Robust Rotation-Invariant Texture Classification Using a Model Based Approach

Patrizio Campisi, *Member, IEEE*, Alessandro Neri, *Member, IEEE*, Gianpiero Panci, and
Gaetano Scarano, *Member, IEEE*

Abstract—In this paper, a model based texture classification procedure is presented. The texture is modeled as the output of a linear system driven by a binary image. This latter retains the morphological characteristics of the texture and it is specified by its spatial autocorrelation function (ACF). We show that features extracted from the ACF of the binary excitation suffice to represent the texture for classification purposes. Specifically, we employ a moment invariants based technique to classify the ACF. The resulting proposed classification procedure is thus inherently rotation invariant. Moreover, it is robust with respect to additive noise. Experimental results show that this approach allows obtaining high correct rotation-invariant classification rates while containing the size of the feature space.

Index Terms—Moment invariants, texture analysis, texture classification.

I. INTRODUCTION

TEXTURE classification has been widely investigated during the past decades. In fact it plays an important role in many applications such as remote sensing, robot vision, crop classification, automatic tissue recognition in medical imaging, content based access to image databases, to cite only a few.

The first stage of the texture classification problem is the training phase. It consists in the extraction of a certain number of texture features from each texture belonging to a training set; the extracted features are then collected in a feature vector. In the subsequent classification stage, the feature vector of the texture to classify is evaluated and then compared to all the feature vectors of the textures training set. Therefore, two main issues must be coped with. The first consists in properly selecting the texture features, which have to be chosen in order to represent the content of each different texture in the database. Moreover their number must be optimized. The second one is the definition of a suitable similarity function.

Several approaches for texture features extraction have been proposed in literature. Most of them are based on statistical, signal processing, and model-based techniques. Early works where based on the analysis of some second order statistical properties of the texture [1], such as the co-occurrence matrix [2]. The signal processing techniques are mainly based on texture filtering followed by energy evaluation. A review of major filtering approaches and a comparative study can be found in [3]. In the recent literature the use of a filter bank instead of a single filter has been proposed, giving rise to several multichannel texture analysis systems [4], [5]. In particular, Gabor filtering has been extensively studied [6], [7]. Moreover, in the last decade, wavelet theory has been widely used for texture classification purposes [8]–[10]. Several stochastic models have also been proposed for texture modeling and classification; they include Gaussian Markov random fields models [11], [12], moving average (MA), autoregressive (AR), and autoregressive moving average (ARMA) models [13]–[15]. In [16] a fractal model has been proposed, and in [17], [18] statistical and harmonic features have been combined. Many of the aforementioned methods, although allow obtaining good classification performances, present a high misclassification rate when the texture is rotated and contaminated with additive Gaussian noise. In [19] and [20] rotation invariant texture classification wavelet based methods are presented.

In this contribution we describe a novel rotation-invariant model based texture classification method. Moreover, a performance comparison is performed with respect to classification methods using a robust rotation invariant wavelet based approach [19] as well as the co-occurrence matrix, which are briefly summarized in Appendices A and B respectively.

The model here employed was originally proposed in [21] for texture synthesis purposes. This model has shown a remarkable versatility since it has been used, in an extended form, in [22] for the synthesis of color textures. In [23] the model has been successfully employed for a model based texture coder scheme. Moreover, stemming from the one proposed in [21], a hybrid approach for texture synthesis, which operates partially in the spatial domain and partially in a multiresolution domain, has been presented in [24].

Hence, according to the aforementioned model, a texture can be represented using

- a proper binary excitation, which is designed to retain the morphological characteristics of the texture sample;
- a linear filter;
- a zero-memory nonlinearity performing a histogram matching.

The rationale behind our approach is that the binary excitation itself suffices to represent the texture for classification purposes. Moreover, the use of the binary excitation makes the classification algorithm more robust with respect to additive Gaussian
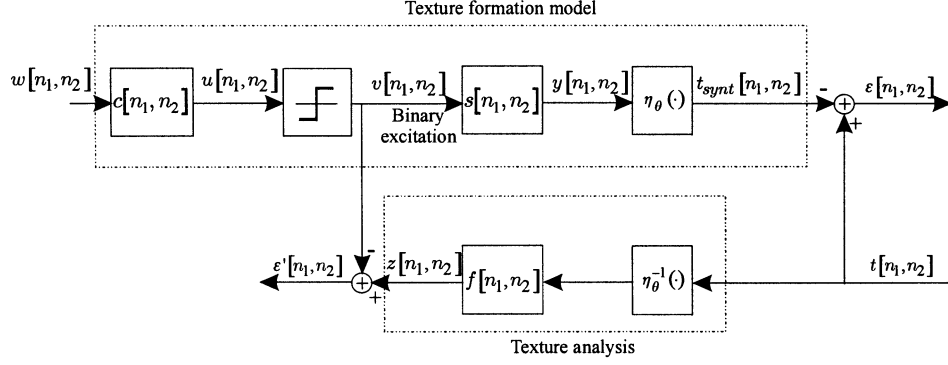
Fig. 1. Texture formation model and texture analysis.

noise. Therefore, after "projecting" the texture onto the reduced complexity binary space, we resort to classify the binary image so obtained using the *rotation invariant moment* [25], [26] of its ACF.

The paper is organized as follows. In Section II the employed texture model is described. In Section III the algorithm for the texture features extraction is detailed. The classification algorithm is described in Section IV. In Section V the experimental results are presented. Ultimately, conclusions are drawn in Section VI.

II. TEXTURE MODEL

With reference to the scheme depicted in Fig. 1, a given texture $t[n_1, n_2]$ is modeled according to the following factorization:

$$t[n_1, n_2] = t_{synt}[n_1, n_2] + \epsilon[n_1, n_2]$$
$$t_{synt}[n_1, n_2] = \eta_{\boldsymbol{\theta}}\big(y[n_1, n_2]\big)$$
$$y[n_1, n_2] = s[n_1, n_2] * v[n_1, n_2] \tag{1}$$

where

- $v[n_1, n_2]$ is a random binary excitation, tailored to retain the zero-crossing locations of the texture prototype, thus capturing the basic morphology of the texture;
- $s[n_1, n_2]$ is a invertible shaping filter, with inverse denoted by $f[n_1, n_2]$, that adds gray-scale details to the binary image $v[n_1, n_2]$;
- $\eta_{\boldsymbol{\theta}}(\cdot)$ is an invertible zero-memory nonlinearity, with inverse denoted by $\eta_{\boldsymbol{\theta}}^{-1}(\cdot)$. It performs a histogram modification, and its behavior is described by the parameters collected in the vector $\boldsymbol{\theta}$. Specifically, for the 256 gray-scale level images under analysis, $\eta_{\boldsymbol{\theta}}(\cdot)$ operates as a look-up-table (LUT) of size 256 values, which transforms each gray-scale input level in such a way that the output image has the desired histogram. The vector $\boldsymbol{\theta}$ simply collects the 256 output values of the LUT;
- $\epsilon[n_1, n_2]$ is a realization of a random field that accounts for any possible model mismatching;
- $y[n_1, n_2]$ is the reproduced texture before the final histogram matching operation. Note that, since the transformation between $y[n_1, n_2]$ and $t_{synt}[n_1, n_2]$ can be inverted, we can also write $y[n_1, n_2] = \eta_{\boldsymbol{\theta}}^{-1}\big(t_{synt}[n_1, n_2]\big)$.

In essence, the model replaces usual independent and identically distributed (i.i.d.) sources employed in linear texture models

[13]–[15] with a non i.i.d. binary source. The binary excitation is designed in order to have the same ACF of the binarized version of the texture sample under analysis. Thus binary field gives the zero-crossing location of the texture prototype. Since most of the structure of a texture is contained in its zero-crossing locations [27], the basic visual morphology of the prototype is then captured. Specifically, the binary excitation $v[n_1, n_2]$ is obtained by filtering a realization $w[n_1, n_2]$ of a white Gaussian random field by a linear filter $c[n_1, n_2]$, and then by hard-limiting the output. The filter $c[n_1, n_2]$ is designed in order to generate a binary excitation $v[n_1, n_2]$ having the same spatial ACF of the corresponding binarized texture obtained in the analysis stage.

The rationale behind this model relies on psycho-visual experiments conducted on texture discrimination [28], which have pointed out that the human visual system, although astonishingly capable of discriminating between different patterns, cannot distinguish between textures that differ only in third and higher order statistics. The effectiveness of this model has been discussed in [21].

As already pointed out in the Introduction, we resort to use the binary excitation to reduce the dimensionality of the space where the texture is represented. Therefore, the texture classification problem reduces to the classification of the spatial ACF of the binary image, obtained in the analysis stage, as described in the next section.

III. TEXTURE FEATURES EXTRACTION

According to the model described in the previous section, the texture prototype can be represented by the spatial ACF $\hat{R}_{\tilde{v}\tilde{v}}[k_1, k_2]$, the filter $s[n_1, n_2]$, obtained as the inverse of the estimation of the filter $f[n_1, n_2]$, and the zero-memory nonlinearity $\eta_{\boldsymbol{\theta}}(\cdot)$. However for classification purposes a texture can be effectively represented by $\hat{R}_{\tilde{v}\tilde{v}}[k_1, k_2]$ obtained assuming the model depicted in Fig. 1 in the blind deconvolution procedure described in Section III-A. Then, $\hat{R}_{\tilde{v}\tilde{v}}[k_1, k_2]$ is characterized by means of the moment invariants as described in Section III-B.

A. Texture Parameters Identification Algorithm

With reference to Fig. 1, let us denote with $\mathbf{v}$, $\mathbf{y}$, $\mathbf{t}$, the column lexicographically ordered arrays associated with $v[n_1, n_2]$, $y[n_1, n_2]$, $t[n_1, n_2]$, and with $\mathbf{f}$ the filter coefficients associated with the filter $f[n_1, n_2]$. Our task is to find the triplet

$(\tilde{\mathbf{v}}, \tilde{\mathbf{f}}, \tilde{\boldsymbol{\theta}})$ by maximizing the log-likelihood function ($\tilde{\ }$ denotes "estimate" of the corresponding vector):

$$(\tilde{\mathbf{v}}, \tilde{\mathbf{f}}, \tilde{\boldsymbol{\theta}}) = \arg \underset{\mathbf{v}, \mathbf{f}, \boldsymbol{\theta}}{\mathrm{Max}} \underset{\mathbf{T}|\mathbf{V};\mathbf{F},\boldsymbol{\Theta}}{\ln p(\mathbf{t}|\mathbf{v};\mathbf{f},\boldsymbol{\theta})}. \tag{2}$$

However, since the cost function of the optimization problem expressed by (2) is highly nonlinear, the numerical search of the solution requires a significantly heavy computational burden. In order to reduce the overall complexity of the optimization, we resort to a sub-optimum criterion consisting of an iterative maximization of the log-likelihood function with respect to each of the vectors $\mathbf{v}$, $\mathbf{f}$, $\boldsymbol{\theta}$, separately, that is:

$$\tilde{\mathbf{f}} = \arg \underset{\mathbf{f}}{\mathrm{Max}} \underset{\mathbf{T}|\mathbf{V};\mathbf{F},\boldsymbol{\Theta}}{\ln p(\mathbf{t}|\tilde{\mathbf{v}};\mathbf{f},\tilde{\boldsymbol{\theta}})} \tag{3}$$

$$\tilde{\mathbf{v}} = \arg \underset{\mathbf{v}}{\mathrm{Max}} \underset{\mathbf{T}|\mathbf{V};\mathbf{F},\boldsymbol{\Theta}}{\ln p(\mathbf{t}|\mathbf{v};\tilde{\mathbf{f}},\tilde{\boldsymbol{\theta}})} \tag{4}$$

$$\tilde{\boldsymbol{\theta}} = \arg \underset{\boldsymbol{\theta}}{\mathrm{Max}} \underset{\mathbf{T}|\mathbf{V};\mathbf{F},\boldsymbol{\Theta}}{\ln p(\mathbf{t}|\tilde{\mathbf{v}};\tilde{\mathbf{f}},\boldsymbol{\theta})}. \tag{5}$$

To further simplify the optimization, as pointed out in [21], we note that, in our estimation problem, $\mathbf{t}$ can be replaced by

$$\mathbf{z} \overset{\mathrm{def}}{=} \mathbf{F}\{\eta_{\boldsymbol{\theta}}^{-1}(\mathbf{t})\} = \mathbf{v} + \boldsymbol{\epsilon}'$$

being $\mathbf{F}$ the linear operator associated with the filter $\mathbf{f}$ and $\boldsymbol{\epsilon}'$ a realization of a random noise process representing the deconvolution error (see Fig. 1). In fact, since the transformation $\mathbf{F}\{\eta_{\boldsymbol{\theta}}^{-1}(\cdot)\}$ is invertible, $\mathbf{z}$ and $\mathbf{t}$ span the same population.

Using the same argumentation as in [29], we may argue that the residual $\boldsymbol{\epsilon}'$ tends to be white and Gaussian distributed, uncorrelated with $\mathbf{v}$ as soon as the iterations proceed. Therefore, by equating the gradient of $\underset{\mathbf{Z}|\mathbf{V};\mathbf{F},\boldsymbol{\Theta}}{\ln p(\mathbf{z}/\tilde{\mathbf{v}};\mathbf{f},\tilde{\boldsymbol{\theta}})}$ w.r.t. $\mathbf{f}$ to zero and solving with respect to $\mathbf{f}$, we obtain [21]:

$$\tilde{\mathbf{f}} = \widehat{R}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}^{-1} \cdot \hat{r}_{\tilde{\mathbf{v}}\tilde{\mathbf{y}}} \tag{6}$$

where $\widehat{R}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}} \overset{\mathrm{def}}{=} \mathbf{Y}^{\mathrm{T}}\mathbf{Y}$ is the sample auto-correlation vector of $\tilde{\mathbf{y}}$, and $\hat{r}_{\tilde{\mathbf{v}}\tilde{\mathbf{y}}} \overset{\mathrm{def}}{=} \mathbf{Y}^{\mathrm{T}}\tilde{\mathbf{v}}$ is the sample cross-correlation matrix of $\tilde{\mathbf{v}}$ and $\tilde{\mathbf{y}}$, being $\mathbf{Y}$ the Toeplitz matrix representing the linear operator associated to the filter with impulse response equal to $\mathbf{y}$.

Moreover, at the neighborhood of the desired solution, $\tilde{v}[n_1, n_2]$ assumes the form

$$\tilde{v}[n_1, n_2] = \mathrm{sign}\left\{\eta_{\boldsymbol{\theta}}^{-1}(t[n_1, n_2]) * f[n_1, n_2]\right\}. \tag{7}$$

Finally, note that the value $\tilde{\boldsymbol{\theta}}$ that maximizes (5), maximizes also the quantity

$$\left(\eta_{\boldsymbol{\theta}}^{-1}(\mathbf{t}) - \mathbf{F}\tilde{\mathbf{v}}\right)^{\mathrm{T}}\left(\eta_{\boldsymbol{\theta}}^{-1}(\mathbf{t}) - \mathbf{F}\tilde{\mathbf{v}}\right). \tag{8}$$

Therefore to achieve a good approximation the transformation $\eta_{\boldsymbol{\theta}}(\cdot)$ is chosen to provide a histogram matching between $\mathbf{F}\tilde{\mathbf{v}}$ and $\mathbf{t}$.

The iterative identification algorithm is depicted in Fig. 2 in its flow graph form, and it is summarized in the following. The initialization step consists of assuming that the nonlinearity $\eta_{\boldsymbol{\theta}}^{(0)}(\cdot)$ and the inverse filter $f^{(0)}[n_1, n_2]$ do not affect the original texture, i.e., they are set equal to the identity operators, $\eta_{\boldsymbol{\theta}}^{(0)}(\cdot) = (\cdot)$, and $f^{(0)}[n_1, n_2] = \delta[n_1, n_2]$, being $\delta[n_1, n_2]$ the bidimensional unit sample sequence.

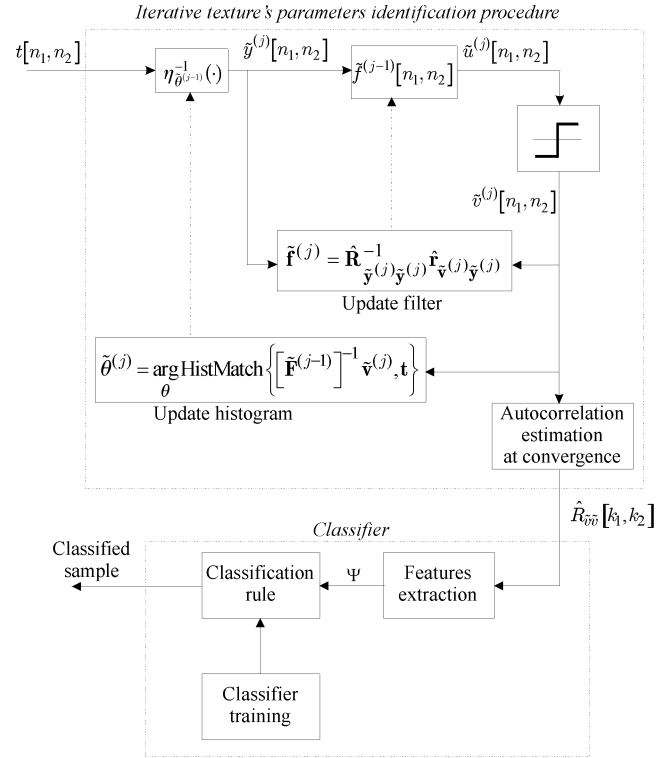Fig. 2.　Texture sample classification procedure.

Then, at the generic jth iteration the simultaneous estimation of the binary excitation $\tilde{v}^{(j)}[n_1, n_2]$ and of the filter $f^{(j)}[n_1, n_2]$ is performed by modifying the histogram of the original texture according to the nonlinearity $\eta_{\boldsymbol{\theta}}^{(j)}(\cdot)$, and by deconvolving the so obtained histogram modified texture $\tilde{y}^{(j)}[n_1, n_2] = \eta_{\boldsymbol{\theta}}^{(j)}(t[n_1, n_2])$ through the previously estimated inverse filter $f^{(j-1)}[n_1, n_2]$. The deconvolved texture $\tilde{u}^{(j)}[n_1, n_2] = y^{(j)}[n_1, n_2] * f^{(j-1)}[n_1, n_2]$ is then hard-limited in order to give the estimate $\tilde{v}^{(j)}[n_1, n_2]$ of the binarized version of the prototype; this latter is finally used to update both the estimates of the inverse filter $f[n_1, n_2]$ and of the zero-memory nonlinearity $\eta_{\boldsymbol{\theta}}(\cdot)$.

The deconvolution algorithm reaches an equilibrium point at the jth iteration when the cross-correlation between $\tilde{y}^{(j)}[n_1, n_2]$ and $\tilde{u}^{(j)}[n_1, n_2]$ is proportional to the cross-correlation between $\tilde{y}^{(j)}[n_1, n_2]$ and $\tilde{v}^{(j)}[n_1, n_2]$, i.e., $R_{\tilde{y}^{(j)}\tilde{u}^{(j)}}[k_1, k_2] \propto R_{\tilde{y}^{(j)}\tilde{v}^{(j)}}[k_1, k_2]$. A detailed discussion on the convergence issue can be found in [24].

When the convergence has been reached, a sample estimate $\hat{R}_{\tilde{v}\tilde{v}}[k_1, k_2]$ of the statistical autocorrelation $R_{\tilde{v}\tilde{v}}[k_1, k_2] \overset{\mathrm{def}}{=} \mathrm{E}\{v[n_1 + k_1, n_2 + k_2]v[n_1, n_2]\}$ is calculated and the moment invariants, detailed in the next section, are evaluated on it.

B. Moment Invariants

Moment invariants were introduced by Hu in his pioneering paper [25], where a set of seven features invariant to the rotation of 2-D objects was introduced. They are summarized in Appendix C. Several works have been done on the definition of various kind of moment invariants, and their performances have been investigated [30], [31]. However, the proposed invariant

sets are algebraically dependent, and this fact is highly undesirable since our goal is the definition of a low dimensionality feature space.

In our classification method, we rely on the moment invariants derived following the approach proposed in [26]. The rational behind this choice relies on the possibility of deriving moment invariants of any order and of defining "bases" of invariants. The moment invariants are defined through the complex moments. A complex moment c_{pq} of order $(p+q)$ of a two-dimensional probability density function $f(x, y)$ is defined as

$$c_{pq} = \sum_{k=0}^{p} \sum_{j=0}^{q} \binom{p}{k} \binom{q}{j} (-1)^{q-j} i^{p+q-k-j} m_{k+j, p+q-k-j}$$
(9)

where $i \overset{\text{def}}{=} \sqrt{-1}$ denotes the imaginary unit and

$$m_{u,v} = \int_{-\infty}^{-\infty} \int_{-\infty}^{-\infty} x^u y^v f(x, y)\, dx dy$$
(10)

is the ordinary noncentral moment of order (u, v).

Following the approach proposed in [26], in order to represent the bidimensional autocorrelation function $\hat{R}_{\tilde{v}\tilde{v}}[k_1, k_2]$ of the deconvolved binary texture $\tilde{v}[n_1, n_2]$, we resort to use the second and third order basis:

$$\begin{aligned}
\psi_1 &= \phi_1, \\
\psi_2 &= \phi_4, \\
\psi_3 &= \phi_6, \\
\psi_4 &= \text{Im}(c_{20} c_{12}^2), \\
\psi_5 &= \phi_5, \\
\psi_6 &= \phi_7
\end{aligned}$$
(11)

and the fourth order basis:

$$\begin{aligned}
\psi_7 &= c_{22}, \\
\psi_8 &= \text{Re}(c_{31} c_{12}^2), \\
\psi_9 &= \text{Im}(c_{31} c_{12}^2), \\
\psi_{10} &= \text{Re}(c_{40} c_{12}^4), \\
\psi_{11} &= \text{Im}(c_{40} c_{12}^4)
\end{aligned}$$
(12)

where $\phi_j, j = 1, \ldots, 7$ are the Hu's moment invariants, defined in Appendix C.

The moments $\psi_j, j = 1, \ldots, 11$ are collected in a vector $\boldsymbol{\Psi}$, which represents the texture feature vector. Here we must immediately point out that an increase in the dimension of the feature vector has not led to a classification performance improvement.

IV. TEXTURE CLASSIFICATION ALGORITHM

The texture classification algorithm, which consists in the training phase and the classification phase [32], is detailed in the following.

Let us consider N texture classes. In the training phase, the feature vector $\boldsymbol{\Psi}_n$ of the nth class is generated as described in the following.

We first select M "typical" finite samples of the texture, that is M different parts all capturing the same relevant characteristics of the texture itself. In other words, these M samples are such to be considered M different realizations of the stationary

TABLE I
CLASSIFICATION PROCEDURE

Learning phase

1. $N = 15$ texture classes are considered. For each class, samples at angles $0°$, $30°$, $45°$, and $60°$ are taken into account. For each angle, $M = 30$ nonoverlapping samples, of dimension 64×64 pixels, are considered.

2. A texture sample is deconvolved according to the algorithm described in Section III to obtain, at convergence, the autocorrelation $\hat{R}_{\tilde{v}\tilde{v}}[k_1, k_2]$ of the deconvolved binary excitation.

3. The moment invariants in (11), (12) are computed on $\hat{R}_{\tilde{v}\tilde{v}}[k_1, k_2]$, and collected in the feature vector $\boldsymbol{\Psi}$.

4. Steps 2-3 are repeated for the number of training samples belonging to the considered class.

5. The feature vector for the class under examination is obtained by averaging the vectors obtained in the previous steps.

6. Steps 2-5 are repeated for the considered classes.

Classification phase

1. The texture sample feature vector is computed as described in steps 2-3 of the Learning phase.

2. The Bayes decision function D_n is evaluated for each class:

$$D_n = (\boldsymbol{\Psi} - \boldsymbol{\Psi}_n)^T C_n^{-1} (\boldsymbol{\Psi} - \boldsymbol{\Psi}_n) + \ln | C_n |$$

being C_n is the covariance matrix of the feature set for the texture class n.

3. The texture is assigned to the class k if $D_k < D_h$ for all $k \neq h$.

random field to which the texture is considered to belong to. Hence, their dimensions are chosen in order to capture a significant number of textons composing the texture.

Then, given a texture sample belonging to a specific class, its autocorrelation function $\hat{R}_{\tilde{v}\tilde{v}}[k_1, k_2]$, which summarizes the morphological structure of the texture prototype, is estimated as described in Section III. Then, the moment invariants ψ_j with $j = 1, 2, \ldots, 11$, are evaluated from the following sequences of samples, normalized to have unitary sum

$$p[k_1, k_2] = \frac{\hat{R}_{\tilde{v}\tilde{v}}[k_1, k_2]}{\sum_{k_1} \sum_{k_2} \hat{R}_{\tilde{v}\tilde{v}}[k_1, k_2]}$$
(13)

according to the formulas (11) and (12), where $m_{u,v}$ are numerically evaluated by discretizing (10) as follows:

$$m_{u,v} = \sum_{k_1} \sum_{k_2} k_1^u k_2^v\, p[k_1, k_2].$$
(14)

Moreover, a normalization of the axes x, y is performed in order to maintain the dynamic range of the moment values consistent for images of different dimensions [32].
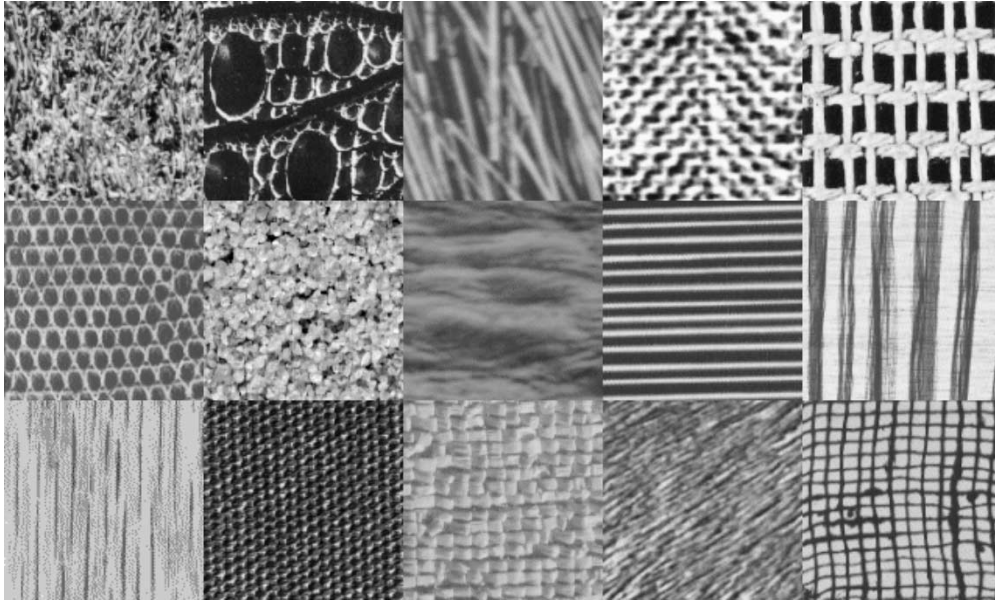
Fig. 3. Texture samples. From left to right. First row: D9 (Grass lawn), D10 (Crocodile skin), D15 (Straw), D17 (Herringbone weave), D20 (French canvas). Second row: D22 (Reptile skin), D29 (Beach sand), D37 (Water), D69 (Wood grain), D51 (Raffia woven). Third row: D68 (Wood grain), D77 (Cotton canvas), D84 (Raffia), D93 (Fur), D103 (Loose burlap).

TABLE II
PERCENTAGE OF CORRECT CLASSIFICATION FOR ROTATED TEXTURES SAMPLES (AT DIFFERENT ANGLES) USING MOMENT INVARIANTS (MI), THE WAVELET BASED METHOD [19] (WA), THE CO-OCCURRENCE MATRIX METHOD (CO)

	20°			70°			90°			120°			150°		
Texture	MI	Wa	Co	MI	Wa	Co	MI	Wa	Co	MI	Wa	Co	MI	Wa	Co
D9	80	85	70	80	75	80	75	70	90	80	75	85	80	75	85
D10	90	80	90	80	90	85	70	80	85	90	80	60	80	85	75
D15	70	80	75	90	65	60	75	75	90	75	75	70	80	70	55
D17	100	80	75	90	90	85	95	90	95	95	75	85	85	80	85
D20	100	95	95	100	80	80	100	70	85	95	80	80	85	95	85
D22	65	65	80	85	75	80	90	65	95	90	65	70	70	65	70
D29	95	70	90	90	70	85	90	40	95	95	70	85	95	60	85
D37	95	60	55	95	70	75	95	60	60	80	70	50	80	75	65
D49	80	90	35	85	80	35	85	90	80	70	90	60	70	90	40
D51	90	85	40	90	65	60	95	70	80	80	80	45	80	65	40
D68	90	70	55	80	70	40	85	80	80	95	80	50	95	70	50
D77	85	85	60	70	80	65	95	90	90	90	80	40	90	90	25
D84	85	45	50	85	60	45	80	50	60	75	50	45	80	55	40
D93	75	80	65	85	85	90	85	70	90	70	85	70	80	65	75
D103	75	65	25	75	60	75	85	85	75	80	65	30	80	65	30
Average	85	75.7	64	85.3	74.3	69.3	86.7	72.3	83.3	84	74.7	61.8	82	73.7	60.3

The M samples composing the class should be chosen to yield each a small model mismatching error $\epsilon[n_1, n_2]$; the effect of this latter on the features vector can be further reduced by averaging over the M samples composing the class, so finally obtaining the mean features vector $\mathbf{\Psi}_n$, which "globally" represents the texture class under analysis.

The classification phase consists in extracting from the texture under analysis the feature vector Ψ and comparing it with the representative feature map Ψ_n, $n = 1, 2, \ldots, N$, using a suitable distance metric criterion. The class at minimum distance with respect to the feature vector of the analyzed texture is decided to be the one the texture belongs to. In our approach the Bayes classifier [32], which is the optimal one when the texture features are assumed to have a Gaussian density distribution, has been used. The classification procedure is summarized in Table I.

V. CLASSIFICATION RESULTS

Specifically, for the classification experiments we have used a set of textures extracted from the Brodatz album [33]. Fifteen classes have been considered ($N = 15$); samples are shown in Fig. 3. For each class, the training set was composed by textures at angles $0°$, $30°$, $45°$, and $60°$. For each angle, $M = 30$ non-overlapping samples of dimension 64×64 pixels were considered.

The textures to classify, having size 128×128 with 256 gray levels, are obtained both from the not rotated images and after a rotation at angles $20°$, $70°$, $90°$, $120°$, and $150°$. The training textures set and the set composed by the textures to classify are disjoint. For each rotation angle, 20 classification tests have been performed. In Table II the percentages of correct classification, using the proposed method (MI), for rotated texture samples, are detailed both with respect to the different textures and to the different angles. The average percentages of correct classification for rotated texture are reported in Table III as well as the ones for not rotated textures.

For the sake of comparison, in the same tables, are reported the classification results obtained using the rotation-invariant wavelet method presented in [19] (Wa), and the ones obtained using the co-occurrence matrix method (Co).

A rough comparison between the time required for the features extraction, for the three considered methods, is also provided. The simulations have been performed using a nonoptimized C++ code on an Intel Pentium III 1.2 GHz based PC, hosting Windows 2000. Approximately, the time required for the extraction of a single set of features is 3.2 s/features vector for the proposed method, 0.4 s/features vector for the wavelet based method, and 1.1 s/features vector for the cooccurrence based method.

A. Performance on Noisy Images

The proposed classification algorithm has been tested for noisy images as well, since real images often contain random noise introduced either by the transmission through a noisy channel or by the imaging process. To this end, Gaussian random noise, with zero mean and variance depending on the wanted Signal to Noise Ratio (SNR), has been added to the images of the classification set, which were at angles $20°$, $70°$, $90°$, $120°$, and $150°$. Four SNR values have been considered (30 dB, 20 dB, 15 dB, 10 dB). Some examples of

TABLE III
AVERAGE PERCENTAGE OF CORRECT CLASSIFICATION FOR NON-ROTATED AND ROTATED TEXTURES SAMPLES USING MOMENT INVARIANTS (MI), THE WAVELET BASED METHOD [19] (WA), THE CO-OCCURRENCE MATRIX METHOD (CO)

Texture	Not rotated			Rotated		
	MI	Wa	Co	MI	Wa	Co
D9	82	88	90	79	76	82
D10	82	82	83	82	83	79
D15	85	78	80	78	73	70
D17	96	95	98	93	83	85
D20	98	90	92	96	83	85
D22	86	88	85	80	67	79
D29	96	70	98	93	62	88
D37	90	69	66	89	67	61
D49	84	86	92	78	88	50
D51	90	83	85	87	73	53
D68	92	85	93	89	74	55
D77	90	90	95	86	85	56
D84	86	87	75	81	52	48
D93	84	85	96	79	77	78
D103	84	72	85	79	68	47
Average	**88.3**	**83.2**	**87.7**	**84.6**	**74.1**	**67.7**

noisy images, at the considered values of SNR, are given in Fig. 4. In Table IV the percentages of correct classification for noisy rotated textures using moment invariants (MI), the wavelet based method [19] (Wa), and the co-occurrence matrix method (Co) are detailed. By comparing the results presented in Tables III and IV for the proposed method, we observe performance degradation starting at SNR values lower that 15 dB. Therefore, the proposed approach is highly reliable even with respect to severe degradations of the image.

VI. CONCLUSIONS

A novel model based rotation invariant texture classification scheme has been proposed. The texture is modeled as the output of a linear system fed by a binary excitation. The texture feature set extraction is constituted by two steps. The first one consists in estimating the binary excitation, which is assumed to efficiently represent the texture for classification purposes, by means of a blind deconvolution procedure. In the second step, a basis of moment invariants is employed to characterize the ACF of the binary excitation. The proposed procedure has been tested to be robust to rotations of the texture sample. Moreover the proposed scheme has been shown to maintain a high correct classification rate even in presence of noise.
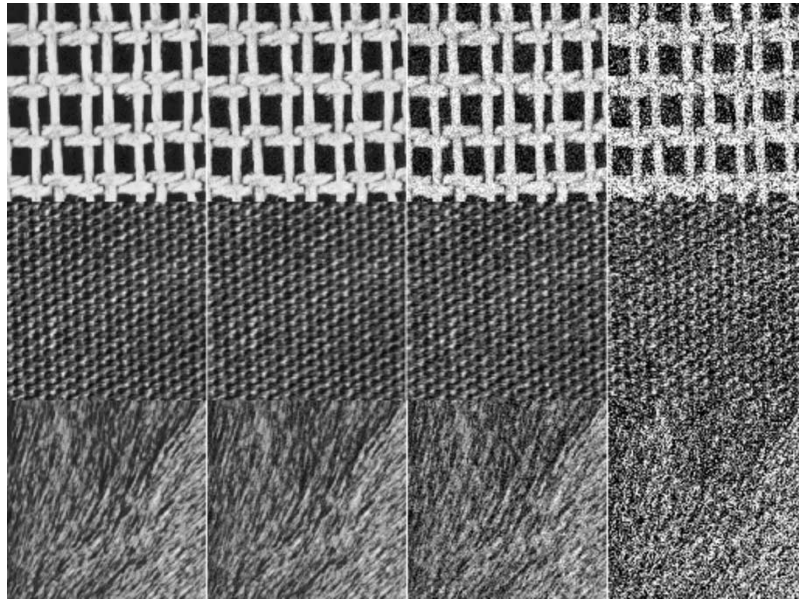
Fig. 4.   Texture samples (First row: D20. Second row: D77. Third row: D93) with added noise. From left to right: original texture, 30 dB, 20 dB, 10 dB.

TABLE IV
PERCENTAGE OF CORRECT CLASSIFICATION FOR NOISY ROTATED TEXTURES SAMPLES USING MOMENT INVARIANTS (MI), THE WAVELET BASED METHOD [19] (WA), AND THE CO-OCCURRENCE MATRIX METHOD (CO)

Texture	Rotated MI				Rotated Wa				Rotated Co			
	30 dB	20 dB	15 dB	10 dB	30 dB	20 dB	15 dB	10 dB	30 dB	20 dB	15 dB	10 dB
D9	79	79	75	59	76	76	72	58	82	80	80	70
D10	82	82	76	57	81	81	78	60	76	76	72	54
D15	78	74	74	66	73	73	73	64	70	68	64	55
D17	93	93	88	75	83	77	77	65	83	80	80	68
D20	96	96	96	77	83	83	83	70	83	80	78	65
D22	76	76	76	48	65	65	62	48	77	77	75	60
D29	88	88	88	60	60	58	58	40	85	85	81	64
D37	85	83	83	71	65	65	65	55	58	58	56	44
D49	78	78	74	58	88	88	86	70	50	50	48	38
D51	87	83	83	61	73	73	70	64	53	50	48	40
D68	85	85	85	58	74	74	72	58	55	55	55	41
D77	86	86	86	60	85	85	85	65	54	52	48	37
D84	81	77	77	52	50	45	45	37	48	48	45	34
D93	79	75	79	67	77	75	75	64	78	78	75	60
D103	79	79	79	67	68	66	63	50	45	45	45	36
Average	**83.5**	**82.3**	**81.3**	**62.4**	**73.4**	**72.3**	**70.9**	**57.8**	**66.5**	**65.4**	**63.3**	**51**

APPENDIX A

In this Appendix, the wavelet-based texture classification scheme described in [19] is briefly summarized. A three-level wavelet decomposition is performed on the texture sample, thus obtaining, for each level, image information of low horizontal frequency but low vertical frequency (LL channel), low horizontal frequency but high vertical frequency (LH channel),

high horizontal frequency but low vertical frequency (HL channel), high horizontal frequency but high vertical frequency (HH channel). Since the magnitudes of the wavelet coefficients in a particular channel assume high values for images with strong textural content in that channel, in [19] a feature vector comprising the average wavelet coefficients magnitude in each channel is considered. Specifically, the HH channels are not used as they can degrade the classification performance since they contain the majority of noise in the image. Moreover, rotation invariance is achieved by jointly combining the LH and HL channels, that is the diagonally opposite wavelet channels, at each level of decomposition in order to obtain a single feature. To summarize, a four dimensional feature vector comprising the average wavelet coefficients magnitude of the three combined LH and HL channels and of the remaining LL channel is used. Then a minimum distance classifier using the Mahalanobis distance [32] was employed to perform the texture classification.

APPENDIX B

Given an image $x[n_1, n_2]$ with a countable number of gray levels, for a fixed distance δ in direction α, the co-occurrence matrix $C^{(\delta,\alpha)}(i,j)$ is the bidimensional histogram measured on the image itself. The value assigned to the generic entry $C^{(\delta,\alpha)}(i,j)$ is the percentage of the pixel pairs $x_1 \overset{\text{def}}{=} x[n_1, n_2]$ and $x_2 \overset{\text{def}}{=} x[n_1 - m_1, n_2 - m_2]$, taken such that $\delta^2 = m_1^2 + m_2^2$, $\tan \alpha = m_2/m_1$, and with values $x_1 = i$ and $x_2 = j$ [2]. Eight common used co-occurrence features, which have been employed in this paper to design a classification method used for performance comparison with the proposed approach, are reported in Table V. It is worth mentioning that these features depend on the inter-pixel distance δ, which in our experiments has been set equal to 1, and on the angle α. Therefore in order to make the features rotation tolerant, in [2] it is suggested to quantize the orientation α in four directions: $0°$, $45°$, $90°$, and $135°$. In practice, the resulting values for the four directions are averaged out, thus providing robustness to rotation.

APPENDIX C

In [25], Hu introduced his seven moment invariants to the rotation of 2-D objects that have been widely used in the past for object recognition. They are reported in the following:

$$\phi_1 = \mu_{20} + \mu_{02},$$
$$\phi_2 = (\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2,$$
$$\phi_3 = (\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2,$$
$$\phi_4 = (\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2,$$
$$\phi_5 = (\mu_{30} - 3\mu_{12})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] +$$
$$+ (3\mu_{21} - \mu_{03})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2],$$
$$\phi_6 = (\mu_{20} - \mu_{02})[(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] +$$
$$+ 4\mu_{11}(\mu_{30} + \mu_{12})(\mu_{21} + \mu_{03}),$$
$$\phi_7 = (3\mu_{21} - \mu_{03})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] +$$
$$- (\mu_{30} - 3\mu_{12})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] \quad \text{(C.1)}$$

TABLE V
EIGHT COMMON FEATURES FROM A CO-OCCURRENCE MATRIX $C(i,j)$, WHERE $M_x = \sum_{i,j=0}^n iC(i,j)$, $M_y = \sum_{i,j=0}^n jC(i,j)$, $S_x(i) = \sum_{j=0}^n C(i,j)$, $S_y(i) = \sum_{i=0}^n C(i,j)$, $H_{xy} = -\sum_{i,j=0}^n C(i,j) \log(S_x(i)S_y(j))$, $H_x = -\sum_{i=0}^n S_x(i) \log S_x(i)$, AND $H_y = -\sum_{j=0}^n S_y(j) \log S_y(j)$

inertia	$C_1 = \sum_{i,j=0}^n (i-j)^2 C(i,j)$
total energy	$C_2 = \sum_{i,j=0}^n C(i,j)^2$
entropy	$C_3 = -\sum_{i,j=0}^n C(i,j) \log C(i,j)$
local homogeneity	$C_4 = \sum_{i,j=0}^n \frac{1}{1 + (i-j)^2} C(i,j)$
max. probability	$C_5 = \max_{i,j} C(i,j)$
cluster shade	$C_6 = \sum_{i,j=0}^n (i - M_x + j - M_y)^3 C(i,j)$
cluster prominence	$C_7 = \sum_{i,j=0}^n (i - M_x + j - M_y)^4 C(i,j)$
info. measure of correlation	$C_8 = \frac{C_3 - H_{xy}}{\max(H_x, H_y)}$

where

$$\mu_{pq} = \int_{-\infty}^{-\infty} \int_{-\infty}^{-\infty} (x - m_{1,0})^p (y - m_{0,1})^q f(x,y) \, dx dy \quad \text{(C.2)}$$

is the central moment of order (p, q) of the object $f(x, y)$.

As stated in [26], it is possible to verify that Hu's invariants (C.1) can be expressed in terms of (9) as follows:

$$\phi_1 = c_{11},$$
$$\phi_2 = c_{20}c_{02},$$
$$\phi_3 = c_{30}c_{03},$$
$$\phi_4 = c_{21}c_{12},$$
$$\phi_5 = \text{Re}(c_{30}c_{12}^2),$$
$$\phi_6 = \text{Re}(c_{20}c_{12}^2),$$
$$\phi_7 = \text{Im}(c_{30}c_{12}^2). \quad \text{(C.3)}$$

However, Hu's invariants are dependent, since it is easy to prove that

$$\phi_3 = \frac{\phi_5^2 + \phi_7^2}{\phi_4^3}. \quad \text{(C.4)}$$

Moreover, they do not represent a basis for the second and third order invariants.

REFERENCES

[1] P. C. Chen and T. Pavlidis, "Segmentation by texture using correlation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-5, pp. 64–69, Jan. 1983.

[2] R. M. Haralik, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-3, pp. 610–621, 1973.

[3] T. Randen and J. H. Husøy, "Filtering for texture classification: A comparative study," *IEEE Pattern Anal. Machine Intell.*, vol. 21, pp. 291–310, Apr. 1999.

[4] M. Unser and M. Eden, "Multiresolution feature extraction and selection for texture segmentation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, pp. 717–728, 1990.

[5] A. C. Bovik, M. Clark, and W. S. Geisler, "Multichannel texture analysis using localized spatial filters," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, pp. 55–73, 1990.

[6] A. Teuner, O. Pichler, and B. J. Hostica, "Unsupervised texture segmentation of images using tuned matched Gabor filters," *IEEE Trans. Image Processing*, vol. 4, pp. 863–870, June 1995.

[7] B. S. Manjunath and W. Y. Ma, "Texture features for browsing and retrieval of image data," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 18, pp. 837–842, 1996.

[8] T. Chang and C.-C. J. Kuo, "Texture analysis and classification with tree-structured wavelet domain," *IEEE Trans. Image Processing*, vol. 2, pp. 429–441, Oct. 1993.

[9] M. Unser, "Texture classification and segmentation using wavelet frames," *IEEE Trans. Image Processing*, vol. 4, no. 11, pp. 1549–1560, Nov. 1995.

[10] G. Van de Wouwer, P. Scheunders, and D. Van Dyck, "Statistical texture characterization from discrete wavelet representation," *IEEE Trans. Image Processing*, vol. 8, pp. 592–598, Apr. 1999.

[11] G. R. Cross and A. K. Jain, "Markov random field texture models," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-5, pp. 25–39, Jan. 1983.

[12] R. Chellappa and S. Chatterjee, "Classification of texture using Gaussian Markov random fields," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 959–963, Apr. 1985.

[13] J. A. Cadzow, D. M. Wilkes, R. A. Peter II, and X. Li, "Image texture synthesis-by-analysis using moving average models," *IEEE Trans. Aerospace Electr. Syst.*, vol. 29, pp. 1110–1121, Oct. 1993.

[14] R. Chellappa and R. L. Kashyap, "Texture synthesis using 2-D non-causal autoregressive models," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 33, pp. 194–203, Feb. 1985.

[15] T. E. Hall and G. B. Giannakis, "Bispectral analysis and model validation of texture images," *IEEE Trans. Image Processing*, vol. 4, pp. 996–1009, July 1995.

[16] L. M. Kaplan, "Extended fractal analysis for texture classification and segmentation," *IEEE Trans. Image Processing*, vol. 8, pp. 1572–1585, Nov. 1999.

[17] F. Liu and R. W. Picard, "Periodicity, directionality, and randomness, wold features for image modeling and retrieval," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 18, July 1996.

[18] P. Campisi, G. Jacovitti, and A. Neri, "Optimized wold-like decomposition of 2D random fields," in *Proc. Eur. Sig. Proc. Conf. EUSPICO '98*, Island of Rhodes, Greece, Sept. 8–11, 1998, pp. 1681–1684.

[19] R. Porter and N. Canagarajah, "Robust rotation-invariant texture classification: Wavelet, Gabor filter and GMRF based schemes," *Proc. Inst. Elect. Eng.*, vol. 144, no. 3, pp. 180–188, June 1997.

[20] G. M. Haley and B. S. Manjunath, "Rotation-invariant texture classification using a complete space-frequency model," *IEEE Trans. Image Processing*, vol. 8, pp. 255–269, Feb. 1999.

[21] G. Jacovitti, A. Neri, and G. Scarano, "Texture synthesis-by-analysis with hard limited Gaussian process," *IEEE Trans. Image Processing*, vol. 7, pp. 1615–1621, Nov. 1998.

[22] P. Campisi, A. Neri, and G. Scarano, "Reduced complexity modeling and reproduction of colored textures," *IEEE Trans. Image Processing*, vol. 9, pp. 510–518, Mar. 2000.

[23] P. Campisi, A. Neri, and D. Hatzinakos, "A perceptually preattentive lossless, model based, texture compression technique," *IEEE Trans. Image Processing*, vol. 9, pp. 1325–1336, Aug. 2000.

[24] P. Campisi and G. Scarano, "A multiresolution approach for texture synthesis using the circular harmonic functions," *IEEE Trans. Image Processing*, vol. 11, pp. 37–51, Jan. 2002.

[25] M.-K. Hu, "Visual pattern recognition by moment invariants," *IRE Trans. Inform. Theory*, vol. IT-8, pp. 179–187, Feb. 1962.

[26] J. Flusser, "On the independence of rotation moment invariants," *Pattern Recognit.*, vol. 33, pp. 1405–1410, 2000.

[27] S. Curtis, S. Shitz, and V. Oppenheim, "Reconstruction of nonperiodic two dimensional signals from zero-crossing," *IEEE Trans. Acoust., Speech, Signal Proc.*, vol. 35, pp. 890–893, June 1987.

[28] B. Julesz and R. Bergen, "Textons, the fundamental elements in preattentive vision and perception of textures," *Bell Syst. Tech. J.*, vol. 62, Jul-Aug. 1983.

[29] R. Godfrey and F. Rocca, "Zero-memory nonlinear deconvolution," *Geophys. Prosp.*, vol. 29, pp. 189–228, 1981.

[30] S. O. Belkasim, K. J. Breeding, and M. Ahmadi, "Pattern recognition with moment invariants: A comparative study and new results," *Pattern Recognit.*, vol. 24, pp. 1117–1138, 1991.

[31] S. X. Liao and M. Pawlak, "On image analysis by moments," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 18, pp. 254–266, 1996.

[32] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*. New York: Academic, 1999.

[33] T. Brodatz, *Textures: A Photographic Album for Artists and Designers*. New York: Dover, 1966.

[34] P. Campisi, A. Neri, and G. Scarano, "Model based rotation-invariant texture classification," in *IEEE Int. Conf. on Image Proc. 2002 (ICIP 2002)*, Rochester, Y, Sept. 22–25, 2002.

Patrizio Campisi (M'99) received the "Laurea" degree in electrical engineering, summa cum laude from the University of Roma "La Sapienza," Roma, Italy, and the Ph.D. degree in electrical engineering from the University of Roma "Roma Tre," Roma, Italy, in 1995 and 1999, respectively.

He is an Assistant Professor with the Department of "Elettronica Applicata", University of Roma "Roma Tre," where he has been also lecturer for the graduate course "Signal Theory" since 1998. From September 1997 until April 1998, he was a visiting research associate with the Communication Laboratory, University of Toronto, Toronto, ON, Canada, and from July 2000 until December 2000, he was a post doctoral fellow with the same laboratory. From October 1999 to October 2001, he held a post doctoral position at the University of Roma "Roma Tre." From March 2003 until June 2003 he was a visiting researcher at the Beckman Institute, University of Illinois at Urbana-Champaign, IL, USA. His research interests are in the area of digital signal and image processing with applications to multimedia communications.

Dr. Campisi has been a member of the Technical Committees of several IEEE conferences (ICME 2002, ICME 2003, ISPA 2003, MMSP 2004). He was the organizer for the special session on "Texture analysis and synthesis" for the IEEE International Symposium on Image and Signal Processing and Analysis 2003 (ISPA 2003). He is a Member of the IEEE Communications and Signal Processing Society and the Italian Professional Engineers association.

Alessandro Neri (M'82) was born in Viterbo, Italy, in 1954. He received the doctoral degree in electronic engineering from the University of Rome "La Sapienza," Rome, Italy, in 1977.

In 1978, he joined the Research and Development Department of Contraves Italiana S.p.A., where he gained a specific expertise in the field of radar signal processing and in applied detection and estimation theory, becoming the chief of the advanced systems group. In 1987, he joined the INFOCOM Department, University of Rome "La Sapienza," as an Associate Professor of signal and information theory. In November 1992, he joined the Electronic Engineering Department of the University of Rome III as Associate Professor of electrical communications and became Full Professor of telecommunications in September 2001. Since 1992, he has been responsible for coordination and management of research and teaching activities in the telecommunication field at the University of Rome III. Since 1998, he has also been responsible for planning and design activities related to university campus telecommunication systems and services. His research activity has mainly been focused on information theory, signal theory, and signal and image processing and their applications to both telecommunications systems and remote sensing. His current research is focused on third- and fourth-generation cellular systems and multimedia communications. Since 1997, he has also been involved in several research programs connected with the use of information technologies in distance learning.

Giampiero Panci was born in Palestrina, Roma, Italy. He received the "Laurea" degree in Telecommunications Engineering in 1996 and the Ph.D. in Communication and Information Theory in 2000, both from Università "La Sapienza" of Rome, Italy.

He has worked in the area of signal processing for communication and radar systems and in the 1999 he was involved in the design of the Signum Coded Sar Processor for the Shuttle Radar Topographic Mission. He currently holds an Associate Research position at the Dipartimento di Scienza e Tecnica dell'Informazione e della Comunicazione (INFO-COM) of the University "La Sapienza" of Rome, where he is also Lecturer for a graduate course in Communications. His current research interests include statistical signal processing, blind identification and equalization and array processing.

Gaetano Scarano (M'00) was born in Campobasso, Italy. He received the "Laurea" degree in Electronic Engineering from Università "La Sapienza" di Roma, Italy, in 1982.

In 1982, he joined the Istituto di Acustica of the Consiglio Nazionale delle Ricerche, Roma, Italy, as Associate Researcher. Since 1988 he has been teaching Digital Signal Processing at University of Perugia, Perugia, Italy where in 1991 he became Associate Professor of Signal Theory. In 1992, he joined the Dipartimento di Scienza e Tecnica dell'Informazione e della Comunicazione (INFO-COM) of the Università "La Sapienza" di Roma, first as Associate Professor of Image Processing, then as Professor of Signal Theory. His research interests lie in the area of signal and image processing, communications, estimation and detection theory, and include channel equalization and estimation, spread spectrum communications and CDMA, texture synthesis and image restoration.

Mr. Scarano has served as Associate Editor of the IEEE SIGNAL PROCESSING LETTERS.