THE TWO-DIMENSIONAL WOLD DECOMPOSITION FOR SEGMENTATION AND INDEXING IN IMAGE LIBRARIES

Radu Stoica   Josiane Zerubia *

INRIA Sophia Antipolis
2004, route de Lucioles, BP 93
06902 Sophia Antipolis Cedex, France

Joseph M. Francos †

Dept. Elec. & Comp. Eng.
Ben-Gurion University
Beer-Sheva 84105, Israel

ABSTRACT

This paper presents a method for indexing and retrieval of multimedia data through texture segmentation, using the Wold decomposition. The texture field is assumed to be a realisation of a regular homogeneous random field. On the basis of a 2-D Wold-like decomposition, the field is represented as the sum of a purely indeterministic component, a harmonic component and a countable number of evanescent fields. A new rigourous distance measure between textures is derived, using Wold parameters. Adopting the MRF framework, we construct a segmentation procedure using the Wold parameters.

1. INTRODUCTION

The access to digital image information is becoming an integral part of many multimedia applications today. Efficient tools for storage, search and navigation in multimedia liberaries are essential [2, 8]. In this paper we address one aspect of this complex problem, namely, the development of algorithms for indexing (labeling) and retrieval of multimedia data, based on the properties of the imagery components of the stored data record. Indexing and retrieval of the data are performed using parametric modeling techniques of the imagery data.

2. THE TEXTURE MODEL

The presented texture model is based on the results of the Wold-type decomposition of 2-D regular and homogeneous random fields, [3]. Let $\{y(n,m),(n,m)\in\mathcal{Z}^2\}$, be a real valued, regular, and homogeneous random field. Then $y(n,m)$ can be uniquely represented by the orthogonal decomposition

$$y(n,m) = w(n,m) + v(n,m) \qquad (1)$$

* This work was supported in part by AFIRST (Association Franco-Israelienne pour Recherche Scientifique et Technique) Grant 960 MRAI 01

† This work was supported in part by the Israel Ministry of Science and the French Ministry of Research and Technology under Grant 8814196

The field $\{w(n,m)\}$ is purely indeterministic and has a unique white innovation driven moving average representation. The field $\{v(n,m)\}$ is a deterministic random field. It can be shown that it is possible to exhibit a family of NSHP total-order definitions such that the boundary line of the NSHP is of rational slope. Let α,β be two coprime integers, such that $\alpha\neq 0$. The angle θ of the slope is given by $\tan\,\theta=\beta/\alpha$. Each of these supports is called *rational non-symmetrical half-plane* (RNSHP). We denote by O the set of all possible RNSHP definitions on the 2-D lattice, (i.e. the set of all NSHP definitions in which the boundary line of the NSHP is of rational slope). The introduction of the family of RNSHP total-ordering definitions results in the following countably infinite orthogonal decomposition of the field's deterministic component, [3]:

$$v(n,m) = p(n,m) + \sum_{(\alpha,\beta)\in O} e_{(\alpha,\beta)}(n,m) \qquad (2)$$

The random field $\{p(n,m)\}$ is a *half-plane deterministic* field, and $\{e_{(\alpha,\beta)}(n,m)\}$ is the evanescent field corresponding to the RNSHP total-ordering definition $(\alpha,\beta)\in O$. Since for practical applications we can exclude singular-continuous spectral distribution functions from the framework of our treatment, a model for the evanescent field which corresponds to the RNSHP defined by $(\alpha,\beta)\in O$ is given by

$$e_{(\alpha,\beta)}(n,m) = \sum_{i=1}^{I^{(\alpha,\beta)}} s_i^{(\alpha,\beta)}(n\alpha-m\beta)\cos\left(2\pi\frac{\nu_i^{(\alpha,\beta)}}{\alpha^2+\beta^2}(n\beta+m\alpha)\right) + t_i^{(\alpha,\beta)}(n\alpha-m\beta)\sin\left(2\pi\frac{\nu_i^{(\alpha,\beta)}}{\alpha^2+\beta^2}(n\beta+m\alpha)\right) \qquad (3)$$

where the 1-D purely-indeterministic processes $\{s_i^{(\alpha,\beta)}(n\alpha-m\beta)\}$, $\{s_j^{(\alpha,\beta)}(n\alpha-m\beta)\}$, $\{t_k^{(\alpha,\beta)}(n\alpha-m\beta)\}$, $\{t_\ell^{(\alpha,\beta)}(n\alpha-m\beta)\}$ are mutually orthogonal for all $i,j,k,\ell, i\neq j, k\neq\ell$, and for all i the processes $\{s_i^{(\alpha,\beta)}(n\alpha-m\beta)\}$ and $\{t_i^{(\alpha,\beta)}(n\alpha-$

$m\beta)\}$ have an identical autocorrelation function. Hence, the "spectral density function" of each evanescent field has the form of a countable sum of 1-D delta functions which are supported on lines of rational slope in the 2-D spectral domain. In the following, we assume that the modulating 1-D processes $\{s_i^{(\alpha,\beta)}(n^{(\alpha,\beta)})\}$ and $\{t_i^{(\alpha,\beta)}(n^{(\alpha,\beta)})\}$ of each evanescent field can be modeled by a finite order AR model.

In this paper, we also assume that the half-plane deterministic field consists only of the harmonic random field :

$$h(n,m) = \sum_{p=1}^{P} C_p \cos 2\pi(n\omega_p + m\nu_p) + D_p \sin 2\pi(n\omega_p + m\nu_p) \quad (4)$$

In general, P is infinite. This component generates the 2-D delta functions of the "spectral density". The parametric modeling of deterministic random fields whose spectral measures are concentrated on curves other than lines of rational slope, or discrete points in the frequency plane, is still an open question to the best of our knowledge.

As stated earlier, the most general model for the purely indeterministic component $w(n,m)$ is the MA model. However, if its spectral density function is strictly positive on the unit bicircle and analytic in some neighborhood of it, a 2-D AR representation for the purely indeterministic field exists as well [4]. In the following, we assume that the above requirements are satisfied. Hence the purely indeterministic component's *autoregressive model* is given by

$$w(n,m) = -\sum_{(0,0)\prec(k,\ell)} b(k,\ell)w(n-k,m-\ell) + u(n,m) \quad (5)$$

where $\{u(n,m)\}$ is the 2-D white innovations field, whose variance is σ^2. In the practical estimation problem, the model support is assumed finite.

Hence, the observed texture field $\{y(n,m)\}$ is uniquely represented by the orthogonal decomposition $y(n,m) = w(n,m) + h(n,m) + \sum_{(\alpha,\beta)\in O} e_{(\alpha,\beta)}(n,m)$. Thus, the problem of estimating the texture model parameters, becomes one of estimating the parameters of the harmonic and evanescent components of the field in the presence of an unknown colored noise generated by the purely-indeterministic component, jointly with estimating the purely-indeterministic component parameters.

3. DISTANCE MEASURES FOR TEXTURES

In previous papers [5, 6] we showed that assuming the texture field is a realization of a Gaussian random field with mixed spectral distribution, essentially indistinguishable replicas of the original texture are synthesized from the estimated parameters. We therefore adopt the Gaussian assumption for the derivation of a distance measure, as well as for the image segmentation application. In this framework it is assumed that each observed texture is a realization of a Gaussian random field whose mean is related to the deterministic component of the Wold decomposition and the structure of its covariance matrix is determined by the parameters of the purely indeterministic component.

More specifically, in this work we have chosen the Kullback distance [7] between two Gaussian models of textures

$$\begin{aligned} D(P_1,P2) = {} & \tfrac{1}{2}(\mu_2-\mu_1)^T(\textstyle\sum_1^{-1}+\sum_2^{-1})(\mu_2-\mu_1) \\ & +\tfrac{1}{2}\mathrm{tr}(\textstyle\sum_1^{-1}\sum_2+\sum_1\sum_2^{-1}-2I) \end{aligned} \quad (6)$$

Clearly, one is interested in being able to define a distance measure, wich is invariant to translation, rotation, and scaling. Since the observed textures are homogeneous, translation affects only the mean function, but not the covariance. Scaling and rotation affect both the mean and covariance. To achieve the desired invariance, we have implemented the following procedure:

- Estimate the parametric model of each texture patch introduced to the system.
- Assuming that in each texture the deterministic component has only harmonic components, or only evanescent components, but not both (which is the case, in practice), every texture we inspect is first rotated so that its dominant harmonic component is aligned with a predefined direction (the x-axis, for example). Similar procedure is applied in the presence of evanescent components.
- When scaling is involved, the ratio of the dominant harmonic (or evanescent) frequencies of the two textures being considered will not be one. Subsampling the texture whose dominant frequency is lower, we obtain two textures with dominant components of identical frequencies.
- Assuming the phase of the texture dominant component is not zero, we crop a sub-picture of the original in which the phase of the dominant component is zero.
- Re-estimate the texture parameters to find the parameters of all the model components.

The next step is to construct the mean and the covariance matrix of the Gaussian field,using the Wold parameters, in order to use the Kulback distance. The construction of the mean was done, using the following procedure :

- Using the deterministic parameters, we are synthetising the mean image.
- Arranging this image in a column vector, it will give the mean vector of the Gaussian field.

The construction of the covariance matrix is as follows :

- Using only the AR parameters, we are building the power sepctral density estimate.

$$S(\omega,\nu) = \frac{\sigma^2}{|\sum_{(0,0)\leq(k,l)} b(k,l)\exp[2\pi j(k\omega + l\nu)]|^2} \quad (7)$$

- Taking the Fourier transform of the power spectral density, will give us the autocorrelation fonction.

$$r(k,l) = \int\int_{-1/2}^{1/2} S(\omega,\nu)exp[2\pi j(k\omega + l\nu)]d\omega d\nu \quad (8)$$

- Building the covariance matrix with the autocorella-tion function.

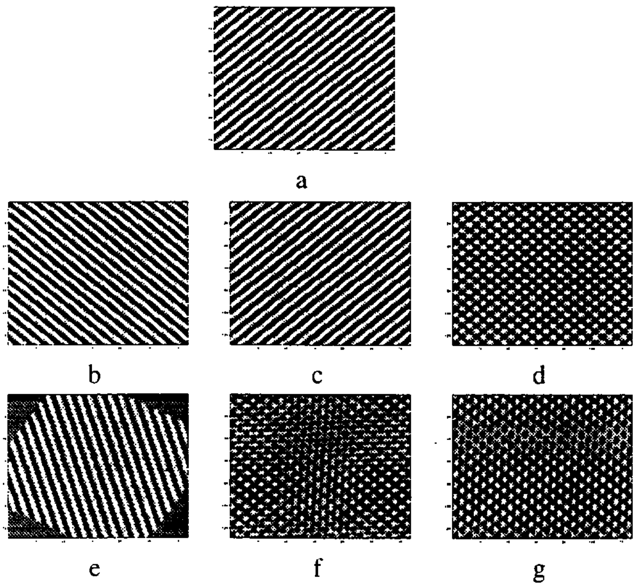

Figure 1: Measuring the distance between synthetic textures using Kullback's distance measure. In decreasing order : a) Texture H1 – the reference texture; b) Texture H1, 90 degree rotated, c) Texture H1, translated d) Texture H2, e) Texture H1 ,60 degree rotated, f) Texture H4, g) Texture H3

4. UNSUPERVISED IMAGE SEGMENTATION

Our goal, in this part of the research, is to employ the Wold-based texture model as a tool for segmenting the image into its distinct textured regions [1]. The proposed unsupervised segmentation procedure models the image using a doubly stochastic Gaussian (DSG) model [9]. A MRF model is applied to model the label field (which indicates the texture

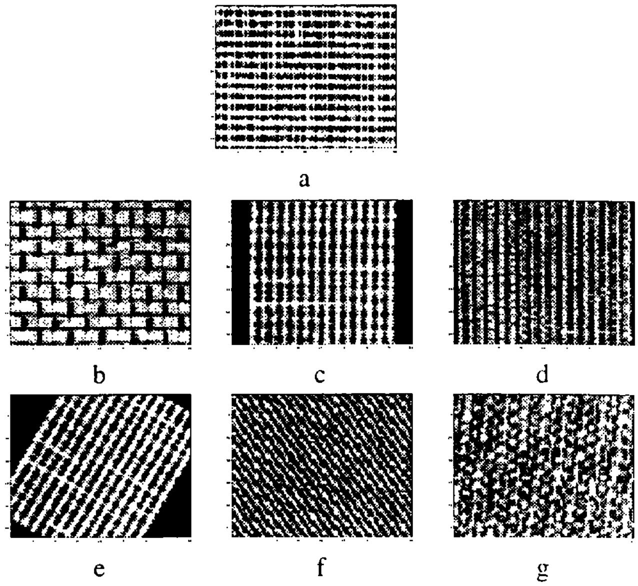

Figure 2: Measuring the distance between Brodatz textures using Kullback's distance measure. In decreasing order : a) Texture T7 – the reference texture; b) Texture T8; c) Texture T7, 90 degree rotated; d) Texture T4; e) Texture T7, 60 degree rotated; f) Texture T6; g) Texture T1.

to which each pixel belongs) and the 2-D Wold decomposition based model is used to model each of the textures (the texture model parameters are conditioned on the label field).

Assume the entire image is composed of an *a-priori* known number, K, of texture patches. An initial coarse parameter estimate is obtained by dividing the image into small blocks, estimating the texture AR model parameters in each block, and clustering them in the parameter space. The clustering is performed using the K-means algorithm. On the regions given by the K-means algorithm we are estimating the Wold parameters. These parameter estimates of the individual textures are then used for the construction of an energy function. This energy function is used in a simulated annealing algorithm to obtain maximum *a-posteriori* (MAP) estimates of the label field. The label field estimation is iteratively performed until the algorithm converges, *i.e.*, until the number of label changes in one iteration falls below a preset threshold. Example results are shown in figure 3.

The optimal estimates of the label field, from the energetic point of view is given by :

$$f^* = min_{f\in F}\{U(f|d) = U(d|f) + U(f)\} \quad (9)$$

where f are the labels and d are the data. The energy $U(d|f)$ is defined as follows :

- At each site s, we are taking the patch centered in s with respect to a neighboring system $N(s)$, and we are computing its power density spectrum F_d.
- At the same site s we are looking for its label f_s. Using the corresponding Wold parameters we are synthetising the texture with respect to the same neighboring system $N(s)$, then the power density spectrum F_s of the synthetic texture is computed.
- The conditional energy is :

$$U(d|f) = \|F_d - F_f\|_2 \tag{10}$$

This energy is equivalent to the model error. Working in the spectral domain we are eliminating the phase error. The regularisation energy is defined using a Potts model :

$$U(f) = -K \sum_{s \in N(r)} \delta(f_r, f_s) \tag{11}$$

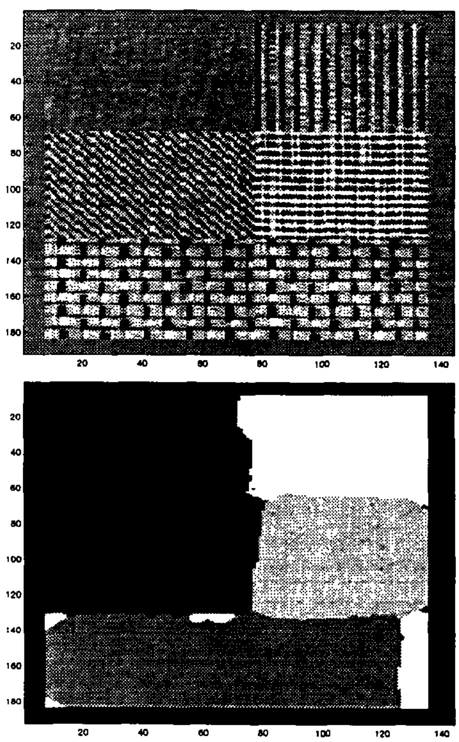

Figure 3: Unsupervised segmentation of Brodatz textures: Original image (top); Final segmentation, after 100 iterations (bottom).

5. CONCLUSION

The indexing and retrieval application will be implemented by integrating the texture segmentation and parameterization algorithms described above, with the texture distance measure. Thus, the distance measure will be applied in the retrieval application, in order to determine how "close" is a user supplied texture to a texture in one of the images already stored in the multimedia library. In the proposed system, each image stored in the library will be segmented to its textured and non-textured regions. The parameters of each textured region will be stored, and will serve as the "indices" to be compared with those of the query texture, using the distance measure we have already developed.

6. REFERENCES

[1] R. Stoica, Etude de la decomposition de Wold 2D et son application aux images texturees, DEA Thesis, ENSEEIHT, Toulouse University, Sept. 1997

[2] R. Mohr, P Gros, B. Lamiroy, S. Picard, C. Schmid, "Indexation et recherche d'images," Proc. GRETSI, Grenoble, Sept. 15-19, 1997.

[3] J. M. Francos, A. Z. Meiri, B. Porat, "A Wold-Like Decomposition of 2-D Discrete Homogeneous Random Fields," *Annls. Appl. Prob.*, vol. 5, pp. 248-260, 1995.

[4] M. P. Ekstrom, J. W. Woods, "Two-Dimensional Spectral Factorization with Applications in Recursive Digital Filtering,"*IEEE Trans. Acoust., Speech and Signal Process.*, vol. 24, pp. 115-128, 1976.

[5] J. M. Francos, A. Narashimhan, J. W. Woods, "Maximum-Likelihood Parameter Estimation of Discrete Homogeneous Random Fields with Mixed Spectral Distribution," *IEEE Trans. Signal Process.*, vol. 44, pp. 1242-1255, May 1996.

[6] J. M. Francos, A. Narashimhan, J. W. Woods, "Maximum-Likelihood Estimation of Textures Using a Wold Decomposition Model", *IEEE Trans. Image Process.*, vol. 4, pp. 1655-1666, Dec. 1995.

[7] M. Basseville, "Distance Measures for Signal Processing and Pattern Recognition,"*Signal processing*, 18(4), pp. 349-369, Dec. 1989.

[8] F. Liu, R. Picard, "Periodicity, Directionality and Randomness: Wold Features for Image Modeling and retrieval," *IEEE Trans. on PAMI*, 18(7), pp. 722-733, July 1996.

[9] F. C. Jeng, J. W. Woods, "Simulated Anealing in Compound Gaussian Random Fields,"*IEEE Trans. on Information Theory*, 36(1), pp. 94-107, Jan. 1990