Compact representation of planar curves based on a wavelet shape descriptor for multimedia applications

Daniel Gatica-Pérez* and Francisco García-Ugalde+

* Human Interface Technology Laboratory. University of Washington. Box 352142. Seattle, Washington 98105, USA

+ División de Estudios de Posgrado. Facultad de Ingeniería. Universidad Nacional Autónoma de México. Apdo. Postal 70-256. México, D.F., 04510. México

e-mail: danielgp@hitl.washington.edu, garcia@verona.fi-p.unam.mx.

Abstract

We present an analysis of a 2-D planar shape descriptor based on the Discrete Wavelet Transform (DWT), as an approach for compact object representation in multimedia applications. To generate the wavelet shape descriptor (WD), a partial 1-D DWT is applied on the vertical and horizontal components of a parameterized closed curve that represents the contour of an object of interest. With this shape representation tool and based on a multiresolution analysis, the curve can be reconstructed on a scale-by-scale basis to the desired degree of approximation. We show that contours of typical multimedia video objects are well represented with a small number of the largest magnitude wavelet coefficients, allowing for an efficient object shape representation. Some issues on the selection of the coefficients are studied. A comparison between the performance of the WD using different wavelet filters and a Fourier shape descriptor is presented, for synthetic and natural contours. The potential use of this descriptor in a multimedia application framework is discussed.

1. Introduction

An efficient representation of shape information is a basic task in many areas of electronic imaging (image and video processing and analysis, and computer graphics). Furthermore, the recent MPEG-4 and MPEG-7 multimedia standards require segmentation and tracking of objects of interest of arbitrary shape, to provide a set of so-called *content-based functionalities* [4]: selective compression, interactive video editing, and retrieval functions. This new paradigm in representation of multimedia information -storing, accessing, exchanging- imposes a need to find reliable and compact shape representation tools that can describe either the spatial characteristics or the temporal evolution of the objects depicted in an image sequence. In particular, for image and video content-based retrieval, the ultimate objective is to integrate these shape representations with several other visual feature descriptors -color, motion, texture- into intelligent models that allow for robust queries in large visual material collections [12].

For several years, wavelet-based techniques have represented a solid area in signal and image processing [7] and have been very succesfully applied in image and video compression [1]. Wavelet algorithms based on matching pursuits [9] have been included for coding of the motion compensated residual signal in the version 2 of MPEG-4 standard, and some others based on embedded zerotree coding will be integrated in the upcoming JPEG standard for still image compression (JPEG-2000). Therefore, it is not suprising that the interest in wavelet approaches for compact object representation, detection and recognition in content-based multimedia applications is increasing [10].

Recently, it was presented in [3] a multiscale descriptor of planar shapes based on periodized wavelets analysis in the continuous metric space $L^2([0,1])$. In this paper we formulate the problem directly using the DWT, and present an analysis of such a descriptor as an efficient contour-oriented representation tool of objects of interest in a multimedia applications framework. Wavelets capabilities for representing local features of a curve, and signal energy compaction are analyzed, and a comparative study of the wavelet descriptor characteristics and performance with respect to a Fourier shape descriptor is presented.

The rest of the paper is organized as follows: section 2 reviews the DWT planar curve descriptor theory; section 3 describes the method for generation of contours of objects of interest in typical multimedia applications; section 4 presents the experimental study on the wavelet descriptor and its comparison with the Fourier descriptor; section 5 discusses the potential of the wavelet descriptor as a compact object representation tool, and section 6 concludes the paper.

2. The wavelet descriptor

Let $\{Z_t^2\}$ be a discrete parameterized closed planar curve -that represents the contour of an object of interest- consisting of N points (say clockwise-oriented) of the form [1]

$$\{Z_t^2\} = \begin{bmatrix} \{X_t\} \\ \{Y_t\} \end{bmatrix} \quad (1)$$

or also, in matrix notation

$$\mathbf{Z^2} = \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix} \quad (2)$$

where t is the arc length, $t = 0, 1, ..., N-1$. Each of the components of $\{X_t^2\}$ and $\{Y_t^2\}$ of $\{Z_t^2\}$ represents a discrete periodic unidimensional function of t. If a level J_0 partial 1-D DWT is applied independently to each component, we can describe the planar curve in terms of the multiresolution analysis (additive decomposition) of $\{X_t^2\}$ and $\{Y_t^2\}$:

$$\begin{aligned} \mathbf{Z^2} &= \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix} = \begin{bmatrix} \mathcal{W}^T\mathbf{W_X} \\ \mathcal{W}^T\mathbf{W_Y} \end{bmatrix} \\ &= \begin{bmatrix} \mathcal{V}_{J_0}^T\mathbf{V_X}_{J_0} \\ \mathcal{V}_{J_0}^T\mathbf{V_Y}_{J_0} \end{bmatrix} + \sum_{j=1}^{J_0} \begin{bmatrix} \mathcal{W}_j^T\mathbf{W_X}_j \\ \mathcal{W}_j^T\mathbf{W_Y}_j \end{bmatrix} \end{aligned} \quad (3)$$

where

- $\mathcal{W}$ is an $N \times N$ real-valued matrix defining the partial DWT [11],

$$\mathcal{W} = \begin{bmatrix} \mathcal{W}_1 \\ \mathcal{W}_2 \\ \vdots \\ \mathcal{W}_j \\ \vdots \\ \mathcal{W}_{J_0} \\ \mathcal{V}_{J_0} \end{bmatrix}$$

- $\mathbf{W_X}, \mathbf{W_Y}$ are column vectors of length N in which the first $(1-2^{-J_0})N$ elements are the wavelet coefficients, and the last $2^{-J_0}N$ elements are the scaling coefficients. In matrix notation, $\mathbf{W_X}_j$, $\mathbf{W_Y}_j$, $j = 1, ..., J_0$ are the first J_0 subvectors of $\mathbf{W_X}$ and $\mathbf{W_Y}$ respectively, and $\mathbf{V_X}_{J_0}$, $\mathbf{V_Y}_{J_0}$ are the corresponding final subvectors. Equivalently, in terms of details and smooths,

$$\mathbf{Z^2} = \begin{bmatrix} \mathcal{S_X}_{J_0} \\ \mathcal{S_Y}_{J_0} \end{bmatrix} + \sum_{j=1}^{J_0} \begin{bmatrix} \mathcal{D_X}_j \\ \mathcal{D_Y}_j \end{bmatrix} \quad (4)$$

[1] In this paper, $\{Z_t^2\}$ indicates that Z is bidimensional, not power operations. The same holds for matrix notation.

where $\mathcal{D_X}_j \equiv \mathcal{W}_j^T\mathbf{W_X}_j$, $j = 1, ..., J_0$, are the detail vectors whose elements are related to changes in the $\mathbf{X}$ component at scale 2^{j-1}, and the smooth $\mathcal{S_X}_{J_0} \equiv \mathcal{V}_{J_0}^T\mathbf{V_X}_{J_0}$ represents averages at a scale 2^{J_0} (the same applies to the $\mathbf{Y}$ component).

The *wavelet descriptor WD* for the planar curve $\mathbf{Z^2}$ is defined as the coefficients of the partial DWT for $\mathbf{X}$ and $\mathbf{Y}$, that is

$$WD(\mathbf{Z^2}) = \begin{bmatrix} \mathbf{W_X} \\ \mathbf{W_Y} \end{bmatrix} \quad (5)$$

With this definiton, a multiresolution representation of the closed curve $\mathbf{Z^2}$ has been obtained. This allows for an efficient representation of the shape, because by using only a certain number of DWT coefficients -those that correspond to the coarser scales- we can approximate the original curve on a scale-by scale basis. Therefore, if J' denotes the level of approximation ($-1 \leq J' \leq J_0 - 1$), the reconstructed curve $\hat{\mathbf{Z}}^2$ of $\mathbf{Z^2}$ is given by

$$\begin{aligned} \hat{\mathbf{Z}}^2 &= \begin{bmatrix} \hat{\mathbf{X}} \\ \hat{\mathbf{Y}} \end{bmatrix} \\ &= \begin{bmatrix} \mathcal{V}_{J_0}^T\mathbf{V_X}_{J_0} \\ \mathcal{V}_{J_0}^T\mathbf{V_Y}_{J_0} \end{bmatrix} + \sum_{j=J_0-J'}^{J_0} \begin{bmatrix} \mathcal{W}_j^T\mathbf{W_X}_j \\ \mathcal{W}_j^T\mathbf{W_Y}_j \end{bmatrix} \\ &= \begin{bmatrix} \mathcal{S_X}_{J_0} \\ \mathcal{S_Y}_{J_0} \end{bmatrix} + \sum_{j=J_0-J'}^{J_0} \begin{bmatrix} \mathcal{D_X}_j \\ \mathcal{D_Y}_j \end{bmatrix} \end{aligned} \quad (6)$$

where $J' = -1$ corresponds to the coarsest level - only the smooth term or scaling coefficients- and $J' = J_0 - 1$ corresponds to perfect reconstruction (by definiton, $\sum_{j=J_0+1}^{J_0} \mathcal{D_X}_j = 0$).

In [3], the derivation of the wavelet descriptor is based on the analysis of periodized wavelets in the continuous metric space $L^2([0,1])$, which leads to the derivation of (3). As we have shown, the same descriptor can be obtained from the direct use of the DWT, working only in the discrete domain.

2.1 Properties of the wavelet descriptor

The two properties of wavelets that are exploited for shape characterization are the capability for detecting and representing local features, and the energy compaction in the transform coefficients. The WD has invariance, uniqueness and stability properties, assuming that the parameterization of curves has the same starting point [3]. This assumption is a consequence of a well known fact: the discrete wavelet transform coefficients and the multiresolution analysis are not invariant with respect to circular shifts. Due to the scanning-based method for extraction of the first point of the planar contour, for some particular kind of objects (walking people), the assumption can be somewhat reasonable -the upper-most point (in image coordinates) can correspond to the contour starting point-. Methods for selection of the starting point of a contour are discussed in [2]. In practice, this problem is usually addressed by maximum correlation schemes.

Given the above assumption and properties, the WD can be used for recognition purposes based on the wavelet coefficients. A method for creating a geometric invariant (rotation, translation and change of size) WD is presented in [3]. Furthermore, these geometric transformations can be interpreted as global rigid motion of the object represented by its contour in an image sequence.

2.2 The Fourier descriptor

A wavelet descriptor is an alternative to a Fourier descriptor (*FD*), a shape representation tool that has been used for years in computer vision and graphics [8]. The definiton of such a descriptor is straightforward. For a planar parameterized curve $\mathbf{Z^2}$, a FD can be defined as

$$FD(\mathbf{Z^2}) = \begin{bmatrix} \mathbf{F_X} \\ \mathbf{F_Y} \end{bmatrix} \tag{7}$$

where $\mathbf{F}$ denotes the Orthogonal Discrete Fourier Transform (ODFT). This transform also allows for the approximation of $\mathbf{Z^2}$ in terms of an additive decomposition [11]:

$$\hat{\mathbf{Z}}^\mathbf{2} = \begin{bmatrix} \hat{\mathbf{X}} \\ \hat{\mathbf{Y}} \end{bmatrix} = \sum_{j=0}^{K'} \begin{bmatrix} \mathcal{D}_{\mathcal{F}_\mathcal{X}}, j \\ \mathcal{D}_{\mathcal{F}_\mathcal{Y}}, j \end{bmatrix} \tag{8}$$

where $\mathcal{D}_{\mathcal{F}_\mathcal{X}}, j$ and $\mathcal{D}_{\mathcal{F}_\mathcal{Y}}, j$ are the jth order Fourier details for $\mathbf{X}$ and $\mathbf{Y}$, and where $K' = 0, ..., \lfloor N/2 \rfloor$.

3. Object contour generation in natural images

For multimedia applications, automatic and semi-automatic segmentation tools can be used for object contour definition. In this paper, we generate moving object contours in digital video by using the following procedure:

1. Markov Random Field motion-based segmentation based on two consecutive frames [5]. A binary image I representing the moving object is generated. The extracted motion borders tend to be close to the real object borders, although there are errors in the segmentation due to noisy motion estimation.

2. Mathematical morphology processing stage on the motion segmentation mask to obtain the contour of the moving object. An opening-closing operation followed by a morphological gradient by erosion generate the object contour.

3. Contour sampling to obtain a small number of contour points. 256 points for synthetic shapes and 128 for natural contours are used in this paper. A smaller number is usually enough to describe the original curves, depending on the size and shape complexity of the moving object. However, in this study we use this number of points to ilustrate the effects of the multiresolution descomposition using a reasonable number of levels.

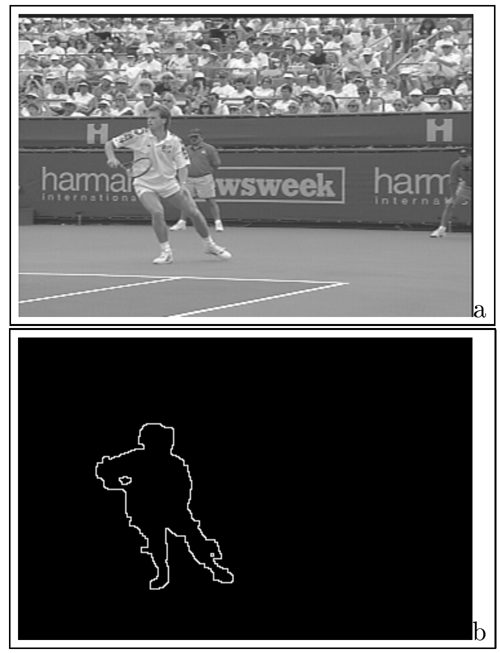

Figure 1: (a) *Stefan* image; (b) final contour.

Fig. 1 shows an example of the extraction contour procedure in *Stefan* image sequence. It can be seen that the contour does not perfectly correspond to the real frontiers of the *object of interest*, as accurate segmentation of arbitrary moving objects in complex scenes is a difficult task. There are some evident visual artifacts in this contour.

4. Experimental analysis

The WD was efficiently implemented using the Mallat pyramid algorithm [11]. As test images, we used synthetic planar curves and also contours of real objects (Fig. 2). For the computation of the partial DWT we chose $J_0 = 6$, in order to have 4 coefficients for each component of $\hat{\mathbf{Z}}^\mathbf{2}$ at the coarsest level, for the 256-point curves (less than 4 wavelet coefficients do not reconstruct meaningful curves); for the *Stefan* curve the coarsest scale corresponds to 2 coefficients as $N = 128$.

4.1 On the selection of the wavelet filter

We present in Fig. 3 the results of the multiresolution approximation using the *WD* for the *Star* and *Phantom* curves, for $J' = -1, 0, 1$ and 2 (that correspond to 4, 8, 16 and 32 wavelet coefficients for each component of $\hat{\mathbf{Z}}^{\mathbf{2}}$), using three different wavelet filters: Haar, Daubechies class D(4) and least asymmetric (LA(8)). In all the results presented in this paper, the x and y axes represent the horizontal and vertical location of the planar curves in the image in pixels (for the synthetic curves, this position has been normalized). We can observe that the well known visual effect of each wavelet filter is evident: the Haar's blocky but symmetric effect traduces into a piecewise polygonal approximation for the 2-D shape; the asymmetric shark fin effect of D(4) is also present, and a less asymmetric and smoother approximation is observed for LA(8). From the tested filters, the best approximation is produced, then, by the LA(8) wavelet, and it requires between 16 and 32 coefficients to properly represent the original curves ($J' = 2$).

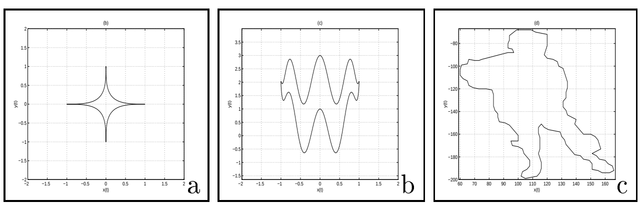

Figure 2: Test planar curves.(a) *Star*; (b) *Phantom*; (c) *Stefan*.

We also show in Fig. 3 the results obtained for the *Stefan* curve for all levels of reconstruction. Again the best approximation is obtained with LA(8). For $J' = 3$ (32 wavelet coefficients for each component of $\hat{\mathbf{Z}}^{\mathbf{2}}$) we obtain a contour in which the general shape of the human body is well described, but some details have been eliminated (this is advantageous as some of the shape artifacts have been removed or diminished, see again Stefan's legs). The superimposition of this approximated contour with the original image is shown in Fig. 4(b). As a simple quantitative approximation measure, we computed the mean euclidean distance between the original contour and its wavelet approximation; the approximation error is equal to 1.46 pixels. We consider that this simplified contour could be useful for shape recognition purposes, and also as the initial point for a procedure of adjustment of the contour to the real frontiers of the object of interest using a deformable contour [6].

In [3], a discussion on the election of the wavelet bases for shape representation is presented, based on several criteria. A large number of wavelets are investi-

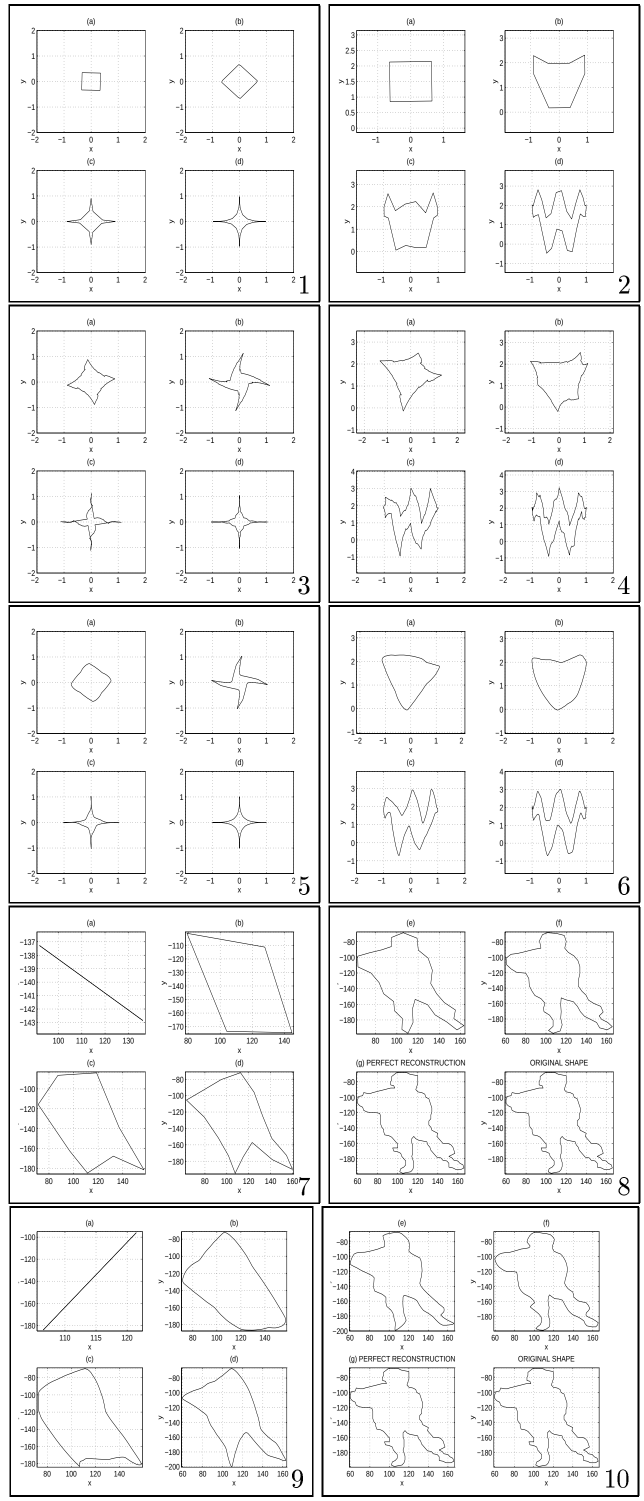

Figure 3: Multiresolution approximation using the WD. For *Star* and *Phantom* curves, the first four levels are shown using (1)-(2) Haar. (3)-(4) D(4). (5)-(6) LA(8) wavelet. For *Stefan* curve, all levels of approximation are used: (7)-(8) Haar; (9)-(10) LA(8) wavelet.

gated. The authors come to three conclusions: (1) for symmetric patterns, wavelets with symmetric or almost symmetric scaling functions (biorthogonal spline-based or Coiflets) produce better results; (2) the Daubechies class, with a relatively small computational cost, preserves corners in non smooth curves; and (3) bases with high vanishing moments, like Coiflets, produce better approximations but are more time-consumming. In fact, even if many natural objects present a certain degree of symmetry, this can be easily broken due to deformable motion and change of viewpoint. In the rest of this paper, we will present results with the LA(8) wavelet filter, that represented a good compromise between computational speed and shape representation.

Figure 4: (a) Original *Stefan* contour; (b) approximated contour using the wavelet descriptor.

For comparison purposes, we show the results obtained using the Fourier descriptor in Fig. 5. The FD generates symmetric approximations of the curves, and, from a given level of reconstruction, it produces similar results to those generated by the WD when an equivalent number of coefficients to synthesize the curves is used [2]. A superimposition of the original contour and its Fourier approximation using 32 coefficients is shown in the last sketch of Fig. 5. The approximation error in this case is 1.33 pixels, which is slightly smaller than the wavelet case.

[2]Note that the shape reconstruction using the WD is done on a scale-by-scale basis, and on an increasing frequency basis for the FD. That is to say, no reordering of the transform coefficients in terms of their magnitude has been done up to this moment.

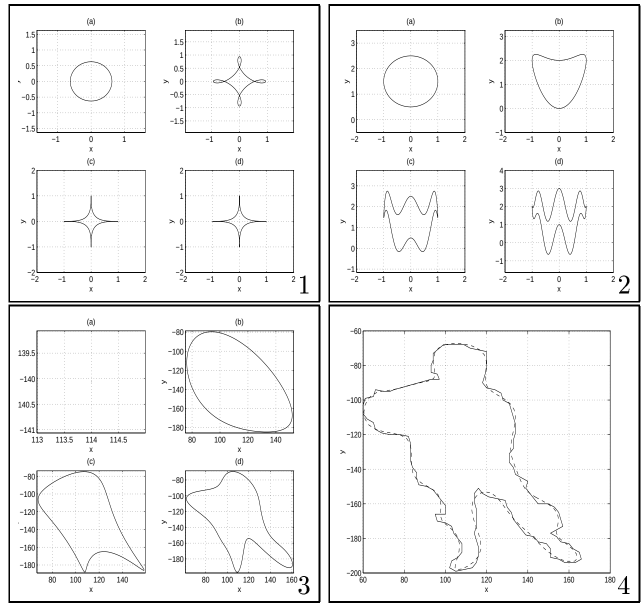

Figure 5: Reconstruction using the FD. The aproximations are done using $K' = 0, 1, 2$ and 3. For *Stefan*, the superimposition of the original contour (solid line), and the Fourier approximation for $K' = 4$ (dashed line) is also shown.

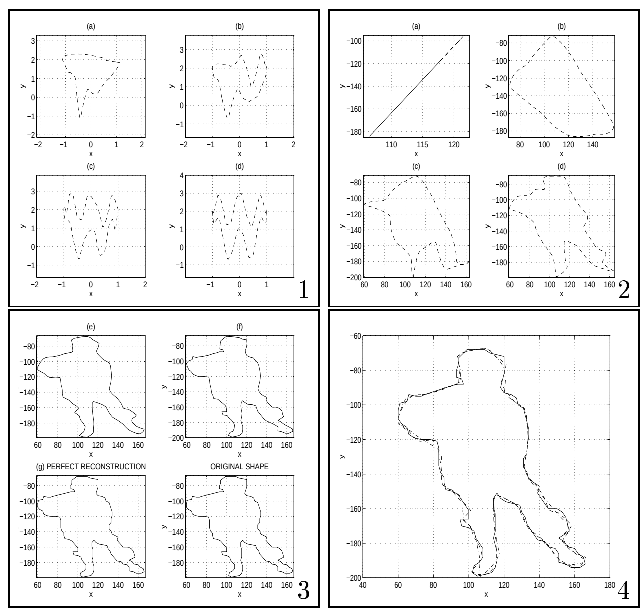

Figure 6: Approximation using the M largest LA(8) wavelet coefficients. For *Phantom*, the aproximations use the same number of coefficients as in section 4.1; for the *Stefan* curve all levels are shown. In (4) the superimposition of the original contour (solid line), and the two approximations using the scale-by-scale approach (dash-dotted line), and the M largest coefficients method (dashed line) is shown.

4.2 On the selection of the wavelet coefficients for shape representation

One of the main advantages of the WD over the FD is its capability of representing and detecting local features of a curve in view of the localization property of wavelets. An improvement in the approximations using the WD is obtained when the curves are selectively reconstructed using the DWT coefficients of largest energy. In Fig. 6 we show the results of reconstructing the test curves using the LA(8) wavelet and the M largest wavelet coefficients, retaining for each case the same number of transform coefficients as in the previous section. Each component $\mathbf{W_X}$ and $\mathbf{W_Y}$ are independently thresholded. The process of reconstruction looks quite different from the one obtained using the scale-by-scale reconstruction. In Fig. 6.4 we present a superimposition of *Stefan* contour and the contours approximated by both the multiresolution synthesis and the M largest coefficients. It can be seen that the M largest coefficients approach provides a more precise approximation. The approximation error is equal to 1.16 pixels, which means that we can achieve a greater data compression by using the M largest coefficients approach. On the other hand, the multiresolution contour obtained in the previous section is smoother.

Other schemes for the coefficient thresholding are possible. The last test illustrates what happens when we synthesize the curve using the coefficients that correspond to the M largest joint magnitude

$$W_{X\,j,t}^{\,2} + W_{Y\,j,t}^{\,2} \tag{9}$$

The results are shown in Fig. 7. The approximation error for the case we are analyzing is 1.31 pixels. We also show a superimposition between the original contour and its M largest joint coefficients approximation. To evaluate the effect of the selection method of the wavelet coefficients for shape reconstruction, we show the reconstruction error as a function of the number of coefficients using the M independent largest coefficients method and the M largest joint coefficients approach. The performance using the independent largest coefficients is better than the corresponding to the joint case. For comparison purposes, we also show the described Fourier descriptor performance, which is in general inferior than the wavelets methods. This figure can be used as a simple quantitative rate-distorsion indicator. For instance, we see that with 20 coefficients using the M independent largest coefficients method, we obtain a reconstruction error of about 2 pixels. We also see that, in order to obtain the same error, we need about 25 coefficients using the M largest joint coefficients approach, and about 27 using the Fourier descriptor. When using more than 30-35 coefficients, all methods perform in a similar way. It is evident that these experiments are not completely conclusive. We are currently investigating the correlation between local 2D shape and wavelet coefficients using the WD.

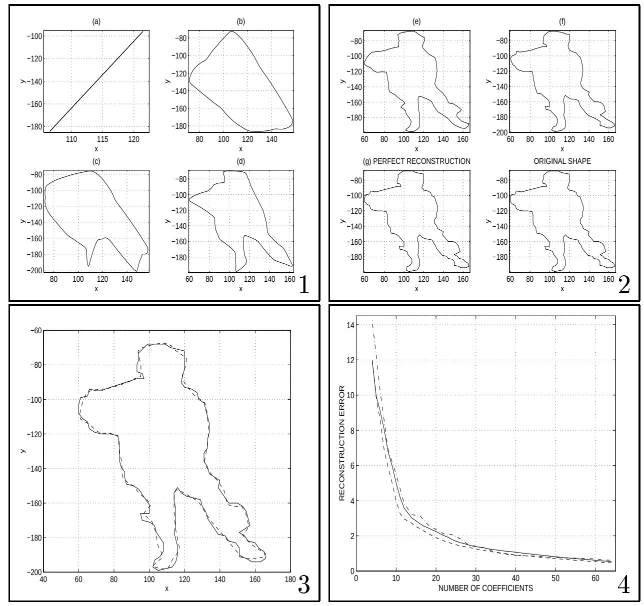

Figure 7: Approximation using the M joint largest $W_{X\,j,t}^{\,2}+W_{Y\,j,t}^{\,2}$ wavelet coefficients, for all levels of reconstruction. In (3) we superimpose the original contour (solid line), and the approximation using this approach (dashed line). In (4) we show a comparison of the reconstruction error for a variable number of coefficients using the M independent largest coefficients (dashed line), the M largest joint coefficients (solid line), and the Fourier descriptor (dash-dotted line).

5. Use of the wavelet descriptor in multimedia applications

There are several applications of the WD in a multimedia framework, as a compact shape representation tool for complex object description and characterization in digital video. Three research lines in which the WD can find applications related to the content-based approach proposed in MPEG-4 and prevailing in MPEG-7 standards are:

1. The WD allows for the modeling of time-varying planar shapes, as we can divide this dynamic behaviour into global rigid motion and local non-rigid deformations. With respect to this local characterization, in [3] it was proposed a deformable wavelet descriptor in which the wavelet coefficients -modeled as random variables- are used to adjust an initial contour to the real borders of an ob-

Figure 8: Application of the WD for retrieval functions in video databases. Two video commands (*hand1* and *hand2*) are shown in (a)-(b) and (e)-(f). The contours ((d), (h)) extracted from the motion segmentation masks ((c), (g)) are represented by the WD. In (i) and (j) the reconstruction for the four first levels is shown.

ject of interest based on optimization techniques (bayesian methods). The main advantage of using the WD is that the adjustment of parameters can be hierarchically implemented, based on the multiscale representation: coarser scales are used for the first adjustment stages, and then local deformations are handled at finer resolutions. However, up to this moment, only very simple object shapes have been analyzed under this perspective. Its applicability to more complex objects needs to be validated.

2. For model object detection, segmentation and tracking, the features provided by the WD could be introduced in learning deformable models [2]. Typical schemes use the position of the discrete contour points to compute a shape vector, by approximating the contour by a parametric curve (B-spline). These shape vectors are normalized and used as training shapes for Principal Component Analysis (eigenshapes). A different possibility would be to use the WD as the features from which the shape model is learned.

3. The WD can also find applications in generic model-free content-based video retrieval. In Fig. 8 we present some of our current experiments. A set of hand commands for human-machine communication is depicted in a video database (Fig. 8(a)-(b), (e)-(f)). The procedure described in this paper -generation of contour of the moving object, computation of the wavelet descriptor- is applied to each of the video sequences in the database. $N = 128$ points are extracted for each contour. A good representation of the hand shapes, useful for shape-based retrieval, is obtained with only a few wavelet coefficients. The multiresolution reconstruction for two hand commands is shown in Fig. 8(i)-(j) for $J' = -1, 0, 1$ and 2 (that correspond to 2, 4, 8, and 16 wavelet coefficients for each component of $\hat{\mathbf{Z}}^{\mathbf{2}}$). The optimum number of coefficients might depend on the complexity of the indexed command alphabet. In Fig. 9 we show the contour recontruction error as a function of the number of coefficients, for the two hands sequences. We notice again that the M independent largest coefficient method produces a more compact representation. Issues on retrieval efficiency using the WD by itself and in combination with other descriptors (object color, motion, described trajectory along the video sequence) are part of our current work.

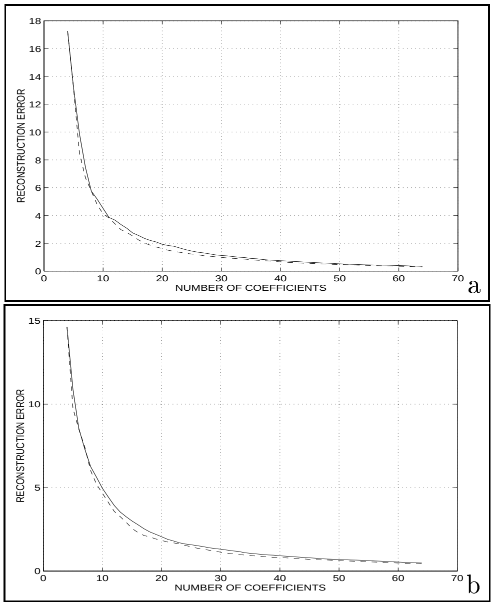

Figure 9: Approximation of (a) *hand1* and (b) *hand2* contours using the WD as a function of coefficients using the M independent largest coefficients (dashed line), the M largest joint coefficients (solid line).

6. Conclusions

We analyzed a planar curve descriptor based on the DWT, as a tool for compact object representation in multimedia applications. Our experiments included testing of different wavelet filters on natural and synthetic 2-D contours, and a comparison between three shape reconstruction methods (multiresolution, M independent x and y largest coefficients and M joint largest coefficients). A comparison with a well known Fourier shape descriptor was also performed. From the experiments, we conclude that the use of the M independent largest coefficient methods produces more compact representations, although the precise correspondence between wavelet coefficient magnitudes and local changes in shape is still a problem to be addressed. Our current work includes a thorough quantitative evaluation of the compactness and performance of the wavelet descriptor in video retrieval applications.

7. Acknowledgments

The authors sincerely thank Prof. Donald B. Percival for his support on wavelet theory. The first author also thanks the National University of Mexico and the CONACYT Scholarship Program. The second author thanks the IRISA support for the presentation of this paper.

8. References

[1] M. Barlaud, ed. *Wavelets in Image Communication*, Elsevier, 1994.

[2] A. Baumberg and D. Hogg, "Learning deformable models for tracking the human body". In M. Shah and R. Jain (eds.), *Motion based recognition*, Dordrecht, Kluwer, 1997. pp. 39-60.

[3] G.C.-H. Chuang and C.-C.J. Kuo, "Wavelet descriptor of planar curves: theory and applications", *IEEE Trans. on Image Processing*, Vol. 5, No. 1, pp. 56-70, January 1996.

[4] T. Ebrahimi, "MPEG-4 video verification model: a video encoding/decoding algorithm based on content representation", *Signal Processing: Image Communication*, Vol. 9, pp. 367-384, 1997.

[5] D. Gatica-Pérez, F. García-Ugalde and V. García Garduño , "Segmentation algorithm for image sequences from a pel-recursive motion field", in *VCIP '97*, Proc. SPIE 3024, pp. 1152-1163, 1997.

[6] A. K. Jain, Y. Zhong and S. Lakshmanan, "Object matching using deformable templates", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 18, No. 3, pp. 267-278, March 1996.

[7] S. Mallat, *A wavelet tour of signal processing*, Academic Press, 1998.

[8] F. Marqués, B. Llorens and A. Gasull, "Prediction of Image Partitions using Fourier descriptors: applications to segmentation-based coding schemes", *IEEE Trans. on Image Processing*, Vol. 7, No. 4, pp. 529-542, April 1998.

[9] R. Neff and A. Zakhor, "Very low bit-rate video coding based on matching pursuits", *IEEE Trans. on Circ. and Sys. for Video Tech.* Vol. 7, No. 1, pp. 158-171, February 1997.

[10] S. Panchanathan, M. Mandal and T. Aboulnasr, "Video Indexing in the Wavelet Compressed Domain" *Proc. of the IEEE ICIP*, pp. 546-550, Chicago, 1998.

[11] D. Percival and A. Walden, *Wavelet Methods for Time Series Analysis*, Cambridge University Press, 1999 (to appear).

[12] Yong Rui, Thomas S. Huang, and Shih-Fu Chang, "Image Retrieval: Current Techniques, Promising Directions and Open Issues", *Journal of Visual Communication and Image Representation*, 10, 1-23 (1999).