RECOGNITION OF ROTATED AND SCALED TEXTURES USING 2-D AR MODELING AND THE FOURIER-MELLIN TRANSFORM

C. Cariou, O. Alata, C. Rosenberger, J.-M. Ogier, K. Chehdi
LASTI - Groupe Image,
ENSSAT - BP 447,
22305 Lannion Cedex, France
Tel: +33 2 96465030 ; Fax: +33 2 96466675
e-mail: cariou@enssat.fr

ABSTRACT

In this communication, we address the problem of the recognition and classification of rotated and scaled stochastic textures. We propose an extension of previous works, which are mostly based upon the use of the Fourier transform for the derivation of translation invariant statistics. More precisely, we suggest the sequential use of a 2-D high resolution spectral estimate, called Harmonic Mean power spectrum density (HM PSD), which presents a good tradeoff between reliability and complexity, with the Fourier-Mellin transform from which rotational and scaling invariants can be derived. Experimental results on rotated and scaled textures are presented that show the efficiency of this new technique.

1 INTRODUCTION

Among the different tasks to be performed in computer vision, from image acquisition towards image interpretation, some of them are particularly crucial. For example, while segmenting a real outdoor scene, one would like to recognize as a same region the perspective projection of the ground, which is often a textured region. Unfortunately, this region of the scene is intrinsically inhomogeneous, i.e. has statistical moments that are globally translation variant. Few results exist concerning the segmentation/classification of textures under this transformation group, especially when no prior information about the scene acquisition geometry is available.

In this communication, we restrict our attention to the particular image transformation group of positive similarities. We address the problem of the recognition of possibly rotated and scaled (zoomed) textures. Many works exist concerning the recognition of rotated and scaled wide-sense patterns, and a dense literature is available. Two main approaches are available. On one hand we find structural methods, principally used in handwritten character recognition. On the other hand are analytic methods, mostly based upon harmonic analysis, i.e. the use of the Fourier transform adapted to specific transformation groups, with applications for example in image registration and pattern recognition (see for example [8], [9]).

Recent developments in the field of harmonic analysis ([2], [5]) allow the construction of complete and convergent sets of invariants for 2-D patterns which are also well adapted to texture recognition. These methods are based upon the fitting of the Fourier transform to specific transformation groups (e.g. similitude, displacement).

Our contribution is based upon the latter technique : using some translation invariant statistical description of the texture, say the autocorrelation function (ACF) or its Fourier representation, the power spectrum density (PSD), it is possible to find a set of features which are characteristic and invariant under any rotation and any scaling of the texture. This idea is not a new one, and is known to provide relatively poor results in terms of texture recognition [10]. However, most techniques make use of either sample ACFs, or PSDs computed with the 2-D Discrete Fourier Transform (DFT). As far as we know, few of them take advantage of so-called 2-D modern spectral analysis methods and of their performances versus DFT-based techniques. An exception is the work developed by Cohen *et al.* [4] that use a 2-D Gauss-Markov Random Field (G-MRF) in order to model the texture : the objective is the same but the numerical treatment requires the optimization of a highly non-linear likelihood function.

The approach developed herein is two-fold : firstly is computed a high resolution spectrum estimate called Harmonic Mean (HM) PSD, which presents a good tradeoff between accuracy and computational load. Secondly, the HM PSD is taken as entry of the Fourier-Mellin transform from which scale and rotation invariants features can be derived.

2 THE PROPOSED METHOD

In the following, we will only be concerned by with homogeneous random textures, i.e. realisations of stationary colored (correlated) random fields, although this method may be applied with some restrictions to deterministic textures. The derivation of a texture recognition method that is invariant with respect to some trans-

formation group first leads to the problem of finding a reliable *translation invariant* description of the texture : this description must be phase blind, in order to avoid the search for generic tiles of the texture before applying some segmentation/classification scheme.
In this work we used the power spectrum density (PSD) function, that is the Fourier transform of the autocorrelation function (ACF). The choice of this description is justified non only by its phase blindness, but also because it is possible, by using modern spectral estimation techniques, to compute the 2-D PSD within short term windows, which may be of great use for a further segmentation of distorted (i.e. locally scaled and rotated) objects. The other problem lies in the development of invariant features. For this, we used the Fourier-Mellin transform, which provides rotation and scale invariant features for a 2-D pattern.

2.1 PSD Estimation

Many techniques have been developed since the mid seventies within the framework of modern spectral estimation. While some of them rely on some deep principles (e.g. maximum entropy), many others use some direct modeling of the statistical interactions between neighbours on the 2-D lattice (e.g. MRF, autoregressive (AR), autoregressive-moving average (ARMA) modeling). Although their performances are often close together, their implementation may require in some cases the optimization of a highly non-linear functional.
Here we chose a causal AR modeling of the texture because of its computational tractability, and also for the fact that a class of reliable DSP estimates can be derived from it [1].
More precisely, let Y be a random process on the $\mathbb{Z}^2$ lattice, and $\{y(m,n)\}_{0\leq m\leq M-1; 0\leq n\leq N-1}$ be a finite realisation of Y. In order to simplify the notation, the 2-D sequence $\{y(m,n)\}$ is assumed to be zero mean. Otherwise the mean should be removed from the image. Then the first quadrant (Q_1) AR modeling with order (P,P) of $\{y(m,n)\}$ is given by :

$$\begin{aligned} y(m,n) &= \sum\sum_{(k,l)\,\in\pi_1} a_1(k,l)\; y(m-k,n-l) \\ &+ \; w_1(m,n) \quad , \end{aligned} \tag{1}$$

where $\{a_1(k,l)\}$ are the Q_1 AR parameters, $\pi_1 = \{(k,l) \mid 0 \leq k \leq P-1; 0 \leq l \leq P-1; (k,l) \neq (0,0)\}$ is the first quarter plane support and $w_1(m,n)$ is a driving noise sequence with zero mean and variance σ_1^2. Then the corresponding Q_1 PSD is given by

$$P_{Q_1}(f_1,f_2) = \frac{\sigma_1^2}{\left|1 - \sum\sum_{(k,l)\,\in\pi_1} a_1(k,l)\; \exp(-j\; 2\pi\,[f_1\; k + f_2\; l])\right|^2} \quad , \tag{2}$$

where $(f_1,f_2) \in [-0.5,0.5]\times[-0.5,0.5]$ are normalised frequencies. It is well known (see [1], [3], [7]) that the Q_1 PSD estimate is severely biased due to the choice of such an anisotropic support for (k,l). However, this drawback may be considerably reduced by use of the second quadrant (Q_2) AR modeling :

$$\begin{aligned} y(m,n) &= \sum\sum_{(k,l)\,\in\pi_2} a_2(k,l)\; y(m-k,n-l) \\ &+ \; w_2(m,n) \quad , \end{aligned} \tag{3}$$

where $\pi_2 = \{(k,l) \mid 0 \leq k \leq P-1; -P+1 \leq l \leq 0; (k,l) \neq (0,0)\}$, leading to the Q_2 PSD $P_{Q_2}(f_1,f_2)$. Both PSDs can be combined in the following manner to yield the so-called *Harmonic Mean* (HM) PSD estimate [6]:

$$\frac{1}{P_{HM}(f_1,f_2)} = \frac{1}{2}\left[\frac{1}{P_{Q_1}(f_1,f_2)} + \frac{1}{P_{Q_2}(f_1,f_2)}\right] \quad . \tag{4}$$

This spectral estimate has been shown in [3] to provide results that are comparable with the Maximum Entropy method, but with a much reduced computational burden. Actually, the set of first and second quadrant parameters $\{\{a_1(k,l)\}_{(k,l)\in\pi_1}, \sigma_1^2; \{a_2(k,l)\}_{(k,l)\in\pi_2}, \sigma_2^2\}$ are obtained through the inversion of symmetric covariance matrices.
One important skill of the use of a parametric spectral description lies in the fact that the quarter plane PSD estimates may be computed directly in polar coordinates, which is important for the use of the Fourier-Mellin transform. For example, we have for the first quadrant DSP :

$$P_{Q_1}(\rho,\theta) = \frac{\sigma_1^2}{\left|1 - \sum\sum_{(k,l)\,\in\pi_1} a_1(k,l)\; \exp(-j\; 2\pi\;\rho\;[k\cos\theta + l\sin\theta])\right|^2} \tag{5}$$

with $\rho \in [0,\frac{1}{2}]$ and $\theta \in [0,2\pi]$. Moreover, due to the fact that the 2-D PSD is an even function (i.e. symmetric w.r.t. the null frequency), only one half of the spectral domain is necessary to describe it, and as a consequence $P_{HM}(\rho,\theta)$ is π-periodic w.r.t. θ.

2.2 The Fourier-Mellin transform and derived invariants

Definition 1 (Fourier-Mellin Transform)

Let $f: \begin{array}{ccc} \mathbb{R}^{+*}\times[0,2\pi) & \longrightarrow & \mathbb{R} \\ (\rho,\theta) & \mapsto & f(\rho,\theta) \end{array}$,

the Fourier-Mellin transform is defined as :

$$M_f(\nu,q) = \int_{\rho=0}^{+\infty}\int_{\theta=0}^{2\pi} \rho^{-i\nu}\;\exp(-iq\theta)\; f(\rho,\theta)\;\frac{d\rho}{\rho}\; d\theta\;, \tag{6}$$

with $q \in \mathbb{Z}$, $\nu \in \mathbb{R}$.

The Fourier-Mellin (FM) integral converges only under strong hypothesis for $f(\rho,\theta)$. For this reason, in [5], Ghorbel proposed an analytic prolongation of the FM transform, called AFMT.

Definition 2 (AFMT) *With the same notation as above, the analytic prolongation of the Fourier-Mellin transform is defined as :*

$$\tilde{M}_f(\nu,q) = \int_{\rho=0}^{+\infty}\int_{\theta=0}^{2\pi} \rho^{-i\nu+\sigma_0}\ \exp(-iq\theta)\ f(\rho,\theta)\ \frac{d\rho}{\rho}\ d\theta \tag{7}$$

with $q \in \mathbb{Z}$, $\nu \in \mathbb{R}$, and $\sigma_0 \in \mathbb{R}^{+\star}$ is a given parameter.

The AFMT is used to avoid the divergence of the Fourier-Mellin integral for almost all patterns.
The properties of the AFMT are the same as the FM transform, as shown below.

Property 1 (Effect of rotation and scaling) *If $g(\rho,\theta) = f(\alpha\rho,\theta+\beta)$ is a scaled and rotated version of $f(\rho,\theta)$, then*

$$\tilde{M}_g(\nu,q) = \alpha^{-\sigma_0+i\nu}\ \exp(iq\beta)\ \tilde{M}_f(\nu,q) \tag{8}$$

This property can be viewed as the application to the similitude group of the generalized shift theorem of the Fourier transform. Note that α and β can be retrieved by :

$$\alpha = \left[\frac{\tilde{M}_f(0,0)}{\tilde{M}_g(0,0)}\right]^{1/\sigma_0} \quad ; \quad \beta = \arg\left[\frac{\tilde{M}_g(0,1)}{\tilde{M}_f(0,1)}\right] \tag{9}$$

Taking the modulus of both terms in (8) leads to features which are invariant under any rotation of the pattern but not under a scaling transformation.
Here we propose to use of the following set of invariants, inspired from [5].

Proposition 1 *Let $f(\rho,\theta)$ be a 2-D pattern and $\tilde{M}_f(\nu,q)$ its AFMT. Then*

$$I_f(\nu,q) = \left|\tilde{M}_f(\nu,q)\ [\tilde{M}_f(0,0)]^{-1+i(\nu/\sigma_0)}\right| \tag{10}$$

defines a set of invariants under any rotation and scaling of the pattern.

This particular set of real invariant features does not have the completeness property (i.e. different patterns may have the same set of invariants), nor the inversibility property (i.e. $f(\rho,\theta)$ cannot be reconstructed from $I_f(\nu,q)$). Nevertheless, a set of complex, complete and inversible invariants can be used (see [5]).
Also for implementation it appears necessary to discretise the AFMT by sampling the ν variable. Thus are computed $\tilde{M}_f(p,q)$ and $I_f(p,q)$ with $p \in \mathbb{Z}$. Moreover, it can be shown easily that $\tilde{M}_f(-p,-q) = \tilde{M}_f^*(p,q)$ where * denotes the complex conjugate, and as a consequence we have $I_f(-p,-q) = I_f(p,q)$.

2.3 Application of the AFMT to the HM PSD

We propose here to take the above HM PSD as entry of the AFMT. One can note that another important effect of using an AR modeling for the texture is the fact that the corresponding parametric PSD is a regular (C^∞) function in the frequency plane. As such, it can be described by a small set of invariants in the FM representation.
However, it should be insured before applying this technique, that the PSD decreases sufficiently quickly to zero as $\rho \longrightarrow \frac{1}{2}$. In other terms, the analysis of textures with too high frequencies should be avoided. Moreover, deterministic textures with purely harmonic components for which the theoretic PSD shows a Dirac distribution should also be avoided, because the effect of a scaling does not dilate Dirac functions in the frequency plane.
For each texture to be classified, the corresponding $P_{HM}(\rho,\theta)$ is computed, and the set of $I_{P_{HM}}(p,q)$ is derived. Finally, a simple Euclidean distance applied to sets of $I_{P_{HM}}(p,q)$ obtained for different textures is used to evaluate the similarity between textures.

3 EXPERIMENTAL RESULTS

We now present some experimental results obtained with this method. Figure 1 shows eight pairs of textures taken from the Brodatz album, each pair with different relative orientations and/or scaling factors. The objective is to recognize pairs of identical textures among the whole set of textures, and, in the case of a correct recognition, to provide estimates of the relative angle and the scaling factor between both textures. The recognition of a texture pair is made through the minimization of the Euclidean distance between all sets of invariants.
For each texture, 24 parameters ($P = 5$ in (1)) were chosen for the quarter plane causal AR modeling. Letting $-3 \leq p \leq 3; 0 \leq q \leq 3$ and taking into account the even symmetry of invariants gave us a set of 25 features for each texture. In this experiment we chose $\sigma_0 = 1.0$ for the computation of the AFMT.
Figure 2 summarises the results. For all correct recognition, a double arrow is drawn between textures. Jointly are given true values and estimates of the scaling factor and relative angle between pairs of correctly recognized textures. At first, one can remark that all pairs of textures were correctly recognized, except for the original D17 texture ; this is probably due to the highly structural properties of this texture. Secondly, estimates of scale factors for correctly recognized textures are often in agreement with true values, except for highly periodic textures like D68 or D77. However, scale factors for more random textures like D24 are much correctly estimated. All these facts seem to confirm the above remarks relative to the general use of this method. As for the relative angle, one can note that most estimates lie in the 10 degrees error range, with for example an estimate of $\hat{\phi} = 5^0$ in place of the true $\phi = 0^0$ for the

isotropic D29 texture.

4 CONCLUSION AND PERSPECTIVES

In this communication we have presented a new approach for rotational and scale invariant recognition of stochastic textures. This method is based upon the combination of, on one hand, high resolution spectrum estimation (AR modeling and HM PSD), and, on the other hand, the use of the Fourier-Mellin transform and derived invariants. Preliminary experimental results are encouraging, and will probably be improved by use of a more efficient spectrum estimate, such as the HMHV PSD [1]. A first perspective of this work concerns the development of a segmentation process that is invariant under rotation and scaling, which is to be very useful in future computer vision systems. Another perspective lies in the research for invariants with respect to other image transformations like stretching or skewing.

References

[1] O. Alata, P. Baylou, M. Najim, 'A new 2-D spectrum estimate using multichannel AR approach of 2-D fast RLS algorithms,' *Proc. ICIP*, pp. 442-445, Santa Barbara, CA, 1997.

[2] R. D. Brandt, F. Lin, 'Representations that uniquely characterize images modulo translation, rotation and scaling,' *Pattern Recognition Letters*, **17**, pp. 1001-1015, 1996.

[3] C. Cariou, J.-M. Boucher, R. Garello, 'Wavelength and orientation resolution comparison of spectral estimators for ocean wave systems on SPOT images,' *Proc. IGARSS*, pp. 1809-1812, Washington DC, 1990.

[4] F. S. Cohen, Z. Fan, M. A. Patel, 'Classification of rotated and scaled textured images using Gaussian Markov Random Field models,' *IEEE Trans. Pattern Analysis Mach. Intell.*, **13**, pp. 192-202, 1991.

[5] F. Ghorbel, 'A complete invariant description for gray-level images by the harmonic analysis approach,' *Pattern Recognition Letters*, **15**, pp. 1043-1051, 1994.

[6] L. B. Jackson, H. C. Chien, 'Frequency and bearing estimation by two-dimensional linear prediction,' *Proc. ICASSP*, Washington DC, pp. 665-668, 1979.

[7] S. M. Kay, *Modern Spectral Estimation - Theory and Application*, Prentice-Hall, Englewood Cliffs, NJ, 1988.

[8] G. Ravichandran, M. M. Trivedi, 'Circular-Mellin features for texture analysis,' *IEEE Trans. Image Processing*, **4**, no 12, pp. 1629-1640, 1995.

[9] B. S. Reddy, B. N. Chatterji, 'An FFT-based technique for translation, rotation, and scale-invariant image registration,' *IEEE Trans. Image Processing*, **5**, no 8, pp. 1266-1271, 1996.

[10] A. Rosenfeld, A. C. Kak, *Digital Picture Processing 2*, Academic Press, San Diego, CA, 1982 (second edition).

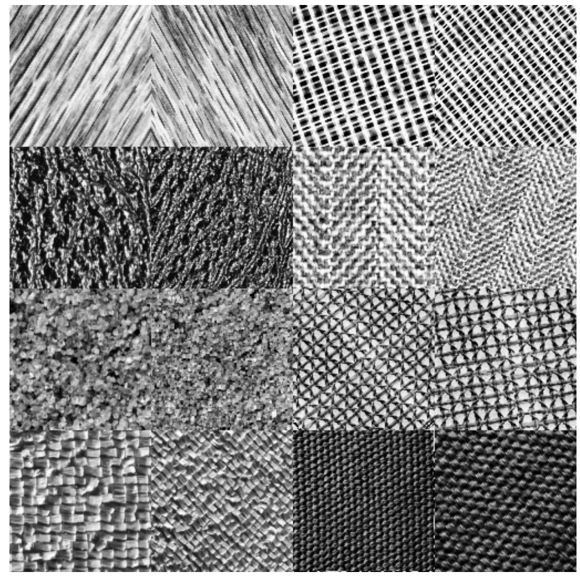

Figure 1: Pairs of rotated and scaled textures

D68	D68 $\alpha = 1.5$ $\hat{\alpha} = 0.81$ $\beta = 60^o$ $\hat{\beta} = 58^o$	D21	D21 $\alpha = 1.5$ $\hat{\alpha} = 1.51$ $\beta = 20^o$ $\hat{\beta} = 68^o$
D24	D24 $\alpha = 1.5$ $\hat{\alpha} = 1.57$ $\beta = 30^o$ $\hat{\beta} = 22^o$	D17	D17
D29	D29 $\alpha = 1.5$ $\hat{\alpha} = 1.21$ $\beta = 0^o$ $\hat{\beta} = 5^o$	D52	D52 $\alpha = 0.5$ $\hat{\alpha} = 0.45$ $\beta = 30^o$ $\hat{\beta} = 30^o$
D84	D84 $\alpha = 1.5$ $\hat{\alpha} = 2.20$ $\beta = 45^o$ $\hat{\beta} = 37^o$	D77	D77 $\alpha = 0.5$ $\hat{\alpha} = 1.11$ $\beta = 50^o$ $\hat{\beta} = 43^o$

Figure 2: Recognition and estimation of relative angles and scale factors