Pattern Recognition Vol. 14, Nos. 1 6, pp. 375 381, 1981.
Printed in Great Britain.

0031 3203/81/070375 07 $02.00/0
Pergamon Press Ltd.

ON THE RELATIONSHIPS BETWEEN SVD, KLT AND PCA

JAN J. GERBRANDS
Department of Electrical Engineering, Delft University of Technology,
P.O. Box 5031, 2600 GA Delft, The Netherlands

(*Received* 9 *January* 1980; *in revised form* 1 *May* 1980; *received for publication* 22 *December* 1980)

Abstract – In recent literature on digital image processing much attention is devoted to the singular value decomposition (SVD) of a matrix. Many authors refer to the Karhunen–Loeve transform (KLT) and principal components analysis (PCA) while treating the SVD. In this paper we give definitions of the three transforms and investigate their relationships. It is shown that in the context of multivariate statistical analysis and statistical pattern recognition the three transforms are very similar if a specific estimate of the column covariance matrix is used. In the context of two-dimensional image processing this similarity still holds if one single matrix is considered. In that approach the use of the names KLT and PCA is rather inappropriate and confusing. If the matrix is considered to be a realization of a two-dimensional random process, the SVD and the two statistically defined transforms differ substantially.

Image processing   Statistical analysis   Statistical pattern recognition   Orthogonal image transforms
Singular value decomposition   Karhunen–Loeve transform   Principal components

I. INTRODUCTION

In recent literature on digital image processing much attention is devoted to the singular value decomposition (SVD) of a matrix, which can be viewed as a separable orthogonal transform or as an expansion of the matrix in rank-1 matrices. The matrix can represent a discrete image or a two-dimensional operator. Applications of the SVD lie in the fields of image restoration and enhancement as well as in image coding.[1–8]

In Huang,[6] Andrews states that SVD and principal components analysis are very similar, but he cautions his readers not to conclude that the SVD and the Karhunen–Loeve expansion are identical. Ahmed and Rao[9] treat the Karhunen–Loeve transform in the context of data compression and refer to the similar technique of principal components analysis in statistics. Their definition of the Karhunen–Loeve transform is identical with the definition of principal components given by Anderson.[10] However, Gnanadesikan[11] introduces the SVD as a way to compute the principal components, and Taylor[12] implies that SVD and principal components analysis are identical techniques. We found these remarks very confusing. In this paper we will try to unravel the situation. In Section II we will define the singular value decomposition and mention some of its properties and applications. In Section III the Karhunen-Loeve transform will be given the same treatment, followed in Section IV by principal components analysis. In Section V the relationships between the three techniques are investigated and their similarities and dissimilarities are exposed.

II. SINGULAR VALUE DECOMPOSITION (SVD)

A well-known result in matrix theory is the fundamental matrix decomposition theorem.[13] According to this theorem any real $n \times m$ matrix $[G]$ of rank p can be decomposed as

$$[G] = [U][\Lambda^{1/2}][V]^t \qquad (1)$$

where t denotes matrix transposition. The $n \times m$ matrix $[\Lambda^{1/2}]$ is of the following form

$$[\Lambda^{1/2}] = \begin{bmatrix} \lambda_1^{1/2} & & & & 0 \\ & \ddots & & & \\ & & \lambda_p^{1/2} & & \\ & & & 0 & \\ & & & & \ddots \\ 0 & & & & 0 \end{bmatrix} \qquad (2)$$

The orthogonal $n \times n$ matrix $[U]$ and the orthogonal $m \times m$ matrix $[V]$ are defined by

$$[G][G]^t = [U]\begin{bmatrix} \lambda_1 & & & & 0 \\ & \ddots & & & \\ & & \lambda_p & & \\ & & & 0 & \\ & & & & \ddots \\ 0 & & & & 0 \end{bmatrix}[U]^t \qquad (3)$$

$$[G]^t[G] = [V]\begin{bmatrix} \lambda_1 & & & 0 \\ & \ddots & & \\ & & \lambda_p & \\ & & & \ddots \\ 0 & & & 0 \end{bmatrix}[V]^t \qquad (4)$$

So the columns of $[U]$ are the eigenvectors of $[G][G]'$ and the columns of $[V]$ are the eigenvectors of $[G]'[G]$. The diagonal entries of $[\Lambda^{1/2}]$ are the square roots of the eigenvalues of $[G][G]'$ as well as $[G]'[G]$ and are called the singular values of $[G]$. The decomposition (1) is usually referred to as the singular value decomposition (SVD) of $[G]$ and has found a number of applications in digital image processing.[1-8] With

$$[U] = [\mathbf{u}_1\mathbf{u}_2\ldots\mathbf{u}_n] \tag{5}$$

$$[V] = [\mathbf{v}_1\mathbf{v}_2\ldots\mathbf{v}_m] \tag{6}$$

equation (1) can be written as

$$[G] = \sum_{i=1}^{p} \lambda_i^{1/2}\mathbf{u}_i\mathbf{v}_i' \tag{7}$$

where the product $\mathbf{u}_i\mathbf{v}_i'$ is an $n\times m$ matrix, being the product of the $n \times 1$ matrix $\mathbf{u}_i$ and the $1 \times m$ matrix $\mathbf{v}_i'$. Equation (7) shows that the SVD constitutes an expansion of the $n\times m$ matrix [G] into a sum of separable rank-1 matrices. If the singular values are ordered according to

$$\lambda_1^{1/2} \geqslant \lambda_2^{1/2} \geqslant \ldots \geqslant \lambda_p^{1/2} \tag{8}$$

and we consider the truncated series approximation of $[G]$ given by

$$[G_k] = \sum_{i=1}^{k} \lambda_i^{1/2}\mathbf{u}_i\mathbf{v}_i' \tag{9}$$

with $k < p$, p being the rank of $[G]$, then the approximation or truncation error norm is given by

$$\|[G] - [G_k]\|^2 = \sum_{i=k+1}^{p} \lambda_i \tag{10}$$

where the matrix norm is defined as $\|[A]\|^2 = \mathrm{tr}[A]'[A]$ and where tr denotes the trace or the sum of the diagonal entries of $[A]'[A]$. With the singular values placed in non-ascending order, the SVD is the two-dimensional separable orthogonal transform with minimum least-squares truncation error. In general, a two-dimensional separable orthogonal transform is formulated as

$$[F] = [C][G][R] \tag{11}$$

where the orthogonal matrices $[C]$ and $[R]$ operate on the columns and on the rows separately, carrying $[G]$ into the transform domain.[7] The coefficients in this domain are the entries of $[F]$. Well-known transforms are the Fourier, Walsh–Hadamard, discrete cosine and Slant transforms. The SVD is the transform that results in a diagonal matrix of transform coefficients.

This can be seen by rewriting equation (1) as

$$[\Lambda^{1/2}] = [U]'[G][V]. \tag{12}$$

Notice that the SVD necessitates the computation of unique transform matrices $[U]$ and $[V]$ for each matrix $[G]$, as opposed to the other well-known transforms.

The SVD may constitute an efficient way of matrix or image representation.[3] Description of $[G]$ in its matrix form requires nm numbers or computer memory locations. Applying the SVD as in equation (7) and incorporating the singular values λ_i into the vectors $\mathbf{u}_i$ results in a description of $[G]$ with $p(n + m)$ numbers. So if $p(n + m) < nm$ the SVD description is more economic. This advantage becomes even more clear if we accept some error and apply equation (9) with $k < p$. More savings could be obtained by exploiting the orthonormality of the vectors $\mathbf{u}_i$, which is a consequence of the orthogonality of the matrix $[U]$. The same is true for the vectors $\mathbf{v}_i$. Notice that these savings are achieved through complex and time-consuming computations.

If $[G]$ represents the impulse response of a two-dimensional spatially-invariant linear filter, the two-dimensional convolution of an image with $[G]$ itself may be replaced by repeated one-dimensional convolutions with $\mathbf{u}_i$ and $\mathbf{v}_i$ through equations (7) or (9).

Under the same restrictions on the rank p of $[G]$, substantial savings in processing time can be achieved. The computation of the SVD has to be performed only once for every filter $[G]$, so this approach may be very attractive if many images must be processed with the same filter.[8]

III. KARHUNEN-LOEVE TRANSFORM (KLT)

In the literature on digital signal processing the Karhunen–Loeve transform is derived as the optimum orthogonal transform for signal representation (data compression) with respect to the mean-square error criterion. This transform results in uncorrelated coefficients in the transform domain. The proof (see e.g. Ahmed and Rao[9]) is along the following lines.

Consider a vector $\mathbf{x}$ with $\mathbf{x}' = [x_1x_2\ldots x_n]$ and the orthogonal transform $[T]'$ given by $[T] = [\boldsymbol{\phi}_1\boldsymbol{\phi}_2\ldots\boldsymbol{\phi}_n]$ with $\boldsymbol{\phi}_i' = [\phi_{i1}\phi_{i2}\ldots\phi_{in}]$. Applying this transform to the vector $\mathbf{x}$ results in

$$\mathbf{y} = [T]'\mathbf{x}. \tag{13}$$

Because $[T]'$ is orthogonal it follows that

$$\begin{aligned}\mathbf{x} &= [T]\mathbf{y} = [\boldsymbol{\phi}_1\boldsymbol{\phi}_2\ldots\boldsymbol{\phi}_n]\mathbf{y}\\ &= y_1\boldsymbol{\phi}_1 + y_2\boldsymbol{\phi}_2 + \ldots + y_n\boldsymbol{\phi}_n.\end{aligned} \tag{14}$$

If we truncate this expansion after k terms, $k < n$, and replace the remaining y_i, $i = k + 1,\ldots,n$ by preselected constants, we obtain

$$\hat{\mathbf{x}} = \sum_{i=1}^{k} y_i\boldsymbol{\phi}_i + \sum_{i=k+1}^{n} c_i\boldsymbol{\phi}_i. \tag{15}$$

Now we look for those constants c_i and those $\boldsymbol{\phi}_i$ that minimize the mean-square error

$$\varepsilon(k) = E\{(\mathbf{x} - \hat{\mathbf{x}})'(\mathbf{x} - \hat{\mathbf{x}})\} \tag{16}$$

where E denotes the mathematical expectation. Without proof we give the result. The optimum values for the constants c_i are

$$c_i = E\{y_i\} = E\{\boldsymbol{\phi}_i^t \mathbf{x}\} = \boldsymbol{\phi}_i^t E\{\mathbf{x}\} \quad i = k+1, \ldots, n \quad (17)$$

while the optimum $\boldsymbol{\phi}_i$ turn out to satisfy

$$[K_x]\boldsymbol{\phi}_i = \lambda_i \boldsymbol{\phi}_i \quad (18)$$

where $[K_x]$ is the covariance matrix of $\mathbf{x}$. This implies that the $\boldsymbol{\phi}_i$ are eigenvectors of $[K_x]$ and the λ_i the corresponding eigenvalues. The resulting mean-square error equals

$$\varepsilon(k) = \sum_{i=k+1}^{n} \lambda_i \quad (19)$$

which is minimum for all values of k if the λ_i are in non-ascending order.

Summarizing, the KLT is defined as

$$\mathbf{y} = [T]^t \mathbf{x} \quad (20)$$

or, equivalently, as the expansion

$$\mathbf{x} = [T]\,\mathbf{y} \quad (21)$$

with the transform matrix $[T]$ defined by

$$[K_x] = [T] \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} [T]^t = [T][\Lambda][T]^t. \quad (22)$$

It can easily be derived that the covariance matrix $[K_y]$ of the vector $\mathbf{y}$ of coefficients in the transform domain equals

$$[K_y] = [T]^t[K_x][T] = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} = [\Lambda] \quad (23)$$

which implies that the entries of $\mathbf{y}$ are uncorrelated.

Now we turn our attention to the Karhunen–Loeve transform of an image matrix $[X]$. In general one has to scan $[X]$ obtaining a vector $\mathbf{x}$ with covariance matrix $[K_x]$. Here we consider column scanning, that is, we place the element x_{kl} of the $n \times m$ matrix into the element $(l-1)n + k$ of the $nm \times 1$ vector $\mathbf{x}$.

If the image statistics are such that all columns of $[X]$ have the same covariance matrix $[K_c]$ and all rows have covariance matrix $[K_r]$, the covariance matrix of the stacked vector $\mathbf{x}$ can be written as[7]

$$[K_x] = [K_r] \otimes [K_c] \quad (24)$$

where $\otimes$ denotes the Kronecker matrix product (7),

$$[A] \otimes [B] = \begin{bmatrix} a_{11} \ldots a_{1n} \\ \vdots \quad\quad \vdots \\ a_{n1} \ldots a_{nn} \end{bmatrix} \otimes [B] = \begin{bmatrix} a_{11}[B] \ldots a_{1n}[B] \\ \vdots \quad\quad \vdots \\ a_{n1}[B] \ldots a_{nn}[B] \end{bmatrix} \quad (25)$$

The KLT of the column-stacked vector $\mathbf{x}$ is based on

$$[K_x] = [T][\Lambda][T]^t. \quad (26)$$

Now if

$$[K_c] = [T_c][\Lambda_c][T_c]^t \quad (27)$$

and

$$[K_r] = [T_r][\Lambda_r][T_r]^t \quad (28)$$

then

$$\begin{aligned}[K_x] &= \{[T_r][\Lambda_r][T_r]^t\} \otimes \{[T_c][\Lambda_c][T_c]^t\} \\ &= \{\{[T_r][\Lambda_r]\} \otimes \{[T_c][\Lambda_c]\}\}\{[T_r]^t \otimes [T_c]^t\} \\ &= \{[T_r] \otimes [T_c]\}\{[\Lambda_r] \otimes [\Lambda_c]\}\{[T_r] \otimes [T_c]\}^t. \end{aligned} \quad (29)^*$$

So we have

$$[T] = [T_r] \otimes [T_c] \quad (30)$$

and equation (20) can be rewritten as (7)

$$[Y] = [T_c]^t[X][T_r] \quad (31)$$

where $\mathbf{y}$ in equation (20) is the column-stacked vector corresponding with $[Y]$ in equation (31).

Equations (24), (27), (28) and (31) constitute the two-dimensional separable KLT, in general resulting in a full matrix $[Y]$, where on average the largest amount of 'energy' is concentrated in the minimum number of entries Y_{ij} of $[Y]$. Equation (31) can be rewritten as

$$[X] = [T_c][Y][T_r]^t = \sum_i \sum_j y_{ij} \mathbf{t}_{c_i} \mathbf{t}_{r_j}^t \quad (32)$$

where the $\mathbf{t}_{c_i}$ are the columns of $[T_c]$ and the $\mathbf{t}_{r_j}$ the columns of $[T_r]$. This implies that the KLT can be described as a decomposition into 'eigen images' defined by the covariance matrix. The transform matrices have to be computed only once for each class of images with one common covariance matrix.

IV. PRINCIPAL COMPONENTS ANALYSIS (PCA)

Principal components analysis is a dimensionality reduction technique in multivariate statistical analysis.[10,11] Multivariate statistical analysis deals with the analysis of data that consist of measurements on a number of individuals or objects. If there are m objects and each object is described by n variables or features, we have a $n \times m$ data matrix. Often the number of variables is too large to handle. A way of reducing this number is to take linear combinations of the variables and discard the combinations with small variances. Now the first principal component is the linear combination with maximum variance, the second has maximum variance subject to being orthogonal to the first, and so on. Notice the resemblance between this problem and the problem of efficient signal representation of Section III. In fact, the definitions of the principal components transform and the Karhunen–Loeve transform will be shown to be almost identical.

Anderson[10] proves that if the $n \times 1$ vector $\mathbf{z}$ has $E\{\mathbf{z}\} = \mathbf{o}$ and covariance matrix $[K_z]$, there exists an orthogonal transform

$$\mathbf{y} = [B]^t \mathbf{z} \quad (33)$$

---

*Here we used the following properties of the Kronecker product:[7,13]

$$\{[A] \otimes [B]\}\{[C] \otimes [D]\} = \{[A][C]\} \otimes \{[B][D]\}$$
$$\{[A] \otimes [B]\}^t = [A]^t \otimes [B]^t.$$

such that the covariance matrix of y is $E\{\mathbf{y}\mathbf{y}'\} = [\Lambda]$ and

$$[\Lambda] = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \quad (34)$$

where $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n$ are the eigenvalues of $[K_z]$ and the columns of $[B]$ are the corresponding eigenvectors. The rth element of y, $y_r = \mathbf{b}_r'\mathbf{z}$ has maximum variance of all normalized linear combinations uncorrelated with $y_1, \ldots, y_{r-1}$. This vector y is defined as the vector of principal components of z.

If we have an input vector x with $E(\mathbf{x}) \neq \mathbf{0}$ we define

$$\mathbf{z} = \mathbf{x} - E(\mathbf{x}) \quad (35)$$

and compute the principal components transform of x as

$$\mathbf{y} = [B]'\mathbf{z} = [B]'(\mathbf{x} - E(\mathbf{x})) \quad (36)$$

with $[B]$ defined by

$$[K_z] = [B][\Lambda][B]'. \quad (37)$$

However, since

$$[K_z] = E\{\mathbf{z}\mathbf{z}'\} = E\{(\mathbf{x} - E(\mathbf{x}))(\mathbf{x} - E(\mathbf{x}))'\} = [K_x] \quad (38)$$

we arrive at the following definition of the principal components transform

$$\mathbf{y} = [B'](\mathbf{x} - E(\mathbf{x})) \quad 39)$$

with

$$[K_x] = [B][\Lambda][B]'. \quad (40)$$

This definition is in agreement with the literature.[10,11,12] In principal components analysis the above vector x is the $n \times 1$ measurement or feature vector of one particular object. In the case of m objects we arrange the m vectors into the $n \times m$ data matrix $[X]$. Under the assumption that all m vectors are realizations of the same random process, the principal components transform becomes

$$[Y] = [B]'\{[X] - E[X]\} \quad (41)$$

where the m output vectors are the columns of $[Y]$ and $E[X]$ is the matrix of mean vectors, which implies that all columns of $E[X]$ are identical.

Due to the orthogonality of $[B]$, equation (41) can be rewritten as

$$[X] - E[X] = [B][Y] \quad (42)$$

where the matrices are $n \times m$, $n \times m$, $n \times n$ and $n \times m$, respectively. In general, principal components analysis results in a dimensionality reduction, to be described as

$$\{[X] - E[X]\}^* = [B'][Y'] \quad (43)$$

where the matrices are $n \times m$, $n \times m$, $n \times q$ and $q \times m$, respectively, and where the approximation $\{[X] - E[X]\}^*$ contains as much of the 'variability' of the original data as possible. Closely related to principal components analysis is the field of factor analysis. In this paper we will not deal with factor analysis separately. Similarities and differences between principal components analysis and factor analysis are discussed in Gnanadesikan,[11] Taylor[12] and Harman.[14]

V. RELATIONSHIPS

In this section we will explore the relationships between the Karhunen–Loeve and principal components transforms and then the relationships between these transforms and the singular value decomposition. The transforms may be discussed in terms of orthogonal expansions of discrete random processes.

Consider first a continuous stochastic process $\xi(t)$, where t is a one-dimensional parameter. The process $\xi(t)$, defined in a time domain $(0, T)$, can be expressed[16–18] in a linear combination of basis functions $\psi_i(t)$, which are orthonormal in the interval $(0, T)$, according to

$$\xi(t) = \sum_{i=1}^{\infty} \alpha_i \psi_i(t) \quad 0 \leqslant t \leqslant T. \quad (44)$$

An infinite number of basis functions is required in order to form a complete set. Now consider a $n \times 1$ random vector $\mathbf{x} = [x_1 x_2 \ldots x_n]'$ as in the previous sections. This vector may be expressed in a linear combination of orthonormal basis vectors $\boldsymbol{\phi}_i = [\phi_{i1}\phi_{i2}\ldots\phi_{in}]'$ as

$$x_k = \sum_{i=1}^{n} y_i \phi_{ik} \quad k = 1, 2, \ldots, n \quad (45)$$

which is equivalent to

$$\mathbf{x} = [T]\mathbf{y} \quad (46)$$

where $[T] = [\boldsymbol{\phi}_1 \boldsymbol{\phi}_2 \ldots \boldsymbol{\phi}_n]$. Note that now only n basis vectors are required for completeness.

The Karhunen–Loeve transform was originally developed and discussed as the expansion of the stochastic process $\xi(t)$ into a complete set of deterministic time functions $\psi_i(t)$ such that the random variables α_i in equation (44) become uncorrelated. This is achieved if the $\psi_i(t)$ are the eigenfunctions of the covariance function $C(t, s)$ of the process $\xi(t)$. The discrete Karhunen–Loeve transform discussed in Section III is based on the eigenvectors of the covariance matrix $[K_x]$ of the vector x.

In equations (20) and (22) the Karhunen–Loeve transform of the input vector x was given as

$$\mathbf{y}_1 = [T]'\mathbf{x} \quad (47)$$

$$[K_x] = [T][\Lambda][T]'. \quad (48)$$

In equations (39) and (40) the principal components transform was defined as

$$\mathbf{y}_2 = [B]'(\mathbf{x} - E(\mathbf{x})) \quad (49)$$

$$[K_x] = [B][\Lambda][B]'. \quad (50)$$

From equations (48) and (50) it is clear that $[T] = [B]$, which implies that the transform matrices of both transforms are identical. It can easily be derived that the covariance matrices for the output vectors $\mathbf{y}_1$ and $\mathbf{y}_2$ are identical and equal to $[\Lambda]$. The only difference lies in the mean vectors $E(\mathbf{y}_1)$ and $E(\mathbf{y}_2)$. For the Karhunen–Loeve transform we have

$$E(\mathbf{y}_1) = [T]'E(\mathbf{x}) \quad (51)$$

and for the principal components

$$E(\mathbf{y}_2) = [T]'E(\mathbf{x} - E(\mathbf{x})) = 0. \quad (52)$$

A geometric interpretation of the Karhunen–Loeve transform is in terms of a rotation of the coordinate system. The principal components method can be described as a shift of the origin of the coordinate system to the point $E(\mathbf{x})$, followed by the same rotation. So for single vectors we conclude that the Karhunen–Loeve and principal components transforms differ only in this shift of the origin.

In the context of multivariate statistical analysis and statistical pattern recognition we have m vectors $\mathbf{x}_i$, $i = 1, \ldots, m$, arranged in a $n \times m$ data matrix $[X]$.

All columns $\mathbf{x}_i$ have the same covariance matrix $[K_c]$. The Karhunen–Loeve transform is

$$[Y_1] = [T]'[X] \quad (53)$$

and the principal components transform

$$[Y_2] = [T]'\{[X] - E[X]\} \quad (54)$$

with

$$[K_c] = [T][\Lambda][T]'. \quad (55)$$

The correspondence between both methods remains the same as in the case of single vectors. In a practical situation one has to estimate the column covariance matrix $[K_c]$ from the data. An unbiased estimate is the sample covariance matrix of the columns of $[X]$

$$[S_c] = \frac{1}{m-1} \sum_{i=1}^{m} (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})' \quad (56)$$

with

$$\bar{\mathbf{x}} = \frac{1}{m} \sum_{i=1}^{m} \mathbf{x}_i. \quad (57)$$

Equation (56) can easily be rewritten as

$$[S_c] = \frac{1}{m-1} \{[X] - [\bar{X}]\}\{[X] - [\bar{X}]\}' \quad (58)$$

where all columns of $[\bar{X}]$ are identical and equal to $\bar{\mathbf{x}}$.

The transform matrix $[T]$ of both the Karhunen–Loeve and the principal components transforms is then derived from

$$[S_c] = [T][\Lambda][T]'. \quad (59)$$

Now compare these equations with the singular value decomposition

$$[G] = [U][\Lambda^{1/2}][V]' \quad (60)$$

with $[U]$ defined by

$$[G][G]' = [U][\Lambda][U]'. \quad (61)$$

If the column covariance matrix is estimated from the data through equation (58), the transform matrix $[T]$ in the Karhunen–Loeve and principal components transforms of the data matrix $[X]$ corresponds with the transform matrix $[U]$ in the SVD of $\{[X] - [\bar{X}]\}$.

The factor $(m - 1)$ in (58) does not influence the orthonormal eigenvectors. If we rewrite equations (53) and (54) as

$$[X] = [T][Y_1] \quad (62)$$

and

$$[X] - [\bar{X}] = [T][Y_2] \quad (63)$$

it is also clear that the matrices $[Y_1]$ and $[Y_2]$ can be written in a form corresponding with the product $[\Lambda^{1/2}][V]'$ in equation (57) with $[V]$ defined by

$$[G]'[G] = [V][\Lambda][V]' \quad (64)$$

where $[G]'[G]$ is related to the sample covariance matrix of the rows of the data matrix. However, principal components analysis is usually described in terms of the column covariance matrix only, due to the fact that one is interested in a reduction of the number of features by exploiting the statistical dependency between these features, i.e. between the elements in a column of the data matrix.

We conclude that, using the estimated column covariance matrix of equation (58), the Karhunen–Loeve and principal components transforms of a data matrix $[X]$ correspond with the SVD of $\{[X] - [\bar{X}]\}$. Now it is clear why Gnanadesikan[11] and Taylor[12] mention the SVD as a way to compute the principal components. It is also clear why Andrews in the book edited by Huang[6] in his treatment of the SVD states: "For those familiar with principal components analysis, the matrices $[G][G]'$ and $[G]'[G]$ can be likened to sample column and row covariance matrices, respectively....".

In the context of multivariate statistical analysis the m columns of the data matrix are viewed as m realizations of a discrete stochastic process with a one-dimensional parameter space. It is reasonable to estimate the covariance matrix from these m realizations. When we turn our attention to digital image processing we come upon an essentially different situation. Let the matrix $[X]$ with elements x_{kl} represent a digital image. The concept of stochastic processes can be generalized to two dimensions by the introduction of a random field (Wong,[19] Rosenfeld[20]), i.e. a random process with a two-dimensional parameter space. The random field $\xi(s,t)$, defined in a region S in a two-dimensional Euclidean space, can be expressed as a linear combination of basis functions $\psi_{ij}(s,t)$, which are orthonormal in S, according to

$$\xi(s,t) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \alpha_{ij}\psi_{ij}(s,t) \quad (s,t) \in S. \quad (65)$$

This equation constitutes the continuous Karhunen–Loeve expansion of the process if the $\psi_{ij}(s,t)$ are the eigenfunctions of the covariance function $C(s_1, t_1, s_2, t_2)$ of the process $\xi(s,t)$. In the discrete case under consideration here, we obtain the expansion of the matrix $[X]$ with elements x_{kl} into a linear combination of orthonormal basis matrices $[\phi_{ij}]$ with elements ϕ_{ijkl}

$$x_{kl} = \sum_{i=1}^{n} \sum_{j=1}^{m} y_{ij}\phi_{ijkl} \quad \begin{array}{l} k = 1,2,\ldots,n \\ l = 1,2,\ldots,m \end{array} \tag{66}$$

or, equivalently, as

$$[X] = \sum_{i=1}^{n} \sum_{j=1}^{m} y_{ij}[\phi_{ij}]. \tag{67}$$

If the matrices $[\phi_{ij}]$ are the eigenmatrices of the covariance $C(k_1, l_1, k_2, l_2)$ of $[X]$, these equations constitute the Karhunen–Loeve transform of the discrete image $[X]$. If the matrices $[\phi_{ij}]$ are determined from the eigenproblems of $[X][X]^t$ and $[X]^t[x]$, they are separable and equation (67) constitutes the singular value decomposition of $[X]$ as given in equation (7).

The solution of the eigenproblem of $C(k_1, l_1, k_2, l_2)$ is greatly facilitated by arranging the elements of this covariance into the $nm \times nm$ matrix $[K_x]$ which was used in Section III. This procedure is discussed in detail in Rosenfeld.[20] This matrix has to be estimated from a large number of image matrices. The KLT based on this estimate bears no resemblance with the SVD which is uniquely defined by and for each matrix $[X]$ individually.

In the special case that the covariance $C(k_1, l_1, k_2, l_2)$ is separable as

$$C(k_1, l_1, k_2, l_2) = C_c(k_1, k_2)\, C_r(l_1, l_2) \tag{68}$$

the eigenmatrices become separable as well and we obtain the two-dimensional separable KLT discussed in Section III. In this case, $[K_x]$ is separable into the Kronecker product of a column covariance matrix $[K_c]$ and a row covariance matrix $[K_r]$. Unbiased estimates of $[K_c]$ and $[K_r]$ are $[S_c]$ and $[S_r]$, respectively, which one might estimate from the n rows and the m columns of one image matrix $[X]$ as

$$[S_r] = \frac{1}{n-1}\{[X] - [\bar{X}]\}^t\{[X] - [\bar{X}]\} \tag{69}$$

$$[S_c] = \frac{1}{m-1}\{[X] - [\bar{X}]\}\{[X] - [\bar{X}]\}^t. \tag{70}$$

The two-dimensional separable KLT of $[X]$ based on these estimates is identical with the SVD of $\{[X] - [\bar{X}]\}$. However, accepting $[S_r]$ and $[S_c]$ as estimates for $[K_r]$ and $[K_c]$, respectively, implies that $[K_x]$ can be estimated from one single image matrix $[X]$, that is, from one realization of the two-dimensional random process. Obviously, the number of degrees of freedom of the matrix $[K_x]$ is greatly diminished by the assumption of separability, which means that the number of parameters one has to estimate in the separable case is much smaller than in the non-separable case. Even then, it is evident that the name KLT should not be used for a transform based on the estimates from equations (69) and (70). In fact, the correct name is SVD. A second problem is the separability of the covariance matrix $[K_x]$. Given a class of images, one has to estimate $[K_x]$ from a large number of these images before any statements can be made about its separability. One could argue that there are popular two-dimensional auto-regressive image models which can be proven to lead to separable covariance matrices. In that situation the problem is shifted to the necessity to investigate the validity of such models, given the class of images.

From the discussion above we conclude that in the context of multivariate statistical analysis or statistical pattern recognition the SVD and the KLT and principal components transforms are very similar if certain estimates of the covariance matrices are used. Under certain conditions all three transforms become identical. In the context of two-dimensional image processing the similarity between the three techniques still holds if only one matrix $[X]$ is considered.

However, we feel that the use of statistical terms like KLT or PCA is rather inappropriate and confusing in such a situation. No such confusion arises if the name SVD, which is defined in deterministic terms, is used. If the matrix $[X]$ is considered to be one realization of a two-dimensional random process, the statistical transforms and the SVD are very different methods.

VI. CONCLUSIONS

For single vectors, the KLT as defined by Ahmed and Rao[9] and the PCA as defined by Anderson[10] are identical except for a possible shift of the origin of the coordinate system. For an $n \times m$ matrix in the context of multivariate statistical analysis or statistical pattern recognition, the m columns of the matrix are regarded as m realizations of a random process and the similarity between KLT and PCA still holds. If the column covariance matrix is estimated from the m realizations the KLT and PCA of the matrix $[X]$ become identical with the SVD of $\{[X] - [\bar{X}]\}$. In the context of two-dimensional image processing this similarity only applies if one single matrix $[X]$ is considered. To prevent confusion, statistical terms like KLT or PCA should not be used in this situation. The correct name is SVD, which is defined in deterministic terms only. If the image $[X]$ is considered to be a realization of a two-dimensional random process, the covariance matrices for the KLT and PCA should be estimated from a number of realizations of that process, that is, from a number of images. In that situation there is a substantial difference between the statistically defined transforms and the SVD. The statistical transforms are the same for all realizations in the class of images under consideration, the deterministic SVD transform is uniquely defined by each image matrix itself. While the approximation or truncation error of the statistical transforms is minimum in the mean-square sense, the SVD truncation error is minimum in the least-squares sense. Thus in the context of two-dimensional statistical signal processing, the KLT and SVD differ substantially, both in theory and in practice.

SUMMARY

In recent literature on digital image processing much attention is devoted to the singular value decomposition (SVD) of a matrix. Many authors refer to the Karhunen–Loeve transform (KLT) and principal components analysis (PCA) while treating the SVD. In this paper we give definitions of the three transforms and investigate their relationships.

For single vectors, the KLT, as defined by Ahmed and Rao[9] and the PCA, as defined by Anderson,[10] are identical except for a possible shift of the origin of the coordinate system. For an $n \times m$ matrix in the context of multivariate statistical analysis or statistical pattern recognition, the m columns of the matrix are regarded as m realizations of a random process and the similarity between KLT and PCA still holds. If the column covariance matrix is estimated from the m realizations the KLT and PCA of the matrix $[X]$ become identical with the SVD of $\{[X] - [\bar{X}]\}$. In the context of two-dimensional image processing this similarity only applies if one single matrix $[X]$ is considered. To prevent confusion, statistical terms like KLT or PCA should not be used in this situation. The correct name is SVD, which is defined in deterministic terms only. If the image $[X]$ is considered to be a realization of a two-dimensional random process, the covariance matrices for the KLT and PCA should be estimated from a number of realizations of that process, that is, from a number of images. In that situation there is a substantial difference between the statistically defined transforms and the SVD. The statistical transforms are the same for all realizations in the class of images under consideration, the deterministic SVD transform is uniquely defined by each image matrix itself. While the approximation or truncation error of the statistical transforms is minimum in the mean-square sense, the SVD truncation error is minimum in the least-squares sense. Thus in the context of two-dimensional statistical signal processing, the KLT and SVD differ substantially, both in theory and in practice.

REFERENCES

1. H. C. Andrews and C. L. Patterson, Singular value decompositions and digital image processing, *IEEE Transactions on Acoustics, Speech, and Signal Processing* **24**, 26–53 (1976).
2. H. C. Andrews and C. L. Patterson, Outer product expansions and their uses in digital image processing, *IEEE Trans. Comput.* **25**, 140–148 (1976).
3. H. C. Andrews, C. L. Patterson, Singular value decomposition (SVD) image coding, *IEEE Trans. Communs* **24**, 425–432 (1976).
4. H. C. Andrews and B. R. Hunt, *Digital Image Restoration*. Prentice-Hall, Englewood Cliffs, NJ (1977).
5. T. S. Huang and P. M. Narendra, Image restoration by singular value decomposition, *Appl. Opt.* **14**, 2213–2216 (1975).
6. T. S. Huang, ed., *Picture Processing and Digital Filtering*. Springer, Berlin (1975).
7. W. K. Pratt, *Digital Image Processing*. John Wiley, New York (1978).
8. S. Treitel and J. L. Shanks, The design of multistage separable planar filters, *IEEE Trans. Geosci. Electron.* **9**, 10–27 (1971).
9. N. Ahmed and K. R. Rao, *Orthogonal Transforms for Digital Signal Processing*. Springer, Berlin (1975).
10. T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*. John Wiley, New York (1958).
11. R. Gnanadesikan, *Methods for Statistical Data Analysis of Multivariate Observations*. John Wiley, New York (1977).
12. C. C. Taylor, Principal component and factor analysis, *The Analysis of Survey Data*, Volume I, *Exploring Data Structures* C. A. O'Muircheartaigh and C. Payne, eds. John Wiley, New York (1977).
13. C. Lanczos, *Linear Differential Operators*. Van Nostrand, New York (1961).
14. H. H. Harman, *Modern Factor Analysis*. University of Chicago Press, Chicago (1967).
15. R. Bellman, *Introduction to Matrix Analysis*. McGraw-Hill, New York (1960).
16. W. B. Davenport and W. L. Root, *An Introduction to the Theory of Random Signals and Noise*. McGraw-Hill, New York (1958).
17. A. Papoulis, *Probability, Random Variables and Stochastic Processes*. McGraw-Hill, New York (1965).
18. K. Fukunaga, *Introduction to Statistical Pattern Recognition*. Academic Press, New York (1972).
19. E. Wong, *Stochastic Processes in Information and Dynamical Systems*. McGraw-Hill, New York (1971).
20. A. Rosenfeld and A. C. Kak, *Digital Picture Processing*. Academic Press, New York (1976).

About the Author – JAN J. GERBRANDS was born in 1948 in 's-Gravenhage, The Netherlands. He received the degree of Ingenieur in electrical engineering from Delft University of Technology, Delft, The Netherlands, in 1974, and is presently working towards the Ph.D. degree. He is currently with the Information Theory Group of the Electrical Engineering Department of Delft University of Technology. His research interests include digital image processing and analysis and pattern recognition, and medical applications.