Signal Processing 11 (1986) 61-79
North-Holland

LOCAL LINEAR TRANSFORMS FOR TEXTURE MEASUREMENTS

Michael UNSER* (Member EURASIP)

Laboratoire Traitement des Signaux, École Polytechnique Fédérale de Lausanne, 16 Chemin de Bellerive, CH-1007 Lausanne, Switzerland

Received 18 September 1984
Revised 3 September 1985

Abstract. The Nth order probability density function for pixels in a restricted neighborhood may be characterized by a set of N histograms (or some corresponding moments) computed along appropriately chosen axes. The projections on those axes are obtained from a local linear transform of the local neighborhood vector. This approach is closely related to filter bank analysis methods and gives a statistical justification for the extraction of texture properties by means of convolution operators or local matches. Optimal and sub-optimal linear operators are derived for texture analysis and classification. Experimental results indicate that the method is robust, flexible, and that it performs as well as standard co-occurrence based methods for texture classification. The proposed approach enables texture characterization with a lower number of features and it is also computationally more appealing.

Zusammenfassung. Das Verfahren der Bildanalyse mit Hilfe lokaler linearer Transformationen erlaubt es, die N-dimensionale Verteilungsdichtefunktion der Punkte eines begrenzten Bildausschnitts durch N Histogramme anzunähern, die entlang geeignet gewählter Achsen aufgestellt werden. Die Projektionen auf diese Achsen werden mit Hilfe einer linearen Transformation des sogenannten Nachbarschaftsvektors berechnet. Diese Annäherung entspricht der Analyse mit Hilfe einer Filterbank und gibt eine statistische Rechtfertigung für die Extraktion von Eigenschaften der Bildtextur mit Hilfe lokaler Merkmalsfilter. Optimale und suboptimale Lösungen für die Wahl derartiger linearer Filter werden vorgeschlagen für Texturanalyse and Klassifikation. Wie Experimente mit realen Bildtexturen zeigen, ist die Methode unempfindlich, anpassbar, und ebenso zuverlässig wie Standardmethoden die Pixelpaarebeziehungen für Texturklassifikation verwenden. Die vorgeschlagene Methode ermöglicht Texturbeschreibungen mit weniger Parametern und benötigt weniger Computeroperationen.

Résumé. La méthode d'analyse de texture par transformation linéaire locale permet une caractérisation partielle d'une densité de probabilité d'ordre N par N histogrammes calculés selon des axes convenablement choisis. Cette approche est équivalente à une analyse par banc de filtres et apporte une justification statistique quant au principe de l'extraction de propriétés de texture par des masques de convolution. Des solutions optimales et sous-optimales pour le choix des filtres sont proposées pour l'analyse et la classification de textures. On montre expérimentalement que la méthode est robuste et flexible. Pour la classification de textures, elle permet d'atteindre des performances aussi bonnes que les méthodes usuelles se basant sur des mesures de co-occurrences. De plus, elle donne lieu à une caractérisation des propriétés de texture avec un nombre moindre d'attributs; elle se prête également à un calcul plus aisé.

Keywords. Texture analysis, texture classification, linear transform, linear operators, filter bank, Karhunen-Loève transform.

1. Introduction

Texture is the term used to qualify the surface of a given object or phenomenon and is undoubtedly one of the main features used in image processing and pattern recognition. Texture must be regarded as a neighborhood property of an image point and has been widely studied during the last decade. A number of analysis methods which all try in some way to describe pixel neighborhood relationships have been suggested [16]. A fruitful approach, which has been studied by different authors, is to extract local neighborhood

* Present affiliation: Biomedical Engineering and Instrumentation Branch, Room 3W13, Building 13, National Institutes of Health, Bethesda, MD 20892, U.S.A.

information by means of linear filtering operators [1, 2, 11, 15, 24, 25, 31]. Thus, the local texture properties of a given region in an image can be characterized by a set of energy measures computed at the output of a filter bank; these convolution masks may be local detectors of elementary structures such as 'flats', 'spots', or 'edges' [24, 25], directional filters which attempt to model the human visual system [11, 15, 31] or linear operators obtained from a principal component analysis of the underlying texture field ('eigen-filters') [1, 2].

This work presents a statistical vector space formulation of this problem which leads to an extension and a unification of previous reported approaches. It suggests representing the Nth order probability density function (PDF) of the pixels in a restricted neighborhood by a set of N histograms evaluated after suitable linear transformation. The transform coefficients are computed over a sliding window and the resulting structure is equivalent to filtering the image with a bank of filters. The choice of a statistical argumentation enables the definition of optimal and sub-optimal linear operators for texture analysis and classification. This new formulation allows the comparison of various transforms (including Laws' set of operators [24]) in their ability to extract valuable texture information.

An important aspect of this research is the experimental evaluation of the proposed method, which is presented in Section 4. First of all, it appears that the performance in texture analysis and classification is not very different among various sub-optimal transforms. Second, the comparison with other well-known techniques indicate that equivalent—and sometimes improved—performance in texture classification can be obtained using the statistical information provided by channel histograms, instead of measures based on pairs of pixels (e.g., correlation coefficients, co-occurrence matrices). Concerning this last aspect, however, we question the experimental procedures and conclusions of some authors who state that "texture energy measures" used on their own are more powerful than measures based on pairs of pixels, in that all possible pairwise configurations in the considered neighborhood were not accounted for [24, 25].

Another important point of the present formulation is that it established the link between a filter bank analysis method and the transform techniques for which a large variety of fast algorithms are available. It should be emphasized, however, that the approach presented in this paper differs, in a number of ways, from the other digital transform methods that have previously been applied to texture analysis [5, 23]. The first and essential difference is that the statistics associated with individual transform coefficients are used as texture descriptors instead of the coefficients themselves (or some functions of the coefficients). The second difference concerns the choice of a relatively small domain to transform (typically 3×3 or 5×5). Finally, the transform coefficients are computed within a sliding window while in the other methods they are typically computed on non-overlapping square sub-images.

2. Texture description

Most statistical approaches to texture analysis provide a partial characterization of the joint probability of the pixels in a restricted neighborhood. In this section, the problem of describing local texture properties is restated as that of finding an adequate statistical description of the distribution of a random vector variable: the *local neighborhood vector*. This formation logically introduces a new method for texture characterization based on a local linear transformation of the local neighborhood vector. The motivation of such an operation is that it allows access to relatively compact statistical measures that are strongly related to the structure of the texture as it will be shown here.

2.1. Local neighborhood vector

A discrete texture image defined on a $K\times L$ rectangular grid is denoted by $\{x_{k,l}\}$, $(k=1,2,\ldots,K,\ l=1,2,\ldots,L)$ and is considered to be

the realization of a bidimensional stationary and ergodic stochastic process. Since texture is a neighborhood property, it is practical to consider the pixels in a given neighborhood index by (k, l) as the components of a local feature vector. For simplicity, the coupling neighborhood is defined as a rectangular $N_x \times N_y$ domain centered on (k, l). The N components of the local neighborhood vector $\boldsymbol{x}_{k,l} = [x_1, \ldots, x_N]^T$ are the sequentially ordered pixels belonging to the $N_x \times N_y$ rectangular domain indexed by (k, l). This formalism transforms the original grey level image $\{x_{k,l}\}$ in a multivariate sequence $\{\boldsymbol{x}_{k,l}\}$ $(k = 1, \ldots, K,\ l = 1, \ldots, L)$ as illustrated in Fig. 1.

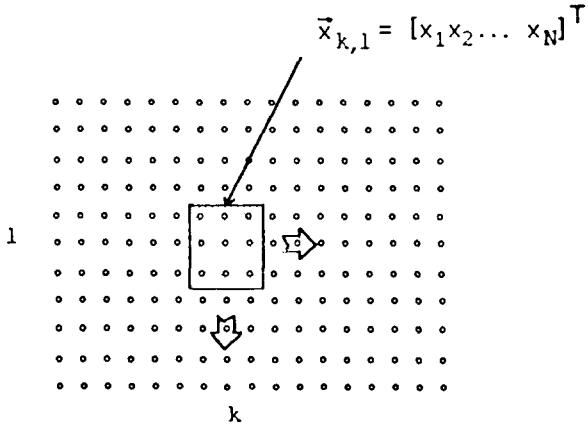

Fig. 1. Definition of the local neighborhood vector.

Within this framework, most commonly used texture measurements may be viewed as estimates of statistics associated with the local neighborhood vector. These statistics are usually estimated, by spatial averaging, from a large number of realizations of the local neighborhood obtained from a region of homogeneous texture in an image. These quantities can—at least conceptually—be obtained by integration (or summation) of the N-dimensional probability density function $p(\boldsymbol{x})$. For example, the grey level histogram provides an estimate of the first order PDFs of the individual components of the neighborhood vector. Note that these distributions are all the same as a direct consequence of stationarity. Estimates of the second order PDFs associated with different pairs of components of the neighborhood are the well-known co-occurrence matrices [17]. Because of stationarity, the total number of distinct second order PDFs for a $N_x \times N_y$ neighborhood reduces to $2N_xN_y - N_x - N_y$. The fact that this type of representation measures all visually useful textural information has been demonstrated by Gagalowicz [14]. Measurements based on co-occurrence statistics between pairs of pixels have been used successfully in a large number of applications. The main computational drawback of such methods is the requirement for large memory storage which depends on the second power of the total number of grey levels.

For practical reasons, it is almost impossible to estimate higher order PDFs unless a parametric model is used. The set of second moments can be used to construct the spatial covariance matrix which is defined as

$$\boldsymbol{C}_x = E\{(\boldsymbol{x} - E\{\boldsymbol{x}\}) \cdot (\boldsymbol{x} - E\{\boldsymbol{x}\})^T\}, \tag{1}$$

and which provides a sufficient statistic for a multivariate Gaussian model. This quantity is closely related to the spatial autocorrelation (or covariance) function of $\{x_{k,l}\}$. Because of stationarity, the components of the local neighborhood mean vector are all the same and the covariance matrix exhibits a close-to-Toeplitz structure with only $2N_xN_y - N_x - N_y + 1$ different entries. This approach to texture characterization is more or less equivalent to autocorrelation methods or Fourier power spectrum approaches [7, 10, 14, 22, 23, 27].

2.2 Local linear transform

An interesting alternative to the previously mentioned approaches is obtained from a linear transformation of the local neighborhood vector [28]. A local linear property extractor (or local linear transform) is defined by

$$\boldsymbol{y}_{k,l} = \boldsymbol{T}_N \cdot \boldsymbol{x}_{k,l}, \tag{2}$$

where $\boldsymbol{T}_N = [\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N]^T$ is a nonsingular $N \times N$ square matrix. This equation may lead to the two

following interpretations. First of all, it can be interpreted as a change of basis in the original pixel space: the y_i's are the coefficients of the local expansion of the neighborhood vector obtained with N linearly independent signals defined from the columns of the inverted matrix $\boldsymbol{T}_N^{-1}$. When $\boldsymbol{T}_N$ is an orthogonal matrix (orthogonal and normalized columns), this operation corresponds to a rotation or a reflection in the original pixel space. Secondly, the local linear transformation, being defined for every index (k, l), is equivalent to an N channel correlation (or convolution). The different rows of the transform matrix $\boldsymbol{T}_N$ define a set of masks that are used to filter the input texture $\{x_{k,l}\}$ producing N output images $\{y_{k,l}^{(i)}\}$ $(i = 1, \ldots, N)$. The corresponding system is depicted in Fig. 2; it is equivalent to a bank of N finite impulse response filters. Every channel will extract a particular aspect of local texture property. As it will be shown below, the efficiency of this analysis method will depend on the choice of the transform matrix $\boldsymbol{T}_N$.

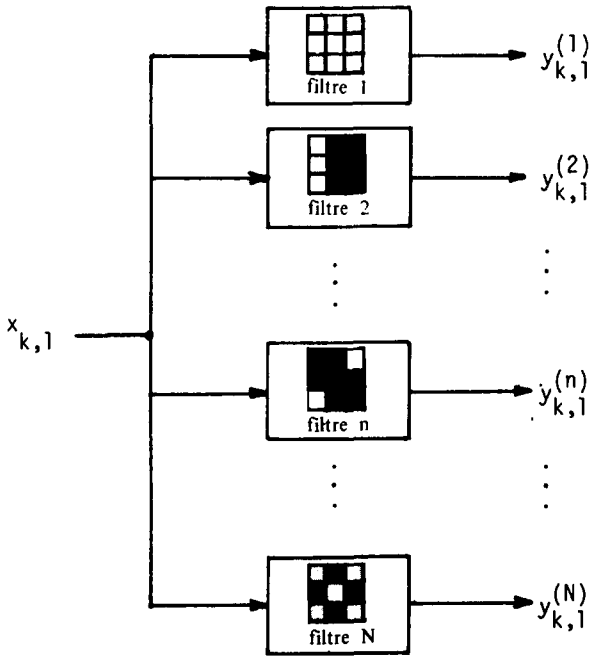

Fig. 2. Bank of N FIR filters for the analysis of texture by local linear transform.

2.3. A new set of texture measurements

First order PDFs of the neighborhood vector in the original basis do not convey any information about local structure or spatial arrangement (this assertion is supported by the fact that the grey scale of an image can be modified, by nonlinear mapping, to fit almost any specified histogram without intrinsically changing the local texture properties). In addition, because of stationarity, these quantities are all the same. After transformation, the first order PDFs of the components of $y_{k,l}$, being strongly affected by the dependence relationships between pixels, will be noticeably different from one another. The statistics for both representation of the local neighborhood vector are related to each other by the following relationships,

$$p_y(\boldsymbol{y}) = |\boldsymbol{T}_N|^{-1} \cdot p_x(\boldsymbol{x} = \boldsymbol{T}_N^{-1} \cdot \boldsymbol{y}), \tag{3}$$

$$\boldsymbol{C}_y = E\{(\boldsymbol{y} - E\{\boldsymbol{y}\}) \cdot (\boldsymbol{y} - E\{\boldsymbol{y}\})^{\mathrm{T}}\} = \boldsymbol{T}_N \boldsymbol{C}_x \boldsymbol{T}_N^{\mathrm{T}}, \tag{4}$$

where $p_x(\boldsymbol{x})$ and $p_y(\boldsymbol{y})$ are the joint PDFs, and $\boldsymbol{C}_x$ and $\boldsymbol{C}_y$ are the covariance matrices of the neighborhood vector before and after transformation. Equation (4) shows that the variances along the new axes are heavily dependent upon the covariance structure of the texture field. It can therefore be suspected that first order statistics after suitable transformation must provide a useful texture description. These quantities, which are estimated by spatial averaging at the output of the filter bank, will result in a simplified characterization of an unknown multidimensional PDF by means of its measured projections on different preselected axes. The *channel histogram* corresponding to the transform coefficient y_i quantified with N_c-grey levels is an estimate of the set of probabilities $\{P_q^{(i)}\}$ $(q = 1, \ldots, N_c)$:

$$P_q^{(i)} = \mathrm{Prob}\{y_i = q\}$$

$$(i = 1, \ldots, N,\ q = 1, \ldots, N_c). \tag{5}$$

An equivalent characterization in terms of moment

estimates can be used. The theoretical channel variances are defined by

$$\sigma_i^2 = \mathrm{Var}\{y_i\} = E[(y_i - E[y_i])^2] = \boldsymbol{t}_i^{\mathrm{T}} \boldsymbol{C}_x \boldsymbol{t}_i \quad (i = 1, \ldots, N) \tag{6}$$

and provide a sufficient description in the case of an underlying gaussian random field. The corresponding set of measurements characterizes the energy distribution at the output of the filter bank. The normalized and centered pth moments are defined by

$$\mu_p^{(i)} = E[(y_i - E[y_i])^p]/\sigma_i^p. \tag{7}$$

Of practical interest are the third (skewness) and fourth (Kurtosis) moments which respectively provide useful measures of the amount of skewness and peakedness in a distribution.

3. Transform selection

Local texture properties may be extracted using well-known transforms such as the discrete sine (DST), cosine (DCT), Hadamard (DHT) or Karhunen-Loève (KLT) transforms. It will be seen here that the selection of a given transform can be related to the neighborhood statistics of the textures to be treated. Two different problems for which optimal solutions will be derived are considered next. The first one is concerned with texture analysis where one wishes to find a representation that describes as well as possible the local texture properties. From a statistical view point, this is equivalent in choosing the set of first order statistics that provide the 'best' characterization of the Nth order PDF of the local neighborhood vector. The second problem is concerned with texture classification where the objective is to define a set of operators that can discriminate as well as possible between various texture fields. Of practical interest is the existence of sub-optimal transforms, which provide satisfactory results for both of these problems.

3.1. Optimal transform for texture analysis

The selection of a transform that is the most suited for this problem has to be evaluated with respect to some performance criterion. In this section, the discussion is restricted to the class of energy preserving transforms satisfying the condition:

$$\mathrm{tr}(\boldsymbol{C}_y = \boldsymbol{T}_N \cdot \boldsymbol{C}_x \cdot \boldsymbol{T}_N^{\mathrm{T}}) = \mathrm{tr}(\boldsymbol{C}_x) = N\sigma^2, \tag{8}$$

where σ^2 is the variance of the texture $\{x_{k,l}\}$. Two different aspects of optimality in the context of texture analysis are considered next.

- *Entropy criterion*: As has been mentioned before, the first order statistics associated with the components of the neighborhood vector, expressed in the usual basis, are all the same and therefore relatively uninteresting. Thus, we should choose the transform which produces first order statistics as 'different' as possible. This solution, which has the greatest difference in variance distribution, is the one that minimizes the entropy criterion [30]

$$H(\boldsymbol{T}_N) = -\sum_{i=1}^{N} \gamma_i \log(\gamma_i), \tag{9}$$

where the γ_i's measure the relative channel energy contributions:

$$\gamma_i = \mathrm{Var}\{y_i\}/\mathrm{tr}(\boldsymbol{C}_x) = \sigma_i^2/N\sigma^2$$

with $0 \leq \gamma_i \leq 1$.

It also follows from the well-known properties of entropy measures that $H(\boldsymbol{T}_N)$ will be maximum when the γ_i's are all equal; this least favorable case corresponds to the unity transform or any of its permutations: $H_{\max} = H(\boldsymbol{I}_N) = \log\{N\}$.

- *Energy criterion*: Another approach would be to choose the transform such that it produces non-correlated variables. This condition is fulfilled when the following energy criterion is maximized and equal to one:

$$E(\boldsymbol{T}_N) = \sum_{i=1}^{N} \mathrm{Var}\{y_i\}^2/\|\boldsymbol{C}_x\|^2 = \sum_{i=1}^{N} \sigma_i^4/\|\boldsymbol{C}_y\|^2, \tag{10}$$

where $\|\cdot\|^2$ represents the Hilbert-Schmidt norm (or energy) of a matrix and is invariant to any similarity transformation. As uncorrelatedness is a necessary condition—but not always sufficient—for independence, the Nth order PDF of the local neighborhood vector may be approximated by a product of first order PDFs obtained from this optimal representation. This approximation, which is exact in the case of a multivariate Gaussian distribution, has the same moments of degree 1 and 2 (mean and covariances) and also the same projections along the principal axes. On the other extreme, criterion (10) is minimized in the initial basis which corresponds to the most correlated representation, as a direct consequence of stationarity.

Both criteria are particular cases of a more general criterion function that has been introduced in [29]. Under the constraint of an energy preserving transform, it has been shown in [29] that the transform that optimizes this general criterion is given by $\boldsymbol{U}_N = [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_N]^T$, where the $\boldsymbol{u}_i$'s are solution of the characteristic equation:

$$\boldsymbol{C}_x \cdot \boldsymbol{u}_i = \lambda_i \cdot \boldsymbol{u}_i. \tag{11}$$

The optimal basis vectors are the eigenvectors of the spatial covariance matrix $\boldsymbol{C}_x$ and the optimal variances are the corresponding eigenvalues: $(\lambda_i, i = 1, \ldots, N)$. Thus, the optimal transform is the well-known Karhunen-Loève transform (KLT) which enables a decorrelated data representation. On the other hand, the worst representation, with respect to the general criterion, is obtained when the variance along the axes are all the same, which is precisely the case for the initial representation $\boldsymbol{x}_{k,l}$ associated to the identity transform. An important consequence is that any nontrivial transform $\boldsymbol{T}_N$ will always improve the efficiency when compared with the initial representation.

The set of eigenfilters obtained from the principal axes of the spatial covariance matrix has initially been proposed by Ade [1] for principal component analysis in the context of texture analysis. The ability of this set of masks to extract the different aspects and constituents of a given texture have been demonstrated in [2].

3.2. Optimal transform for texture classification

It is reasonable to assume that the most efficient channels for the discrimination between two textures ω_1 and ω_2 are those in which the corresponding energy contributions are as different as possible. Following this idea, a performance criterion associated with a $N \times N'$ transform matrix $\boldsymbol{T}_{N'} = [\boldsymbol{t}_1, \ldots, \boldsymbol{t}_{N'}]^T$ with $N' \leq N$ is introduced as

$$J_1(N') = \sum_{j=1}^{N'} (\sigma_{2j}^2 - \sigma_{1j}^2)^2 / (\sigma_{1j}^2 \sigma_{2j}^2)$$

$$= \sum_{j=1}^{N'} (\sigma_{1j}^2/\sigma_{2j}^2 + \sigma_{2j}^2/\sigma_{1j}^2 - 2) \geq 0, \tag{12}$$

where σ_{1j}^2 and σ_{2j}^2 $(j = 1, \ldots, N')$ are the channel variances or energy contributions corresponding to texture ω_1 and ω_2, respectively. In equation (12), the term corresponding to the relative energy contribution in channel no. j is symmetrical—with respect to indexes 1 and 2—and invariant to any scaling factor of the corresponding basis vector. This quantity is minimum and equal to zero when the variances σ_{1j}^2 and σ_{2j}^2 are the same; its importance grows as their respective values differ. Note that for a given $N \times N$ transform $\boldsymbol{T}_N$, the performance criterion (12) is maximized for any $N' < N$ when the row vectors $\boldsymbol{t}_i$ are indexed as

$$\sigma_{11}^2/\sigma_{21}^2 + \sigma_{21}^2/\sigma_{11}^2 \geq \cdots$$
$$\geq \sigma_{1N}^2/\sigma_{2N}^2 + \sigma_{2N}^2/\sigma_{1N}^2.$$

It is shown in Appendix A that the optimal transform of dimension $N \times N'$ that maximizes the criterion function $J_1(N')$, for $N' = 1, \ldots, N$, is given by $\boldsymbol{U}_{N'} = [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{N'}]^T$ where the row vectors $\boldsymbol{u}_i$ $(i = 1, \ldots, N)$ satisfy

$$\boldsymbol{C}_{x1} \cdot \boldsymbol{u}_i = \gamma_i \cdot \boldsymbol{C}_{x2} \cdot \boldsymbol{u}_i \quad (i = 1, \ldots, N), \tag{13}$$

where $\boldsymbol{C}_{x1}$ and $\boldsymbol{C}_{x2}$ are the texture spatial covariance matrices and where the eigensolutions of this equation $(\boldsymbol{u}_i, \gamma_i)$ $(i = 1, \ldots, N)$ are ordered in the

following manner:

$$\gamma_1 + 1/\gamma_1 \geq \gamma_2 + 1/\gamma_2 \geq \cdots \geq \gamma_N + 1/\gamma_N. \quad (14)$$

The corresponding maximum value of the criterion function $J_1(N')$ is

$$\max\{J_1(N')\} = \sum_{i=1}^{N'} (\gamma_i + 1/\gamma_i - 2). \quad (15)$$

A particular solution of equation (13) is obtained with a transform matrix V_N of dimension $N \times N$ which simultaneously diagonalizes the two spatial covariance matrices C_{x1} and C_{x2}. It is easily verified, by substitution, that the row-vectors of such a transform satisfy equation (13) and that the scalar values γ_i $(i = 1, \ldots, N)$ are then determined by

$$\gamma_i = \lambda_i^{\langle 1 \rangle} / \lambda_i^{\langle 2 \rangle} \quad (i = 1, \ldots, N), \quad (16)$$

where $\lambda_i^{\langle 1 \rangle}$ and $\lambda_i^{\langle 2 \rangle}$ are the channel variances after transformation for texture ω_1 and ω_2. It can be demonstrated that such a nonsingular transform matrix V_N exists for any symmetrical matrices C_{x1} and C_{x2}. However, this matrix is generally non-orthogonal and therefore not unique. An interesting solution for the simultaneous diagonalization of two covariance matrices has been proposed by Fukunaga and Koontz [13]; it has the remarkable property of producing complementary eigenvalues such that $\lambda_i^{\langle 1 \rangle} + \lambda_i^{\langle 2 \rangle} = 1$ $(i = 1, \ldots, N)$.

3.3. Sub-optimal transforms

The optimal sets of masks introduced in the preceding sections are primarily of theoretical value. They depend on the covariance structure of the textures to be analysed or classified and are usually cumbersome to determine: both solutions require the estimation of spatial covariance matrices and the use of standard—computationally expensive—numerical eigenvectors extraction methods. In the texture analysis problem, the optimal solution given by the KLT will be generally different from one texture field to another. This is also true for the solution given by (13) which, in addition, is difficult to generalize when more than two textures have to be considered. These approaches are therefore difficult to apply in practical applications such as texture classification or image segmentation. For these reasons, sub-optimal transforms should be used which can considerably simplify the feature extraction procedure. Another advantage is that fast algorithms are available for most of these transforms.

Usually, separable transforms, such as the discrete cosine (DCT) [4], sine (DST) [19, 20], real even and odd Fourier (DREFT and DROFT) [29] transforms, provide a close approximation of the KLT for a wide sense stationary process. A direct consequence is that these transforms will approximately diagonalize the spatial covariance matrix of a very large class of textures. It can therefore be suspected that their performance in texture analysis or classification is very close to the one that would be obtained with the optimal solutions given by (11) and (13), respectively. Furthermore, separable transforms can be computed by successive filtering along the rows and columns. This is illustrated in Fig. 3, which shows a separable filter bank that computes a running DST (discrete sine transform) in a 3×3 neighborhood.

The sets of masks associated with different transforms in the case of 3×3 are shown in Fig. 4. The set of masks introduced by Laws [24] is also represented. It is worthwhile to note the similarity of this particular set of operators and the sub-optimal DCT or DROFT with which it shares two basis vectors out of three. This fact is quite surprising as Laws followed a quite different reasoning than the one presented here and designed his filter bank empirically, using a combination of flat ($|1\ 2\ 1|$), spot ($|1\ -2\ 1|$) and edge ($|1\ 0\ -1|$) detectors; the corresponding basis vectors also define a local linear—but nonorthogonal—transform as defined in Section 2.2.

4. Experimental results

This section presents an experimental evaluation of the proposed method. The first part is concerned

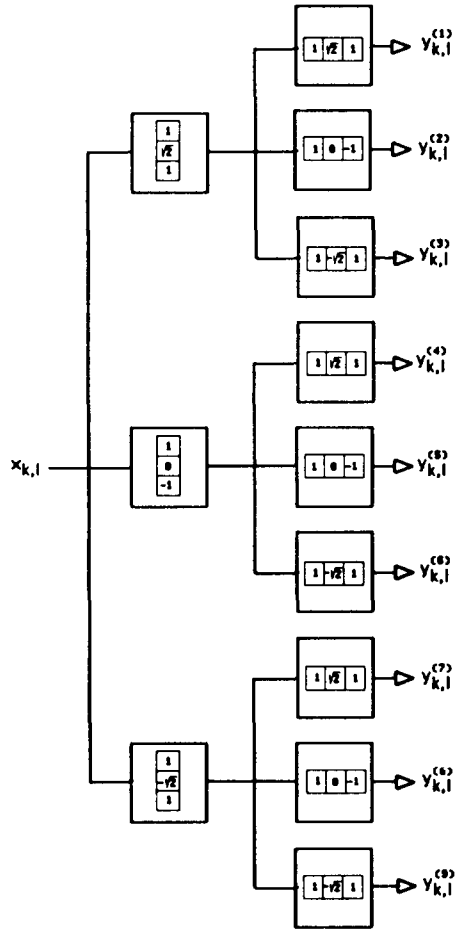

Fig. 3. Computation of a running 3×3 DST by successive row and column filtering.

with texture analysis and provides an illustration of the principle of extraction of textural properties by local linear transformation. The second part presents a detailed evaluation in the context of texture classification. In a first step, local texture properties are characterized by a set of N channel variances only. The performances of different sub-optimal transforms are then compared for various neighborhood sizes (2×2 up to 5×5). The effects of adding features such as third and fourth moments are also taken into account. Finally, the proposed method is shown to compare favorably with other well-known techniques such as correlation and co-occurrence based methods.

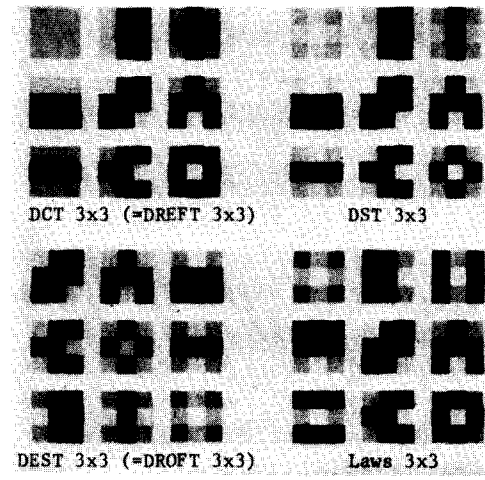

Fig. 4. Representation of the set of masks for various 3×3 sub-optimal separable transforms. The definitions of the DREFT (discrete real even Fourier transform), DROFT (discrete real odd Fourier transform) and DEST (discrete even sine transform) are given in [29].

The 12 Brodatz textures that have been used in this study and which are displayed in Fig. 5 were taken from [6]. The original photographs were digitalized and converted into 256×256 picture arrays quantified into 256 grey levels. Spatial transducer nonuniformities were compensated by a local normalization procedure over a 64×64 window [28]. A standard histogram flattening procedure was then performed producing output images with 32 equiprobable grey level values. The experimental data was therefore not distinguishable on the basis of first order statistics only.

4.1. Texture analysis

In these series of experiments a subset of the initial data (labeled (A), (B), and (C) in Fig. 5) has been considered more extensively. In a first step, the spatial covariance matrices were estimated over a 2×2 neighborhood and the associated KLT computed. The corresponding masks are shown in Fig. 6 and appear to be very similar to those obtained for the various sub-optimal transforms that have been mentioned previously, which in this case are all equivalent to a 2×2 discrete Hadamard transform (DHT). These sets of masks

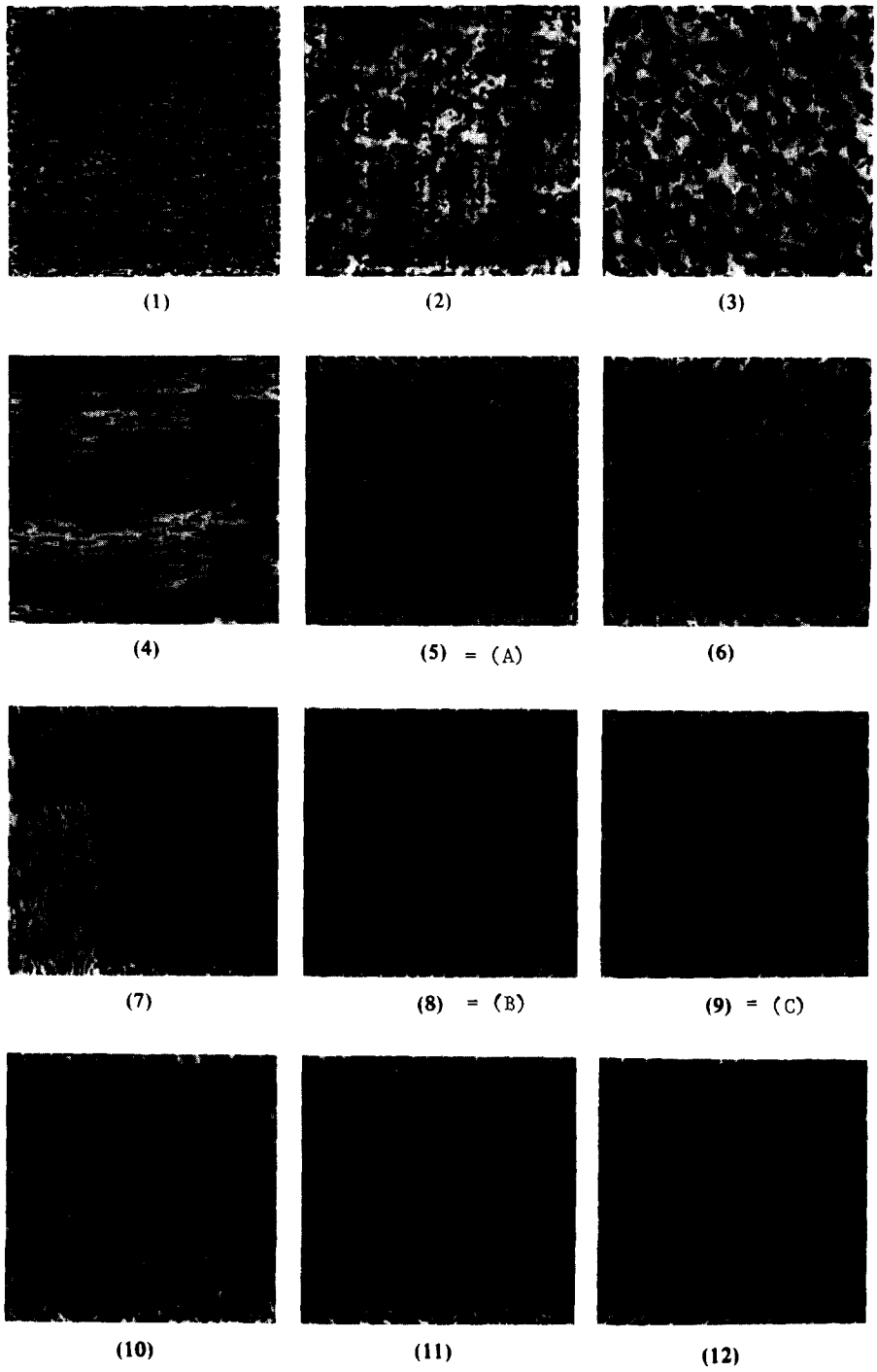

Fig. 5. Preprocessed Brodatz textures used for experiments (256×256 pixels with 32 equiprobable grey levels).

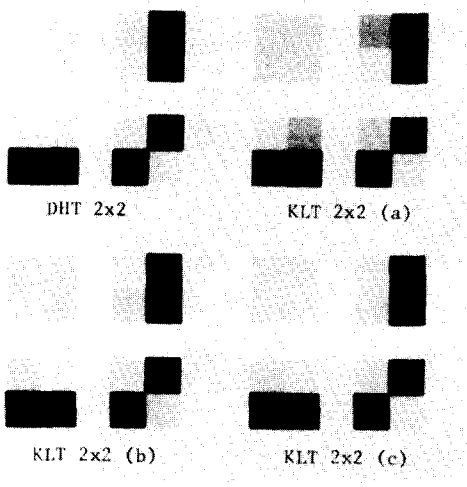

Fig. 6. Sub-optimal and optimal sets of masks (2×2) for analysis of textures (A), (B), and (C) in Fig. 5.

are all formed from a 'flat' detector (lowpass filter usually associated with the first component) and various edge detectors in the vertical horizontal and diagonal directions. The images were then filtered by a running 2×2 DHT; the result of this processing for texture (A) as well as the channel histograms are shown in Figs. 7 and 8, respectively. Also shown are the channel variances and the corresponding eigenvalues. From this experiment, it appears very clearly that the 2×2 DHT provides an excellent substitute for the KLT for a 2×2 neighborhood. This fact has also been verified by the testing of other textures. The quality of the approximation of the KLT have been measured by computing the quantities defined in equations (9) and (10). For both entropy and energy criteria, a relative performance measure has been defined as

$$\eta(\boldsymbol{T}_N) = \frac{\zeta^* - \zeta(\boldsymbol{T}_N)}{\zeta^* - \zeta(\boldsymbol{I}_N)}, \tag{17}$$

where ζ^*, $\zeta(\boldsymbol{T}_N)$, and $\zeta(\boldsymbol{I}_N)$ are the criterion values obtained with the KLT, the transform matrix $\boldsymbol{T}_N$ and the identity matrix, respectively. In the case of 2×2 neighborhood, typical values from 99.6% up to 100% have generally been obtained.

Various sub-optimal transforms were also compared with the optimal KLT for larger neighborhoods (3×3 up to 5×5). Table 1 gives the results of this computation in the case of a 3×3 neighborhood for textures (A), (B), and (C) in Fig. 5; more detailed results may be found in [28]. For these textures, the DST was always found to perform better than the other transforms and nearly as good as the optimal KLT. A coefficient close to 100% indicates that the channel variances are as differently distributed as possible as well as the transform coefficients being nearly decorrelated. The good performance of the DST in data representation is not surprising, as it is well known that this transform provides a good approximation of the KLT of a separable first order Markov when $\rho \geq 0.5$ (see, for example, [29]), that is, when the process is essentially highpass.

Table 1

Entropy and energy criterions computed on textures (A), (B), and (C) for various 3×3 local linear transforms

Texture (A)					
Transform		$H(T)$	$h(T)$	$E(T)$	$e(T)$
Identity	3×3	2.197	0.0%	40.298%	0.0%
DCT (DREFT)	3×3	1.598	91.738%	94.541%	90.856%
DEST (DROFT)	3×3	1.676	79.811%	89.035%	81.633%
DST	3×3	1.553	98.559%	99.491%	99.148%
KLT	3×3	1.544	100%	100%	100%
Texture (B)					
Transform		$H(T)$	$h(T)$	$E(T)$	$e(T)$
Identity	3×3	2.197	0.0%	49.994%	0.0%
DCT (DREFT)	3×3	1.805	90.840%	94.878%	89.758%
DEST (DROFT)	3×3	1.851	80.182%	90.221%	80.444%
DST	3×3	1.779	97.009%	98.481%	96.963%
KLT	3×3	1.766	100%	100%	100%
Texture (C)					
Transform		$H(T)$	$h(T)$	$E(T)$	$e(T)$
Identity	3×3	2.197	0.0%	56.634%	0.0%
DCT (DREFT)	3×3	1.904	94.730%	97.690%	94.687%
DEST (DROFT)	3×3	1.951	79.657%	90.347%	77.724%
DST	3×3	1.893	98.4411%	99.2094%	98.1756%
KLT	3×3	1.888	100%	100%	100%

This method has also been applied successfully to the control of textured surfaces and the detection of defects in textiles [3]. In this last study, it was found that sub-optimal transforms such as the DST

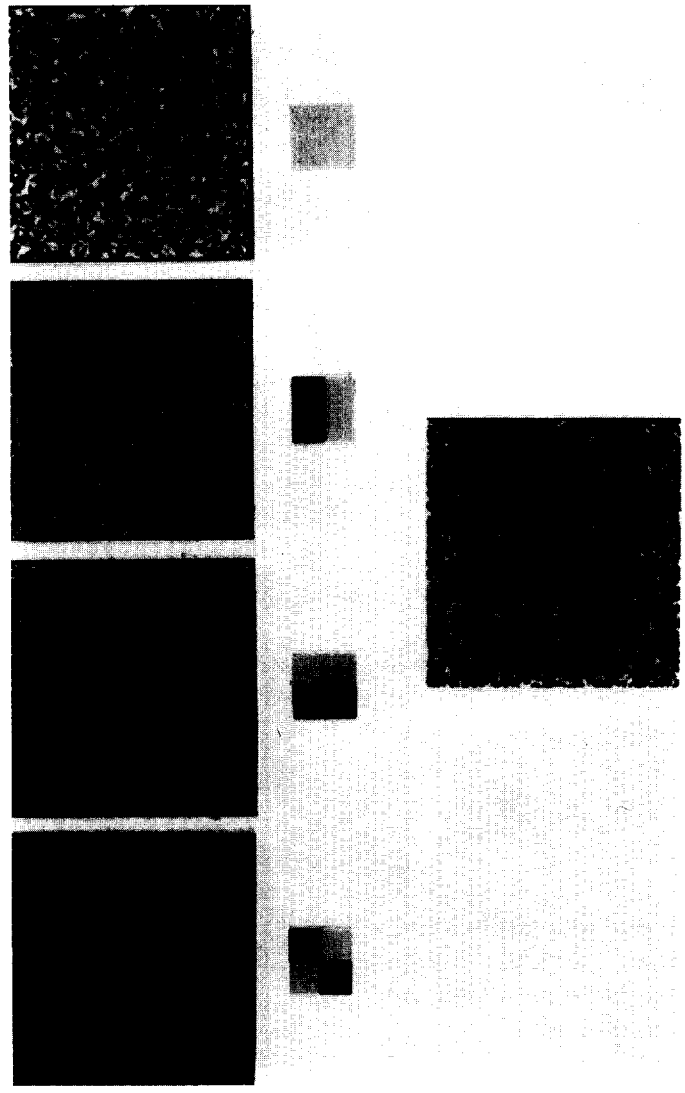

Fig. 7. Sub-optimal 2×2 DHT processing of texture (C): local transform coefficient images.

performed as well—and sometimes better—than the KLT for the detection of defects in materials. Very similar results were obtained with all sets of masks that have been tried; this indicates that the method is robust.

4.2. Texture classification

For classification purposes, the twelve textures images shown in Fig. 5 were then divided in square regions with 50% overlap. The experimental data set for each class consisted of 961 texture samples of dimension 16×16, 225 samples of dimension 32×32, and 49 samples of dimension 64×64.

An optimal Bayesian decision rule was used to classify the texture samples, based on their measured feature values. The class conditional probability density functions of the M-dimensional feature vector z were assumed to be multivariate Gaussian distributions with mean vectors and covariance matrices: (μ_i, C_i) $(i = 1, \ldots, 12)$. Under such an assumption, the Bayes

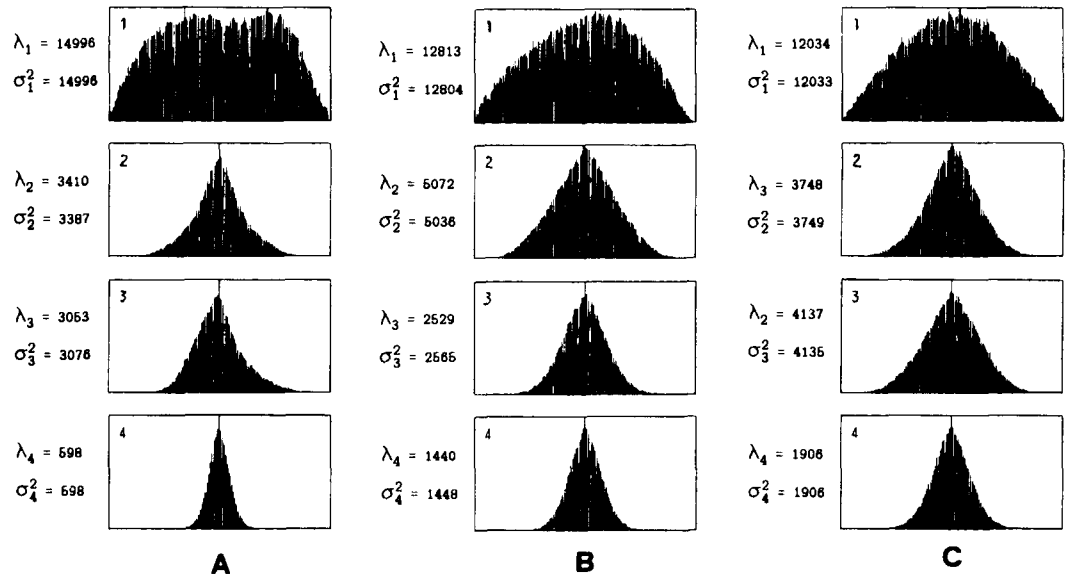

Fig. 8. Sub-optimal 2×2 DHT processing of textures (A), (B), and (C) in Fig. 5: channel histograms.

classifier, which minimizes the total probability of error, is equivalent to assign a texture sample with feature vector z to the class with minimum distance value [12]:

$$d_i(z) = (z-\mu_i)^T \cdot C_i^{-1} \cdot (z-\mu_i) + \log\{|C_i|\} \quad (i = 1, \ldots, 12). \tag{18}$$

For each pattern that has been tested, the training was performed on the remaining samples ('leaving-one-out method') using maximum likelihood estimates of the distribution parameters.

4.2.1. *Classification with channel variances*

Various sub-optimal local transforms have been compared in their ability to discriminate between the twelve textures in Fig. 5. The definition of the corresponding basis vectors may be found in [20, 26, 29]. The statistics of N-dimensional local neighborhood vector were characterized by a set of N estimated channel variances. The classification results for different neighborhood sizes are shown in Tables 2, 3, 4 and 5. As can be suspected, classification becomes more accurate as the size of the texture samples increases. For a given size of the local neighborhood, the performances of the transforms that have been considered

Table 2

Classification of the twelve textures in Fig. 5 using channel variances estimated after a local linear sub-optimal transform in a 2×2 neighborhood. M is the number of features, P_c the percentage of correct classification, and N_c the number of samples that have been correctly classified

Transform	M	Size	P_c	N_c
DHT 2×2	4	16×16	81.86%	9440 out ot 11 532
DHT 2×2	4	32×32	94.85%	2561 out of 2700
DHT 2×2	4	64×64	99.49%	585 out of 588

Table 3

Classification of the twelve textures in Fig. 5 using channel variances corresponding to various sub-optimal transform in a 3×3 neighborhood

Transform		M	Size	P_c	N_c
DST	3×3	9	16×16	88.66%	10 224 out of 11 532
DCT	3×3	9	16×16	88.58%	10 216 out of 11 532
LAWS	3×3	9	16×16	88.45%	10 200 out of 11 532
DROFT	3×3	9	16×16	88.41%	10 195 out of 11 532
DST	3×3	9	32×32	98.26%	2653 out of 2700
DCT	3×3	9	32×32	98.18%	2651 out of 2700
LAWS	3×3	9	32×32	98.18%	2651 out of 2700
DROFT	3×3	9	32×32	98.18%	2651 out of 2700
DST	3×3	9	64×64	100%	588 out of 588
DCT	3×3	9	64×64	100%	588 out of 588
LAWS	3×3	9	64×64	100%	588 out of 588
DROFT	3×3	9	64×64	100%	588 out of 588

Table 4

Classification of the twelve textures in Fig. 5 using channel variances corresponding to various sub-optimal transform in a 4×4 neighborhood

Transform		M	Size	P_c	N_c
DREFT	4×4	16	16×16	90.78%	10 469 out of 11 532
DROFT	4×4	16	16×16	90.61%	10 450 out of 11 532
DST	4×4	16	16×16	90.56%	10 444 out of 11 532
DCT	4×4	16	16×16	90.41%	10 426 out of 11 532
DEST	4×4	16	16×16	90.30%	10 413 out of 11 532
DREFT	4×4	16	32×32	99.07%	2675 out of 2700
DCT	4×4	16	32×32	98.96%	2672 out of 2700
DEST	4×4	16	32×32	98.93%	2671 out of 2700
DST	4×4	16	32×32	98.77%	2667 out of 2700
DROFT	4×4	16	32×32	98.74%	2666 out of 2700
DREFT	4×4	16	64×64	100%	588 out of 588
DCT	4×4	16	64×64	100%	588 out of 588
DEST	4×4	16	64×64	100%	588 out of 588
DST	4×4	16	64×64	100%	588 out of 588
DROFT	4×4	16	64×64	100%	588 out of 588

Table 5

Classification of the twelve textures in Fig. 5 using channel variances corresponding to various sub-optimal transform in a 5×5 neighborhood

Transform		M	Size	P_c	N_c
DST	5×5	25	16×16	92.03%	10 613 out of 11 532
DCT	5×5	25	16×16	91.96%	10 605 out of 11 532
DREFT	5×5	25	16×16	91.94%	10 602 out of 11 532
DROFT	5×5	25	16×16	91.93%	10 601 out of 11 532
DEST	5×5	25	16×16	91.78%	10 584 out of 11 532
LAWS	5×5	25	16×16	91.45%	10 546 out of 11 532
DST	5×5	25	32×32	99.63%	2690 out of 2700
DREFT	5×5	25	32×32	99.63%	2690 out of 2700
DCT	5×5	25	32×32	99.59%	2689 out of 2700
DROFT	5×5	25	32×32	99.59%	2689 out of 2700
DEST	5×5	25	32×32	99.56%	2688 out of 2700
LAWS	5×5	25	32×32	99.48%	2686 out of 2700
DST	5×5	25	64×64	100%	588 out of 588
DREFT	5×5	25	64×64	100%	588 out of 588
DCT	5×5	25	64×64	100%	588 out of 588
DROFT	5×5	25	64×64	100%	588 out of 588
DEST	5×5	25	64×64	100%	588 out of 588
LAWS	5×5	25	64×64	100%	588 out of 588

are very similar; this tends to indicate that the method is quite robust. In the case of 3×3 and 5×5 neighborhoods the best results are obtained with the sine (DST) and cosine (DCT) transforms. The good performance of the 4×4 DREFT—which in this particular case is equivalent to a 4×4 Hadamard transform—is quite surprising; this can be of practical interest because of the availability of very fast computation algorithms. The detailed results (confusion matrices) of the classification of 32×32 texture samples for various size of the local DST are given in Table 6. The effect of increasing the size of the neighborhood is to improve the classification results without introducing new errors. It is also worthwhile to note that classification errors usually occur in a symmetrical way, between texture pairs which are the most difficult to discriminate visually.

4.2.2. Classification with second, third, and fourth moments

In the present formulation, the use of second order moments or channel variances is only an interesting alternative to other well-known techniques which are based—in a more or less explicit way—on measurements that are directly related to the spatial covariance matrix. This category includes all methods that make use of power or energy measurements in the spatial [24, 25] or spectral domain [7], as well as autoregressive or predictive methods [10, 22, 27]. These techniques are especially well suited for the analysis and classification of Gaussian processes.

Unlike the previous mentioned approaches, the proposed method enables the measurement of some higher order statistics. A finer description of the local texture properties may be obtained by computing additional features such as the estimated normalized skewness and kurtosis defined on the channel histograms (equation (7)). The corresponding classification results are shown in Table 7 in the case of a 2×2 neighborhood. The addition of the third or fourth moments improves the performances significantly. As can be expected, the best results are obtained in using

Table 6

Confusion matrices for the classification of the twelve textures in Fig. 5 using channel variances estimated after a running 2×2, 3×3, and 4×4 discrete sine transform (DST)

	1	2	3	4	5	6	7	8	9	10	11	12
1	225	0	0	0	0	0	0	0	0	0	0	0
2	0	217	0	0	0	0	0	0	0	8	0	0
3	0	2	219	0	0	4	0	0	0	0	0	0
4	0	0	0	225	0	0	0	0	0	0	0	0
5	0	0	0	0	215	0	0	0	0	0	10	0
6	0	0	2	0	0	213	0	0	0	7	0	3
7	0	0	0	0	0	0	225	0	0	0	0	0
8	0	0	0	0	0	0	0	225	0	0	0	0
9	1	0	0	0	0	0	0	1	219	0	0	4
10	0	4	0	0	0	16	0	0	0	184	0	21
11	0	0	0	0	26	0	0	0	0	0	199	0
12	0	0	0	0	0	12	0	0	5	13	0	195

DST 2×2: Correct classification 2561 out of 2700.
Total score = 94.85%.

	1	2	3	4	5	6	7	8	9	10	11	12
1	225	0	0	0	0	0	0	0	0	0	0	0
2	0	224	0	0	0	0	0	0	0	1	0	0
3	0	0	223	0	0	2	0	0	0	0	0	0
4	0	0	0	225	0	0	0	0	0	0	0	0
5	0	0	0	0	225	0	0	0	0	0	0	0
6	0	0	3	0	0	218	0	0	0	2	0	2
7	0	0	0	0	0	0	225	0	0	0	0	0
8	0	0	0	0	0	0	0	225	0	0	0	0
9	0	0	0	0	0	0	0	0	225	0	0	0
10	0	1	0	0	0	6	0	0	0	200	0	18
11	0	0	0	0	0	0	0	0	0	0	225	0
12	0	0	0	0	0	1	0	0	1	10	0	213

DST 3×3: Correct classification 2653 out of 2700.
Total score = 98.26%.

	1	2	3	4	5	6	7	8	9	10	11	12
1	225	0	0	0	0	0	0	0	0	0	0	0
2	0	224	0	0	0	0	0	0	0	1	0	0
3	0	0	225	0	0	0	0	0	0	0	0	0
4	0	0	0	225	0	0	0	0	0	0	0	0
5	0	0	0	0	225	0	0	0	0	0	0	0
6	0	0	1	0	0	223	0	0	0	0	0	1
7	0	0	0	0	0	0	225	0	0	0	0	0
8	0	0	0	0	0	0	0	225	0	0	0	0
9	0	0	0	0	0	0	0	0	224	0	0	1
10	0	0	0	0	0	3	0	0	0	205	0	17
11	0	0	0	0	0	0	0	0	0	0	225	0
12	0	0	0	0	0	0	0	0	1	8	0	216

DST 4×4: Correct classification 2667 out of 2700.
Total score = 98.77%.

Table 7

Classification of the twelve textures in Fig. 5 using channel variances as well as third and fourth moments estimated after a local linear sub-optimal transform (DHT) in a 2×2 neighborhood

Transform	M	Size	P_c	N_c
		2nd order moments		
DHT 2×2	4	16×16	81.86%	9440 out of 11 532
DHT 2×2	4	32×32	94.85%	2561 out of 2700
DHT 2×2	4	64×64	99.49%	585 out of 588
		2nd and 3rd order moments		
DHT 2×2	8	16×16	86.17%	9937 out of 11 532
DHT 2×2	8	32×32	98.33%	2655 out of 2700
DHT 2×2	8	64×64	100%	588 out of 588
		2nd and 4th order moments		
DHT 2×2	8	16×16	84.23%	9713 out of 11 532
DHT 2×2	8	32×32	98.04%	2647 out of 2700
DHT 2×2	8	64×64	100%	588 out of 588
		2nd, 3rd, and 4th order moments		
DHT 2×2	12	16×16	87.63%	10 105 out of 11 532
DHT 2×2	12	32×32	99.19%	2678 out of 2700
DHT 2×2	12	64×64	100%	585 out of 588

conjointly second, third, and fourth moments. This shows very clearly that the textures used in these experiments are non-Gaussian and that, in such a case, there is an advantage in using higher order statistical information to improve the results in classification.

Various nonnormalized moments have also been tested on their own. It was found that even moments—the set of channel variances being the most powerful—always performed better than odd moments.

4.2.3. *Comparison with correlation methods*

Equation (6) shows that the knowledge of the spatial covariance matrix enables the determination of the power at the output of any FIR filter, as long as its support is entirely contained in the region defining the local neighborhood. This is also true when these quantities are estimated by spatial averaging. Therefore, the statistical information conveyed by the set of all correlation coefficients should be more complete than the one associated to the channel variances. Formally, these descriptions are only equivalent when the local linear transform diagonalizes the spatial covariance matrix. It is also important to mention that the channel variances provide a more compact description of second order statistical information (N channel variances versus $2(N-\sqrt{N})$ correlation coefficients for a $\sqrt{N}\times\sqrt{N}$ neighborhood).

In order to measure the loss of information due to sub-optimal transform processing, the texture samples have been classified based on measurements of the correlation coefficients. These quantities were normalized with respect to the texture variance; this last feature is meaningless for our purpose owing to the preprocessing which imposed the same first order statistics for all test images. These results are shown in Table 8 and must be considered as a upper bound of what can be achieved by means of second order statistics only. As would be expected, the performances are slightly better than those based on measurements of channel variances (Tables 2 and 3). However, this difference is very small, especially for a 2×2 neighborhood. In the case of 3×3 neighborhood, the loss of performance when using channel variances is only of the order of 1%, which is a very good result if one takes in account that nine features are used instead of twelve. This demonstrates experimentally that the sub-optimal transforms that have been considered are nearly as powerful as a hypothetical solution that would be optimal—in the sense of equation (13)—for all possible pairs of textures in Fig. 5. Furthermore, it appears that the results obtained in Section 4.2.2, using higher order moments, are much better than what could have been obtained by means of second order statistical measures only.

4.2.4. *Comparison with co-occurrence methods*

Spatial grey-level dependence or co-occurrence matrices are one of the most popular and powerful sources of features for texture characterization [17]. These matrices are usually considered as intermediate measures and are used to define more global texture features such as correlation, contrast, etc. If one considers a local neighborhood of dimension $N = N_x \times N_y$, it is possible to

Table 8

Classification of the twelve textures in Fig. 5 using the different correlations coefficients associated to a 2×2 and 3×3 neighborhood

Features	M	Size	P_c	N_c
2×2 Correlation coefficient	4	16×17	81.93%	9448 out of 11 532
2×2 Correlation coefficient	4	32×32	96.04%	2593 out of 2700
2×2 Correlation coefficient	4	64×64	99.49%	585 out of 588
3×3 Correlation coefficient	12	16×16	89.65%	10 338 out of 11 532
3×3 Correlation coefficient	12	32×32	99.37%	2683 out of 2700
3×3 Correlation coefficient	12	64×64	100%	588 out of 588

compute $2N_xN_y - N_x - N_y$ different co-occurrence matrices, each of them corresponding to a given relative displacement, which is usually specified by a distance value d and an orientation θ. These matrices provide estimates of the joint probability density functions for all different pairs of pixels in the considered domain. Obviously, this type of statistical characterization is more powerful than the corresponding set of correlation coefficients. The main drawback of this type of approach is its requirement for excessive memory storage and computation time.

In this study, we have restricted ourselves to a 2×2 neighborhood. The four corresponding symmetric co-occurrence matrices ($d = 1$ and $\theta = 0, 45, 90, 135$ degrees) have been computed on every texture sample and used to evaluate two different sets of features. The first set consists of the most commonly used features, namely Haralick et al.'s 'energy', 'entropy', 'correlation', and 'inertia' [17]. The second set of features, namely, 'cluster shade', 'cluster prominence', 'homogeneity', 'inertia', and 'correlation', has been proposed more recently by Conners et al. [8, 9]. As a matter of fact, both of these texture descriptions are more complete than the one that would be obtained measuring second moments only. The results in texture classification are given in Table 9 and have to be compared with the values in Table 7 obtained by local linear transformation. It is not surprising that co-occurrence features are more powerful than correlation or texture energy measures alone. However, it is worthwhile noticing that the percentages of correct classification obtained with the 2×2 DHT, using second, third, and fourth moments, compare favorably with those achieved with set no. 1. This is even more remarkable since the number of features—as well as the computational effort—is

Table 9

Classification of the twelve textures in Fig. 5 using co-occurrence features computed from the four co-occurrence matrices defined in a 2×2 neighborhood

Features	M	Size	P_c	N_c
Set 1 ($d = 1$, $\theta = 0, 45, 90, 135°$)	16	16×16	88.82%	10 243 out of 11 532
Set 1 ($d = 1$, $\theta = 0, 45, 90, 135°$)	16	32×32	97.74%	2639 out of 2700
Set 1 ($d = 1$, $\theta = 0, 45, 90, 135°$)	16	64×64	99.83%	587 out of 588
Set 2 ($d = 1$, $\theta = 0, 45, 90, 135°$)	24	16×16	91.19%	10 516 out of 11 532
Set 2 ($d = 1$, $\theta = 0, 45, 90, 135°$)	24	32×32	99.63%	2690 out of 2700
Set 2 ($d = 1$, $\theta = 0, 45, 90, 135°$)	24	64×64	100%	588 out of 588

significantly less important in the first case. The excellent results of set no. 2 are not surprising if one takes in account the effort that many researchers have spent, during the last decade, to improve feature extraction based on co-occurrence matrices. It is believed that a similar improvement could be achieved with an optimized set of measurements based on a linear transformation of the local neighborhood vector.

From these last experiments, it can be concluded that the proposed approach, providing that higher moments are computed, is almost as efficient for texture discrimination as methods based on co-occurrence measurements. Moreover, it seems preferable, in practical applications, to extract local texture properties using local linear transforms rather than co-occurrence matrices, for the following reasons:

(i) It enables a more compact description of local texture properties. It is also easy to adapt for various sizes of the local neighborhood.

(ii) The evaluation of the channel histograms or some associated moments is computationally less demanding.

(iii) Because of its parallel structure, this algorithm is especially suited for an implementation on a specialized parallel image processor; it is therefore easily applicable in the context of real time image processing.

(iv) The principle for the extraction of texture features by linear filtering is in agreement with recent physiological and psychological findings on the visual system [18, 21]. Furthermore, classification errors generally occur between textures that are the most difficult to differentiate visually; this is not necessarily the case when co-occurrence measurements are used.

5. Conclusion

It has been shown that local linear transforms can efficiently be applied to texture analysis and classification. The proposed method gives access to higher order statistical information by means of simple histogram or moment computation along selected axes in the space of pixel values in a specified neighborhood. Optimal sets of masks have been derived for texture analysis and two class texture classification. The practical importance of sub-optimal transform processing has been emphasized as it considerably simplifies computation and is applicable to cases where more than two textures have to be considered (N class texture classification problem or segmentation).

From experimental evaluation and comparison with other well-known techniques, texture measurements obtained from the channel histograms appear to have the following advantages which should make them useful in most practical applications:

- *Good performance*: Channel variances alone were found to be almost as powerful as the corresponding set of correlation coefficients. If higher order statistics are also extracted, using third and fourth moments, the resulting set of features yields as good a classification as standard co-occurrence measures based on pairs of pixels.
- *Compactness of the description*: Usually, the proposed approach uses less features than methods based on pairs of pixels. The number of channel histograms is always smaller than the number of different correlation coefficients or co-occurrence matrices than can be defined for the same neighborhood (for example, sixteen channel variances vs. twenty four correlation coefficients for a 4×4 neighborhood).
- *Robustness*: Performance in texture analysis or classification is not greatly affected by the exact shape of the masks which indicates that the method is robust.
- *Flexibility*: The quality of the texture description can be improved, according to the needs, by simply increasing the size of the local neighborhood.
- *Simple algorithmic structure*: Very fast and efficient algorithms are available for most sub-optimal transforms that have been considered. In addition, because of its vectorial form, the method is especially well suited for parallel implementation.

Acknowledgment

The author is grateful to Profs. Murat Kunt, Frederick De Coulon and Murray Eden for their support and valuable suggestions.

Appendix A. Optimal transform for two class texture classification

A generalized form of the performance criterion defined by equation (12) with a $N \times N'$ ($N' \leq N$) transform matrix $\boldsymbol{T} = [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{N'}]^T$ is given by

$$J = \sum_{j=1}^{N'} G\{\sigma_{2j}^2/\sigma_{1j}^2 + \sigma_{1j}^2/\sigma_{2j}^2\}, \tag{A.1}$$

where $G\{x\}$ is an increasing function for $x \geq 2$. Clearly, the term corresponding to the transformed coefficient j is minimum for $\sigma_{1j}^2 = \sigma_{2j}^2$. The more variances will differ from each other, the easier it will be to distinguish between textures ω_1 and ω_2. From this point of view, the most favorable transform is the one that maximizes $J(N')$.

The optimal basis vectors are solutions of

$$\partial J/\partial \boldsymbol{u}_j = 0 \quad (j = 1, 2, \ldots, N). \tag{A.2}$$

Because of the condition that $\partial G\{x\}/\partial x > 0$ for $x \geq 2$, the evaluation of the partial derivatives of $J(N)$ yields:

$$\boldsymbol{C}_{x1} \cdot \boldsymbol{u}_j / (\boldsymbol{u}_j^T \cdot \boldsymbol{C}_{x1} \cdot \boldsymbol{u}_j) = \boldsymbol{C}_{x2} \cdot \boldsymbol{u}_j / (\boldsymbol{u}_j^T \cdot \boldsymbol{C}_{x2} \cdot \boldsymbol{u}_j) \quad (j = 1, 2, \ldots, N) \tag{A.3}$$

which is equivalent to

$$\boldsymbol{C}_{x1} \cdot \boldsymbol{u}_j = \gamma_j \cdot \boldsymbol{C}_{x2} \cdot \boldsymbol{u}_j \quad (j = 1, 2, \ldots, N). \tag{A.4}$$

This is easily verified in multiplying this expression by $\boldsymbol{u}_j^T$. Thus, the value of γ_j is given by

$$\gamma_j = (\boldsymbol{u}_j^T \cdot \boldsymbol{C}_{x1} \cdot \boldsymbol{u}_j)/(\boldsymbol{u}_j^T \cdot \boldsymbol{C}_{x2} \cdot \boldsymbol{u}_j) \quad (j = 1, 2, \ldots, N), \tag{A.5}$$

which is in agreement with equation (A.3).

The optimal criterion value is found by substitution in (A.1) of the quantities given by equations (A.3) and (A.5):

$$\max\{J(N')\} = \sum_{j=1}^{N'} G\{\gamma_j + 1/\gamma_j\} = \sum_{j=1}^{N'} G\{\boldsymbol{u}_j^T \cdot (\boldsymbol{C}_{x1}^{-1} \cdot \boldsymbol{C}_{x2} + \boldsymbol{C}_{x2}^{-1} \cdot \boldsymbol{C}_{x1}) \cdot \boldsymbol{u}_j\}, \tag{A.6}$$

where the solutions $(\boldsymbol{u}_j, \gamma_j)$ $(i = 1, 2, \ldots, N)$ are ordered according to equation (14).

References

[1] F. Ade, "Application of principal component analysis to the inspection of industrial goods", *SPIE Conf.*, Geneva, 1983.

[2] F. Ade, "Characterization of texture by 'eigenfilter'", *Signal Processing*, Vol. 5, No. 5, September 1983, pp. 451-457.

[3] F. Ade, N. Lins and M. Unser, "Comparison of various filter sets for defect detection in textiles", *Proc. 7th Internat. Conf. on Pattern Recognition*, Montréal, Canada, July 30-August 2, 1984, pp. 428-431.

[4] N. Ahmed, T. Natarajan and K.R. Rao, "Discrete cosine transform", *IEEE Trans. Comput.*, Vol. C-23, January 1974, pp. 90-93.

[5] R. Bajcsy and L. Lieberman, "Texture gradient as a depth cue", *Computer Graphics and Image Process.*, Vol. 5, 1976, pp. 52-67.

[6] P. Brodatz, *Textures—A Photographic Album for Artists and Designers*, Dover, New York, 1966.

[7] C.H. Chen, "A study of texture classifications using spectral features", *Proc. 6th Internat. Conf. on Pattern Recognition*, München, 1982, pp. 382-385.

[8] R.W. Conners, "Toward a set of statistical features which measure visually perceivable qualities of textures", *Proc. Pattern Recognition Image Processing Conf.*, August 1979, pp. 382-390.

[9] R.W. Conners and C.A. Harlow, "Towards a structural textural analyser based on statistical methods", *Comput. Graphics and Image Process.*, Vol. 12, 1980, pp. 224-256.

[10] P. De Souza, "Texture recognition via autoregression", *Pattern Recognition*, Vol. 15, No. 6, 1982, pp. 471-475.

[11] O.D. Faugeras, "Texture analysis and classification using a human visual model", *Proc. Internat. Joint Conf. on Pattern Recognition*, Kyoto, Japan, November 1978, pp. 549-559.

[12] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, Academic Press, New York, 1972.

[13] K. Fukunaga and W.L. Koontz, "Application of the Karhunen-Loève expansion to feature selection and ordering", *IEEE Trans. Comput.*, Vol. C-19, No. 4, April 1970, pp. 311-318.

[14] A. Gagalowicz, "Vers un modèle de textures", Thèse de Doctorat des Sciences, Université de Paris VI, France, 1983.

[15] G.H. Granlund, "Description of texture using the general operator approach", *Proc. 5th Internat. Conf. on Pattern Recognition*, 1980, pp. 776-779.

[16] R.M. Haralick, "Statistical and structural approaches to texture", *Proc. IEEE*, Vol. 67, 1979, pp. 786-804.

[17] R.M. Haralick, K. Shanmugan and I. Dinstein, "Textural features for image classification", *IEEE Trans. Systems, Man and Cybern.*, Vol. SMC-8, No. 6, November 1973, pp. 610-621.

[18] D.H. Hubel and T.N. Wiesel, "Brain mechanism of vision", *Scientific American*, Vol. 241, No. 3, September 1979, pp. 130-146.

[19] A.K. Jain, "A fast Karhunen-Loève transform for a class of stochastic processes", *IEEE Trans. Commun.*, Vol. COM-24, September 1976, pp. 1023-1029.

[20] A.K. Jain, "A sinusoidal family of unitary transforms", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. PAMI-1, No. 4, October 1979, pp. 356-365.

[21] B. Julesz and R. Bergen, "Textons, the fundamental elements in preattentive vision and perception of textures", *The Bell System Tech. J.*, Vol. 62, No. 6, July/August 1983, pp. 1619-1645.

[22] R.L. Kashyap, R. Chellappa and A. Khotanzad, "Texture classification using features derived from random field models", *Pattern Recognition Lett.*, Vol. 1, 1982, pp. 43-50.

[23] L. Kirvida, "Texture measurements for automatic classification of imagery", *IEEE Trans. EMC*, Vol. 18, No. 2, 1976, pp. 38-42.

[24] K.I. Laws, "Textured image segmentation", Ph.D. Thesis, Rept. 940, Image Processing Institute, Univ. of Southern California, January 1980.

[25] M. Pietikaïnen, A. Rosenfeld and L.S. Davis, "Experiments with texture classification using averages of local pattern matches", *IEEE Trans. on Systems, Man and Cybernet.*, Vol. SMC-13, No. 3, 1983, pp. 421-426.

[26] W.K. Pratt, *Digital Image Processing*, Wiley-Interscience, New York, 1978.

[27] C.W. Therrien, "Linear filtering models for texture classification and segmentation", *Proc. 5th Internat. Joint Conf. on Pattern Recognition and Image Processing*, Miami, 1980, pp. 1132-1135.

[28] M. Unser, "Description statistique de textures: Application à l'inspection automatique", Ph.D. Thesis No. 534, Ecole Polytechnique Fédérale de Lausanne, Switzerland, 1984.

[29] M. Unser, "On the approximation of the Karhunen-Loève transform of stationary processes", *Signal Processing*, Vol. 7, No. 3, December 1984, pp. 231-249.

[30] S. Watanabe, "Karhunen-Loève expansion and factor analysis", *Trans. 4th Prague Conf. on Information Theory*, 1965, Czechoslovak Academy of Sciences, Prague, 1967.

[31] D. Wermser and C.-E. Liedtke, "Texture analysis using a model of the visual system", *6th Internat. Conf. on Pattern Recognition*, Münich, Germany, October 19-22, 1982, pp. 1078-1081.