論文 97-22-8-12

Computationally Efficient Wavelet Transform for Coding of Arbitrarily-Shaped Image Segments

Eui-Sung Kang*, Jae-Yong Lee*, Jong-Han Kim**, and
Sung-Jea Ko* *Regular Members*

ABSTRACT

Wavelet transform is not applicable to arbitrarily-shaped region (or object) in images, due to the nature of its global decomposition. In this paper, the arbitrarily-shaped wavelet transform(ASWT) is proposed in order to solve this problem and its properties are investigated. Computational complexity of the ASWT is also examined and it is shown that the ASWT requires significantly fewer computations than conventional wavelet transform, since the ASWT processes only the object region in the original image. Experimental results show that any arbitrarily-shaped image segment can be decomposed using the ASWT and perfectly reconstructed using the inverse ASWT.

Ⅰ. Introduction

Various transform-based coding techniques have been proposed for the reduction of great amount of redundancy in images [1]. The discrete cosine transform (DCT) has been widely used due to its decorrelative property and its simplicity for VLSI implementation. On the other hand, wavelet coding schemes have several important properties for image coding, namely: *i*) they offer a flexible multiresolution image representation with localization in both time and frequency, *ii*) they avoid the "*blocking effect*" associated with block transform coding due to its global decomposition characteristics, and *iii*) human visual characteristics can be incorporated for efficient compression [2],[3].

However, the aforementioned transforms are difficult to apply to object-oriented coding methods based on the segmentation of the image into foreground (arbitrarily-shaped object) and background, because these transforms are performed only on the rectangular region. In order to solve this problem, a variety of the arbitrarily-shaped transforms have been proposed [4]-[7]. The "shape-adaptive DCT algorithm" has been proposed to apply the DCT to the arbitrarily-shaped region [4]. This algorithm aligns all row and column pixels next to each other in order to transform the arbitrarily-shaped region. This technique is computationally efficient, but it may cause severe distortions between pixels within the segment (*i.e.*, break-ups or holes within a segment are eliminated). The "shape-adaptive wavelet transform (SAWT)" is also proposed to apply the wavelet transform to the arbitrarily-shaped region [5], [6]. This algorithm transforms the arbitrarily-shaped region by the method similar to the shape-adaptive DCT, and thus it has the same problem as the shape-adaptive DCT has. Another method, named "object-based wavelet transform", converts the arbitrarily-shaped object to a rectangular region to apply the

---

*고려대학교 대학원 전자공학과
**삼성전자 정보통신연구소
論文番號:97181-0529
接受日字:1997年 5月 29日

ordinary wavelet transform [7]. This method is not computationally efficient since the decomposition procedure requires extrapolation and low pass filtering.

In this paper, we present the ASWT which can be applied to the arbitrarily-shaped region (or object) in the image. The decomposition by the proposed ASWT is performed by applying a filter bank to the extended object region instead of the rectanglular image. With the proposed method, it is possible to obtain the maximum decorrelation of the image segment by preserving the original shape of the object in each subband. The inverse ASWT can perfectly reconstruct the original shape of the segmented region. The ASWT requires significantly fewer computations than conventional wavelet transform, since the ASWT processes only the object region in the original image. The proposed algorithm can be utilized for the object-oriented coding schemes which segment image object from image backround.

The organization of this paper is as follows. The ASWT is proposed in Section II, and its propoerties are investigated in Section III. Experimental results are given in Section IV. Finally, conclusion remarks are given in Section V.

II. Arbitrarily-Shaped Wavelet Transform

In this section, a new implementation method of the ASWT which is applicable to the arbitrarily-shaped object in images is proposed. Before we introduce the proposed method, we first review the shape-adaptive wavelet transform (SAWT).

A. SAWT

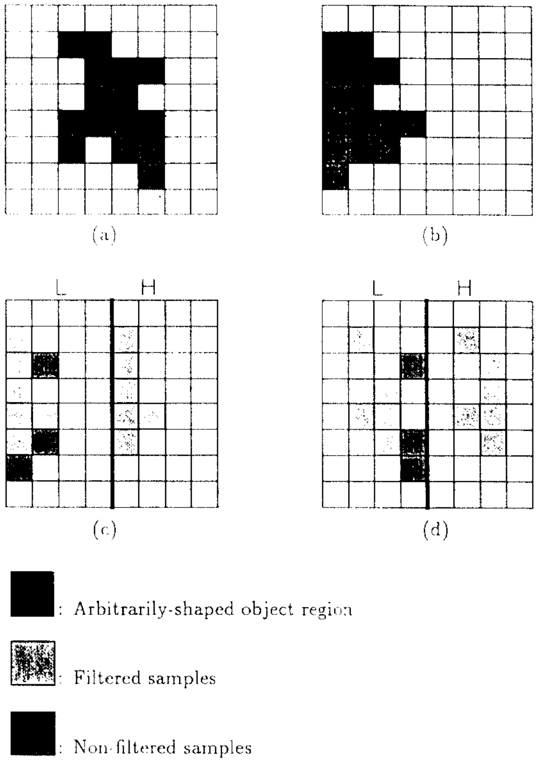
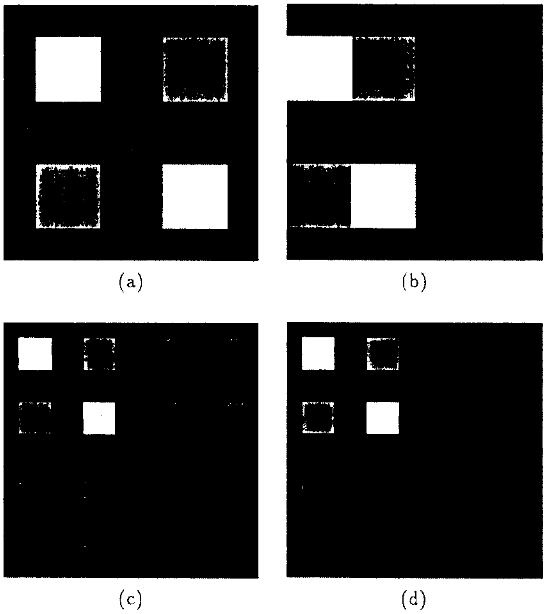


Fig. 1 Processing steps for the SAWT, (a) the original object, (b) the alignment of the object region in the horizontal direction, (c) the horizontal decomposition result of the aligned object, and (d) the placement of the filtered samples into the subbands such that each subband preserve original shape of the object region.

The procedure of the SAWT is shown in Fig. 1. Consider an arbitrarily-shaped object in Fig. 1(a). First, all row pixels in the arbitrarily-shaped object are aligned next to each other as shown in Fig. 1(b). The aligned pixels are then decomposed by a filter bank in the horizontal direction. Since a filter bank constraints the input samples to have the size equal to a multiple of the number of bands, the non-filtered samples can be generated and are placed into the low frequency band, as shown in Fig. 1(c). After the decomposition, the filtered samples are placed into the subbands such that each subband preserves the original shape of the region as shown in Fig. 1(d). The vertical decomposition is performed in the same method but the all column pixels are aligned next to each other. This method can give rise to some problems in the following case. For example, consider the separate objects, which have the different gray level (white and gray), as shown in Fig. 2(a). Since all row pixels are first aligned next to each other for the hori-


(a) (b)

(c) (d)

Fig. 2 Example of the SAWT, (a) the original object, (b) the aligned object in the horizontal direction, (c) the result of the horizontal followed by vertical decomposition using SAWT, and (d) the decomposition result using the proposed ASWT.

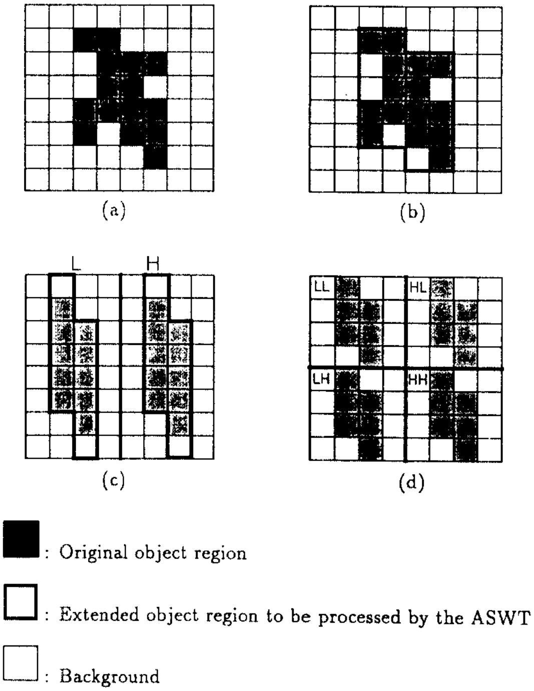
zontal decomposition, the gap between the objects is eliminated as shown in Fig. 2(b). As a result, the unexpected edges are generated within the object in the low frequency band. Fig. 2(c) shows the first-stage decomposition (horizontal followed by vertical decomposition) result. Unnecessary edge information is also produced in the high frequency bands. Therefore, this method is inefficient for further compression of each subband.

This problem can be solved by using the proposed ASWT to be described in the next subsection. With the proposed ASWT, we can not only reduce the edge information in the high frequency bands, but also preserve the original shape of the object in the low frequency band(see Fig. 2(d)).

B. The proposed ASWT

In general, wavelet transform decomposes the input data into N bands by using a FIR filter bank. For the filter bank, the length of input data must be equal to a multiple of the number of bands. For example, for $N=2$, the length of input data is a multiple of two. Under this constraint, the decomposition using the filter bank is performed in the unit of two pixels.

In Fig. 3, the proposed ASWT algorithm is illustrated. The operation goes as follows. It is performed in a separable way as in conventional decomposition schemes. As a first step, before applying the FIR filter in the horizontal direction, each of two pixels in the image is examined to determine whether these pixels belong to an image object or not. If at least one of these two pixels is a part of the object, both of these pixels are included in the extended object region. The original and extended object regions are shown in



■ : Original object region

□ : Extended object region to be processed by the ASWT

□ : Background

Fig. 3 Processing steps for the ASWT, (a) the original object, (b) the extended object, (c) the result of the horizontal followed by vertical decomposition, and (d) the vertical decomposition result.

Fig. 3(a) and (b), respectively. It is seen that the extended object region has pixels which do not belong to the original object. These pixels are replaced by an adjacent pixel value (see the next section for more details). After the extended object region is determined, the wavelet filter bank composed of the low and high pass filters is applied to that region.

Boundary problems can occur at the right end of the object region, when the length of the filter is larger than 2. In order to solve the boundary problem, either symmetric or periodic extension can be used [8],[9]. In this paper, the periodic extension is used to handle the boundaries. With the periodic extension, the left boundary samples within the extended object region are appended to the right end of sequence as follows;

$$\underbrace{s_0\ s_1\ s_2\ \cdots\ s_{n-2}\ s_{n-1}\ s_n}_{\text{a line in the extended object}}\ \underbrace{s_0\ s_1\ s_2\ \cdots}_{\text{appended}} \qquad (1)$$

where s_i represents pixels (samples) in the image. The filtered samples of the extended object region are placed in a subband such that the subband preserves the original shape. Fig. 3(c) shows the horizontal decomposition result. In the same manner, the vertical decomposition is performed on the horizontal decomposed image, and the decomposition result with four subbands(LL, LH, HL, and HH) is shown in Fig. 3(d). Note that the ASWT preserves the original shape of the image object in each subband. This is important for further efficient coding such as the embedded zerotree wavelet [10] or the multiresolution motion estimation/compensation, since it can avoid some distortion between pixels within the image and keep the self-similarity of the object in each subband.

Table I. Variance of subbands according to the selection of the extra pixel value.

	sample variance of each subband			
extra pixel value	LL band	LH band	HL band	HH band
0	1008.3	1231.9	1017.7	214.3
255	176.4	2743.4	1033.8	217.7
adjacent pixel value	0	0	0	0

Ⅲ. Properties of the ASWT

In this section, we investigate some properties of the ASWT such as effect of the extra pixels. It is shown that the ASWT can reduce the computational complexity of wavelet transform significantly.

A. Effect of the extra pixels

The proposed ASWT algorithm processes the extended object region which may be somewhat larger than the original object, which is necessary for preserving the original shape of the object. Extra pixels which are not a part of the original object are included in the extended object region. The value of these extra pixels is important, since it affects the coding performance in each subband. For example, consider an image segment of arbitrary shape as shown in Fig. 3. Before applying the filter to the object region, the extra pixels have to be replaced by a certain value. We can investigate how the extra pixel value affects statistical properties of each subband by computing the variances of the transformed coefficients in each subband. In this simulation, assume that all the pixels in the object region have the same gray level(= 127) and background pixel values are zero. The extended object regions were generated by appending different values of the extra pixels and processed by the ASWT. The sample variances of the transformed coefficients in each subband(LL, LH, HL, and HH) according to the selection of the extra pixel value are compared in Table I. It is seen that the usage of the adjacent pixel value for the extra pixel value gives the variance equal to zero in each subband. This result indicates that efficient compression can be yielded by the usage of the adjacent pixel value, since the edge informations in the high frequency bands can be minimized.

B. Computational complexity

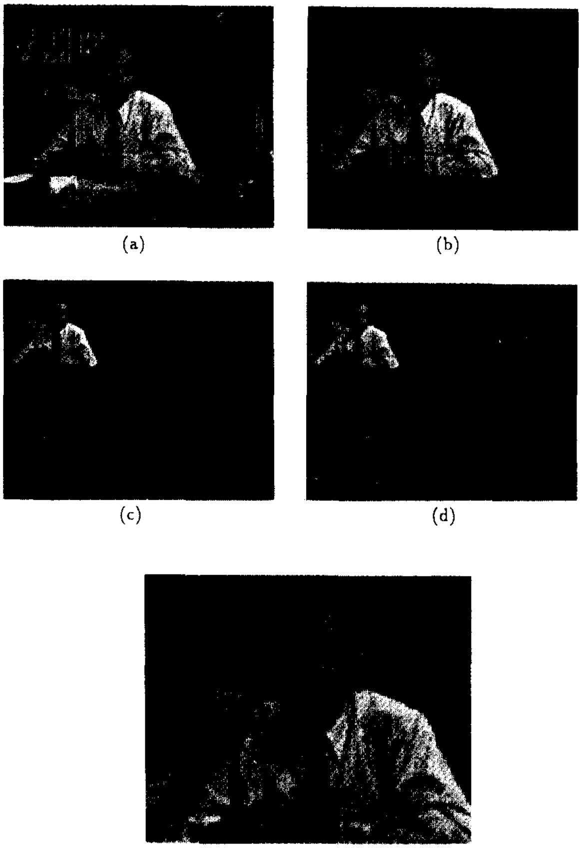
To compare the complexity of ASWT with that of conventional wavelet transform, the number of operations required for these two transforms is evaluated. Assume that the size of the image is $N \times N$ and the length of the wavelet filter is equal to P. If M and A, respectively, denote multiplication and addition operations, the computation required to obtain the filter output at each pixel is $P \cdot M + (P-1) \cdot A$. To decompose an $N \times N$ image using the filter bank composed of the low and high pass filters in the horizontal direction, $[P \cdot M + (P-1) \cdot A] \cdot N^2$ operations are required. The successive application of the filter bank in vertical direction to the horizontally decomposed image produces four subbands, which requires also $[P \cdot M + (P-1) \cdot A] \cdot N^2$ operations. Therefore, the total operations required for the first-stage decomposition are $2 \cdot [P \cdot M + (P-1) \cdot A] \cdot N^2$. The proposed ASWT can significantly reduce the number of multiplications and additions. However, the decomposition by the ASWT requires additional $\frac{3}{2} \cdot N^2$ comparison operations to identify the object region in the image (N^2 comparisons for the horizontal decomposition and $\frac{1}{2} \cdot N^2$ comparisons for the vertical decomposition). Assuming that the object region occupies $1/k$ of the image, the number of multiplications and additions is reduced to the $1/k$ of that for the conventional wavelet transform. That is, the decomposition can be performed by $2 \cdot P \cdot N^2 \cdot \frac{1}{k}$ multiplications, $2 \cdot (P-1) \cdot N^2 \cdot \frac{1}{k}$ additions, and $\frac{3}{2} \cdot N^2$ comparisons. Table II summarizes the comparison of the computational complexity.

Table Ⅱ. Comparison of the computational complexity between conventional wavelet transform and the proposed ASWT.

	Multiplications	Additions	Comparisons
Conventional WT	$2 \cdot P \cdot N^2$	$2 \cdot (P-1) \cdot N^2$	–
ASWT	$2 \cdot P \cdot N^2 \cdot \frac{1}{k}$	$2 \cdot (P-1) \cdot N^2 \cdot \frac{1}{k}$	$\frac{3}{2} \cdot N^2$

Ⅳ. Experimental Results

For the experiment, a Daubechies 4-tap FIR filter, which has orthogonal basis, is used as the wavelet filter. To apply the proposed ASWT to the real image, the segmented image is needed. For the experiments, the


(a) (b)

(c) (d)

(c)

Fig. 4 Salesman, (a) the original image, (b) the segmented image, (c) the decomposed image using the proposed ASWT, (d) the decomposed image using the SAWT, and (e) the reconstructed image using the inverse ASWT.

"Salesman" image shown in Fig. 4(a), whose format is QCIF with 256 gray levels, is manually segmented into the object and background as shown in Fig. 4(b). The "object" consists of the person's head and the upper part of body, and the "background" is filled with a constant value of zero to generate a perfect and well defined segment border. Fig. 4(c) shows the decomposed image obtained by the proposed ASWT. For comparison, the decomposed image using SAWT is shown in Fig. 4(d). It is seen that the ASWT preserves the original shape of the image object while producing the small amount of high frequency informations. In addition, the comparison with the result of the ASWT shows that the proposed ASWT produced smaller high frequency information than the SAWT. The reconstructed image by the inverse ASWT is shown in Fig. 4(e). The reconstructed image using SAWT is omitted since it is identical to that in Fig. 4 (e). The MSE between the original segmented and reconstructed image is almost zero(1.945×10^{-10}). This result shows that the inverse ASWT can also perfectly reconstruct the original image.

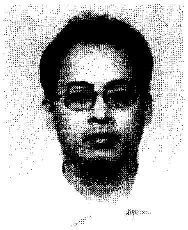
V. Conclusion

In this paper, the ASWT was proposed to transform the arbitrarily-shaped image segments and its properties were investigated. It was shown that any arbitrarily-shaped region in the original image can be decomposed using the ASWT and perfectly reconstructed using the inverse ASWT. The proposed ASWT can not only reduce the edge informations in the high frequency bands, but also preserve the original shape of the object in the low frequency band. It was also shown that the proposed ASWT requires significantly fewer computations than conventional wavelet transform. With the proposed ASWT, the number of multiplications and additions is reduced to $1/k$ of that for the wavelet transform, when the object regions to be processed occupy the $1/k$ of the image. The ASWT can be a computationally efficient alternative to the conventional wavelet transform for the low bitrate coding such as object-oriented coding.

References

1. M. Rabbani and P. Jones, "Digital Image Compression Techniques," *SPIE Tutorial Texts in Optical Engineering*, vol. TT7, 1991.
2. S. Mallt, "A Theory for Multiresolution Signal Decomposition : The Wavelet Representation," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 11, no. 7, pp. 674-693, 1989.
3. S. Mallt, "Multifrequency Channel Decomposition of Images And Wavelet Models," *IEEE Trans. Acoust, Speech, Signal Processing*, vol. 37, no. 12, pp. 2091-2110, Dec. 1989.
4. T. Sikora and B. Makai, "Shape-Adaptive DCT for Generic Coding of Video," *IEEE Trans. Circuits and Systems for Video Tech.*, vol. 83, no. 2, pp. 272-287, Feb. 1995.
5. O. Egger, P. Fleury, and T. Ebrahimi, "Shape Adaptive Wavelet Transform for Zerotree Coding," *Proceedings of the European Workshop on Image Analysis and Coding*, Rennes, France, vol. 1, pp. 201-208, Feb. 1996.
6. O. Egger, T. Ebrahimi, and M. Kunt, "Arbitrarily-Shaped Wavelet Packets for Zerotree Coding," *Proc. ICASSP*, vol. IV, pp. 2335-2338, May 1996.
7. T. Aono, N. Ito, H. Katata, and H. Kusao, "Object Based Wavelet Transform Tool for Video," *ISO/IEC JTC1/SC29/WG11, MPEG-4 Proposal*, Nov. 1995.
8. G. Karlsson and M. Vetterli, "Extension of Finite Length Signals for Subband Coding," *IEEE Trans. Signal Processing*, vol. 17, pp. 161-168, Jun. 1989.
9. R. H. Bamberger, S. L. Eddins, and V. Nuri, "Generalized Symmetric Extension for Size-Limited Multirate Filter Banks," *IEEE Trans. Image Processing*, vol. 3, no. 1, pp. 82-86, Mar. 1993.
10. J. M. Shapiro, "Embedded Image Coding Using Zerotees of Wavelet Coefficients," *IEEE Trans.*

Signal Processing, vol. 41, no. 12, pp. 3445-3462, Dec. 1993.



강 의 성(Eui-Sung Kang) 정회원

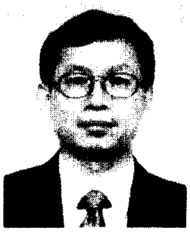
1991년 2월 : 고려대학교 공과대학 전자공학과 졸업(학사)
1995년 2월 : 고려대학교 대학원 전자공학과 졸업(공학석사)
1995년 3월~현재 : 고려대학교 대학원 전자공학과 박사과정
※관심분야 : 신호처리, 영상 개선, 영상 압축, 웨이블렛

이 재 용(Jae-Yong Lee) 정회원

1993년 2월 : 고려대학교 공과대학 전자공학과 졸업(학사)
1995년 2월 : 고려대학교 대학원 전자공학과 졸업(공학석사)
1995년 3월~현재 : 고려대학교 대학원 전자공학과 박사과정
※관심분야 : 비디오코딩 및 전송, 멀티미디어통신

김 종 한(Jong-Han Kim) 정회원

1995년 2월 : 고려대학교 공과대학 전자공학과 졸업(학사)
1997년 2월 : 고려대학교 대학원 전자공학과 졸업(공학석사)
1997년 3월~현재 : 삼성전자 정보통신연구소 요소부품전문그룹 연구원
※관심분야 : 영상 압축, 멀티미디어 통신



고 성 제(Sung-Jea Ko) 정회원

1980년 2월 : 고려대학교 공과대학 전자공학과 졸업(학사)
1986년 5월 : State Univ. of New York at Buffalo, 전기 및 컴퓨터공학과(공학석사)
1988년 8월 : State Univ. of NEw York at Buffalo, 전기 및 컴퓨터공학과(공학박사)
1981년 8월~1983년 12월 : 대한전선 중앙연구소 연구원
1988년 8월~1992년 5월 : The University of Michigan-Dearbon, 전기 및 컴퓨터공학과 조교수
1992년 3월~현재 : 고려대학교 전자공학과 부교수
1996년 11월 : IEEE best paper award
※관심분야 : 신호 및 영상 처리, 영상 압축 및 통신, 멀티미디어 통신 등.