ELSEVIER

Signal Processing: *Image Communication* 6 (1994) 83–99

SIGNAL PROCESSING: IMAGE COMMUNICATION

3-D spatiotemporal subband decompositions for hierarchical compatible video coding by mathematical morphology

Soo-Chang Pei[a,*], Fei-Chin Chen[b]

[a] *Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan, ROC*
[b] *Department of Electrical Engineering, National Taiwan Institute of Technology, Taipei, Taiwan, ROC*

Received 4 October 1992

Abstract

An efficient hierarchical 3-D spatiotemporal subband decomposition scheme using mathematical morphology is proposed in this paper. It can support a hierarchy of video standards including both progressively scanned and interlaced signals. The compatibility among different video and distribution standards is increasingly important because of a multitude of video and multimedia services within broadband communication networks. This hierarchical coding scheme also can provide compatibility by splitting the source signal into several hierarchical layers available for different consumer receiver sets. Recently, a sophisticated spatiotemporal linear filtering bank is proposed for this 3-D subband decomposition. However, morphological filtering can be used here to reduce its complexity greatly for its simplicity and efficiency in parallel/pipeline hardware implementation. Some experiments, using a 2-D frequency sweep as a test pattern, are given to show the effectiveness of this approach.

Key words: Subband decomposition; Video coding; Morphological filtering

1. Introduction

Videotransmission via broadband integrated service digital network (B-ISDN) will become a major service of the future video communication networks. Emerging broadband communication networks [2] facilitate the existence of various video services, ranging from high definition TV (HDTV), enhanced TV to videotelephone (VT) and videoconferencing. Compatibility between these different services is clearly required [8]. In fact, these services correspond to a possible hierarchy of video scanning formats and associated bit-rates as shown in Table 1 [8,3,4]. It consists of 5 different video formats, including both progressive-scan and interlaced signals with various spatial formats. By splitting the source signal into several hierarchical video standard layers, a receiver selects and decodes only those layers suitable for its display monitor. A subband coding approach is to decompose a signal into a low-resolution version and several bands containing high-frequency detail information. By combining the low-resolution version with some of the high-frequency bands, a higher-resolution version of the video signal is obtained.

* Corresponding author.

SSDI 0923-5965(93) E0037-5

Table 1
Possible hierarchy of video services [3, 4, 8]

Video service	Scanning parameters	Aspect ratio	Sampling parameters		Bit-rate (Mbit/s)
			Luminance Y	Chrominance UV	
HDP	1250/50/1:1	16:9	144 MHz/1920	72 MHz/960	2304
HDI	1250/50/2:1	16:9	72 MHz/1920	36 MHz/960	1152
EDP	625/50/1:1	16:9	36 MHz/960	18 MHz/480	576
EDI	625/50/2:1	16:9	18 MHz/960	9 MHz/480	288
VT	312.5/25/1:1	16:9	4.5 MHz/480	2.25 MHz/240	72

HDP: progressive high definition
HDI: interlaced high definition
EDP: progressive enhanced definition
EDI: interlaced enhanced definition
VT: videotelephone format

Table 1 shows the possible hierarchy of video standards. The spatiotemporal spectrum of each standard video signal is shown in Fig. 1(a). Clearly, a 3-D spatiotemporal filtering is necessary to obtain one of the HDI (Fig. 1(b)), EDP (Fig. 1(c)), EDI (Fig. 1(d)), VT (Fig. 1(e)) signals from the HDP signal (Fig. 1(f)). This requires a nonrectangular 3-D spatiotemporal subband decomposition. Recently, a sophisticated spatiotemporal linear filtering bank is proposed for this 3-D subband decomposition [3, 1, 4]. It is designed and implemented in a more complicated polyphase form. This polyphase filter's bank is composed of simple modulators, polyphase filters, adders, delays, upsamplers and downsamplers [3, 4]. The objective of this paper is to explore whether an alternative effective method, called morphological filtering bank scheme, can do these jobs. Morphological filters are relatively simple and very effective for subband decomposition and directional filtering. Some experiments, using a 2-D frequency sweep as a test pattern, are given to show the effectiveness of this approach.

In this paper, Section 2 reviews mathematical morphological operations on binary and gray-level images. Section 3 illustrates two spatiotemporal and one spatial subband decompositions which can be used for hierarchical video schemes. The first spatiotemporal subband decomposition, called reduced temporal hierarchy [3, 4], supports only the HDP, EDP and VT progressive-scan services, whereas the second spatiotemporal subband decomposition, called full temporal hierarchy [3, 4] and the spatial hierarchy, are capable of handling both interlaced and progressive-scan video signals. Section 4 gives some subband decomposition experimental results using morphological filtering, and conclusions are made in Section 5.

2. Mathematical morphology

Mathematical morphology is an approach to image processing based on set-theoretic concepts of shape [10]. An image can be represented by a set of pixels, the morphological operations deals with two images: the original data to be analyzed and a structuring element, which is analogous to the kernel of a convolution operation. Each structuring element has a shape which can be thought of as a parameter to the operation. The four basic operations in mathematical morphology are dilation, erosion, opening and closing. This science was introduced by Matheron; Serra presents these concepts in detail [10]. A tutorial is found in Haralick et al. [6].

2.1. Binary image morphology

Assume X (original image) and B (structuring element) are subsets of a 2-D Euclidean space, let

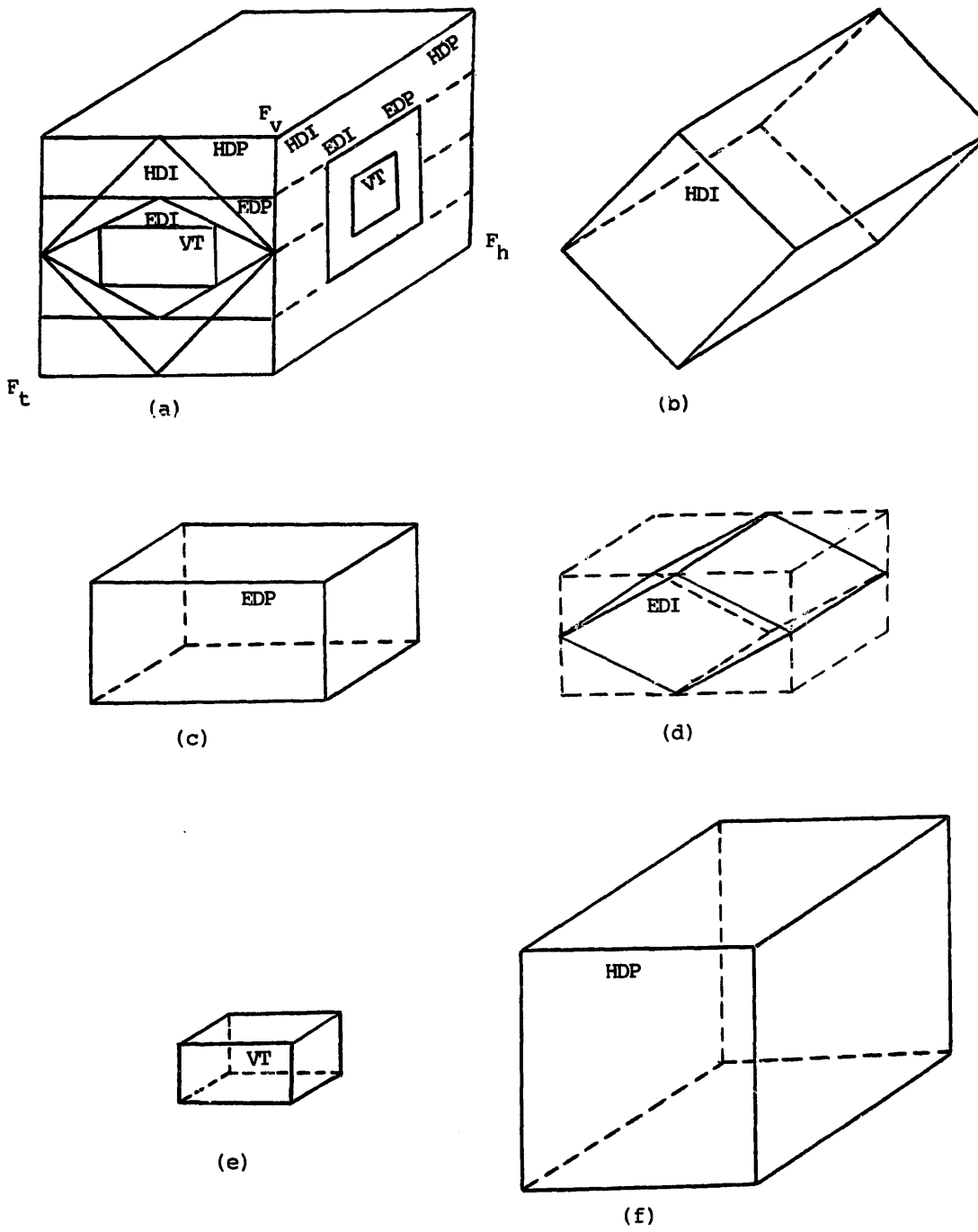


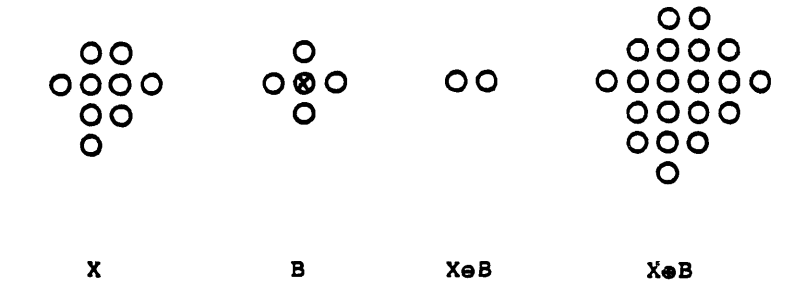
Fig. 1. (a) Resolution levels in the spatiotemporal Fourier space, (b) HDI, (c) EDP, (d) EDI, (e) VT, (f) HDP.

$(X)_b$ denote the translation of X by the vector b,

$$(X)_b = \{y | y = x + b, x \in X\}, \tag{1}$$

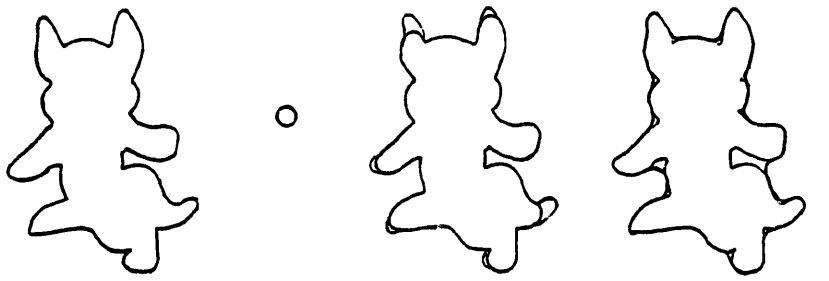
then the two most fundamental morphological operations, dilation and erosion, can be defined as follows:

$$X \oplus B = \bigcup_{b \in B} (X)_b = \{y | y = x + b, x \in X, b \in B\}, \tag{2}$$



⊗ denotes the origin of the structuring element B.

(a)



(b)

Fig. 2. Example of binary image morphology. (a) Erosion and dilation of X by B. (b) Opening and closing of X by B. (The dark solid curves correspond to the boundaries of the transformed sets.)

$$X \ominus B = \bigcap_{b \in B} (X)_{-b} = \{y \mid b \in B, \text{ implies } (y + b) \in X\}. \tag{3}$$

This means that a dilation of X by B, $X \oplus B$, can be described as the union of translations of X by all points b contained in the structuring element B. For erosion, $X \ominus B$, the result is an intersection of translation by all points $-b$, where b is contained in B. Fig. 2(a) shows the example of erosion and dilation of X by B. Dilation and erosion cause the expanding or shrinking of areas when the structuring element has a disklike shape. Although dilation and erosion are complementary morphological operations, they are noninvertible. This means that the original image in general cannot be recovered performing dilation to its eroded version or vice versa.

An opening is defined as an erosion followed by a dilation by the same structuring element and is shown as

$$X \circ B = (X \ominus B) \oplus B. \tag{4}$$

On the other hand, we define a closing as a dilation followed by an erosion as

$$X \bullet B = (X \oplus B) \ominus B. \tag{5}$$

Openings on an image with a structuring element B can be pictured by moving B inside all the shapes in an image and marking only those places where

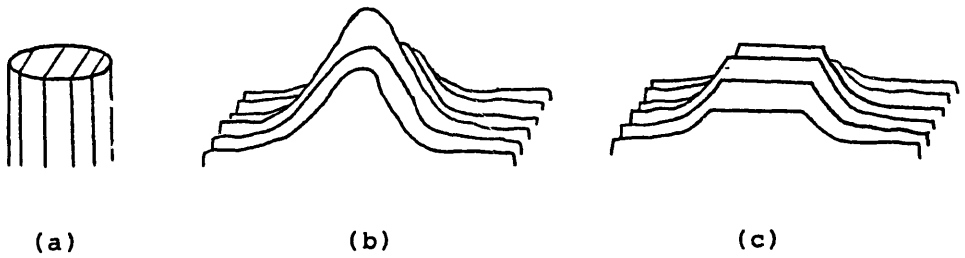


Fig. 3. Example of gray-level image morphology. (a) Disk-shaped structuring element. (b) Gray-level image landscape. (c) Opening by a disk-shaped structuring element.

B fits. Similarly, closings on an image with a structuring element B can be pictured by moving B around the outside of an image with the result that the concave corners are rounded and the convex corners remain square. The example of opening and closing of X by B is shown in Fig. 2(b).

2.2. Gray-scale morphology

Morphological concepts can be extended to gray-scale images. In the gray-scale morphology [6], gray-scale images are visualized as 3-D landscapes. The structuring elements of the gray-scale morphology are 3-D shapes like spheres or cylinders. Performing morphological operations on a gray-scale image by a spherical structuring element is equivalent to sliding a sphere across the gray-level surface.

Let the image $X(x)$ be represented as a function of coordinates x. The analytical definitions of the gray-level morphology operations are as follows:

$$X \oplus B = D(x) = \max_{b \in B}[X(x - b) + B(b)], \tag{6}$$

$$X \ominus B = E(x) = \min_{b \in B}[X(x + b) - B(b)], \tag{7}$$

where the $B(b)$'s are weights that are a function of b.

Notice that all of the relationships for binary image morphology are preserved here in a form in which intersection is replaced by min, and union is replaced by max. The opening and closing operations for gray-level images are defined as in binary case, i.e.

$$X \circ B = (X \ominus B) \oplus B, \tag{8}$$

$$X \bullet B = (X \oplus B) \ominus B. \tag{9}$$

In gray-level morphology the disk-shaped structuring element shown in Fig. 3(a) is 3-D. The result of a 3-D opening is to move the top areas under the top surface of the landscape in Fig. 3(b) defined by the image, and keep those areas where the disk-shaped structuring element fits. Strongly peaked areas with widths less than the diameter of the disk-shaped structuring element will flatten out as illustrated in Fig. 3(c).

3. The hierarchical spatiotemporal and spatial subband decomposition using mathematical morphology

The video hierarchy shown in Table 1 can be separated into a hierarchy in the vertical–temporal frequency plane and the vertical–horizontal frequency plane, as shown in Fig. 1(a). These two hierarchies are referred to as the temporal and the spatial hierarchy. Denoting the spatial hierarchy spectrum of a video service by the subscript s and argument w, the ordering of video services on the basis of their frequency content is described as

$$\begin{aligned} \mathrm{VT}_s(w) \subset \mathrm{EDI}_s(w) &= \mathrm{EDP}_s(w) \subset \mathrm{HDI}_s(w) \\ &= \mathrm{HDP}_s(w). \end{aligned} \tag{10}$$

This relation ensures that a subband decomposition of the spatial hierarchy exists and is simple using a morphological subband decomposition [9].

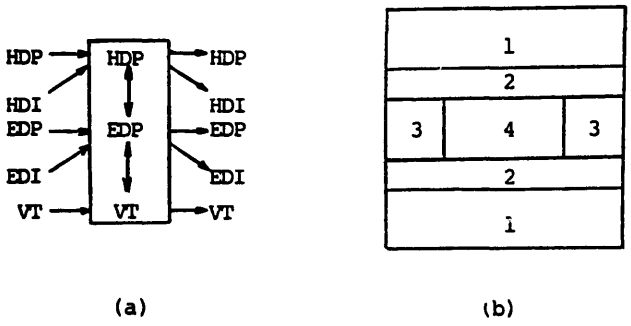


Fig. 4. Reduced temporal hierarchy strategy [3,4].

On the other hand, the decomposition of the temporal hierarchy is more complicated because the HDI and EDP vertical-temporal spectra overlap each other partially as shown in Fig. 1(a). None of the two spectra can be extracted from the other. If we exclude either the HDI or EDP video service, we can get the following incomplete temporal hierarchies as

$$VT_t(w) \subset EDI_t(w) \subset EDP_t(w) \subset HDP_t(w) \quad (11)$$

and

$$VT_t(w) \subset EDI_t(w) \subset HDI_t(w) \subset HDP_t(w), \quad (12)$$

where the subscript t refers to temporal hierarchy.

3.1. Reduced temporal hierarchy

Only with the reduced temporal hierarchy for progressive-scan video services, compatibility between the VT, EDP and HDP signals can be fully

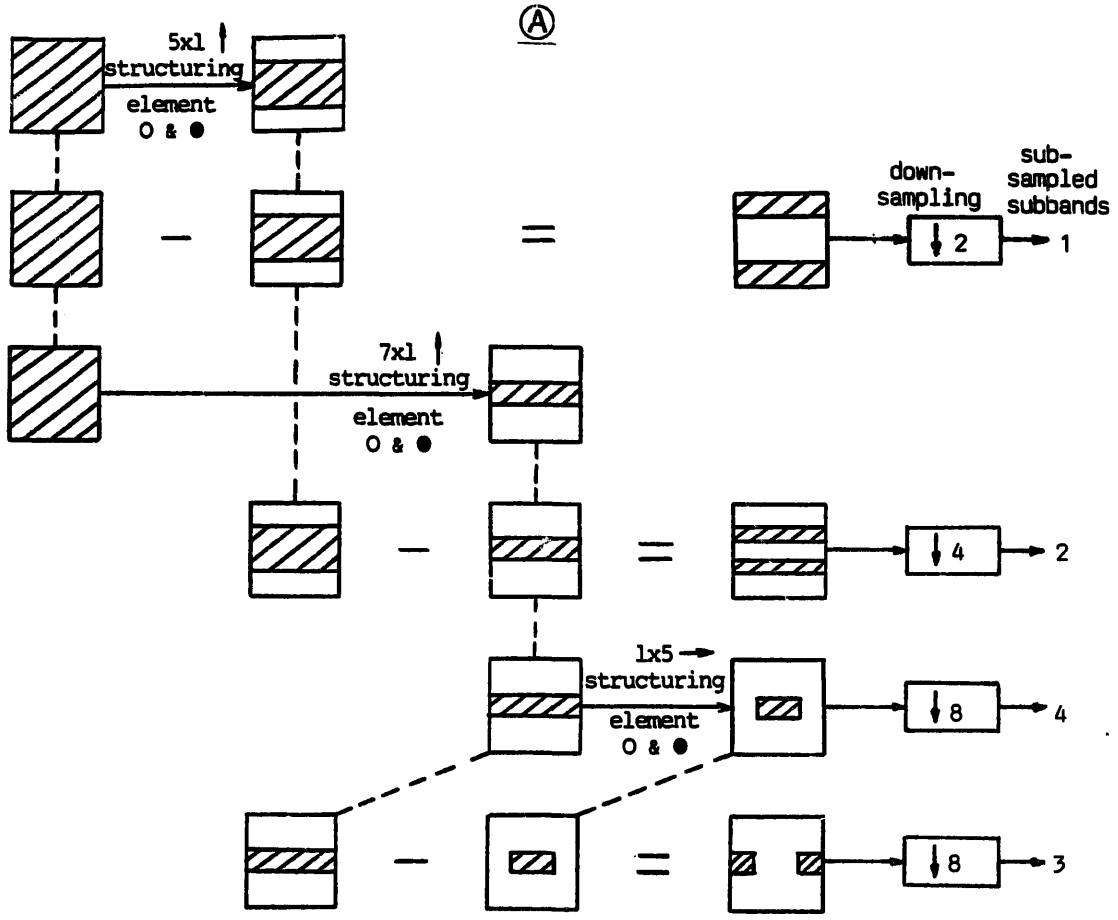


Fig. 5. (a) System diagram of subband decomposition for the reduced temporal hierarchy using morphological filter banks. ○ and ● denote opening and closing, respectively. The passband/stopband are located in the shaded/blank areas. The same data are shared and indicated by the dashed lines.

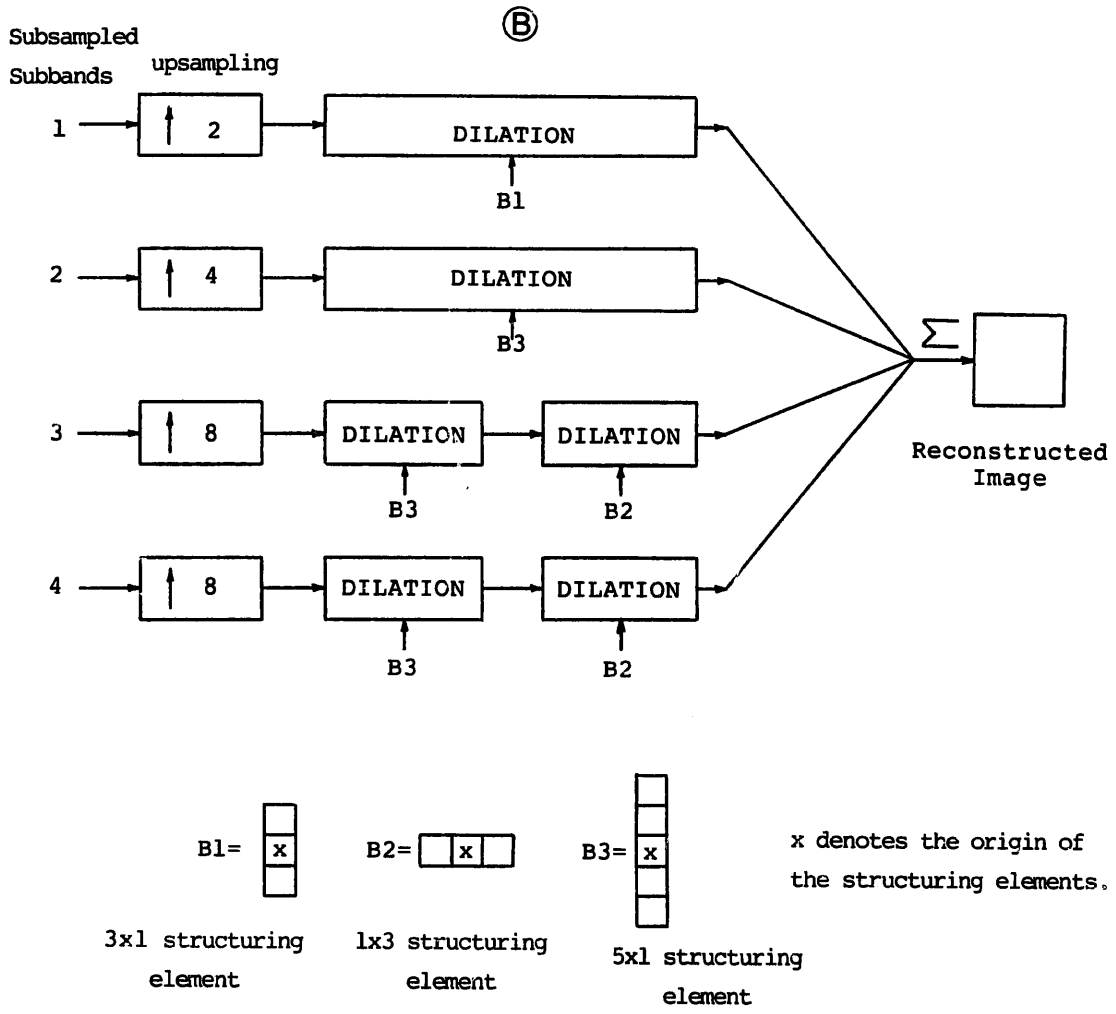


Fig. 5. (b) System diagram of reconstruction for the reduced temporal hierarchy using morphological filter banks.

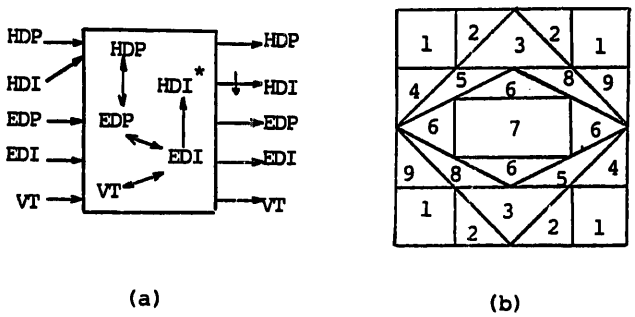


Fig. 6. Full temporal hierarchy strategy [3,4].

warranted. For interlaced signals, first an upconversion in Fig. 4(a) is done to become progressive-scan services before transmission. Then at the receiver side a downconversion will get back to interlaced signals again. This is illustrated in Fig. 4(a) [3,4], while the partitioning of the vertical–temporal spectrum of the HDP signal is shown in Fig. 4(b) [3,4]. It is clear that the advantage of this 'reduced temporal hierarchy' is its regularity and complexity for subband decomposition.

The morphological lowpass filters are used here instead of the linear band splitting filters. Sequential alternating application of the morphological operations of opening and closing by means of the same structuring elements removes details of the image that are small relative to this structuring element. We call these alternating sequential filters morphological lowpass filters. The 1-D highpass

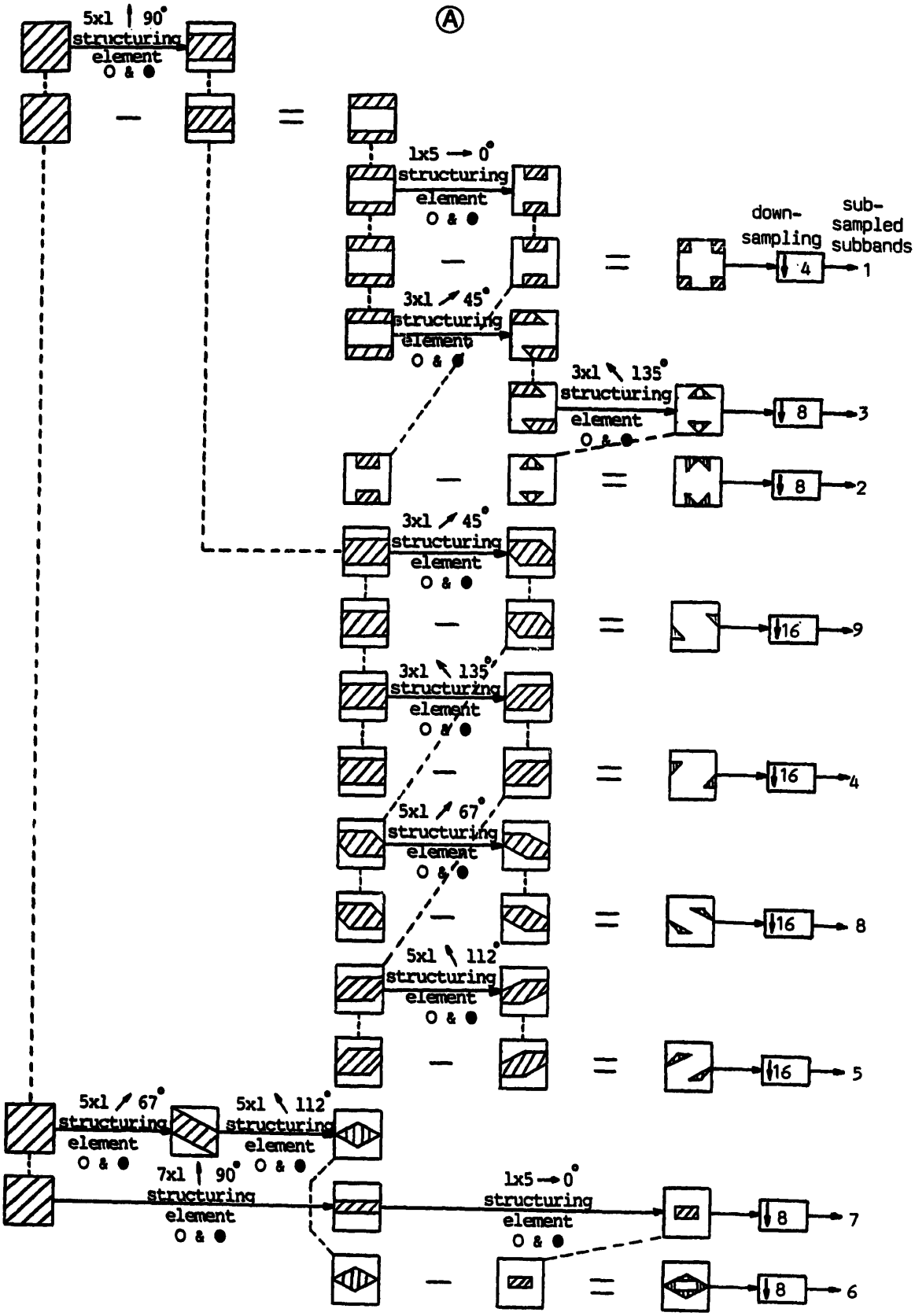


Fig. 7. (a) System diagram of subband decomposition for the full temporal hierarchy using morphological filter banks.

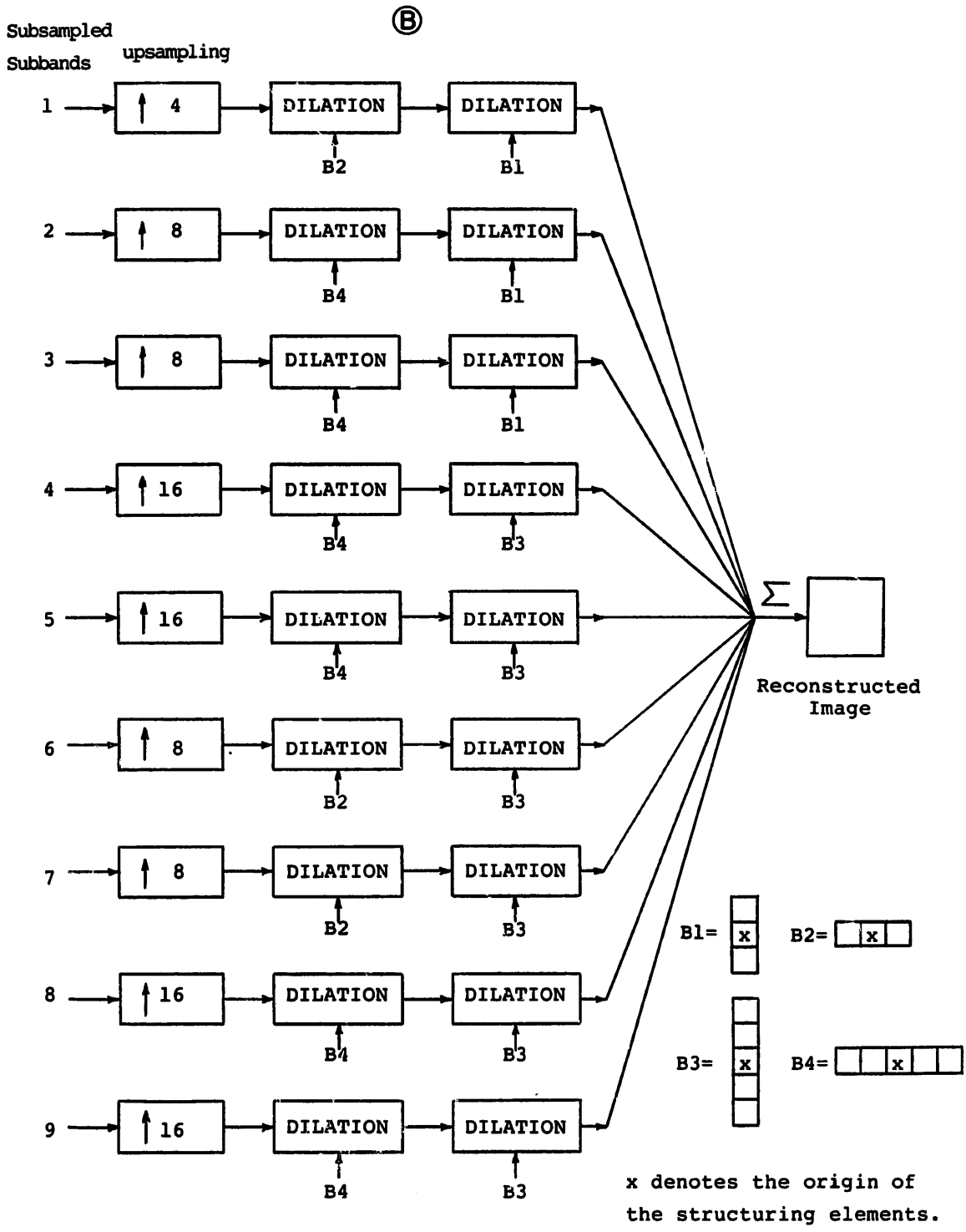


Fig. 7. (b) System diagram of reconstruction for the full temporal hierarchy using morphological filter banks.

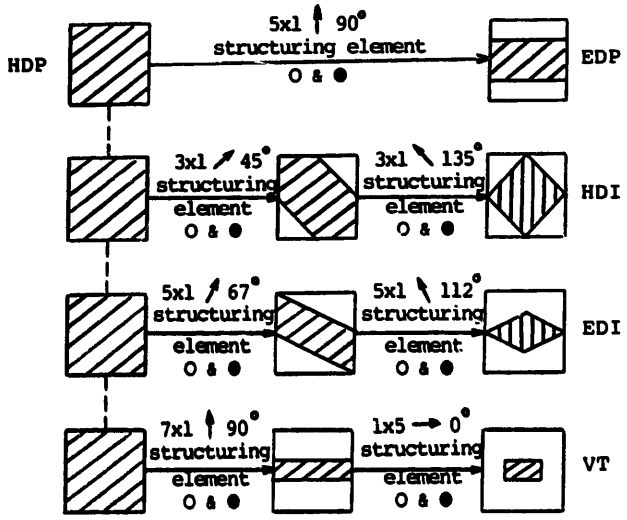


Fig. 8. System diagram of subband decomposition for the conversion from HDP signal to EDP, HDI, EDI and VT signal, respectively, in the vertical-temporal frequency plane using morphological filter banks.

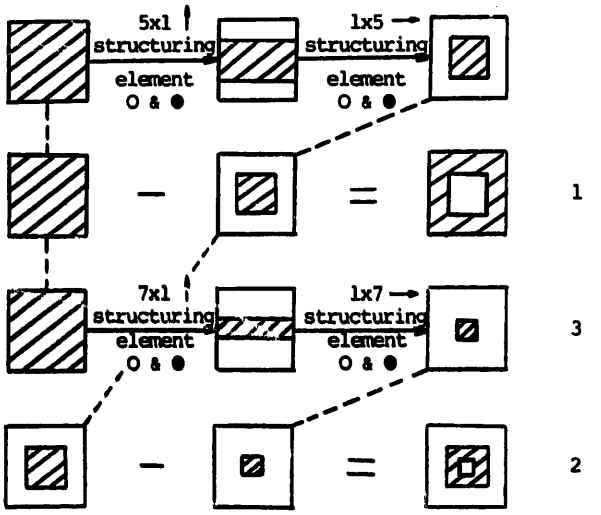


Fig. 9. System diagram of subband decomposition for the spatial hierarchy using morphological filter banks.

filters can be constructed with the original signal and the complement of the lowpass filtering $[-H_0(x)]$ as follows:

1-D lowpass filter H_0: closing [opening (x)], (13)

1-D highpass filter H_1: $X - H_0(x)$. (14)

By using morphological filter banks, the system diagram of the reduced temporal hierarchy strategy

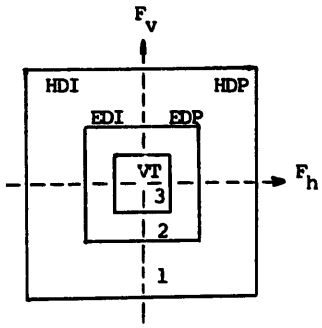


Fig. 10. Spatial hierarchy strategy.

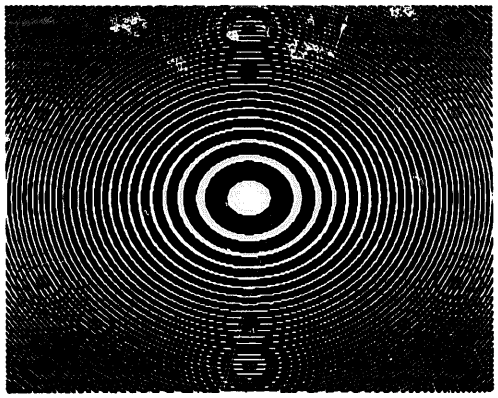


Fig. 11. Test pattern, 2-D frequency sweep.

is shown in Fig. 5(a). The vertical–temporal plane is decomposed into four subbands. The layer associated with the VT signal includes subband number 4; the layer associated with the EDP signals includes subbands 2–4, while the layer associated with the HDP signals includes the whole subbands 1–4 (Fig. 4(b)) totally. The reconstruction process is shown in Fig. 5(b) and the interpolation structuring elements are also shown for reference.

3.2. *Full temporal hierarchy*

A more complete temporal hierarchy called the 'full temporal hierarchy' performs subband decomposition of the vertical–temporal spectra such that all video services can be reconstructed directly as shown in Fig. 6(a) [3,4]. By HDI* in Fig. 6(a), we denote the HDI signal represented on a spatiotemporal orthogonal sampling lattice. The

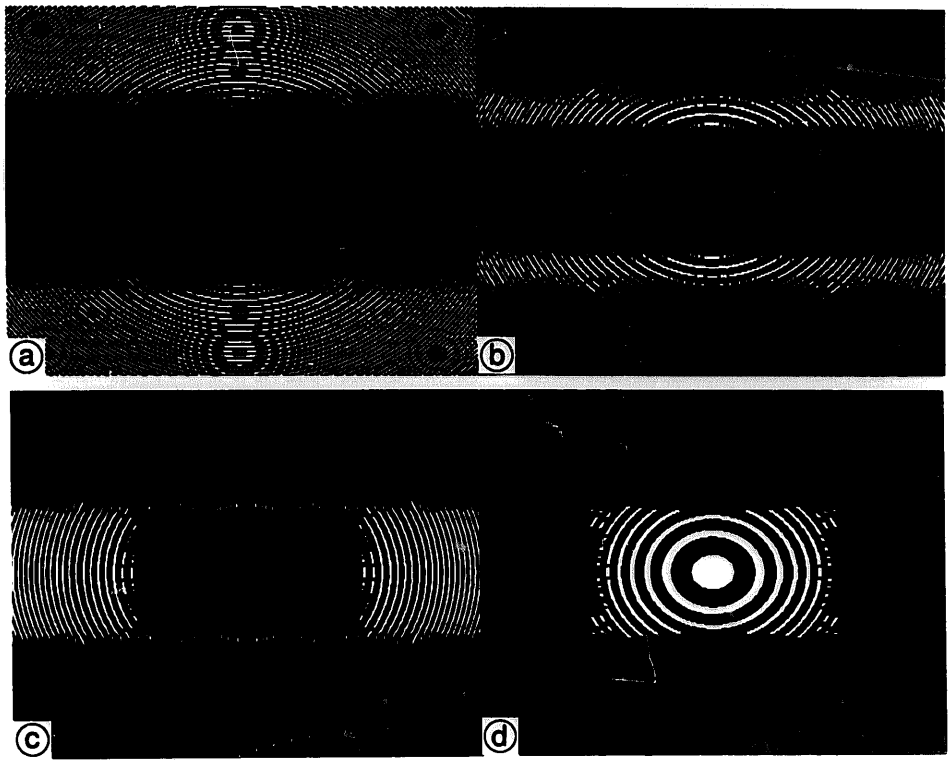


Fig. 12. Resultant subbands of the reduced temporal hierarchy: (a) Subband 1, (b) Subband 2, (c) Subband 3 and (d) Subband 4.

resulting subband decomposition of the vertical–temporal frequency plane is shown in Fig. 6(b) [3,4]. For example, the EDI spectrum is composed of subbands 6 and 7. As compared to the reduced temporal hierarchy, the full temporal hierarchy strategy is significantly more complex due to additional vertical–temporal decomposition stages.

Recently, a sophisticated spatiotemporal linear filtering bank scheme [3,1,4] is proposed to implement this full temporal hierarchy decomposition. Several nonseparable diamond shaped and parallelogram 2-D filtering are necessary to extract the wedge-shaped parts of the signal spectrum; also it requires two π-modulators and several complicated non-orthogonal quincunx and directional decimations to be involved in the above scheme (see Fig. 7 in [3]). In our approach, very simple morphological filters can replace the above 2-D linear filters and the nonorthogonal decimations. By means of mathematical morphology filters, the complexity of filter design and implementation is greatly reduced because only summation/subtraction operations and maximum/minimum decisions are needed. Also, the processing speed is very high because the basic morphology operations can be implemented in the pipeline processing form of parallelism [5].

The full temporal hierarchy strategy using morphological filter banks is shown in Fig. 7(a). The 2-D directional filter banks for wedge-shaped subband decomposition are designed by a separable product of 1-D directional morphological filters with six angle orientations 0°, 45°, 67°, 90°, 112° and 135° and their complements. These morphological filters are much easier to design and implement than the above 2-D diamond-shaped and parallelogram linear filters. For the conversion from HDP signal to EDP, HDI and VT signal, respectively, the simple morphological filter banks in Fig. 8 can achieve this downward compatibility very easily. The reconstruction process for the full temporal hierarchy is shown in Fig. 7(b) and the interpolation structuring elements are also shown for reference.

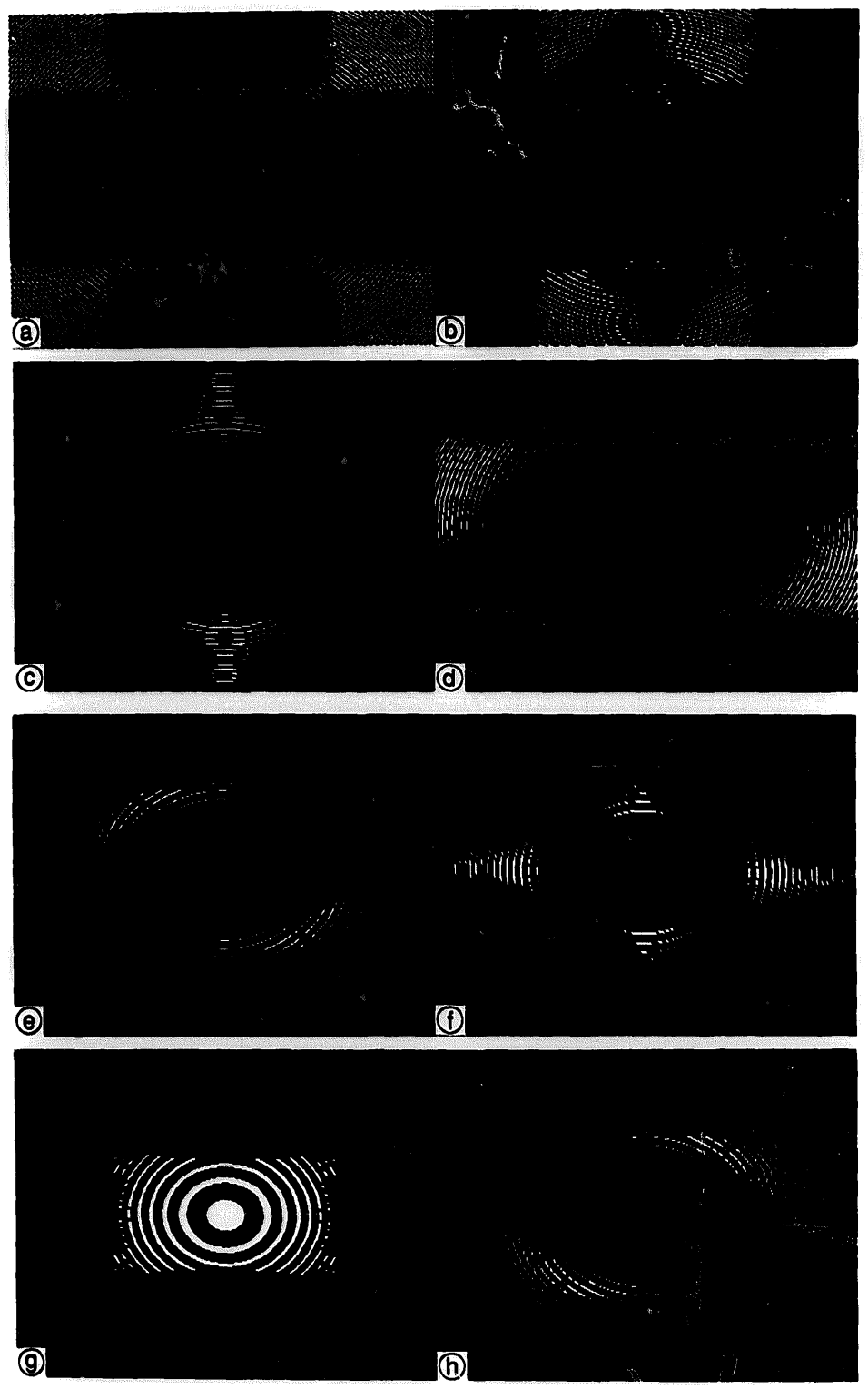


Fig. 13. Resultant subbands of the full temporal hierarchy: (a) Subband 1, (b) Subband 2, (c) Subband 3, (d) Subband 4, (e) Subband 5, (f) Subband 6, (g) Subband 7, (h) Subband 8.

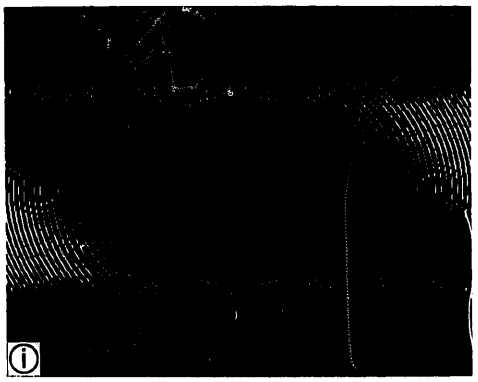
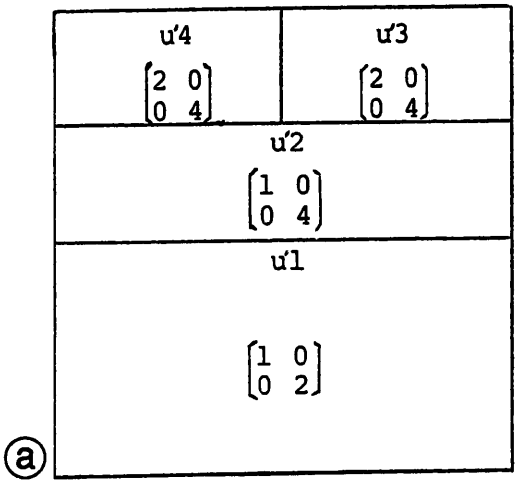
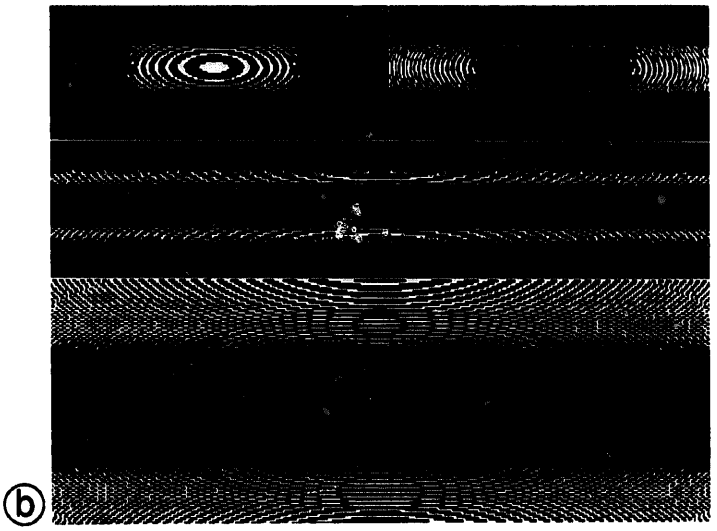


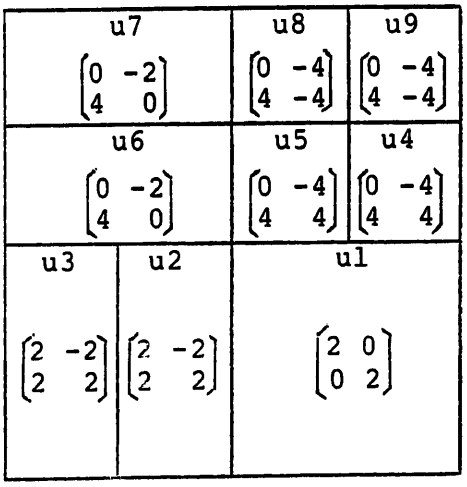
Fig. 13. (i) Subband 9.

3.3. Spatial hierarchy

The subband decomposition in the vertical–horizontal frequency (F_v–F_h) plane is called spatial hierarchy. As shown in Fig. 1(a), the spatial hierarchy is well defined since the ordering of video services based on their frequency spectra is straightforward in Eq. (10). The resulting subband decomposition of the vertical–horizontal frequency plane is shown in Fig. 10. The layer associated with the VT signal consists of subband 3, and the layer associated with the EDP or EDI signal includes subbands 2 and 3; while the layer associated with







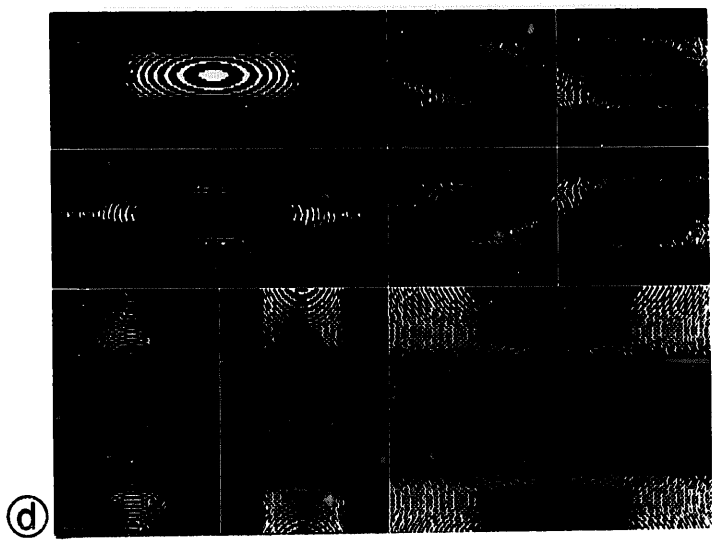


Fig. 14. Resultant subbands and their corresponding subsampling matrices of (a, b) the reduced temporal hierarchy, (c, d) the full temporal hierarchy.

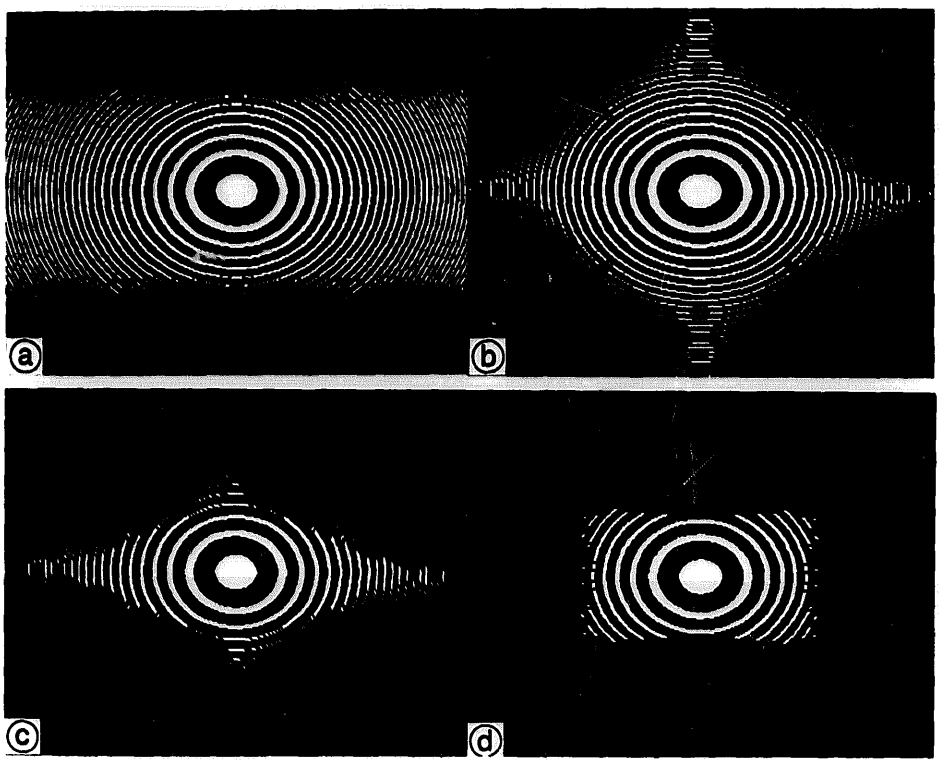


Fig. 15. Result of subband decomposition for the conversion from HDP signal to (a) EDP, (b) HDI, (c) EDI and (d) VT signals.

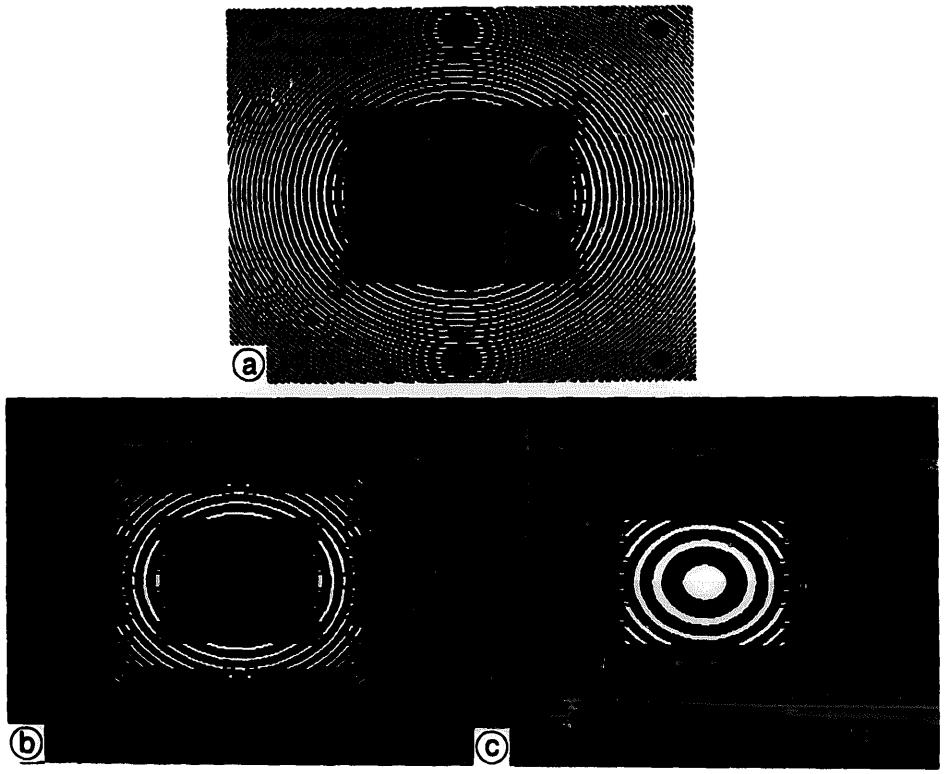


Fig. 16. Resultant subbands of spatial hierarchy: (a) Subband 1, (b) Subband 2 and (c) Subband 3.

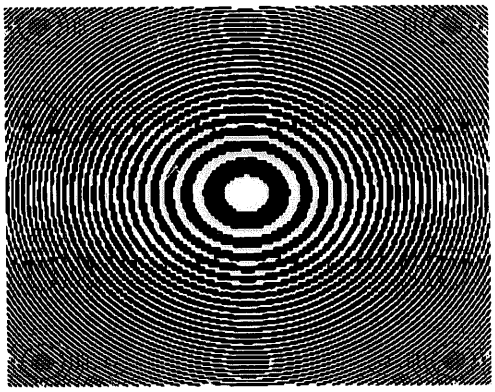


Fig. 17. Reconstructed image of the test pattern for the reduced temporal hierarchy.

the HDP or HDI signal contains subbands 1–3. Fig. 9 shows the system diagram of spatial hierarchy using morphological filter banks.

4. Experimental results

The test pattern, 2-D frequency sweep, used in our experiments is shown in Fig. 11. We present the experimental results obtained from the two vertical–temporal subband decompositions and the vertical–horizontal subband decomposition in this section. In a practical coding environment these vertical–temporal decompositions have to operate in the vertical and temporal dimensions and have to perform the required filtering at an image sequence. However, for simplicity we have applied the decomposition to the 2-D frequency sweep (Fig. 11) in our experiments. Since this image contains all the 2-D frequencies, we consider it as being the vertical–temporal or vertical–horizontal plane of an HDP signal; the morphological operations are done in the pixel domain while the results are interpreted as in the frequency domain.

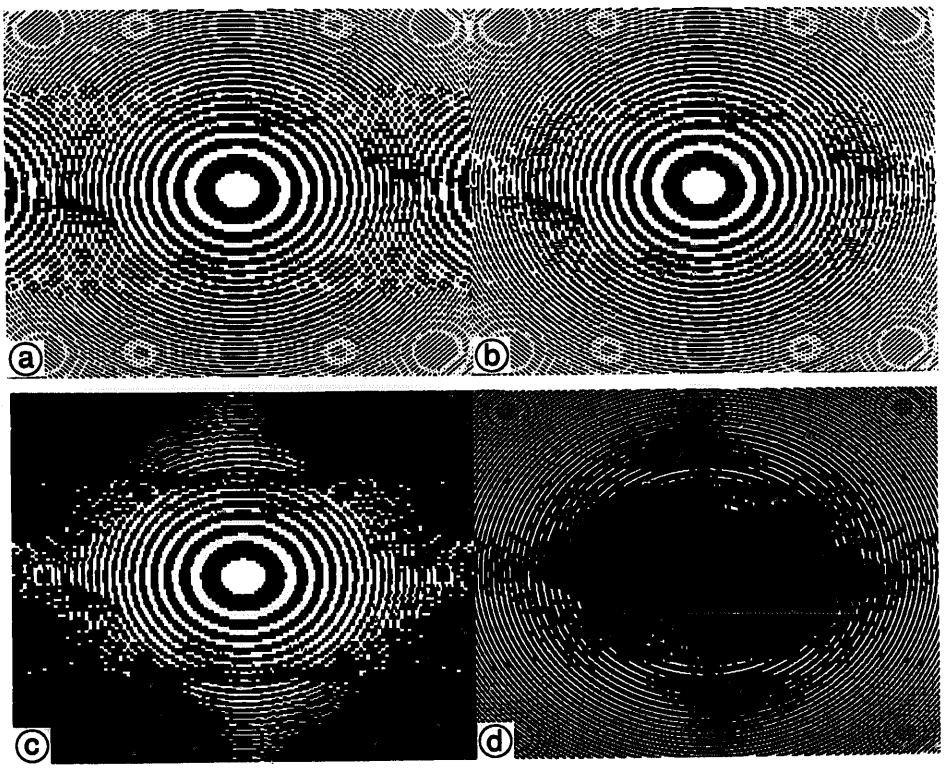


Fig. 18. (a, b) Reconstructed image of the test pattern for the full temporal hierarchy. (c) Reconstructed image and (d) the absolute error image of the HDP signal using subbands 3, 5, 6, 7, 8 only.

The first experiment, which is implemented by the procedures illustrated in Fig. 5(a), investigates the frequency contributions of the different subbands of the reduced temporal hierarchy subband decomposition. The experimental result shown in Fig. 12 matches with the partitioning of the vertical–temporal spectrum of the HDP signal (Fig. 4(b)). The resultant subbands and their corresponding subsampling matrices of the reduced temporal hierarchy are shown in Figs. 14(a) and 14(b). Only one decimation (u'_i, $i = 1, 2, 3, 4$) is needed to perform at each final subband output stage, the previous several decimations needed in linear filtering bank in [3] are combined here into one decimation, for example: $u'_4 = u_0 \times u_{10} \times u_{20}$.

In the second experiment, we use the procedures illustrated in Fig. 7(a) to investigate the frequency contributions of the different subbands of the full temporal hierarchy subband decomposition. The experimental result shown in Fig. 13 demonstrates the successful partitioning of the vertical–temporal spectrum of the HDP signal for the full temporal hierarchy (Fig. 6(b)). Figs. 14(c) and 14(d) show the resultant subbands and their (corresponding) subsampling matrices of the full temporal hierarchy. The final directional decimations u_i, $i = 1, 2, \ldots, 9$, are obtained in the same way as above.

The procedures shown in Fig. 8 implement the third experiment. The experimental result shown in Fig. 15 illustrates the downward compatibility from HDP signal to EDP, HDI, EDI and VT signal, in the vertical–temporal frequency plane.

The fourth experiment is implemented by the procedures shown in Fig. 9. The experimental result illustrated in Fig. 16 shows the successful partitioning of the vertical–horizontal spectrum of the HDP signal hierarchy (Fig. 10).

In the last two experiments, we use the procedures illustrated in Figs. 5(b) and 7(b) to perform the reconstruction of the test pattern from the subsampled subbands for the reduced temporal hierarchy and the full temporal hierarchy, respectively, using morphological filter banks. The experimental results given in Figs. 17 and 18 show the successful reconstruction of the reduced temporal hierarchy and the full temporal hierarchy, respectively. The reconstruction of the reduced temporal hierarchy is nearly exact as shown in Fig. 17. The reconstructed image of the test pattern for the full temporal hierarchy has some aliasing patterns in left and right boundary areas corresponding to bands 4 and 9, as shown in Fig. 18(a), since the original test pattern also has some aliasing pattern in these areas. If the subsampling and upsampling rates of bands 4 and 9 are reduced from 16 to 4, the aliasing in these areas is reduced and the reconstruction is much improved as shown in Fig. 18(b). The reconstructed image and the error image of the HDP signal are shown in Figs. 18(c) and 18(d), respectively, after it has been reconstructed using subbands 3, 5, 6, 7, 8 only. In Fig. 18(d), the black and white areas indicate small and large errors, respectively. As expected, the frequencies are well reconstructed in the desired diamond-shaped area shown in Fig. 18(c) with some similar distortion as in [3, 4]. Experimental results show that this distortion is less severe for natural sequences since the frequencies involved in the distorted areas are less important in natural sequences [3, 4].

5. Conclusions

This paper has described two vertical–temporal and one vertical–horizontal subband decompositions, which can be used for hierarchical video coding applications, using mathematical morphology. It supports the spatial and temporal hierarchy of 5 different video standards including both progressively scanned and interlaced signals. Experimental results confirm the selectivity in frequency of the different subbands in the spatial and temporal hierarchy.

This 3-D subband decompositions method using mathematical morphology have great potential for high definition television (HDTV) compatible coding and multimedium video compression, etc. Meanwhile, the morphological operations involved are very simple, fast and well suited for parallel/pipeline hardware (VLSI) implementation [5]. Some experiments are given to show the effectiveness of this approach.

References

[1] R.H. Bamberger and M.J.T. Smith, "A filter bank for the directional decomposition of images: Theory and design", *IEEE Trans. Signal Process.*, Vol. 40, No. 4, April 1992, pp. 882–893.

[2] J.A. Bellisio and K.H. Tzou, "HDTV and the emerging ISDN network", *SPIE Visual Communications and Image Process Conf.*, Boston, 1988, pp. 772–786.

[3] F. Bosveld, R.L. Lagendijk and J. Biemond, "3D subband decompositions for hierarchical video coding", *SPIE Vol. 1605 Visual Communications and Image Processing '91: Visual Communication*, 1991, pp. 769–780.

[4] F. Bosveld, R.L. Lagendijk and J. Biemond, "Compatible spatio-temporal subband encoding of HDTV", *Signal Processing*, Vol. 28, No. 3, September 1992, pp. 271–289.

[5] R.M. Haralick and X. Zhuang, "Pipeline architectures for morphological image analysis", *Machine Vision and Applications*. Vol. 1, 1988, pp. 23–40.

[6] R.M. Haralick, S.R. Sternberg and X. Zhuang, "Image analysis using mathematical morphology", *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. PAMI-9, No. 4, 1987, pp. 532–550.

[7] J.D. Johnston, "A filter family designed for use in quadrature mirror filter banks", *Internat. Conf. Acoust. Speech Signal Process. '80*, 1980, pp. 291–294.

[8] M. Pecot, P.J. Tourtier and Y. Thomas, "Compatible coding of television images. Part 2. Compatible system", *Signal Processing: Image Communication*, Vol. 2, No. 3, October 1990, pp. 259–268.

[9] S.C. Pei and F.C. Chen, "Subband decomposition of monochrome and color images by mathematical morphology", *Optical Engineering*, Vol. 30, No. 7, July 1991, pp. 921–933.

[10] J. Serra, "The convex set model", in: *Image Analysis and Mathematical Morphology*, Academic Press, London, 1982, pp. 93–124.

[11] M. Vetterli, J. Kovačević and D.J. LeGall, "Perfect reconstruction filters banks for HDTV representation and coding", *Signal Processing: Image Communication*, Vol. 2, No. 3, October 1990, pp. 349–363.