PERGAMON

Pattern Recognition 32 (1999) 443–451

PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY

Wavelet correlation signatures for color texture characterization

G. Van de Wouwer*, P. Scheunders, S. Livens, D. Van Dyck

Vision Lab, Department of Physics, University of Antwerp, Groenenborgerlaan 171, 2020 Antwerpen, Belgium

Received 19 December 1997

Abstract

In the last decade, multiscale techniques for gray-level texture analysis have been intensively used. In this paper, we aim to extend these techniques to color images. We introduce wavelet energy-correlation signatures and we derive the transformation of these signatures upon linear color space transformations. Experiments are conducted on a set of 30 natural colored texture images in which color and gray-level texture classification performances are compared. It is demonstrated that the wavelet correlation features contain more information than the intensity or the energy features of each color plane separately. The influence of image representation in color space is evaluated. © 1999 Pattern Recognition Society. Published by Elsevier Science Ltd. All rights reserved.

Keywords: Texture analysis; Classification; Color spaces; Feature extraction; Wavelet signatures

1. Introduction

For image analysis, color and texture are two of the most important properties, especially when one is dealing with real-world images. Classical image analysis schemes only take into account the pixel graylevels, which represents the total amount of visible light at the pixel's position. The performance of such schemes can be improved by adding color information [1]. The color of a pixel is typically represented with the RGB tristimulus values, each corresponding to the red, green and blue frequency bands of the visible light spectrum. Color is then a feature in the three-dimensional RGB color space, which contains information regarding the spectral distribution of light complementary to the gray-level information.

An important topic when processing color images is their representation. The RGB representation is frequently being transformed into other color spaces. A large variety of (linear and non-linear) transforms and standard color spaces can be found in the literature [2,3]. The performance of an image analysis system can strongly depend on the choice of the color representation [4,5]. However, there does not appear to be a systematic means of determining an optimum color-coordinate system for a particular task.

In the analysis of color images, the description of image regions has mainly been performed using color histograms [3,6]. When local spatial correlations are important to characterize a region, color histograms no longer suffice. The extra information needed to adequately describe the image regions is commonly known as "texture". Texture has been studied extensively and many texture analysis schemes have been proposed [7,8]. The fundamental property which they all have in common is

*Corresponding author. Fax: 32 3-218.03.18 E-mail: wouwer@ruca.ua.ac.be

PII: S0031-3203(98)00035-1

that they exploit spatial interactions between the pixels of a neighborhood.

A rather limited number of systems use combined information of color and texture, and even when they do, both aspects are mostly dealt with using separate methods [9,10]. It is only recently that attempts are being made to combine both aspects in a single method, by extending gray-level texture analysis methods to color images. This combination can be made more formal by defining "color-texture" as "the set of local statistical properties of the colors of image regions". Efficient characterization of color texture requires the exploitation of spatial correlations as well as correlations between color bands.

A number of researchers have performed experiments on color texture analysis. Caelli and Reye have proposed a method in which they extract features from three spectral channels by using three multiscale isotropic filters [11]. A related approach by Tan and Kittler extracts features from three channels with a discrete cosine transform [12]. Both methods did not consider correlations between spectral bands. Some recent methods have tried to include this. The well-known occurrence matrix method has been investigated by Hauta-Kasari et al. [13]. A Markov random field model for color textures has been proposed by Panjwani and Healey [14]. Some work closely related to this has elaborated on invariants with respecto the model [15,16]. Kittler et al. have studied pseudo-Wigner distributions and chromato-structural features to detect defects in colored texture surfaces [17].

Despite these examples, work on color texture analysis is still very small. The importance of a joint color-texture characterization is, however, expected to grow rapidly in the near future, e.g. for indexing image databases. At present the color extensions of several major texture analysis methods are still unexplored. We will investigate those methods, namely the wavelet multiresolution decomposition.

Multiresolution representations give rise to an interesting class of texture analysis methods. Strong arguments for their use can be found in psychovisual research, which offers evidence that the human visual system processes images in a multiscale way [18]. Wavelets provide a convenient way to obtain a multiresolution representation [19,20], from which texture features are easily extracted. The so-called energy signatures have proven to be very powerful for texture analysis [21–24].

We propose a scheme for the classification of colored texture images. Feature extraction using wavelet decomposition is described. Wavelet correlation signatures are defined which contain the energies of each color plane, together with the cross correlation between different planes. While the first have already been successfully used for texture characterisation, the second represents the coupling between texture and color. We will show that these features transform linearly upon linear color space transformation. The experiments will demonstrate the usefulness of correlation signatures as texture features. The influence of the choice of color space representation on classification performance will be investigated.

The outline of this paper is as follows: in the next section the use of the wavelet transform for feature extraction is discussed and wavelet signatures are introduced. Section 3 elaborates on various color space transforms and investigates the behavior of the wavelet signatures upon such transforms. Section 4 briefly explains the techniques used for image classification and in Section 5 several experiments are conducted and their results discussed. In the last section some conclusions are drawn.

2. Wavelet signatures

The (continuous) wavelet transform of a 1-D signal $f(x)$ is defined as

$$(W_a f)(b) = \int f(x)\psi^*_{a,b}(x)\,\mathrm{d}x, \tag{1}$$

where the wavelet $\psi_{a,b}$ is computed from the *mother* wavelet ψ by translation and dilation

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}}\psi\left(\frac{x-b}{a}\right) \tag{2}$$

The mother wavelet ψ has to satisfy the admissibility criterion to ensure that it is a localized zero-mean function. This transform can be discretized by restraining a and b to a discrete lattice ($a = 2^n$, $b \in Z$). Typically, some more constraints are imposed on ψ to ensure that the transform is non-redundant, complete and constitutes a multiresolution representation of the original signal. This has led to an efficient real-space implementation of the transform using quadrature mirror filters.

The extension to the 2-D case is usually performed by using a product of 1-D filters. In practice, the transform is computed by applying a separable filter bank to the image:

$$L_n(\mathbf{b}) = [H_x * [H_y * L_{n-1}]_{\downarrow 2,1}]_{\downarrow 1,2}(\mathbf{b}) \tag{3}$$

$$D_{n1}(\mathbf{b}) = [H_x * [G_y * L_{n-1}]_{\downarrow 2,1}]_{\downarrow 1,2}(\mathbf{b}) \tag{4}$$

$$D_{n2}(\mathbf{b}) = [G_x * [H_y * L_{n-1}]_{\downarrow 2,1}]_{\downarrow 1,2}(\mathbf{b}) \tag{5}$$

$$D_{n3}(\mathbf{b}) = [G_x * [G_y * L_{n-1}]_{\downarrow 2,1}]_{\downarrow 1,2}(\mathbf{b}) \tag{6}$$

where $\mathbf{b} \in R^2$, $*$ denotes the convolution operator, $\downarrow 2,1$ ($\downarrow 1,2$) sub-sampling along the rows (columns) and $L_0 = I(\mathbf{x})$ is the original image. H and G are a low- and band-pass filter, respectively. L_n is obtained by low-pass filtering and is therefore referred to as the low-resolution image at scale n. The D_{ni} are obtained by band-pass

filtering in a specific direction and thus contain directional detailed information at scale n; they are referred to as the *detail images*. The original image I is thus represented by a set of subimages at several scales; $\{L_d, D_{ni}\}_{i=1,2,3}^{n=0,\ldots,d-1}$ which is a *multiscale representation of depth d* of the image I.

The energy of a subimage D_{ni} is defined as

$$E_{ni} = \int (D_{ni}(\mathbf{b}))^2 \, d\mathbf{b}. \tag{7}$$

The *wavelet energy signatures* $\{E_{ni}\}_{n=0,\ldots,d-1,i=1,2,3}$ reflect the distribution of energy along the frequency axis over scale and orientation and have proven to be very useful for gray-level texture characterization. Since most relevant texture information has been removed by iterative low-pass filtering, the energy of the low-resolution image L_d is generally not considered a texture feature.

The most straightforward extension of the wavelet energy signatures to color images is to transform each color plane separately and extract the energies of each transformed plane; i.e. replace I by the R,G and B-plane consecutively in Eqs. (3)–(6). We denote such an energy by $E_{ni}^{X_j}$ where the X_j indicates the color plane. This triples the amount of features w.r.t. the gray-level case.

Let us define

$$C_{ni}^{X_jX_k} = \int D_{ni}^{X_j}(\mathbf{b}) D_{ni}^{X_k}(\mathbf{b}) \, d\mathbf{b} \tag{8}$$

and call the set $\{C_{ni}^{X_jX_k}\}_{n=0,\ldots,d-1,i=1,2,3}^{j,k=1,2,3,j\leq k}$ the *wavelet covariance signatures*. They include the energies for $j = k$; the others represent the covariance between different color planes and consequently the coupling between the color and texture properties of the image.

The covariance signatures, however, are by definition proportional to the energies. To remove this redundant information, the covariance signatures are normalized in the following way:

$$\tilde{C}_{ni}^{X_jX_k} = \begin{cases} E_{ni}^{X_j} & j = k, \\ \dfrac{C_{ni}^{X_jX_k}}{E_{ni}^{X_j}E_{ni}^{X_k}} & j \neq k. \end{cases} \tag{9}$$

The features $\{\tilde{C}_{ni}^{X_jX_k}\}_{n=0,\ldots,d-1,i=1,2,3}^{j,k=1,2,3,j\leq k}$ are the *wavelet correlation signatures*.

3. Color space transforms

For compression purposes, transformations to different color spaces are often employed to achieve image bandwidth reduction without significantly degrading image quality. However, since our goal is to efficiently characterize texture, the choice of color space should enable the extraction of useful features rather than visual image representation. Linear as well as non-linear color space transforms can be applied. Non-linear ones are mainly employed to obtain a color space in which the coordinates have an intuitive meaning (mostly a luminance, a saturation and a hue component) [25]. They typically introduce some non-removable singularities, which is very impractical for further processing. We will limit ourselves to linear color space transforms, i.e.

$$X' = MX, \tag{10}$$

where $X = (X_1(\mathbf{x})X_2(\mathbf{x})X_3(\mathbf{x}))^\tau$ contains the original components of the signal ($^\tau$ means transpose), M is a 3×3 invertible transformation matrix and X' contains the transformed signal.

Three color space transforms (for which $X = (R\,G\,B)^\tau$) are studied in this paper:

(1) The UVW-space; CIE uniform chromaticity scale (V = Y = luminance):

$$\begin{pmatrix} U \\ V \\ W \end{pmatrix} = \begin{pmatrix} 0.405 & 0.116 & 0.133 \\ 0.299 & 0.587 & 0.114 \\ 0.145 & 0.827 & 0.627 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}. \tag{11}$$

This is a "perceptually uniform" space constructed so that equal changes in the space are experienced as equal changes in color by human perception.

(2) The YIQ-space (NTSC transmission primaries):

$$\begin{pmatrix} Y \\ I \\ Q \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.253 & 0.312 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}. \tag{12}$$

The Y signal is the image luminance and the I and Q signals carry the chrominance information.

(3) The K–L space (Karhunen–Loève transform):

$$\begin{pmatrix} K_1 \\ K_2 \\ K_3 \end{pmatrix} = \begin{pmatrix} 0.333 & 0.333 & 0.333 \\ 0.500 & 0.000 & -0.500 \\ -0.500 & 1.000 & -0.500 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}. \tag{13}$$

It may look surprising that a fixed transform is called Karhunen–Loève. Indeed, the K–L transform is formed by the eigenvector of the correlation matrix of a specific image and is thus image-dependent. However, as reported by Ohta [4], the eigenvector remain approximately the same for a large set of natural color images. This was confirmed experimentally on our own image database. As expected, we found that the largest part of the variance is concentrated around the K_1-axis (which, in fact, represents the intensity). The remainder is almost completely found along the K_2-axis, leaving the last axis without significant variance. Eq. (13) transforms an image into an orthogonal basis in which the axes are statistically uncorrelated, and in that sense decorrelates the information present in RGB space.

3.1. Effect of linear color space transform on the wavelet signatures

We now investigate how the wavelet covariance signatures transform under a linear color space transform. For $a = 2^n$, let us rewrite Eq. (1) in 2-D for the case of a separable wavelet:

$$(W_{2^n,i} f)(\mathbf{b}) = \int f(\mathbf{x}) \varphi^{i*}_{2^n,b}(\mathbf{x})\, \mathrm{d}x, \tag{14}$$

where $\{\varphi^i_{2^n,b}\}_{i=1,2,3}$ are the wavelets to obtain the detailed images D_{ni}. Let us fix n and i for the remainder of this paragraph. For a color image $X = (X_1(\mathbf{x})X_2(\mathbf{x})X_3(\mathbf{x}))^\tau$, Eq. (14) reads in vector notation:

$$(W_{2^n,i} X)(\mathbf{b}) = \int X(\mathbf{x})\, \varphi^{i*}_{2^n,\mathbf{b}}(\mathbf{x})\, \mathrm{d}\mathbf{x}. \tag{15}$$

Further, define C_{ni} as a matrix with the wavelet covariance signatures as elements:

$$C_{ni} = \int (W_{2^n,i} X)(\mathbf{b})(W_{2^n,i} X)(\mathbf{b}))^\tau\, \mathrm{d}\mathbf{b}. \tag{16}$$

After a color space transformation $X' = MX$ the covariance signatures become

$$\begin{aligned}
C'_{ni} &= \int (W_{2^n,i} X')(\mathbf{b})(W_{2^n,i} X')(\mathbf{b}))^\tau\, \mathrm{d}b \\
&= \left(\int X'(\mathbf{x}) \varphi^{i*}_{2^n,\mathbf{b}}(\mathbf{x})\, \mathrm{d}\mathbf{x} \right) \left(\int X'(\mathbf{x}) \varphi^{i*}_{2^n,\mathbf{b}}(\mathbf{x})\, \mathrm{d}\mathbf{x} \right)^\tau \mathrm{d}\mathbf{b} \\
&= \int \left(\int MX(\mathbf{x}) \varphi^{i*}_{2^n,\mathbf{b}}(\mathbf{x})\, \mathrm{d}\mathbf{x} \right. \\
&\quad \left. \times \left(\int MX(\mathbf{x}) \varphi^{i*}_{2^n,\mathbf{b}}(\mathbf{x})\, \mathrm{d}\mathbf{x} \right)^\tau \mathrm{d}\mathbf{b} \right. \\
&= M \left[\int \left(\int X'(\mathbf{x}) \varphi^{i*}_{2^n,\mathbf{b}}(\mathbf{x})\, \mathrm{d}\mathbf{x} \right) \right. \\
&\quad \left. \times \left(\int X'(\mathbf{x}) \varphi^{i*}_{2^n,\mathbf{b}}(\mathbf{x})\, \mathrm{d}\mathbf{x} \right)^\tau \mathrm{d}\mathbf{b} \right] M^\tau \\
&= M C_{ni} M^\tau
\end{aligned} \tag{17}$$

Or, explicitly

$$C^{X_j,X_k}_{ni} = \sum_{r,s=1}^{3} m_{jr} m_{ks} C^{X_r X_s}_{ni}. \tag{18}$$

For the energies this means (taking the RGB-space for X):

$$\begin{aligned}
E^{X_j}_{ni} = C^{X_j,X_j}_{ni} = \sum_{r,s=1}^{3} m_{jr} m_{js} C^{X_r X_s}_{ni} &= m^2_{j1} E^R_{ni} \\
&+ m^2_{j2} E^G_{ni} + m^2_{j3} E^B_{ni} + 2 m_{j1} m_{j2} C^{RG}_{ni} \\
&+ 2 m_{j1} m_{j3} C^{RB}_{ni} + 2 m_{j2} m_{j3} C^{GB}_{ni}.
\end{aligned} \tag{19}$$

These formulas offer an interesting insight into the effect of linear color space transform on the wavelet signatures. Eq. (18) shows that a linear color space transform implies a linear transform of the covariance signatures. However, from Eq. (19) it follows that this is not true for the energy signatures. The first three terms reveal that the "new" energy features are linearly obtained from the "old" ones; the next three terms, however, depend on the covariances between the R, G and B planes for the same subimage. There is no clear connection between the energies in the original and transformed color space; to compute the latter the wavelet covariance signatures are required.

Eq. (19) also shows that performing a simple linear transform from RGB space to another color space results in a clearly different feature set. Hence, the quality of the features (i.e. their ability to characterise and discriminate between color textures) shall be heavily dependent on the choice of color space. This shall be demonstrated in the experimental section.

When one experiments using several color transforms, a practical advantage of relation (18) comes into play. It is sufficient to perform the wavelet transform once (for the R, G and B planes) and to compute the covariance signatures. The new wavelet signatures are then obtained using Eq. (18) without the need for performing several wavelet transforms.

For the correlation signatures, the simple relation (18) does not hold. To transform correlation signatures into other color spaces, it is therefore convenient to transform the covariance signatures first and to normalize them afterwards.

4. Classification methods

Once an appropriate set of texture features is computed, the next step is to adopt a suitable classification algorithm to assess the features' discriminative power. This section describes briefly the methods used to design and evaluate a *knn*-classifier. These subjects are treated more in detail in standard pattern recognition literature [26,27].

For each different feature set, a *k-nearest neighbor classifier* (*knn*-classifier) was designed. This is a *supervised* classifier, i.e. it requires a set of labeled feature vectors derived from the available data samples as a *design set*. Classification of feature vector $\mathbf{x}$ is performed by searching the k closest design vectors according to some metric $d(\mathbf{x}, \mathbf{y})$. The vector $\mathbf{x}$ is assigned to that class to which the majority of these k nearest neighbors belong. More complex classification schemes such as neural networks could be considered, but these typically depend on initialisation and learning time, while *knn* classifiers do not. Since the emphasis in this work is on the feature extraction stage, *knn* provides an efficient and robust classification scheme

for evaluation of classifier performance and comparison of feature sets.

If the distance measure d is chosen to be Euclidean, then some features (the ones with largest variances) tend to dominate this measure. This could be resolved by employing other distance measures (e.g. the Mahalanobis distance), but we have chosen to normalize the features so that they all have equal variances.

To evaluate the performance of the *knn*-classifier, one desires to know how well it classifies "unseen" data, i.e. data not used to design the classifier. One approach is to divide all data in a design set used for designing the classifier and a *test set* used for performance evaluation. This is the *hold-out* method that is known to have a rather big positive bias on the error estimation and does not use the available data very economically since a portion of it is not used for designing the classifier. As an alternative, we have employed the *leave-one-out method*, which sequentially picks each available data sample and classifies it (by the *knn* rule) using the remaining samples. This ensures that all data (but one sample) is used for designing and that an independent test set is kept.

Suppose that there are c possible classes with N_i (labeled) feature vectors from the ith class. Each available sample is thus employed once as a test sample. Classifier performance is measured using the *error counting* approach. The *class error rate* $\hat{\varepsilon}_i$ is defined as

$$\hat{\varepsilon}_i = \frac{\text{nr. of falsely classified test samples from class } i}{N_i} \tag{20}$$

and the *mean error rate* $\hat{\varepsilon}$ as

$$\hat{\varepsilon} = \frac{1}{c}\sum_{i=1}^{c}\hat{\varepsilon}_i = \frac{\text{total nr. of falsely classified test samples}}{N}, \tag{21}$$

where $N = \sum_{i=1}^{c} N_i$ is the total number of samples. This number estimates the percentage of test samples classified correctly and is used as a measure for the performance of the classifier (whenever we mention "classification performance" we refer to the mean error rate).

However, since $\hat{\varepsilon}$ is estimated using a random sample of feature vectors it is actually a random number which is a statistical estimate of the *true mean error* ε. It is therefore meaningful to give a confidence interval. Defining a *variance* σ, a 95% confidence interval of the mean error is given by [27]:

$$\hat{\varepsilon} - \sigma < \varepsilon < \hat{\varepsilon} + \sigma \quad \text{with} \quad \sigma = 1.96\sqrt{\frac{\hat{\varepsilon}(1-\hat{\varepsilon})}{N}}. \tag{22}$$

The classification performance of a classifier will be expressed by $\hat{\varepsilon} \pm \sigma$.

Well known in pattern recognition literature is the curse of dimensionality phenomenon, which dictates that classification performance not necessarily increases with an increasing number of features (given a fixed amount of data samples). Therefore, given a feature extraction scheme and a finite number of design samples, there exists an optimal number of features for a particular task. This becomes inherently important when dealing with colored images, since the number of extracted features is much larger compared to the gray-level case. Therefore, it is crucial to adopt a feature selection (or extraction) scheme to find a (sub-) optimal set of features. In this work the floating forward feature selection scheme (FFFS) [28] is adopted, which has recently been found to outperform other selection schemes [29]. This algorithm is initialised by taking the best feature ("best" is defined here as giving the lowest mean error). The selection then continues by iteratively adding (or deleting) a feature in each step to obtain a subset of all available features which gives the lowest mean error.

5. Experiments and discussion

Thirty real-world (512×512) RGB color images from different natural scenes [30], presented in Fig. 1, were selected. A database of 1920 color image regions of 30 texture classes was constructed by subdividing each image into 64 non-overlapping 64×64 subimages. Each image region was decomposed into a wavelet basis of depth 4 using a biorthogonal spline wavelet of order 2 [31]. The following feature sets were conducted:

(1) Intensity (gray-level) images were generated by computing the luminance, thereby discarding color information. A wavelet transform of depth 4 was performed and energy signatures were computed for each of the 12 detailed images. (total: 12 features)
(2) Each R, G and B component was wavelet transformed (depth 4) and energy signatures were computed from each detailed image. (total: 36 features)
(3) Each R, G and B component was wavelet transformed (depth 4) and correlation signatures were computed from each detailed image. (total: 72 features)
(4) 72 correlation signatures were computed using (18) for the three color spaces mentioned in Section 3:
    (a) UVW space,
    (b) YIQ space,
    (c) K-L space.

Each feature set was used to design a *knn*-classifier and to estimate its performance using the methodology outlined in Section 4.

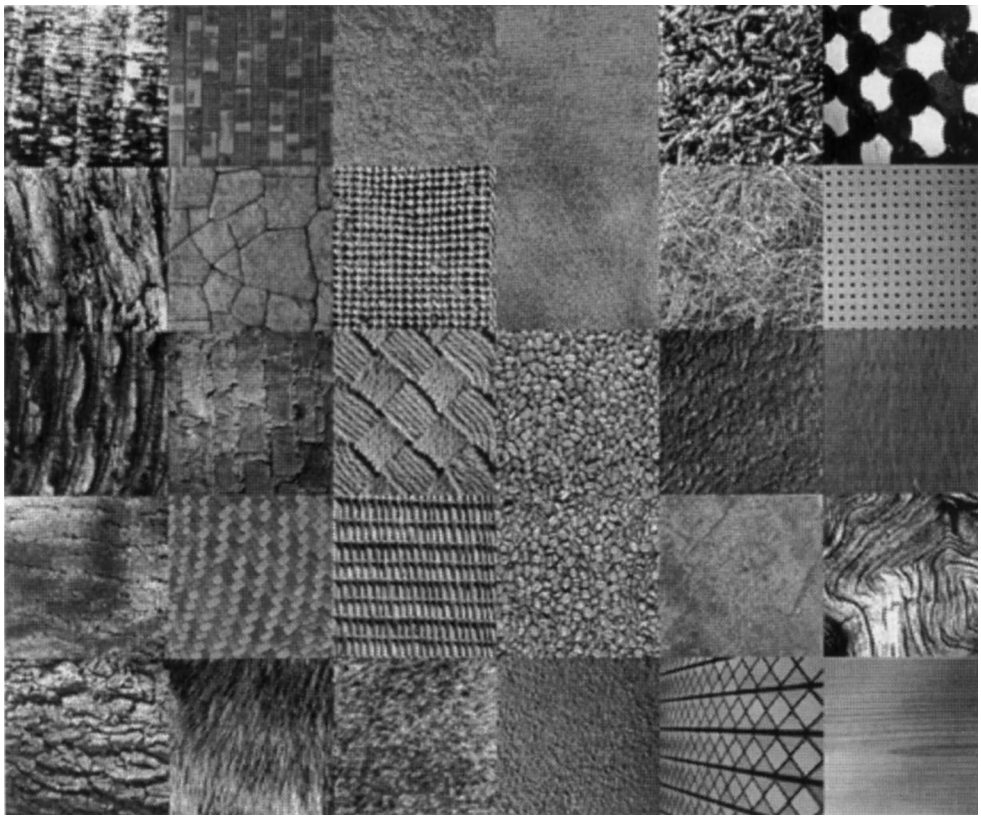

Fig. 1. Selection of images from the VisTex-database: from top to bottom and left to right: Bark0, Bark4, Bark6, Bark8, Bark9, Brick1, Brick4, Brick5, Fabric0, Fabric4, Fabric7, Fabric9, Fabric11, Fabric13, Fabric16, Fabric17, Fabric18, Food0, Food2, Food5, Food8, Grass1, Sand0, Stone4, Tile1, Tile3, Tile7, Water6, Wood1, Wood2.

Recall that the feature selection algorithm finds the (sub-) optimal feature subset for each dimensionality. This is depicted in Fig. 2 which displays the mean-error rate for each classifier in the function of the dimensionality of the used feature set. This representation allows a quick comparison between classifiers. Fig. 2a shows that features from color images significantly improve classification performance with respect to the gray-level features. It also shows that correlation signatures perform better than (color) energy signatures. Fig. 2b compares the performance of the correlation signatures in different color spaces, which shows that features in the K–L space perform best.

One observes from Fig. 2 that the mean errors for all classifiers saturate about a dimensionality of 10, at which point the class-error rates for each texture class were investigated. For Fabric0-7-17-18, Food5, Sand0, Tile7 and Wood2, the class errors were 0% for all classifiers and for Fabric9-16 and Water6, they were 0–3%. In Table 1, the remaining classes, for which class error rates differed more than 3% between classifiers are reported. Our further discussion is based on these error rates.

On the intensity images a mean error of 16% is achieved. The error rates show that intensity alone contains sufficient information to characterize some textures (e.g. Fabric0-7, Sand0, ...), but fails to do so on others (e.g. Bark8-9, Food0-2, Wood1). This can be expected since for instance Food0 and Food2 are recordings of two different species of beans (lima and coffee) which have very similar appearance.

Experiment 2 shows that adding color information does indeed significantly increase classification performance (a decrease of 7% in mean-error rate). The classifier confuses on the same texture classes as in experiment 1, but in a lesser degree. Especially, the discrimination between Food0-2 and Bark9 has improved substantially.

Comparing columns 2 and 3 from Table 1 shows that the correlation signatures offer a clear advantage over the energies. The error rate drops for most classes (excluding Stone4 and Tile3) resulting in a decrease of about 5% in mean-error rate. Especially, the confusion between Food0-2 has decreased substantially.

Analyzing the results of experiment 4 (Table 1, cols. 4a–4c), it is apparent that the error for the UVW space

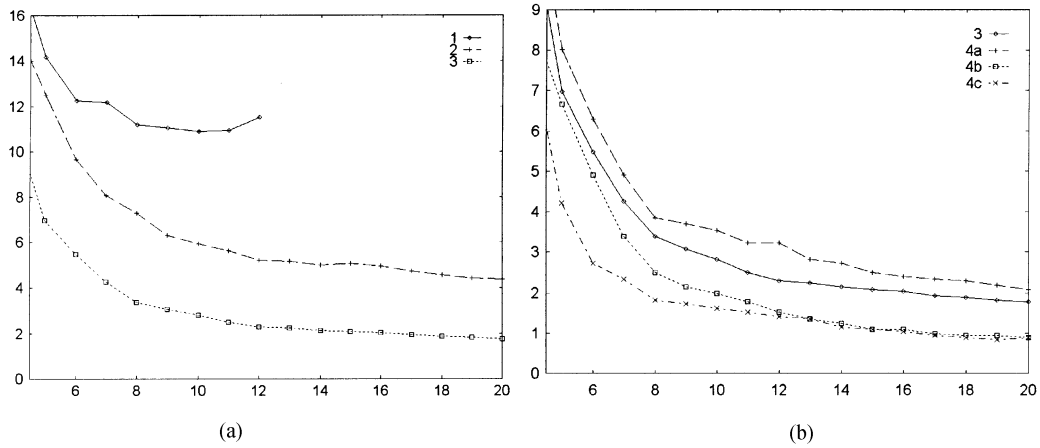

Fig. 2. Mean-error rate (in perc.) versus feature set dimensionality graphs for different classifiers. (a) for different signatures: intensity (1), energy RGB (2), correlation RGB(3). (b) correlation signatures in different color spaces: RGB (3), UVW (4a), YIQ (4b), K–L (4c). The variances σ on the mean errors vary in each point but are bounded by $0.4\% \leqslant \sigma \leqslant 1.6\%$.

Table 1
Class-error rates (in perc.) of the remaining texture set (the ones with different error rates per classifier) using the selected best 10 features. (1) gray-level energies, (2) energy in RGB space, (3) correlation in RGB space, (4a) Correlation in UVW space, (4b) correlation in YIQ space, (4c) correlation in K–L space

	Energy			Correlation		
	INTEN[1]	RGB[2,3]		UVW[4a]	YIQ[4,b]	K–L[4,c]
Bark0	21.9	20.3	17.2	15.6	0.0	0.0
Bark4	15.6	4.7	0.0	3.1	3.1	3.1
Bark6	20.3	7.8	1.6	6.2	0.0	0.0
Bark8	20.3	12.5	0.0	9.4	9.4	4.7
Bark9	40.6	7.8	6.2	12.5	7.8	6.2
Brick1	1.6	4.7	0.0	1.6	0.0	0.0
Brick4	14.1	9.4	1.6	1.6	6.2	3.1
Brick5	10.9	4.7	1.6	3.1	1.6	3.1
Fabric4	4.7	4.7	1.6	0.0	1.6	0.0
Fabric11	3.1	1.6	1.6	4.7	0.0	1.6
Fabric13	6.2	1.6	0.0	0.0	0.0	0.0
Food0	29.7	7.8	0.0	0.0	1.6	0.0
Food2	42.2	26.6	0.0	3.1	1.6	0.0
Food8	9.4	3.1	0.0	1.6	0.0	0.0
Grass1	4.7	6.2	6.2	4.7	4.7	6.2
Stone4	3.1	1.6	3.1	3.1	1.6	3.1
Tile1	6.2	6.2	1.6	4.7	1.6	6.2
Tile3	26.6	26.6	35.9	26.6	17.2	9.4
Wood1	25.0	15.6	1.6	1.6	1.6	1.6
Mean error $\pm$ variance	16.1 ± 2.1	9.1 ± 1.6	4.2 ± 1.1	5.4 ± 1.3	3.1 ± 1.0	2.5 ± 0.9

(6%) is drastically higher than for the K–L and YIQ spaces (3%) and also higher than for the RGB space. We believe that this is due to the fact that the UVW transform maps the RGB-components to a non-orthogonal coordinate system. The YIQ transform is nearly orthogonal and the K–L space is exactly orthogonal.

The best results are obtained with the K–L transform where a mean error of less than 3% is achieved.

Note that the first component of the K–L transform is the intensity, which indeed proved to be an important feature (cf. experiment (1)). The other two axes then represent the image information which is statistically uncorrelated with intensity. These features can be interpreted as the extra information which is not present in texture or color separately, and which can be denoted as color texture information.

The conducted experiments demonstrate that color texture can adequately be described by the wavelet correlation signatures. These features are not only suited for image classification, but can easily be employed for other color texture analysis tasks. For instance, for segmentation the wavelet signatures are computed over a (small) local window centered on each pixel of the image, resulting in one feature vector per pixel. Each pixel is then assigned to a particular image region, e.g. by (unsupervised) clustering techniques in the space of feature vectors.

6. Conclusion

We have discussed the extraction of features which combine image color and texture information to adequately describe the concept of color texture.

We introduced the wavelet correlation signatures, which are extracted from a multiscale representation of the color image. We discussed the importance of the color space in which feature extraction is performed. We also proved that for a linear color space transform, the correlation signatures can be obtained by means of a linear transform, a property which does not hold for the energy signatures.

Experimental results using a set of real-world colored textures, demonstrated the usefulness of wavelet signatures in color texture image analysis. In classification experiments, it was shown that color texture features increase performance with respect to the gray-level case. Best classification performances were obtained in the Karhunen–Loève space.

References

[1] G.J. Klinker, S.A. Shafer, T. Kanade, A physical approach to color image understanding, Int. J. Comput. Vision 4 (1990) 7–38.

[2] G. Wyszecki, W.S. Stiles, Color Science, Concepts and Methods, Quantitative Data and Formulas, 2nd eds. J Wiley, New York, 1982.

[3] Q.-T. Luong, Color in computer vision, C.H. Chen, L.F. Pau, P.S.P. Wang, eds. Handbook of Pattern Recognition & Computer Vision, Ch. 2.3, pp. 311–368. World Scientific, Singapore, 1993.

[4] Y.I. Ohta, T. Kanade, T. Sakai. Color information for region segmentation, Comput. graphics Image Process. 13 (1980) 222–241.

[5] W.K. Pratt, Spatial transform coding of color images, IEEE Trans. Comm. Technol 19(6) (1971) 980–992.

[6] C.L. Novak, S.A. Shafer, Methods for estimating scene parameters from color histograms, J. Opt. Soc. Am. A. 11 (1994) 3020–3036.

[7] R.M. Haralick, K. Shanmugan, I. Dinstein, Texture for image classification, IEEE Trans. Systems Man Cybernet 3(3), (1973) 610–621.

[8] T.R. Reed, J.M.H. du Buf, A review of recent texture segmentation and feature extraction techniques, CVGIP: Image Understanding 57(3) (1993) 359–372.

[9] D. Lee, R. Barber, W. Niblack, M. Flickner, J. Hafner, D. Petkovic, Indexing for complex queries on a query-by-content image database, Proc. of the 12nd IAPR Int. Conf. on Pattern Recognition. Vol. I, pp. 142–146, Jerusalem, Israel, 1994.

[10] J.R. Smith, S. Chang, Local color and texture extraction and spatial query, IEEE Proc. Int. Conf. on Im. Proc. Vol. III, 1996, pp. 1011–1014.

[11] T. Caelli, D. Reye, On the classification of image regions by colour texture and shape, Pattern Recognition 26(4) (1993) 461–470.

[12] S.C. Tan, J. Kittler, On color texture representation and classification, Proc. of the 2nd Int. Conf. on Image Processing, 1992, pp. 390–395.

[13] M. Hauta-Kasari, J. Parkkinen, T. Jaaskelainen, R. Lenz, Generalized co-occurrence matrix for multispectral texture analysis, Proc. of the 13th Int. Conf. on Pattern Recognition, 1996, pp. 785–789.

[14] D.K. Panjwani, G. Healey, Markov random field models for unsupervised segmentation of textured color images, IEEE Trans. Pattern Anal. Machine Intell 17(10) (1995) 939–954.

[15] R. Kondepudy, G. Healey, Use of invariants for recognition of three dimensional textures, J. Opt. Soc. Am. 11(11) (1994) 3037–3049.

[16] G. Healy, Lizhi Wang, Illumination-invariant recognition of texture in color images, J. Opt. Soc. Am. A 12(9) (1995) 1877–1883.

[17] J. Kittler, R. Marik, M. Mirmedhi, M. Petrou, J. Song, Detection of defects in colour texture surfaces, Proc. of the IAPR Workshop on Machine Vision Applic., pp. 558–567. Tokyo, Japan, 1994.

[18] T.S. Lee, Image repregentation using 2d gabor wavelets, IEEE Trans. Pattern Anal. Mach. Intell. 18(10) (1996) 959–971.

[19] I. Daubechies, Ten Lectures on Wavelets. Capital City Press, Montpellier, Vermont 1992.

[20] S. Mallat, A theory for multiresolution signal decomposition: the wavelet representation, IEEE Trans. Patt. Anal. Mach. Intell. 11(7) (1989) 674–693.

[21] T. Chang, C.-C.J. Kuo, Texture analysis and classification with tree-structured wavelet transform, IEEE Trans. Image. Process. 2(4) (1993) 429–441.

[22] A. Laine, J. Fan, Texture classification by wavelet packet signatures, IEEE Trans. Patt. Anal. Mach. Intell. 15(11) (1993) 1186–1190.

[23] M. Unser, Texture classification and segmentation using wavelt frames. IEEE Trans. Image Process. 4(11) (1995) 1549–1560.

[24] O. Pichler, A. Teuner, B.J. Hosticka, A comparison of texture feature extraction using adaptive gabor filtering, pyramidal and tree structured wavelet transforms, Pattern Recognition 29(5) (1996) 733–742.
[25] H. Levkowitz, G.T. Herman, Glhs: a generalized lightness, hue and saturation model, CVGIP 11(11) (1994) 3011–3019.
[26] R.O. Duda, P.E. Hart, Pattern Classification and Scene Analysis. Wiley, New York, 1973.
[27] P.A. Devijver and J. Kittler, Pattern Recognition: A Statistical Approach. Prentice-Hall, Englewood Cliffs, NJ, 1982.
[28] P. Pudil, J. Novovicova, J. Kittler, Floating search methods in feature selection. Pattern Recognition Lett 15 (1994) 1119–1125.
[29] A. Jain, D. Zongker, Feature selection — evaluation, application, and small sample performance, IEEE Trans. Pattern Anal. Mach. Intell. 19(2) (1997) 153–158.
[30] VisTex, Color image database, http://www-white.media.mit.edu/vismod/imagery/Vision Texture. MIT Media Lab, 1995.
[31] M. Unser, A. Aldroubi, M. Eden, A family of polynomial spline wavelet transforms, Signal Process. 30 (1993) 141–162.

About the Author—GERT VAN DE WOUWER obtained a degree in theoretical physics from the University of Ghent, Belgium, in July 1994. Since then, he is a research assistant at the Vision Lab, department of Physics at the University of Antwerp, Belgium. Among his research interests are image processing and pattern recognition. In particular, he has been involved in several subjects in the field of wavelet texture analysis and is finishing to write his Ph.D. thesis on this subject.

About the Author—PAUL SCHEUNDERS graduated in physics at the University of Antwerp (UIA) in 1986, and received a Ph.D. in Physics at the same university in 1990, with work in the field of statistical mechanics. Since 1991 he is a research associate at the University of Antwerp (RUCA) at the Vision Lab of the department of Physics. He is currently working in the field of image processing and pattern recognition.

About the Author—STEFAN LIVENS received his B.Sc. in Physics from the University of Antwerp, Belgium in 1992. Since then, he works as a research assistant at the VisieLab research group of the same university. In the fall of 1997, he defends his Ph.D. thesis entitled "Image Analysis for Material Characterisation". He has a special interest in multiresolution texture analysis. Related to this topic, he moderates an electronic newsletter called "Wavelets for Texture Analysis". Currently he is involved in a project in collaboration with AGFA Gevaert, where he deals with colour and shape analysis of optical microscopy images.

About the Author—DIRK VAN DYCK received a Ph.D. in Physics from the University of Antwerp in 1977 and is currently a Professor of Theoretical Mechanics, Information Theory and Image Procession at the same university. His research interests are the theoretical aspects of dynamic electron diffraction and imaging, holographic reconstruction and structural retrieval, image processing and pattern recognition, and artificial intelligence.