SYNTHESE ET ANALYSE DE TEXTURE UTILISANT DES FILTRES A MINIMUM D'INCERTITUDE

TEXTURE SYNTHESIS AND ANALYSIS USING MINIMUM UNCERTAINTY FILTERS

Roland Wilson

Department of Electrical & Electronic Engineering, The University of Aston in Birmingham

RESUME

Un modèle nouveau de texture est proposé, fondé sur l'idée d'un processus vectoriel des impulses qui pousse un opérateur vectorel linéaire, duquel les composants représentent régions distincts du domaine de fréquence spatiale. La méthode a flexibilité et a de la ressemblance avec le système visuel.

SUMMARY

A new texture model is proposed, based on the notion of a vector impulse process driving a vector filter, whose components are chosen to represent optimally disjoint regions of the spatial frequency domain. The method is shown to have great flexibility and its relation to the visual system is discussed.

SYNTHESE ET ANALYSE DE TEXTURE UTILISANT DES FILTRES A MINIMUM D'INCERTITUDE

TEXTURE SYNTHESIS AND ANALYSIS USING MINIMUM UNCERTAINTY FILTERS

(Roland Wilson)

1. Introduction

The problem of texture description is one that has received an enormous amount of attention in the literature (e.g. refs. 1-7) and yet remains a topic of debate. The debate centres around both what precisely is meant by 'texture' and how such regions should be generated and analysed.

Of the methods proposed so far, statistical techniques predominate. The most common stochastic models of texture are either linear, for example using auto-regressive (AR) or auto-regressive-moving-average (ARMA) models, [4], [5], or non-linear, among which Markovian models tend to be favoured (see e.g. [6], [7]). These are natural choices, offering a range of model complexities, from simple 1-d linear causal models to 2-d models with control of the n-tuple probability densities. While the more complex, nonlinear models offer generality, it is by no means clear that such generality is necessary, in view of the findings of Gagalowicz and others, [1], [2], which suggest that 2nd order spatial correlation and first order statistics may be sufficient, at least as far as the human visual system is concerned.

The common element in all of these models is of a system driven by a suitable noise process. This is summarised in the general model of Faugeras et al [2]. While the model described here is similar in general form to that of [2], it has some important differences from the implementations described by those authors. The principal differences are in the choice of driving process and the choice of filters. Rather than use a continuous process, an impulse process is used. This gives a greater range of image types, ranging from isolated features and periodic patterns to textures indistinguishable from those produced by a continuous process. Rather than use a single filter with a set of variable parameters, such as in an ARMA model, a set of filters is used, each filter being driven by a separate noise process. The filters are selected to give a maximally efficient division of the frequency domain for a given spatial window. The advantage of this method is its flexibility: the filters can be so chosen as to synthesise any desired microstructure, ranging from functions which bear a strong similarity to those found in the mammalian visual cortex [8], [9], to pure sinusoids. The filters are based on solution of one of the eigenvalue problems associated with the uncertainty principle [10], [11]. Similar filters have found application in areas as diverse as spectrum estimation [12] and image feature extraction [13], [11]. They thus represent a variation of the approach to image processing first proposed by Granlund, based on operators which are local both spatially and in the frequency domain [14] and which form the basis of the GOP image processor [15], [16]. Furthermore, corresponding to the synthesis method is one of analysis, which consists of estimating the local energy for each filter in the set. The method is shown to be both general and effective. A description of the theoretical basis of the method is followed by a discussion of some preliminary experimental results and the relation of the methods employed to early visual processing.

2. A Model of Texture

The general model of texture proposed here is illustrated in Fig.1. A vector impulse process $s(x,y,u,v)$ is generated according to some specified probability law. For example, a spatially stationary process can be defined by

$$\begin{aligned} &\mathrm{Prob}\{s(x-x_1,y-y_1,u_1,v_1) = 1\} \\ &= \mathrm{Prob}\{s(x,y,u_1,v_1) = 1\} \\ &= 1 - \mathrm{Prob}\{s(x,y,u_1,v_1) = 0\} \end{aligned} \tag{1}$$

In other words the process $s(x,y,u,v)$ is defined as a binary process on a 4-dimensional discrete lattice, two of whose dimensions represent the spatial co-ordinates x and y. As another illustrative example, a 1-d spatially periodic process can be defined with period X by

$$\mathrm{Prob}\{s(x-kX,y,u_1,v_1) = s(x,y,u,v)\} = 1 \quad \forall k \tag{2}$$

Thus, while $s(x,y,u,v)$ will typically be a stationary process with independent increments, a large class of processes can be generated in this way.

The vector process $s(x,y,u,v)$ is spatially convolved with a 4-dimensional, shift-invariant operator $h(x,y,u,v)$, the outputs being summed to produce the texture field $f(x,y)$

$$f(x,y) = \sum_{u\varepsilon U}\sum_{v\varepsilon V}\sum_{x\varepsilon X}\sum_{y\varepsilon Y} h(x_1,y_1,u,v) . s(x-x_1,y-y_1,u,v) \tag{3}$$

where U,V,X,Y are the regions of support of the operator $h(x,y,u,v)$.

It follows immediately from the central limit theorem [17] that if $s(x,y,u,v)$ is a process which is independent in the four dimensions

$$\mathrm{Prob}\{s(x_1,y_1,u_1,v_1) \,|\, s(x\varepsilon X; x\neq x_1, y\varepsilon Y; y\neq y_1, u\varepsilon U; u\neq u_1, v\varepsilon V; v\neq v_1)\} = \mathrm{Pr}\{s(x_1,y_1,u_1,v_1)\} \tag{4}$$

then the probability density of $f(x,y)$ becomes Gaussian as U,V,X and Y tend to infinity. Thus one of the commonest texture processes, the Gaussian ([1]-[5]), may also be obtained as a limiting case of the above model.

Having defined a 4-dimensional driving process, it is now necessary to consider a suitable class of linear filters. While many choices are possible, one has special properties which commend its use. The filters are based on the solution of the eigenvalue problem (5)

$$T_s F^* T_f F \psi = \lambda \psi \tag{5}$$

where F is the N^2 dimensional 2-d discrete Fourier transform operator

SYNTHESE ET ANALYSE DE TEXTURE UTILISANT DES FILTRES A MINIMUM D'INCERTITUDE

TEXTURE SYNTHESIS AND ANALYSIS USING

MINIMUM UNCERTAINTY FILTERS

(Roland Wilson)

$$f_{klmn} = N^{-1}\exp\{-\frac{j2\pi(km+ln)}{N^2}\} \quad (6)$$

and T_s and T_f are diagonal truncation operators, defined by

$$t_{sklmn} = \delta_{km}\delta_{ln} \quad (k,l)\in\Omega_s \quad (7)$$
$$= 0 \quad (k,l)\in\Omega_s$$

and

$$t_{fklmn} = \delta_{km}\delta_{ln} \quad (k,l)\in\Omega_f \quad (8)$$
$$= 0 \quad (k,l)\in\Omega_f$$

where Ω_s and Ω_f define regions of arbitrary shape in the spatial and frequency domains respectively. Typically, either Ω_s or Ω_f or both are defined in cartesian separable terms [11]. In the case where both are cartesian separable, the eigenvalue problem (5) becomes separable and the eigenvectors are the Kronecker product of the corresponding 1-d eigenvectors. The solutions of (5) and their elegant symmetry properties have been discussed at length in [11], where a more complete list of references, including the original work of Slepian [10] can be found. For the present purposes, it suffices to note that the following are true:

2.1 The eigenvectors ψ will be real provided Ω_f is symmetrical.

2.2 The eigenvalues λ satisfy $0 \le \lambda < 1$ and represent the fraction of energy remaining after the spatially limited eigenvector ψ is bandlimited (by F^*T_fF) and truncated (by T_s). The eigenvector ψ^0 corresponding to the largest eigenvalue, λ_0, therefore has the largest simultaneous concentration of energy inside the two regions Ω_s, Ω_f of any function.

2.3 The related problem (9) below can be used to define pairs of bandpass filters which are in quadrature [11]:

$$FT_sF^*T_f\phi = \lambda\phi \quad (9)$$

where

$$\phi = F\psi \quad (10)$$

The two extreme cases of Eq. (5) are of some interest because they exhibit the flexibility of the method. Consider first the case of maximal concentration in the frequency domain $\Omega_f = (k_1, l_1)$ then with

$$T_f = \delta_{km}\delta_{ln}\delta_{kk_1}\delta_{ll_1}$$
$$T_s = I \quad (11)$$

it follows that

$$\psi^0_{mn} = N^{-1}\exp -\{\frac{j2\pi(k_1m+l_1n)}{N^2}\}, \lambda_0 = 1 \quad (12)$$

and $\lambda_k = 0$, $1 \le k < N^2$. Thus at one extreme the eigenvectors are pure complex exponentials. Similarly, if the roles of T_s and T_f are reversed

$$T_s = \delta_{km}\delta_{ln}\delta_{kk_1}\delta_{ll_1}$$
$$T_f = I \quad (13)$$

then

$$\psi^0_{kl} = \delta_{kk_1}\delta_{ll_1} \quad \lambda_0 = 1 \quad (14)$$

and $\lambda_0 = 0$, $1 \le k < N^2$, yielding an eigenvector of maximal spatial concentration. Obviously, making intermediate choices of operator implies an eigenvector of intermediate concentration in the two domains. Nonetheless, the eigenvector ψ^0 always represents the most effective trade-off between the two domains, a most desirable feature, both theoretically and in the light of recent physiological work (c.f. Section 4 below).

Having completed this summary of the properties of the functions ψ^0, it remains to incorporate them in the texture model. This is done by assigning to each lattice point (u,v) a distinct region of the frequency domain $\Omega_f(u,v)$ such that

$$\Omega_f(u,v) \cap \Omega_f(u^1,v^1) = \phi \quad u=u^1 \; v=v^1 \quad (15)$$

In this way a mapping is established between disjoint regions of the frequency domain and the filter functions which best represent them within a given spatial window Ω_s, which is common to all the functions in the set. Thus the filters h(x,y,u,v) are simply the 0th eigenvectors ψ^0 of the eigenvalue problem (5) defined by Ω_s and $\Omega_f(u,v)$

$$h(x,y,u,v) = \psi^0_{xy}(u,v) \quad (16)$$

where

$$T_sFT_f(u,v)F^*\psi^0(u,v) = \lambda_0(u,v)\psi^0(u,v) \quad (17)$$

Recalling the limiting cases of Eqs. (11) to (14), it can be seen that choosing $T_f(u,v)=I$ for some (u,v), a single impulse results which can reproduce any component of the original impulse process, while if

$$T_f(u,v) = \delta_{km}\delta_{ln}\delta_{ku}\delta_{lv} \quad 0\le u,\ v<N^2$$
$$T_s = I \quad (18)$$

then the filters become complex exponentials and the synthesis method becomes pure spectral synthesis [3].

Thus the texture model is both powerful and general, covering the full range of images normally classed as 'texture'. Furthermore, given a texture image to be described, simple filtering operations with the same set of filters yield estimates of the relative energy in each region of the frequency domain. Specifically, a complex filterwhose components are in quadrature can be defined by use of the Eq. (9)

$$FT_sF^*T_f(u,v)\phi^0(u,v) = \lambda_0(u,v)\phi^0(u,v) \quad (19)$$

where the truncation operator $T_f(u,v)$ contains only 'positive' frequencies, e.g.

$$T_f(u,v) = 0 \quad u<0 \quad (20)$$

The envelope of the spatial convolution of $\phi^0(u,v)$ with the image then provides an estimate of the instantaneous energy within the region $\Omega_f(u,v)$, $e(x,y,u,v)$

$$e(x,y,u,v) = |f(x,y)*\phi^0(x,y,u,v)|^2 \qquad (21)$$

The function $e(x,y,u,v)$ thus provides a measure of the relative distribution among the regions $\Omega_f(u,v)$ of the frequency domain and may thus be used as a texture feature vector [1]. Furthermore, the disjointness of the regions makes the estimates for different (u,v) practically uncorrelated, provided a suitable spatial window size is chosen. The method therefore provides an optimal compromise between descriptions based solely on localised features, like edges, and spectral or correlation methods [3].

3. Experiments

In order to test the ideas, a set of test images have been generated using a thresholded white Gaussian noise process to produce the impulse process which was input to the set of filters shown in Fig. 2. It is clear from these responses that polar separable forms of the truncation operators have been used, rather than a cartesian separable form. The reasons for this are related to known properties of mammalian visual cortical neurons (c.f. next section) and are also discussed in [16]. Figures 3 and 4 show a couple of textures generated with these methods, illustrating the range of texture types that can be produced. Figure 5 shows the effect of the analysis step on the texture of Fig. 4. The near orthogonality of the filters makes the distinction between the regions both easy to see and sufficient for classification.

4. Discussion: Relation to Recent Findings on the Visual Cortex

It has been known for many years that cells in the mammalian visual cortex are sensitive to features of different orientations. More recently, however, there had been a considerable controversy over the exact nature of these 'feature detectors' or even whether they are feature detectors at all. The debate centred around the bandwidth of the cells: a high bandwidth (> 1.5 octave) and spatial concentration being seen as evidence for feature detection, while a low bandwidth was used as support for a 'Fourier analysis' model of vision. Recent physiological work suggests that the truth lies somewhere in between, a finding which is supported by measured receptive field profiles which bear close resemblance to those shown in Fig. 4 [8], [9]. Moreover, the 4-dimensional arrangement of the filters used here exhibits some common features with the 'hypercolumn' arrangement found in the mammalian cortex [18]. While a close correspondence between the 'filters' used in a general texture synthesis/analysis scheme and those used by the visual system is not essential, it is nonetheless useful to employ a system which allows for that possibility. A fuller discussion of these ideas can be found in [18].

5. Conclusions

A general model of texture has been presented which gives a great degree of flexibility in the range of textures it can describe and offers the potential for employing methods which have at least a superficial resemblance to those employed by the visual system.

While the results obtained so far are of a preliminary nature, they have been sufficiently encouraging to warrant further work. In particular, the possibility of employing these methods in a hierarchical fashion appears to offer real conceptual and computational advantages.

Acknowledgment

The work reported here was supported by the U.K. S.E.R.C. and I.B.M. (U.K.) Ltd. My thanks are also due to Professor Gosta Granlund and Dr Hans Knutsson, whose help and encouragement have been invaluable.

References

1. Pratt, W K, Faugeras, O D, Gagalowicz, A
'Applications of Stochastic Texture Field Models to Image Processing'
Proc IEEE, 69, 5, pp 542-551, 1981

2. Faugeras, O D, Pratt, W K
'Decorrelation Methods of Texture Feature Extraction'
IEEE Trans on Pattern Anal and MI, 2, 4, pp 323-332, 1980

3. Haralick, R M
'Statistical and Structural Approaches to Texture'
Proc IEEE, 67, 5, pp 786-803, 1979

4. Jain, A K
'Advances in Mathematical Models for Image Processing'
Proc IEEE, 69, 5, pp 502-528, 1981

5. Chellappa, R, Kashyap, R L
'Synthetic Generation and Estimation in Random Field Models of Images'
Proc IEEE Conf on Pattern Rec and Image Proc, pp 577-583, Dallas, 1981

6. Garber, D D, Sawchuk, A A
'Texture Simulation using a Best-fit Model'
Proc IEEE Conf on Pattern Rec and Im Proc, pp 603-608, Dallas, 1981

7. Schmitt, F, Massaloux, D
'Texture Synthesis using a Bidimensional Markov Model'
Proc IEEE Conf on Pattern Rec and Im Proc, pp 593-596, Dallas, 1981

8. Watson, A B
'Summation of Grating Patches Indicates Many Types of Detector at one Retinal Location'
Vision Res, 22, pp 17-26, 1982

9. Movshon, J A, Thompson, I D, Tolhurst, D O
'Spatial Summation in the Receptive Fields of Simple Cells in the Cat's Striate Cortex'
J Physiol, 283, pp 53-77, 1978

10. Slepian, P
'Prolate Spheroidal Wave Functions, Fourier Analysis and Uncertainty, IV: Extensions to Many Dimensions'
BSTJ, 43, pp 3009-3057, 1964

11. Wilson, R
'Uncertainty, Eigenvalue Problems and Filter Design'
Linkoping University Int'l Tech Report
LiTH-ISY-I-0580, 1983

12. Thomson, D J
'Spectrum Estimation and Harmonic Analysis'
Proc IEEE, 70, pp 1055-1096, 1982

13. Shanmugam, K, Didley, F M, Green, J A
'An Optimal Frequency Domain Filter for Edge Detection in Digital Pictures'
IEEE Trans on Pattern Anal and MI, 1, pp 37-69, 1979

14. Granlund, G H
'In Search of a General Picture Processing Operator'
Comput Graphics and Image Proc, 8, pp 155-178, 1978

15. Lindgren, K, Antonsson, D, Granlund, G H
'GOP, a Different Architecture for Fast and Flexible Image Processing'
Proc Nat Comput Conf, Chicago, 1981

16. Knutsson, H
'Filtering and Reconstruction in Image Processing'
Linkoping University, Ph D thesis, 1982

17. Papoulis, A
'Probability, Random Variables and Stochastic Processes'
New York, McGraw Hill, 1965

18. Wilson, R
'The Uncertainty Principle in Vision'
Linkoping University Int'l Tech Report
LiTH-ISY-I-0581, 1983

SYNTHESE ET ANALYSE DE TEXTURE UTILISANT DES FILTRES A MINIMUM D'INCERTITUDE

TEXTURE SYNTHESIS AND ANALYSIS USING MINIMUM UNCERTAINTY FILTERS

(Roland Wilson)

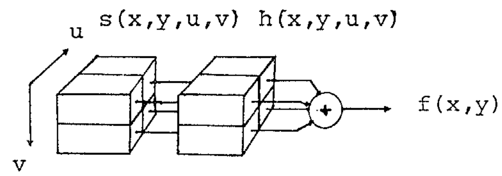
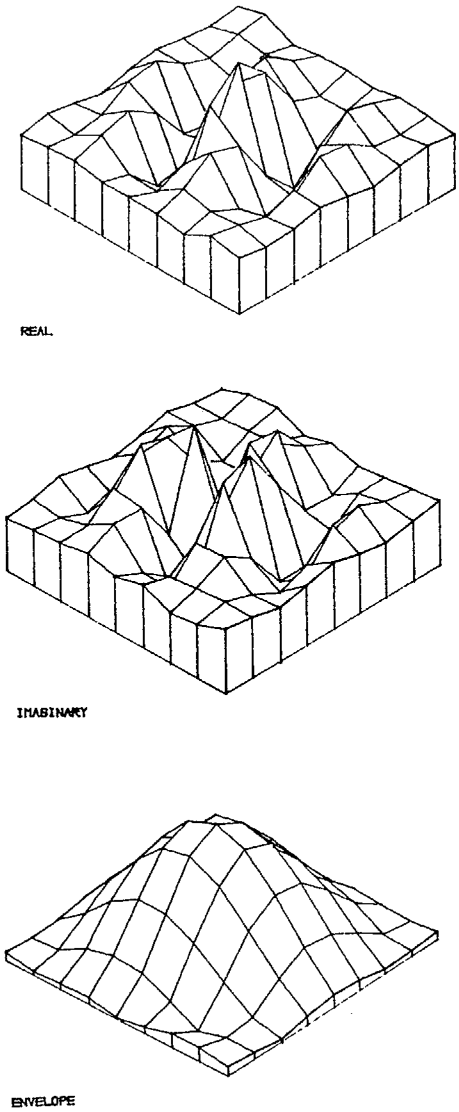


Fig. 1: General model of texture synthesis

Fig. 2: Spatial responses of one of directional filters used in producing Figs. 3-5



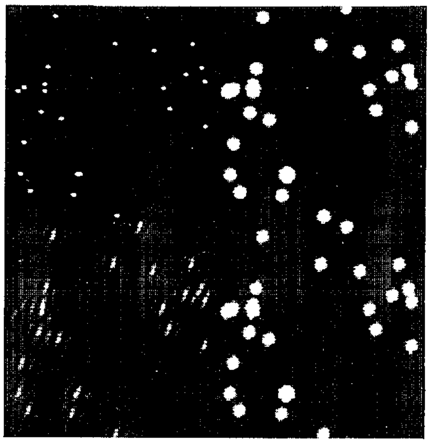


Fig. 3: Textures generated with low impulse frequency:
a) original impulse field b) circularly symmetric l.p. filter c) directional l.p. filter d) sum of b) and c)

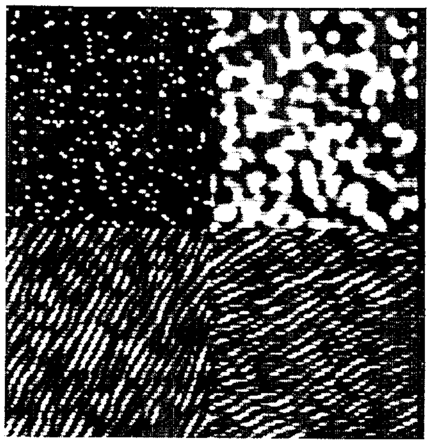


Fig. 4: Textures generated with high impulse frequency:
a) original impulse field b) circularly symmetric l.p. filter c) d) directional b.p. filters

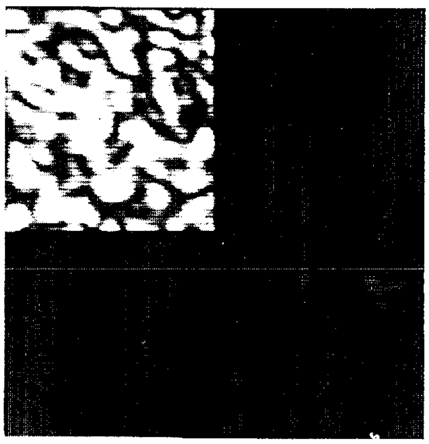


Fig. 5: Estimate of energy in one of 4 directions made on textures at a) 0^o b) 45^o c) 90^o d) 135^o to estimating filter direction