Synthesis–by–Analysis of Complex Textures

Patrizio Campisi, Alessandro Neri*, Gaetano Scarano***

*Electronic Engineering Dept., University of Rome III
via della Vasca Navale 84, I-00146 Rome, Italy
Tel:+39.6.5517.7004, Fax:+39.6.5579.078,
email: neri@infocom.ing.uniroma1.it

**INFOCOM Dept., University of Rome "La Sapienza"
via Eudossiana 18, I-00184 Rome, Italy
Tel:+39.6.4458.5500, Fax:+39.6.4873.300,
email: gaetano@infocom.ing.uniroma1.it

ABSTRACT

A technique for unsupervised texture synthesis by analysis is presented. It is based on a stochastic approximation of a textured field obtained by nonlinearly transforming a complex white Gaussian random field. The nonlinear transformation is constituted by two linear filters connected by a complex hard-limiter. The identification of the texture model is performed by means of a Bussgang blind deconvolution algorithm exploiting a generalization to the complex case of the Van Vleck rule. After a theoretical discussion of the method typical examples are provided.

1 Introduction

Generation of real images resembling some given prototypes is of great interest in many applications, including synthesis of textures intended for very low bit rate video coders and performance assessment by simulation of both computer vision algorithms and remote sensing systems. However, *electromagnetic* images produced by microwave sensors (SAR) as well as *acoustic coherent* images are usually constituted by a pair of (real) images produced by the in–phase and in–quadrature channels of the receiver, better represented in a complex framework. On the other hand, recent advances in signal theory (*e.g.* scale and orientation subbands wavelets decompositions, steerable pyramids [1], circular harmonic wavelets [2]) have shown the worthiness of a complex approach for the representation of real images.

Here, we address the problem of synthesis of complex textures, by extending the nonlinear unsupervised texture synthesis by analysis recently proposed in [3].

While, in principle, current texture synthesis techniques could separately be applied to the real and imaginary components of the complex image, such an approach would not exploit the inter–channel correlation.

To overcome this limitation, linear techniques employing rational models (MA,AR,ARMA) excited by independent identically distributed (i.i.d) random fields could be also adapted [4, 5, 6], as well as other techniques not based on particular statistical assuptions or model [7]. However, the linear model based identification procedure presents an inherent instability, due to the i.i.d. assumption of the excitation, so that it is often necessary the aid of human intervention (see for instance [5]).

In our approach, the synthetic complex texture is obtained by filtering a *phase-only* image, which retains most of the morphological properties of the original texture. This in turn implies that, usually, the phase-only image is spatially correlated. We note that the phase of a complex image includes the zero–crossing information of both the real and the imaginary parts that, in principle, is sufficient for a complete recovery of these components [8]. The continuous phase information assures a better overall quality of the synthesis due to a lower sensitivity to noise w.r.t. the zero–crossing information alone.

The proposed method relies on the statistical equivalence, up to the second–order (autocorrelation), between the phase of the (filtered) reference texture and the phase of the synthetic one. This choice is motivated by the simplicity of the generation of a phase-only image with desired autocorrelation. In fact, this image is just the phase of a complex Gaussian random field whose autocorrelation is determined by the generalization of the Van Vleck rule to the complex case (hyphergeometric law) [9].

In summary, the overall canonical model is a nonlinear scheme composed of a linear system excited by a white Gaussian random field, cascaded with a phase–extractor and a final linear system, as shown in Fig.1. This model arises

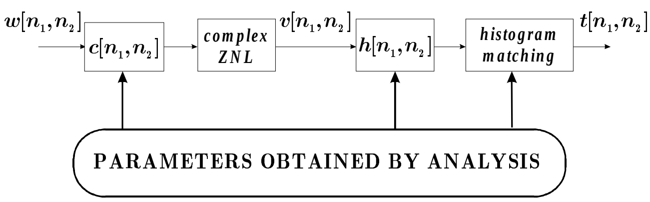

Figure 1: *Synthesis*

from the analysis of the texture performed using a complex Bussgang deconvolution [3, 10]. A general scheme of the Bussgang deconvolution algorithm is shown in Fig.2, where, at convergence, the process $z[n_1, n_2]$ can be suitably replaced by the process $v[n_1, n_2]$, which is actually generated in the synthesis. By choosing the complex zero–

memory nonlinearity (CZNL) in a proper fashion, we can use Gaussian excitations in the synthesis phase, obtaining a texture model as the output of a linear system driven by a non–iid, non–Gaussian random field.

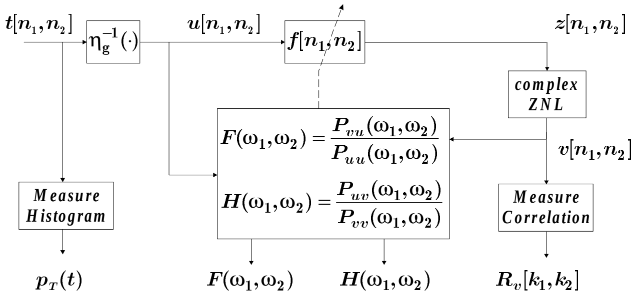

Figure 2: *Analysis*

For real textures, the *signum* function has been employed in [3], where several textures have been succesfully mimed by using the Bussgang approach.

Here, we propose the use of a complex extension of the *signum* function, namely a nonlinearity which retains only the phase, setting the magnitude to a constant value. (This corresponds to choose ± 1 for real values.) In this way, by invoking the so–called *hypergeometric law* described in [9], we can determine the autocorrelation of the Gaussian process which generates the non–Gaussian process $v[n_1, n_2]$ through the complex hard–limiting nonlinearity (see Fig.1).

In summary, the Bussgang approach allows for simple parametrization of the texture model; in fact, while the first order statistics of the non–Gaussian excitation $v[n_1, n_2]$ is determined by the selected CZNL, its second–order statistics can be easily controlled on the basis of the autocorrelation of the observed reference texture, in order to mimic the spatial characteristics in a more detailed fashion.

In this contribution, we describe the generalization to the complex case of the synthesis–by–analysis technique presented in [3]. Then, results of texture synthesis–by–analysis are shown by means of some classical examples.

2 THE IDENTIFICATION METHOD

As illustrated in Fig.1, the proposed method assumes that a complex textured image $\mathbf{t} = \{t[n_1, n_2]\}$ can be reconstructed by a phase only image $\mathbf{v} = \{v[n_1, n_2]\}$ by means of an possibly nonminimum phase linear filter $h[n_1, n_2]$ cascaded with a memoryless non linearity $\eta(\cdot)$, *i.e.* the following factorization applies:

$$t[n_1, n_2] = \eta(h[n_1, n_2] * v[n_1, n_2]) \quad (1)$$

with the constraint

$$|v[n_1, n_2]| = 1 \quad (2)$$

We note that, as enlightened by psycophisiological experiments, in order to achieve a good subjective quality, synthetic real images must present the same pointwise (first order) statistics and the same autocorrelation of the original textures. In our scheme the memoryless nonlinearity $\eta(\cdot)$ accomplishes the task of the histogram matching between the reference texture and its reconstructed version.

Since infinitely many factorizations are possible, including the trivial solution with a phase-only image constituted by a unitary pulse, in order to assure a generalization capability during the synthesis phase, we interprete the reference texture $\mathbf{t}$ as a typical realization of a random field, and impose that the factorization (1) applies to all the typical realizations of that random field. To this aim, at first we partition the reference image into M smaller blocks, regarded now as different realizations, and we impose the constraint that the reconstruction filters and the memoryless non linearities of any block coincide.

The identification procedure here employed, represents an extention to the complex 2-D case of the blind deconvolution method employed for blind equalization of communication channels excited by binary data sequences [11, 16]. This procedure can be summarized as follows. Let $\mathbf{t}_k$ be the 1-D array obtained from the k-th block of the reference texture by rearranging the columns in lexicographic order, let H be the linear operator associated to the filter $\mathbf{h} = h[n_1, n_2]$ and $\eta_{\mathbf{g}}(\cdot)$ a set of invertible zero-memory non linearities, depending on a certain number of parameters collected in the vector $\mathbf{g}$, then $\mathbf{t}_k$ can be written as:

$$\mathbf{t}_k = \eta_{\mathbf{g}}\left(H\mathbf{v}_k\right) + \mathbf{n}_k; \quad k = 1, \cdots, M \quad (3)$$

where $\mathbf{v}_k$ is the unitary magnitude excitation field and $\mathbf{n}_k$ is a residual random field tacking into account eventual model mismatching.

Then, we look for the triplet $\left(\tilde{\mathbf{v}}, \tilde{\mathbf{h}}, \tilde{\mathbf{g}}\right)$ that maximizes the log-likelihood of the given sample textures $\mathbf{t}_k$, *i.e.*:

$$\left(\tilde{\mathbf{v}}, \tilde{\mathbf{h}}, \tilde{\mathbf{g}}\right) = \arg\mathrm{Max}\sum_k \ln p(\mathbf{t}_k/\mathbf{v}; \mathbf{h}, \mathbf{g})$$

Due to the spatial dependence introduced by the filter, the optimization problem associated to the above functional is extremely difficult and we resort to a close approximation obtained by applying a recursive estimation scheme to the weaker criterion [12]:

$$\tilde{\mathbf{v}} = \arg\mathrm{Max}\sum_k \ln p(\mathbf{z}_k/\mathbf{v}; \tilde{\mathbf{f}}, \tilde{\mathbf{g}})$$
$$\tilde{\mathbf{f}} = \arg\mathrm{Max}\sum_k \ln p(\mathbf{z}_k/\tilde{\mathbf{v}}; \mathbf{f}, \tilde{\mathbf{g}})$$
$$\tilde{\mathbf{g}} = \arg\mathrm{Max}\sum_k \ln p(\mathbf{z}_k/\tilde{\mathbf{v}}; \tilde{\mathbf{f}}, \mathbf{g})$$

where $\mathbf{f} = \{f[n_1, n_2]\}$ are the coefficients of the inverse filter $F \stackrel{def}{=} H^{-1}$, $\mathbf{z}_k \stackrel{def}{=} F\left[\eta_{\mathbf{g}}^{-1}(\mathbf{t}_k)\right]$, and $\eta_{\mathbf{g}}^{-1}(\cdot)$ the inverse of $\eta_{\mathbf{g}}(\cdot)$.

Setting $\mathbf{u}_k = \eta_{\mathbf{g}}^{-1}(\mathbf{t}_k)$, and denoting with $D_{\mathbf{u}_k}$ the linear operator corresponding to the filter with impulse response equal to $\mathbf{u}_k$, the observation $\mathbf{z}_k$ can be written as follows

$$\mathbf{z}_k = D_{\mathbf{u}_k} \cdot \mathbf{f} = \tilde{\mathbf{v}}_k + \mathbf{n}'_k$$

where $\mathbf{n}'_k$ is the realization of a random noise representing the deconvolution error. Assuming as in [13] that the deconvolution error tends to a Gaussian white noise uncorrelated with the useful signal and equating to zero the gradient of $\sum_k \ln p(\mathbf{z}_k/\tilde{\mathbf{v}};\mathbf{f},\tilde{\mathbf{g}})$ with respect to $\mathbf{f}$, we obtain

$$\tilde{\mathbf{f}} = \left[\sum_k (D^{\mathrm{T}}_{\tilde{\mathbf{u}}_k}\cdot D_{\tilde{\mathbf{u}}_k})\right]^{-1}\left[\sum_k D^{\mathrm{T}}_{\tilde{\mathbf{u}}_k}\cdot \mathbf{v}_k\right]$$
$$\stackrel{def}{=} [\mathbf{R}_{\tilde{\mathbf{u}}}]^{-1}\,\mathbf{R}_{\tilde{\mathbf{v}}\tilde{\mathbf{u}}}$$

where $\mathbf{R}_{\tilde{\mathbf{u}}}$ is the sample correlation matrix of $\tilde{\mathbf{u}}$ and $\mathbf{R}_{\tilde{\mathbf{v}}\tilde{\mathbf{u}}}$ is the correlation matrix of $\tilde{\mathbf{v}}$ and $\tilde{\mathbf{u}}$.

On the other hand, near convergence the noise $\mathbf{n}'_k$ becomes isotropic and the optimal estimate of the unitary magnitude excitation reduces to a classical likelihood estimate, so that

$$\tilde{v}[n_1,n_2] = \arg(z[n_1,n_2])$$

Finally, it can be verified that the estimate $\tilde{\mathbf{g}}$ of $\mathbf{g}$ reduces to the least square (parametric) fit of $\tilde{\mathbf{v}}_k$, and can, in practice, be approximated by means of an histogram matching between $\tilde{\mathbf{v}}_k$ and $\tilde{F}^{-1}[\mathbf{t}_k]$.

In summary, at each iteration, we first modify the histogram of the texture sample in accordance to the non-linearity $\eta_{\mathbf{g}}^{-1}$ estimated in the previous iteration, then we extract the phase of the image obtained by deconvolving the equalized image through the estimated inverse filter. The phase-only image is then employed to update both the filter and the non-linearity. The algorithm stops when the non-linearity and the filter do not change, apart a possible scale factor.

As a second step of the identification procedure we have to estimate the impulse response $c[n_1,n_2]$ of the first stage of the model. To this aim we recall that given a stationary, ergodic, zero mean, Gaussian random field $\mathbf{s} = \{s[n_1,n_2]\}$ with autocorrelation function $\sigma_s^2\rho_s[k_1,k_2]$, the normalized autocorrelation function (a.c.f.) $\rho_v[k_1,k_2]$ of the image $\mathbf{v}$ obtained by retaining only its phase is related to the normalized a.c.f. $\rho_s[k_1,k_2]$ by the hypergeometric law [9]:

$$\rho_v[k_1,k_2] = \frac{\pi}{4}\rho_s[k_1,k_2]\cdot {}_2F_1\left(\frac{1}{2},\frac{1}{2};2;|\rho_s[k_1,k_2]|^2\right) \qquad (4)$$

where ${}_2F_1(\cdot,\cdot;\cdot;\cdot)$ is the Gaussian hyphergeometric function. Thus, we determine the impulse response $c[n_1,n_2]$ by constraining its autocorrelation to satisfy the hypergeometric law (4).

The synthesis is schematized in Fig.1, where $w[n_1,n_2]$ is a realization of a complex white Gaussian field, the CZNL is a phase exctractor and the nonlinearity $\eta(\cdot)$ is a simple histogram matching. The analysis provides also the filters parameters; the colouring filter $c[n_1,n_2]$ through the use of the hypergeometric law, and the shaping filter $h[n_1,n_2]$ as the best (in a minimum mean square error sense) filter transforming the (almost phase–only) image $z[n_1,n_2]$ into the (histogram matched) texture $u[n_1,n_2]$ (see Fig.2).

3 RESULTS AND CONCLUSION

To show the capabilities of the proposed method, real textures have been synthetized using a complex representation. Exploiting the simmetries in the Fourier spectrum of real 2–D signals, several definitions of 2–D *analytical* signal are possible (see for instance [14, 15]). Here, we have adopted the following definition of 2–D Hilbert transform [15]:

$$\mathcal{H}\{x[n_1,n_2]\} = \frac{1}{2}(h_1[n_1,n_2]+h_2[n_1,n_2])*x[n_1,n_2] \qquad (5)$$

where $h_i[n_1,n_2] = \dfrac{2}{\pi n}\sin^2\dfrac{\pi}{2}n_i$ are the impulse responses of the 1–D Hilbert transformers along the horizontal $(i=1)$ and vertical $(i=2)$ directions, respectively. Note that, using the definition (5), the Hilbert transform of a real signal is a real signal.

The complex analytical signal $x[n_1,n_2]+j\mathcal{H}\{x[n_1,n_2]\}$ has Fourier spectrum vanishing in the III quadrant (both ω_1 and ω_2 negative), twice the Fourier transform of the real 2–D signal $x[n_1,n_2]$ in the I quadrant (both ω_1 and ω_2 positive) and equal to the Fourier transform of $x[n_1,n_2]$ in the II and IV quadrants (ω_1 and ω_2 of opposite sign). For the class of analytical textures, the excitation noise in Fig.1 is an *analytical* Gaussian random field rather than white.

The performance of the whole synthesis–by–analysis procedure is illustrated in Figs.3,4 and 5 for typical textures extracted from Brodatz's collection. In Fig.3 the following real textures are displayed : D93–Fur (top–left), D84–Raffia (top–right), D68–Wood (bottom–left) and D77–Cotton Canvas (bottom–right). From Fig.4 we see that most of the morphological structure is retained by the (deconvolved) phase–only textures. Note the capability of the procedure to copy even structured textures, as it can be seen from Fig.5 where the synthetic textures are shown.

In conclusion, the method is conceptually similar to other techniques based on linear modeling. The basic difference is that we have explicitly inserted in the excitation process visually relevant morphological information. This in turn implies that the excitation cannot be simply modeled by an i.i.d. random field. Using the phase–only information allows for simple generation of such non i.i.d. fields, through the generalization of the Van Vleck (arcsinus) law to the complex case, known as the hyphergeometric law.

Moreover, the use of a complex representation of real signals allows for taking into account also medium–term correlation carried out by the 2–D Hilbert transformation. Another advantage of the complex representation is constituted by a better extraction of the zero–crossing of the texture using the continous phase.

Current research is devoted to exploit other complex representations as well as the use of other simple CZNL's in the proposed procedure.

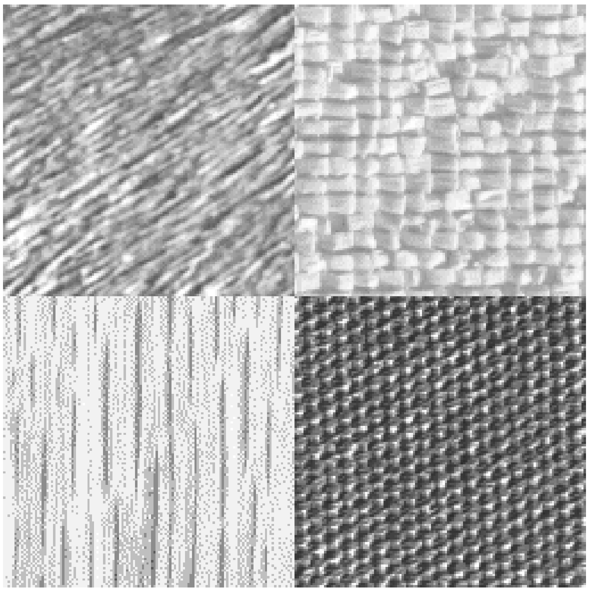

Figure 3: *Real textures from Brodatz's collection: D93–Fur (top–left), D84–Raffia (top–right), D68–Wood (bottom–left) and D77–Cotton Canvas (bottom–right).*

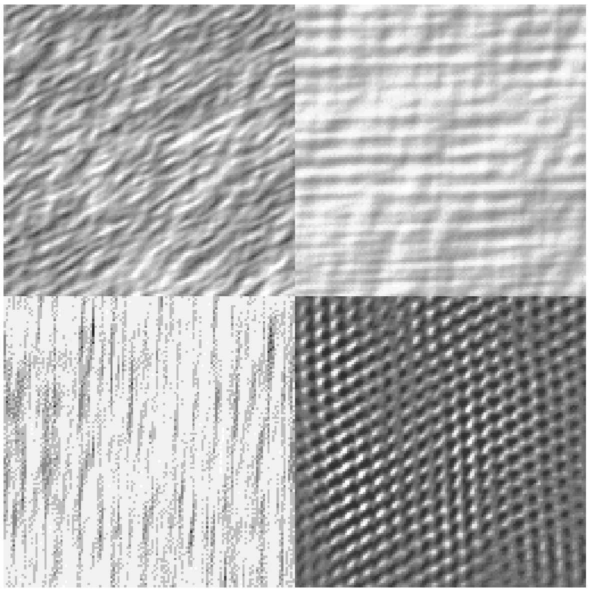

Figure 5: *Textures of Fig.3 synthetized using complex analytical textures.*

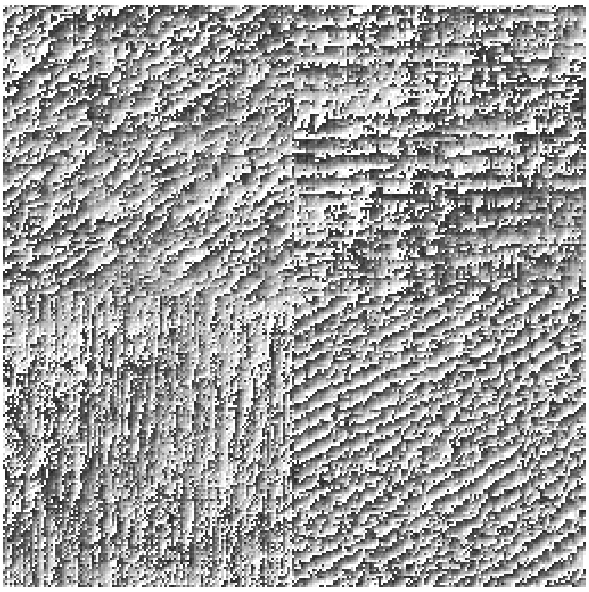

Figure 4: *Phase–only textures drawn from analytical textures associated to textures of Fig.3.*

REFERENCES

[1] E.P. Simoncelli, W.T. Freeman, "The steerable pyramid: a flexible architeture for multi–scale derivative computation", *IEEE Int. Conf. on Image Processing*, Washington (DC), USA, October 1995.

[2] G. Jacovitti, A. Neri, "Multiscale image features analysis with circular harmonic wavelets", *Wavelet Applications in Signal and Image Processing III, A.F. Laine, M.A. Unser eds, Proc. SPIE 2569*, 1995.

[3] A. Neri, G. Jacovitti, G. Scarano, "Texture Synthesis with Bussgang Techniques", *IEEE Int. Conf. on Image Processing*, Washington (DC), USA, October 1995.

[4] J.K. Tugnait, "Estimation of linear parametric models of non–Gaussian discrete random fields with applications to texture synthesis", *IEEE Trans. Image Proc.*, vol. 3, no. 2, March 1994.

[5] J.A. Cadzow, D.M. Wilkes, R.A. Peter II, X. Li, "Image texture synthesis–by–analysis using moving average models", *IEEE Trans. on Aerospace and Electronic Systems*, vol. 29, no. 4, Oct. 1993.

[6] T.E. Hall, G.B. Giannakis,"Bispectral analysis and model validation of texture image" *IEEE Trans. Image Proc.*, vol. 4, no. 7, July, 1995.

[7] D.J. Heeger, J.R. Bergen, "Pyramid-Based Texture Analysis/Synthesis", *IEEE ICIP*, Washington (DC), USA, 1995.

[8] S. Curtis, S. Shitz, V. Oppenheim, "Reconstuction of non-periodic two dimensional signals from zero-crossing", *IEEE Trans. ASSP*, vol. 35, 1987.

[9] G. Jacovitti, A. Neri, "Estimation of the autocorrelation function of complex Gaussian stationary processes by amplitude clipped signals", *IEEE Trans. Information Theory*, vol.40, no.1, January 1994.

[10] A. Neri, G. Scarano, G. Jacovitti, "Bayesian iterative method for blind deconvolution", *Adaptive Signal Processing*, S. Haykin ed., *Proc. SPIE 1565*, 1991.

[11] S.A. Billings, S.Y. Fakhoury, "Identification of a class of non-linear system using correlation analysis", *Proc. Inst. Elec. Eng.*, vol. 125, no. 7, 1978.

[12] S. Lakshmanan, H. Derin, "Simultaneous parameter estimation and segmentation of Gibbs random fields using simulated annealing", *IEEE Trans. PAMI*, vol. 11, Aug. 1989.

[13] R. Godfrey, F. Rocca, "Zero memory non-linear deconvolution", *Geophisical Prospecting*, vol. 29, 1981.

[14] F. Peyrin, Y.M. Zhu, R. Goutte, "Extension of the notion of analytic signal for multidimensional signals: Application to images", **Signal Processing III: Theories and Applications**, I.T. Young and alii eds., North–Holland, Amsterdam, 1986.

[15] Y.M. Zhu, F. Peyrin, R. Goutte, "The use of a two–dimensional Hilbert transform for Wigner analysis of 2–dimensional real signals", *Sig. Proc.*, vol. 19, 1990.

[16] Y. Sato, "A method of self-recovering equalization for multi-level amplitude modulation", *IEEE Trans. Commun.*, vol. 23, June 1975.