

TEXTURE ANALYSIS USING PROBABILISTIC MODELS OF THE UNIMODAL AND MULTIMODAL STATISTICS OF ADAPTIVE WAVELET PACKET COEFFICIENTS

Roberto Cossu, Ian H. Jermyn, Josiane Zerubia

Ariana (joint research group of CNRS/INRIA/UNSA)
INRIA, 2004 route des Lucioles, B. P. 93
06902 Sophia Antipolis, France
Email: `firstname.lastname@inria.fr`

ABSTRACT

Although subband histograms of the wavelet coefficients of natural images possess a characteristic leptokurtotic form, this is no longer true for wavelet packet bases adapted to a given texture. Instead, three types of subband statistics are observed: Gaussian, leptokurtotic, and interestingly, in some subbands, multimodal histograms. These subbands are closely linked to the structure of the texture, and guarantee that the most probable image is not flat. Motivated by these observations, we propose a probabilistic model that takes them into account. Adaptive wavelet packet subbands are modelled as Gaussian, generalized Gaussian, or a constrained Gaussian mixture. We use a Bayesian methodology, finding MAP estimates for the adaptive basis, for subband model selection, and for subband model parameters. Results confirm the effectiveness of the proposed approach, and highlight the importance of multimodal subbands for texture discrimination and modelling.

1. INTRODUCTION

Texture is important in many image processing applications, and wavelet packets have played an important role in its description and analysis [1]. In addition to the use of fixed bases, several methods have been proposed for selecting the ‘best’ wavelet packet basis for a texture, but these have been *ad hoc* both in the criterion used to define ‘best’, and in the ‘distance’ used to distinguish textures. In particular, no underlying texture model has been used. Brady *et al.* [2] address this problem within a Bayesian framework. For each texture class, they assume a Gaussian distribution for images of that class, with inverse covariance lying in the large set of operators diagonal in some wavelet packet basis. An exact MAP estimate of the covariance is then found, which has the side effect of specifying a preferred wavelet packet basis for each class, in which the covariance is diagonal.

The Gaussian ansatz appears to go against the observation that in standard wavelet bases, the subband histograms assume a leptokurtotic form [3, 4]. However, the latter can no longer be assumed to hold for coherent textures as opposed to natural images, or for different bases. In this situation, a Gaussian is a useful minimal assumption. The failure of leptokurtosis for adaptive bases and coherent textures was confirmed in [5]. Although many of the adaptive subbands are leptokurtotic, many of the remainder appear Gaussian. Most interestingly, in some subbands a completely different behaviour is observed: the subband histograms are bimodal. (Examples are shown in figure 2.) These subbands are usually narrow in frequency content; the presence in the subband histograms of maxima at non-zero coefficient values thus indicates the likely presence of approximate periodicities running throughout images of that texture class, with frequencies in the support of these subbands. The importance of this fact for texture modelling is clear. In the absence of such phenomena, *i.e.* if all histogram modes are at zero, the most probably image of that texture class is flat, and thus ‘untextured’. This conflicts strongly with our prior knowledge: the most probably images of raffia, or of forest in a remote sensing image for example, are not flat.

Motivated by this empirical evidence, we extend the approach proposed in [2] by relaxing the Gaussian assumption to consider three possible models for each adaptive subband: Gaussian, generalized Gaussian, and a constrained mixture of three Gaussians. A Bayesian methodology is used to compute MAP estimates for the adaptive basis, for the subband models, and for the subband model parameters. Results confirm the existence of multimodal subbands, demonstrate the accuracy of the modelling, and highlight the importance of the new multimodal statistics and models for texture analysis and description.

The paper is organized as follows. Section 2 describes the models and how we estimate their parameters. Section 3 describes how we select the model for each subband. The MAP estimate of the adaptive basis is addressed in sec-

This work was partially funded by EU project Mournir, HP-99-108, www.mournir.org.

tion 4. Experimental results are reported in section 5. In section 6 we discuss the results, and draw conclusions.

2. MODELS FOR DESCRIBING SUBBAND DISTRIBUTIONS

While the adaptive wavelet packet bases in [5] were intimately connected to the Gaussian assumption, it is clear that the usefulness of the adaptive wavelet packet decomposition extends further. It allows the analysis to focus on important regions of the Fourier domain, while maintaining spatial localization; and adaptivity means that it is reasonable to assume that the different subbands so identified are somewhat independent. In any case, that is the assumption we will make here. (Note that this does not mean that the standard wavelet coefficients are independent. Indeed one interpretation of the adaptive wavelet packet models is as a way of encoding long-range dependencies between standard wavelet coefficients.)

The distributions that we will consider as models for a given texture class are defined 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; for each subband in T , one of three models: Gaussian (G), generalized Gaussian (GG), or a constrained mixture of three Gaussians (MoG) to model the multimodal statistics; and for each subband, the appropriate model parameters. Below we describe the estimation procedures for GG and MoG, the case of G being familiar.

2.1. Generalized Gaussian Model

The generalized Gaussian distribution for a subband is

$$\Pr(w_\alpha | f_\alpha, s_\alpha) = Z_\alpha^{-1} e^{-f_\alpha \sum_{i \in \alpha} |w_{\alpha,i}|^{s_\alpha}}, \quad (1)$$

where α is the index of the subband, and i indexes the coefficients within that subband; $w_{\alpha,i}$ is a wavelet packet coefficient, and w_α is the set of all coefficients in subband α ; s_α is called the shape factor, while f_α controls the width of the distribution; and Z_α is a normalization constant that depends on f_α , s_α , and the size of the subband. To estimate the parameters of the generalized Gaussian model, we use invariant MAP estimation [6] with Jeffreys' prior, which is equivalent to maximum likelihood. This gives the equations

$$N_\alpha - s_\alpha f_\alpha \sum_{i \in \alpha} |w_{\alpha,i}|^{s_\alpha} = 0 \quad (2)$$

$$N_\alpha \ln f_\alpha - N_\alpha \Psi \left(1 + \frac{1}{s_\alpha} \right) + s_\alpha^2 f_\alpha \sum_{i \in \alpha} |w_{\alpha,i}|^{s_\alpha} \ln |w_{\alpha,i}| = 0, \quad (3)$$

where Ψ is the digamma function. We solve equations (2) and (3) numerically for f_α and s_α using the algorithm in [4].

2.2. Constrained Mixture of Gaussians

We model subbands showing multimodal behaviour with a constrained mixture of three Gaussians. One of the components has mean $\mu_{\alpha,0} = 0$, variance $\sigma_{\alpha,0}^2$, and prior probability $p_{\alpha,0}$ in the mixture. The other two components have means $\mu_{\alpha,1} = -\mu_{\alpha,2}$, variances $\sigma_{\alpha,1}^2 = \sigma_{\alpha,2}^2$, and prior probabilities $p_{\alpha,1} = p_{\alpha,2}$. The constraints ensure that the mean of the distribution is zero. The mixture of Gaussians is given by

$$\Pr(w_\alpha | p_{\alpha,j}, \mu_{\alpha,j}, \sigma_{\alpha,j}) = \prod_{i \in \alpha} \left[\sum_{j=0}^2 p_{\alpha,j} (2\pi\sigma_{\alpha,j}^2)^{-1/2} e^{-\frac{1}{2} \left(\frac{w_{\alpha,j} - \mu_{\alpha,j}}{\sigma_{\alpha,j}} \right)^2} \right].$$

The invariant MAP estimates of the model parameters are computed using the EM algorithm [7].

3. MODEL SELECTION

To perform model selection for a given subband α in decomposition T , we again use MAP estimation. Let θ_α denote the set of parameters for the model $M_\alpha \in \{G, GG, MoG\}$ for subband α . The probability of M_α and θ_α is given by

$$\Pr(\theta_\alpha, M_\alpha | w_\alpha) \propto \Pr(w_\alpha | \theta_\alpha) \Pr(\theta_\alpha | M_\alpha) \Pr(M_\alpha). \quad (4)$$

We use Jeffreys' prior for the θ_α . Under suitable assumptions about the range of each parameter, invariant MAP estimation is equivalent to setting $\Pr(\theta_\alpha | M_\alpha) = e^{-d_m \ln A}$ in equation (4), where d_m is the dimensionality of the model. That is, it is as if the parameter distributions were all uniform on some (large) interval A . Note that this term penalizes model complexity. A is not determined automatically, but the result of model selection is not very sensitive to A once an order of magnitude has been set. We suppose that $\Pr(M_\alpha) = 1/3$, *i.e.* that all models are equally likely *a priori*. The probability is maximized by first maximizing over θ_α for each model M_α , as already described, and then picking the model with the highest value of the above probability.

4. TREE DECOMPOSITION

Dyadic partitions are in bijective correspondence with quad-trees, so to estimate the adaptive wavelet packet basis we use a depth-first search optimal subtree algorithm, a form of which was also used in [8]. This is optimal for weighted trees for which the optimal weight of a subtree is independent of the rest of the tree. The probability of the adaptive

basis, the model, and its parameters is

$$\begin{aligned} \Pr(\Theta, T|w) &\propto \Pr(w|\Theta, T)\Pr(\Theta|T)\Pr(T) \\ &= \Pr(T) \prod_{\alpha \in T} \Pr(w_\alpha|\Theta_\alpha, T)\Pr(\Theta_\alpha|T) \quad , \quad (5) \end{aligned}$$

where w is the training data, *i.e.* all wavelet packet coefficients; Θ denotes the set of models and parameters for all subbands; and Θ_α denotes those for subband α . We choose the prior probability of a basis to be

$$\Pr(T) = B^{-1} e^{-\beta|T|} = B^{-1} \prod_{\alpha \in T} e^{-\beta} \quad ,$$

where $|T|$ is the number of subbands in the basis, and B is a normalization constant. This prior penalizes bases with many subbands, and effectively regularizes the functions θ_α viewed as functions of frequency. With this choice, the logarithm of equation (5) satisfies the criterion for the optimal subtree algorithm, where the weight of a leaf vertex is given by equation (4) evaluated at the MAP estimates. We thus find exact MAP estimates for Θ and T .

5. EXPERIMENTAL ANALYSIS

We applied the above modelling framework to a number of textures in the Brodatz album [9]. The results reported here are for Raffia and Herring, shown in figure 1, but similar results were obtained for other textures.

The optimal decompositions in figure 1(b) and (d) were obtained with $A = 5$, $\beta = 300$, which produced reasonable results for all the textures that we analysed. As expected, the multimodal subbands focus on frequencies related to the structure in the texture. Figure 2 plots the wavelet packet coefficient histograms and the fitted models for some subbands, showing that the model selection and parameter estimation procedures work well. The interested reader can find further results on Brodatz textures, as well as results on texture classes taken from remote sensing images, in [10].

5.1. Multimodal subbands and texture discrimination

For all unimodal subbands, the most probable value of the wavelet packet coefficients is zero, *i.e.* the most probable image composed of these subbands is untextured. In contrast, even though the multimodal subbands have zero mean, the most probable value for the coefficients is non-zero, and therefore textured. Since the multimodal subbands typically have narrow frequency support, we can think of them as capturing the principal periodicities in the texture, in a manner analogous to the Wold decomposition [11]. This is made clear by the positions of the MoG subbands in the decomposition in figure 1(b). Because of their close relation to structure, multimodal subbands provide a powerful tool for texture discrimination as well as description.

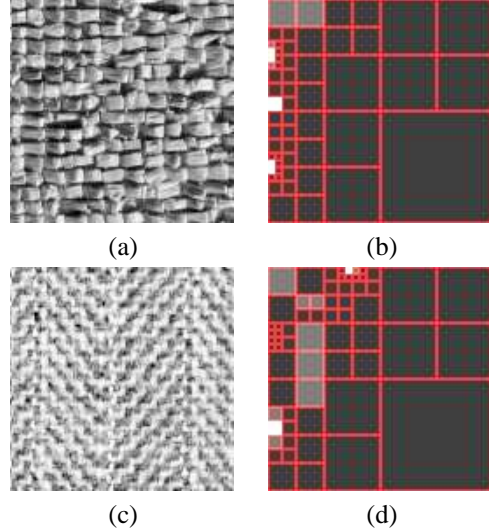


Fig. 1. a) Raffia; b) its decomposition; (c) Herring; (d) its decomposition. Colour denotes the model for each subband: black, GG; grey, G; white, MoG.

Consider again the Raffia (R) and Herring (H) textures. Let $S_i : i \in \{R, H\}$ denote the set of multimodal subbands for each texture. Let $W_{i,j}$ denote the set of $|S_i|$ -dimensional vectors of undecimated wavelet packet coefficients from the subbands S_i , computed from texture j . For fixed i , $W_{i,R}$ and $W_{i,H}$ can be plotted as differently coloured points in $|S_i|$ dimensions. The first row of figure 3 shows projections from such a plot onto two pairs of subbands from S_R . Red/dark patterns correspond to coefficients from Raffia; green/light patterns correspond to coefficients from Herring. As can be seen, the two textures are well separated in this space, with the coefficients of Raffia forming a ‘ring’ around the coefficients of Herring. For the sake of comparison, the second row of figure 3 shows similar plots of the wavelet packet coefficients from unimodal subbands. The figure shows clearly that for these subbands, the two textures are much less easily separable.

6. DISCUSSION AND CONCLUSION

The model developed herein can be seen in a number of different ways. Viewed as a modelling tool, adaptive wavelet packet bases allow the description of long-range dependencies amongst standard wavelet coefficients in a given subband, while maintaining the advantages of independent subbands. In this sense, they are complimentary to, for example, hidden Markov tree models, which capture well dependencies across scale. The modelling of the multimodal subbands means that the distributions no longer have the peculiar property that the most probable image is untextured.

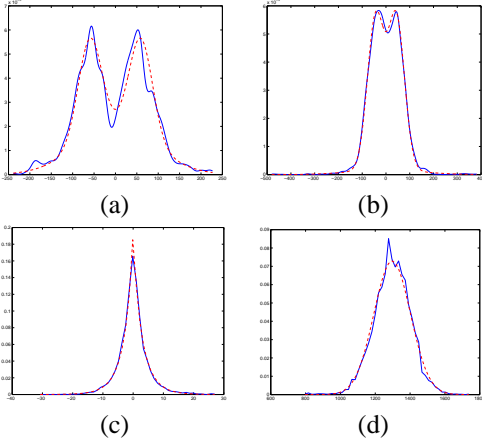


Fig. 2. Subband histograms (solid blue) and fitted models (dashed red) for four subbands from figure 1(b).

Because they are better models, it seems likely that their use will enable performance improvements in applications such as classification and denoising.

Viewed as an analysis tool, the models developed here can be seen as part of a process of separating mixture distributions into components that are closer to physical quantities. The models do this in two ways. First, they focus on classes of image that can be regarded as coherent textures. Second, they separate the multimodal distributions of certain wavelet packet subbands from the generic leptokurtic distributions of standard wavelet subbands. The benefits of performing this type of ‘microscopy’ have already been demonstrated by the fact that the Gaussian models of [2] led to the identification of the multimodal subbands. The degree of structure that can be observed in figure 3 indicates that the new models described here may well lead to other novel insights.

7. REFERENCES

[1] S. Livens, P. Scheunders, G. Van de Wouwer, and D. Van Dyck, “Wavelets for texture analysis, an overview,” in *Proc. IEE Int’l Conf. Im. Proc. and its Applications*, Dublin, Ireland, July 1997, pp. 581–585.

[2] K. Brady, I. H. Jermyn, and J. Zerubia, “Texture analysis: An adaptive probabilistic approach,” in *Proc. IEEE Int’l Conf. Image Processing*, Barcelona, Spain, September 2003.

[3] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, 2nd edition, 1999.

[4] M. N. Do and M. Vetterli, “Wavelet-based texture retrieval using generalized Gaussian density and

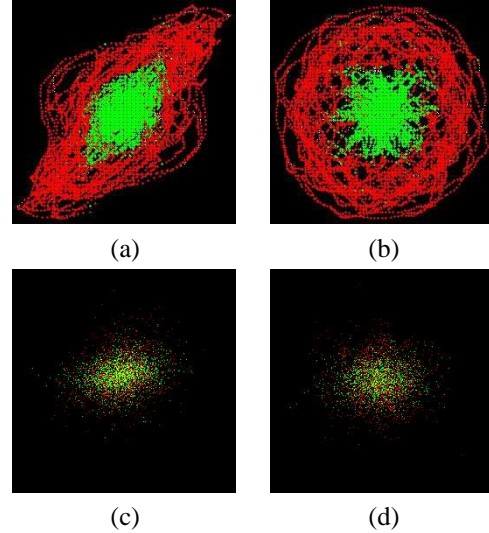


Fig. 3. Scatterplot of wavelet packet coefficients of Raffia and Herring, taken from pairs of subbands. In row 1, both subbands are multimodal for Raffia but unimodal for Herring, while in row 2, the subbands are unimodal for both textures. Red/dark corresponds to Raffia; green/light corresponds to Herring.

Kullback-Liebler distance,” *IEEE Trans. Image Processing*, vol. 11, pp. 146–158, 2002, 2.

[5] K. Brady, I. Jermyn, and J. Zerubia, “A probabilistic framework for adaptive texture description,” Research Report 4920, INRIA, Sept. 2003.

[6] I. H. Jermyn, “On Bayesian estimation in manifolds,” Research report 4607, INRIA, November 2002.

[7] A. P. Dempster, N. M. Laird, and D. B. Rubin, “Maximum likelihood from incomplete data via the EM algorithm,” *J. Roy. Stat. Soc.*, vol. 39, pp. 1–38, 1977.

[8] R. R. Coifman and M. V. Wickerhauser, “Entropy-based methods for best basis selection,” *IEEE Trans. Info. Th.*, vol. 38, no. 2, pp. 713–718, 1992.

[9] P. Brodatz, *Textures: A Photographic Album for Artists and Designers*, Dover, New York, USA, 1966.

[10] 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.

[11] J. M. Francos, A. Z. Meiri, and B. Porat, “A unified texture model based on a 2d Wold-like decomposition,” *IEEE Trans. Signal Processing*, vol. 41, pp. 2665–2678, 1993.