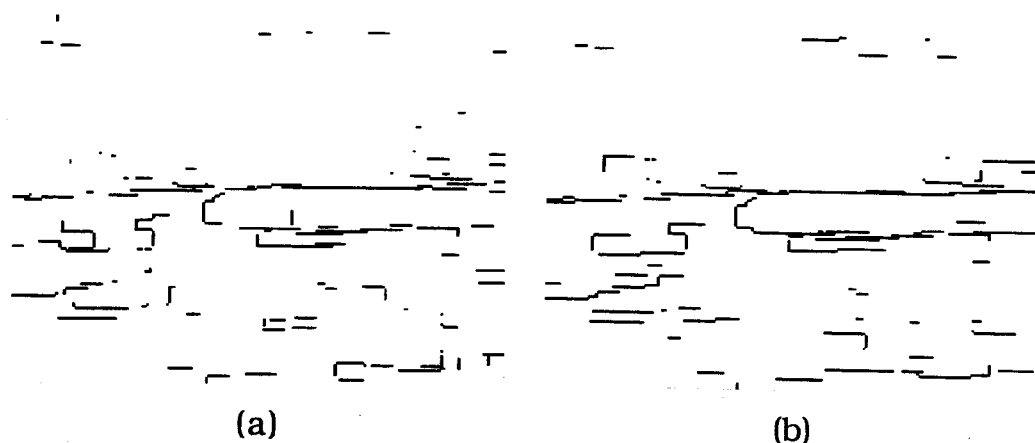
(a)   (b)

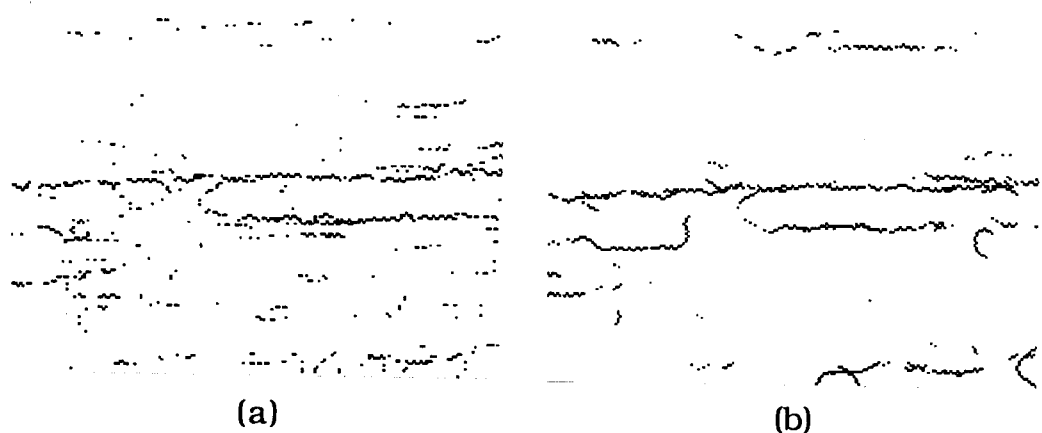Figure 4-5 Results of processing W1-J using (a) rec(9,5), (b) rec(13,5).

(a)   (b)

Figure 4-6 Results of processing W1-J using (a) disc 3, (b) disc 4.

(a)   (b)
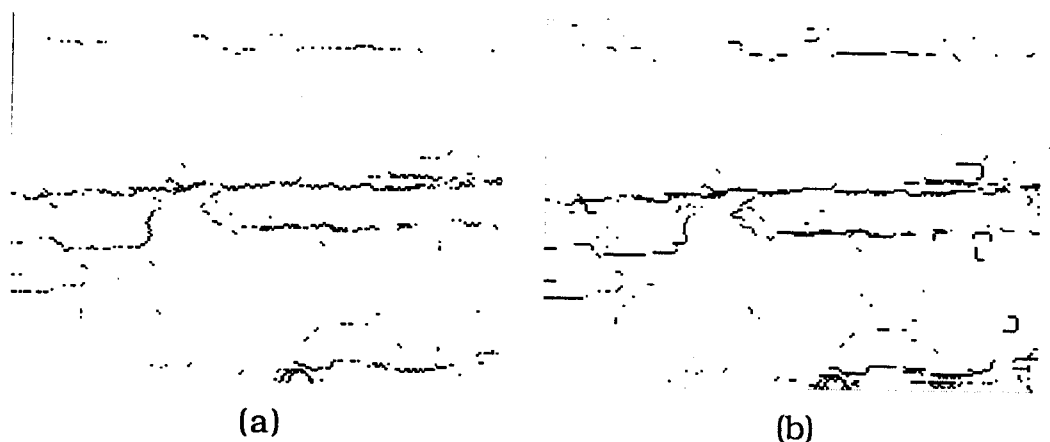
Figure 4-7 Results of processing W1-J using (a) dome with height 5, (b) dome with height 8.

4.2.2 Comparison of Various Mapping Methods

In this section we compare the results of processing various color components (Table 4.1) using the structuring elements discussed in Section 4.2.1. We begin with Figures 4-8, 4-9, 4-10, and 4-11 which show each component transformed by NAD-CVM, NAD-CRM, and total principal component method for Wood 3. All images shown are linearly scaled in amplitude to provide a good visualization.

Figures 4-9a, 4-10a, and 4-11a show the *best* components selected for the object. We cannot clearly identify the advantages and disadvantages of each mapping from these figures. However, through the experiment, we show the advantage of NAD-CVM method to separate the object from the background.

Figures 4-12 and 4-13 show the results of processing Wood 2 and Wood 3, respectively. Since even grayscale processing of Wood 2 and Wood 3 provide somewhat reasonable results as shown in Figures 4-12e and 4-13e, our discussion is focused on the performance of the background (wood grains) edge removal in each component. The results show that NAD-CVM and NAD-CRM mapping methods yield better performance than any other component processing. Since Wood 2 and Wood 3 have some amount of sample mean vector discriminant potential between object and background, NAD-CVM and NAD-CRM behave similarly. Table 4.2 shows the transformation vectors (c_i of matrix C in Equation (2-29)) for NAD-CVM and NAD-CRM methods. The similarities between those vectors further explain similar behavior of the two mappings for Wood 2 and Wood 3.

Table 4.2 Transformation vectors for NAD-CVM and NAD-CRM methods.

	Wood 1	Wood 2	Wood 3
NAD-CVM	$[0.209 -0.155\ 0.088]^T$	$[0.112 -0.154\ 0.060]^T$	$[0.045 -0.066\ 0.021]^T$
NAD-CRM	$[0.125 -0.183\ 0.050]^T$	$[0.089 -0.138\ 0.055]^T$	$[0.040 -0.058\ 0.018]^T$

Figure 4-14 shows the result of processing Wood 1. Only Figure 4-14b shows connected edges of blue stain strip. Any other component does not show a good detection of blue stain, while suppressing the wood grains in background. The comparison of NAD-CVM and NAD-CRM is very interesting in this case. Since the sample mean vectors of the object and the background are located closely, NAD-CRM mapping attempts to get a large sample mean vector discrimination as well as sample variance discrimination. The NAD-CVM tries to obtain large sample variance ratio of the object to the background. The resulting transformation vectors for Wood 1 are shown in Table 4.2, and as we expected they are dissimilar. As shown in Figure 4-14c, this is the case when the NAD-CRM fails. However NAD-CVM does not provide a perfect result either. Figure 4-14b shows that lighter and smaller stains, which are located below the main blue stain strip, could not be detected. In conclusion we state that the NAD-CVM method comes the closest to solving the blue stain detection problem.

(a)

(b)
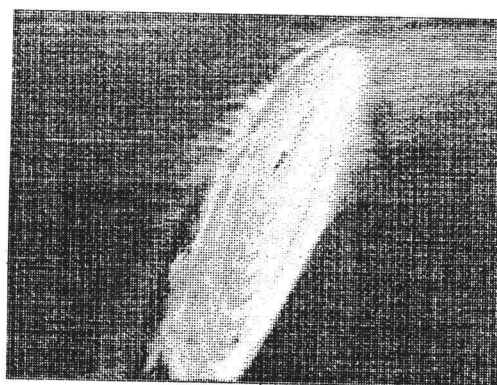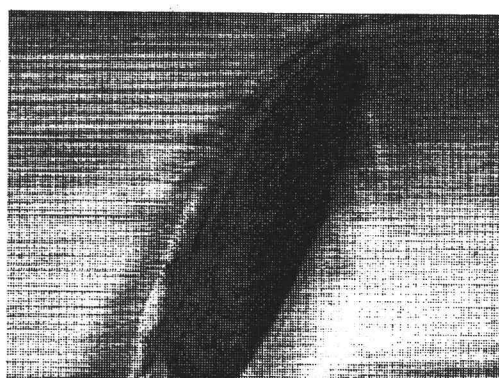
(c)

Figure 4-8 Original RGB components of Wood 3 : (a) R, (b) G, (c) B.

(a)

(b)

(c)

Figure4-9 Components transformed by NAD-CVM method: (a) W3-H, (b) W3-I, (c) W3-J.
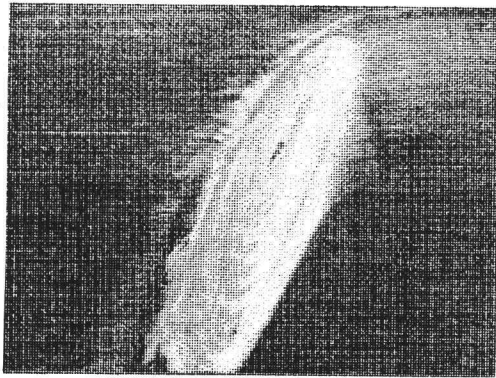
(a)
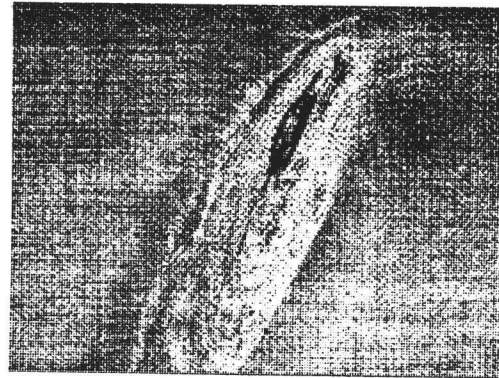
(a)
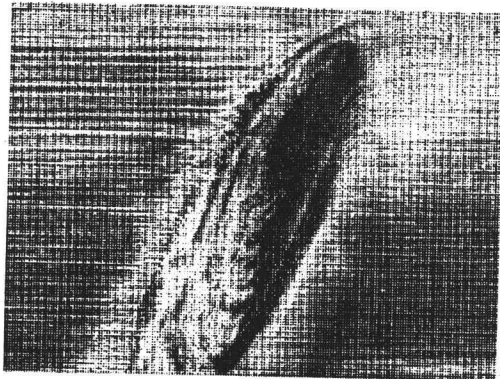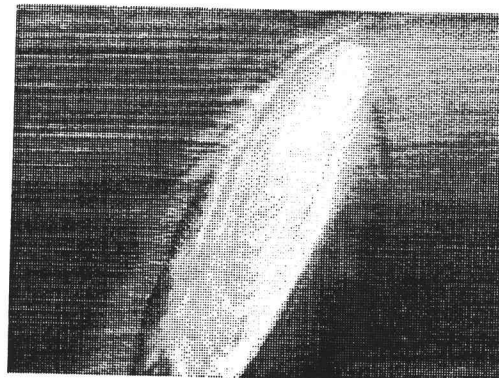

(b)
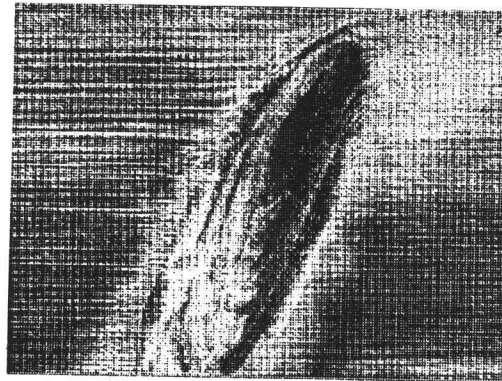
(b)
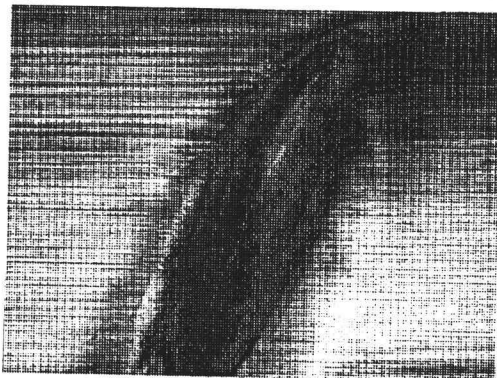
(c)

(c)

Figure4-10 Components trans-
formed by NAD-CRM method:
(a) W3-K, (b) W3-L, (c) W3-M.

Figure4-11 Components trans-
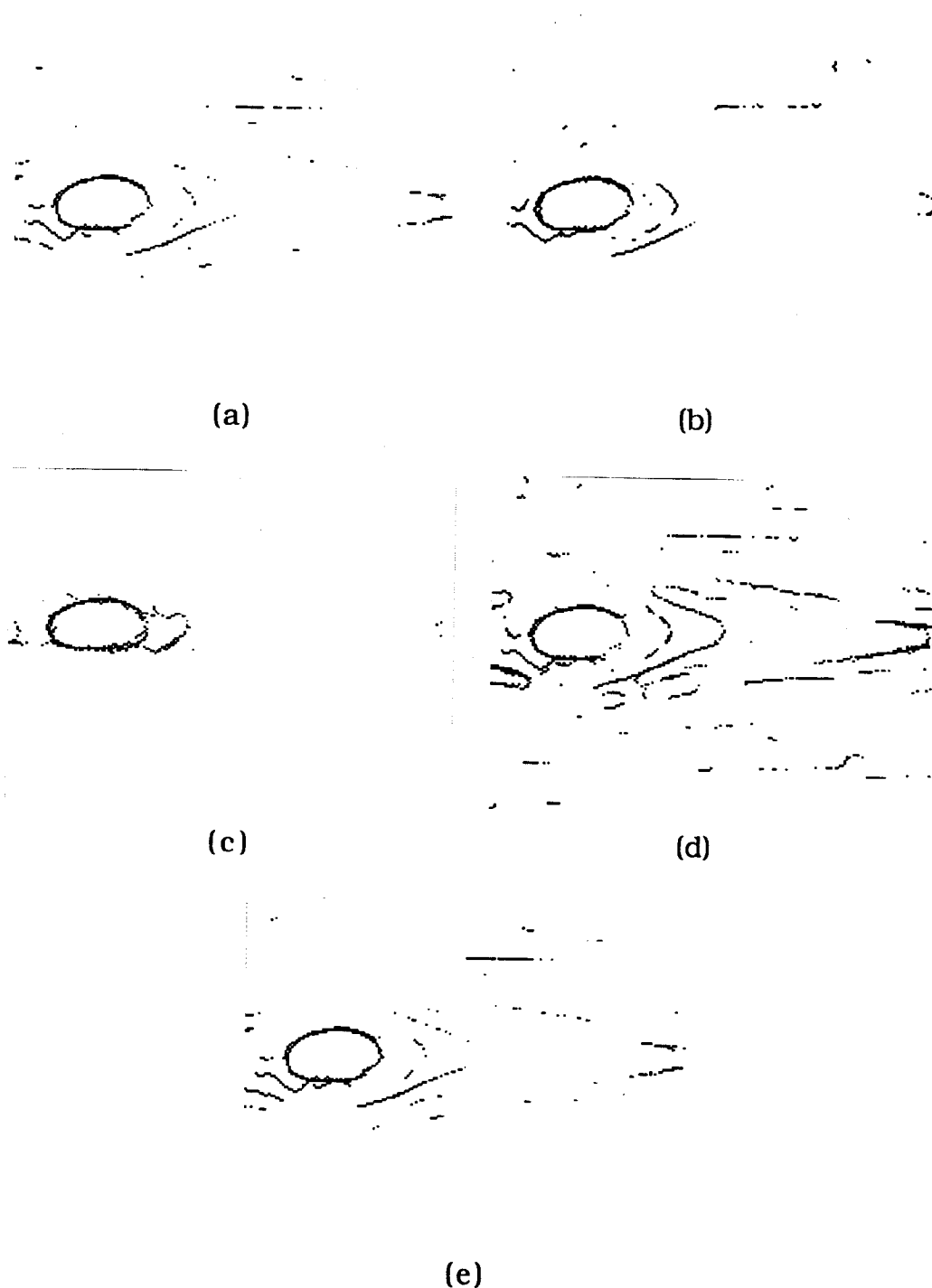formed by K-L method:
(a) W3-N, (b) W3-O, (c) W3-P.

(a)

(b)

(c)

(d)

(e)

Figure4-12 Results of processing Wood 2 for the components: (a) W2-G, (b) W2-I, (c) W2-M, (d) W2-P, (e) W2-Gray.

(a)

(b)

(c)

(d)

(e)
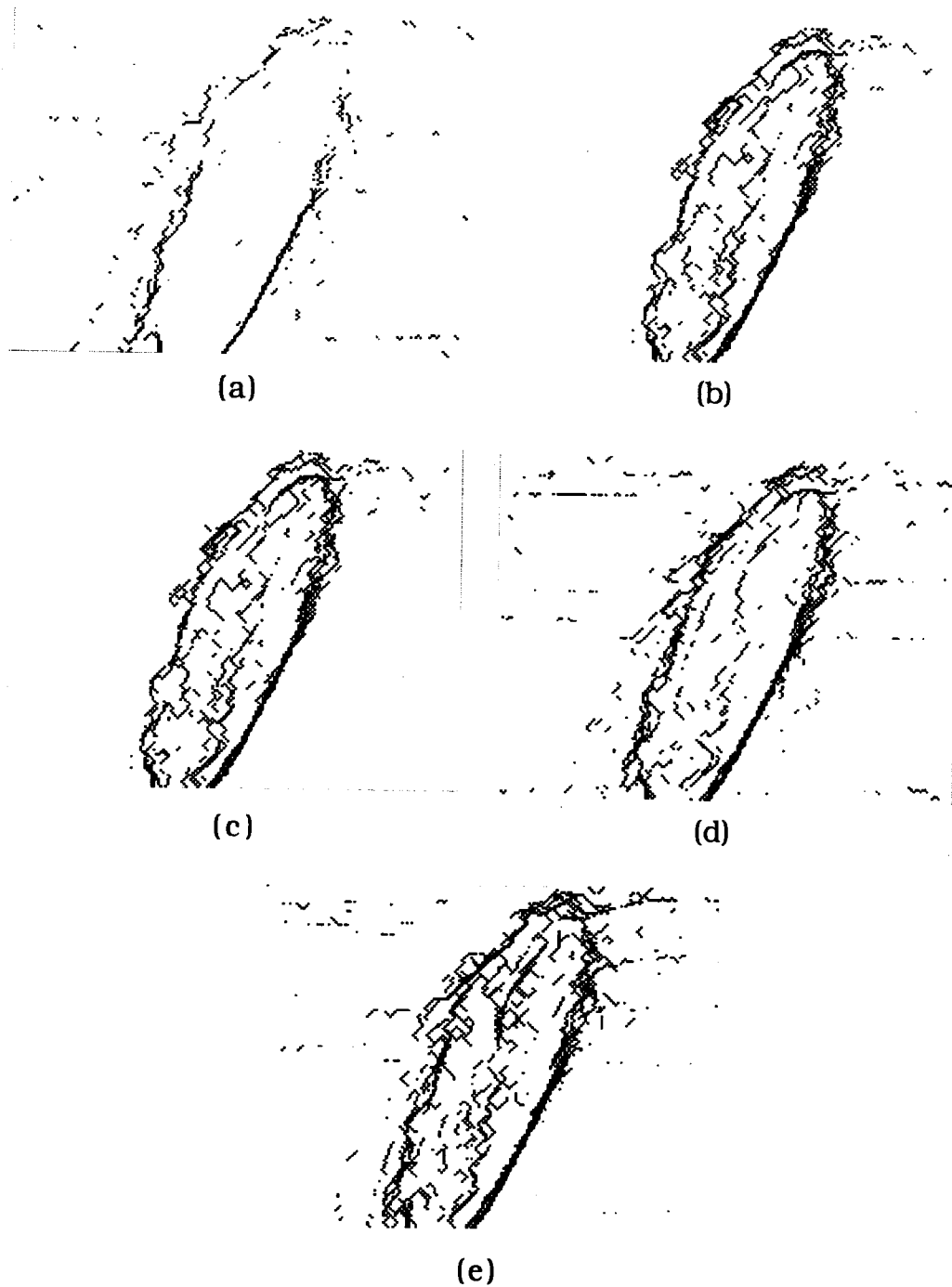
Figure4-13 Results of processing Wood 3 for the components: (a) W3-B, (b) W3-H, (c) W3-K, (d) W3-P, (e) W3-Gray.

(a)

(b)

(c)

(d)

(e)
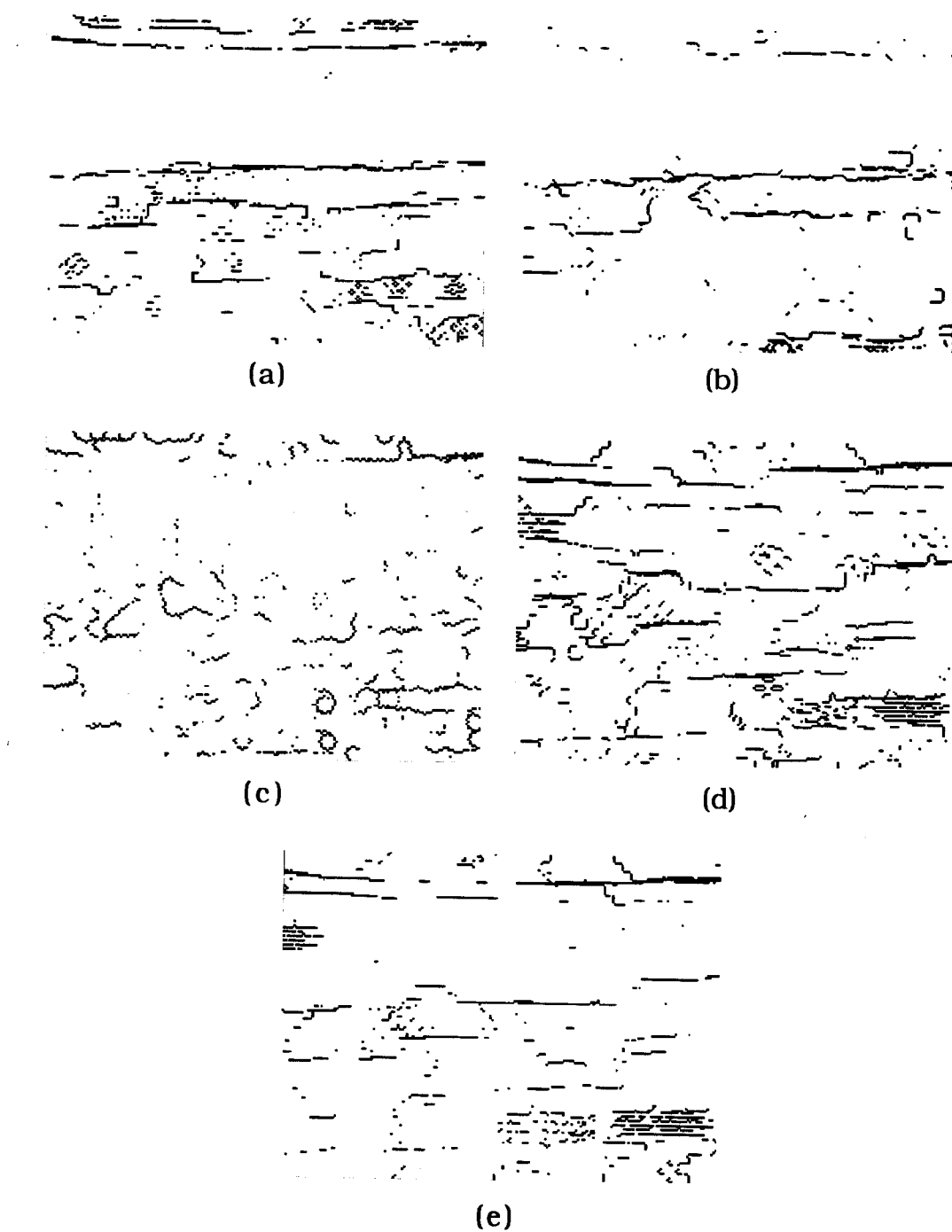
Figure4-14 Results of processing Wood 1 for the components: (a) W1-R, (b) W1-J, (c) W1-L, (d) W1-P, (e) W1-Gray.

4.2.3 Grayscale, Color and Mapping Method

In this section, we discuss the need for color image processing and mapping methods. Figures 4-15a and 4-15b show two thresholded images of grayscale Wood 1 image (W1-Gray). Figure 4-15a shows a threshold level which removes not only wood grains in the background but also a large part of the blue stain. Figure 4-15b shows the result of using slightly higher threshold level. Comparing the two images, we note that graylevels of the object (blue stain) and the background (wood grains) are mixed, so grayscale image can not be used for detecting blue stains. Furthermore the blue stain texture is almost the same as that of background grains. We consequently turn our attention to color image analysis. However the RGB components by themselves do not provide a good threshold separation as shown in Figures 4-16 and 4-17. Figure 4-17 is a set of histograms of R, G, and B components of Wood 1, and shows that blue and green components can not be separated by thresholding while red component may provide a better result. In Figure 4-16 it is shown that thresholding the red component provides the best result among all three components. Here we note that the color of "blue" stain is not pure blue, but a combination of red, green, and blue. The color of grain is also a combination of red, green, and blue with similar blue and green-levels to those of the blue stain. Thus we conclude that the difference of color between blue stain and wood grain is mainly caused by red rather than blue or green component. Still the result of thresholding the red component is unacceptable, since the grains in the upper part of image cannot be removed. Referring to the discussion in Section 4.2.2, a transformation of the RGB components is necessary and NAD-CVM mapping provides the required separation.
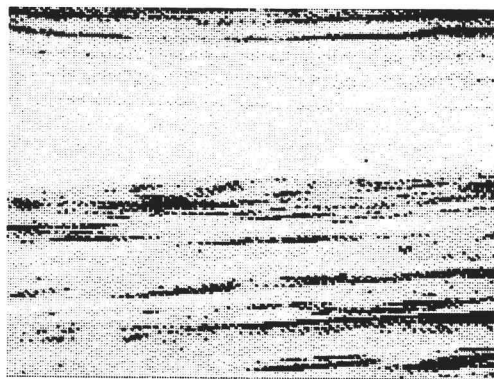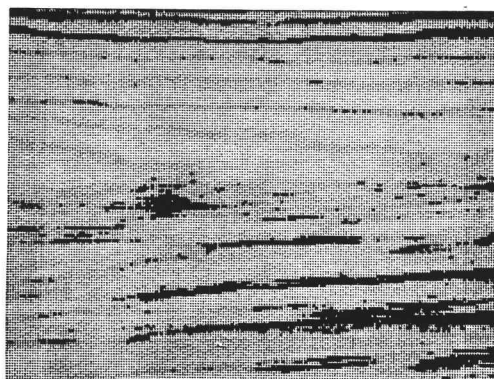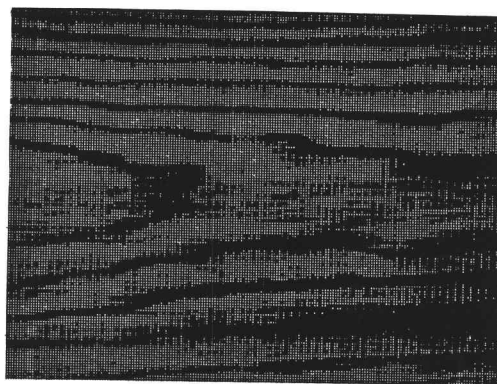
(a)

(a)

(b)
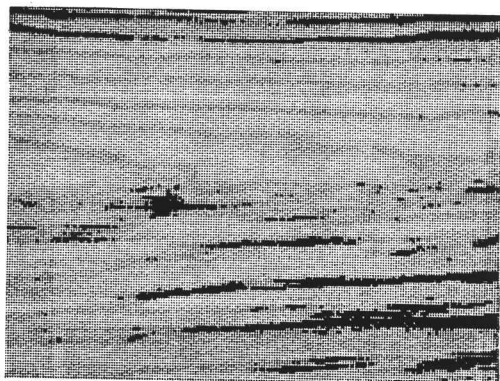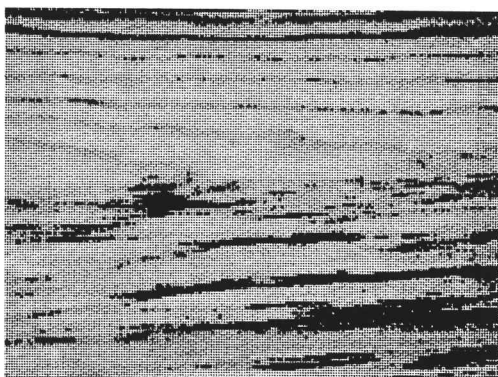
(b)

Figure 4-15 Thresholded images
of W1-Gray : threshold level at (a)
190, (b)195.

(c)

Figure 4-16 Thresholded
images of (a) W1-R, (b) W1-G,
(c) W1-B

R-Total
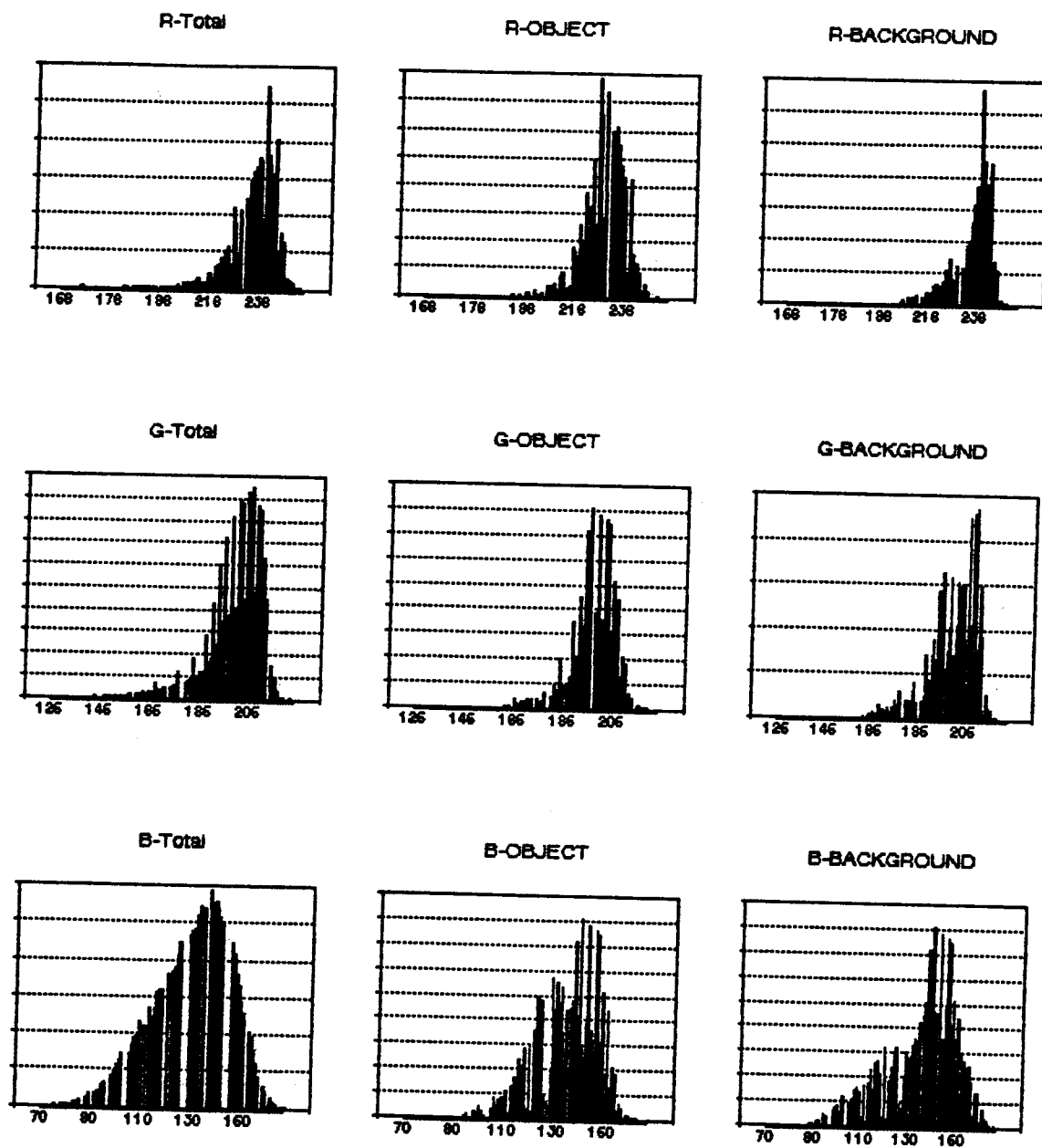
R-OBJECT

R-BACKGROUND

168 176 196 216 236

168 176 196 216 236

168 176 196 216 236

G-Total

G-OBJECT

G-BACKGROUND

126 145 165 185 205

126 145 165 185 205

126 146 166 186 205

B-Total

B-OBJECT

B-BACKGROUND

70 90 110 130 160

70 90 110 130 160

70 90 110 130 160

Figure4-17 Histograms of R, G, B components of Wood 1 image.

4.3 Experiments on Remotely Sensed Ocean Data and Color Mandrill Picture

To show various applications of algorithms discussed in section 4.1.2, a remotely sensed ocean surface temperature data and a mandrill picture were used as the test images.
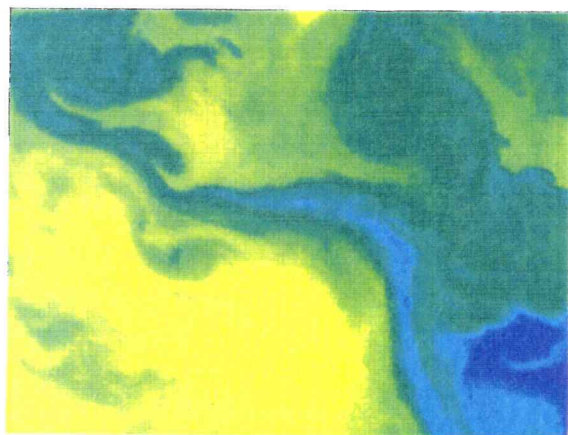
Figure 4-18a shows a pseudo-colored original ocean data. The data represents remote measurements of ocean surface temperature, which in turn is related to the plankton concentration. In this experiment, morphological edge detector with open-close smoothing provides simple boundary between high temperature region and low temperature region. Such a boundary can be further used for analyzing movement of boundaries of successively measured data sets. In this application, we use relatively large structuring element (disc 4), and large threshold level. Figures 4-18b, 4-18c, and 4-18d show the results of edge detection in the form of binary images for : no open-close smoothing (Figure 4-18b), open-close smoothing by flat structuring element disc 4 (Figure 4-18c), and open-close smoothing by dome-like structuring element with disc 4 support (Figure 4-18d). It should be noted that Figure 4-18d represents the closest result to the intended one however the edges are slightly underdetected. Comparison of the three figures should focus on the effectiveness of open-close smoothing. Consequently Figure 4-18b shows small, unwanted details and Figure 4-18c shows underdetected edges.

The above results indicate that analyzing ocean data is a good application for selective filtering property of morphological filters. Assuming that we need to analyze multi-variate ocean data, measured by multi-spectral sensors, the above result and availability of consistent background sample covariance matrix justify application of NAD-CVM method and morphological filtering.

Figure 4-19 shows color mandrill picture which is tested for various coordinate transformation methods. We selected this picture since it represents a very complex color image and thus is a very difficult test for proposed here techniques. For NAD-CVM and NAD-CRM methods mandrill's nose was assigned as the object, and the rest

as the background. Figures 4-20, 4-21, and 4-22 show the results of NAD-CVM, NAD-CRM, and K-L transformation methods, respectively. Figure 4-20a and 4-20c represent the best and the poorest components for the object, respectively. They also show maximum sample variance ratio separation of the object and the background. Figure 4-20b and 4-20d are binary edges of Figure 4-20a and 4-20c, respectively, obtained by morphological edge detector with open-close smoothing. They confirm the properties of NAD-CVM and selective morphological filtering. The results of NAD-CRM (Figure 4-21) are similar to those of NAD-CVM except that Figure 4-21d shows better object boundary and less background details. These differences between NAD-CVM and NAD-CRM methods agree with the discussion of Chapter II and Section 4.2.2. Figure 4-22 shows the results of K-L method. Figure 4-23a, 4-23b, and 4-23c show the binary edge images of R, G, B components, respectively. Comparing Figures 4-22 and 4-23, we note that the first principal component (Figure 4-22b) shows the most details in both, the object and the background.
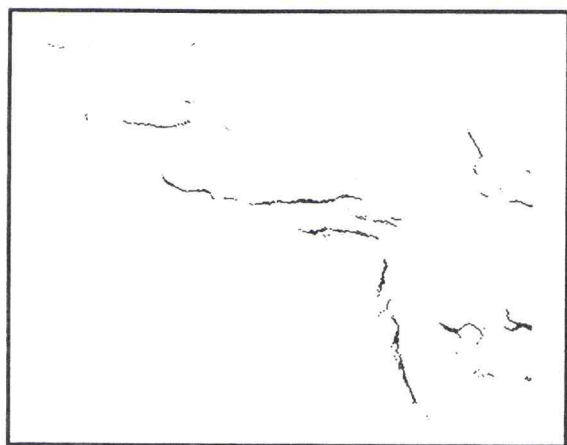
We use the results of this section in Chapter 5, while discussing the thesis contributions.
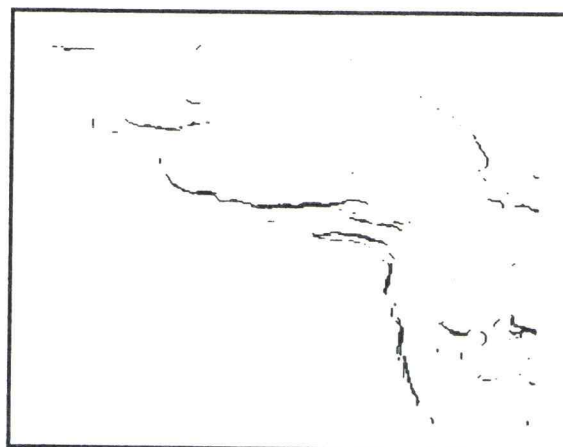
(a)
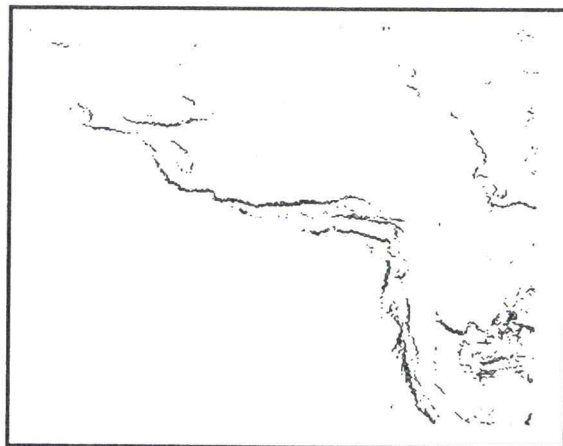
(b)

(c)

(d)

Figure 4-18 (a) Pseudo-colored original ocean data, (b) edge image of (a) without using open-close smoothing, (c) edge image of (a) with open-close smoothing by flat disc 4, (d) edge image of (a) with open-close smoothing by dome shape structuring element with disc 4 support.
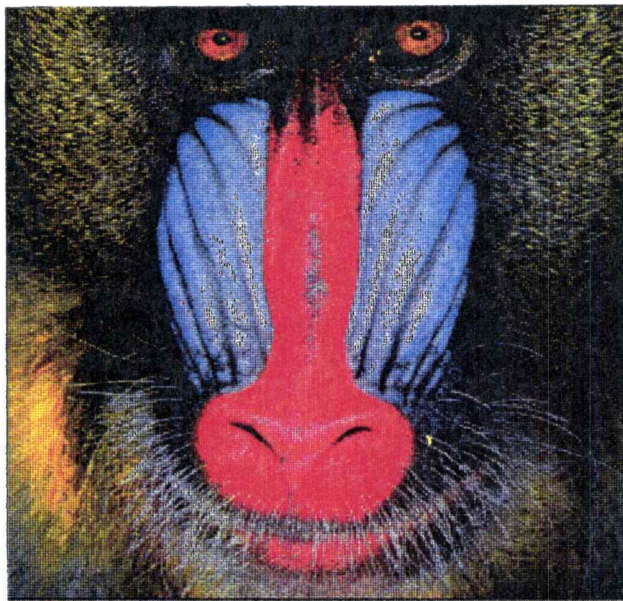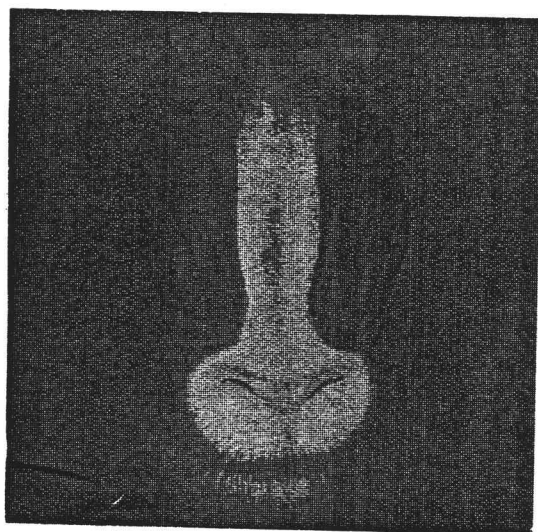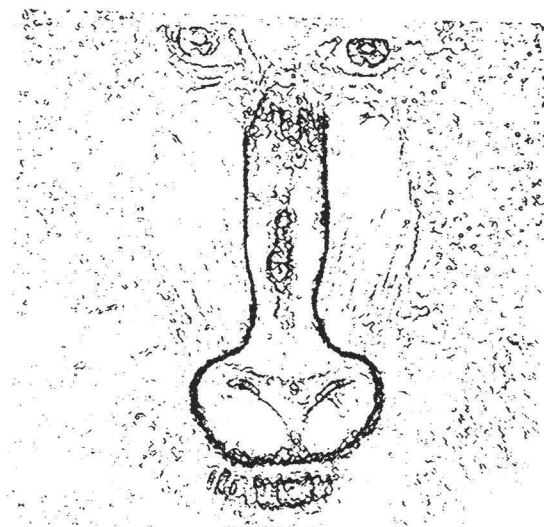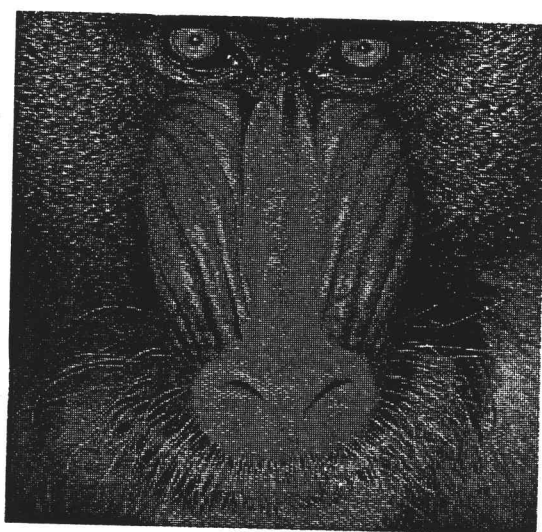
Figure 4-19 Original mandrill image.

(a)

(b)

(c)

(d)

Figure 4-20 Results of NAD-CVM method : (a) the best component for the object, (b) edge image of (a), (c) the best component for the background, (d) edge image of (c).
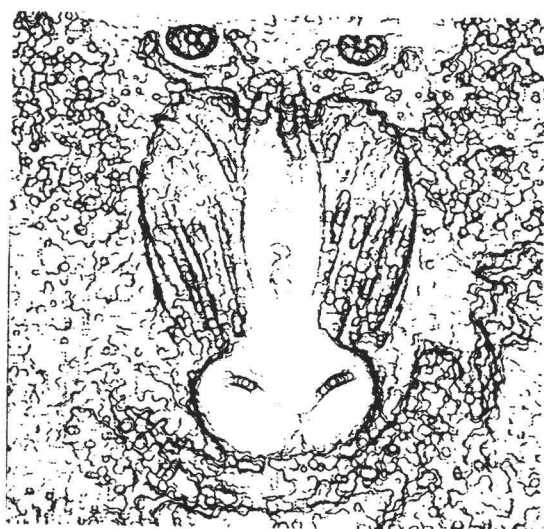
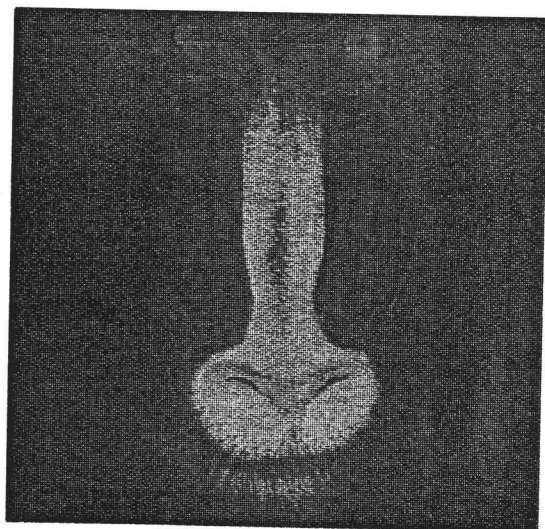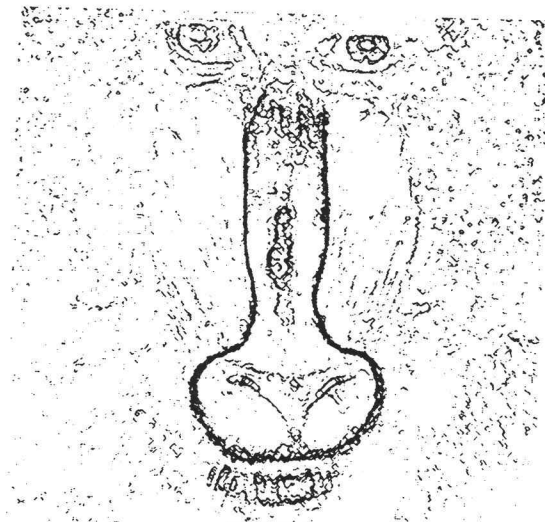(a)

(b)

(c)

(d)

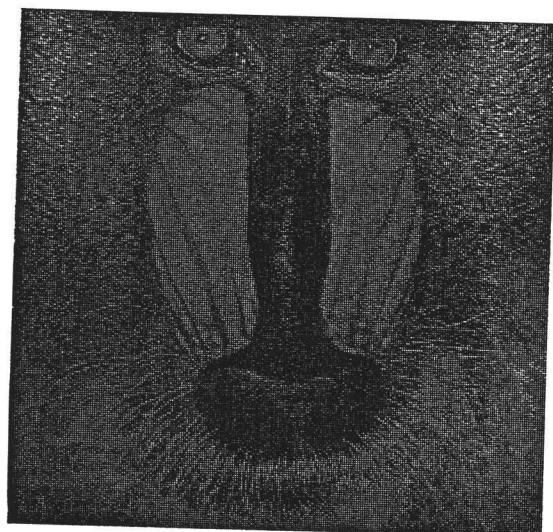Figure 4-21 Results of NAD-CRM method : (a) the best component for the object, (b) edge image of (a), (c) the best component for the background, (d) edge image of (c).

(a)

(b)

(c)

(d)
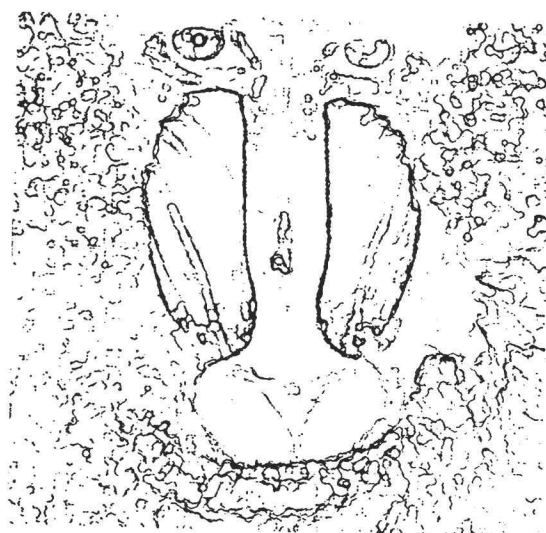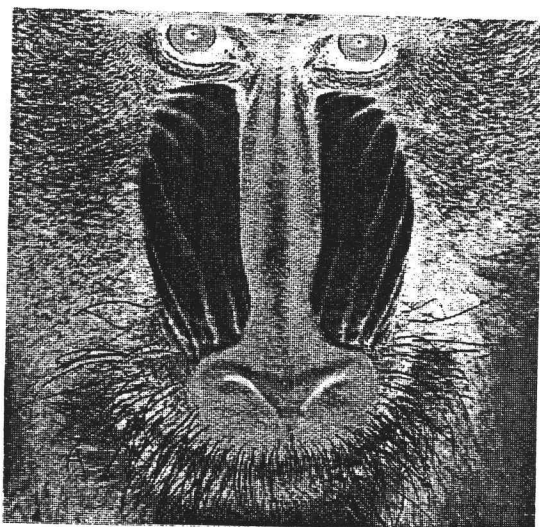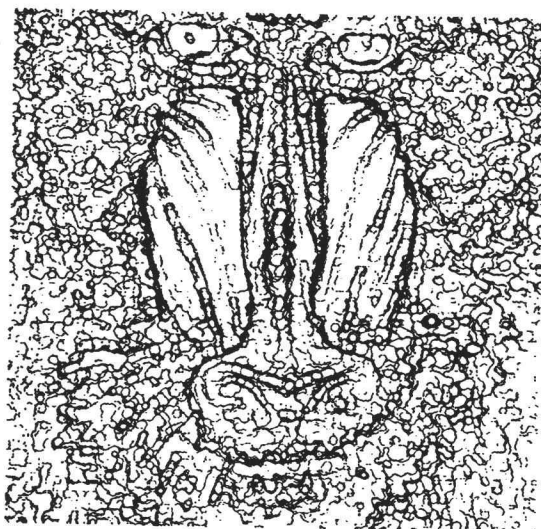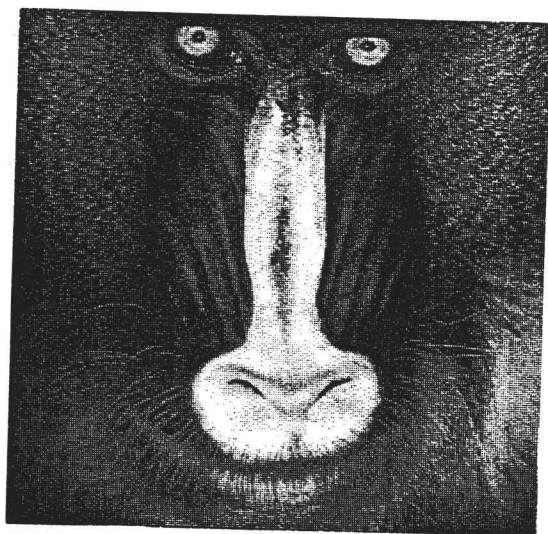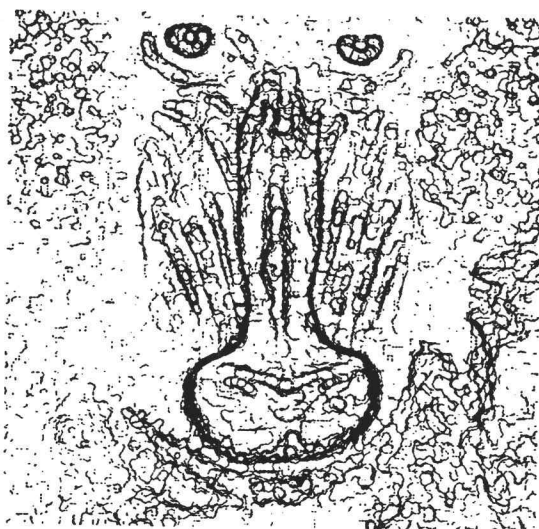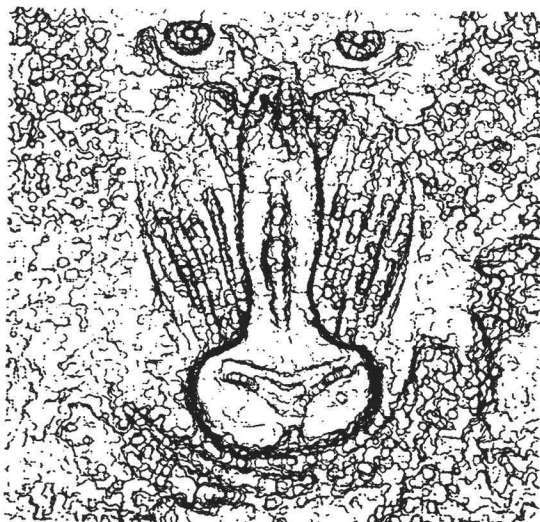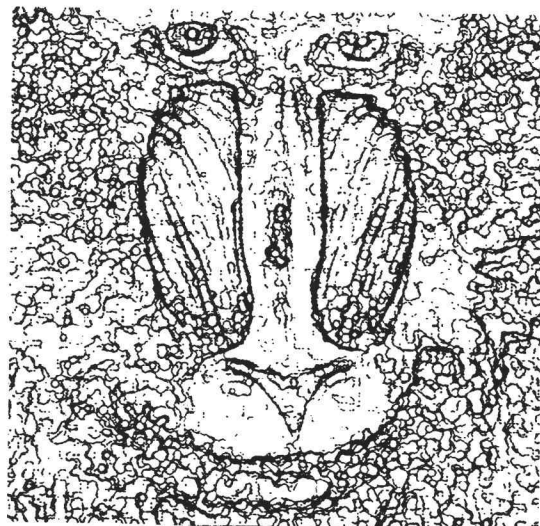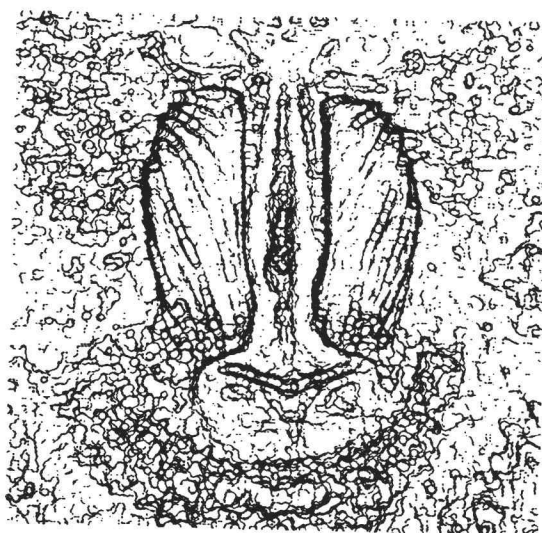
Figure 4-22 Results of K-L method : (a) the first principal component, (b) edge image of (a), (c) the second principal component, (d) edge image of (c).

(a)

(b)

(c)

Figure 4-23 Edge images of (a) R-component, (b) G-component
(c) B-component.

5. CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

The goal of this study was to apply morphological filtering to the process of multi-variate signal analysis. For reason of the lack of ordering in vector spaces, direct approaches to the extension of the theory of single-component morphological filters to multi-variate cases have been failures. Thus, in the current study, a two-stage processing technique, consisting of the maximum separation of the object from background feature and separate morphological filtering of each component, has been proposed.

To separate the object from the background, the NAD-CVM mapping method for multi-component signal transformation was considered. The morphological measure interpretation of sample variance indicated that the NAD-CVM mapping constituted an excellent preprocessing tool for morphological filtering of multi-variate signals. An unsupervised implementation of NAD-CVM was tested experimentally and the theoretical results were confirmed.

5.1.1 Mapping Methods

The study of mapping methods is popular in the pattern recognition area, particularly for discriminating among individual observations between two classes. In the current study, a mapping method for the separation of local data structures was used as a complement to the nature of morphological (neighborhood set) operations. The fact of separation maximized the ratio of variances between two classes. This approach is novel in the sense that mapping techniques provide for maximum local data structure

separation as a preprocessing for multi-variate signals prior to morphological filtering. The normalization and simultaneous diagonalization of sample covariance matrices (NAD-CVM) was identified as a promising mapping technique. The desirable properties of NAD-CVM mapping were studied theoretically, then verified experimentally.

5.1.2 Mathematical Morphology in Relation to Mapping Methods

To relate the NAD-CVM separation quality to a morphological measure, the *average height* of opening by a structuring element was defined. It was demonstrated that the first absolute central moment was a reasonable substitute for sample variance, and was related to the *average height* of opening. In addition, the increase (or decrease) of sample variance by the application of NAD-CVM mapping increased (or decreased) *average height* of opening. This relationship between sample variance and *average height* was used to select a proper structuring element for a morphological filtering algorithm, a process which enabled the elimination of small background details while preserving discontinuities within the object. Thus, the results of this investigation have justified the degree of effort required to construct a two-stage morphological processing algorithm for multi-variate signals.

5.1.3 Generalization of Mathematical Morphology

Although the study of mapping methods provides a constructive means for morphological filtering solutions of multi-variate signals, it would still be more appropriate to generalize uni-variate mathematical morphology. This generalization process was attempted by defining an umbra and a top-surface of multi-variate signal based upon partial-ordering. In general, for a partial-ordering approach, *multi-variate morphology* is reduces to processing each signal component separately. However, as illustrated in this study with a simple example, if an order in the signal range can be defined, then

generalization of the umbra concept will provide new and intuitive results for *multi-variate morphology*.

5.1.4 Experimental Results

Since the NAD-CVM can be applied to any two-class, multi-variate signal analysis problem, there are a number of applications of the technique considered in this study. As a test of algorithm performance, a morphological edge detector with open-close smoothing was simulated. The edge detector, in combination with the NAD-CVM unsupervised mapping method, performed well while detecting defects in color board images. Promising results were also obtained for the detection of blue stains in wooden boards, an image analysis problem which cannot be solved using uni-variate analysis. In addition, edge detection was also tested for ocean images and for a complex color image of a mandrill. The result of these tests further verified the basis for the theoretical claims presented in this investigation.

5.2 Suggestions for Future Study

It has been demonstrated that finding a proper partial-ordering in vector spaces should be the first step toward generalization of single-component morphology to multi-variate signals. Such a partial-ordering can be obtained by using the concepts of convex, cone-like set inclusions. Simultaneously, the development of a multi-variate version of one-to-one correspondence between signal and its umbra, based upon proper modification of minimum and maximum operations, should be further developed.

Linear transformation techniques, as a vehicle to generalize morphological filtering, was investigated. Since nonlinear techniques may be effective for some specific cases, it would be of interest to investigate nonlinear mapping properties for preprocessing of multi-

variate signals. This study is further justified by the fact that the morphological filters are nonlinear.

Average height was introduced as a consideration in the relation of sample variance and structuring element size for the deterministic observations. Based upon the assumption of random observations, a theory for the relation of the original data statistics to the (morphologically) filtered data should be developed. The second order statistics of morphological filtering have been examined by Zhu [43]. Using stochastic image models, sample means and variances can be replaced by their statistical counterparts, thus enabling a precise and formal mathematical analysis. This type of statistical framework can also be used to define an objective index for the evaluation of the performance of the algorithm proposed in the current study.

BIBLIOGRAPHY

[1] G. Biswas, A.K. Jain, and R.C. Dubes, "Evaluation of projection algorithms," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. PAMI-3, pp.701-708, 1981.

[2] I. Destival, "Mathematical morphology applied to remote sensing," *Acta Astronautica*, Vol.13, No.6/7, pp.317-385, 1986.

[3] D. Foley and J. Sammon, "An optimal set of discriminant vectors," *IEEE Trans. on Computers*, Vol. C-24, No. 3, pp.281-289, 1975.

[4] D.A. Forsyth, "A Novel Approach to Color Constancy," *Proc. of 2-nd Intl. Conf. on Computer Vision*, Tampa, Fl., pp. 9-18, Dec. 1988.

[5] K. Fukunaga and W.L.G. Koontz, "Application of the Karhunen-Loeve expansion to feature selection and ordering," *IEEE Trans. on Computers*, Vol. C-19, No. 4, pp.311-318, 1970.

[6] E. Gelsema and R. Eden, "Mapping Algorithms in ISPAHAN," *Pattern Recognition*, Vol. 12, pp.127-136, 1980.

[7] R. Gershon, "Aspect of Perception and Computation in Color Vision," *Computer Vision, Graphics, and Image Processing*, vol. 32, 1985, pp.244-277.

[8] C.R. Giardina and E.R. Dougherty, *Morphological Methods in Image and Signal Processing*, Prentice-Hall, 1988.

[9] R. Haralick, S.R. Sternberg, and X. Zhuang, "Image analysis using Mathematical Morphology," *IEEE Tran. on Pattern Analysis and Machine Intelligence*, Vol. PAMI-9, No.4, July, 1987.

[10] G. Healey, "A Color Reflectance Model and its Use for Segmentation," CH2664-1/IEEE, 1988, pp.460-466.

[11] G. Healey and T.O. Binford, "A Color Metric for Computer Vision,"*Proceedings of Image Understanding Workshop*, Cambridge, MA, April 1988, pp.854-861.

[12] A.K. Jain, *Fundamentals of Digital Image Processing*, Prentice-Hall, 1989.

[13] J. Kittler and P. Young, "A new approach to feature selection based on Karhunen-Loeve expansion," *Pattern Recognition*, Vol. 5, pp.335-352, 1973.

[14] G. Klinker, S. Shafer, and T. Knade, "Image Segmentation and Reflection Analysis Through Color, *Proceedings of Image Understanding Workshop*, Cambridge, MA, April 1988, pp.838-853.

[15] J. Lee, R. Haralick, and L. Shapiro, "Morphologic edge detection," *Proc. of 8th Int. Conf. on Pattern Recognition*, Paris, pp.369-373, Oct. 1986.

[16] P. Maragos, *A Unified Theory of Translation-Invariant Systems with Applications to Morphological Analysis and Coding of Images*, Unpublished Doctoral Dissertation, School of Elec. Eng., Georgia Institute of Technology, GA, July 1985.

[17] P. Maragos and R.W. Schafer, "Morphological Filters-Part I : Their Set-Theoretic Analysis and Relations to Linear Shift-Invariant Filters," *IEEE Trans. Acoustic, Speech, and Signal Processing*, vol.35, Aug. 1987, pp.1153-1169.

[18] G. Matheron, *Random Sets and Integral Geometry*, John Wiley and Sons, New York, 1975.

[19] Y. Ohta, *Knowledge-based interpretation of outdoor natural color scenes*, Pitman Publishing Inc., 1985.

[20] W.K. Pratt, *Digital Image Processing*, John Wiley and Sons, 1978.

[21] M. Savoji and R. Burge, "Note : On different methods based on the Karhunen-Loeve Expansion and used in image analysis," *Computer Vision, Graphics, and Image Processing*, Vol. 29, pp. 259-269, 1985.

[22] J. Serra, Image Analysis and Mathematical Morphology, Academic Press, New York, 1982.

[23] W. Sieddlecki, K. Siedlecka, and J. Sklansky, "An overview of mapping techniques for exploratory pattern analysis," *Pattern Recognition*, Vol. 21, No. 5, pp.411-429, 1988.

[24] S.R. Sternberg, "Biomedical Image Processing," *IEEE Computer*, Jan. 1983, pp.22-34.

[25] S.R. Sternberg, "Greyscale Morphology," *Computer Vision, Graphics, and Image Processing*, vol.35, 1986, pp.333-355.

[26] S.S. Wilson, "Theory of Matrix Morphology," Techinical Report No. 39, Applied Intelligence Systems, Inc., Oct. 1989.

[27] C. Wright, E. Delp, and N. Gallagher," Nonlinear target enhancement for the hostile nuclear environment," *IEEE Trans. on*

Aerospace and Electronic Systems, Vol.26, No.1, pp.122-148, Jan. 1990.

[28] *AVS User's Guide*, Stardent Computer Inc., April, 1991.

[29] *Selects-Finish / Commons - Boards : Vol.2*, Western Wood Products Association, 1989.

[30] J. Fehlauer and B.A. Eisenstein, "A declustering criterion for feature extraction in pattern recognition," *IEEE Trans. on Comp.*, Vol. c-27, pp.261-266, 1978.

[31] R.A. Fisher, "The use of multiple measurements in taxonomic problems," *Ann. Eugenics*, Vol. 7, pp. 178-188, 1936.

[32] G. Strang, *Linear Algebra and its Applications (2-nd Ed.)*, Academic Press, 1980.

[33] M. Celenk and S.H. Smith, "Color image segmentation by clustering and parametric histogramming technique," *Proc. of 8th Intl. Conf. on Pattern Recognition*, Paris, pp.883-886, 1986.

[34] M. Celenk, "A color clustering technique for image segmentation," *Computer Vision, Graphics, and Image Processing*, Vol.52, pp.145-170, 1990.

[35] A. Gottschalk and G. Buchbaum, "Information theoretic aspects of color signal processing in visual system," *IEEE Trans. on System, Man, and Cybernetics*, Vol. SMC-13, No.5, Sep./Oct. 1983.

[36] K.C.Gowda, "A feature reduction and unsupervised classification algorithm for multispectral data," *Pattern Recognition*, Vol.17, No.6, pp.667-676, 1984.

[37] Y. Ohta, T.Knade, and T. Sakai, "Color information for region segmentation," *Computer Graphics and Image Processing*, Vol.13, pp.222-241, 1980.

[38] J. Parkkinen and T. Jaaskelainen, "Color representation using statistical pattern recognition," *Applied Optics*, Vol.26, No.19, pp.4240-4245, Oct. 1987.

[39] J. Parkkinen, T. Jaaskelainen, and M.Kuittinen, "Spectral representation of color images," *Proc. on Intl. Conf. on Pattern Recognition*, pp.933-935, 1988.

[40] M. Pietikainen and D. Harwood, "Segmentation of color images using edge-preserving filters," *Advances in Image Processing and Pattern Recognition*, V. Cappellini and R. Marconi (editors), North-Holland, pp.94-99, 1986.

[41] S. Tominaga, "A mapping method for computer vision," *Proc. of 7th Intl. Conf. on Pattern Recognition*, pp.650-652, 1984.

[42] B.A. Wandell, "The synthesis and analysis of color images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. PAMI-9, No.1, Jan. 1987.

[43] F. Zhu, *Stochastic Properties of Morphological Filters*, Doctoral Dissertation, Oregon State University, Corvallis, OR, July 1991.

[44] A. Berman and A. Ben-Israel, "A note on pencils of hermitian or symmetric matrices," *SIAM J. Appl. Math.*, Vol. 21, No. 1, pp. 51-54, July 1971.

[45] K.N. Majindar, "Linear combinations of hermitian and real symmetric matrices," *Linear Algebra and its Applications*, Vol. 25, pp. 95-105, 1979.

[46] W.K. Pratt, "Spatial transform coding of color images," *IEEE Trans. on Communication Theory*, Vol. COM-19, No. 6, Dec. 1971.

[47] R. Nevatia, "A color edge detector and its use in scene segmentation," *IEEE Trans. on System, Man, and Cybernetics*, Vol. SMC-7, No. 11, pp. 820-826, Nov. 1977.

[48] M. Pietikainen and D. Harwood, "Edge information in color images based on histograms of differences," *8-th Intl. Conf. on Pattern Recognition*, Paris, pp. 594-596, Oct. 1986.

[49] R.N. Strickland, C.S. Kim, and W.F. Mcdonell, "Luminance, hue, and saturation processing of digital color images," *Applications of Digital Image Processing IX*, SPIE Vol. 697, 1986.

[50] J. Franklin, *Matrix Theory*, Prentice Hall, 1968.

[51] D. Steinberg, *Computational Matrix Algebra*, McGraw Hill, 1974.

[52] R. Duda and P. Hart, *Pattern Classification and Scene Analysis*, John Wiley and Sons, Inc., 1973.

[53] A. Rosenfeld and A. Kak, *Digital Picture Processing (Vol. 1 and 2)*, Academic Press, 1982.

[54] K. Stuart and S. Ord, *Kendall's Advanced Theory of Statistics, Vol.1: Distribution Theory, (5-th ed.),* Oxford Univ. Press, 1987.

[55] S. Watanabe, *Pattern Recognition - Human and Mechanical,* John Wiley and Sons, Inc., 1985.

[56] E. Oja, *Subspace Methods of Pattern Recognition,* Research Studies Press, 1983.

[57] R. Gnanadesikan, *Methods for Statistical Data Analysis of Multivariate Observations,* John Wiley and Sons, Inc., 1977.

[58] K. Fukunaga, *Introduction to Statistical Pattern Recognition,* Academic Press, 1972.

[59] G. Goulb and C. Van Loan, *Matrix Computations,* John Hopkins Univ. Press, 1983.

[60] E. Cech, *Topological Spaces,* John Wiley and Sons, Inc., 1966.

[61] G. Choquet, *Topology,* Academic Press, 1966.