ELSEVIER

Signal Processing: *Image Communication* 8 (1996) 131–148

SIGNAL PROCESSING: IMAGE COMMUNICATION

The bi-orthogonal decomposition in image processing: Signal analysis and texture segmentation

J.A. Dente[a,*], R. Vilela Mendes[a], A. Lambert[b], R. Lima[b]

[a]*Laboratório de Mecatrónica, DEEC, Instituto Superior Técnico, Av. Rovisco Pais 1, 1096 Lisboa Codex, Portugal*
[b]*Centre de Physique Théorique, CNRS, Luminy Case 907, F13288 Marseille Cedex, France*

Received 25 February 1994

Abstract

Decomposition of multi-dimensional signals into orthogonal modes, generated by the data itself, has proven to be a powerful method for the identification of the structures that underlie complex dynamics. This technique is proposed as an image processing tool. The potential applications of the bi-orthogonal decomposition are data compression, noise-resistant communications and texture segmentation. The energy and information content of the bi-orthogonal modes as well as some of the proposed applications are illustrated by several examples.

1. Introduction

The bi-orthogonal decomposition [1], a technique specially adapted to the analysis of two-dimensional signals, is a natural candidate for image processing. After reviewing the theory of the bi-orthogonal decomposition (Section 2) we show how this technique characterizes the information content and the non-Gaussiuan nature of real world images (Section 3).

Contour information and small details are associated to very low energy modes. Redistributing the energy among all information carrying modes one obtains images with improved noise resistance (Section 4).

The entropy associated to the bi-orthogonal decomposition is an important quantity close to the actual information content of the image. Of particular interest is the local block entropy which characterizes local correlations and is the basis for a new method for texture segmentation (Section 5).

2. The bi-orthogonal decomposition

The decomposition into orthogonal modes is a well known procedure in signal analysis referred to as the *Karhunen–Loève decomposition* [6, 10] or *principal component analysis*. Given an N-component random vector $x_i(t)$ $\{i = 1, \ldots, N\}$, the covariance matrix $Q = E[xx^T]$ is diagonalized and the random vector $x(t)$ expressed as

$$x_i(t) = \sum_{k=1}^{N} c_k(t)\,\phi_i^{(k)}, \tag{2.1}$$

where each $\phi_i^{(k)}$ is an eigenvector of Q, that is a column of the matrix A that diagonalizes

*Corresponding author.

SSDI 0923-5965(95)00041-0

Q ($Q = A\lambda A^T$, λ diagonal). Using the N eigenvectors as a basis, the best (mean-square) P-component approximation to the signal x is obtained by keeping the c_k coefficients associated with the largest P eigenvalues. This property is useful for data compression applications. The Karhunen–Loève technique has been used for image processing [9, 14, 15]. The image is divided in small blocks, each block is taken as a sample of a statistical signal and the labeling of the blocks plays the role of a time variable. The expectation value of the covariance $E[xx^T]$ is then taken over these blocks.

There is, however, another way of looking at the Karhunen–Loève decomposition that may be useful for image processing. Let the eigenvalues of Q be α_k^2. Then, defining $c_k(t) \equiv \alpha_k \psi_k(t)$, Eq. (2.1) is rewritten as

$$x_i(t) = \sum_{k=1}^{N} \alpha_k \psi_k(t) \phi_i^{(k)}. \tag{2.2}$$

The signal appears now as a sum of products of two families of orthonormal functions of t and of the discrete variable i. This way of looking at the Karhunen–Loève decomposition is called *the bi-orthogonal decomposition* and is particularly useful for the analysis of data depending on two variables.

In an image, an important part of the information content is related to geometrical correlations. They concern the variation of the gray level along particular directions and contain the contour and shape information of the image. This suggests the use, in Eq. (2.2), of the space coordinates x and y as the two independent variables.

In all generality the bi-orthogonal decomposition analyses signals $u(x,y)$ that depend on variables defined in two distinct spaces ($x \in X$, $y \in Y$). We summarize the main results concerning this decomposition and refer to [1] for more details. Let the signal $u(x, y)$ be a measurable complex-valued function defined on $X \times Y$, where X and Y are either $\mathbb{R}^n$ or $\mathbb{Z}^n$ or subsets of one of these. The signal itself is used as a kernel to define a linear operator $U: L^2(Y) \to L^2(X)$ by

$$(U\psi)(x) = \int_Y u(x, y)\psi(y)\,dy \quad \forall \psi \in L^2(Y), \tag{2.3a}$$

with adjoint operator $U^\dagger: L^2(X) \to L^2(Y)$

$$(U^\dagger \phi)(y) = \int_X u^*(x, y)\phi(x)\,dx \quad \forall \phi \in L^2(X). \tag{2.3b}$$

With these definitions the analysis of the signal $u(x,y)$ is the spectral analysis of the operator U. In general the spectrum contains continuous and point spectral components. However, for simplicity, we will assume that $u \in L^2(X \times Y)$ or that X and Y are compact and u continuous. This implies that U is a compact operator and the spectrum consists of a countable set of isolated points. Then, there is a canonical decomposition of $u(x, y)$ such that

$$u(x, y) = \sum_{k=1}^{\infty} \alpha_k \phi_k(x) \psi_k^*(y) \tag{2.4}$$

is norm-convergent and

$$\alpha_1 \geqslant \alpha_2 \geqslant \cdots > 0, \quad (\phi_i, \phi_j) = (\psi_i, \psi_j) = \delta_{i,j}.$$

The functions $\Phi_k(x)$ are the eigenfunctions of the operator $L = UU^\dagger$, and the $\Psi_k(y)$ are the eigenfunctions of $R = U^\dagger U$. These functions are related by

$$\phi_k = \alpha_k^{-1} U \psi_k. \tag{2.5}$$

The operators L and R are non-negative operators with kernels $l(x_1, x_2)$ and $r(y_1, y_2)$ which are the X and Y-correlation functions of the signal

$$l(x_1, x_2) = \int_Y u(x_1, y)u^*(x_2, y)\,dy, \tag{2.6}$$

$$r(y_1, y_2) = \int_X u^*(x, y_1)u(x, y_2)\,dx. \tag{2.7}$$

In the decomposition, the eigenvectors Φ_k and Ψ_k of the L and R operators appear coupled to the same eigenvalue α_k^2. The products $\Phi_k \Psi_k$ are therefore the independent X,Y-structures that compose the signal. This decoupling of structures occurs because, as opposed to other methods of signal

analysis (Fourier, wavelets, etc.), the functional basis decomposing u is produced by u itself.

From the bi-orthogonal decomposition several global quantities may be constructed. The square of the norm of the signal in $L^2(X \times Y)$, which we call the *energy*, is the sum of the eigenvalues

$$E(u) = \int_{X \times Y} u(x, y)u^*(x, y)\mathrm{d}x\,\mathrm{d}y = \sum_k \alpha_k^2. \tag{2.8}$$

The *dimension* of a signal is defined to be the dimension of the range of U. For the compact case this is the number of non-zero eigenvalues α_k^2. The *ε-dimension* of the signal is the number of eigenvalues larger than ε. The size of the eigenvalues is a good characterization of the degree of approximation in the sense that, truncating the U operator to

$$U_P = \sum_{k=1}^{P} \alpha_k \phi_k \psi_k$$

the norm of the error $\| U - U_p \|$ is smaller than the first neglected eigenvalue.

Associated to the eigenvalue structure of the bi-orthogonal decomposition one also defines an *entropy* by

$$H_{\mathrm{BO}}(u) = -\sum_{k=1}^{N} p_k \log p_k, \tag{2.9}$$

with

$$p_k = \frac{\alpha_k^2}{\sum_k \alpha_k^2}. \tag{2.10}$$

For the continuum case there will be an infinite number of modes and the sum in (2.9) may be replaced by $-\lim_{N\to\infty} (1/\log N) \sum_{k=1}^{N} p_k \log p_k$.

Notions of entropy relate to the information content and are useful for image processing [5, 11]. The entropy measures the amount of disorder in a system and is sensitive to the spread of possible states which the system can take. For an image, the simplest idea would have been to make the states correspond to the possible values which individual pixels can take. Then, the entropy (associated to the gray level histogram) is given by

$$H_{\mathrm{GL}} = -\sum_{j=0}^{M-1} P(j) \log P(j), \tag{2.11}$$

where $P(j)$ is the probability of pixel value j and M is the number of different values which the pixels can take. However, Eq. (2.11) represents the actual information content of the image only if the pixels are all uncorrelated. This is not the case in real world images. Instead of the original image, consider, for example, an image formed by the differences of neighbouring pixels. The original image can be reconstructed from the difference image together with the value of the first pixel. Therefore, they contain the same information. However, one usually finds that the entropy of the gray level histogram of the *difference image* is smaller than the one for the original image. This occurs because the difference image extracts some of the space correlations existing in the image, hence its entropy is closer to the actual information content of the image.

The bi-orthogonal decomposition extracts the normal modes of the image taking into account correlations along two coordinate axis. Therefore we expect the entropy H_{BO} associated to the weights of the bi-orthogonal modes to be even closer to the actual information content of the image. To test this conjecture we have computed the gray level histogram entropy, the entropy of the difference images and the entropy of the bi-orthogonal decomposition for real world and synthetic images. In general, the entropy of the difference image is smaller than the gray level histogram entropy. In all cases tested, we have however found that the bi-orthogonal entropy H_{BO} is the smaller of them all. In Table 1 we list the computed values for a real world image ("bears", Fig. 13) and the synthetic image used for the textures test in Section 5 (Fig. 8).

Table 1
Entropy values for the *textures* and *bears* images

Image	Bi-orthogonal	Histogram	Difference
Textures	1.1134	5.1965	2.7945
Bears	0.4260	3.8528	3.4427

3. Signal analysis by the bi-orthogonal decomposition

An image is a continuous function $u(x, y)$ of two variables defined on a rectangular region of the plane. The value at the point (x, y) is called the *brightness* or *gray level*. Black and white images or the components of a colour image are therefore scalar functions represented by $[M \times N]$ pixel matrices U. To carry out the bi-orthogonal decomposition of U, the matrix $UU^{\dagger}$ is diagonalized (if $M < N$) and the eigenvectors $\Psi_k(y)$ of $U^{\dagger}U$ obtained from (2.5). The image becomes

$$u(x, y) = \sum_{k=1}^{M} \alpha_k \phi_k(x) \Psi_k(y). \tag{3.1}$$

We then compute the global energy $E(u)$ and the normalized contribution of each one of the structures to the energy (p_k) and the entropy ($p_k \log p_k$). In general, one finds that most of the information is contained in a number P of leading modes, P being much less than M. The remaining structures reconstruct a signal that is indistinguishable from noise. Notice, however, that the signal is highly non-Gaussian and important information is contained in the low energy tail of the retained P modes (see below).

Because only a fraction of the $\phi_k \Psi_k$ structures is sufficient to reconstruct the image, the bi-orthogonal decomposition has a potential as a data compression technique. For the compression to be effective the components of the first P eigenvectors that are kept, together with their eigenvalues, must have less bits than the whole image. If, in addition, the eigenvectors vary little from image to image, in a sequence of frames, it would be enough to transmit the new eigenvalues together with a special code indicating that the last transmitted set of eigenvectors is to be used.

To carry out the bi-orthogonal decomposition involves a larger amount of computation than an expansion in a fixed basis as, for example, the discrete cosine transform. Notice, however, that the expansion in a basis generated by the data itself is a shorter and more natural way to deal with the data, whereas a collection of Fourier coefficients is not particularly enlightning because in practice

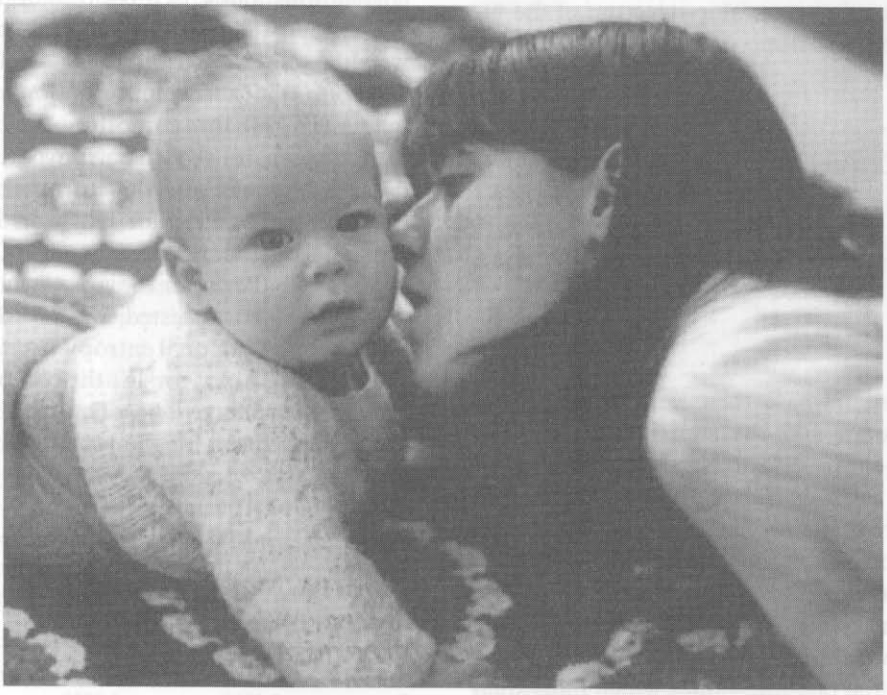

Fig. 1. A woman and a baby.

a sine wave does not look much like any image signal.

As an example consider a $[624 \times 800]$ pixel image U taking values in the range $[0, 255]$ (Fig. 1). To carry out the bi-orthogonal decomposition the 624×624 matrix $UU^{\dagger}$is diagonalized and one obtains an expansion as in Eq. (3.1) with $M = 624$ modes. The normalized contributions of the $\phi_k \Psi_k$ structures to the energy (p_k) and to the entropy ($p_k \log p_k$) are plotted in Fig. 2(a, b). The

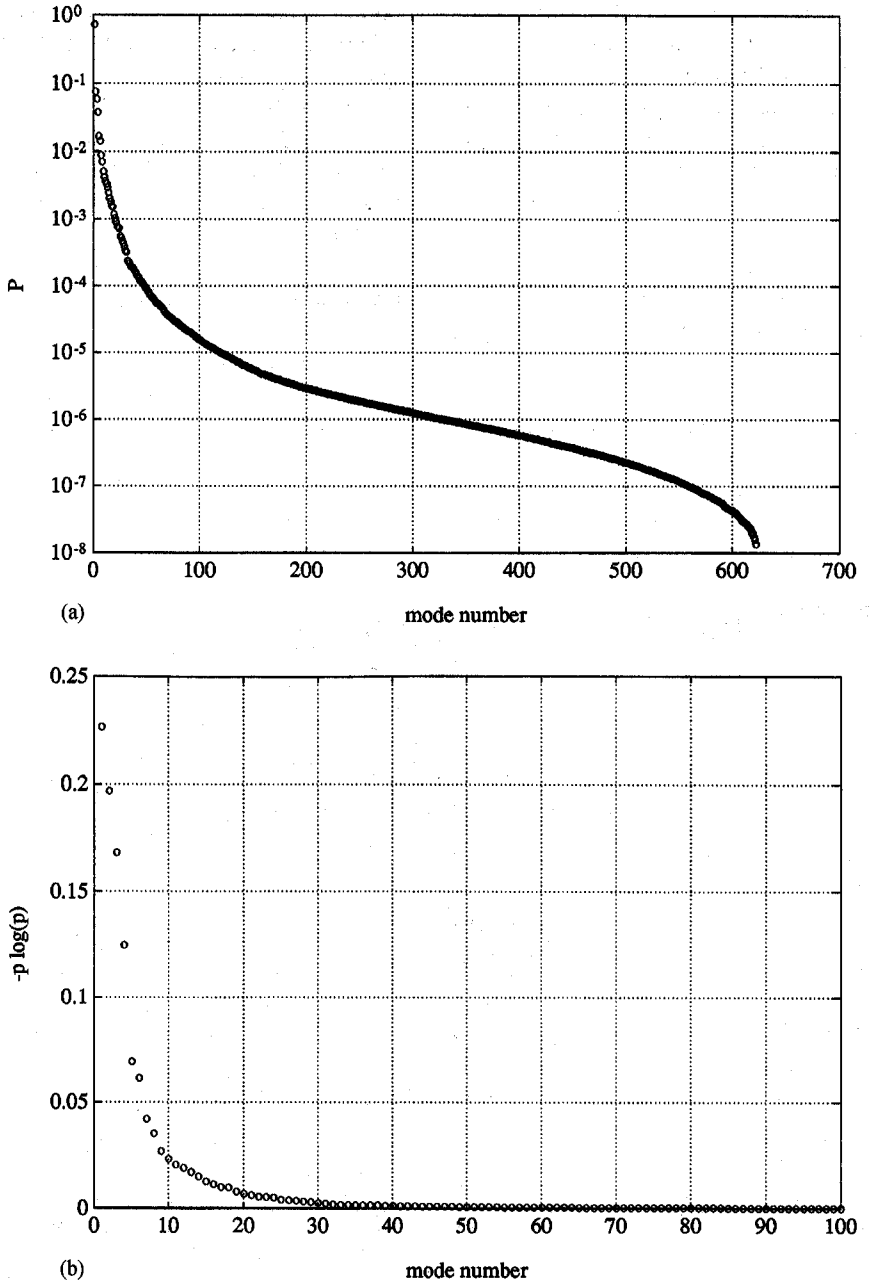

Fig. 2. (a) Normalized energy of the structures. (b) Normalized entropy of the structures.

non-Gaussian nature of the signal is well apparent from these plots. We find that one of the eigenvalues is much larger than the others. This means that a large amount of the global energy is associated with a small number of structures with the largest eigenvalues. In this example 80% of the global energy is associated with two structures. The energy concentration in these structures does not however relate to its information content. In Fig. 3 we show the image that corresponds to the two structures with the largest energy. One sees that, as far as the intelligibility of the picture is concerned, the information associated with these structures is very low, because it is not possible to recognize anything important. These structures look like the output of a low-pass filter. They characterize the local average gray level throughout the picture, but all information about details is elsewhere. This becomes clear in Fig. 4, where we show the image obtained without the two largest eigenvalues, i.e. with only 20% of the energy. Notice that the original picture has gray level values between 0 and 255. However, after the diagonalisation, the components of the structures $\Phi_k \Psi_k$ take positive and negative real values. For the graphical representation of the truncated images we have always shifted and rescaled the gray level to make it lie in the range [0, 255].

The *effective dimension* of the signal is the number of structures needed to reconstruct the information contained in the image. In Fig. 5(a, b) we show the images formed with the 20 and 120 largest eigenvalues. Important contour information is still contained in the (weak) eigenvectors between the 21st and the 120th. However, the image formed by the first 120 eigenvectors is virtually indistinguishable from the original image. Even for this image, which contains many small details, one fourth of the structures is sufficient to reconstruct the image in an accurate fashion. The normalized square error

$$\mathrm{NSE} = \frac{\sum (u(x, y) - \hat{u}(x, y))^2}{\sum u^2(x, y)}$$

between the actual image and the sum of the first 120 modes is 0.427% and its energy is 99.9% of the total energy. For the 20 modes picture of Fig. 5(a) the energy is 98.8% and the NSE is 5.5%.

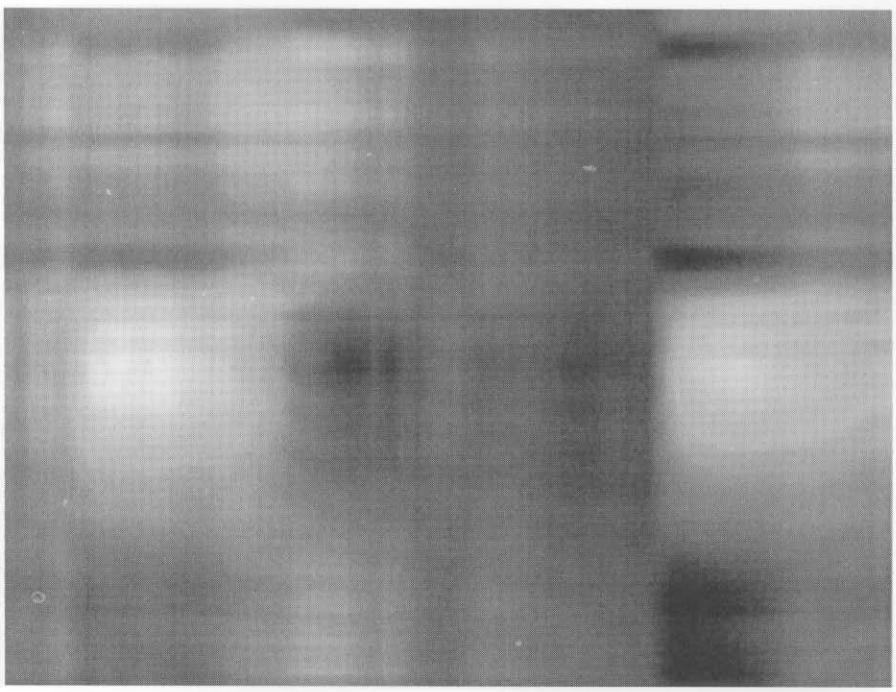

Fig. 3. Image of the two largest energy modes.

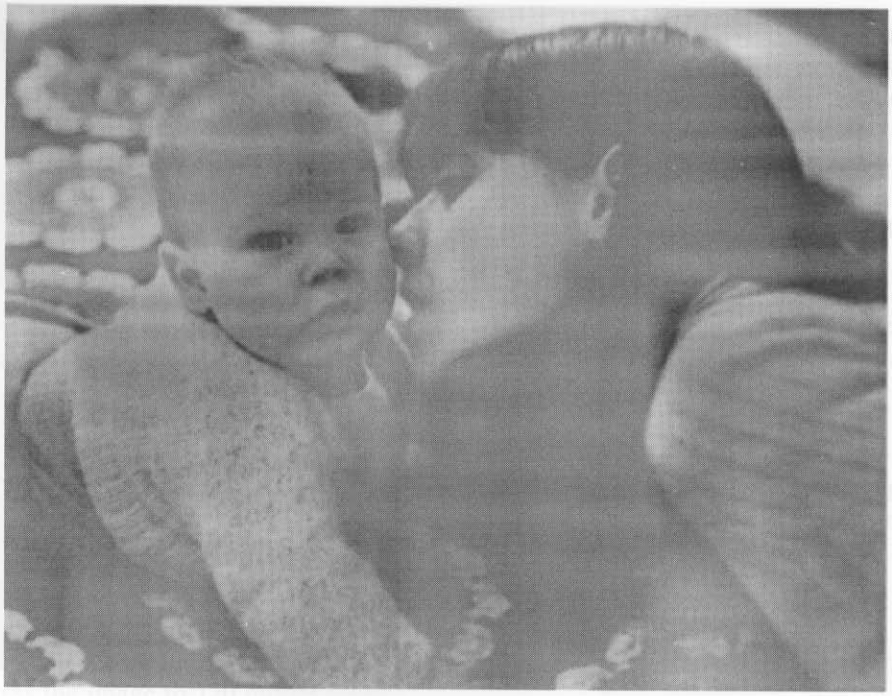

Fig. 4. Image excluding the two largest energy modes.

Efficient signal processing should take into account the statistical structure of the signal. An important point that comes up from the analysis of this example, and is true in general for real world images, is the non-Gaussian nature of the signal. It is well known that gray level intensity histograms of real world images are also non-Gaussian. The bi-orthogonal decomposition is however more specific in pinning down the meaning of the several parts of the spectrum, in particular the important contour information contained in the small energy components. This implies in particular that image signals are very sensitive to noise, a subject we discuss in more detail in the next section.

4. Noise and the bi-orthogonal decomposition

Using the bi-orthogonal decomposition the effect of noise in an image becomes a problem of perturbation of linear operators. If noise is added to an image $u(x, y)$, the resulting signal

$$v(x, y) = u(x, y) + e(x, y) \tag{4.1}$$

is associated, by (2.3a), to an operator

$$V = U + \mathscr{E}. \tag{4.2}$$

If $e(x, y)$ is small compared with $u(x, y)$, the operator $\mathscr{E}$ is a small perturbation added to U. Therefore the bi-orthogonal decomposition of $v(x, y)$ is related to the one of $u(x, y)$ by the usual techniques for perturbations of linear operators [7]. According to a well known criterion, the eigenvalues and eigenvectors of the perturbed operators $VV^\dagger$ and $V^\dagger V$ are regular functions of the parameter ε that defines the size of the perturbation, if the following condition holds

$$\Delta = \inf_k |\alpha_k^2 - \alpha_{k-1}^2| > |\varepsilon|, \tag{4.3}$$

where $|\varepsilon| = C \sup_{x,y} |e(x, y)|$ for some positive constant C. As long as (4.3) is fulfilled, no crossing of eigenvalues occurs and the only modifications are small deformations of the original eigenvectors and eigenvalues of $UU^\dagger$.

However, in a real image, $|\alpha_k^2 - \alpha_{k-1}^2|$ is very small for large k. Therefore any small amount of noise will modify the structures associated to the

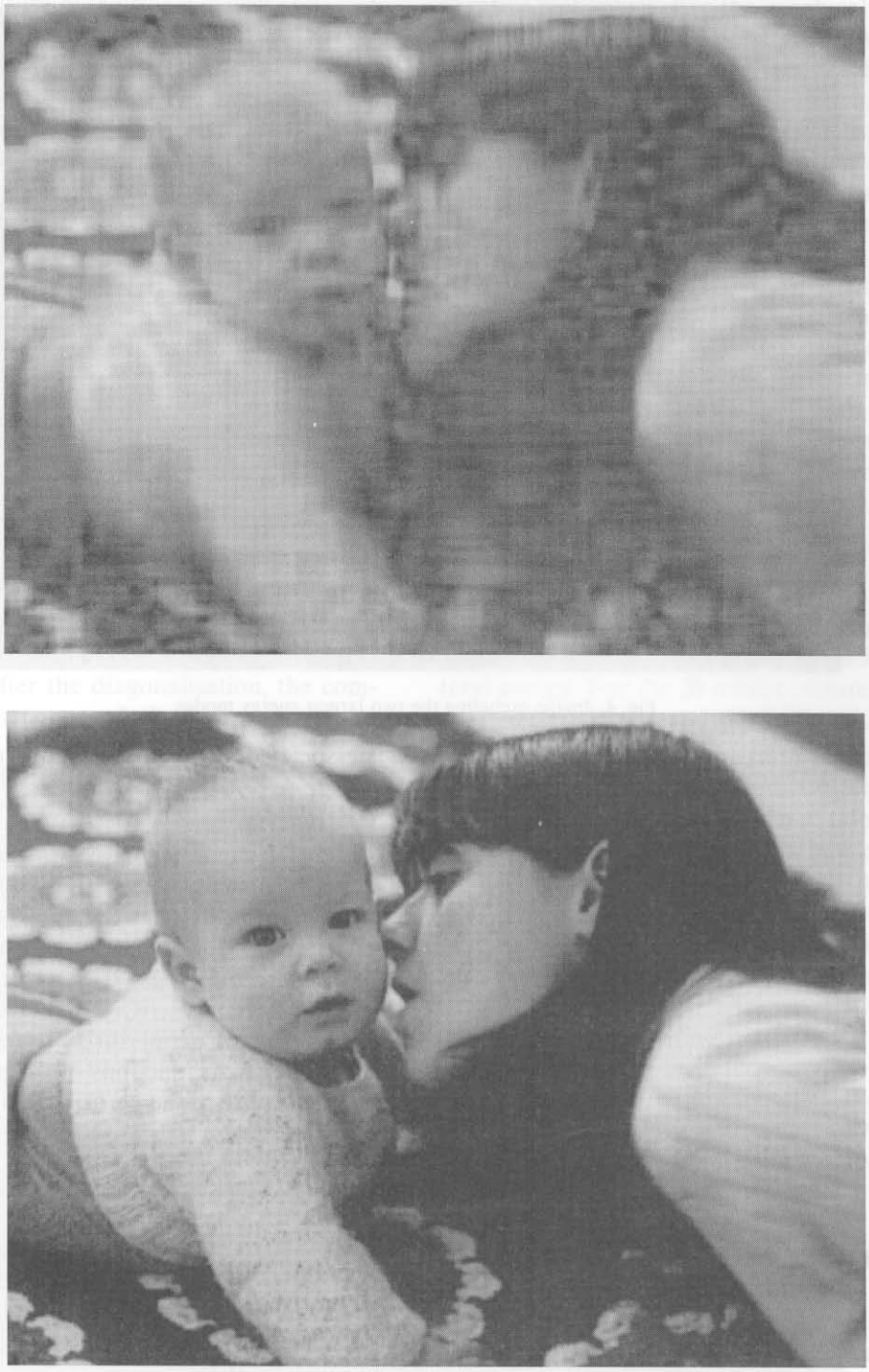

Fig. 5. (a) Image with the 20 largest eigenvectors. (b) Image with the 120 largest eigenvectors.

small eigenvalues. As we have seen in Section 2 these are the structures that contain information about contours and small details which, as a consequence, are very sensitive to noise. Suppose, however, that to the image $u(x, y)$ we associate a *spectral harmonic* counterpart, defined by

$$\tilde{u}(x, y) = \sum_{k=1}^{P} \theta_k \phi_k(x) \Psi_k(y), \tag{4.4}$$

where now the parameter θ_k^2 is a linear function of k

$$\theta_k^2 = A - \lambda k. \tag{4.5}$$

With $A = \lambda(P + 1)$ (P being the number of modes kept in the compressed image) one insures that $\tilde{\Delta} = |\theta_k^2 - \theta_{k-1}^2| = \lambda > 0$, the eigenvalues are equally spaced and, in the sense of perturbations of linear operators, $\tilde{u}$ is maximally robust for small noise perturbations.

We have tested this effect in the image studied in Section 3 (Fig. 1), namely:

(1) We perform the bi-orthogonal decomposition and compress the image to 120 modes

$$u_P(x, y) = \sum_{k=1}^{120} \alpha_k \phi_k(x) \psi_k(y).$$

(2) Then we form the image $\tilde{u}$, as in Eq. (4.4), θ_k^2 being the linear function (4.5) with coefficients chosen to preserve the same total energy, that is,

$$\sum_{k=1}^{120} \alpha_k^2 = \sum_{k=1}^{120} \theta_k^2. \tag{4.6}$$

(3) The image $\tilde{u}$ is corrupted by additive noise

$$v = \tilde{u} + e.$$

(4) The bi-orthogonal decomposition of the corrupted image $v(x, y)$ is obtained:

$$v(x, y) = \sum_{k=1}^{P} \theta'_k \phi'_k(x) \Psi'_k(y).$$

(5) An image u' is reconstructed using the bi-orthogonal modes of the corrupted image together with the original eigenvalues

$$u'(x, y) = \sum_{k=1}^{P} \alpha_k \phi'_k(x) \psi'_k(y).$$

For any given amount of channel noise compatible with the perturbation argument sketched above, the difference between the $\theta'_k \psi'_k$structures and the original ones $\phi_k \psi_k$ is expected to be smaller

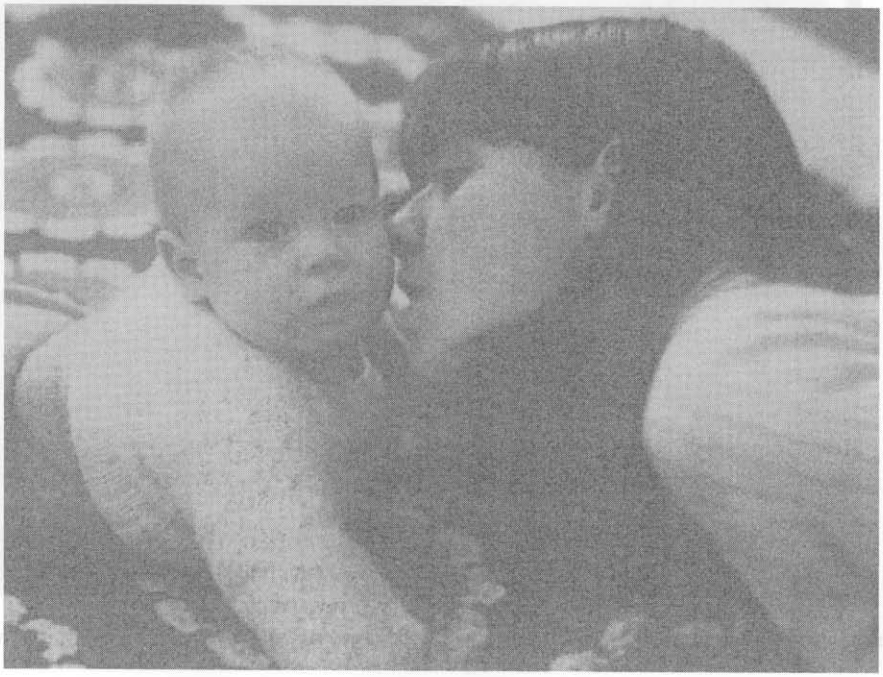

Fig. 6. Original image corrupted by Gaussian noise.

(a)

(b)

Fig. 7. (a) Image reconstructed from the harmonic picture exposed to the same noise intensity. (b) Reconstructed image without the first mode.

when $\tilde{u}$ is used than when the original picture is corrupted directly by the same amount of noise. Likewise the order of the eigenvalues will, with higher probability, be preserved. The results that are obtained have some interesting features. Whereas in the original image the effect of the noise is an overall blurring of contours and details, in the modified $\tilde{u}$ image one seems to have mostly a shift in the modes leading to the appearance of faint bands directed along the coordinate axis. By

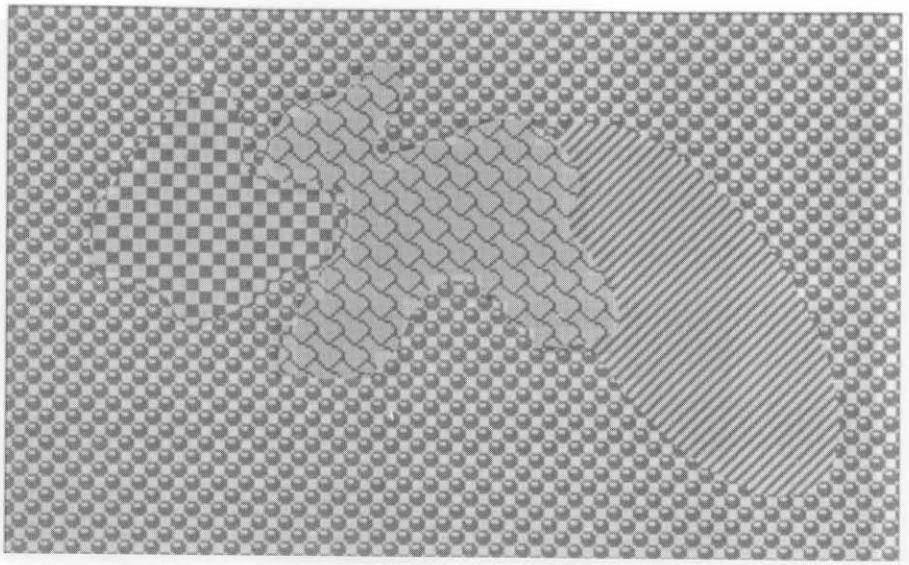

Fig. 8. Textures image.

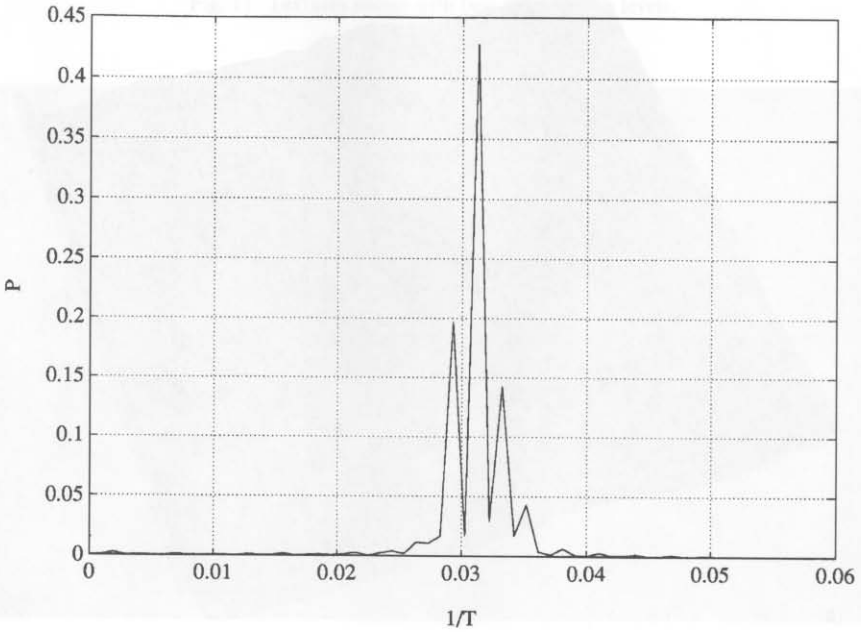

Fig. 9. Fourier spectrum of the ϕ_2 bi-orthogonal mode of *textures*.

contrast the contours and intelligibility of small details is not much affected.

According to Eq. (4.3), when the function α_k^2 is replaced by θ_k^2, preserving the same total energy, nose-resistance will increase for the small modes and will decrease for the large energy modes. Hence, eliminating some of the large energy modes the spurious bands should become weaker without affecting the contours. These effects are illustrated in Figs. 6 and 7(a, b). Fig. 6 shows the original image corrupted by Gaussian additive noise with energy equal to 10% of the image

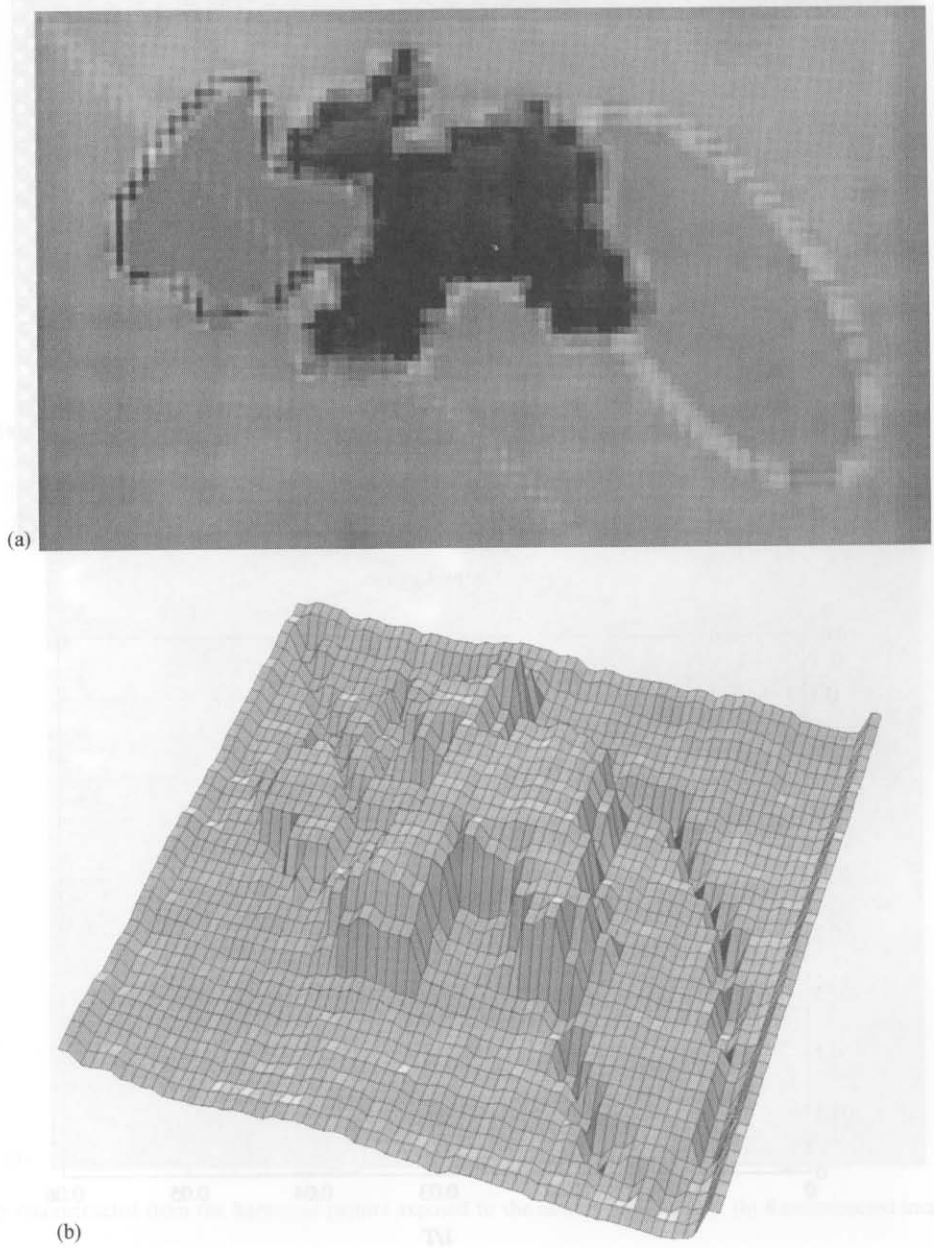

Fig. 10. (a, b) Block entropy image for *textures*.

energy. Fig. 7(a) is the reconstructed image u' for the same amount of noise and Fig. 7(b) is u' without the first mode.

The control of noise effects, by the transformations discussed above, might be used for image communications. The amount of computation needed in the processing of an image using this method, makes it unpractical for high rate transmission, at the present time at least. Nevertheless for remote transmission and when speed is not the main issue, the proposed correction of the non-Gaussian nature of the signal might still be useful.

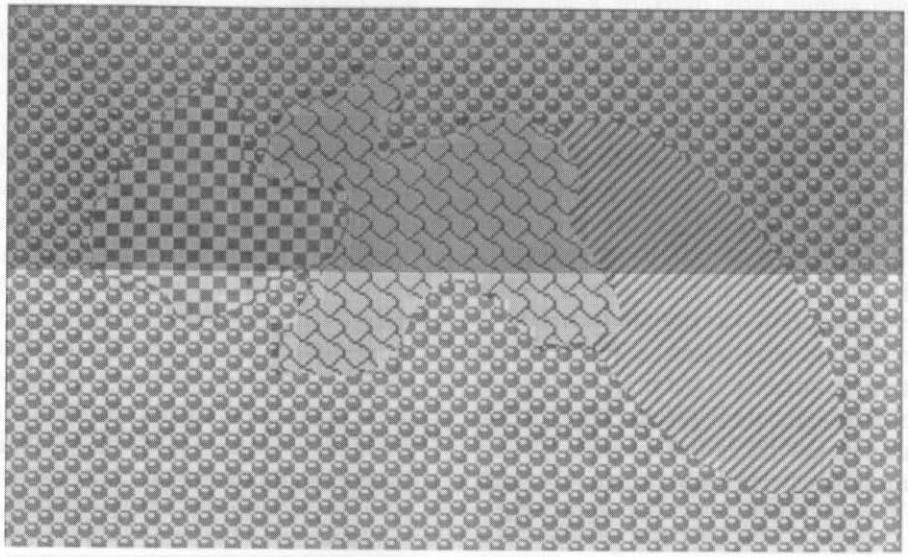

Fig. 11. Textures image with two illumination levels.

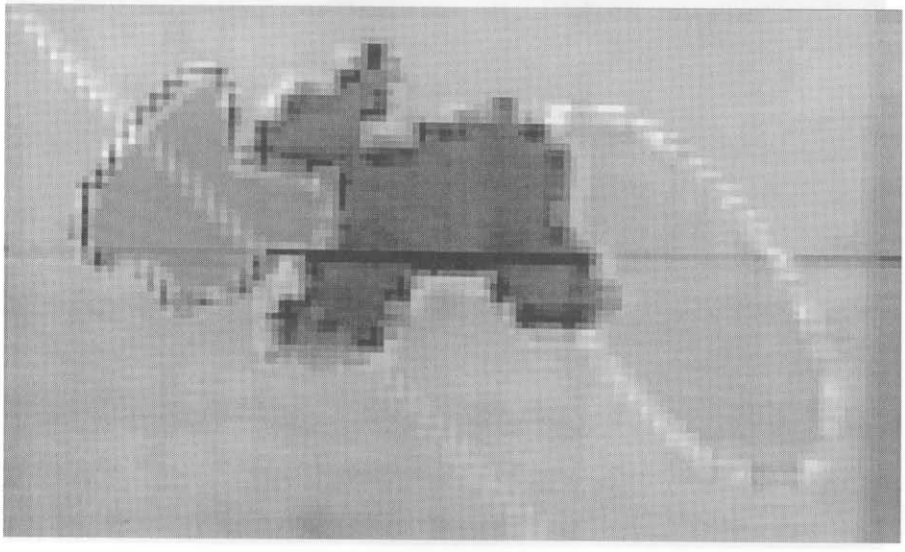

Fig. 12. Block entropy image with two illumination levels.

5. Texture segmentation by local bi-orthogonal decomposition

Image segmentation is the partition of a plane image into exclusive regions which, in some sense, are homogeneous. When the purpose of segmentation is to distinguish some object from a background and the brightness of the object and the background are significantly different, segmentation by *gray level thresholding* is a possibility. This works well for example in automated manufacturing processes if the assembly parts are kept dark against a bright background but, in most images of three-dimensional objects, different illumination levels in different parts of the same object make gray level thresholding a poor discriminating technique.

A more frequently used technique is *edge detection* by convolution of the image with a discrete difference operator, followed by a contour filling algorithm to decide which pixels belong to each one of the segmented regions. Edge detection also faces serious difficulties because factors such as illumination may either hide physical boundaries or, through shadows, cause brightness discontinuities which are not related to any real boundaries.

The reason why segmentation in computer vision is such a difficult problem, as compared with the ease with which the 'eye plus brain' system performs this task, is because in the brain a huge amount of information is stored concerning the way the real world looks like. Then, based on a few external stimuli, like brightness levels and a few contours, the bulk of the segmentation process in the brain is likely to be mostly an exercise in pattern matching of the external stimuli with our 'image of the world' data basis. While our computers are not equipped with a data basis of comparable size and complexity as the brain, computers must rely on a refinement in the analysis of the external stimuli part of the process. This means that quantitative characterizations of global and local properties of the image must be developed, which might even have to be finer and more accurate than those performed by the human eye. Only then, might we compensate for the weakness of the data basis in computer vision.

Fig. 13. Bears.

The most difficult of all segmentation problems occurs when different regions of the image cannot be distinguished by gray level nor by sharp boundaries, but only by a difference in texture. Texture refers to the local characteristics of the image. A local gray level histogram is a local statistical parameter. However, for the texture, what matters most are the local spatial correlations between pixel intensities. The brain will probably perform the texture segmentation task by pattern matching with its data basis of textures but, in the computer, the only alternative is to attempt an objective mathematical characterization of what texture means. Several quantities have been proposed as a measure of texture. For example the gray level co-occurrence matrix [3, 4], the local autocorrelation function [13], the autocorrelation spread measures [2], the number of edges in a neighborhood [13], the local Fourier spectrum [8, 12], the moments of the gray-level histograms of small windows and texture primitives together with grammar rules to generate the patterns.

In this section we propose the idea that the entropy associated to the bi-orthogonal decomposition of local blocks is an appropriate tool to characterize textures in an image. The method for texture segmentation by local bi-orthogonal decomposition (LBO) contains three steps:

(i) *Identify the texture average scale.* Compute the Fourier transform of a few randomly chosen lines and columns of the image. Then identify the first peak in the spectrum after the peak around zero (which corresponds to the average pixel intensity and long-range slow variations). In typical images the first large peak away from zero is the lowest texture frequency ω_T. A block size $M \times N$ is then chosen where M and N correspond to the average $1/\omega_T$ along the lines and the columns. Instead of using the Fourier transform of a set of lines and columns, we may use the Fourier transform of one of the modes in a global bi-orthogonal decomposition of the image.

(ii) *Construct the entropy image.* The image is now divided into blocks of size $M \times N$ and the bi-orthogonal entropy (Eq. (2.9)) of each block is computed. Assigning to each block its entropy value one obtains a *block entropy image.*

(iii) *Segmentation from the entropy image.* The entropy image is smoothed by some standard algorithm and contour tracing from the smoothed entropy image completes the process of texture segmentation by LBO.

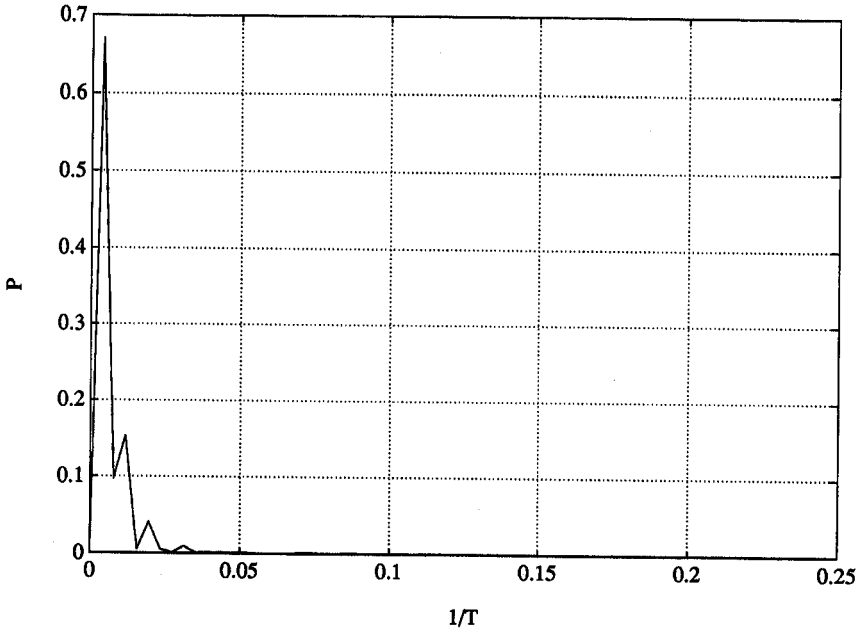

Fig. 14. Fourier spectrum of the ϕ_2 bi-orthogonal mode of *bears.*

The algorithm was tested on the synthetic image shown in Fig. 8. This image has several textures which were constructed in such a way that the local (in 32×32 blocks) gray level average is everywhere the same and no boundary lines exist separating the different textures. In this sense this example presents a pure case of segmentation by textures.

The average texture scale was found by computing the Fourier transform of the eigenfunctions ϕ_k of the global bi-orthogonal decomposition. Fig. 9 shows the spectrum of ϕ_2. A large peak may be seen, that corresponds to a block of dimension 32×32 pixels. The image is then divided into blocks of size 32×32 and the bi-orthogonal entropy of each block is computed to obtain the block entropy image. To obtain an entropy image with better resolution we might compute the entropy in the neighbourhood of every pixel in the original image. However, this procedure is time consuming. It suffices to generate an entropy image using the block entropy for only a smaller number of pixels in the original image. Fig. 10(a, b) shows the entropy image evaluated using 32×32 blocks separated by 8 pixels. In the entropy image, zones with different textures are well separated by the entropy values. This enables us to use a simple gradient algorithm to find the contours, thus performing the texture segmentation.

If the local entropy actually characterises the local texture, it should not be too sensitive to illumination levels in the image. We have tested this feature by changing the intensity in one half of our test image (Fig. 11). The relative insensitivity of the block entropies to illumination levels is apparent from the entropy image shown in Fig. 12.

In the 'pure textures' example described above, texture segmentation by local bi-orthogonal decomposition seems to work efficiently. In real world images, however, we see some difficulties and limitations of the method in particular when different materials have the same local correlations (same texture) or when there are many different texture frequencies ω_T. We show the results for a difficult case in Figs. 13–16. The first difficulty occurs in the choice of the block size. Different texture scales do occur, hence there is no unique block size appropriate for all texture features. Fig. 14 shows the spectrum of the global ϕ_2 eigenfunction of the 'bears' image. A large dimension is suggested for the dominant block size. This corresponds to large areas of

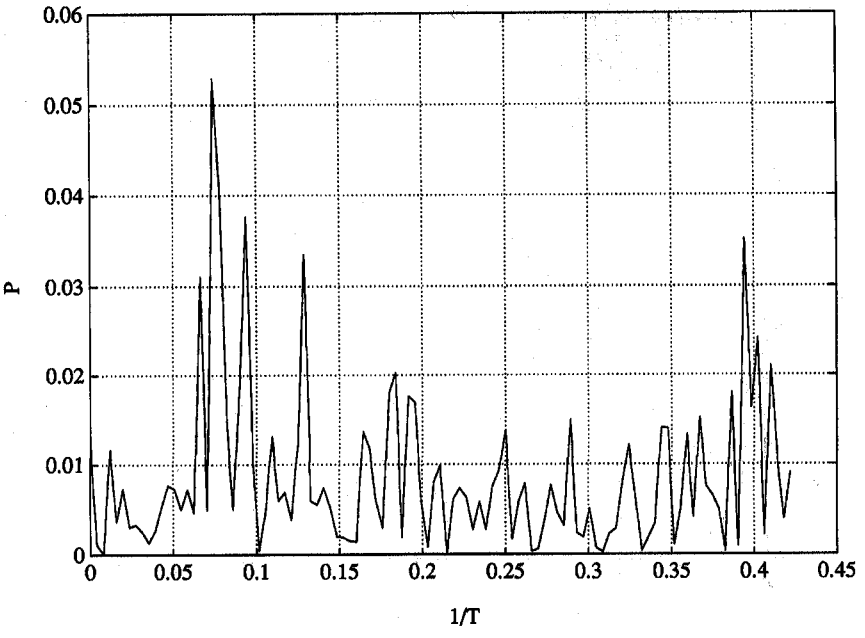

Fig. 15. Fourier spectrum of the ϕ_{50} bi-orthogonal mode of *bears*.

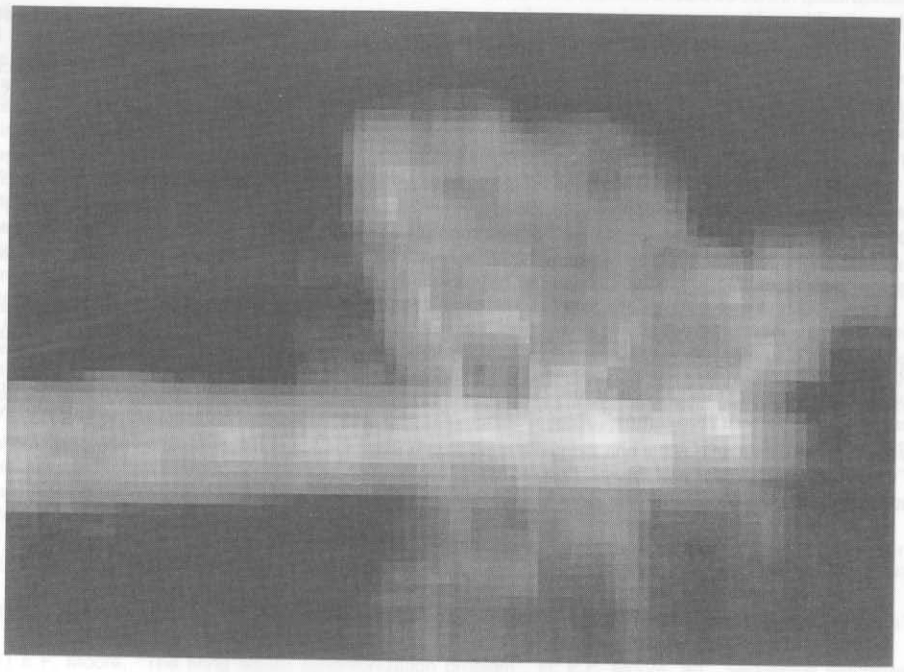

Fig. 16. Block entropy image for *bears*.

the image (the water and the mountain) which are not the main objects in the image. If other eigenfunctions are used, other characteristic block sizes are found, as shown in Fig. 15. The low block sizes are associated to low energy levels but, in spite of their small intensities, they are important to define the details of the image, that is, the microstructures that define the texture. An entropy image corresponding to a block size 16×16 is shown in Fig. 16.

6. Conclusions

As shown in the preceding sections, the bi-orthogonal decomposition, a technique specially adapted to handle functions depending on two variables, is an interesting tool for image processing. The main weak point, standing on the way to practical applications, is the amount of computation needed to process an image using this technique. The computation requirements involve mostly the diagonalization of large matrices or the diagonalization of many small ones. At least with the computing power available at the present time, the technique is probably not appropriate for high rate transmissions, although it might be useful for off-line processing or for low rate transmission under noisy conditions.

After many years of exploration of clever techniques to process images without much prior information, the image processing field is now moving in the direction of creation and use of a large data basis of patterns, shapes and textures. Pattern matching with such a large data basis will probably be the coming future in image segmentation. Here also the bi-orthogonal decomposition, with its intrinsic capabilities for the extraction of the typical space correlations, might be useful in the definition of a library of patterns for real world images.

References

[1] N. Aubry, R. Guyonnet and R. Lima, 'Spatio-temporal analysis of complex signals: Theory and applications', *J. Stat. Phys.*, Vol. 64, 1991, p. 683.

[2] O.D. Fangeras and W.K. Pratt, "Decorrelation methods of texture feature extraction", *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 2, 1980, p. 323.

[3] R.M. Haralick, K. Shanmugan and I. Dinstein, "Texture features for image classification", *IEEE Trans. Systems Man Cybernet.*, Vol. 3, 1973, p. 610.

[4] R.M. Haralick and L.G. Shapiro, "Glossary of computer vision terms", *Pattern Recognition*, Vol. 24, 1991, p. 69.

[5] J.N. Kapur, P.K. Sahoo and A.K.C. Wong, "A new method for gray level picture thresholding using the entropy of the histogram", *Comput. Graph. Vision Image Process.*, Vol. 29, 1985, p. 273.

[6] K. Karhunen, "Zur Spektral Theorie stochastischer Prozesse", *Ann. Acad. Sci. Fennicae, Ser. A*, Vol. 1, 37, 1944, p. 34.

[7] T. Kato, *Perturbation Theory of Linear Operators* (Springer, Berlin 1966).

[8] G.G. Lendaris and G.L. Stanley, "Diffraction pattern sampling for automatic pattern recognition", *Proc. IEEE*, Vol. 58, 1970, p. 198.

[9] J.S. Lim, *Two-dimensional Signal and Image Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1988.

[10] M. Loève, *Probability Theory*, Van Nostrand, Princeton, NJ, 1955.

[11] T. Pun, "A new method for grey-level picture thresholding using the entropy of the histogram", *Signal Processing*, Vol. 2, No. 3, July 1980, pp. 223–237.

[12] A. Rosenfeld, "Automatic recognition of basic terrain types from aerial photographs", *Photogrammic Engrg.*, Vol. 28, 1962, p. 115.

[13] A. Rosenfeld and E.B. Troy, Visual texture analysis, Proc. UMR-Mervin J. Kelly Communications Conference Section 10-1, 1970.

[14] T.D. Sanger, Optimal unsupervised learning in feedforward neural networks, MIT Artificial Intelligence Laboratory Tech. Report 1086, 1989.

[15] T.D. Sanger, "Optimal unsupervised learning in a single-layer linear feedforward network", *Neural Networks*, Vol. 2, 1989, p. 459.