A COMPUTATIONALLY SIMPLE PROCEDURE FOR IMAGERY DATA COMPRESSION
BY KARHUNEN-LOEVE METHOD

K. SHANMUGAM, Member IEEE and R.M. HARALICK, Member IEEE
University of Kansas Center for Research
Lawrence, Kansas 66044

ABSTRACT

Of the several methods that have been proposed for imagery data compression, the Karhunen-Loeve procedure minimizes the mean square error between the original and reconstructed imagery data. In spite of its optimality property, the Karhunen-Loeve procedure has not been widely used because of its computational complexity. The main difficulty is in the computation of the eigenvectors and the eigenvalues of the covariance matrix of the imagery data since the dimension of the covariance matrix is usually large.

We are presenting in this paper a computationally short procedure for calculating the eigenvalues and eigenvectors of the covariance matrix. We show that the eigenvalues and eigenvectors of the NXN bisymmetric covariance matrix can be obtained from the eigenvalues and eigenvectors of two N/2 x N/2 submatrices. Since the eigenvector calculations are proportional to the third power of the matrix dimension, the proposed procedure reduces the computations by a factor of four.
1. **INTRODUCTION**

Imagery data in general contain a large amount of redundant information because of the high positive correlation between the grey levels of spatially adjacent image elements. Several imagery data compression techniques have been proposed recently for removing this redundant information \( (1,2,3,4,\ldots) \). Of these methods, the principal component method (based on the Karhunen-Loeve expansion) minimizes the mean square error between the original and compressed imagery data.

In the principal components method, the image is first split into a number of small mutually exclusively spatial regions or windows and the grey levels of these regions are treated as \( N \) dimensional vectors \(^1\). The image is then a collection of these vectors. These \( N \) dimensional vectors \( X_1, X_2, \ldots, X_k \) are then projected into some smaller \( r \)-dimensional subspace having maximal variance. In this way the \( N \) components of the original data may be expressed in terms of \( r \) components thus achieving a data compression of \( N/r \).

An optimal basis for the \( r \) dimensional subspace is the set of \( r \) eigenvectors \( V_1, V_2, \ldots, V_r \) corresponding to the \( r \) largest eigenvalues of the sample covariance matrix of \( X_1, X_2, \ldots, X_k \). The reconstructed value of the imagery data in the \( j \)th subimage region (window) is given by

\[
X_j^* = \sum_{i=1}^{r} [V_i^T X_j] V_i
\]

(1)

The principal component procedure minimizes the mean square error

\[
E = \frac{1}{k} \sum_{i=1}^{k} \left\| X_i - X_i^* \right\|^2
\]

(2)

\(^1\)These vectors are assumed to have a mean of zero. If not, the mean vector can be calculated and subtracted from each of these vectors.
The minimum value of $E$ for $r$ projections is given by

$$E_r = \sum_{i=r+1}^{N} \lambda_i$$  \hspace{1cm} (3)$$

where $\lambda_i$ are the $N-r$ smallest eigenvalues of the covariance matrix.

In spite of its optimality, the principal component method has not been widely used because of its computational complexity (5). The main difficulty is in the computation of the eigenvalues and the eigenvectors of the NXN covariance matrix. The window sizes for large images range from 4X4 to 10X10, leading to 16x16 to 100x100 covariance matrices. The calculations for the eigenvalues and eigenvectors of these large sized matrices require considerable amount of computation time and storage.

We are presenting in this paper a computationally short procedure for calculating the eigenvalues and eigenvectors of the covariance matrix. We show that the eigenvalues and eigenvectors of the NXN bisymmetric covariance matrix can be obtained from the eigenvalues and eigenvectors of two N/2 X N/2 submatrices. Since the eigenvector calculations are proportional to the third power of the matrix dimension, the proposed procedure reduces the computations by a factor of four.

II. CONSTRUCTION OF THE SAMPLE COVARIANCE MATRIX

The entries in the covariance matrix of a given image are obtained by calculating the average covariance of all the elements in the image that has the same spatial relationship as the entry being considered. This procedure can best be illustrated using an example.

The 4X4 array shown in Figure 1.a represents a small image whose data are to be compressed. The image elements are labeled from 1 to 16. The size of the window for this example is 2X2 and the arrangement of components of the data vector $X$ within each window is shown in Figure 1.b. For this image, the average covariance array and the covariance matrix are computed as follows.

Each element in the average covariance array (Figure 1.c.) is the average covariance of all the elements in the original image, having the same spatial
A. Original Image

B. Arrangement of Variables within a Window

C. Average Covariance Array

D. Covariance Matrix.
relationship to each other as the element of the covariance array has to the lower center element. For example, the element c in the covariance array is located along the 45° diagonal line from the lower center (reference) element e. Accordingly the entry ε is the average covariance of 9 pairs of similarly spatially related image elements (5,2), (6,3), (7,4), (9,6), (10,7), (11,8), (13,10), (14,11), and (15,12). Other elements of the average covariance array are calculated using similar spatial relationships. The entries in the 4X4 covariance matrix (Figure 1.d) for the image can be obtained from the data contained in the average covariance array. For example, the entry at second row third column of the covariance matrix represents the covariance of x₂ and x₃. Elements x₃ and x₂ are along the 45° diagonal of the window and the element c in the average covariance array bears the same relationship to the lower center (reference) element e. Hence, the entry c is used to fill in the (2,3) element in the covariance matrix. Similarly the remaining entries in the covariance matrix are obtained.

The procedure described in the preceding paragraphs can be used for square windows of any size MXM, with M less than the overall dimension of the image itself. In the remainder of the paper we will restrict our attention to windows of size MXM where M is even.

The sample covariance matrix obtained from the average covariance array has a bisymmetric form. Also, the M²X²M² covariance matrix consists of M submatrices of dimension MXM. These submatrices appear in a bisymmetric form within the covariance matrix. Let us consider the following example to further illustrate the bisymmetric properties of C.

Figure 2.a shows the arrangement of the components of X within a 4X4 window, and the 16X16 covariance matrix of X is shown in a partitioned form in Figure 2.b.

---

2 Def: An NXN matrix A is bisymmetric if and only if:

\[ a_{ij} = a_{ji}, \quad i, j = 1, 2, \ldots, N \quad \text{and} \quad a_{ij} = a_{N+1-i, N+1-j}, \quad i, j = 1, 2, \ldots, N. \]

It follows from this definition that \( a_{ij} = a_{ji} = a_{N+1-i, N+1-j} = a_{N+1-j, n+1-i} \), \( i, j = 1, 2, \ldots, N \).
The 4X4 submatrices $A_1$ represent the covariance of the elements of one row of the window with the elements of another row of the window. The matrices $A_1$, $A_6$, $A_{11}$ and $A_{16}$ represent the covariance of the elements of a row with the elements of the same row. The spatial relationships existing between the elements of row 1 is the same as the spatial relationships between the elements of row 2 or row 3 or row 4. Since the entries in all these matrices are obtained from the average covariance array using spatial relationship between the elements, $A_1$, $A_6$, $A_{11}$ and $A_{16}$ are identical. Next, let us consider the matrices $A_2$, $A_7$, $A_{12}$, $A_5$, $A_{10}$ and $A_{15}$ which represent the covariance between the element of one row of the window with the elements of an adjacent window. For instance, $A_2$ represents the covariance between the elements $x_1$, $x_2$, $x_3$, $x_4$ and $x_5$, $x_6$, $x_7$ and $x_8$. Referring to figure 2.a, the spatial relationships between $x_1$, $x_2$, $x_3$, $x_4$ and $x_5$, $x_6$, $x_7$, $x_8$ and the spatial relationships between $x_5$, $x_6$, $x_7$, $x_8$ and $x_9$, $x_{10}$, $x_{11}$, $x_{12}$. Hence, the entries in $A_2$ and $A_7$ will be the same. Extending this reasoning, it is easy to see that $A_2$, $A_7$, $A_{12}$, $A_5$, $A_{10}$ and $A_{15}$ are identical. Similarly, it can be shown $A_7$, $A_{12}$, $A_9$ and $A_{14}$ are the same and that $A_4$ and $A_{13}$ are also identical. Hence, the form of the covariance matrix becomes

$$
C = \begin{bmatrix}
A_1 & A_2 & A_3 & A_4 \\
A_2 & A_1 & A_2 & A_3 \\
A_3 & A_2 & A_1 & A_2 \\
A_4 & A_3 & A_2 & A_1 \\
\end{bmatrix}
$$
This example illustrates that the $M^2 \times M^2$ covariance matrix consists of $M$ submatrices of dimension $M \times M$ arranged in a bisymmetric form. The submatrices are not symmetric; however, for each matrix, $a_{ij} = a_{M+1-i, M+1-j}$.

The symmetry properties of the covariance matrix lead to a computationally simple procedure for the eigenvalue-eigenvector calculations. The procedure we will develop in the next section is similar to a method given by Ray and Driver (6) for decomposition of the Karhunen-Loève series representation of stationary random process.

III. EIGENVALUES AND EIGENVECTORS OF THE COVARIANCE MATRIX.

We now show that the eigenvalues and eigenvectors of the $2m \times 2m$ covariance matrix, $C$, can be obtained by calculating the eigenvalues and eigenvectors of two $m \times m$ submatrices of $C$. This simplification results from the bisymmetry properties of $C$ and the simplified procedure is developed through the following lemmas.

Lemma 1. The covariance matrix $C$, of dimension $2m \times 2m$, can be partitioned into $m \times m$ submatrices of the following form:

$$
C = \begin{bmatrix}
A & \text{BP} \\
\text{PB} & A
\end{bmatrix}
$$

where $A$ and $B$ are $m \times m$ submatrices of $C$ and $P$ is an $m \times m$ matrix with 1's along the NW-SE diagonal and 0's elsewhere. I.e., the $(i,j)^{th}$ element of $P$ is given by

$$
(P)_{ij} = 1 \quad \text{for } j = m + 1 - i
$$

$$
= 0 \quad \text{otherwise.}
$$

The proof of lemma 1 follows from the construction of $C$.

Lemma 2. The eigenvectors $V_i$, $i = 1, \ldots, 2m$ of $C$ has either one of the following two forms:
where $v_i$ is an $m \times 1$ column vector.

**Proof:** The characteristic equation of $C$ is given by

$$CY = \lambda Y$$

where $Y$ is an eigenvector of $C$ corresponding to the eigenvalue $\lambda$. We want to prove that the $i$th component $y_i$, denoted by $y_i^\prime$, satisfies

$$y_i = \pm y_{2m+1-i}.$$  \hspace{1cm} (5)

We begin by writing the characteristic equation in the form

$$\lambda y_i = \sum_{j=1}^{2m} c_{ij} y_j,$$  \hspace{1cm} (6)

and

$$\lambda y_{2m+1-i} = \sum_{j=1}^{2m} c_{2m+1-i,j} y_j.$$  \hspace{1cm} (7)

where $c_{ij}$ is the $(i,j)^{th}$ element of $C$. We may also write (7) as

$$\lambda y_{2m+1-i} = \sum_{j=1}^{2m} c_{2m+1-i,2m+1-j} y_{2m+1-j}$$

$$i = 1, 2, \ldots, 2m$$

The bisymmetric property of $C$ yields $c_{2m+1-i,2m+1-j} = c_{ij}$ and hence we may write the preceding equation as

$$\lambda y_{2m+1-i} = \sum_{j=1}^{2m} c_{ij} y_{2m+1-j}.$$  \hspace{1cm} (8)
Equations (6) and (8) both represent 2m equations in 2m unknowns. Letting
\[ y_{2m+1-i} = Z_i \] (8) becomes
\[ A Z_i = \sum_{j=1}^{2m} c_{ij} Z_j \] (9)

From (6) and (9) it is obvious that the solution for the \( y_i \)'s are also the solution for the \( Z_i \)'s, i.e. \( y_i \) and \( y_{2m+1-i} \) have the same solution.

Since the sign of the eigenvectors are not unique, forcing the norm of the eigenvectors to 1 makes
\[ y_i = \pm y_{2m+1-i} \quad i=1, \ldots, 2m \]

Thus we establish (4) and hence the proof of the lemma.

Lemma 3. The 2m eigenvalues \( \lambda_1, \lambda_2, \ldots, \lambda_{2m} \) and the corresponding eigenvectors \( V_1, V_2, \ldots, V_{2m} \) of \( C \) divide into the following two groups:

I. \[ \lambda_i = \lambda_i^+ \]
\[ V_i = \begin{bmatrix} v_i^+ \\ p_i^+ \end{bmatrix} \quad i = 1, 2, \ldots, m \] (10)

and

II. \[ \lambda_i = \lambda_i^- \]
\[ V_{m+i} = \begin{bmatrix} v_i^- \\ -p_i^- \end{bmatrix} \quad i = 1, 2, \ldots, m \] (11)

where \( \lambda_i^+ \) and \( v_i^+ \) are the eigenvalues and eigenvectors\(^3\) of the \( m \times m \) submatrix \( A+B \),

---

\(^3\)The vectors \( v_i^+ \) and \( v_i^- \) are normalized to give \( \| v_i^+ \|^2 = \| v_i^- \|^2 = 1/2 \) so that \( \| V_i \|^2 = 1 \).
and \( \lambda_i^- \) and \( \nu_i^- \) are the eigenvalues and eigenvectors of the \( m \times m \) submatrix \( A-B \), i.e.

\[
(A+B) \nu_i^+ = \lambda_i^+ \nu_i^+ \quad ; \quad i = 1, 2, \ldots, m \tag{12}
\]

\[
(A-B) \nu_i^- = \lambda_i^- \nu_i^- \quad ; \quad i = 1, 2, \ldots, m \tag{13}
\]

**Proof:** The characteristic equation of \( C \) is

\[
CV_i = \lambda V_i
\]

Using the partitioned form of \( C \), we may write the preceding equation as

\[
\begin{bmatrix}
A & BP \\
PB & A
\end{bmatrix} V_i = \lambda_i V_i
\]

(14)

Lemma 2 gives two forms of \( V_i \) and substituting the first form given in (4), we can write equation (14) as

\[
\begin{bmatrix}
A & BP \\
PB & A
\end{bmatrix} \begin{bmatrix}
\nu_i \\
PV_i
\end{bmatrix} = \lambda_i \begin{bmatrix}
\nu_i \\
PV_i
\end{bmatrix}
\]

(15)

The first \( m \) equations of (15) yield

\[
A \nu_i + BP PV_i = \lambda_i \nu_i , \text{ or}
\]

\[
(A+B) \nu_i = \lambda_i \nu_i
\]

(16)

since \( PP = I \).

Comparing (16) with the characteristic equation of the matrix \( A+B \) given in (12), we see that \( \lambda_i = \lambda_i^+ \) and \( \nu_i = \nu_i^+ \); and hence the proof of the first part of the lemma. Similarly, by taking the second form of \( V_i \) given in (4) we can prove part II of lemma 3. This completes the proof of lemma 3 which shows that the eigenvalues and eigenvectors of the \( 2m \times 2m \) covariance matrix \( C \) can be obtained from the eigenvalues and eigenvectors of two \( m \times m \) submatrices \( A+B \) and \( A-B \).
IV. CONCLUSIONS

We have presented a procedure which simplifies the computational complexity involved in calculating the eigenvectors and eigenvalues of the covariance matrix of imagery data. The procedure is based on the decomposition of the covariance matrix C as

$$C = \begin{bmatrix} A & BP \\ PB & A \end{bmatrix}$$

We have shown that the eigenvalues and eigenvectors of the $2m \times 2m$ covariance matrix can be obtained from the eigenvalues and eigenvectors of the $m \times m$ submatrices $A+B$ and $A-B$. Since the eigenvector calculations are proportional to the third power of the dimension of the matrix, the proposed procedure reduces the computations by a factor of four.

Also, the eigenvectors of the covariance matrix has to be stored, or transmitted to the receiver for reconstructing the imagery data according to (1). The symmetry property of the eigenvectors given in (10) and (11) enables us to store or transmit only half of the components of the eigenvectors. This results in a considerable saving in transmission time, especially if the dimension of the eigenvector is large.
REFERENCES


