A Wavelet based Contour Representation for fitting and tracking applications.

FERNANDO PÉREZ NAVA[1], ANTONIO FALCÓN MARTEL[2]
[1] Dep. de Estadística, Inv. Op. y Computación. Universidad de La Laguna. Tenerife. Spain
[2] Dep. de Informática y Sistemas, Universidad de Las Palmas de Gran Canaria.Gran Canaria.Spain
e-mail: [1]fdoperez@ull.es, [2]afalcon@dis.ulpgc.es

Abstract

A technique is presented to construct a multiscale representation of planar contours based on the wavelet transform (WT). To generate this wavelet description, a partial 1-D discrete wavelet transform (DWT) is applied to the vertical and horizontal components of a length-parametrized planar curve. This multiscale representation descomposes the curve into different levels of resolution and allows to reconstruct it to a desired degree of aproximation. The results show that typical objects are well represented by a small number of wavelet coefficients allowing for a compact object shape representation. An aplication of this compact representation for matching and tracking purposes is presented and its results are analized.

Keywords: pattern recognition, wavelets, shape description, matching, tracking.

1. Introduction

A powerful propery for distinguising an object from its surroundings in an image is overall shape. Shape can be used to complete the information provided by other local properties in an image such as gray level, texture or color. Therefore an efficient representation of shape information is a basic task in many areas of computer vision, video processing and analysis and computer graphics.

Many shape representations that are potentially useful for shape description have been developed [1]. Two-dimensional shape can be represented using a real or complex 1-D function. From these representations, shape descriptors can arise in the form of Chain codes, Polygonal aproximations, Elliptic fourier descriptors, B-Splines or Multiscale gaussian descriptions.

Recently it was presented in [2] a multiscale descriptor of planar shapes based on periodized wavelet analysis in the continous metric space $L^2([0,1])$. In [3] an analysis of such description was presented based on the Discrete Wavelet Transform. It was shown that shapes can be represented with a small approximation error, allowing for an efficient shape representation.

However, in order to use a wavelet representation in practical applications it is necesary to obtain an invariant wavelet representation. In this work we show how such representation can be obtained and we use such representation to the contour matching and tracking problem.

2. The Wavelet description.

A wavelet basis uses translations and dilations of a scaling function ϕ and a wavelet functions φ. A 1-D function f can be expresed as:

$$f(x) = \sum_{k \in Z} c_{J,k}\, 2^{\frac{J}{2}} \phi(2^J x - k) + \sum_{j=1}^{J} \sum_{k \in Z} d_{J,k}\, 2^{\frac{J}{2}} \varphi(2^J x - k)$$

Let then $\mathbf{r}(s)$ = (x(s), y(s)) be a discrete parametrized closed planar curve that represents the shape of an object of interest. If the wavelet transform is applied independently to each of the x(s), y(s) functions, we can describe the planar curve in terms of a decomposition of $\mathbf{r}(s)$:

$$\mathbf{r}(s) = \sum_{k \in Z} \mathbf{c}_{J,k}\, 2^{\frac{J}{2}} \phi(2^J s - k) + \sum_{j=1}^{J} \sum_{k \in Z} \mathbf{d}_{J,k}\, 2^{\frac{J}{2}} \varphi(2^J s - k)$$

$$\mathbf{c}_{j,k} = \begin{pmatrix} c_{\ j,k;\mathrm{x}} \\ c_{\ j,k;\mathrm{y}} \end{pmatrix}, \quad \mathbf{d}_{j,k} = \begin{pmatrix} d_{\ j,k;\mathrm{x}} \\ d_{\ j,k;\mathrm{y}} \end{pmatrix}.$$

where subindex x and y represent function pertenence, obtaining a multiresolution representation of shape where coarser scales provide a simplified representation of shape and finer scales add more detail to the contour representation (Fig 1):

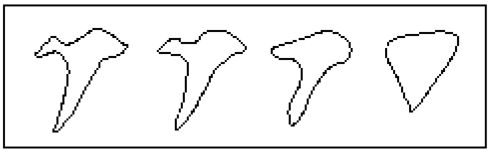


Figure 1 Multiscale representation of shape

2.1 Properties of the wavelet descriptors

The main reason in using wavelets is its capability of detecting and representing local features. The WDT has invariance, uniqueness and stability properties asuming that the parametrization of curves has the same starting point . This is a consequence of a well known fact: the WDT coefficients are not invariant to parametrization shifts.

3. Wavelet descriptors

The WDT and the multivavelet discrete transform (MWDT) can be efficiently implemented in linear time using a pyramid algorithm. Since our goal is to obtain a compact representation of contour it is necesary to develop a contour simplification strategy.

In [3] it is proposed to threshold independently the $c_{j,k,x}$ and $c_{j,k,y}$ coeficients, and their results are discused. Although this strategy provides a compact representation it is not invariant. It can be easily show that a contour rotation does not conserve the magnitude of these coefficients and therefore a small coeficient that has to be thresholded at one angle may be made not to be thresholded by a suitable rotation. It is therefore necesary to obtain an invariant thresholding rule. We propose to threshold the $\mathbf{c}_{j,k}$ vector based on its norm. It can be show that this yields an invariant simplification. In order to test the simplification capability of this rule we have aplied this thresholing rule to a set of synthetic planar curves (see Fig 2). For each image we only used the outline of the figure, interior contours were not processed. All contours were resampled in order to obtain 256 points and decomposition was taken to the coarsest level. Then the set of the *n* most important coefficients (coefficients with maximum amplitude) were obtained for different values of *n*=32, 64, 128.

The mean distance error between the reconstructed curve curve from this set of coefficients and the original curve is then show for distinct wavelet families: Haar, Daubechies class D4 and least asymmetric LA8, and distinct multiwavelet families (Geronimo-Hardin-Massopust GHM, Chu-Lian CL, and Shen-Tan-Tham SA(4)) for comparison purposes the Elliptic Fourier Transform and B-Splines are also included.

The results of the mean error distance (in pixel units) for the complete set of images where:

	n= 32	n= 64	n=128
Fourier	**4.0**	**1.8**	**0.8**
B-Splines	**3.9**	**2.1**	**0.9**
Haar	**6.1**	**3.3**	**1.7**
D4	**2.8**	**1.2**	**0.6**
La8	**2.4**	**1.0**	**0.4**
GMH	**2.5**	**1.1**	**0.5**
CL	**2.7**	**1.1**	**0.5**
SA(4)	**2.3**	**1.0**	**0.5**

Table 1 Mean distance error for test images

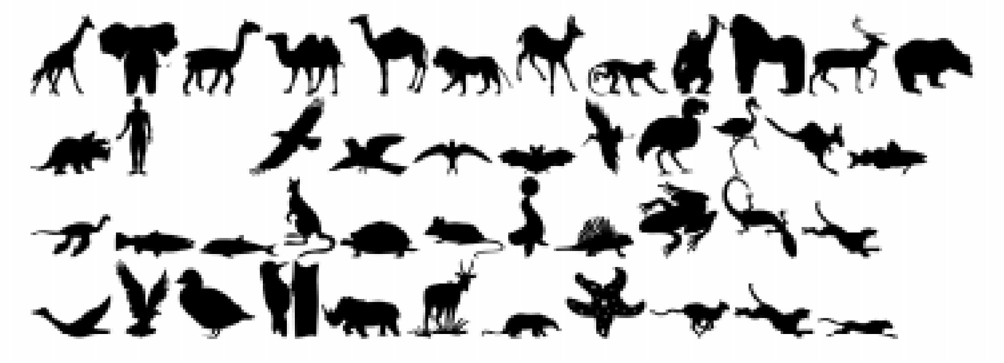


Figure 2 Test images

Numerical results show in general, that best compresion is achieved with the SA(4) multiwavelet and the LA8 wavelet. B-spline results where obtained from subsampling the contour. In Figure 3 it is shown the mean error distance (n=32) for all test data for Fourier elliptic descriptors (thick line), LA8 wavelet (discontinous line) and SA(4) multiwavelet (thin line). Test contours are numbered left to right and top to bottom. The worst contour for both SA(4) and LA(8) is show in Figure 4 where the original contour (thick line), is shown with the LA8 wavelet approximation (discontinous line) and SA(4) multiwavelet approximation (thin line).

In shape processing it is generally desirable not only a description with a small mean distance. Often it is useful to obtain a small maximum distance between the original contour and its approximation. To test the capacity of this description to obtain a low maximum error, the 95% percentile of the maximum distance was computed for the LA8 and SA(4) multiwavelet for all images in the database obtaining the results in Table 2. Therefore resultss show that we have obtained a compact contour representation that is also invariant. This propery will be used in the next sections.

	n= 32	n= 64	n=128
La8 wavelet	**4.14**	**1.72**	**0.76**
SA(4) multiwavelet	**4.09**	**1.80**	**0.79**

Table 2. The 95% percentile of the distance error

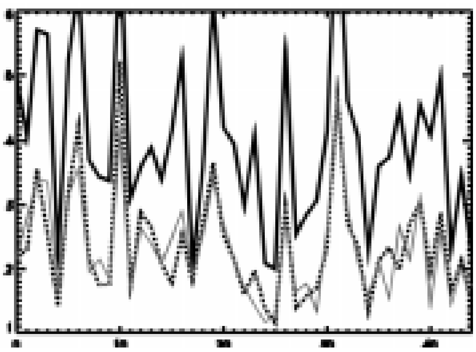


Figure 3 Error comparison

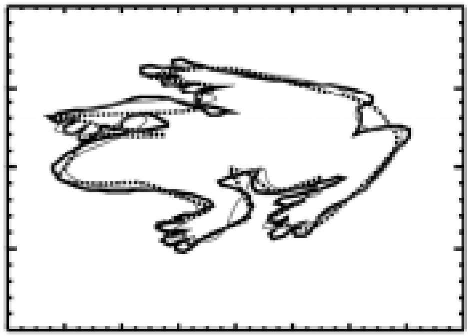


Figure 4 Maximum test image error

4. Wavelet based Contour fitting.

The contour fitting problem can be stated as [5]: $\min_{\mathbf{r(s)}} \ \alpha\left\|\mathbf{r}-\overline{\mathbf{r}}\right\|^2+\left\|\mathbf{r}-\mathbf{r_f}\right\|^2$ where $\overline{\mathbf{r}}$ is a reference shape and $\mathbf{r}_f$ is the detected shape on image. The first term is a regularization expresion that biases the fitted curve toward a mean shape to a degree determined by a regularization constant α. The second term biases the solution to the contour detected in image. Using the orthogonality of the wavelet coefficients the problem can be stated as:

$\min_{\mathbf{C}} \ \alpha\left\|\mathbf{C}-\overline{\mathbf{C}}\right\|^2+\left\|\mathbf{C}-\mathbf{C_f}\right\|^2$ where $\mathbf{C}$ is a vector with the wavelet coeficients $c_{i,j,x}$ and $c_{i,j,y}$.

In practice it is desirable that the regularizing term is invariant against rotations and translations so that it only influences the shape of the fitted curve and not its position or orientation. This can be obtained by projecting the shape on the subspace of deformations by a suitable projection matrix $\mathbf{S}$. The problem can then be reformulated as:

$$\min_{\mathbf{C}} \ \alpha\left(\mathbf{C}-\overline{\mathbf{C}}\right)^T\mathbf{S}\left(\mathbf{C}-\overline{\mathbf{C}}\right)+\left(\mathbf{C}-\mathbf{C_f}\right)^T\left(\mathbf{C}-\mathbf{C_f}\right)$$

this formulation and the wavelet description allows a fast, inversion-free evaluation of the optimal solution $\hat{\mathbf{C}}$

$$\begin{aligned}
\hat{\mathbf{C}} &= (\alpha\,\mathbf{S}+\mathbf{I})^{-1}(\alpha\,\mathbf{S}\,\overline{\mathbf{C}}+\mathbf{C_f}) = \\
&((\alpha+1)\mathbf{S}+\mathbf{I}-\mathbf{S})^{-1}(\alpha\,\mathbf{S}\,\overline{\mathbf{C}}+\mathbf{C_f}) = \\
&\left(\left(\frac{1}{\alpha+1}\right)\mathbf{S}+(\mathbf{I}-\mathbf{S})\right)(\alpha\,\mathbf{S}\,\overline{\mathbf{C}}+\mathbf{C_f}) = \\
&\left(\left(\frac{1}{\alpha+1}\right)\mathbf{S}(\alpha\,\overline{\mathbf{C}}+\mathbf{C_f})+(\mathbf{I}-\mathbf{S})\mathbf{C_f}\right) = \\
&\mathbf{S}\left(\frac{\alpha}{\alpha+1}\overline{\mathbf{C}}+\frac{1}{\alpha+1}\mathbf{C_f}\right)+(\mathbf{I}-\mathbf{S})\mathbf{C_f}
\end{aligned}$$

Note how the preceeding expresion allows an easy interpretation: In the deformation subspace a weighted mean between the mean shape and the image shape is carried out. In the translation and rotation subspace (corresponding to the **I-S** matrix) all information is taken from the contour in the image.

An experiment showing the fitting algorithm is shown below, in Figure 5 we can see the initial position of the contour (mean shape) , then in Figure 6 we can see the fitted contour.

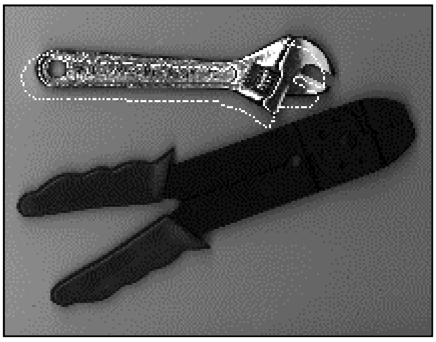


Figure 5 Mean Shape

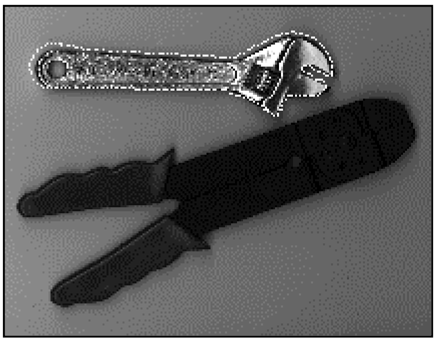


Figure 6 Fitted Contour

5.- Wavelet based contour tracking

In order to use the previous contour fitting algorithm to more complex applications such as tracking an important limitation appears: it is necesary to know the entire shape of the curve in the image. This makes that the algorithm has a limited application to noise free and non cluttered images. If this is not the case, it is necesary to infer the wavelet coefficients from the partial shape in image. This makes the fitting minimization expression more complex so that matrix inversion can not be avoided. The main benefit in this case comes from the compact representation of the contour which reduces significatively the dimensionality of the problem using the thresholding rule in section 3.

The benefits from this compact representation can also be used where a parametric contour formulation is needed. We now show a tracking application where the active contour[5] parametric formulation has been used.

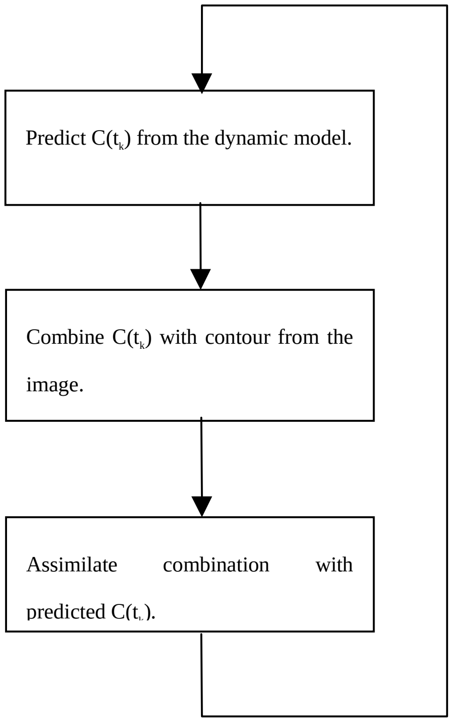
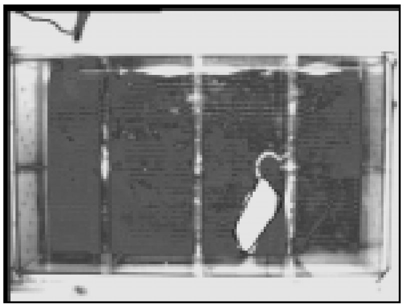


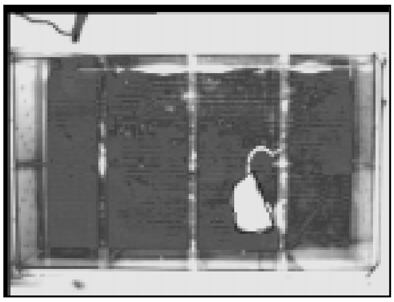
Figure 7 Kalman filter

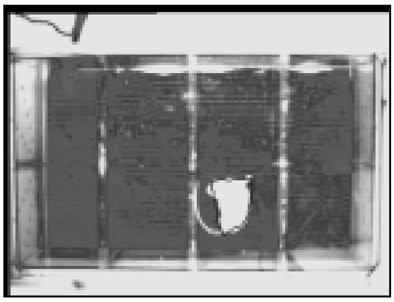
In this formulation the wavelet coeficients evolution is modeled as an second order AR process:

$$\mathbf{C}(t_k) - \overline{\mathbf{C}} = \mathbf{A}_2(\mathbf{C}(t_{k-2}) - \overline{\mathbf{C}}) + \mathbf{A}_1(\mathbf{C}(t_{k-1}) - \overline{\mathbf{C}}) + \mathbf{B}_0\mathbf{w}_k$$

where $\mathbf{w}_k$ is independent gaussian noise with zero mean and $\overline{\mathbf{C}}$ is the mean shape defined in the previous section. Contour tracking is then carried out with a Kalman filter using a propagation loop (Figure 7). In Figures 8-10 we show an implementation of the tracker. Wavelet coefficient dynamics are descomposed in three subspaces: translation, rotation and deformation, and different parametrizations are defined n them reflecting the different evolution on these subspaces. The wavelet decomposition used has been Daubechies LA(8).







Figures 8,9,10 Shape tracking. Contour is overimposed on image with color black

Results show how the mouse is correctly tracked even in the presence of severe deformations of the shape using only 8 wavelet coefficients.

6. Conclusions

In this paper we have presented an invariant compact representation of the contour. Its ability to simplify contour representation has been show by several examples from a contour database. Moreover this representation has been used to solve fitting and tracking problems where complexity has been reduced by means of the wavelet representation of the contour. This provides an efficient representation capable to provide a set of so-called content-based functionalities such as tracking required by the recent MPEG-4 and MPEG-7 multimedia standards.

7. References

[1] S. Loncavic, A Survey of Shape Analysis Techniques, Pattern Recognition, Vol 31, nº8, pp 983-1001(1998)

[2] G.C.-H. Chuang and C.-C.J. Kuo, “Wavelet description of planar curves”, IEEE Trans. on Image Procesing, Vol 5, Nº 1, pp 56-70, (1996)

[3] D. Gatica-Pérez and F. García Ugalde, “Compact Representation of planar curves based on wavelet shape descriptors for multimedia applications”, Vision Interface 99 pp- 604-611 (1999)

[4] G.Plonka and V. Strela, From Wavelets to Multiwavelets, Math. Methods for Curves and Surfaces, Vanderbilt Univ. Press, pg-375-399 (1998).

[5] A. Blake and M. Isard, *Active Contours*, Springer (1998)