A SET OF MULTIRESOLUTION TEXTURE FEATURES SUITABLE FOR UNSUPERVISED IMAGE SEGMENTATION

Ioannis Matalas, Stephen Roberts and Harry Hatzakis
Department of Electrical and Electronic Engineering
Imperial College of Science, Technology and Medicine
London SW7 2BT, U.K.
email: imatal@ic.ac.uk

Abstract

We propose a set of multiresolution features for texture description. Image smoothing at multiple scales using the fast smoothing B-spline transform is performed and a number of features, such as the local area, the normal vector dispersion and the gradient orientation, are computed from each scale. A simple disparity function is applied to assess the discriminative power of these features with comparison to other texture methods. Being effective even for small observation windows, the proposed features are suitable for high-resolution texture segmentation.

1 Introduction

Texture is an important element in vision and has been analysed in its own right for the last three decades by researchers in computer vision. Several measures of texture have been used to discriminate between different objects and to segment scenes. Many researchers have focused on texture features such as coarseness, "linelikeness", contrast, regularity, disorder, directionality, complexity, busyness, strength etc. Tamura *et al* [8] and more recently Rao and Lohse [7], in an attempt to find a general description common to all Brodatz textures, considered a large set of high-level features and found a strong correlation between certain features in this set (e.g., linearity and directionality). Both concluded that the most significant features turned out to be coarseness, contrast and directionality. It still remains, however, to decide upon the computational measures that can be used to quantify these features.

In this work we propose such a set of texture measures. This consists of the area of the image surface, the gradient direction, the vector dispersion and the grey level at each pixel of the image. All of these measures are computed at different degrees of smoothing. Such a multiresolution approach (see also [10, 3]) is justified by the fact that texture has scale dependent properties. The "disparity" measure used to test the homogeneity (or closeness) between a pair of regions is base on the Kolmogorov-Smirnov distance. Experiments with Brodatz textures and comparison with three other texture description techniques (Laws filters, Gabor filters and Gaussian MRF model) demonstrate the effectiveness of the proposed feature set.

2 The proposed set of texture features

As it should already be clear, our intention is to use a small set of features which captures the basic properties used by humans [8, 7] for texture discrimination, i.e., coarseness, directionality and contrast. The procedure to compute our texture features is the following:

Initially, the image intensity field is approximated by a number of B-spline surfaces using the fast *smoothing B-spline transform* (see appendix). Given this B-spline analytical representation, it is quite easy to compute estimates of the *surface area*, the *gradient direction*, the *normal vector dispersion* and the *smoothed grey level* at each pixel. The normal vector dispersion (see [2], p.470) is a measure of how dispersed the unit vectors normal to the image surface are in a neighbourhood (say 7×7) around the given pixel. It takes values close to 1 for even surfaces and close to 0 for very rough surfaces.

These four features can be computed at any scale of smoothing (determined by a smoothing coefficient λ) providing four "feature images" $\mathbf{y}_{\lambda}^{(i)}$, $i = 1, ..., 4$ for each scale. We found that smoothing at three different scales is, in general, adequate to describe most Brodatz textures.

Clearly, texture characterises regions and not individual pixels. Using the precalculated "feature images" we can easily compute the first order statistics (i.e., histograms) for each feature in any region (of any shape) in the image. The histograms of "surface area" and "vector dispersion" indicate how coarse or fine a texture is at a given smoothing scale. The histogram of the gradient direction (from 0° to 180°) denotes the predominant directions if any (a flat histogram corresponds to a non-directional texture). Finally, the grey level and surface area statistics characterise the contrast of the texture.

There is, however, a strong correlation between any two features $\mathbf{y}_{\lambda_\alpha}^{(i)}$ and $\mathbf{y}_{\lambda_\beta}^{(i)}$ of the same type at two different scales. A simple way to reduce this correlation is to use *differential features* $\mathbf{y}_{d_{\alpha\beta}}^{(i)} = \mathbf{y}_{\lambda_\alpha}^{(i)} - \mathbf{y}_{\lambda_\beta}^{(i)}$ More-

over, these differential features carry information about the cross-scale behaviour of textures which is often very important for texture discrimination. We found experimentally that the introduction of differential features for the surface area and the smoothed intensity, generally, improves the discriminative capability of our set. On the other hand, for the gradient direction and the vector dispersion, the differential features cannot be said to affect this capability in a positive way.

It should be stressed that our ultimate goal is unsupervised image segmentation and not simply texture classification and in order to achieve accurate segmentations at high resolution we need to use features that can be both computed reliably in irregularly shaped regions and also perform reasonably well when computed in small regions. For instance, spatial frequency methods as the Fourier power spectrum can only be computed in rectangular blocks and thus are not preferred.

In section 4 we compare the performance of the proposed feature set with three other texture description approaches: (a) The Laws masks [4] of size 5×5 to demonstrate that even a small set of 4 features (at one scale only) can characterise texture better than the much larger set of the 25 Laws masks, (b) The even symmetric Gabor filters (see [3]) with 4 orientations and 5 frequencies, and (c) The Gaussian Markov Random Field (GMRF) model which is a popular method to model textured random fields and has been extensively used in the context of unsupervised segmentation (see [11, 5, 6]). This comparison shows the more consistent behaviour of our features and an ability to characterise a much larger range of textures. Moreover, another advantage of our method with respect to the GMRF model is that the maximum likelihood estimates of the GMRF parameters cannot easily be calculated in irregular regions, while the computation of the pseudo-likelihood estimates is significantly more expensive than the simple computation of feature histograms (as in our case).

In our experiments, we observed that we can achieve a significant improvement in texture discrimination ability when we include second order statistics of the 4 features in our set. More specifically, for each pixel site s, we consider a neighbourhood N_s and for each "feature image" $\mathbf{y}^{(i)}$ we find the absolute difference $z_r^{(i)} = |y_s^{(i)} - y_{s+r}^{(i)}|$, $(r \in N_s)$. Thus, for each neighbour r and for each feature (i) we can find the histogram of $z_r^{(i)}$ and use it as an additional feature. Using this method, however, the number of features increases excessively. We attempted to reduce their number by considering groups of neighbours, forming "rings" of increasing size. For instance, when referring to the 7×7 neighbourhood of fig. 2 we compute three second order histograms corresponding to the neighbours labelled with 1, 2 and 3. Thus, their number reduces from 36 to only 3.

3 The disparity function

Many successful texture classification methods use a large number of (typically) correlated features (such as those calculated by the co-occurence matrices or local linear transforms) and then apply a feature space transformation in order to select a new feature space space of small dimensionality with the best possible discriminative power. However, this feature space transformation is texture dependent and requires the calculation of covariance matrices for the textures. When the task in hand is unsupervised segmentation, where training samples are not available, the calculation of the statistics of each texture class is computationally expensive since it has to be done "on-line" with the segmentation (similar to the K-means algorithm).

Our viewpoint is that instead of trying to find exactly the "right" features with the maximum discriminative power, one should find a simple and fast way of integrating multiple or even redundant features. A way to do this is by using a "disparity" (or homogeneity) function based on the Kolmogorov-Smirnov (KS) distance (see [1]). When applied to the histograms $h_1(i)$ and $h_2(i)$ of two data sets $\mathbf{v}_1$ and $\mathbf{v}_2$ the KS distance provides a simple test statistic for the hypothesis that the two sets are samples from the same underlying distribution. The KS distance is the maximum (vertical) distance between the cumulative histograms $F_k(t) = \sum_{i=-\infty}^{t} h_k(i)$, $(k = 1, 2)$ of the two data sets, i.e.,

$d(\mathbf{v}_1, \mathbf{v}_2) = max_{-\infty<t<\infty}|F_1(t) - F_2(t)|$

The KS distance always lies in $[0, 1]$ and has a desirable invariance property: it is invariant to strictly monotone transformations of the data. We can now define as the "disparity" measure between two region D_1 and D_2 (which can be of any shape and are not necessarily connected) the function

$$\Delta(D_1, D_2) = max_{1 \le i \le m} d(\mathbf{y}^{(i)}(D_1), \mathbf{y}^{(i)}(D_2)) \qquad (1)$$

where $\mathbf{y}^{(i)}(D)$ denotes the data in a region of the i^{th} "feature image". Again $0 \le \Delta(D_1, D_2) \le 1$ and a value close to 0 (1) means that the likelihood of homogeneity is very large (small).

Apart from its invariance to strictly monotone transformations this disparity measure has the property of having the same interpretation anywhere in the image (i.e., it is automatically calibrated to the differing statistical properties of different areas of the image).

We have chosen to use this type of disparity function in our experiments, even though it may not be the best possible way to perform texture classification, because it provides a good basis to compare different feature sets. Note that the histogram is a compact representation of all the first order information of a random set, whereas first order moments (mean, variance etc.) carry only partial first order information.

4 Experimental results

In order to evaluate the discriminative power of our features we performed some simple classification experiments using Brodatz textures. We used one set of 7 Brodatz microtextures (256×256) and trained the classification system using our method, the Laws filters, the Gabor filters and the GMRF model of 5^{th} order (for this order of the GMRF model we found the best classification results). Then we subdivided each pattern into non-overlapping blocks of sizes (32×32), (16×16) and (8×8). For the first three methods, we used the disparity function of section 3 and for the GMRF model we used the joint conditional PDF of the data in a block (see [11, 6]) in order to classify each block to the "closest" class.

Figures 3(a)-(g) show the correct classification rates of the methods for different block sizes. It can be seen that the proposed set has, in general, a better and more consistent behaviour than either the Laws masks (e), the Gabor filters (f) or the GMRF model (g). Its performance is further improved if we consider the second order histograms of the proposed features (b) resulting from the neighbourhood configuration of fig. 2. Improved classification rates are obtained when multiple smoothing scales are used (c), especially when differential features are included (d).

Notice that the GMRF does not have very consistent behaviour; while it models some textures very well it seriously fails for others (e.g., D9). We also experimented with macrotextures and we found that the proposed set performed even more favourably compared to the linear filtering methods, while the GMRF model again presented an inconsistent behaviour.

Figures 1(a)-(g) show the segmentation of a collage of 5 Brodatz textures (a). The correct segmentation is shown in (b). The images in (c), (d) and (e) are the results of unsupervised segmentation using the proposed features at one smoothing scale (c), three smoothing scales (d), and three smoothing scales with differential features (e). Figure 1(f) shows the unsupervised segmentation using the Laws features and fig. 1(g) shows the supervised segmentation using the GMRF model. Notice that the GMRF model, although applied in a supervised mode, performs worse than our features. In all cases, we applied an hierarchical segmentation algorithm (similar to the one in [11]) where the region label process is assumed to be a MRF.

5 Conclusions

A set or texture features based on the approximation of the image intensity field by a cubic B-spline surface is proposed here. These features can be computed at any degree of smoothness of the approximating surface. The performance of our set improves when differential and second order features are included. The proposed features provided better classification rates than the Laws, the Gabor and GMRF texture description methods. Also, they are particularly suitable for segmentation because they have a good discriminative power even for relatively small observation windows.

APPENDIX: Smoothing B-spline transform

Unser et al [9] showed that for a signal on N equidistant data points the smoothing cubic B-spline transform can be written as the product of of one causal and one anticausal filter:

$$\mathcal{S}_\lambda(z) = (1-a_1-a_2)^2 \frac{1}{1 - a_1 z^{-1} - a_2 z^{-2}} \frac{1}{1 - a_1 z^{1} - a_2 z^{2}} \tag{2}$$

where

$$a_1 = 2\rho\cos(\omega), \quad a_2 = -\rho^2$$

$$\rho = \left(\frac{24\lambda - 1 - \sqrt{\xi}}{24\lambda}\right)\left(\frac{48\lambda + 24\lambda\sqrt{3+144\lambda}}{\xi}\right)^{1/2}$$

$$\omega = \tan^{-1}\left(\sqrt{\frac{144\lambda - 1}{\xi}}\right), \xi = 1 - 96\lambda + 24\lambda\sqrt{3+144\lambda}$$

and $\lambda > 0$ is called the *smoothing* or *regularisation coefficient* which establishes a sort of compromise between approximation and smoothness. By employing the decomposed transfer function (2), the smoothing B-spline filter can be implemented recursively with as few as four multiplications and four additions per sample point. First the causal component is applied from left to right, and then the anticausal is applied from right to left. The extension to the 2D case is straightforward.

References

[1] D. Geman, S. Geman, C. Graffigne, and P. Dong. Unsupervised texture segmentation using Markov random field models. *IEEE Trans. Patt. Anal. and Machine Intell.*, vol. PAMI-12:609–628, 1990.

[2] R. M. Haralick and L. G. Shapiro. *Computer and Robot Vision*, volume 1. Addison-Wesley, 1992.

[3] A.K. Jain and F. Farrokhnia. Unsupervised texture segmentation using Gabor filters. *Patt. Recognition*, vol. 24:1167–1186, 1991.

[4] K.I. Laws. Goal-directed textured image segmentation. *Appl. Artif. Intell., SPIE*, vol. 548:19–26, 1985.

[5] B.S. Manjunath and R. Chellappa. Unsupervised texture segmentation using Markov random field models. *IEEE Trans. Patt. Anal. and Machine Intell.*, vol. PAMI-13:478–482, 1991.

[6] B.S. Manjunath, T. Simchony, and R. Chellappa. Stochastic and deterministic networks for texture segmentation. *IEEE Trans. Ac. Sp. Sign. Proc*, vol. ASSP-38:1039–1049, 1990.

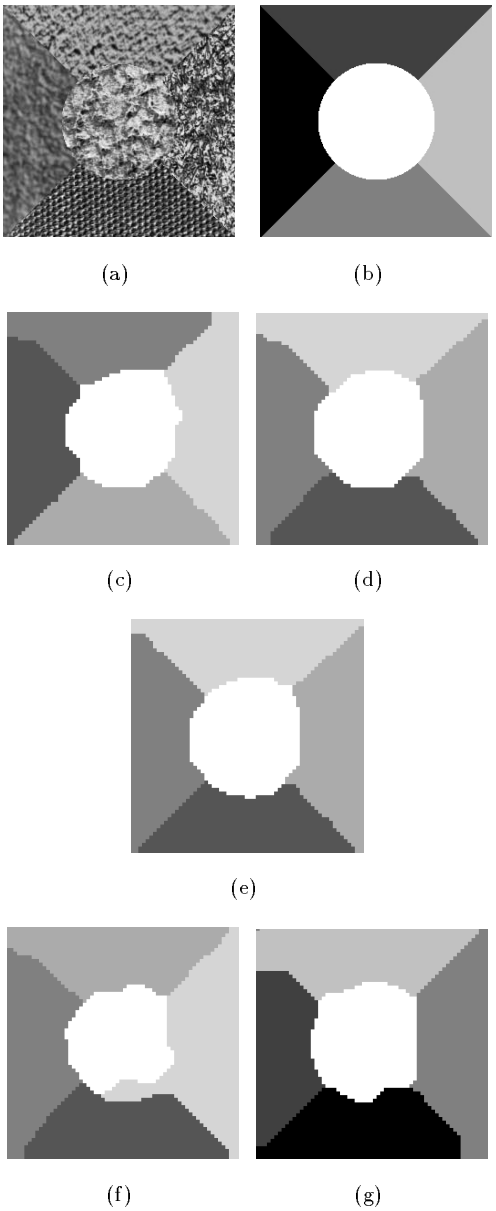
[7] A.R. Rao and G.L. Lohse. Identifying high level features of texture perception. *CVGIP: Grap. Models Im. Proc.*, vol. 55:218–233, 1993.

[8] H. Tamura, S. Mori, and T. Yamawaki. Textural features corresponding to visual perception. *IEEE Trans. Systems Man Cyber.*, vol. SMC-8:460–473, 1978.

[9] M. Unser, A. Aldroubi, and M. Eden. B-spline signal processing. *IEEE Trans. Sig. Proc.*, vol. 41:821–48, 1993.

[10] M. Unser and M. Eden. Multiresolution feature extraction and selection for texture segmentation. *IEEE Trans. Patt. Anal. and Machine Intell.*, vol. PAMI-11:717–728, 1989.

[11] C.S. Won and H. Derin. Unsupervised segmentation of noisy and textured images using Markov random fields. *CVGIP: Grap. Models Im. Proc.*, vol. 54:308–328, 1992.



(a) (b) (c) (d) (e) (f) (g)

Figure 1: Segmentation of a 5 region Brodatz collage.

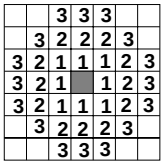


Figure 2: Neighborhood configuration for calculation of second order histograms.

proposed features, one scale

pattern	32x32	16x16	8x8
D17	0.937	0.895	0.768
D24	1.0	0.883	0.755
D4	1.0	0.973	0.917
D57	0.937	0.828	0.521
D77	1.0	0.980	0.923
D9	0.844	0.676	0.385
D92	0.984	0.809	0.544
average	0.957	0.863	0.689

(a)

prop. set- 1 scale + 2nd order

pattern	32x32	16x16	8x8
D17	1.0	0.973	0.800
D24	1.0	0.910	0.755
D4	1.0	0.980	0.920
D57	0.953	0.896	0.695
D77	1.0	0.992	0.966
D9	0.937	0.868	0.457
D92	0.906	0.723	0.623
average	0.970	0.906	0.745

(b)

proposed feat., three scales

pattern	32x32	16x16	8x8
D17	0.984	0.954	0.882
D24	0.953	0.895	0.756
D4	1.0	0.984	0.940
D57	0.984	0.887	0.643
D77	0.937	0.937	0.896
D9	0.765	0.602	0.371
D92	0.953	0.781	0.508
average	0.939	0.876	0.714

(c)

proposed set with diff. feat.

pattern	32x32	16x16	8x8
D17	1.0	0.961	0.868
D24	1.0	0.887	0.769
D4	1.0	0.988	0.950
D57	1.0	0.953	0.767
D77	1.0	0.969	0.921
D9	0.937	0.820	0.550
D92	1.0	0.949	0.720
average	0.991	0.932	0.792

(d)

Laws masks (5x5)

pattern	32x32	16x16	8x8
D17	0.891	0.867	0.663
D24	0.937	0.863	0.764
D4	0.984	0.887	0.695
D57	1.0	0.969	0.723
D77	0.891	0.914	0.860
D9	0.703	0.574	0.400
D92	0.844	0.582	0.266
average	0.893	0.808	0.619

(e)

Gabor (5 freq., 4 orient)

pattern	32x32	16x16	8x8
D17	1.0	0.953	0.889
D24	1.0	0.922	0.800
D4	0.937	0.843	0.688
D57	0.937	0.812	0.495
D77	0.875	0.910	0.874
D9	0.750	0.640	0.481
D92	0.812	0.500	0.261
average	0.902	0.797	0.641

(f)

GMRF model, 5th order

pattern	32x32	16x16	8x8
D17	1.0	1.0	0.932
D24	1.0	0.976	0.858
D4	1.0	0.964	0.956
D57	1.0	0.996	0.974
D77	1.0	0.973	0.895
D9	0.562	0.410	0.334
D92	0.859	0.703	0.563
average	0.917	0.860	0.787

(g)

Figure 3: Classification rates for a set of 7 Brodatz textures.