

# Wavelets for Multiresolution Shape Recognition

Maria Grazia Albanesi, Luca Lombardi

University of Pavia, Dipartimento di Informatica e Sistemistica

Via Ferrata 1, I-27100 Pavia, Italy

Phone: +39 +382-505350

Fax: +39 +382-525638

e-mail: albanesi@ipvmv1.unipv.it, luca@ipvvis.unipv.it

**Abstract.** This paper describes a new method for 2-D shape recognition based on a multiresolution characterisation of the shape. From the Wavelet coefficients, features are extracted in order to perform a translation-rotation-scaling invariant recognition. Wavelets and multiresolution are exploited in order to reduce complexity of the matching task between the input image and the set of models. In the paper, motivations and performance of the algorithm are presented. Experimental results are also reported in several tests, including noise addition. The approach is quite general, and it could be extended to texture analysis, thus providing a unified paradigm for shape and texture recognition.

## 1 Introduction

Researchers have been using the simplest case of Wavelet, i. e. the Haar transform, since the beginning of this century [1], but a complete framework is more recent [2,3,4,5], in particular for what concerns the construction of Wavelet bases and efficient algorithms for its computation. For this reason, some concepts of Wavelet theory seem to be very old and familiar, but a complete investigation of all the aspects of Wavelet theory may reveal new and interesting aspects. These novelties seem very appealing in several and very different applications: digital signal filtering [6] and analysis [7], image [8], audio [9], and video [10] compression, noise suppression [11], analysis of astronomical images [12], and differential equation systems [13]. At the beginning, the study of Wavelet theory can be misleading, because the same concepts can be considered under three (*very different!*) points of view: (a) in the mathematical framework, as a technique for an expansion of a signal, in terms of orthonormal (or bi-orthogonal) basis; (b) in the theory of pyramidal transformation and coding, as a variant (or alternative) to the classical Burt and Adelson [14] scheme; and (c) in the theory of filter banks, as a simple relationship exists between the Wavelet functions and the transfer function of the filters. Thanks to this variety of points of view, various branches of modern science build up successful Wavelet - based algorithms. However, it is difficult to use the same technique in different disciplines, such as engineering, mathematics and physics. For this reason, very recently some researchers [5] gave a unified view of Wavelet theory, with the aim of suggesting new understanding also in some classical problems of signal processing.

In the last two or three years, Wavelets have become very popular, and new interest is rising on this topic. However, they are very often used in recognition tasks on images (pattern, texture or object localisation and/or recognition). In order to give a possible explanation of this fact, some considerations may be useful. First of all, time - frequency decomposition of Wavelets has some analogies with *classical* transformations (such as Windowed Fourier, Gabor and Wigner); for this reason, Wavelets could be a valid alternative to such transformations which have a dominant role in filtering and analysis of signals, but a less impact in recognition methods. Moreover, the strict relationship between Wavelets and their nearest relatives (subband coding and filter banks) makes more evident the role of Wavelets in some classical problems of coding and filtering (noise suppression, extraction of coefficients to decorrelate data, lossless and lossy compression).

Besides, recognition problem has different aspects which involve not only low-level transformation and/or digital signal processing, but also combines very different tasks: edge extraction and/or smoothing, contour extraction and measurements of geometric features (area, bounding box, etc...). In the following section, this topic is further investigated: some considerations about the usefulness of Wavelets in shape recognition are underlined, and the characteristics of the technique proposed in this paper are reported.

## **2 Wavelets and Shapes: Existing Approaches and the Novelty of this Contribution**

Historically, shape description has a very important role in pattern recognition [15]; moreover, some emerging topics in research have improved interest about this problem, for example, efficient querying in large database. In fact, many content-based indexing and retrieval methods combine shape description and characterisation with other indices (colour, texture) to retrieve the right image from database [16].

In shape recognition, some common choices are the description of the shape with its edge or contour [17] or of its skeleton [18]. Other methods combine these descriptions in more complex tasks, which involve the extraction of feature parameters and the computation of statistical data on the features to drive the matching with the model [19]. In the proposed algorithm, Wavelet transforms are used both in shape description and in matching task. The reasons of this choice can be summarised by the following ideas:

- 1) The main characteristic of Wavelets (if compared to other transformations) is the possibility to give a multiresolution analysis of the image in the form of coefficient matrices. Multiresolution is a powerful tool which enables to adopt the *active vision* paradigm [20,21] in the recognition task. In other words, the matching with the model of the shape is not performed on the entire image, but at different resolutions, thus reducing complexity and erroneous behaviour of the method.

- 2) The Wavelets enable to have an invariant description of the shape at different resolutions, even if the investigated shape undergoes scaling and/or rotation, because the band of details of Wavelets generally characterise different physical structures of shapes.

3) Computational complexity of Wavelets is linear with the number ( $N$ ) of computed coefficients ( $O(N)$ ) while other transformations, also in their fast implementation, lead to a  $N \log_2(N)$  complexity. This is very important because the recognition process is always a complex task, and reducing only a part of it may be meaningful.

4) Wavelets are flexible: several bases exist, and one can imagine to choose the basis which is best suitable to the shape (or class of target shapes). This motivation is very appealing, but it is totally useless if we have no criteria to find the right basis for a given shape. We think that this is still an open problem, and up to now only experimental considerations rule the choice of a Wavelet. However, the choice of an appropriate basis can be useful in the next point.

5) Wavelets are appealing also for dedicated hardware design. If the recognition task has impressive real - time constraints, the possibility of embedding part of the recognition process (i. e., the transformation) in hardware is very interesting. In this solution some advantages, which are peculiar of the Wavelets, can be used for chip design: scalability, symmetry, precision and error tolerance. This approach has been successfully adopted in compression tasks, as it is clearly described in the ADV601 chip [22], from Analog Devices, which embedded in a single chip a video Codec - Decoder based on Wavelet Transform (real - time performance, PAL and NTSC format at 50 - 60 frame/sec).

For all these reasons, the choice of Wavelets can be very promising. In literature it has been already exploited in few cases: for character recognition [23], by exploiting the representation of coefficient matrices for the training of a classification neural network; for 2-D- shapes using Wavelet local maxima in the coefficient expansion in order to detect edges and consequently to discriminate shapes from their contour [17]. In our approach, we adopt different criteria: rather than to address the problem of edge and contour determination, we combine Wavelet expansion with moment calculation in order to extract features from the shapes; our algorithm has therefore the following characteristics:

1) No a priori knowledge of the shape is necessary in any phase of recognition task, in order to transform low-level information of the coefficient matrices in shape representation;

2) Multiresolution analysis in terms of coefficient matrices in different bands and level of Wavelet expansion is used to rule recognition at different level of detail.

3) From the rough Wavelet coefficient matrices, descriptors are extracted to characterised area, orientation and scale at different resolutions.

4) Our representation is invariant to translation; invariance to rotation and scaling is also guaranteed by the algorithm, but user can relax this condition. In this case the number of descriptors increases.

5) The experimental results suggest that this method can be applied also to texture (preliminary tests have been carried out with encouraging results), thus providing a unified approach to shape and texture recognition. In the next sections, the algorithm is explained in details, and experimental results are reported.

### 3 The Proposed Algorithm

In our algorithm, the target image contains one binary instance of the shape to be classified. The general scheme is shown in figure (1). First, the Wavelet transform is applied to the input image: a variable number of decompositions is possible, up to a complete  $\log_2(N)$  levels, for a  $N \times N$  input image  $f(x,y)$ .

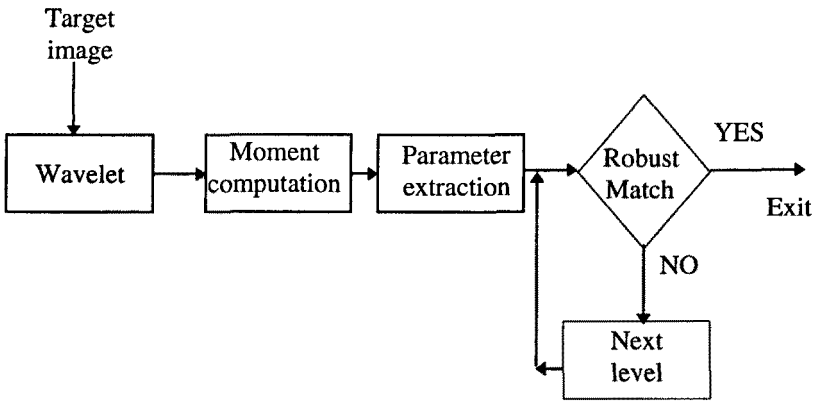


Fig. 1. Flow-chart of the algorithm.

A meaningful choice is to select three levels, as suggested by our experiments, thus giving seven coefficient matrices. For each coefficient matrix, a set of moments is computed, according to the common formula:

$$m_{p,q} = \sum_{j=1}^N \sum_{i=1}^N x^p y^q f(x,y) \quad (1)$$

We have considered up to the second-order moments, and we have extracted a set of two parameters  $F_m^{bk}$  ( $m=1:2$ ) for *each band*  $b$  ( $b = 1:3$ ) and *each level*  $k$  ( $k = 1: \log_2 N$ ) of resolution. As we want a scale or rotation invariant recognition, the parameters are the normalised second-order moments, as expressed in the following formulas:

$$F_1^{bk} = \frac{\tilde{m}_{20}}{m_{00}}$$

$$F_2^{bk} = \frac{\tilde{m}_{02}}{m_{00}} \quad (2)$$

where  $\tilde{m}_{20}$  and  $\tilde{m}_{02}$  are the modified moments computed in the band  $b$  at level  $k$  after a rotation of the coefficient matrices. The rotation assures the condition

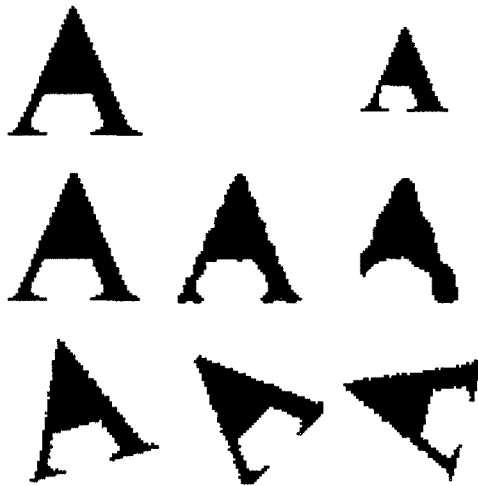
$$m_{11} = 0 \quad (3)$$

If the user wants to relax the invariance to rotation and scaling, other two parameters are available (area and rotation angle). In the matching phase, the set of parameters is compared (using the mean-squared error) with the analogous set of the models. The comparison is performed in multiresolution, starting obviously from the coarser resolution, in order to reduce complexity. Only if the computed distances (errors) from more than one candidate shape are comparable, thus yielding a possible error in recognition, the comparison is performed to the finest resolution.

We have proved the efficacy of this method under different hypotheses (different models, without or with noise added to the input image, with rotation and scaling of the object). In the following section, we describe the set of experiments and results.

## 4 Experimental Results

We have test the algorithm with different classes of shapes. Here we show the result in one case of our test, namely the recognition of shapes of printed characters, even if our method is quite general and it is not limited to character recognition. We have different class of models, corresponding to different fonts (the complete alphabet in Courier, Times and Arial).



**Fig. 2.** An instance of one shape used in the test: in the first row, the model of the shape and an example of its scaled version; in the second row, examples of the model with addition of noise; in the third row, examples of rotation with noise.

We have tested the algorithm with different levels of noise. Noise has been imposed on the image of the model by discarding an increasing percentage (up to 98.4%) of coefficients of the Wavelet Transform (see figure 2), thus simulating the possible effect of a lossy compression algorithm (shape recognition may be possible also on decompressed data of a transform-based compression method) or of the transmission on a very noisy channel. The data refer to the recognition task

performed at the coarser level of resolution, thus better results can be achieved if we iterate the matching at the finer levels. In figure 3 a plot of the error rate in the recognition task, as a function of the increasing noise is shown. The bars in the plot refer to the different Wavelets we use in the experiments. The effect of the noise is very different on the shapes of the characters, and in some cases the added noise is misleading also for a human “recognizer” (see figure 4). As the degradation of the image is a loss of information rather than an addition of noise (in the classical sense), we have use as a quantitative measure of the degradation o the shape the figure:

$$\%_{noise} = \frac{|Area_{model} - Area^{noisy}_{model}|}{A_{model}} \quad (4)$$

where  $A_{model}$  is the area of the shape (in pixels) without noise and  $Area^{noisy}_{model}$  is the area of the same character after degradation. With the figure  $\%_{noise}$  defined in eq. (4) it is possible to compare the recognition of very different characters. The plot of figure 3 refers to different Wavelets: the Haar, the 9/7 bi-orthogonal filters, the cubic-spline case embedding model of the bi-orthogonal human visual system [26], and the bi-orthogonal *binary* case [27] (coefficients are only fraction of power of two). As it can be seen, the choice of Wavelets is quite relevant in the recognition task: for low level of noise, the error rate is zero for 7/9 and binary transforms, which are the best in any case, also for increasing noise. For high level of noise, the error refers only to few specific shapes, namely the “E” and the “T”, which are particularly sensitive to noise (see figure 4).

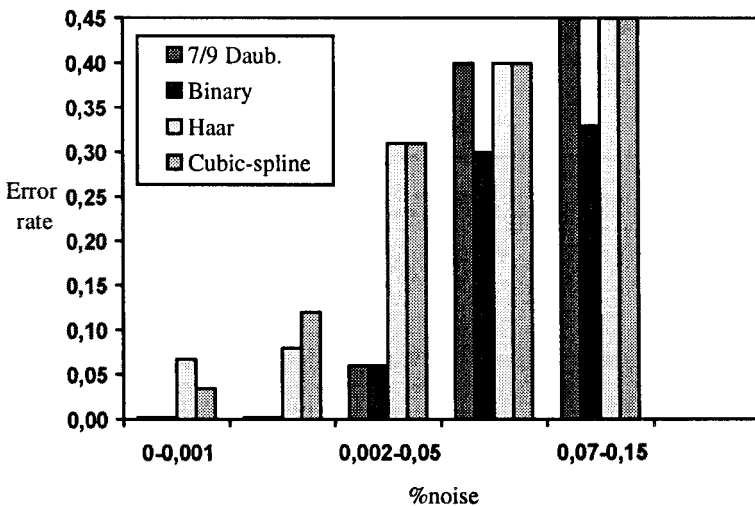


Fig. 3. Error rate in the recognition task, as a function of the increasing noise. (Note: the error rates for 7/9 and binary transforms in the first two classes of %noise are zero)

## 5 Conclusions and Future Works

This paper has proposed a new algorithm for 2-D shape recognition in multiresolution. From the Wavelet coefficients, features of the shapes are extracted in order to compare the input image to the set of models. Wavelets and multiresolution are exploited in order to reduce complexity or to make the process more flexible (choice of the accuracy of the recognition process and of its complexity).

Future works will try to adapt the method to texture analysis and recognition to integrate a unified paradigm of shape/texture recognition in a possible application of querying in image database.



Fig. 4. Examples of worst cases in the recognition task: noise has evidently corrupted the shapes.

## 6 References

1. A. Haar, Zur Theorie der orthogonalen Funktionensysteme, Math. Annal. 69, pp. 331-371, 1910.
2. Daubechies, Orthonormal Bases of Compactly Supported Wavelets, Communications on Pure and Applied Mathematics, **XLI**, pp 909-906, (1988).
3. S. Mallat, A Theory for Multiresolution Signal Decomposition: the Wavelet Representation.
4. Daubechies, The Wavelet Transform, Time-Frequency Localization and Signal Analysis, IEEE Trans. on Information Theory, **36**, N. 5, pp 961-1005, (1990).
5. Vetterli, J. Kovacevic, Wavelets and subband coding, Prentice Hall, 1995.
6. S. Basu, C. Chiang, Wavelets and Perfect Reconstruction Subband Coding with Casual Stable IIR Filters, IEEE Trans. On Circuits and Systems-II: Analog Digital Signal Processing, vol. 42, N. 1, 1995.
7. S. Mallat, Multifrequencies Channel Decompositions of Images and Wavelets Models, IEEE Trans. On Acoustics, Speech and Signal Processing, Vol. 37, N. 12, 1989.
8. J. M. Shapiro, An Embedded hierarchical Image Coder using Zerotrees of Wavelets Coefficients, Proc. of DCC Conference, pp. 214-223, 1993.
9. M. V. Wickerauser, Acoustic signal compression using Walsh-type wave packets, Yale University, August 1989.
10. F. Argenti, B. Benelli, A. Mecocci, Source Coding and Transmission of HDTV Image Compression with the Wavelet Transform, IEEE Journal on Selected Areas in Communications, vol. 11, N. 1, 1993.

11. M. Lang, h. Guo, J. Odegard, C. S. Burrus, R.O. Wells, Nonlinear processing of a shift-invariant DWT for noise reduction, SPIE Conf. on Wavelet Applications, vol. 2491, Orlando, Florida, 1995.
12. F. Rué, A. Bijaoui, A Multiscale Vision Model Applied to Astronomical Images, Vistas in Astronomy, special issue on Vision Modelling and Information Coding, Vol. 40, N. 4, 1996, pp.502.
13. Y. Maday, V. Perrier, C. Ravel, Adaptivité dynamique sur bases d'ondelets pour l'approximation d'équations aux dérivées partielles, C. R. acad. Sci. Paris, 312, Séries I, pp. 45-41, 1991.
14. P. J. Burt, E. Adelson, The Laplacian pyramid as a compact image code, IEEE Trans. Communications, 31(4), pp. 532-540
15. D. h. Ballard, N. C. Brown, Computer Vision, Prentice Hall Englewood Cliffs, NJ. 1982.
16. W. Wu, h. Mo, M. Sakauchi, An Image retrieving System Based on User-specified Recognition model, MVA96, IAPR Works hop on Macine Vision Applications, Nov. 1996, pp. 506-531.
17. Y. Nakamura, T. Yoshida, Learning Two-Dimensional Shapes using Wavelets Local Extrema, Proc. of 12<sup>th</sup> IAPR International Conference on Pattern Recognition, Volume III, pp. 48-52, 1994
18. C. Arcelli, A. Ramella, Sketching a gray tone pattern from its distance transform, Pattern Recognition, Vol. 29, N. 12, 1996, pp. 2033-2045.
19. R. Leagult, C. Y. Suen, Contour tracing and parametric approximations of digitized patterns, A. Kryzak, T. Kasvand, C. Y. Seun eds., pp. 225-240, World Scientific Publishing, Singapore, 1989.
20. AA. VV. Dialgue., *CVGIM*, pp. 65-118, 1994.
21. J. Alomonos, I Weiss, A. Bandyopadyyay, Active Vision, Proc. First IEEE Int. Conference on Computer Vision, 35-54, London 1987.
22. *ADV601p,Pm* data sheet from Analog Devices.
23. W.Lee, C. H. Kim, H. MA, Y.Y.Tang, Multiresolution recognition of unconstrained handwritten numerals with Wavelet Transform and Multilayer Cluster Neural Network, Pattern recognition, Vol. 29, N. 12, pp. 1953-1961, 1996.
24. Cohen, I Daubechies, J. C. Feauveau, Biorthogonal bases of compactly supported Wavelets, Communications on Pure and Applied Mathematics, XLI, pp 485-560, (1992).
25. A. Watson, G. Yang, J. Solomon, J. Villasenor, Visual Thresholds for Wavelet Quantization Error, SPIE Porc. Vol. 2657, H uman Vision and Electronic Imaging, The Society for Imaging Science and Technology, 1996.
26. S. Bertoluzza, M. G. Albanesi, On the coupling of the Human Visual system and Wavelet transform for Image Compression, Proceedings of SPIE Mathematical Imaging: Wavelet Applications in Signal and Image Processing II, pp. 389-397, 1994.
27. H. Chao, P. Fisher, An Approach to Fast Integer Reversible Wavelet Transform for Image Compression, downloaded at <http://www.mathsoft.com/wavelets.html>