

# Recent and prospective decorrelation techniques for image processing

Serge COMES \*  
Stéphane MAES \*  
Marc VAN DROOGENBROECK \*\*

## Abstract

*This paper is devoted to the description of recent and prospective decorrelation techniques used in image processing. Two classes are taken into consideration : the universal or generic methods and the adaptive methods. The first class performs signal processing independently of its specificity. These methods only rely on statistical properties of the signal. They are principally based on orthogonal transforms and subband coding. In opposition, the adaptive techniques take the a posteriori contents of the images into account. Some of them adapt universal techniques to the signal content (i.e. adaptive subband coding or hybrid methods). Others are based on feature extraction processes which include more physical entities like edges and regions. For example, object approaches attempt to describe images as collections of regions characterized by their shape and their aspect.*

**Key words :** Image processing, Decorrelation, Statistical method, Adaptive method, Digital image, Image coding, Orthogonal transformation, Subband decomposition, Mixed method, Non uniform sampling, Pattern extraction, Object oriented method.

---

## TECHNIQUES DE DÉCORRÉLATION RÉCENTES ET FUTURES POUR TRAITEMENT D'IMAGE

---

## Résumé

*L'article présente les techniques de décorrélation actuellement utilisées ou développées en traitement d'images. Elles sont regroupées en deux classes : les méthodes universelles, encore appelées génériques et les méthodes adaptatives. Dans le premier cas, le signal est traité indépendamment de ses propres caractéris-*

*tiques. Ces méthodes sont essentiellement constituées de transformations orthogonales et de décompositions en sous-bandes. Par contre, les techniques adaptatives considèrent davantage le contenu a posteriori de l'image traitée. Certaines résultent d'une simple adaptation de techniques universelles au contenu du signal (i.e. le codage en sous-bandes auto-adaptatif ou les méthodes hybrides). D'autres reposent davantage sur l'extraction de caractéristiques propres, ou d'entités physiques comme des contours ou des régions. Par exemple, l'approche orientée objet essaie de décrire une image sous la forme d'une collection de régions définies par leur forme et leur contenu.*

**Mots clés :** Traitement image, Décorrélation, Méthode statistique, Méthode adaptative, Image numérique, Codage image, Transformation orthogonale, Décomposition sous bande, Méthode mixte, Échantillonnage non uniforme, Extraction forme, Méthode orientée objet.

## Contents

- I. Introduction.
- II. Universal techniques.
- III. Adaptive techniques.
- IV. Conclusion.

References (86 ref.).

## I. INTRODUCTION

In this paper, we are concerned with digital images. They are signals composed of a large amount of samples, usually called pixels or pels, which are coded with

---

\* S. Comes and S. Maes are Research Assistants of the National Fund for Scientific Research, Belgium (FNRS), S. Maes is also member of FYMA (Unité de Physique Théorique et Mathématique, UCL) and of the Math dpt. and CAIP (Center for Computer Aids for Industrial Productivity), at Rutgers University and M. Van Droogenbroeck is under an IRSIA (Institut pour l'Encouragement de la Recherche Scientifique dans l'Industrie et l'Agriculture, Belgium) Grant.

\*\* Laboratoire de télécommunications et télédétection, Université catholique de Louvain, 3, place du Levant, B-1348 Louvain-la-Neuve, Belgium.

a certain accuracy (i.e. on a given amount of bits). For images, like those met in television, medical and satellite applications, the associated binary flows, which would be required for perfect transmission or lossless storage, are usually prohibitive. In the sense of the information theory, different lossy coding techniques (i.e. compression techniques which introduce information losses) enable to reduce the amount of bits which have to be handled, while keeping a still acceptable level of degradation.

A typical source coding scheme consists of the following three steps as described in Figure 1, although in specific applications, some of them might be missing :

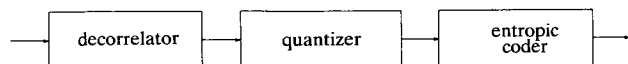


FIG. 1. — Typical source coding scheme.  
It consists of a decorrelator, a quantizer and an entropy coder.

*Schéma typique de codeur de source.  
Il est généralement constitué d'un décorrélateur,  
d'un quantificateur et d'un codeur entropique.*

1) A *decorrelative coder* which transforms the pixels into slightly uncorrelated coefficients [63]. Basically, an important redundancy should exist among the pixels of any usual image. It means that the statistical correlation  $\Gamma_{\text{orig}}(\tau)$  between the pixels in the spatial domain is a rather broadly peaked function of the distance between them  $\tau$ . It is illustrated by the left-hand side plot in Figure 2, in the case of a typical gray scale still image.

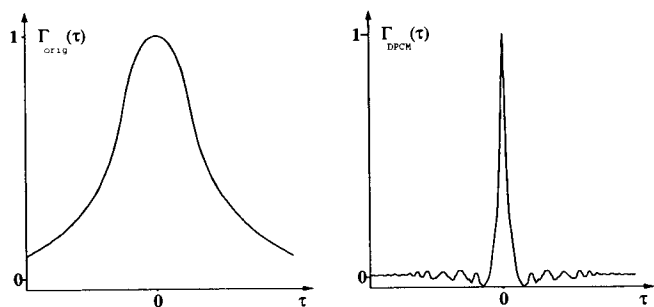


FIG. 2. — Correlation function before and after DPCM coding. The left-hand side plot shows a unidimensional correlation function between the pixels of a typical gray scale image, before any processing, versus  $\tau$ , the distance between those pixels. The right-hand side plot shows the correlation function obtained after application of a DPCM (i.e. at the output of the decorrelator).

*Fonction de corrélation avant et après codage par MICD. Le graphe de gauche représente la fonction de corrélation à une dimension entre pixels d'une image typique, en fonction de  $\tau$ , la distance entre ces pixels. Le graphe de droite représente la fonction de corrélation obtenue après DPCM (en sortie du décorrélateur).*

Therefore, to optimally code the signal (i.e. to use the minimum amount of bits), basic information theory implies to perform a conditional coding which takes all the former spatial history of the signal into account. Unfortunately, such coders are rather difficult to implement and time consuming. It is far more easy and fast to prearrange the pixels into less correlated coefficients which, according to information theory, do not

require any more conditional coders (i.e. the signal can now be fairly considered as generated by independent sources). The correlation function  $\Gamma_{\text{DPCM}}(\tau)$  obtained with a classical decorrelation process (DPCM, see later in the text), is presented in Figure 2 (right-hand side plot). In fact, the coefficients should rather be statistically independent, but decorrelation is the best that can be achieved with second order statistics. As the correlation is reduced, we may also expect that only non negligible coefficients carry relevant information. Therefore, the quantization errors introduced at the second step, should produce fairly minimal distortions on the reconstructed signal. An additional advantage is that it often happens that the decorrelator automatically rearranges the pixels into coefficients with physical (e.g. frequency) or perceptual meanings, and therefore are usually well suited for further processing.

2) A *quantizer* which reduces the accuracy on the coefficients down to a still acceptable quality. It may be implemented with a scalar or a vector quantizer [41, 62, 74], adaptive or not. Although it is sometimes difficult to dissociate from the decorrelator, it is only at this level that losses are introduced. Scalar quantizers reduce the numerical accuracy used to describe the new coefficients (i.e. by discarding least significant bits (LSB)). On the other hand, vector quantization gathers up neighbouring coefficients into vectors which are adaptively approximated by centroids for coding purposes.

3) An *entropy coder* which efficiently codes the quantized signal without any further distortion (i.e. lossless). It relies upon the statistical properties of its input bit flow [4]. The basic principle of an entropy coder is to code more (less) probable coefficients by shorter (longer) words (e.g. Huffmann code [23]). In consequence its complexity highly depends on the efficiency of the two preceding steps.

Although this paper is concentrated on still image decorrelation techniques, for some applications as television broadcasting, videotelephony or HDTV, it is also necessary to consider the temporal correlation which characterizes the redundancy between successive images. It guarantees to further decrease the associated bit rate. As those temporal decorrelative aspects are beyond the scope of this paper, we will just describe, in short terms, the motion estimation and compensation [2, 61].

Motion estimation techniques compute the relative displacement between *areas* present in successive images of a moving sequence. Usually, those areas are blocks of pixels in a block matching technique or the pixels themselves in a pel-recursive technique. They could also be regions produced by an object oriented coder. Only the difference between the current area and the motion-compensated area from the previous frame is then coded. This approach to reduce the temporal redundancy is very similar to the predictive coding described in Section II.1. except that the predictions are given by this motion-compensated areas. The *block matching* algorithm is a deeply experimented technique for mo-

tion estimation. It divides an image into the block pattern which better matches the corresponding blocks in the previous image, following a given criterion like, for instance, the mean of the absolute difference between the pixels of the blocks. Thereafter, we transmit the difference between those blocks and the associated motion vectors.

The purpose of this paper is limited to the description of still image *decorrelative* techniques. Coding schemes involving classical discrete cosine transform (DCT) provide good results for medium compression rates (between 15 and 20). In the same way, subband coding and recursive block coding (RBC) methods present a wider diversity of tools thanks to, among other things, wavelets, perfect reconstruction structures, efficient statistical models or perceptual decompositions. The efficiency of those *universal* techniques which are described in Section II is reduced for larger compression rates (greater than 25). To reach such rates, decorrelative techniques have to take *a posteriori* properties of the signal into account which means that these properties depend on the contents of each image, in opposition to properties from a statistical model based on a representative set of images, which we call *a priori* properties. This leads to the *adaptive* techniques which include methods adapted to the signal characteristics and feature extraction processes. The Section III is devoted to the description of those techniques.

## II. UNIVERSAL TECHNIQUES

The universal techniques (also more recently known as *generic methods*) perform signal processing independently of its specificity (i.e. TV images,...). They rely upon statistical *a priori* models of the signal, especially upon the expected high redundancy between adjoining samples (remember Figure 2). Therefore, they are always adapted to a class of images but never really to a particular image. The spatial and spectral interpretations of these methods contributed to their popularity in the eighties all the more so as their implementation complexity remains fairly acceptable.

Predictive methods were the first techniques proposed to take advantage of the correlation between adjoining pixels. However, the compression ratios were not large enough for a lot of applications. It was mandatory to better rearrange the signal. Stochastic function theory suggests that some linear orthogonal transformation should be powerful techniques : the signal is studied in a new domain, close to the one defined by the Karhunen-Loève transformation (KLT), which theoretically is the optimal decorrelative process. To satisfy the conditions necessary to efficiently approximate the KLT, the image has to be split up into blocks. Unfortunately, because of those blocks, unbearable effects appear at large compression ratios. Improvements implied the introduction of techniques applied on the whole signal (e.g. subband

coding) or on overlapping blocks (e.g. lapped orthogonal transform, recursive block coding). All those techniques allowed to increase the coding efficiency for a given compression ratio.

### II.1. Predictive and pyramidal coding.

Predictive [31] and pyramidal (Laplacian) [7] techniques may be considered as pioneers in this field. In the predictive difference pulse code modulation (DPCM), the signal value at a given moment is used as prediction for the next signal value. Therefore, we just have to transmit the error of prediction. In pyramidal techniques, the signal is split up into coarse approximations (obtained by lowpass filters) and error functions. The error signals are expected to be weakly correlated and/or suitable for further processing. This appears very clearly in Figure 2. Unfortunately, in the case of the pyramidal techniques, the total number of coefficients which have to be transmitted increases with the level of decomposition. It probably explains the reason for the lack of popularity of the pyramidal approaches, till the introduction of hierarchical subband coding which follows the same idea [49] (see Section II.3.).

### II.2. Orthogonal transforms.

The previous techniques have been partially supplanted by orthogonal transforms [1]. These transforms perform a change of basis : the transformed coefficients  $y$  result from the projection of a vector  $x$  made by the pixels of the original signal on an orthogonal basis defined by the matrix  $T$  :

$$(1) \quad y = T \cdot x$$

Orthogonal transforms are linear operations. Besides, thanks to their unitary property, their inverse transforms are defined by the transpose  $T^T$  of the transform matrix and the energy content is preserved. All this concurs to simplify implementation and analysis of orthogonal transformations.

Ideally, we should always perform a Karhunen-Loève transform (KLT) which leads to the optimal decorrelation. Indeed, this transform diagonalizes the covariance matrix of the signal to be coded [65]. As the basis functions are also the eigenvectors of this matrix, there is no fast implementation algorithm. Of course it excludes any practical applications! Therefore, other basis functions have to be chosen on the ground of *a priori* statistical models of the signal, in such a way that the associated coefficients of decomposition should be fairly well decorrelated. Among the wide variety of suitable transforms, we can mention the Hadamard transform, the discrete Fourier transform (DFT), the discrete cosine transform (DCT),...

The discrete cosine transform (DCT) [1, 78] optimally decorrelates first order Markovian highly correlated

signals [11] and it provides a good compromise between decorrelation and implementation complexity. Furthermore, it can be easily interpreted in the frequency space. Each basis function performs a bandpass filtering with a different central frequency. Figure 3 shows those

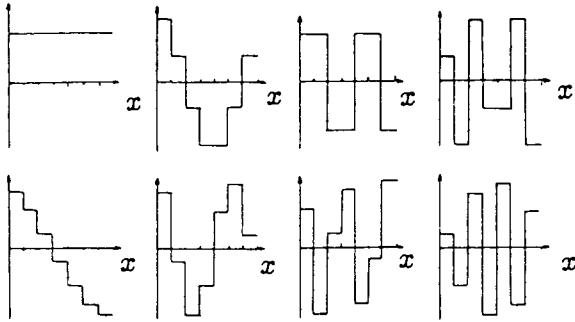


FIG. 3. — Basis functions of an eight order DCT.  
Fonctions de base d'une transformation en cosinus discrète d'ordre 8.

basis functions in the case of an eighth order DCT. Video applications involve *separable block* DCT : unidimensional DCT is successively performed along both coordinate axes of a block of pixels. Block size of 8 by 8 pixels yields good decorrelation and fast implementation. This explains why block DCT is an essential tool in the decorrelative coders of the CCITT JPEG and MPEG standards [2, 34]. Unfortunately, blocking and mirroring effects introduce unbearable distortions at very high compression rates [44].

The LOT (lapped orthogonal transform) [52, 53] has been introduced in order to reduce those parasitic effects. This transform is still linear and orthogonal, but the basis functions of one block slightly overlap on the adjoining blocks as shown in Figure 4. The block effects are eliminated at the cost of a longer computation time as the basis functions are twice more spread.

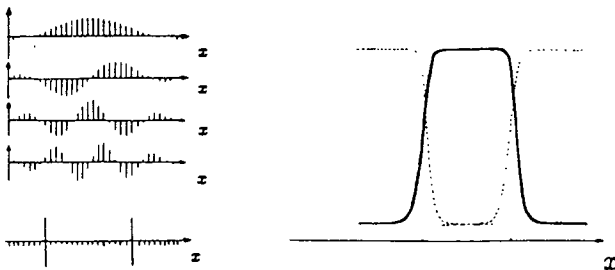


FIG. 4. — Basis functions and overlap window of a LOT. The left-hand side plot shows some basis functions of an 32th order LOT, applied on blocks of 16 pixels. The right-hand side plot shows the overlap window between adjoining blocks performed by the basis functions of the LOT.

Fonctions de base et fenêtre de recouvrement d'une transformation orthogonale à recouvrement (LOT). La figure de gauche montre quelques fonctions de base d'une LOT d'ordre 32, appliquée sur des blocs de 16 pixels. La figure de droite montre la fenêtre de recouvrement entre blocs voisins réalisée par la LOT.

### II.3. Subband coding.

Subband coding [12] has been presented as an efficient extension of the transform methods. The spectral interpretation of the DCT basis functions suggests to decorrelate with filters in the spectral domain. The idea of subband coding is to split up the signal into different frequency subbands which are separately coded. In the case of weakly stationary signals, it can be shown [65] that perfect decorrelation can be achieved between disjointed subbands. However, as there should still remain correlation inside the subbands, their number has to be large enough to ensure an adequate level of decorrelation. The goal of the subband decomposition is to obtain signals well conditioned for further efficient vector quantization or DPCM/PCM decorrelation. As the signal is processed as a whole, compression applied on the different subbands does not lead to blocking effects on the reconstructed signal as encountered with a block DCT transform.

In practice, the signal is split up into different frequency subbands by bandpass filters. The resulting signals are downsampled in order to keep constant the total number of samples. For the reconstruction, the different subbands are interpolated by adequate zero insertion and bandpass filtering. For example, a two-subbands scheme is shown in Figure 5. This scheme is composed by low- and high-pass filters  $H(z)$  and  $G(z)$  followed by downsamplers by 2 in the decomposition stage, and upsamplers by 2 followed by adequate filters in the reconstruction stage.

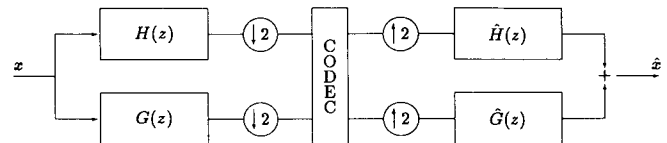


FIG. 5. — Two-channels subband scheme.  
Schéma sous-bande à 2 canaux.

It can be shown that orthogonal decompositions into wavelet packet bases [45, 49] are in fact particular subband decompositions. Wavelet transform and wavelet packet transform are new time-frequency analysis tools [14]. For images, the terms time-frequency have to be understood in the broader sense of space-spectrum. Also, any orthogonal block transform can be easily implemented with a subband coding scheme [78].

In the field of video applications, the subband decomposition is more often implemented with a separable scheme : unidimensional filter banks, like the one shown in Figure 5, successively operate in the horizontal and the vertical directions. However, non-separable decompositions have been recently introduced in order to reduce the weaknesses of separable approaches. Both subband decompositions are described below.

#### II.3.1. Separable subband decomposition.

As separable subband decompositions successively act in the horizontal and the vertical directions, we will

only discuss a unidimensional subband scheme. In fact, the basic tool of such decompositions is the above-mentioned two-subbands scheme (see Figure 5).

The filter banks are designed to cancel out aliasing errors introduced by downsampling which yields an overlapping of the repeated spectra at the output of the finite impulse response (FIR) filters. The quadrature mirror filters (QMF) [12] approach was introduced to fulfill such antialiasing requirement. Their low-pass  $H(\Omega)$  and high-pass  $G(\Omega)$  filters are symmetric around  $\Omega = \pi/2$ . The classical QMF proposed by Crochiere and designed by Johnston [33] produces amplitude modulation. The elimination of residual modulation errors is achieved by lossless subband schemes. The additional perfect reconstruction constraints lead to the design of new lattice filter structures — with or without linear phase — and poly-phase representations [77, 83].

Until recently, the design of filters was commonly based on the half-band approach [12] which minimizes the difference between ideal and designed frequency filter response. Wavelet theory [69], besides the fact that it is the background theory of subband coding, provides a new approach for filter design : spline subband analysis [49] and FIR regular filters [13]. With those filters, it is possible to control the mathematical regularity and stability of the decompositions. It seems to have some important effects on the perceptual quality of the reconstructed images, for a given quantization process.

Separable subband decompositions are practically obtained by cascading the above-mentioned two-bands schemes. There are three types of subband decompositions (see Figure 6) :

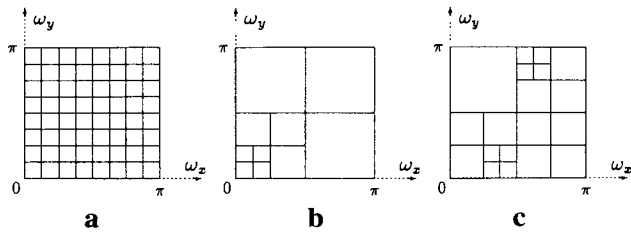


FIG. 6. — Uniform (a), hierarchic (b) and optimal (c) decompositions.

*Décompositions uniforme (a), hiérarchique (b) et optimale (c).*

— *uniform decomposition* : the decomposition scheme is iterated on each subband. The resultant decomposition is fairly similar to the spectral location of the block DCT basis functions ;

— *hierarchical or wavelet decomposition* : at each step, only the low frequency subband is further decomposed. This decomposition is well consistent with the space-scale nature of the human visual perception which seems to decompose the signal into octave scale bands [56]. Furthermore, such transforms provide good decorrelative properties thanks to the finer decomposition produced in the more energetic low frequency areas ;

— *optimal intermediate decompositions* : the different subbands are decomposed provided that a given *a priori*

criterion is fulfilled [79]. This ends up with a pattern intermediate between the two first types.

**II.3.2. Non-separable decomposition.**

Although the separable subband decomposition is quite popular, subband coding does not have to be separable. It can also be achieved by means of direct convolutions with two-dimensional non-separable filters. As the separability condition reduces the number of degrees of freedom, they might be less suited to the 2-D specificities of video signal. Moreover it is well known that visual perception separates signals in channels, called perceptual components, tuned with respect to the position, the frequency and the orientation [64]. Separable subband decomposition does not sufficiently take those perceptual effects into account : the diagonal band contains mixed orientations (i.e.  $+45^\circ$  and  $-45^\circ$ ). Thanks to their more general characteristics, non-separable decompositions enable to discriminate between those orientations.

Two different approaches are taken into consideration : quincunx and hexagonal decompositions. Figure 7 shows such hierarchic decompositions. In the quincunx case [20, 84] (Fig. 7b), the signal is split up into two subbands with a diamond shaped lowpass band. This decomposition better fits the 2-D hyperbolic shape of any video energy spectrum. Hexagonal decompositions [75] (Fig. 7a) provide better orientation analyses. The signal is split up into a low frequency and three high frequency subbands oriented towards  $-30^\circ$ ,  $30^\circ$ , and  $90^\circ$ . Contrary to separable decompositions, the orientations are now well separated.

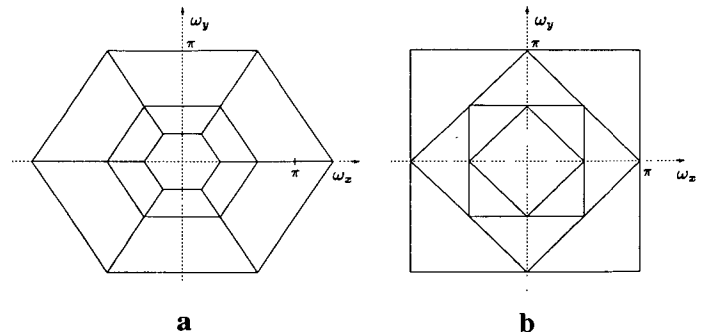


FIG. 7. — Non-separable hierarchical subband decompositions : hexagonal (a) and quincunx (b).

*Décompositions sous-bande hiérarchiques non séparables : hexagonale (a) et quinconce (b).*

The difficulty with those non-separable decompositions is the relative incompatibility between the desired, often even special, shape of the subbands and the perfect reconstruction constraints which are drastic [35]. For example, in the hexagonal case, orthogonal QMF banks [75] are mandatory just in order to cancel any aliasing error. A complex optimization is thereafter required to cancel the remaining modulation errors. Perfect reconstruction lattices provide a graceful way to reduce the complexity of the synthesis process [37].

Other approaches involve orthogonal wavelets [25, 37]. Nevertheless, insufficiently developed techniques for 2-D filter design as well as too complex implementations restrain the performances of non-separable subband decompositions.

At the present time, separable subbands decompositions have reached a degree of maturity which allows their introduction in the next generations of normalized coders (MPEG coder for example). Some researches are now devoted to the design of adaptive filters or decompositions like those described in Section III.1.1. On the other hand, non-separable subbands decompositions rather belong to the experimental domain.

#### II.4. Recursive block coding.

We should also mention the recursive block coding (RBC) technique [19]. It is a general formulation which encompass, among other things, transform coding, predictive coding, quadtree coding, applied on a residual error (of prediction) signal instead of the original signal.

RBC has been proposed as a solution to the blocking effects which appear in a lot of image processing techniques. In those techniques, the blocks were usually introduced in order to satisfy, not only implementation constraints, but also some hypotheses which are only locally valid. Like, for instance, the Markovian behavior, in the case of the DCT [11, 44]. The blocking effect appears when the manipulations applied on adjoining blocks have too unlike effects near their boundaries. Improvement implies that joined information is introduced within those adjoining blocks. In RBC, the image is divided into overlapping blocks. The preceding boundary values, common between adjoining blocks, are used to predict the content of the whole block, with an *a priori* minimum variance representation (MVR) (i.e. a Markov chain model) of the signal to be coded. In other words, as illustrated in Figure 8, two sides of a block are used as input for the predictor of the pixels inside the block while the two others are used as partial prediction of two adjoining blocks. The algorithm of block division

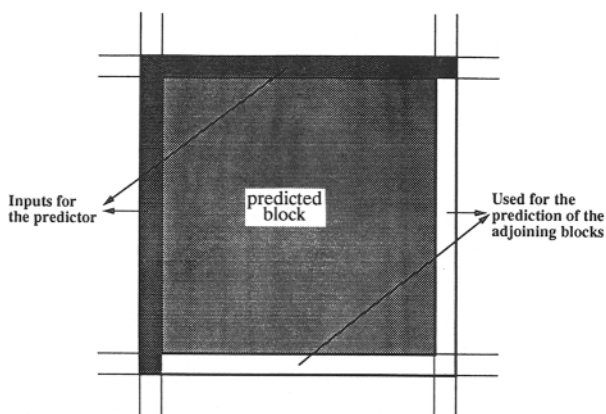


FIG. 8. — The overlap between adjoining blocks in RBC.

*Le recouvrement existant entre les blocs utilisés en codage en blocs récursif.*

and prediction may be adaptive or not. For example in adaptive quadtree RBC shown in Figure 9, the image is initially divided into blocks of a certain size. The error of prediction of each block is computed and coded. If, for some blocks, it requires too many bits, they are further divided into four smaller blocks and the process is iterated on each of those new blocks until the associated predictions are estimated accurate enough.

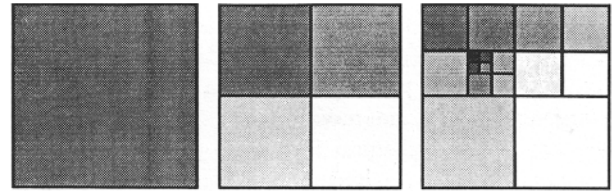


FIG. 9. — Adaptive quadtree RBC. In quadtree RBC, the size of the blocks is adaptive. We start with the block presented on the left. The boundaries are used to predict the block content and the error of prediction is quantized and coded. If it still requires too many bits, the decision is taken to split up the block into smaller blocks (middle). The process is iterated until the prediction is graded satisfactory for each block. On the right, we end up with a block decomposition which is optimal for this RBC algorithm.

*Codage en quadrant récursif adaptatif. Cette approche permet d'auto-adapter la taille des blocs. Nous partons du bloc présenté à gauche. Les bords sont utilisés pour prédire le contenu du bloc et l'erreur de prédiction est quantifiée et encodée. Si cela demande encore trop de bits, la décision est prise de diviser le bloc (centre). La procédure est ainsi itérée, jusqu'à ce que la prédiction soit jugée satisfaisante pour chaque bloc. La configuration ainsi obtenue (droite), est optimale pour l'algorithme utilisé.*

As expected, the main advantage of such techniques is analogous to subband coding and to the LOT : because of the overlapping between the different blocks, most of the distortions inherent in transform coding (essentially the blocking effects) are now dramatically reduced for a same compression rate.

### III. ADAPTIVE TECHNIQUES

Universal techniques provide good results for medium compression rates (between 15 and 20) as those which are used in CCITT standard coders for TV broadcasting [2, 34]. However, videotelephony and HDTV need appropriate decorrelation algorithms adapted to the nature of the information. For several years, image coding covers other fields. For example, very high resolution images occur more often in teledetection or medical analysis. Methods based on feature extraction — like edges, shapes, textures or even more local information (e.g. singularities) — may be more suitable for those new applications. Anyway, this kind of images requires new methods for efficient, lossless or not, storage. Image processing also takes various forms, like object detection and recognition, computer vision, image synthesis,...

The so-called adaptive techniques are decorrelative methods which take the *a posteriori* contents of the images into consideration. They do not any more rely

only upon statistical properties of an image class. Two ways are presented in this paper : the adaptation of universal techniques to the signal contents and feature extraction methods which represent an image like a set of characteristic points, or like a set of objects. The coding process of adaptive techniques is sometimes rearranged : it starts with an analysis, followed by a compression which preserves entities or properties detected during the first step. For example, in radiology, it might happen that we are only interested into the detection of lesions. Once those important regions have been detected, their degradation is minimized, whatever will happen to the other regions. Those very particular techniques are often not any more universal (i.e. not appropriate for any type of image).

**III.1. Adaptation of universal techniques.**

**III.1.1. Adaptive subband coding.**

The design of subband filters is most usually based on the synthesis of bandpass frequency responses. Nevertheless, other processes based on optimization criteria have been recently proposed to adapt the filter design to the signal characteristics [80]. One can mention for instance : entropy minimization [15], lateral lobes limitation, minimization of the product of the variances of the different subbands augmented by the reconstruction errors,... The goal is always the same : to be able to tune the filter bank like a fairly good approximation of the KLT, which implies that output signals should be as decorrelated as possible. We have already mentioned that all the block transforms can be performed as subband decompositions at a relatively low computation cost. In consequence, this adaptive approach should provide the degrees of freedom required to drive out efficiently the KLT.

The adaptive best basis subband decompositions are also promising processes. They are similar to the classical subband coding schemes because, at first, the signal is uniformly decomposed down to an *a priori* determined level as illustrated by the top row in Figure 10. At each step of this process, the intermediate subbands are quantized and they may even be entropy coded. A measure of the coding efficiency, denoted  $\mu(\cdot)$  (e.g. the number of negligible coefficients or an optimal linear combination between the amount of bits after quantization of the subband and its associated error of reconstruction,...) is computed on each of the subbands and stored in memory. In the original algorithm proposed by Wickerhauser [85], the different values of  $\mu(\cdot)$  are stored in a tree structure, where each cell, associated to a subband  $f$  is the father of the cells associated to the new subbands obtained by subband decomposition of  $f$  (see Figure 10, left bottom). When the whole uniform decomposition is completed, the measure functions are compared and the optimal intermediate subband decomposition is determined. If necessary, some subbands are recombined accordingly. It ends up with a subband algorithm where the

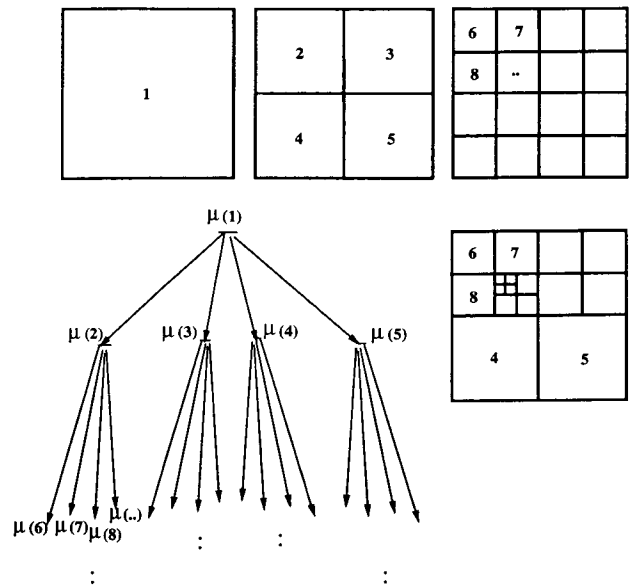


FIG. 10. — Adaptive best basis subband decomposition.  
Décomposition sous-bande optimale adaptative.

different subbands at each level of decomposition have been further split up provided that the given *a posteriori* criterion is fulfilled (Fig. 10, bottom right).

**III.1.2. Hybrid methods.**

The hybrid methods of decorrelation are combinations of predictive methods, RBC, orthogonal transformations and subband decompositions.

An often mentioned hybrid coding consists in coding with DPCM the coefficients which are obtained by transform coding of the original signal [30]. It was initially developed for speech processing with efficient vocoders [38]. The performances rely on the well suited statistical properties of the signal obtained after transformation.

An other example has been introduced for classification purposes of similar still images (e.g. face pictures). It combines an adaptive best basis subband coding algorithm (see Section III.1.1), with a KLT transform applied on the few remaining low energy coefficients while all the other coefficients are set to zero. Images of the same type can be coded on a very small number of coefficients [86].

We also mention optimal hybrid transform-subband coding schemes [46] which are subband best basis decompositions with decision criteria (similar to those used for adaptive best basis subband coding) to apply further subband decompositions on a spatial subband or on its transform. In Figure 11, we see the operators  $\hat{A}$  and  $\hat{B}$ , applied on the subbands. They have to be well chosen in order to allow perfect reconstruction during the synthesis process. In particular, there is the optimal DCT-hybrid subband coding decomposition (i.e. where  $\hat{A}$  and  $\hat{B}$  are simultaneously the identity or DCT operators) which optimally and adaptively splits up the signal in smaller regions (spatial or frequential) with adequate properties

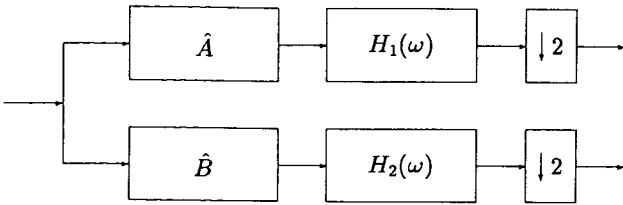


FIG. 11. — Decomposition stage of an hybrid transform-subband coding scheme. It is a classical two-subbands coding stage, but additional operators  $\hat{A}$  and  $\hat{B}$  may be applied on the subbands.

*Etape de décomposition d'un schéma de codage hybride sous-bande-transformation. C'est un étage à deux sous-bandes classique, avec des opérateurs supplémentaires  $\hat{A}$  and  $\hat{B}$  qui peuvent s'appliquer sur les sous-bandes.*

for further decorrelation or coding. The decisions are feedbacks from the entropy coder : a subband is further decomposed and an operator is applied on this subband, only if it significantly reduces the associated bit flow after quantization. It is essentially a best basis algorithm with additional degrees of freedom provided by the subband operators. After optimal decomposition, additional operators may be applied on and/or between the different subbands (e.g. DCT, DPCM,...).

We also have to point out the partial subband predictions [47] as will be described in Section III.2.1. Hybrid subband coding can also be used to model the signal properties (e.g. determination of statistical properties, recognition and identification,...).

In the same way, Mallat [51] has introduced the concept of local time/frequency multilayer orthogonal transforms. A large set of basis functions associated to different classical transformations (e.g. subband coding, Gabor transform, block DCT,...) are stored in a dictionary. The signal is iteratively decomposed into successively best matching subsets of functions belonging to the dictionary. Such an algorithm leads very easily to an optimal combination of the different techniques used to build the dictionary.

**III.1.3. Adaptive non regular sampling.**

Adaptive non regular sampling implies that the signal is sampled at locations which are not regularly spaced. There essentially exist two classical approaches : non-uniform sampling [32] and zero-crossings [29]. The distinction between both of them is somewhat confused. Let us say that for non-uniform sampling, the sample locations are chosen *a priori*, while in zero-crossings, we transmit the positions where some events (decided *a priori*) happen.

Shannon's sampling theorem [38] implies that the uniform sampling rate is determined by the highest frequency contained in the signal. Somehow, it seems intuitively more interesting to try to apply Shannon's sampling theorem on smaller portion of the signal. If uniform sampling was performed in these regions, the bit rate would depend upon the local frequency contents and it would increase whenever the maximum frequency increases and conversely, as it is suggested in Figure 12.

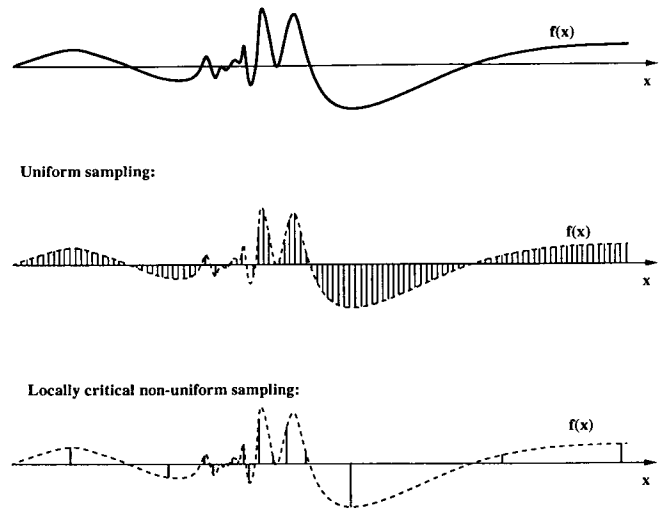


FIG. 12. — Adaptive non-uniform sampling. The signal to sample is shown in the figure at the top. In the case of uniform sampling (figure in the middle), the sampling frequency is determined by the highest frequency in the whole signal. In locally critical non-uniform sampling (figure at the bottom), we try to use a sampling frequency determined only by the local highest frequency content.

*Echantillonnage non uniforme adaptatif. Le signal à échantillonner est illustré à la figure du dessus. Dans le cas de l'échantillonnage uniforme (figure du milieu), la fréquence d'échantillonnage est déterminée par la plus haute fréquence présente dans le signal considéré dans sa totalité. Dans le cas de l'échantillonnage localement critique non uniforme (figure du dessous), nous essayons d'utiliser une fréquence d'échantillonnage déterminée par la plus haute fréquence localement présente.*

We decide to speak of local methods when such an approach is attempted. On the other hand, we speak of global methods when the sampling processes are applied on the whole image.

Global methods lead to a data flow roughly equal to Nyquist's critical density as for uniform sampling. To improve the efficiency, it is mandatory to have a more local approach, according to our intuitive idea presented in Figure 12 (i.e. to sample near the critical rate for the local frequency content [48]). Maxima or other characteristics detection in subband coding, or even after other transformations, is a potential way to increase the sampling rate only where there are relevant singularities [50]. For example, the Figure 13 sketches

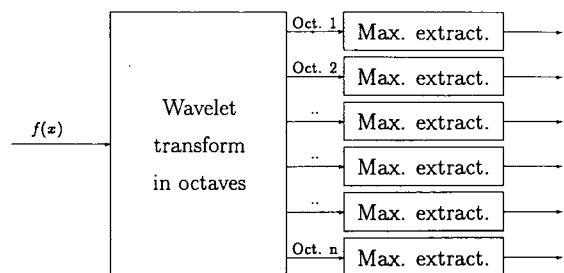


FIG. 13. — Mallat's maxima wavelet representation. An approximation of the locally critical non-uniform sampling is obtained by taking the wavelet transform (at each octave) of the signal. For each octave, only the locations and the amplitudes of the wavelet transform are kept.

*Représentation par maximum des ondelettes de Mallat. Une approximation d'un échantillonnage localement critique non uniforme est obtenue en prenant la transformée en ondelette (à chaque octave) du signal. Pour chaque octave, nous conservons uniquement la position et l'amplitude de la transformée en ondelette.*



Mallat's wavelet maxima representation algorithm. The wavelet transform is computed for each octave, and only the locations and amplitudes of the local maxima of the wavelet transform are kept. This algorithm leads to results very close to our intuitive goal : samples are taken only at the location of the singularities of the image.

This is closely related to Marr's conjecture [55], which claims that all the valuable information is contained in the positions, the value, the steepest slopes and the orientations of the edges in a well smoothed version of the original signal. There exist sets of signal characteristics which ensure almost perfect reconstruction [48, 50]. For instance, we could mention the local maxima of the wavelet transform in the case of the algorithm described in Figure 13. All those approaches require iterative reconstruction processes based on Banach fix point theorem, and such a process is always time consuming. If degradations are tolerated, some subsets of characteristics can be neglected. By progressively relaxing the constraints, it becomes possible to reach huge compression rates for still images (between 80 and 120 and even more!), with degradations which are optimally kept to a minimum or at least well controlled from a perceptual point of view [22, 50]. Efficiency can perhaps still be further improved with hybrid methods.

### III.2. Feature extraction methods.

Attempts have been made to include more physical entities such as edges, objects or regions. It becomes mandatory, in order to reach higher compression rates which imply degradation of the image, to develop algorithms able to detect, classify and select the relevant characteristics.

Here appears a difficulty inherent in all feature oriented coding scheme : there is no way to detect *all* of those features [66]. The problem lies somewhere between image coding, pattern recognition and psychophysics [43]. One naive, but in our context, an extremely efficient way to work, is to consider that the only relevant features are those which deserve a coding purpose. And this set of features is not necessary the one which an observer would *see* in an image.

Increasing attention has been devoted to singularity detection in an image, partially as the eye is very sensitive to abrupt variations [55]. Also, because in domains like medical imaging, robotic vision or remote sensing, the important information, which has to be transmitted, is sometimes only the location or the type of different singularities. The alteration of the global *aspect* of the image is, in that case, of far less importance.

Therefore, in a first approach, singularities are preserved throughout the whole process. Unfortunately, spectral methods consist in narrow filtering the signal to extract the high frequency content as singularity tracks. As singularities usually have a broad spectrum, those spectral methods spread profile changes over adjoining pixels, in a way which is often excessive. In feature

oriented algorithms, the local changes are, by definition, preserved (or even accentuated).

An other approach follows a more global point of view : an image embodies regions with uniform properties (e.g. brightness, color or texture,...). Their boundaries are among the local singularities, defined here above. Therefore, it seems intuitively possible to detect those *edges* by splitting up the signal into different regions with similar properties. This should lead to a division into closed regions. Unfortunately, as it is often not satisfied, typical region algorithms are necessary.

After a presentation of the techniques used for singularity detection and coding, we will discuss the object oriented approaches.

#### III.2.1. Detection and characterization of local singularities.

Since Marr's works [55], it is commonly admitted that a lot of information on the singularities is included in different frequency channels.

Wavelet theory confirms this approach and allows to accurately understand how this information is distributed throughout the frequency domain [27]. This information is also geometrically correlated [47] (i.e. the shape of a singularity and the associated content in the different subbands is correlated) and this can be used to reduce the coefficient flow, in subband coding, by partial predictions of the different subbands corresponding to the different orientations or to the different scale levels (see hybrid subband coding). It implies that similar patterns appear inside the different subbands and this explains the good results obtained with vector quantization applied on the output signal [3].

The human visual system is particularly sensitive to one type of singularities : the *edges* [55]. We define edges as locations of wide variations in intensity, in one dominant direction (corners are intersections of two, or more, edges). There is a wide range of edge detectors (which may be used as a part of segmentation algorithms) : Marr's Laplacian zero-crossings, high frequency filters, directional decomposed based coding methods, non-linear variational methods, local non-uniform sampling, energy content maximization, neural networks, non-linear and morphological filters,... We briefly describe the principle of those methods :

— *Marr's Laplacian zero-crossings* consists in the following process [55]. Firstly, the image is fed through a bandpass gaussian filter. Secondly, the edges are detected as inflection points of the resulting signal (zero-crossings). Actually in his original algorithm, Marr uses zero-crossings on the signal obtained after filtering with the second derivative of the Gaussian, but this remains to the same as what we have presented so far. This method is a particular case of the *local non uniform sampling edge detection* [22, 48] which has been described in Section III.1.3. : the characteristics points (e.g. local maxima of the wavelet transform) used to reconstruct the signal can be considered as the edges of the image [50]. This method does not only characterize the position of the edges but also their profile.

— With *highpass filters* and *directional decomposed based coding methods*, the signal is fed through high-pass filters (more or less directional, depending on the algorithm) and the edges are detected by thresholding [39]. The main informations produced by this method are the location and the orientation of the edges.

— *Local non-linear variational methods* consists in edge detection by minimization of *pseudo-energy functionals* [8, 9]. Those functionals are defined by analogy to the elasticity equations of thin membranes and straps. The energy contains an internal contribution which constrains the smoothness of the curve, a conformal contribution which ensures the conformity between the original signal and its segmented version, and an external contribution which pushes the contours towards the salient image features. An additional contribution is mandatory to limit the cumulated edge length or the number of edges.

— In *energy content maximization*, edges are located at positions where the local energy content is maximum [70, 82]. The difficulty is, of course, to define an appropriate local energy content. Some definitions based on the Hilbert transform have given excellent results. For instance, we can define the local energy content as the sum of the squared signals obtained by convolution with an even and an odd filter [82].

— In *neural network edge detection*, the local spatial average values of the signal are computed. Edges are automatically drawn, by use of a neural network, where the local average changes too fast [57, 58]. It is a way to implement the image pre-processing performed by a human retina. For this application, the neural networks are not used for any type classification or learning purpose, but rather like a non linear filter!

— In signal processing, one usually calls a filter any operation which is linear and translation invariant. It is a well-known result that the filter turned to be the convolution product of the signal by a (generalized) function. In practice, the filters are often transformed into more or less bandpass filters, a process that makes an identical further filtering insignificant in the case of an ideal filter; this property is called *idempotence*. A transformation  $\phi$  is linear when  $\phi(f + g) = \phi(f) + \phi(g)$ . Such a condition is apparent when the structure of the physical phenomenon is itself linear. However, although many signals combine additively, visual signals obey a different law. Furthermore edges often have a large spectral content. As a consequence the filter only smooths the signal, which makes the detection of edges even harder. Instead of linearity, one could require the preservation of inclusion relations which may exist between objects. A very powerful theory was only constructed upon the two axioms of idempotence and inclusion (which is incompatible with linearity) : the theory of *morphological filters* [73]. A very simple morphological filter cascades a *minimum* operator and a *maximum* operator both taken on a  $3 \times 3$  square. If we compare gray values to elevations on a topographic surface, the minimum operation will cut small peaks and the maximum removes all the narrow valleys. To detect

the edges one just has to compute the difference between the original image and the filtered image.

We want to mention that those edge detection algorithms are also often used for image analysis instead of image decorrelation and coding.

### III.2.2. Object approach.

Object approaches attempt to describe an image as a collection of closed sets (also called *regions*) of pixels which share a common property [66, 71]. Each region is characterized by its shape (*contour*) and its aspect (*texture*). These entities seems to be more natural as they coincide with psychological concepts of vision. However, as the objects serve coding purposes, the coder algorithm may detect, code and transmit object informations different from those naturally indicated by a human observer. For some kind of scenes, like a forest scene, it is even impossible to find a unique collection of regions. This is a consequence of the difficulty to define the concept of a region. Nevertheless, in specific applications (e.g. videophony) or for images which are easy to describe, the scientific literature proposes models for regions and give compression ratios [5, 39, 76].

Figure 14 is the block-diagram of a simple object oriented coding scheme for still images. It contains at least three steps : the segmentation, the contour coding and the texture coding.

#### III.2.2.1. Segmentation.

The aim of image segmentation is to divide the image into several meaningful areas, resulting in a set of regions of any shape. The boundaries of those regions are the contours of the object.

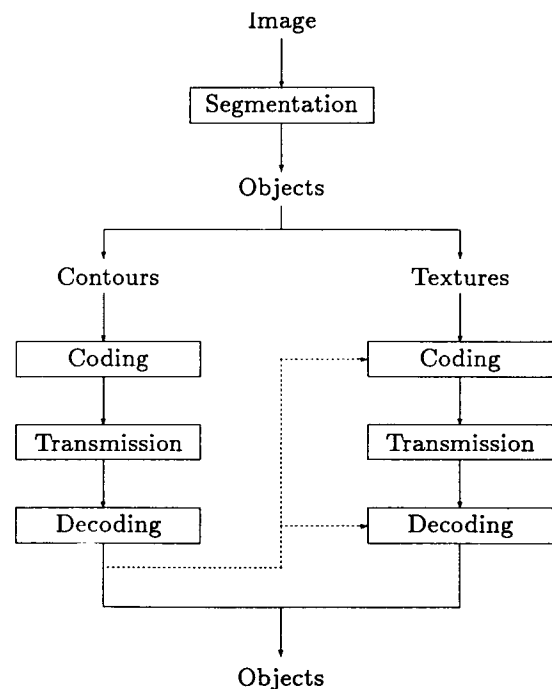


FIG. 14. — Block-diagram of a simple object oriented coding scheme.

*Schéma d'un codeur simple, orienté par la notion d'objet.*

The principal techniques [71] of object oriented segmentation are the *top-down* and the *bottom-up* (also called *region growing* methods). Top-down techniques start with the whole image as a unique region. During the segmentation process a region is split up into smaller regions (e.g. in a quadtree decomposition), whenever the region seems to contain non uniform regions. On the contrary, bottom-up methods regroup regions which have analogous texture contents. Sometimes, they are gathered together like for the *split-and-merge* technique. The uniformity criterion depends upon the nature of the image under study; in most models, it is a combination of different stochastic statistics (mean, variance,...) estimated on each region. We caution against mixing parameters with different units as unfortunately done for many criteria proposed in papers. Images which have to be segmented on the basis of their macroscopic patterns, also called textures, require extremely complex models [36]. Gabor filters are intensively used for the segmentation of patterns that show significant spatial organization characteristics [6, 16]; their use leads to interesting results.

The set-theoretical body of mathematical morphology [26, 72] proposes also different segmentation techniques which are region growing methods. They are all based on the *watershed transform* [59]. The best way to apprehend this tool is to return to the topographic interpretation of an image function : the lighter the gray tone of a pixel, the higher the altitude on the topographic surface. To each minimum is associated a catchment basin. If a drop of water falls at a given location on the surface, it will follow the steepest slope and reach a local minimum. As we expect, a region is defined as the set of all points which have the same attracting minimum and its contours are located on separation lines where a drop of water is equally attracted by two adjoining basins. The major drawback of the method — although it is not limited to this one — is the occurrence of severe over-segmentation. Many techniques are proposed to overcome this common inconvenient (e.g. pre-processing filtering).

#### III.2.2.2. Contour coding.

To achieve significant compression ratios, the coder has to optimize the code related to the contours because it often represents about 80% of the total bit rate. For that reason, a multitude of techniques have been developed for efficient shape description; they belong to two families : the *external* and the *internal* shape description algorithms.

The *external* family consists into contour descriptors : they code along the contour and apply even when a contour is not closed, so that they can also be used for the coding of singularities.

The most widely used coding techniques are *run-lengths* [10] — the ones applied for facsimile transmission [28] — 4- or 8-connected *chain codes* [21, 42, 60] and piecewise linear approximations [17] or *curve approximations* (Fourier descriptors, polynomial approximations, etc).

- A run-length coder transforms sequences of successive identical symbols into numbers which give each sequence length. For example, the one-dimensional binary signal 00111000011 would be transformed in the byte sequence : 2342. In the case of two-dimensional signals like contours, the coder mostly works on the contour direction sequences.

- Chain coders too use the contour direction elements : they relate each direction to the previous one and describe it as an element of a given set of possible directions, i.e. the 4-connected set {forward, right, backward, left}.

- The line or curve approximations of borders constitute the last category among the external methods; the contours are divided and approximated with several straight or curved segments whose relative anchor positions are transmitted by the coder. Although convenient for large regions, the method lacks of efficiency for small regions.

A frequently used reference for the measure of the efficiency is *1.5 bit per contour point* plus starting points [18].

Since they follow lines, the external methods are usually easy to implement. However, a disadvantage of simple boundary coding is that all the regions boundaries are coded to represent region locations. This causes no difficulty for large regions but in the case of small regions, especially for one pixel wide regions or irregular contours, the shape cannot be treated in the same way.

The second class of shape representations is the category of *internal* or *area* descriptors. This time all the efforts are concentrated on coding the surface of each region, whose borders form the contours. Usually the shape is transformed into an alternative spatial representation. Many articles are concerned with the morphological *skeletons* [54, 68, 72]. The skeleton of a region is somewhat like a human skeleton : a thinned version of the external shape. To reconstruct the original border in the Euclidean space, it is sufficient to associate to each skeleton point the radius of the largest circle included in the region (this is not the unique way to define the morphological skeleton). In a digital space, the circle is replaced by a square, a hexagon or any convex set. The contour representation it produces is a collection of triplets  $\{(x, y, r)\}$ , where  $(x, y)$  is the position and  $r$  the size of the scaled form used during the skeleton extraction. Unfortunately the lack of experiments does not really permit to conclude about the efficiency.

#### III.2.2.3. Texture coding.

A good texture coding should interact with the contour coding since these two entities are not totally independent. This is why, as shown in Figure 14, the texture coding branch gets the contour information. The result is that the receiver has to decode first the shape before the texture. As an outcome, texture coding is not a well studied problem in opposition to contour coding and to texture analysis. The latter subject has throughout been explored since the end of the sixties. The aim of

texture analysis is to determine parameters which exactly describe the signal. An associated compression technique transmits the parameter values only and the receiver synthesizes the texture with respect to a model and to these parameter values.

Up to now only a few papers have considered the task of coding the texture of arbitrarily shaped regions. As usual, they work in the spatial domain or in a transform domain.

*Spatial methods*, like *polynomial representations*, are almost the only ones which have been developed [36, 39]. In a polynomial representation, the texture is approximated by a two-dimensional polynomial function; the order of the polynomial function and the coefficient values are chosen to minimize the approximation error. Although this kind of algorithms produces high compression ratios, yet obviously many image details are lost, especially in dense regions. To counter a dissatisfactory restitution we either could increase the amount of regions or the polynomial order. The first solution is very costly since contour coding techniques do not accommodate to small contour segments. An increase of the polynomial order is not a constructive solution because of a higher sensitivity to noise. In practice the polynomials are limited to the second order which means that any texture is described with a maximum of six parameters and that they can not be applied on a large variety of textures.

In opposition to spatial methods and in agreement with the interesting frequency interpretation, we can code textures by means of transform coefficients. However, in order to avoid that the texture spectrum includes part of the border information, spectral methods require an *extrapolation* of the texture on a regular domain (e.g. a rectangle) or the definition of *generalized orthogonal transforms* :

(i) Figure 15 illustrates how the shape spectrum alters the texture spectrum. The upper-left image is the texture defined on his support; it is a 8-bit image enclosed in a square of size  $64 \times 64$ . The observed texture  $o(x, y)$  can be modeled as the product of a *window function*  $w(x, y)$ , which takes the value 1 inside the window and 0 outside, by an extrapolated textured signal  $e(x, y)$  which would cover the whole circumscribing rectangle; the upper-right image of Figure 15 shows such a signal. In the spectral domain, the previous product is equivalent to the convolution product  $\mathcal{O}(u, v) = \mathcal{W}(u, v) \otimes \mathcal{E}(u, v)$ . As clearly indicated by this last equation, the best-suited spectrum  $\mathcal{E}(u, v)$  for coding purposes is affected by the window spectrum  $\mathcal{W}(u, v)$ ; it is why a spatial *extrapolation* technique, which tries to restore  $e(x, y)$  from the observed signal  $o(x, y)$ , is also called a *deconvolution* technique. Images (c) and (d) of Figure 15 compare two coded textures : (c) is the original texture reconstructed with only 50 coefficients of the Fourier transform; the other coefficients are equal to zero. The last image represents the reconstruction with 50 Fourier coefficients after the original texture has been extrapolated by a deconvolution process described in [81]. As it can be seen, if we fix the amount of coefficients or the compression

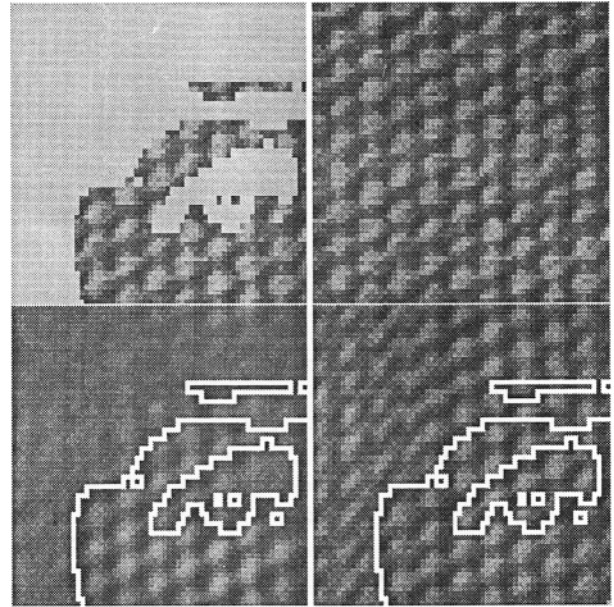


FIG. 15. — Extrapolation of a regular texture to a square. The first signal represents the texture on an arbitrarily shaped region. Picture (b) is a possible extension of the texture which contains no information about the shape. (c) is the signal reconstructed with 2% of the total amount of Fourier coefficients without extrapolation. (d) is the same as (c) but after extrapolation.

*Extrapolation d'une texture régulière à un carré. Le premier signal représente la texture sur un segment de forme arbitraire. La figure (b) est une extrapolation possible où n'est présente aucune information de forme. (c) est le signal obtenu avec 2% des raies du spectre de Fourier, sans extrapolation. (d) correspond à (c) mais avec extrapolation.*

ratio, the extrapolated texture is closer to the original than the reconstructed texture with no extrapolation.

(ii) Gilge *et al.* [24] develop a scheme which decomposes the texture signal into an orthogonal basis adapted to each region whatever its shape is. This is done by redefining the inner product : the *generalized inner product* between functions  $f(x, y)$  and  $g(x, y)$  is  $\int \int_{\mathcal{D}} w(x, y) f(x, y) g(x, y) dx dy$  where  $\mathcal{D}$  and  $w(x, y)$  are respectively the smallest rectangle that circumscribes the region and the window function. The choice of, for example, cosine functions defined on  $\mathcal{D}$  for  $f(x, y)$  and  $g(x, y)$  leads to an orthogonalization process and eventually to a spectral texture description adapted on each region shape. Unfortunately the receiver has to compute the basis functions again because the region dependent function  $w(x, y)$  was used during the orthogonalization process. This is not suitable for real time applications.

Because of the additional overhead needed for the transmission of the region borders, an object oriented coder is not necessarily more efficient than the other techniques presented here above. In addition there is a risk to produce complex algorithms. Powerful representation scheme for contours and textures are mandatory, in order to achieve more efficient compression ratios.

A primal interrogation is the evaluation of the quality in object oriented coding schemes; i.e. how does a human perceive shape deformation ?

An other outlook is the integration of the shape coding into a multiresolution scheme like it is done for progressive transform coding. Theoretical ideas were suggested

but not deeply tested. We also need improvements for texture coding. Leou [40] developed a simple idea : the two-dimensional texture signal is ranged into a row by following the pixels of a region in a counterclockwise direction along the border towards its centroid until all pixels are traced out. Then he applies a polyline description of the new one-dimensional signal.

It would also be fruitful to apply classical spatial coding techniques for the texture representation assuming that the region shapes are given.

### III.2.3. Contextual coding techniques.

Object oriented methods enable to include complex *a priori* information [71]. A new class of methods called model-based image coding has recently attracted attention, especially for facial expression, because of its data compression performances and its various applications. In this coding system, the analysis of input images and the synthesis of the output are performed assuming a prior information on the characteristics of the object to be coded. The inherent difficulty is related to the choice of a good model and efficient parameters for the application which is under consideration. Neural network classifiers are promising candidates for the preliminary analysis part (recognition and classification of the objects) [67].

## IV. CONCLUSION

In this paper, we have shown that the decorrelation methods for image processing are not confined to the classical DCT, RBC or subband decompositions. New methods try to extract relevant information on the image contents in order to improve the efficiency of the whole coding process. We should also keep in mind that the movement between successive images can also be used in order to further increase the compression efficiency (3-D subband coding, block matching, object matching,...).

*Manuscrit reçu le 10 février 1993.*

## REFERENCES

[1] AHMED (N.), RAO (K.). Orthogonal transforms for digital signal processing. *Springer-Verlag* (1975).  
 [2] ANG (P.), RUETZ (P.), AULD (D.). Video compression makes big gains. *IEEE Spectrum* (1991), **28**, n° 10, pp. 16-19.  
 [3] ANTONINI (M.), BARLAUD (M.), MATTHIEU (P.), DAUBECHIES (I.). Image coding wavelet transform. *IEEE Trans. on Image Processing* (1992), **1**, n° 2, pp. 205-220.  
 [4] ASH (R.). Information theory. *Dover* (1990).  
 [5] BIGGAR (M.), MORRIS (O.), CONSTANTINIDES (A.). Segmented-image coding : performance comparison with the discrete cosine transform. *IEE Proceedings F, Radar, and Signal Processing* (1988), **135**, n° 2, pp. 121-132.  
 [6] BOVIK (A.), CLARK (M.), GEISLER (W.). Multichannel texture analysis using localized spatial filters. *IEEE Trans. on Pattern Analysis and Machine Intelligence* (1990), **12**, n° 1, pp. 55-73.  
 [7] BURT (P.), ADELSON (E.). The Laplacian pyramid as a compact

image code. *IEEE Trans. on Communications* (1983), **1**, n° 4, pp. 532-540.  
 [8] CASELLES (V.), *et al.* A geometric model for active contours in image processing. *Céramade* (1992), *Cahier de mathématique de la décision*, N-9210.  
 [9] CHAMBOLLE (A.). Segmentation « optimale » de signaux. *Problèmes non linéaires appliqués : modélisations mathématiques pour le traitement d'images* (1992), Ecole CEA-EDF-INRIA.  
 [10] CHAUDHURI (B.), CHANDRASHEKHAR (S.). Neighboring direction runlength coding : an efficient contour coding scheme. *IEEE Trans. on Systems, Man, and Cybernetics* (1990), **20**, n° 4, pp. 916-921.  
 [11] CLARK (R.). Relation between the Karhunen Loève and cosine transform. *IEE Proceedings-F : Communications, radar and signal processing* (1981), **128**, n° 6, pp. 359-360.  
 [12] CROCHIERE (R.), RABINER (L.). Multirate digital signal processing. *Prentice-Hall* (1983).  
 [13] DAUBECHIES (I.). Orthonormal bases of compactly supported wavelets. *Comm. Pure and Appl. Math.* (1988), **41**, pp. 909-996.  
 [14] DAUBECHIES (I.). Ten lectures on Wavelets. *Siam* (1992).  
 [15] DELSARTE (P.), MACQ (B.), SLOCK (D.). Signal-adapted multiresolution transform for image coding. *IEEE Trans. on Information Theory* (1992), **38**, n° 2, pp. 897-904.  
 [16] DU BUF (J.). Abstract processes in texture discrimination. *Spatial Vision* (1992), **6**, n° 3, pp. 221-242.  
 [17] DUDA (R.), HART (P.). Pattern classification and scene analysis. *Wiley-Interscience* (1973).  
 [18] EDEN (M.), KOCHER (M.). On the performance of a contour coding algorithm in the context of image coding ; part I : Contour segment coding. *Signal Processing* (1985), **8**, n° 4, pp. 381-386.  
 [19] FARELLE (P.). Recursive block coding for image data compression. *Springer-Verlag* (1991).  
 [20] FEAUVEAU (J.). Analyse multirésolution par ondelettes non orthogonales et bancs de filtres numériques. *Ph.D. Thesis* (1990), Université Paris Sud.  
 [21] FREEMAN (H.). On the encoding of arbitrary geometric configurations. *IRE Trans. on Electron. Comput.* (1961), **10**, pp. 260-268.  
 [22] FROMENT (J.), *et al.* Compression d'images et ondelettes. *Problèmes non linéaires appliqués : ondelettes et paquets d'ondelettes* (1991), Ecole CEA-EDF-INRIA.  
 [23] GALLAGER (R. G.). Variations on a theme by Huffman. *IEEE Trans. on Information Theory* (1978), **24**, n° 6, pp. 668-674.  
 [24] GILGE (M.), ENGELHARDT (T.), MEHLAN (R.). Coding of arbitrarily shaped image segments based on a generalized orthogonal transform. *Signal Processing : Image Communications* (1989), **1**, n° 2, pp. 152-180.  
 [25] GUILLEMOT (C.), CETIN (E.), ANSARI (R.). M-channel nonrectangular wavelet representation for 2-D signal : Basis for quincunx sampled signals. *Int. Conf. on Acoustics, Speech and Signal Processing* (1991), pp. 2813-2816.  
 [26] HARALICK (R.), STERNBERG (S.), ZHUANG (X.). Image analysis using mathematical morphology. *IEEE Trans. on Pattern Analysis and Machine Intelligence* (1987), **9**, n° 4, pp. 532-550.  
 [27] HOLSCHNEIDER (M.), *et al.* Pointwise analysis of Riemann's non differentiable function. *Inventiones Math.*, Springer Verlag (1991).  
 [28] HUANG (T.). Coding of two-tone images. *IEEE Trans. on Communications* (1977), **25**, n° 11, pp. 1406-1424.  
 [29] HURT (N.). Phase retrieval and zero crossings. *Kluwer* (1989).  
 [30] HABIBI (A.). Hybrid coding of pictorial data. *IEEE Trans. on Communications* (1974), **22**, n° 5, pp. 614-624.  
 [31] JAIN (A.). Image data compression : a review. *Proc. of the IEEE* (1981), **69**, n° 3, pp. 349-389.  
 [32] JERRI (A.). The Shannon sampling theorem, its various extensions and applications : a tutorial review. *Proc. of the IEEE* (1977), **65**, n° 11, pp. 1565-1596.  
 [33] JOHNSTON (J.). A filter family designed for use in quadrature mirror filter banks. *Int. Conf. on Acoustics, Speech and Signal Processing* (1987), pp. 191-195.  
 [34] JURGEN (R.). Digital video. *IEEE Spectrum* (1992), **29**, n° 3, pp. 24-30.  
 [35] KARLSSON (G.), VETTERLI (M.). Theory of two-dimensional multirate filter banks. *IEEE Trans. on Acoustics, Speech, and Signal Processing* (1990), **8**, n° 6, pp. 925-937.  
 [36] KOEFLER (G.). Texture discrimination : a review of the theory and its algorithms. *Problèmes non linéaires appliqués : modélisations mathématiques pour le traitement d'images* (1992), Ecole CEA-EDF-INRIA.  
 [37] KOVACEVIC (J.), VETTERLI (M.). Nonseparable multidimensional perfect reconstruction filter banks and wavelet bases for  $R^n$ . *IEEE Trans. on Information Theory* (1992), **38**, n° 2, pp. 533-555.  
 [38] KUNT (M.). Traitement numérique du signal. *Dunod* (1981).  
 [39] KUNT (M.), BENARD (M.), LEONARDI (R.). Recent results in high-compression image coding. *IEEE Trans. on Circuits and Systems* (1987), **34**, n° 11, pp. 1306-1336.  
 [40] LEOU (F.-C.), CHEN (Y.-C.). A contour-based image coding technique with its texture information reconstructed by polyline representation. *Signal Processing* (1991), **5**, n° 1, pp. 81-89.

- [41] LINDE (Y.), BUZO (A.), GRAY (R.). An algorithm for vector quantizer design. *IEEE Trans. on Communications* (1980), **28**, n° 1, pp. 84-95.
- [42] LU (C.-C.), DUNHAM (J.). Highly efficient coding schemes for contour lines based on chain code representations. *IEEE Trans. on Communications* (1991), **59**, n° 10, pp. 1511-1514.
- [43] LU (Y.), JAIN (R.). Reasoning about edges in scale space. *IEEE Trans. on Pattern Analysis and Machine Intelligence* (1992), **14**, n° 4, pp. 450-468.
- [44] MACQ (B.). Perceptual transforms and universal entropy coding for an integrated approach to picture coding. *Ph.D. Thesis* (1989), Université Catholique de Louvain.
- [45] MAES (S.). Hints on the relationship between subband coding, wavelet and vaguelette decompositions and cosine sine altered transformations. *European Signal Processing Conference, Brussels, EURASIP* (1992), pp. 1041-1044.
- [46] MAES (S.). Adaptive optimal decompositions with hybrid DCT-subband coding. *Int. Conf. on Signal Processing Time-Frequency and Time-Scale Analysis* (1992), pp. 479-482.
- [47] MAES (S.), VAN DROOGENBROECK (M.). Edge characterization in hierarchical subband coding schemes. *Int. Conf. on Image Processing and its Applications* (1992), pp. 598-601.
- [48] MAES (S.). The H-point theorem applied in signal processing. *Int. Conf. on Signal Processing Time-Frequency and Time-Scale Analysis* (1992), pp. 467-470.
- [49] MALLAT (S.). A theory for multiresolution signal decomposition : the wavelet representation. *IEEE Trans. on Pattern Analysis and Machine Intelligence* (1989), **11**, n° 7, pp. 674-693.
- [50] MALLAT (S.), et al. Wavelet maxima representation. *Wavelets and Applications*, Masson/Springer-Verlag (1991).
- [51] MALLAT (S.), ZHANG (Z.). Adaptive time-frequency decomposition with matching pursuits. *Int. Conf. on Signal Processing Time-Frequency and Time-Scale Analysis* (1992), pp. 7-10.
- [52] MALVAR (H.), STAELIN (D.). The LOT : transform coding without blocking effects. *IEEE Trans. on Acoustics, Speech, and Signal Processing* (1989), **37**, n° 4, pp. 553-559.
- [53] MALVAR (H.). Lapped transforms for efficient transform/subband coding. *IEEE Trans. on Acoustics, Speech, and Signal Processing* (1990), **38**, n° 6, pp. 969-978.
- [54] MARAGOS (P.), SCHAFER (R.). Morphological skeleton representation and coding of binary images. *IEEE Trans. on Acoustics, Speech and Signal Processing* (1986), **34**, n° 5, pp. 1228-1244.
- [55] MARR (D.), HILDRETH (E.). Theory of edge detection. *Proc. of Royal Soc. of London* (1980), **207**, pp. 187-217.
- [56] MARTENS (J.), MAJOR (G.). The perceptual relevance of scale-space image coding. *Signal Processing* (1989), **17**, n° 4, pp. 353-364.
- [57] MEAD (C.). Analog VLSI and neural systems. *Addison Wesley* (1989).
- [58] MEAD (C.), ISMAIL (M.). Analog VLSI implementation of neural systems. *Kluwer Academia* (1989).
- [59] MEYER (F.), BEUCHER (S.). Morphological segmentation. *Journal of Visual Communication and Image Representation* (1990), **1**, n° 1, pp. 21-46.
- [60] MINAMI (T.), SHINOHARA (K.). Encoding of line drawings with a multiple grid chain code. *IEEE Trans. on Pattern Analysis and Machine Intelligence* (1986), **8**, n° 2, pp. 269-276.
- [61] MUSSMANN (G.), PIRSCH (P.), GRALLERT (H.-J.). Advances in picture coding. *Proc. of the IEEE* (1985), **73**, n° 4, pp. 523-548.
- [62] NASRABADI (N.), KING (R.). Image coding using vector quantization : a review. *IEEE Trans. on Communications* (1988), **36**, n° 8, pp. 957-971.
- [63] NETRAVALI (A.), HASKELL (B.). Digital pictures, representations and compression. *Plenum* (1988).
- [64] OLZAK (L.), THOMAS (J.). Seeing spatial patterns, chapter 7 of *Handbook of perception and human performance*, vol. 1. Sensory processes and perception. *John Wiley and Sons* (1986).
- [65] PAPOULIS (A.). Probability, random variable and stochastic processes. *McGraw-Hill* (1984).
- [66] PAVLIDIS (T.). Structural pattern recognition. *Springer-Verlag* (1977).
- [67] PERANTONIS (S.), LISBOA (P.). Translation, rotation and scale invariant pattern recognition of high order neural networks and moment classifiers. *IEEE Trans. on Neural Networks* (1992), **3**, n° 2, pp. 241-251.
- [68] PITAS (I.), ANASTASIOS (N.), VENETSANOPOULOS (A.). Morphological shape decomposition. *IEEE Trans. on Acoustics, Speech, and Signal Processing* (1990), **12**, n° 1, pp. 38-45.
- [69] RIOUL (O.), VETTERLI (M.). Wavelets and signal processing. *IEEE Signal Processing Magazine* (1991), **8**, n° 4, pp. 14-38.
- [70] RONSE (C.). On idempotence and related requirements in edge detection. *Philips Research Laboratory* (1990).
- [71] SCHALKOFF (R.). Digital image processing and computer vision. *Wiley* (1989).
- [72] SERRA (J.). Image analysis and mathematical morphology. *Academic Press* (1982).
- [73] SERRA (J.), VINCENT (L.). An overview of morphological filtering. *Circuits Systems and Signal Process* (1992), **11**, n° 1, pp. 47-108.
- [74] SHANNON (C.). A mathematical theory of communication. *Bell Syst. Tech. J.* (1974), **27**, pp. 379-423 and 623-656.
- [75] SIMONCELLI (E.), ADELSON (E.). Non-separable extensions of quadrature mirror filters to multiple dimensions. *Proc. of the IEEE* (1990), **78**, n° 4, pp. 652-664.
- [76] SORYANI (M.), CLARKE (R.). Segmented coding of digital image sequences. *IEE Proceedings I, Communications, Speech and Vision* (1992), **139**, n° 2, pp. 212-218.
- [77] VAIDYANATHAN (P.). Multirate digital filters, filter banks, poly-phase networks, and applications : a tutorial. *Proc. of the IEEE* (1990), **78**, n° 1, pp. 56-93.
- [78] VANDENDORPE (L.), MACQ (B.). Optimum quality and progressive resolution of video signals. *Ann. Télécommunic.* (1990), **45**, n° 9-10, pp. 487-502.
- [79] VANDENDORPE (L.). Optimized quantization for image subband coding. *Signal Processing* (1991), **4**, n° 1, pp. 65-79.
- [80] VANDENDORPE (L.). Hierarchical coding of digital moving picture. *Ph.D. Thesis* (1991), Université Catholique de Louvain.
- [81] VAN DROOGENBROECK (M.), VANDENDORPE (L.). Integration of the selective deconvolution algorithm into a multiresolution scheme. *Image Processing : Algorithms and Techniques III* (1992), **1657**, pp. 225-235, San Jose, SPIE.
- [82] VENKATESH (S.), OWENS (R.). Implementation detail of a feature detection algorithm. *Technical report 89/12* (1989), Univ. Western Australia, Dpt. Comp. Sc.
- [83] VETTERLI (M.), LE GALL (D.). Perfect reconstruction FIR filter banks : some properties and factorizations. *IEEE Trans. on Acoustics, Speech, and Signal Processing* (1989), **37**, n° 7, pp. 1057-1071.
- [84] VISCITO (E.), ALLEBACH (J.P.). The analysis and design of multidimensional FIR perfect reconstruction filter banks for arbitrary sampling lattices. *IEEE Trans. on Circuits and Systems* (1991), **38**, n° 1, pp. 29-41.
- [85] WICKERHAUSER (M.). Picture compression by best-basis subband coding, Preprint (Yale University, 1990).
- [86] WICKERHAUSER (M.). Talk presented at the symposium on wavelets and signal processing. *CNRS-LMA, Marseille* (5-9 nov 1990).

---

## BIOGRAPHY

Marc VAN DROOGENBROECK was born in Brussels, Belgium, on February 1967. He received the M.S. degree in electrical engineering from the Université Catholique de Louvain in 1990. Since 1991 he has been appointed by the IRSIA (Institut pour l'Encouragement de la Recherche Scientifique dans l'Industrie et l'Agriculture, Belgium). He is currently working towards a Ph.D. thesis on picture coding by means of an object representation both at the University of Louvain and at the Center for Mathematical Morphology of the School of Mines of Paris. His research interests include image processing and mathematical morphology.

Stéphane MAES was born in Brussels, Belgium, on February 1966. He received, with the Highest Honors, the M.S. degree in Electrical Engineering and the B.S degree in Physics from the Université Catholique de Louvain (Belgium) in 1990. In 1991, he received, with the Highest Honors, the degrees of M.S. in Physics from the Université Catholique de Louvain (Belgium) and M.S.S. from the International Space University (USA). He has been funded by the CERN (European Center for Nuclear Research), in 1988 and by the ESA (European Space Agency) in 1991. Since 1990, he is employed as Research Assistant of the Belgian National Fund for Scientific Research (FNRS). He works towards a PhD. in signal processing with the University Catholique de Louvain (Belgium) and Rutgers University (USA).

Serge COMES was born in Brussels, Belgium, on May 1966. He received the M.S. degree in Electrical Engineering from the Université Catholique de Louvain in 1989. Since 1990, he is employed as Research Assistant of the Belgian National Fund for Scientific Research (FNRS). He is currently working towards a PhD thesis on perceptual digital picture coding. His research interests include image processing, non-separable multirate filter banks, picture quality and human visual system modeling.