Texture classification via conditional histograms

Eugenia Montiel [a,*], Alberto S. Aguado [a], Mark S. Nixon [b]

[a] *Electronic and Electrical Engineering, University of Surrey, Guildford, Surrey GU2 5XH, United Kingdom*
[b] *Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, United Kingdom*

Abstract

This paper presents a non-parametric discrimination strategy based on texture features characterised by one-dimensional conditional histograms. Our characterisation extends previous co-occurrence matrix encoding schemes by considering a mixture of colour and contextual information obtained from binary images. We compute joint distributions that define regions that represent pixels with similar intensity or colour properties. The main motivation is to obtain a compact characterisation suitable for applications requiring on-line training. Experimental results show that our approach can provide accurate discrimination. We use the classification to implement a segmentation application based on a hierarchical subdivision. The segmentation handles mixture problems at the boundary of regions by considering windows of different sizes. Examples show that the segmentation can accurately delineate image regions.
© 2005 Elsevier B.V. All rights reserved.

Keywords: Texture classification; Co-occurrence features; Image analysis; Image segmentation; Region detection; Non-linear image analysis

1. Introduction

Previous works have shown that histograms can be used as powerful descriptions for non-parametric classification (Unser, 1986a; Valkealahti and Oja, 1998; Ojala et al., 1996, 2000; Hofmann et al., 1998; Puzicha et al., 1999). In contrast to parametric features (Haralick, 1979), histograms contain all the information of distributions avoiding the problem of feature selection. In general, the development of a method for automatic feature selection is not trivial since optimal performance requires a careful selection of features according to particular types of textures (Ohanian and Dubes, 1992; Jain and Zongker, 1997; Sullis,

---
* Corresponding author. Tel.: +44 1483 68 6044; fax: +44 1483 68 6031.
E-mail addresses: e.montiel@eim.surrey.ac.uk (E. Montiel), msn@ecs.soton.ac.uk (M.S. Nixon).

1990; Ng et al., 1998). However, histograms cannot be directly used as texture descriptors; the computation and dimensionality impose prohibitive computational resources for applications (Rosenfeld et al., 1982; Augusteijn et al., 1995). The computation of histograms of low-dimensionality is considered as an open problem (Unser, 1986a; Valkealahti and Oja, 1998; Rosenfeld et al., 1982; Ojala et al., 2001).

In addition to make texture descriptors useful for applications, the reduction of histogram's dimensionality has two important implications. First, it avoids sparse histograms due to insufficient training data. Secondly, if histograms are compact, it is possible to consider more complex pixels interdependencies increasing the discrimination power. It is important to notice that if the histogram reduction is effective, we should expect good discrimination with similar features than the ones used to codify the high dimension description. As such, it is important to distinguish between the effectiveness of the reduced process and the additional power obtained by including more complex pixel interdependencies.

A powerful approach to histogram reduction is to perform a quantisation to adapt the histogram bins according to the distribution (Puzicha et al., 1999). In (Valkealahti and Oja, 1998; Ojala et al., 2001) the adaptation is defined by using techniques of vector quantisation (Gersho and Gray, 1992). Results on texture discrimination have shown that this is a very powerful technique to reduce multiplex (i.e., >2) co-occurrences. However, although a tree structure can be used to handle the complexity required by the encoder, yet the encoder requires a significant number of operations and sample data. The search uses more memory that a full search vector quantisation and the process can lead to sub-optimal solutions (Gersho and Gray, 1992).

In this paper, we simplify histograms by considering combinations of the random variables defining the joint probability function (i.e., grey tone or colour dependence) (Unser, 1986a; Rosenfeld et al., 1982; Ojala et al., 2001). In (Rosenfeld et al., 1982), and later (Unser, 1986a), histograms define the probability of the differences of grey levels for pixel pairs. The motivation is that these

operations define the principal axes of the second order joint probability function (Unser, 1986a). However, they have had an unpredictable success in applications (Schulerud and Carstensen, 1995; Chetverikov, 1994). The main caveat of this representation is that, in general, random variables in a texture do not define Gaussian independent distributions. Additionally, although the average error of the difference is well approximated by the factorised probability (Ojala et al., 2001), the difference operation loses spatial information and histograms can become bad approximations of joint probability functions. In order to maintain spatial information, we propose to encode the texture's random structure by computing joint probabilities defining the dependence between pixels forming regions sharing intensity or colour properties. To reduce dimensionality, statistical distributions are computed for binary images. We combine joint distributions of binary values and the probability of intensity values to define a collection of histograms. These histograms are normalised, thus they can be used for non-parametric texture discrimination independently of the size of the sampled region (Puzicha et al., 1999).

We use the non-parametric classification to implement a segmentation application based on a hierarchical quadtree scheme. Hierarchical strategies have been very effective for image segmentation. The hierarchical approach has two main advantages. First, it performs a fast partition by considering regions rather than individual pixels in fixed overlapping windows. The partition is defined by considering regions of different sizes at different levels of the hierarchy. The second advantage is that it reduces classification errors due to mixture of features computed in a fixed window size. This is convenient to delineate accurate region borders (Ojala et al., 2000; Hsu, 1978; Dutra and Mascarenhas, 1984; Marceau et al., 1990; Briggs and Nellis, 1991; Ma and Manjunath, 1997). These properties have been exploited in algorithms of segmentation based on intensity (Horowitz and Pavlidis, 1976; Wu, 1992), motion (Szeliski and Shum, 1996; Lee, 1998) and texture information (Ojala et al., 2000; Chen and Pavlidis, 1979). Our segmentation is based on the technique presented

in (Chen and Pavlidis, 1979), but it divides an image using non-parametric classification.

2. Statistical characterisation of textures

The interdependence of pixels in a texture can be defined by the joint probabilities computed for random variables associated to pixel intensities. Given n discrete random variables $c(x_1)$, $c(x_2), \ldots, c(x_n)$ at positions $x_1, x_2, \ldots, x_n$, the nth-order density function defines the probability that the variables take the values $c_1, c_2, \ldots, c_n$, respectively (Papoulis, 1991). That is, $f(c_1, \ldots, c_n; x_1, \ldots, x_n) = P\{c(x_1) = c_1, \ldots, c(x_n) = c_n\}$. Here, c takes values within the range of possible grey levels or colours. Since textures are stationary processes, then distributions can be expressed independently of their position. Thus, $f(c_1, \ldots, c_n; x_1, \ldots, x_n)$ depends on the distance between x_1 and any other variable x_i. That is, $f(c_1, \ldots, c_n; \tau_2, \ldots, \tau_n)$ for $\tau_i = |x_i - x_1|$. This function represents the probability of obtaining the grey level values $c_1, c_2, \ldots, c_n$ at distances $\tau_1, \tau_2, \ldots, \tau_n$ measured from x_1. The general form can be limited to different orders to define alternative texture descriptors. For example, co-occurrence matrices (Haralick et al., 1973a; Kovalev and Petrou, 1996; Strand and Taxt, 1994) are defined by considering only two points. That is,

$$f(c_1, c_2; \tau) \quad \text{for } \tau = |x_2 - x_1| \tag{1}$$

for values of τ defining neighbourhoods of 3×3 or 4×4 pixels. These descriptors have powerful discrimination properties. However, the computation of the probability for each combination c_1 and c_2 requires a large number of samples and computations. Previous works have considered simplifications obtained by replacing the dependence on the values c_1 and c_2 for arithmetical combinations (Unser, 1986a; Rosenfeld et al., 1982; Ojala et al., 2001). For example difference histograms (Ojala et al., 2001) are defined as

$$f(c_1, c_2 - c_1; \tau) \tag{2}$$

This equation characterises the same information as Eq. (1), but the changes in intensities are given relative to the value at x_1. This can be simplified by considering that all the possible values of c_1 have the same probability of occurrence and that the probabilities of $P\{c(x_1) = c_1\}$ and $P\{c_1 - c_2 = \alpha\}$ are independent. In this case

$$f(c_1, c_1 - c_2; \tau) \cong f(\alpha; \tau) \quad \text{for } \alpha = c_1 - c_2 \tag{3}$$

This description can be extended to large neighbourhoods or to differences of higher order. The main advantage is that it is simpler than Eq. (1), thus it can make the classification faster and it requires less training data. The main drawback is that many combinations of values c_1, c_2 map into the same value of α, thus information about permutations is lost. Additionally, dependence on intensity information is lost making classification on small regions difficult. In the next section we present a characterisation that simplifies Eq. (1) keeping dependence in intensity information. Previous works have shown the importance of intensity values for classification (Dubuisson-Jolly and Gupta, 2000).

3. Conditional histograms

We can describe the interdependence of pixels in a texture by considering ideas of binary feature selection. Previous works have shown that binarisation can be very effective to characterise the spatial dependence of pixels in images (Wang and He, 1990; Chen et al., 1995; Hepplewhite and Stonham, 1997; Ojala and Pietikäinen, 1999). As suggested in (Ojala et al., 2000), we use the joint distribution of binary patterns. However, in order to be able to locate a texture embedded in different images, we use a global threshold strategy (Chen et al., 1995). A global threshold divides an image into regions with similar intensity properties. Thus, we can expect that the joint probabilities computed in a binary image to contain much of the information about the spatial structure of the texture.

If γ_1 and γ_2 represent binary features at x_1 and x_2, then the structural information given by the permutations in two locations can be written as

$$f(c_1 | \gamma_1, \gamma_2; \tau) \quad \text{for } \tau = |x_2 - x_1| \tag{4}$$

This represents the probability that a pixel has intensity c_1 conditioned to the intensities of

neighbouring pixels. Since we consider binary features, the second-order spatial and intensity interdependencies of neighbouring pixels are defined by the four histograms

$$
\begin{aligned}
f_1(c_1;\tau) &= f(c_1|0,0;\tau) \\
f_2(c_1;\tau) &= f(c_1|0,1;\tau) \\
f_3(c_1;\tau) &= f(c_1|1,0;\tau) \\
f_4(c_1;\tau) &= f(c_1|1,1;\tau)
\end{aligned}
\tag{5}
$$

This description can be extended to large neighbourhoods or n-order interdependencies. That is,

$$
f(c_1|\gamma_1,\gamma_2,\ldots,\gamma_n;\tau_2,\ldots,\tau_n) \quad \text{for } \tau_i = |x_i - x_1|
\tag{6}
$$

Thus a texture is characterised by 2^n one-dimensional histograms per binary feature. For colour or multispectral images, we can maintain a low dimensionality by considering each colour component separately. Thus, for an image with m chromatic components, we have $m2^n$ histograms. Although it is possible to consider directly the colour components in the definition of the histograms, in general, it is better to perform a pre-processing such that each component has a higher discriminatory ability. For example, if $c_{1,R}$, $c_{1,G}$ and $c_{1,B}$ are the red, blue and green values of an image (i.e., $m = 3$) then the description in Eq. (6) can be extended to

$$
f(c_{1j}|\gamma_1,\gamma_2,\ldots,\gamma_n;\tau_2,\ldots,\tau_n)
\tag{7}
$$

for

$$
\begin{aligned}
c_{1,1} &= (c_{1,R} + c_{1,G} + c_{1,B})/3 \\
c_{1,2} &= c_{1,R} - c_{1,B} \\
c_{1,3} &= (2c_{1,G} - c_{1,R} - c_{1,B})/3
\end{aligned}
\tag{8}
$$

These definitions can obtain a set of components with discriminatory ability as good as that obtained by the Karhunen Loeve transformation (Ohta et al., 1980).

4. Classification

In general, the classification performance depends on the discrimination approach. Numerous discrimination approaches are possible (Devijver and Kittler, 1982; Schalkoff, 1992) and classifiers can improve the results at the expense of complexity, computational resources and requirements in the size and quality of the training data. However, it is beyond the scope of this work to evaluate classification schemes. We are interested in evaluating the discrimination properties of conditional density histograms. We have chosen to use a non-parametric classification based on the dissimilarity between the histograms of the training classes and the histograms of the sample to be assigned to the class. The non-parametric approach is particularly suitable for low dimensionality feature spaces and can provide good classification results with relative low computational resources (Puzicha et al., 1999).

In non-parametric discrimination techniques, histograms define the feature vectors that form the basis for the classification. Thus, each element in the histogram corresponds to the value of a feature and a texture class is characterised by $m2^n$ features per binary operator. An important difference with previous approaches is that in our classifier we assume that training samples define the same distribution. That is, instead of forming different distributions for each training sample, we increment the estimate of a single collection of distributions for each texture class. Thus, the training data of a class characterises a single point in the feature space, making the classification to be more dependent on the selection of good texture characterisation rather than on the power of the classifier. If we included a sophisticate classifier that can distinguish between no-linear separable classes defined by several features, then we would compensate for poor features by using complex discrimination for disperse collections in the feature space. Thus, the performance would be directly related to the size of the training data and to the complexity of the classifier and not to the quality of the features.

An advantage of computing features incrementally is that we can obtain an estimate of the distributions with a small number of training samples avoiding sparse histograms. It is well known that histograms with few entries per bin produce a poor performance in non-parametric tests (Ojala et al., 1996; Puzicha et al., 1999). In some applications, it is important to be able to classify by using reduced training data. For example, in image editing

or region tracking applications, it is important to minimise the number of times the user selects samples to delineate a region of interest. Additionally, in these applications, training cannot be performed off-line, thus we have to minimise the time spent in computing features and in creating the texture database.

In non-parametric classification techniques, there are alternative ways to define the dissimilarity between histograms. Valuable experimental work has evaluated the performance of alternative dissimilarity measures for k-NN classifiers (Ojala et al., 1996; Puzicha et al., 1999). The results suggest that the performance depends on several factors such as the size of the training data, type of features, type of images and particular applications. In general, when histograms are not sparse (e.g., by using marginal distributions or by considering many training samples (Puzicha et al., 1999)), the difference in performance is not significant for most measures. For small sample data, it is better to use measures based on statistics or aggregate measures less sensitive to sample noise. Since we are using only one set of histograms per class, thus we expect (and in fact this is part of the motivation to reduce dimensionality) to have histograms with a significant number of entries per bin. Thus, we choose to use the L_1 norm that provides a good classification with a small computational load. In the case of windows with few pixels, the performance can be maintained if the number of bins is reduced. This can be achieved by adaptive binning, but this is computationally expensive (Ojala et al., 1996). In a simpler strategy, we constrain the similarity measure to bins that have a significant number of entries.

By considering Eq. (6), a test sample S is assigned to the class of the model M_j that minimises the absolute difference between corresponding bins in each histogram.

$$D(S, M_j) = \sum_h^{m2^n} \sum_i |f_{S_h}(c_i|\gamma_1, \gamma_2, \ldots, \gamma_n; \tau_2, \ldots, \tau_n) - f_{M_{j,h}}(c_i|\gamma_1, \gamma_2, \ldots, \gamma_n; \tau_2, \ldots, \tau_n)| \quad (9)$$

The fist summation indicates all the histograms whilst the second summation iterates for each bin. The sub-index h is used to index the histogram

of the sample and the model. This definition measures whether the pixels in two textures have similar intensities with similar spatial organisation. That is, the value of $D(S, M)$ will be small if the intensity values of the two textures are similar and they are grouped into regions of similar contrast.

5. Segmentation

We used the classification to implement a segmentation application based on a top-down hierarchical subdivision. This approach searches for an optimum partition by dividing the image in a quad-tree homogenous decomposition. This comprises three steps. First, a region is classified. Secondly, it is partitioned and each partition is classified. Finally, it is necessary to measure the homogeneity of the partition. If the region is homogenous, then the whole region is assigned to the same class and the subdivision is stopped. If the region is not homogenous, the region is subdivided. The subdivision is repeated until the image region is equal to one pixel.

In a hierarchical approach, the segmentation performance depends on the classification and on the ability of computing an optimum partition. An optimal partition divides the image into regions of roughly uniform texture. Thus, the success depends on performing an appropriate decision about the homogeneity of a region. Unfortunately, there has not been a practical evaluation of homogeneity measurements that could help us to choose an optimum partition framework. We base our criterion of homogeneity in two heuristic rules. First, we consider that classification at boundaries contains a mixture of two textures (i.e., non-homogeneous). Accordingly, we subdivide regions that have at least one neighbouring region of a different class. This criterion of subdivision delineates the boundaries between texture regions. The second criterion of homogeneity is based on the classification in successive levels of the quad-tree. Similar to (Ojala et al., 2000), we measure the uniformity by computing the differences between histograms of the four sub-blocks in the subdivision. However, we do not

measure the variance of the sub-blocks, but the confidence that we have in the result.

We consider the uncertainty in classification as the ratio between the similarity values of the two textures in the database that are most similar to the distribution computed from a region. We denote as j^* the element that minimises $D(S, M_j)$ and we denote as $D(S, M_{j^*})$ the corresponding minimum value. We denote as j^+ the class that minimises $D(S, M_j)$ when the element j^* is not considered. Thus, $D(S, M_{j^+})$ denotes the distance value for the class j^+. The uncertainty of classifying S is defined as

$$U(S, j^*) = \frac{D(S, M_{j^+})}{D(S, M_{j^*})} \tag{10}$$

This measure will be close to one if the classification is vague. The uncertainty in the classification decreases when the difference between classification distances increases. Thus, the subdivision must minimise the uncertainty in the entire image. We perform the subdivision based on the comparison between the uncertainty of the region and the uncertainty of the new regions. However, this comparison cannot be based on the uncertainty between different levels in the quad-tree. Since each level of the quad-tree has regions of smaller size, then the uncertainty in lower levels is always higher.

In order to evaluate the subdivision, we consider the uncertainty when regions are classified by considering the class in the upper level and the new classification in the low level. If S_1, S_2, S_3 and S_4 are the regions in a subdivision, then the change in uncertainty due to a splitting operation can be measured as

$$U(S_1, j^*) + U(S_2, j^*) + U(S_3, j^*) + U(S_4, j^*)$$
$$- (U(S_1, j_1^*) + U(S_2, j_2^*) + U(S_3, j_3^*) + U(S_4, j_4^*)) \tag{11}$$

where j^* minimises in the upper level and, j_1^*, j_2^*, j_3^* and j_4^* minimise each one of the new sub-regions. The sub-division is performed if the absolute difference in Eq. (11) is lower than a fixed threshold. The basic idea is to subdivide the region only if it is composed of several textures. In this case, the classification obtained by smaller regions

composing a mixture has less uncertainty than when we consider a single class for the whole region.

Although uncertainty is capable of giving a useful measure of homogeneity, still it is extremely difficult to classify small blocks in an image. If the window is to small, then it probably does not contain sufficient information to characterise the region, thus increasing the probability of misclassification (Ojala et al., 2000). In order to obtain an accurate delineation of texture regions, we reduce the number of potential classes. When the subdivision is due to a boundary and the window size is smaller than a fixed threshold, the classification is made by considering only the texture classes of current and neighbour regions. That is, we assume that there are not new regions smaller than the fixed threshold, thus the segmentation can be stopped and the subdivision can only be used to delineate the existing coarse regions. The threshold size determines the minimum data necessary to obtain a good classification and it is strongly dependent on the number of texture categories. As more classes are included, the probability of misclassification of small data increases. Thus, the threshold should be increased such that, the classification of regions is made by including enough information. The threshold should be set to the minimum window size for which the classification obtains reliable results. This value is an input parameter of the classification application.

6. Experimental results and examples

6.1. Experimental data

In order to assess the discrimination capabilities, we have performed two experimental tests based on the data presented in (Valkealahti and Oja, 1998; Ojala et al., 1996, 2001; Ohanian and Dubes, 1992). The first test (Valkealahti and Oja, 1998; Ojala et al., 2001) defines 32 texture categories from selected images of the Brodatz collection. The second test (Ojala et al., 1996; Ohanian and Dubes, 1992) defines 16 texture categories from four types of images.

For the 32-category problem, texture samples are obtained from a 256×256 image with 256 grey levels. Each image is subdivided into 16 blocks of 64×64 pixels and each block is transformed into three new blocks by 90° rotation, scale from the 45×45 pixels in the middle and by combining rotation and scaling. This produces 2048 blocks. Half of the data is selected by randomly choosing 8 blocks and the corresponding transformed blocks. This data is used to define class histograms and the remaining blocks to evaluate the classification process.

The 16-category problem uses four types of images containing four distinct textures. Two types of textures are generated from fractal and Gaussian Markov random field models. The other two types were obtained from leather and painted surfaces. Images are divided into 256 non-overlapping regions of 32×32 pixels. Only 200 regions of each image are selected to obtain 3200 samples. Performance is measured by considering the leave-one-out error. Thus, 3184 samples are used for training and 16 to assess the discrimination.

6.2. Computation of features

In contrast to (Valkealahti and Oja, 1998) where feature histograms are computed for each texture block separately by considering a 4×4 neighbourhood, we obtain histograms for each texture category by considering only a 2×2 neighbourhood. Thus, the 32 training blocks of each class are used to compute 4×4 conditional histograms. The first collection of four represents the class without any geometric transformation; the second collection is used for the scale, and the last two for rotation and combination of scale and rotation. To obtain binary features, we use three fixed thresholds with values of 128, 64 and 32. Thus, each class is represented by $4 \times 4 \times 3$ histograms.

6.3. Classification results

Table 1 shows the average classification results obtained for 10 random selected test sets for the first test. The table shows the average for each class.

Table 1
Average classification accuracies (%) over 10 experiments for 32 texture categories

Texture	Accuracy
bark	100.00
beachsand	85.00
beans	100.00
burlap	100.00
d10	61.56
d11	95.63
d4	96.25
d5	96.88
d51	100.00
d52	100.00
d6	100.00
d95	100.00
fieldstone	76.25
grass	99.06
ice	63.13
image09	100.00
image15	96.56
image17	92.50
image19	100.00
paper	99.38
peb54	97.81
pigskin	92.50
pressdcl	100.00
raffia	100.00
raffia2	100.00
reptile	100.00
ricepaper	100.00
seafan	93.13
straw2	100.00
tree	92.19
water	100.00
woodgrain	100.00

In general, the classification performance is very good with exception of the classes: beachsand, D10, ice and fieldstone. A detailed observation of the results showed that most misclassifications for these classes are for blocks obtained by the scale transformation. Fig. 1 shows six examples of these misclassifications. Fig. 1(a) shows a block of beachsand that was misclassified as the class grass. The second texture in Fig. 1(a) shows an example of this class. Fig. 1(b)–(f) shows other misclassifications obtained for the scaling classes. The remarkable similarity between the textures in Fig. 1 can explain the misclassifications. Additionally, since scaling is obtained from digital images, there is some lost in resolution and as consequence aliasing produces
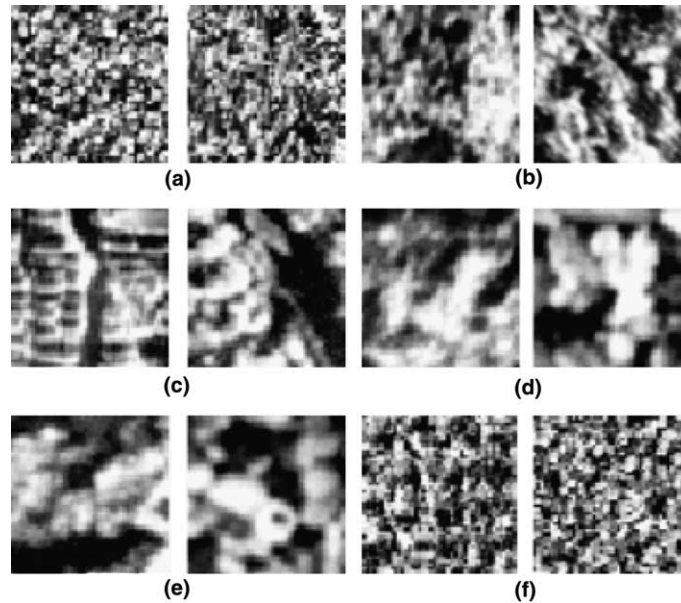
Fig. 1. Examples of misclassification. (a) Beachsand and grass, (b) ice and fieldstone, (c) D10 and tree, (d) fieldstone and peb54, (e) peb54 and fieldstone and (f) pigskin and beachsand.

new blocks with very similar values for neighbouring pixels. Thus, spatial information obtained by considering neighbouring pixels is not so good as the original blocks.

Table 2 shows the global performance of the classification. The table includes the results presented in (Valkealahti and Oja, 1998; Ojala et al., 2001) for signed differences (Valkealahti and Oja, 1998), co-occurrences (Haralick et al., 1973b), absolute differences (Ojala and Pietikäinen, 1996), Gaussian random field model (Van Gool et al., 1985), reduced histograms (Ojala et al., 2001) and channel histograms (Unser, 1986b). In general, the accuracy of the proposed features compares to the most successful techniques. However, it is important to notice that quantisation and signed grey level distributions required 16th and 9th order probabilities. We have used second order joint probabilities in a 2×2 neighbourhood. This makes the complexity adequate for applications requiring online training.

Table 3 shows the result for the second test. The table includes the results presented in (Ojala et al., 1996; Ohanian and Dubes, 1992) for local

binary patterns (Wang and He, 1990), co-occurrences (Haralick et al., 1973b), grey level differences (Unser, 1986a) and classification based on a combination of three different texture features (Ohanian and Dubes, 1992). For the leave-one-out classification, the classification accuracy for conditional histograms was of 100%. In order to highlight the advantage of conditional histograms, we perform the same test by reducing the training data. In our results good performance can be maintained with only 10% to 5% of the data. Table 4 shows the classification results for only 20 training samples and windows of 32×32 and 16×16. For the 32×32 case this represents the 10% of training data. When the window is reduced to 16×16 only 2.5% of the original data is used. Classification performance is maintained with less training data because, as explained in Section 4, histograms are computed by using data incrementally. Thus, when we reduce training data, we are not reducing the number of features, but the accuracy of the features. But since histograms are just one dimensional, then they can be populated with few data.

Table 2
Global performance for 16 texture categories

	Average classification
Signed differences (LPB)	
Second order	93.3
Fourth order	95.7
Eighth order	96.8
Co-occurrence matrices	
Third order	90.8
Fifth order	93.8
Ninth order	94.4
Absolute differences	
Second order	85.3
Fourth order	92.1
Eighth order	93.2
MRF	
Seventh order	71.3
Combined features	90.0
Reduced histograms	
TSOM, cosine transform	93.9
VQ, cosine transform	93.4
TSOM, grey levels	92.8
Channel histograms	
Multi-dimensional	90.4
One-dimensional	78.2
Conditional histograms	
Second order	94.91

Table 3
Global performance for 16 texture categories

Local binary patterns	
LBP	81.40
LPB and contrast	87.62
LBP and covariance	185.03
Co-occurrence features	
4 Features	188.25
9 Features	90.69
Grey level difference	
DIFFX and DIFFY	96.56
Combined features MRF, Gabor, Fractal	
4 Features	191.07
9 Features	195.41
Conditional histogram	
32×32 (199 samples per class)	100.00
32×32 (20 samples per class)	1100.00
16×16 (20 samples per class)	195.69

Table 4
Average classification accuracy (%) for 16 texture categories

Texture	32×32	16×16
Fractal1	100.00	100.00
Fractal2	100.00	81.11
Fractal3	100.00	100.00
Fractal4	100.00	100.00
mrf1	100.00	100.00
mrf2	100.00	92.78
mrf3	100.00	100.00
mrf4	100.00	94.44
Leather1	100.00	100.00
Leather2	100.00	100.00
Leather3	100.00	100.00
Leather4	100.00	84.44
paint1	100.00	100.00
paint2	100.00	100.00
paint3	100.00	78.33
paint4	100.00	100.00

Results obtained by considering 20 training samples per class.

6.4. Segmentation

Figs. 2 and 3 show selected examples of the segmentation application for grey level and colour images. In these examples, training data was obtained by considering two windows of 32×32 for each class. For each example, we show the segmentation regions, the borders between classes and the uncertainty in the classification. Uncertainty is shown as a colour image whose change in brightness represents the confidence of the classification of each pixel. Brighter colours represent a high degree of confidence, whilst darker colours show areas whose classification is more uncertain. We can see that in general large areas of uniform texture are classified in regions of large size with low uncertainty, whilst high detailed areas are divided into small regions with high uncertainty.

The examples show that texture features can be used in applications to obtain well-delineated borders. Notice that as boundaries are refined, regions reduce the confidence in the classification. Figs. 2(a) and 3(a) show two examples of segmentation for synthetic images composed of grey level and colour textures, respectively. Each image has a resolution of 256×256 pixels. We can observe that the refinement of regions provides an analysis capable of producing accurate segmentation results. In these examples, the colour information
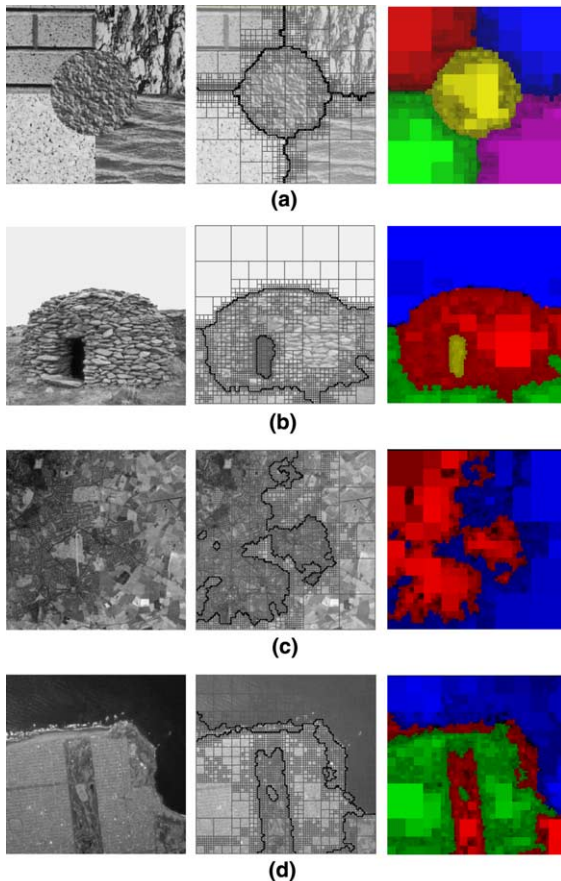
Fig. 2. Examples of texture images. (a) Synthetic image composed of five grey level textures, (b) image containing grey level natural textures, (c) infrared band of a satellite image and (d) intensity image.
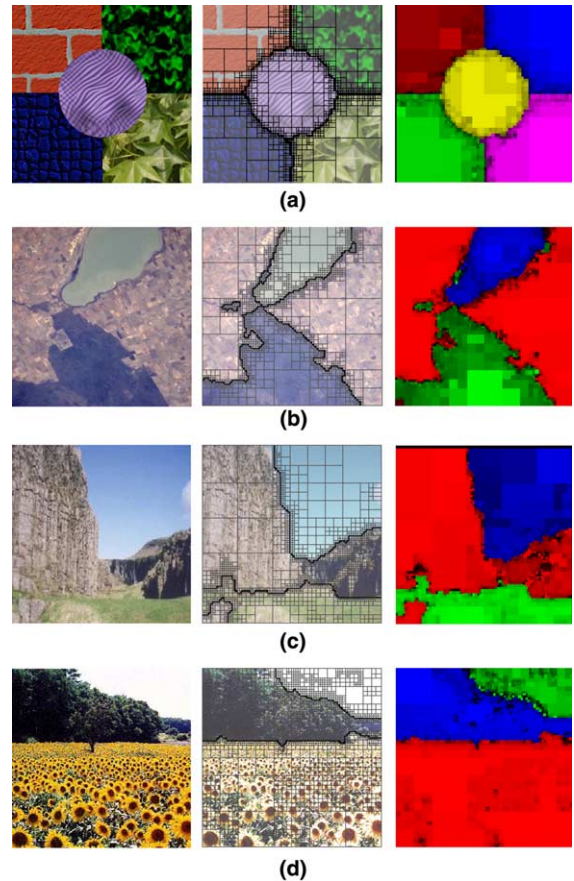

Fig. 3. Examples of colour texture images. (a) Synthetic image composed of five colour textures, (b) satellite image with three colour bands, (c) and (d) images containing colour natural textures.

produces a clearer delineation of regions since small regions are more accurately classified. The detailed evaluation of the segmentation results shown that 98.7% and 99.2% of pixels in the grey level and colour images, were correctly classified. Although the difference between the results of the final classification is very small, we can observe a clear distinction in the uncertainty maps. We can observe that whilst for the colour image a high confidence in the classification results is maintained for regions of reduced size, grey level information tends to produce less contrasted and darker regions. Thus, when regions are small, the lack of information increases the probability of erroneous classification for grey level textures.

However, since the increase in probability is only significant for small regions, then the error of the whole segmentation process is small.

The example in Fig. 2(b) contains a grey level image with four types of regions. Two types of regions have a rather regular appearance with white and black intensities, respectively. In spite of the lack of texture, the intensity component in the characterisation produces a successful classification. Fig. 2(c) corresponds to the infrared band in a multispectral LANDSAT image. A training set was defined by selecting two regions as representative members of urban and agricultural land covers. In this example, we can see that the model of texture developed can be used to accurately

delineate textured regions. The uncertainty map provides useful information for classification. For example, although the top left region was classified as urban area, the low confidence value shows that it probably contains a different type of land cover. In the same manner, small black regions in the red areas indicate possible fields inside an urban area. The results in Fig. 2(d) show an accurate segmentation capable of discriminating between regions of constant intensity and regions with textured patterns.

Fig. 3(b) shows a similar example to Fig. 2(d). In this case, the uncertainty is reduced due to colour data. We can observe an accurate segmentation even for small regions. Fig. 3(c) and (d) show two examples of colour texture landscapes. In Fig. 3(d), the green tones in two texture classes produce a considerable similarity between some regions. Accordingly, the segmentation process subdivides several times the regions that contain green areas. Although small green regions have a large uncertainty with respect to larger areas, the classification results lead to accurately delineated borders. The segmentation of the image in Fig. 3(c) has a more clear distinction of classes. Thus, larger regions with low uncertainty are obtained. We can notice that in regions where rock and grass merge, the uncertainty increases and the subdivision becomes finer in order to identify pure classes.

7. Conclusions and further work

We have proposed a characterisation of textures based on a mixture of colour and contextual information obtained from binary features. The characterisation defines one-dimensional histograms that represent the conditional probability of intensity values given the joint probabilities of pixels in image regions. Experimental results show that a non-parametric classification based on conditional histograms produces a compact and powerful set of features. High classification performance is obtained by considering only second order distributions. The compactness of the representation has three main interests. First, compactness is important to make texture analysis practical. This is particularly relevant for applica-

tions requiring on-line database construction. Secondly, it avoids sparse histograms that can reduce the classification performance. Finally, since the number of bins is reduced, compactness minimises the data required during the training step. We reinforce this last point by considering training data incrementally.

We have included examples that show the application of the classification to region delineation by means of a hierarchical subdivision. Examples show that the classification is useful to obtain well-delineated borders. The dependence of the representation on intensity data is suitable to classify regions of small sizes. Our current work is considering the potential implications of incremental training. We think that distributions can be used to determine when training can be stopped and to detect when training data agree with a single distribution. If data do not agree with a single distribution, then several classes should be used to represent a texture. Additionally, we consider extending the approach to applications on non-supervised segmentation. There are recent studies where efficient unsupervised segmentation is performed using feature distributions (Ojala et al., 2000).

References

Augusteijn, M.F., Clemens, L.E., Shaw, K.A., 1995. Performance evaluation of texture measures for ground cover identification in satellite images by means of a neural network classifier. IEEE Trans. Geosci. Remote Sensing 33 (3), 616–626.

Briggs, J.M., Nellis, M.D., 1991. Seasonal variation of heterogeneity in the tallgrassprairie: A quantitative measure using remote sensing. Photogramm. Eng. Remote Sensing 57 (4), 407–411.

Chen, P.C., Pavlidis, T., 1979. Segmentation by texture using a co-occurrence matrix and a split and merge algorithm. Comput. Vision Graphics Image Process. 10, 172–182.

Chen, Y.Q., Nixon, M.S., Thomas, D.W., 1995. Statistical geometric features for texture classification. Pattern Recognition 28, 537–552.

Chetverikov, D., 1994. GLDH based analysis of texture anisotropy and symmetry: An experimental study. Proc. Internat. Conf. on Pattern Recognition I, 444–448.

Devijver, P.A., Kittler, J., 1982. Pattern Recognition, a Statistical Approach. Prentice Hall, Englewood Cliffs, London.

Dubuisson-Jolly, M.-P., Gupta, A., 2000. Color and texture fusion: Application to aerial image segmentation and GIS updating. Image Vision Comput. 18 (10), 823–832.

Dutra, L.V., Mascarenhas, N.D.A., 1984. Some experiments with spatial feature extraction methods in multispectral classification. Int. J. Remote Sensing 5 (2), 303–313.

Gersho, A., Gray, R.M., 1992. Vector Quantization and Signal Compression. Kluwer Academic Publishers, Dordrecht, Netherlands.

Haralick, R., 1979. Statistical and structural approaches to texture. Proc. IEEE 67, 786–804.

Haralick, R., Shanmugam, K., Distein, I., 1973a. Textural features for image classification. IEEE Trans. Systems Man Cybernet. 3, 610–621.

Haralick, R., Shanmugam, K., Dinstein, I., 1973b. Textural features for image classification. IEEE Trans. Systems Man Cybernet. SMC-3, 610–621.

Hepplewhite, L., Stonham, T.J., 1997. N-tuple texture recognition and the zero crossing sketch. Electron. Lett. 33 (1), 45–46.

Hofmann, T., Puzicha, J., Buhmann, J., 1998. Unsupervised texture segmentation in a deterministic annealing framework. IEEE Trans. Pattern Anal. Mach. Intell. 20 (8), 803–818.

Horowitz, S.L., Pavlidis, T., 1976. Picture segmentation by a tree traversal algorithm. J. ACM 23 (2), 368–388.

Hsu, S., 1978. Texture-tone analysis for automated land-use mapping. Photogramm. Eng. Remote Sensing 44 (11), 1393–1404.

Jain, A., Zongker, D., 1997. Feature selection: Evaluation, application and small sample performance. IEEE Trans. Pattern Anal. Mach. Intell. 19 (2), 153–158.

Kovalev, V., Petrou, M., 1996. Multidimensional co-occurrence matrices for object recognition and matching. Graph. Models Image Process. 58 (3), 187–197.

Lee, J.W., 1998. Joint optimization of block size and quantization for quadtree based motion estimation. Image Process. 7 (6), 909–912.

Ma, W., Manjunath, B., 1997. Edge flow: A framework of boundary detection and image segmentation, Proc. IEEE Conf. on Computer Vision and Pattern Recognition, pp. 744–749.

Marceau, D.J., Howarth, P.J., Dubois, J.M., Grattton, D.J., 1990. Evaluation of the grey-level co-occurrence matrix method for land-cover classification using SPOT imagery. IEEE Trans. Geosci. Remote Sensing 28 (4), 513–519.

Ng, L.S., Nixon, M.S., Carter, J.N., 1998. Combining feature sets for texture extraction. Proc. IEEE Southwest Symposium on Image Analysis and Interpretation, Texas, pp. 103–108.

Ohanian, P.P., Dubes, R.C., 1992. Performance evaluation for four classes of textural features. Pattern Recognition 25 (8), 819–833.

Ohta, Y.-I., Kanade, T., Sakai, T., 1980. Color information for region segmentation. Comput. Graph. Image Process. 13, 222–241.

Ojala, T., Pietikäinen, M., 1996. A comparative study of texture measures with classification based on feature distributions. Pattern Recognition 29 (1), 51–59.

Ojala, T., Pietikäinen, M., 1999. Unsupervised texture segmentation using feature distributions. Pattern Recognition 32, 477–486.

Ojala, T., Pietikäinen, M., Harwood, D., 1996. A comparative study of texture measures with classification based on feature distributions. Pattern Recognition 29, 51–59.

Ojala, T., Pietikäinen, M., Mäenpää, T., 2000. Gray scale and rotation invariant texture classification with local binary patterns, Proc. Sixth European Conf. on Computer Vision, Dublin, Ireland, vol. 1, pp. 404–420.

Ojala, T., Valkealahti, K., Oja, E., Pietikäinen, M., 2001. Texture discrimination with multidimensional distributions of signed gray level differences. Pattern Recognition 34 (3), 727–739.

Papoulis, A., 1991. Probability, Random Variables, and Stochastic Processes. McGraw–Hill.

Puzicha, J., Rubner, Y., Tomasi, C., Buhmann, J.M., 1999. Empirical evaluation of dissimilarity measures for color and texture. Proc. IEEE Internat. Conf. on Computer Vision (ICCV'99), pp. 1165–1173.

Rosenfeld, A., Wang, C.Y., Wu, A.Y., 1982. Multispectral texture. IEEE Trans. Systems Man Cybernet. 12 (1), 79–84.

Schalkoff, R., 1992. Pattern Recognition. Statistical, Structural and Neural Approaches. Wiley.

Schulerud, H., Carstensen, J., 1995. Multiresolution texture analysis of four classes of mice liver cells using different cell cluster representations. Proc. 9th Scandinavian Conf. on Image Analysis, Uppsala, Sweden, pp. 121–129.

Strand, J., Taxt, T., 1994. Local frequency features for texture classification. Pattern Recognition 27 (10), 1397–1406.

Sullis, J.R., 1990. Distributed learning of texture classification. Lect. Notes Comput. Sci. 427, 349–358.

Szeliski, R., Shum, H.-Y., 1996. Motion estimation with quadtree splines. IEEE Trans. Pattern Anal. Mach. Intell. 18 (12), 1199–1210.

Unser, M., 1986a. Sum and difference histograms for texture classification. IEEE Trans. Pattern Anal. Mach. Intell. 8, 118–125.

Unser, M., 1986b. Local linear transforms for texture measurements. Signal Process. 11 (1), 61–79.

Valkealahti, K., Oja, E., 1998. Reduced multi-dimensional co-occurrence histograms in texture classification. IEEE Trans. Pattern Anal. Mach. Intell. 20 (1), 90–94.

Van Gool, L., Dewaele, P., Oosterlinck, A., 1985. Texture analysis anno 1983. Comput. Vision Graphics Image Process. 29, 336–357.

Wang, L., He, D.C., 1990. Texture classification using texture spectrum. Pattern Recognition 23 (8), 905–910.

Wu, X., 1992. Image coding by adaptive tree-structured segmentation. IEEE Trans. Inform. Theory 38 (6), 1755–1767.