

GENERALIZATION OF THE COOCCURRENCE MATRIX FOR COLOUR IMAGES: APPLICATION TO COLOUR TEXTURE CLASSIFICATION

VINCENT ARVIS¹, CHRISTOPHE DEBAIN¹, MICHEL BERDUCAT¹ AND ALBERT BENASSI²

¹Cemagref Clermont-Ferrand, Unité de recherche: «Technologies, Systèmes d'information et procédés pour l'agriculture et l'agroalimentaire - TSCF», 24 Avenue des Landais, BP 50085, 63172 Aubière cedex, France;

²Laboratoire de Météorologie Physique, Université Blaise Pascal, 24 Avenue des Landais, 63177 Aubière cedex, France

e-mail: vincent.arvis@laposte.net, christophe.debain@cemagref.fr, michel.berbucat@cemagref.fr, benassi@univ-bpclermont.fr

(Accepted February 2, 2004)

ABSTRACT

Three different approaches to colour texture analysis are tested on the classification of images from the VisTex and Outex databases. All the methods tested are based on extensions of the cooccurrence matrix method. The first method is a multispectral extension since cooccurrence matrices are computed both between and within the colour bands. The second uses joint colour-texture features: colour features are added to grey scale texture features in the entry of the classifier. The last uses grey scale texture features computed on a previously quantized colour image. Results show that the multispectral method gives the best percentages of good classification (VisTex: 97.9%, Outex: 94.9%). The joint colour-texture method is not far from it (VisTex: 96.8%, Outex: 91.0%), but the quantization method is not very good (VisTex:83.6%, Outex:68.4%). Each method is decomposed to try to understand each one deeper, and computation time is estimated to show that multispectral method is fast enough to be used in most real time applications.

Keywords: colour texture classification, computation time, multispectral, Outex, VisTex.

INTRODUCTION

Texture analysis is a branch of image processing that began to be studied thirty years ago. Although the concept of texture was difficult (and is always) to define, the studies showed that spatial statistics computed on the grey levels of the images were able to give good descriptors of the perceptual feeling of texture (For a review, see Haralick (1979) and Van Gool *et al.* (1983)). Such textural descriptors are still developed today to give more and more powerful tools for classification tasks or segmentation problems (Ojala *et al.*, 1996). But for a decade, the study of texture has been extended to the study of texture in colour images. Approaches used are based on existing grey level methods that are adapted to take into account the colour information. In the literature, three families of approaches can be found.

The first consists in the use of joint colour-texture features (Dubuisson-Jolly and Gupta, 2000; Drimbarean and Whelan, 2001; Mäenpää *et al.*, 2002). Textural features (grey scale) and colour features (moments or

histograms for example) are computed individually and then are used together as the entry of a classifier. The second consists in the reduction of the colour information in the colour image using a quantization method (Chang and Wang, 1996; Hauta-Kasari *et al.*, 1996; Chang and Krumm, 1999; Chen and Chen, 2002). The image obtained is coded like a grey scale image, each grey scale corresponding to a different colour. Classical texture features are then computed on this image. So both approaches use a transformation of the colour image to be able to apply existing greyscale methods, without giving a new definition to colour texture.

The last consists in taking into account the correlations between the colour bands while computing the texture features. Descriptors are computed both within and between channels to give information on the whole colour texture (Rosenfeld *et al.*, 1982; Van de Wouwer *et al.*, 1999; Paschos, 1998;2000;). Even if this approach seems to intimacy link colour and texture in the descriptors to obtain real colour texture features, it is rarely encountered, certainly because it seems computationally heavier than the other methods.

In this paper, we propose such a multispectral method considering the correlations between the colour bands. This is an extension of the method based on the cooccurrence matrices proposed by Haralick *et al.* (1973), which is an old method but always a reference.

To study the efficiency of the method, we test it in a classification problem on the image databases VisTex and Outex available on the internet and used by the computer vision community. Examples of images from these databases are given Fig. 1. The cooccurrence method is also extended according to the two other approaches described above (fusion of texture and colour descriptors and quantization of the colour image) to have a comparison between the three approaches to the texture in colour images.

MATERIAL AND METHODS

IMAGE DATA

Images used for the experiments belonged to two different data sets. The first included 54 images of the VisTex database (MIT Media Lab) where images

were taken under non-specified illumination conditions and imaging geometry, so they are representative of real world conditions. The second included 68 images of the Outex database (Ojala *et al.*, 2002), where images were taken with a fixed imaging geometry and with specified illumination source (a 2856 K incandescent CIE A light source). So differences between images are due only to a difference in medium.

The 54 original VisTex images of resolution 512×512 were split into 16 samples of 128×128. These images are available on the Outex site as test suite Contrib_TC_00006. For each texture, half of the samples were used in the training set and the other half served as testing data. The samples of the training set were the white squares of a draughtboard (beginning in the upper left corner) in order to consider a possible non uniformity of the original images (Fig. 2).

The 68 Outex images were treated similarly, but as their size was 746×538, we obtained 20 samples of 128x128. At the Outex site, this is the test suite Outex_TC_00013. The training and test sets were chosen in the same way as before, thus 680 samples in each set.



Fig. 1. Examples of images from the VisTex and from the Outex database. Vistex gives images taken in real world conditions whereas images from Outex are taken under controlled conditions.

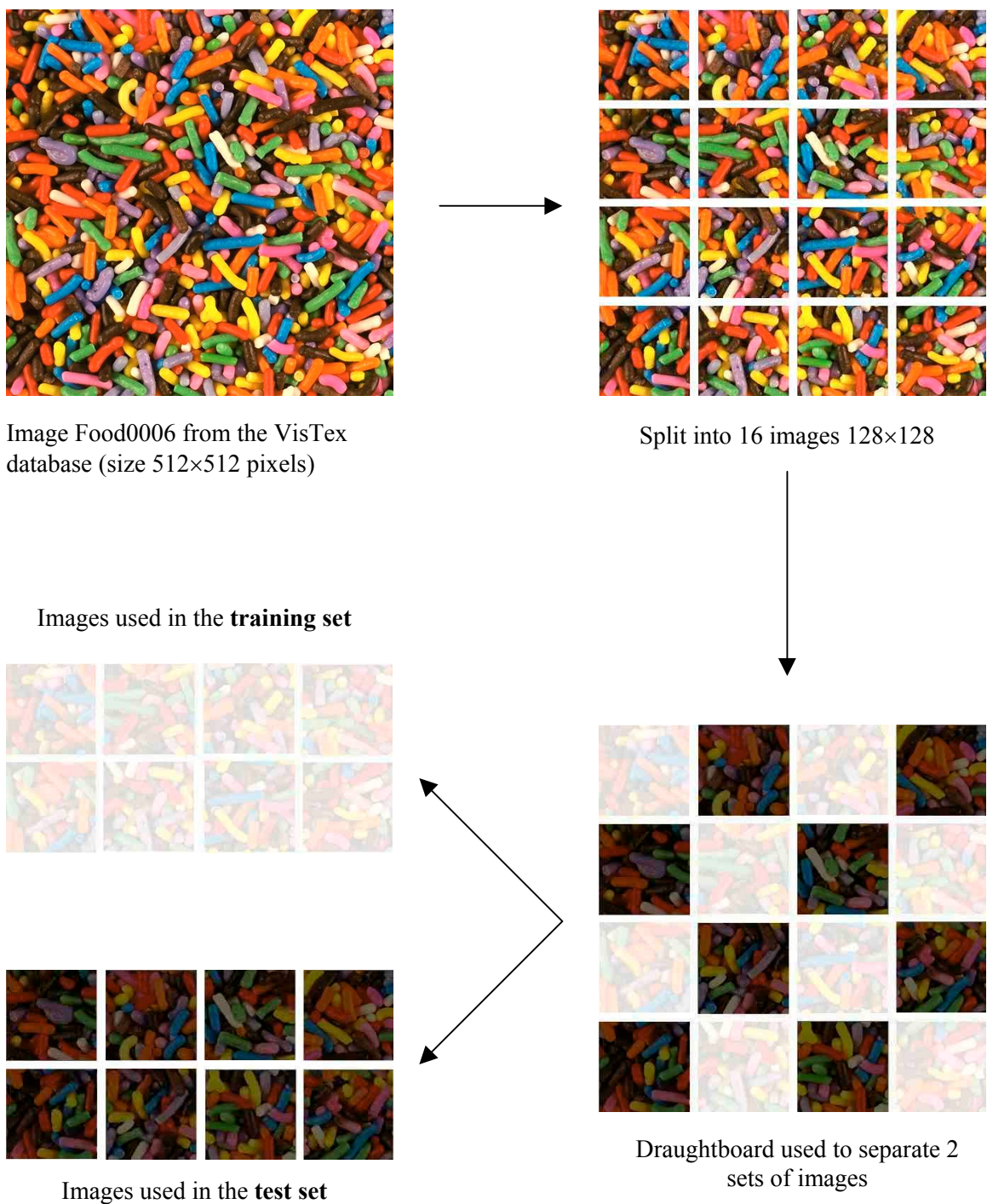


Fig. 2. Distribution of images between the training set and the test set for the class Food0006 of the VisTex database.

FEATURES

All the colour texture methods presented below are based on the cooccurrence matrix and Haralick features. This method is classical in the pattern recognition community and have extensively been used on grey scale images. But let us first briefly

recall the definitions of these. Let I be a greyscale image coded on m grey levels. Let $s=(x, y)$ be the position of a pixel in I and $\mathbf{t}=(\Delta x, \Delta y)$ be a translation vector. The cooccurrence matrix $M_{\mathbf{t}}$ is a $m \times m$ matrix whose (i,j) th element is the number of pairs of pixels separated by the translation vector \mathbf{t} that have the pair of grey levels (i,j) .

$$M_t(i,j) = \text{card} \{(s,s+\mathbf{t}) \in R^2 \mid I[s] = i, I[s+\mathbf{t}] = j\} \quad (1)$$

The choice of the relative position vector was the same as Haralick's. This is a distance of one pixel in eight directions to take into account the eight nearest-neighbours of each pixel. The eight matrices obtained were then summed to obtain a rotation-invariant matrix M . Let us quote that since $M_t(i,j) = M_{-\mathbf{t}}(j,i)$, M is symmetric.

Haralick assumed that the texture information is contained in this matrix, and texture features are then calculated from it. He extracted 14 parameters from the cooccurrence matrix, but only five are commonly used because it was shown that the 14 are very correlated with each other, and that the five sufficed to give good results in a classification task (Connors and Harlow, 1980). The features are homogeneity (E), contrast (C), correlation (Cor), entropy (H) and local homogeneity (LH).

Let us recall their definitions:

$$E = \sum_i \sum_j (M(i,j))^2, \quad (2)$$

$$C = \sum_{k=0}^{m-1} k^2 \sum_{|i-j|=k} M(i,j), \quad (3)$$

$$\text{Cor} = \frac{1}{\sigma_i \sigma_j} \sum_i \sum_j (i - \mu_i)(j - \mu_j) M(i,j) \quad (4)$$

where μ_i and σ_i are the horizontal mean and variance and μ_j and σ_j are the vertical statistics.

$$H = \sum_i \sum_j M(i,j) \log(M(i,j)), \quad (5)$$

$$\text{LH} = \sum_i \sum_j \frac{M(i,j)}{1 + (i-j)^2} \quad (6)$$

The five features were normalized by the number of bins in the cooccurrence matrix M in order to fit between 0 and 1.

Multispectral method

This method (Fig. 3) is an extension of the cooccurrence method to multispectral images, i.e. images coded on n channels. In this case, let $C_1, C_2, \dots, C_u, \dots, C_n$ be the n channels of the image, each coded on m levels. Let $\mathbf{t}=(\Delta x, \Delta y)$ be a translation vector and $(C_u \rightarrow C_v)$ be a couple of channels. $(C_u \rightarrow C_v)$ indicates that in the couples of pixels defined by \mathbf{t} , the first belongs to the channel C_u and

the second to C_v . The generalized cooccurrence matrices are:

$$M_{t,(C_u \rightarrow C_v)}(i,j) = \# \{(s,s+\mathbf{t}) \in R^2 \mid C_u[s] = i, C_v[s+\mathbf{t}] = j\} \quad (7)$$

with one matrix per couple of channels $(C_u \rightarrow C_v)$.

In our case, \mathbf{t} was a translation of 1 pixel in the eight directions and the matrices obtained were summed to obtain $M_{(C_u \rightarrow C_v)}$. Let's quote that $M_{t,(C_u \rightarrow C_v)}(i,j) = M_{-\mathbf{t},(C_v \rightarrow C_u)}(j,i)$, so $M_{(C_u \rightarrow C_v)}$ and $M_{(C_v \rightarrow C_u)}$ are containing the same information. Both were summed to obtain a symmetric matrix $M_{(C_u, C_v)}$, and five Haralick features were computed on this matrix. In our case, colour images are coded on three channels, leading to six different matrices: (R,R), (G,G), (B,B) that are the same as greyscale cooccurrence matrices computed on one channel and (R,G), (R,B), (G,B) that take into account the correlations between the channels. So this method led to a total of 30 texture features.

Fusion of colour and texture features

This method (Fig. 4) consists in a change in the colour space of the images, in order to obtain one channel containing the luminance information and two others containing chrominance information. Texture features are then computed from the luminance channel and other features named colour features are computed from the chrominance channel (Drimbarean and Whelan, 2001). On the intensity channel, texture features were extracted as described above. The cooccurrence matrix was computed on images with 256 grey tones and on images reduced to 32 grey tones by uniform quantization. The aim of this was to obtain an indication of the loss of texture information due to a reduction of grey tones resolution. On the chromaticity channels, colour features were extracted, consisting in the mean and standard deviation of each channel. Thus a total of 9 features characterized one image sample.

First, the HSV (hue, saturation, value) colour space was used. It corresponds better to how people experience colour than the RGB colour space does: hue (H) represents the wavelength of a colour if it would be monochromatic. Hue varies from 0 to 1 when colour goes from red to green then to blue and back to red. H is then defined modulo 1. As colour is seldom monochromatic, saturation (S) represents the amount of white colour mixed with the monochromatic colour. Value (V) does not depend on the colour, but represents the brightness. So H and S are chrominance and V is intensity.

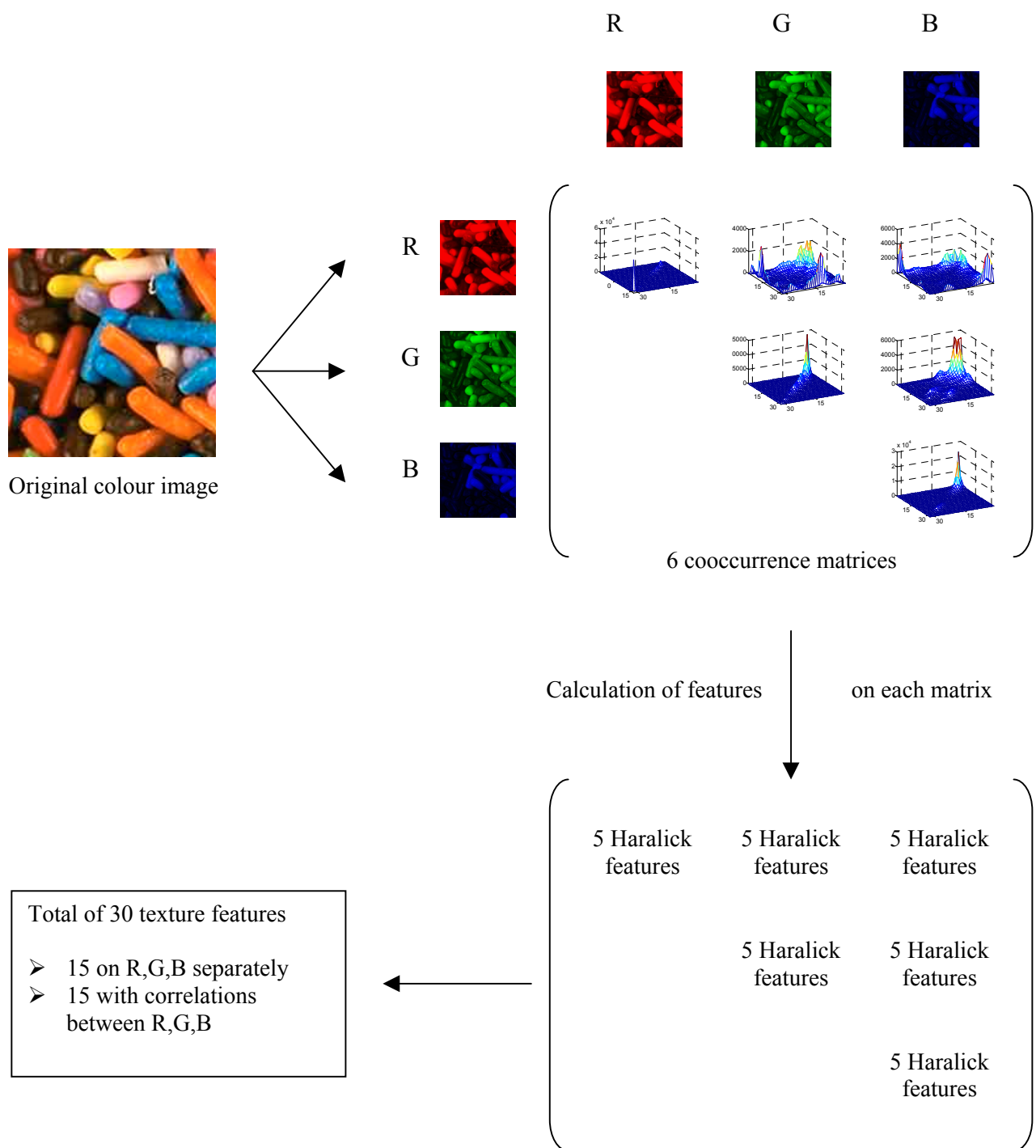


Fig. 3. Illustration of the **multispectral method** applied to an image from the class *Food0006* of the *VisTex* database.

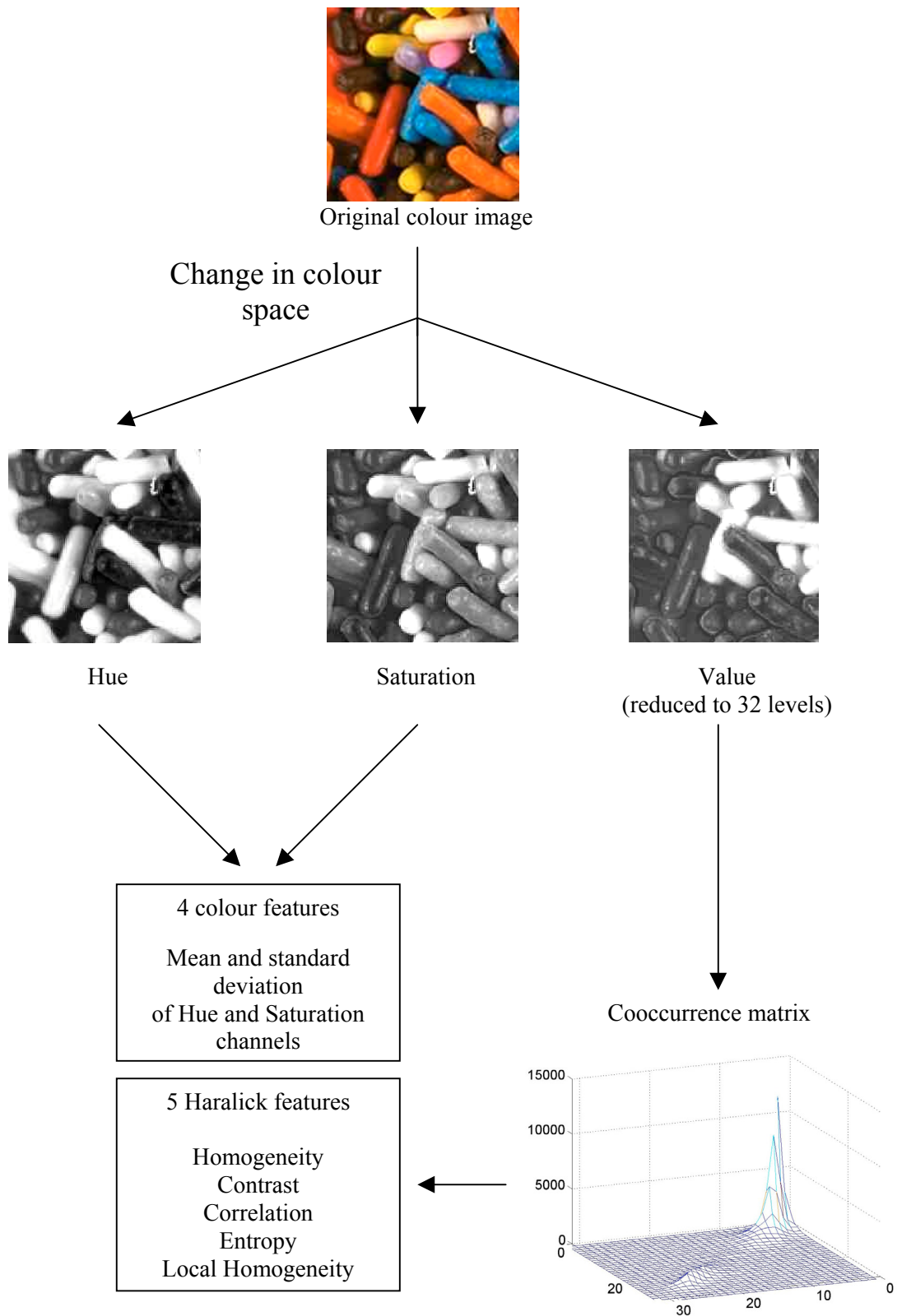


Fig. 4. Illustration of *the fusion of colour and texture features* applied on an image from the class *Food0006* of the *VisTex* database.

The following equations transform RGB in [0,1] to HSV in [0,1]:

$$V = \max(R, G, B), \quad (8)$$

$$S = \frac{V - \min(R, G, B)}{V}, \quad (9)$$

$$H = \frac{G - B}{6S} \quad \text{if } V=R, \quad (10)$$

$$H = \frac{1}{3} + \frac{B - R}{6S} \quad \text{if } V=G, \quad (11)$$

$$H = \frac{2}{3} + \frac{R - G}{6S} \quad \text{if } V=B. \quad (12)$$

A second colour space was tested, for which relations are linear. The space chosen was the YCbCr colour space, which is widely used for digital video. In this format, luminance information is stored as a single component (Y), and chrominance information is stored as two colour-difference components (Cb and Cr). Cb represents the difference between the blue component and a reference value. Cr represents the difference between the red component and a reference value. These features are defined for video processing purposes and so are not meaningful concerning human experience.

These equations transform RGB in [0,1] to YCbCr in [0, 255].

$$Y = 16 + 65.481 R + 128.553 G + 24.966 B \quad (13)$$

$$Cb = 128 - 37.797 R - 74.203 G + 112 B \quad (14)$$

$$Cr = 128 + 112 R - 93.786 G - 18.214 B \quad (15)$$

In addition, experiments with texture features only and colour features only were performed in the HSV colour space to see which of the texture or colour information was most prominent.

Quantization

This method (Fig. 5) is the generalization of the grey level cooccurrence matrix method proposed by Hauta-Kasari *et al.* (1996). Instead of computing the cooccurrence matrix using the value of the grey levels, colour images are quantized to extract several colour classes, and the cooccurrence matrix uses the label of the classes for its computing. Even if there is no natural order between the colour classes, the method of quantization must allow that the same color is labelled in the same way from an image to another. This means that even if some colour classes are not present in an image, the label always exists and the number of pixels with this label is zero. So the used quantization method consists in a separation

of the RGB cube into n^3 equal sub cubes, each one labelled with a number between 0 and n^3-1 .

Cooccurrence matrix was then computed on the image obtained and the five Haralick features were extracted. The effect of the number of classes was tested on three cases: $n = 4$, $n = 6$ and $n = 8$.

CLASSIFICATION

The features described above were used in classifying the two test sets. As it was seen before, each sample image was represented by a vector of p features, with a different p according to the colour texture approach used. One sample is then a point in a p -dimensional space. On the training set, a change of space is necessary to group samples from the same class while separating samples from different classes. This was performed by a Discriminant Factorial Analysis on the p -dimensional space of the training set. In the new space obtained, a small distance between two points means that the corresponding samples are likely to belong to the same class. Then each sample of the test set was projected in the new space obtained, and one sample is affected to the class the more represented in its 5 nearest neighbours. This method is known under the name of k -Nearest Neighbours. The choice of a simple classification method was done to make the results more representative of the effectiveness of the texture features than of the classifier itself.

RESULTS AND DISCUSSION

The analysis of the results (Table 1) shows that the best percentages of good classification are achieved with the multispectral method, both on VisTex and Outex with respectively 97.9% and 94.9%. Behind, the fusion of texture and colour features reaches 96.8% and 91.0% and last the quantization method 83.6% and 68.4%.

So the proposed method seems to be a very good approach for colour texture classification problems, in comparison with the other approach tested, but also with results found in the literature on these databases (Mäenpää *et al.*, 2002). Furthermore, experiments using only the features computed within colour bands (without correlations) and between colour bands (correlations only) shows that both are complementary. Actually, results without correlations and with only correlations reach for example 90.9% and 88.4%, whereas used together, they achieve 94.9%. This result shows that correlations between colour bands introduce new and relevant information and so is worth being used as a colour texture descriptor in addition with features computed within colour bands.

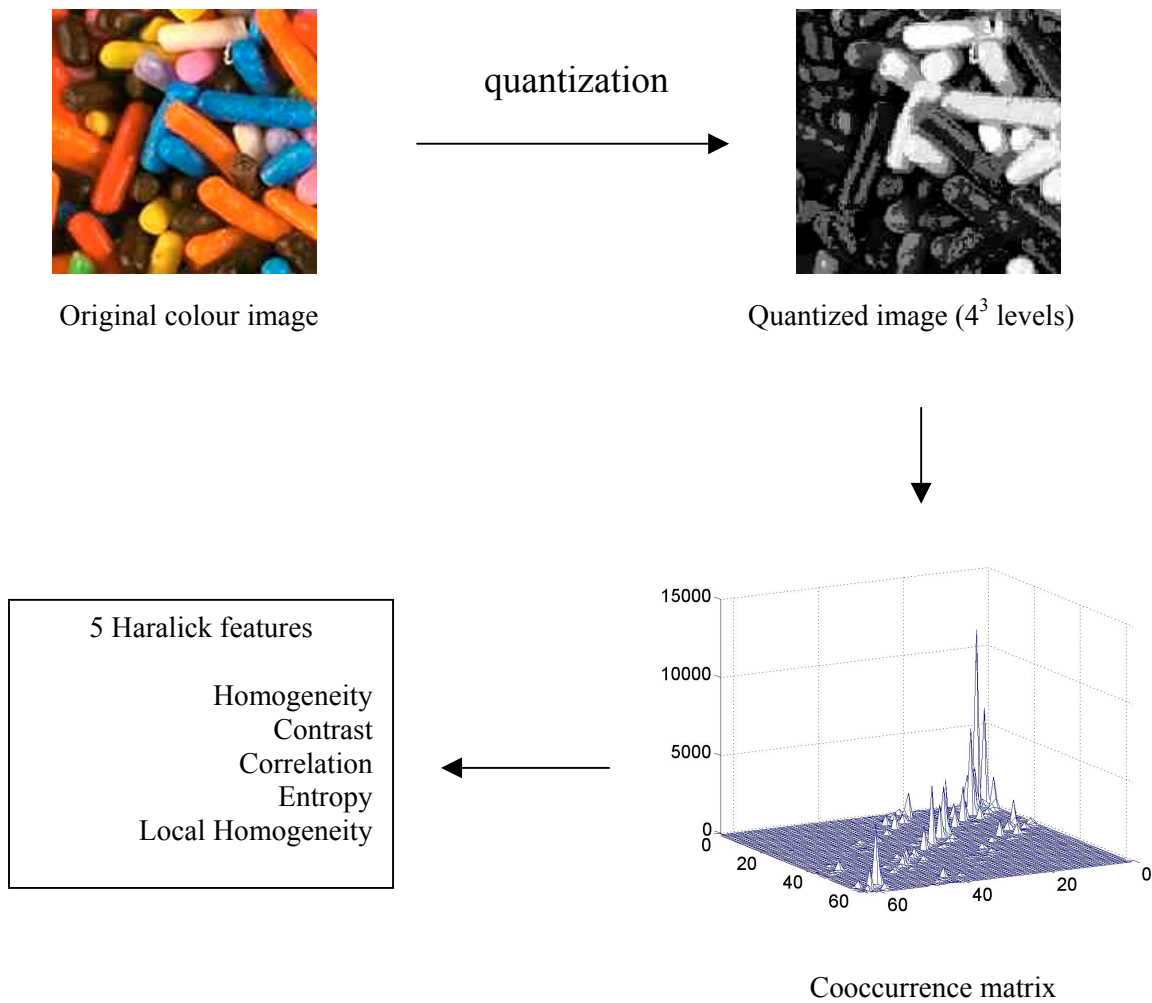


Fig. 5. Illustration of *the quantization method* applied on an image from the class *Food0006* of the *VisTex* database.

Table 1. *Classification results: percentages of good classification.*

	Vistex			Outex		
	32 grey levels	256 grey levels		32 grey levels	256 grey levels	
<i>Multispectral</i>						
All features	97.7%	97.9%		94.9%	94.1%	
No correlation	95.6%	97.5%		90.9%	90.7%	
Correlations only	93.1%	92.6%		88.4%	86.6%	
<i>Fusion Texture+Colour</i>						
HSV colour space	96.1%	94.9%		91.0%	88.4%	
YCbCr colour space	96.3%	96.8%		86.9%	88.2%	
Grey scale Texture	76.4%	75.0%		63.8%	66.0%	
Colour		77.1%			73.1%	
<i>Uniform quantization</i>						
	4 ³ levels	6 ³ levels	8 ³ levels	4 ³ levels	6 ³ levels	8 ³ levels
	83.1%	82.9%	83.6%	54.6%	68.4%	65.6%

The use of joint colour-texture features is a relevant method too, even if the multispectral method is better. Here again, the strength of the method comes from complementarity: classification with colour features only gave 77.1% and texture only 75.0%. When both are used together, the results extensively increase, since they achieve 96.8%. So colour and texture information are complementary and, used together, they are able to give good results of classification, even if the colour features are very simple (only mean and standard deviation of the chrominances). Concerning the choice of a colour space, both give good results: HSV is better for Outex and YC_bC_r is better for VisTex with no obvious reasons, so we cannot conclude about the use of one more than the other. Even so, we could have expected the results to be worse with HSV. Actually, the equations used for its computation are strongly non-linear and introduce discontinuity at 0, 1/3 and 2/3 in the computing of H. Moreover, H is defined modulo 1, so this feature should be treated in a different way than the others, what is not the case here (Hanbury and Serra, 2001; Hanbury, 2002). Even so, the fact that it has an intuitive significance seems to balance the results comparing with YC_bC_r for which relations are linear, but where features are not meaningful concerning human experience.

The quantization method gives bad results in comparison with the other methods. Even so, this method is commonly used and gives very good results in segmentation problems. In our classification test, it is not the case because of the quantization method used. In a segmentation problem, quantization methods are more precise and efficient to assign colour classes because there is only one image and the comparison between textures is done in this image. So all existing colours do not have to be quantized but only colours the contained in the image. This can explain the bad results achieved for the classification with this method, but we can quote that the results are nevertheless better than when using grey scale texture only (VisTex: 76.4%, Outex: 66.0%) so it is an improvement of the grey scale method and a few colour information is taken into account with these features.

Tests of the computation time were performed for each of the three approaches. For this we used a PC with a Intel® Pentium® 4 processor 1.50 GHz and 512 MB Ram, image processing was programmed with Matlab 6, using a C program for the calculation of the cooccurrence matrices. Times are given for the computing of all the features of each method, on images 128×128 previously loaded in the memory. Results are presented on Table 2.

Table 2. *Computation time.*

	Computation time (ms)		
	32 grey levels	256 grey levels	
Multispectral	107	1845	
Fusion Texture+Colour	95	354	
Uniform quantization	4 ³ levels	6 ³ levels	8 ³ levels
	43	193	1537

The fastest method is quantization with 4³ levels, but classification results of this method are bad. After follows the use of joint colour-texture features and the multispectral method with 32 levels. Both are close to 100 ms, which is too great for real time video applications (40 ms), but allows real time applications in many fields in industry and in agriculture. Actually, such applications often use other technologies that are not faster. For example, high-performance GPS work at a frequency of 10 Hz, so the same as our computation. The other computation times are greater, and increase as far as 1845 ms for the multispectral method with 256 levels. Such times are too long for a real time application, but by comparing classification results obtained with 32 and 256 levels, we can see that 32 levels give almost as good or even better results than with 256 levels. So the use of a small level such as 32 is a quite good compromise between classification results and computation time, especially for our multispectral method that allows real time applications.

In conclusion, the multispectral method proposed gives very good results for a classification problem. The use of the cooccurrence matrix method computed both within and between the colour bands allows to achieve such results, so this way to study colour texture seems to be a promising one. Moreover, the computation time for this method is not prohibitive and allows real time applications.

Furthermore, this multispectral method tested here on colour images can easily be extended to other data coming for example from other sensors supplying more than three channels. So the method could find successful applications in domains other than purely image processing, such as materials science or any domain interested in the study of complex media.

REFERENCES

- Chang C-C, Wang L-L (1996). Color texture segmentation for clothing in a computer-aided fashion design system. *Image and Vision Computing* 14:685-702.

- Chang P, Krumm J (1999). Object Recognition with Color Cooccurrence Histograms. IEEE Conference on Computer Vision and Pattern Recognition, Fort Collins, CO, June 23-25, 1999.
- Chen K-M, Chen S-Y (2002). Color texture segmentation using feature distributions. *Patt Recogn Letters* 23(7): 755-71.
- Connors RW, Harlow CA (1980). A theoretical comparison of texture algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2(3): 204-22.
- Drimbarean A, Whelan PF (2001). Experiments in colour texture analysis. *Patt Recogn Letters* 22:1161-7.
- Dubuisson-Jolly M-P, Gupta A (2000). Color and texture fusion: application to aerial image segmentation and GIS updating. *Image and Vision Computing* 18:823-32.
- Hanbury A (2002). The taming of the hue, saturation and brightness colour space. *Proceedings of the Seventh Computer Vision Winter Workshop, Bad Aussee, Austria, February 2002.*
- Hanbury A, Serra J (2001). Mathematical morphology in the HLS colour space. *Proceedings of the British Machine Vision Conference 2001*, 451-60.
- Haralick RM (1979). Statistical and structural Approaches to Texture. *Proceedings of the IEEE* 67:786-804.
- Haralick RM, Shanmugam K, Dinstein I (1973). Textural Features for Image Classification. *IEEE Transactions on Systems, Man and Cybernetics* 3(6):610-21.
- Hauta-Kasari M, Parkkinen J, Jaaskelainen T, Lenz R (1996). Generalized cooccurrence matrix for multispectral texture analysis. *Proceedings of the 13th International Conference on Pattern Recognition, Vienna, Austria, August 25-30, 1996.*
- Mäenpää T, Pietikäinen M, Viertola J (2002). Separating color and pattern information for color texture discrimination. *16th International Conference on Pattern Recognition, August 11-15, Quebec City, Canada, 1: 668-71.*
- MIT Media Lab Vision Texture-VisTex Database. <http://www-white.media.mit.edu/vismod/imagery/VisionTexture/vistex.html>
- Ojala T, Pietikäinen M, Harwood D (1996). A comparative study of texture measures with classification based on feature distributions. *Patt Recogn* 29:51-9.
- Ojala T, Pietikäinen M, Mäenpää T, Viertola J, Killönen J, Huovinen S (2002). Outex - new framework for empirical evaluation of texture analysis algorithms. *16th International Conference on Pattern Recognition, Quebec, Canada, 1:701-6.* <http://www.outex.oulu.fi>
- Paschos G (1998). Chromatic correlation features for texture recognition. *Patt Recogn Letters* 19:643-50.
- Paschos G (2000). Fast color texture recognition using chromaticity moments. *Patt Recogn Letters* 21:837-41.
- Rosenfeld A, Wang C-Y, Wu AY (1982). Multispectral Texture. *IEEE Transactions on Systems, Man and Cybernetics* 12(1):79-84.
- Van de Wouwer G, Scheunders P, Livens S, Dyck DV (1999). Wavelet correlation signatures for color texture characterization. *Patt Recogn* 32:443-51.
- Van Gool L, Dewaele P, Osterlinck A (1983). Texture analysis anno 1983. *Computer Vision Graphics Image Processing* 29(12):336-57.