TEXTURE DISCRIMINATION USING MULTIMODAL WAVELET PACKET SUBBANDS

Roberto Cossu, Ian H. Jermyn, Josiane Zerubia

Ariana (CNRS/INRIA/UNSA joint research group)
INRIA, B. P. 93, 06902 Sophia Antipolis, France
Email: firstname.lastname@inria.fr

ABSTRACT

The subband histograms of wavelet packet bases adapted to individual texture classes often fail to display the leptokurtotic behaviour shown by the standard wavelet coefficients of 'natural' images. While many subband histograms remain leptokurtotic in adaptive bases, some subbands are Gaussian. Most interestingly, however, some subbands show multimodal behaviour, with no mode at zero. In this paper, we provide evidence for the existence of these multimodal subbands and show that they correspond to narrow frequency bands running throughout images of the texture. They are thus closely linked to the texture's structure. As such, they seem likely to possess superior descriptive and discriminative power as compared to unimodal subbands. We demonstrate this using both Brodatz and remote sensing images.

1. INTRODUCTION

The analysis of the statistics of wavelet coefficients has for the most part focused on 'natural images' and standard wavelets [1, 2, 3, 4, 5]. Such statistics are necessarily mixtures from a variety of sources, both because 'natural images' consist of regions corresponding to many different entities, and because standard wavelets mix frequency bands that may have very different behaviours individually. The statistics of coherent sets of images analysed with bases other than standard wavelet bases may therefore be very different from the leptokurtotic histograms found in such mixtures.

The subband histograms of wavelet packet bases adapted to coherent classes of texture [6, 7, 8] confirm this intuition. Brady *et al.* [6] use Gaussian models in which the covariance is assumed diagonal in at least one wavelet packet basis. This basis is learned from examples, and adapts to each texture modelled. Spatial dependencies are captured by the basis itself rather than explicitly as in hidden Markov tree models [4, 5]. The resulting subband histograms fall into three classes. Many subbands still show leptokurtotic behaviour. Others, usually those with a smaller frequency range, show Gaussian behaviour. Most interestingly, some of the subbands with the narrowest frequency content are

multimodal, with no mode at zero. Intuitively, the presence of a mode at zero indicates that the wavelet representation is sparse. This is true of 'natural' images, in which there are large regions that are relatively flat compared to the edges present, but it is no longer true when a phenomenon occurring throughout the images of a given class is focused on by a particular subband.

Although the multimodal subbands were discovered using Gaussian models, these are clearly unsuitable given the observed statistics. In [7, 8], a new model was developed incorporating the multimodal behaviour. This model is briefly described in section 2. It confirms and refines the results of the Gaussian models, and provides hints of even more complex and structured behaviour [7].

The aim of this paper is to demonstrate the existence and discriminatory power of multimodal subbands. First, in section 3, we provide theoretical and empirical evidence for the existence of multimodality and show that multimodality is strictly related to the main periodicities characterizing the texture's structure. Then, in section 4, we demonstrate the discriminatory power of multimodality using both Brodatz textures and textures taken from remote sensing imagery. We summarize in section 5.

2. MODEL FOR MULTIMODAL SUBBANDS

In [8], a model was developed to include the newly observed multimodal subbands. In this model, each subband is modelled by one of three distributions: Gaussian (G); generalized Gaussian (GG); or a constrained mixture of three Gaussians (MoG). The full model is parameterized by the following data: a dyadic partition of one quadrant of the Fourier domain, T, which, given a mother wavelet, defines a wavelet packet basis; a map μ from T to a set M of three models, $\{G, GG, MoG\}$, giving the model used in each subband; a map θ from T to the space of model parameters for each subband. The subbands in T are assumed independent.

The G and GG subband models are familiar. The MoG subband model is a mixture of three Gaussians indexed by $j \in \{0, 1, 2\}$, with means $\mu_{\alpha,j}$, variances $\sigma_{\alpha,j}^2$, and mixing probabilities $p_{\alpha,j}$. (α is the subband index.) These satisfy
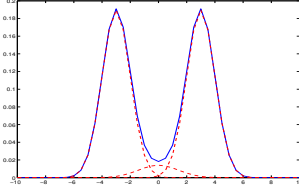
Fig. 1. An example of the MoG model used to model multimodal subbands.

Fig. 2. Some textures, their optimal decompositions (dark grey = GG, light grey = G, white = MoG), and examples of multimodal subband histograms and the models fitted to them.

the constraints: $\mu_{\alpha,0} = 0$; $\mu_{\alpha,1} = -\mu_{\alpha,2}$; $\sigma_{\alpha,1}^2 = \sigma_{\alpha,2}^2$; and $p_{\alpha,1} = p_{\alpha,2}$. These ensure that the mean of the distribution is zero. A example is shown in figure 1.

In [8], we describe how to find exact MAP estimates of T, μ, and θ from training data for a texture class, using an efficient depth-first search algorithm on the space of dyadic partitions. The resulting models capture the different statistics of the adaptive subbands, and in particular the multimodal statistics. Examples will be shown in section 4.

Note that the model is not limited to textures that shown multimodality: it is perfectly capable of modelling textures without them. When they are present however, it is in a position to take advantage of the extra information they provide.

3. EXISTENCE OF MULTIMODAL SUBBANDS

3.1. Theory

Multimodal subband histograms, or more precisely, subband histograms with no mode at zero, are closely connected to the presence of structures running throughout the images of a texture class. The basic cause can be indicated very briefly, with an idealized model.

Suppose that a (1d) texture class consists of all translations of a sinusoid with fixed frequency k_0 and amplitude B. The mean histogram of a subband under a translation invariant distribution is equal to the marginal distribution of any wavelet packet coefficient from that subband. We will suppose, wlog, that the Fourier amplitude of such a wavelet packet coefficient at frequency k_0 is one. Then it is easy to see that the distribution of this coefficient takes the bimodal form

$$H(\omega) = \begin{cases} \frac{1}{\pi}(B^2 - \omega^2)^{-1/2} & \text{if } -B \leq \omega \leq B \\ 0 & \text{otherwise} \end{cases} . \quad (1)$$

Intuitively this arises from the fact that in a sinusoid of amplitude B, the values $\pm B$ occur 'more often' that other values (including zero).

As is demonstrated in [7], more complex models, involving several sinusoids with noise added to the amplitudes, blur but do not destroy the bimodality. These models demonstrate how periodicities in the texture can produce multi-
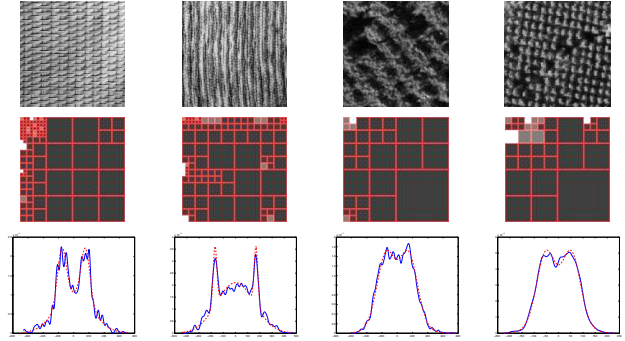
modal subband histograms if the wavelet packet basis is sufficiently focused on the frequencies involved.

3.2. Experiment

Figure 2 shows the results of training the model on four textures, two from the Brodatz album, and two from aerial remote sensing images. The second row shows T and μ in graphical form, the lines indicating the dyadic partition, and the colours in each subband indicating the model assigned to that subband, dark grey being GG, light grey G, and white MoG. As can be seen, the multimodal subbands are closely connected to the periodicities present in the textures. The third row shows the histogram of a selected multimodal subband (blue/full line) and the model fitted to it (red/dashed). The double peaks are not as clear in the remote sensing images. This is partly due to the patches in the two images where there are no trees. If training is performed on samples without these 'holes', a more pronounced bimodality is found.

Further experiments show the existence of multimodal subbands in many Brodatz textures, and in remote sensing textures such as ploughed fields, planted forests, and certain configurations of buildings.

4. DISCRIMINATION AND DESCRIPTION

In all the experiments reported below, the following procedure was used. First, a model was trained for each of the two classes involved. Second, a subband was selected from one of the models, according to criteria that will be detailed in a moment. Call this subband S. A new model of subband S was then trained for the second class. Third, the undecimated wavelet packet coefficients in S were calculated for the test image. Fourth, for each pixel p, class probabilities
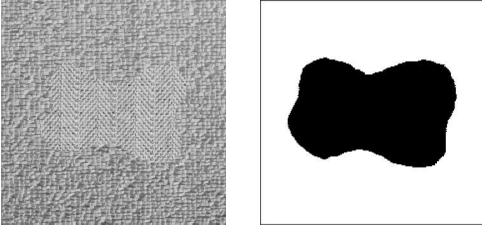
Fig. 3. On the left is a mosaic formed of the Herring and Raffia Brodatz textures. On the right, the result of classification using one subband multimodal for Raffia, but unimodal for Herring.

Fig. 4. On the left, the reconstruction of the mosaic in figure 3 using a single unimodal subband. On the right, the reconstruction of the mosaic using a single multimodal subband.

Fig. 5. On the left, a remote sensing image. On the right, the probability map resulting from the use of one subband multimodal for the 'ploughed field' texture, but unimodal for the 'unploughed field' texture.

were computed from the models of subband S only, using the coefficients in S belonging to a patch centred on p, of size equal to the subband filter size of S. Thus only the data lying in S was used to compute the probabilities. The differences of the log probabilities of each class form the 'probability map'.

Note that these experiments are not designed to produce the best possible classification results. The full model (*i.e.* using all the subbands) always produces a better result. What is remarkable is that the results of the above procedure using a subband that is multimodal for one texture but not for the other are often very good—indeed for Brodatz textures they approach the performance of the full model—whereas the results using a subband that is unimodal for both textures are very poor.

Figure 3 shows a mosaic made of the Herring and Raffia textures from the Brodatz album. On the right is the result of applying the above procedure using a subband multimodal for Raffia, but unimodal for Herring. Similar results were obtained for other multimodal subbands. The same procedure was then repeated using several subbands of the same size, but unimodal for both textures. The results were all very poor, bearing very little if any connection to the mosaic. Indeed, in many cases the whole image was classified as belonging to one class. The average misclassification rates were as follows. Using Raffia multimodal subbands, the average pixel misclassification rate of Raffia was 14.9%, and of Herring was 0.7%. Using Herring multimodal subbands, Raffia was misclassified 5.1%, and Herring at 17.7%. Using unimodal subbands, the misclassification rate for Herring was 80% and for Raffia 6.5%. The last number is low for the reason that has already been mentioned: most of the unimodal results classified most of the image as Raffia.

Figure 4 shows reconstructions of the mosaic in figure 3 using one of the unimodal subbands (on the left) and one of the multimodal subbands (on the right). The left-hand image shows almost no trace of the mosaic structure, which accounts for the poor cl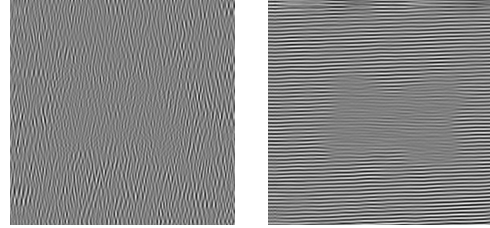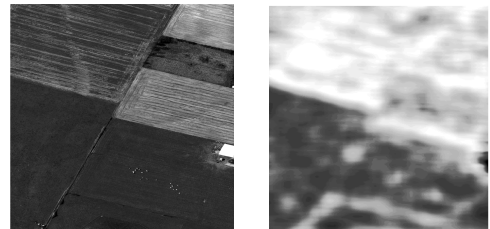assification result. In contrast, the image reconstructed from the multimodal subband clearly shows the mosaic structure, the amplitude of the periodicity being much larger in Raffia than in Herring.

Figure 5 shows a remote sensing image. Models were trained on 'ploughed field' using the upper part of the image, and on 'unploughed field' using the lower part. The above procedure was then followed for a subband multimodal for 'ploughed field' but unimodal for 'unploughed field'. The resulting probability map is shown on the right. The map indicates the ability of the multimodal subband to provide important information for discriminating between the two textures, and in particular to assign a consistently low probability to 'unploughed field'.

Figure 6 shows a second remote sensing image. Models were trained on the 'forest' texture in the lower part of the image, and on a 'background' class consisting of the top-left hand corner. On the right of the figure is shown the result of reconstructing the image using just the multimodal subband in the 'forest' texture. Note how the tree structure of the 'forest' texture has been captured, and how this same structure captures the presence of other trees in the image, while other areas show small response (the image has been linearly stretched for visualization). An exception is the bottom left-hand corner, where the arrangement and size of the buildings represents roughly the same periodicity
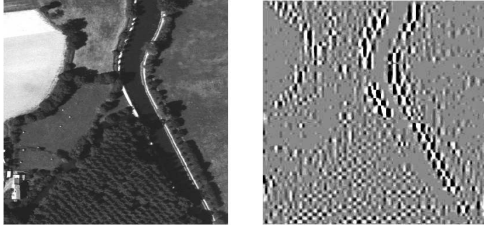
Fig. 6. On the left, a remote sensing image. On the right, the reconstruction using a subband multimodal for the 'forest' texture in the lower part of the image.
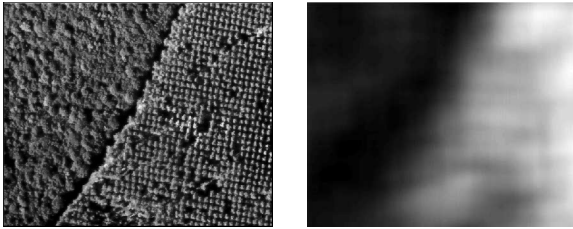

Fig. 7. On the left, a remote sensing image. On the right, the probability map resulting from the use of one subband multimodal for the 'poplar stand' texture on the right and unimodal for the 'forest' texture on the left.

as the forest. Note that the subband used here is a standard wavelet subband. Analyzed using a Gaussian or generalized Gaussian model, this subband is not at all remarkable. Nevertheless, the new model was able to detect the multimodality in this subband, and hence capture the structure.

Figure 7 shows another remote sensing image. The classes were 'poplar stand', on the right of the image, and 'forest' on the left. On the right is shown the probability map resulting from the use of a subband multimodal for 'poplar stand', but unimodal for 'forest'. Note again the consistent assignment of a very low value to 'forest', and a much higher value generally to 'poplar stand'.

5. CONCLUSION

In this paper, we have provided evidence for the existence of multimodal wavelet packet subbands in textures, and have demonstrated their link to the characteristic structure of a texture. We have also demonstrated the descriptive and discriminative power of the model of these multimodal subbands developed in [8]. The new multimodal statistics are only revealed when attention is shifted away from the universal behaviour displayed by the standard wavelet coefficients of whole images, which are necessarily mixtures of many components, towards the behaviour shown by the

coefficients of bases adapted to individual components of these mixtures.

Practically speaking, while the classification maps obtained for real remote sensing images are not as accurate as for synthetic images, the probability maps provide useful information. For example, in cases where the classes are spectrally overlapped or otherwise poorly separable, a complex classification system could jointly exploit multispectral information and the information contained in this probability map to obtain a more accurate classification result.

References

[1] S. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 11, pp. 674–693, 1989.

[2] E. P. Simoncelli and E. H. Adelson, "Noise removal via Bayesian wavelet coring," in *Proc. IEEE ICIP*, Lausanne, Switzerland, September 1996, pp. 379–382.

[3] P. Moulin and J. Liu, "Analysis of multiresolution image denoising schemes using generalized-Gaussian and complexity priors," *IEEE Trans. Info. Th.*, vol. 45, no. 3, pp. 909–919, 1999.

[4] M. Crouse, R. Nowak, and R. Baraniuk, "Wavelet-based statistical signal processing using hidden markov models," *IEEE Trans. Signal Processing*, vol. 46, no. 4, pp. 886–902, 1998.

[5] M. J. Wainwright and E. P. Simoncelli, "Scale mixtures of Gaussians and the statistics of natural images," in *Advances in Neural Information Processing Systems*, S. A. Solla, T. K. Leen, and K.-R. Müller, Eds., Cambridge, MA, 2000, vol. 12, pp. 855–861, MIT Press.

[6] K. Brady, I. H. Jermyn, and J. Zerubia, "Adaptive probabilistic models of wavelet packets for the analysis and segmentation of textured remote sensing images," in *Proc. Brit. Mach. Vis. Conf.*, Norwich, U. K., September 2003.

[7] R. Cossu, I. H. Jermyn, K. Brady, and J. Zerubia, "Probabilistic models of the unimodal and multimodal statistics of adaptive wavelet packet coefficients," Research report, INRIA, February 2003, To appear.

[8] R. Cossu, I. H. Jermyn, and J. Zerubia, "Texture analysis using probabilistic models of the unimodal and multimodal statistics of adaptive wavelet packet coefficients," in *Proc. IEEE ICASSP*, Montreal, Canada, May 2004.