

Automated, transformation invariant, shape recognition through wavelet multiresolution

Brault P. and Mounier H.

Département AXIS, Institut d'Électronique Fondamentale,
Bâtiment 220, Université Paris-Sud, 91405 Orsay, FRANCE.

ABSTRACT

The purpose of this study is to demonstrate the efficiency of Wavelet Multi-Resolution Analysis (W-MRA) applied to automated shape recognition in an automatic vehicle driving scheme. Different types of shapes have to be recognized in this framework. They pertain to most of the objects entering the sensors field of a car. These objects can be road signs, lane separation lines, moving or static obstacles, other automotive vehicles, or visual beacons. The recognition process must be invariant to global, affine or not, transformations which are : rotation, translation, scaling and perspective (wide angle camera lenses). It also has to be invariant to more local, elastic, deformations due to environmental conditions (weather : rain, mist, light reverberation) as well as optical or electrical noise.

To demonstrate our method, an initial shape, with a known contour, is compared to the same contour altered by rotation, translation, scaling and perspective, and is also compared to completely different shape contours. The curvature computed for each contour point is used as a main criterion in the shape matching process. The original part of this work is to use wavelet descriptors, generated with a fast orthonormal W-MRA, rather than Fourier descriptors, in order to provide a multi-resolution description of the contour to be analyzed. In such way, the intrinsic spatial localization property of wavelet descriptors can be used and the recognition process can be speeded up.

Keywords: Shape Recognition, Wavelet Multi-Resolution, contour matching with curvature criterion, dynamic programming, transformation invariance, recognition decision parameter

1. INTRODUCTION

Briefly summarized, the purpose of this work is to realize the matching of two shapes by using a specific simplification criterion for these shapes. Small, local and elastic, deformations, as well as global deformations like rotations, translation, scaling and perspective, are taken into account in this framework. A simplification process is, in general, at the basis of every image analysis, and in particular, at the basis of most pattern recognition techniques. Moreover, we want to quantify the matching constraint and use it as a decision threshold in the recognition process.

In pattern recognition and image analysis, we often need to simplify the structure of the incoming signal in order to carry out, in a faster and more efficient way, any automatic process on the information . In our case, this simplification takes place in a 2-to-1D dimensionality reduction which consists in working on the image contour rather than with the full original image. Once the characterization criterion or "image-cue" is chosen, among several types of "vision cues" (color, move, mass) or "shape cues" (contour), we have to use a relevant descriptor for this criterion. Relevant means that, for a pattern recognition scheme based on the contour, the descriptor used for the 1D-representation of our object will exhibit invariance by the transformations (rotation, translation a.s.o.) applied to the objects.

The possible method for pattern recognition consists in performing a W-MRA of an initial image and then try to detect in it the wavelet decompositions of a database of shapes. Another method could be to compute the correlation between the image Wavelet Decomposition (WD) and the objects WD. This naive approach gives rise to several problems. The action of rotations and dilations on the objects of interest completely modifies the WD of

Further author information: (Send correspondence to Brault.P.)

Brault P.: E-mail: patrice.brault@ief.u-psud.fr

Mounier H.: E-mail: hugues.mounier@ief.u-psud.fr

these objects. But translation variance is probably the first perturbing cause in a discrete wavelet transform. A signal translated by a factor Δb on the position (or time) axis will exhibit totally different wavelet coefficients with respect to its initial position.

Translation, or shift, invariance has given rise to numerous works. A first solution is to remove sub-sampling after filtering (see for example the "orthonormal shell expansion"¹⁷ and the "a trous" algorithm^{19,1,14}). In this case, the approximation and wavelet sub-spaces contain as many samples as the original image and a translated image exhibits the same wavelet decomposition.

Another more general solution applicable to linear transformations has been proposed by.^{8,20} They consider that the total coefficients energy is constant for a fixed level of decomposition, whatever the translation of the signal can be.

On another hand, and in order to reduce the computation complexity, most pattern recognition and image analysis methods operate first by simplifying the image to be analyzed. Generally, the image is analyzed as a whole or by regions of interest, or objects. By selecting a characterization criterion (ring-projection or mass, color, contour, move, a.s.o). and a descriptor (Fourier, wavelets, histogram), one is able to extract a specific and relevant information. Then, recognition, comparison, compression, filtering and other processes can be performed on the image characterized by this information.

The choice of a criterion and of a descriptor serves two main purposes : a reduction of the image complexity, e.g. 2D-1D, by extracting a relevant information, and the ability for some descriptors to preserve this information and, in particular, linear or not transformations. For our case, these transformations are translation, rotation, scale change and perspective (caused by changes in camera focal length).

2. CONTOUR MATCHING WITH WAVELET MULTI-RESOLUTION

In our framework, the fusion of several "image-cues" like vision-cues (move, colour, texture, contrast) and shape-cues (contour) is highly considered for target, road sign or obstacle recognition. With this paper, we start with a recognition made by matching the contours of an object with a database. Existing matching algorithms operate on the image as a whole. This approach is not considered accurate enough here and we prefer to perform contour matching on a detailed point-to-point basis. Hence a small deformation (e.g. a person move) can be detected.

As a starting point, we will consider that both objects to be compared are already segmented in a closed contour. We invite here the reader to refer to the abundant literature on contour detection and in particular to Canny's detector.

The contours are then coded by the matrix of their cartesian coordinates in the image projection plan. Two interesting properties of these coordinates, are that they simplify the curvature computation, and that they provide, when referenced to the contour centroid, the translation invariance. To obtain these cartesian coordinates, the contour is first reduced to one pixel width by using an algorithm based on a minimal direction change. The contour origins are then matched (normalized) by correlation in order to get an optimal fit and match the homologous points between both contours. In such way, these contours will be described by starting with homologous starting points. The directions of description are matched by the same process. The contour dimension is interpolated on 2^n points to make possible the use of a fast wavelet multi-resolution decomposition.

Following,²⁴ the multi-resolution matching possibilities are explored by using a similitude criterion based on the contour curvature.

We then perform a multi-resolution decomposition on the contours and compute the curvature for each contour series $x(n)$ and $y(n)$ at each level of the wavelet decomposition. Yet we do not use Fourier descriptors but wavelet ones. Wavelet descriptors exhibit interesting properties with regard to a multi-resolution approach. They are ... natural, fast computed $O(A \times N)$, where A stands for the filter length, at the maximum highest resolution, with regard to FFT $O(N \log_2(N))$, and offer a localization property. One of the major drawbacks of the Fourier transform is its inadequacy to spatial localization : the transform of a Dirac perturbs the whole frequency spectrum. Then the Dirac localization information cannot be coded in its transform. In our case the singularity in the series $(x(n), y(n))$ of an initial contour perturbs the Fourier transform for any resolution of this contour. Here stands the major difference

with the wavelet transform which easily locates a high frequency pattern.

The wavelet approach can be used in several ways. High magnitude coefficients, which are characteristic of high curvature points serve as position-mark for the matching and reconstruction process. Other methods have been explored and can also be used in this framework. The "M largest coefficients" method, proposed in,⁹ offers a faster and better progressive shape reconstruction, than the usual scale-by-scale method, by using in priority for the reconstruction the maximum amplitude coefficients computed in the complete multi-level decomposition. Another approach using the dyadic (continuous-discrete) transform of Mallat^{13,19} gives a direct access to the contour singularities which are representative of the curvature. Then, object shapes can be discriminated by some characteristic features extracted at high curvature points.¹⁶

In our process, the matching constraint is transmitted from coarse to fine resolution. The matching process itself is done by dynamic programming. One advantage of dynamic programming is its easy implementation. We start with the coarsest resolution without any constraint. The cost function used in the algorithm is the distance between the curvature of homologous points. A supplementary cost for horizontal and vertical transitions can also be used. The result of the first matching provides couples of matched points. The matching at a given level is constrained by keeping a minimum distance with the matching at the immediate coarser resolution. In fact, each step consists in adding slight deformations to a similar global shape.

The $x(n)$ and $y(n)$ signals are reconstructed to the next finer resolution by applying an inverse wavelet transform. The operations are pursued in the same way up to the finest resolution.

The following section shows details of the complete algorithm.

3. DETAILED ALGORITHM

We want to match two objects located in two different images I_1 and I_2 , (See Fig. 1 and 2).

a) Contour Cartesian coding

These objects, called S_1 and S_2 , are first detected and segmented. We do not deal, here, with the segmentation process and make reference to the abundant literature on the subject [e.g. Canny detector]. So we assume that the objects have already been segmented in two contours C_1 and C_2 . An appropriate efficient coding of their contour must be performed. Here, the MDC (minimal direction change) is used as a criterion for selecting the most pertinent pixels of the original segmentation. This results in a one-pixel contour coded by the cartesian coordinates of the pixel series.

b) Relative coordinates

The two contour series are expressed in coordinates relative to the center of the contour.

c) Interpolation to 2^n

The series of contour points are interpolated to a power of 2 in order to perform a wavelet multi-resolution analysis.

d) Contour normalization

Homologous points of both contours then have to be put in correspondence by correlation. This results in a one to one mapping of the contours with correspondence of their starting points.

e) Contour orientation

We re-iterate the correlation process with successive inversions of the x, the y and both x and y coordinates in order to detect an OX, OY or O symmetry between contours C_1 and C_2 . In the practical case of car recognition, we will only consider the Oy symmetry which is the most realistic one. We then find the best correlation between direct (c) case) and symmetric (d) case) orientations.

f) Curvature

The curvature is computed locally for each point $C(n)$ (with coordinates $x(n)$ and $y(n)$) of both contours, by using

an approximation on the first and second derivatives.

$$\kappa(n) = \frac{\dot{x}(n)\ddot{y}(n) - \ddot{x}(n)\dot{y}(n)}{[\dot{x}(n)^2 + \dot{y}(n)^2]^{\frac{3}{2}}} \quad (1)$$

The curvature criterion is based on the use of a distance between two homologous points $C_1(n_1)$ and $C_2(n_2)$ of contours C_1 and C_2 . This distance, based on the local curvature in C_1 and C_2 , is defined as :

$$d(C_1(n_1), C_2(n_2)) = |\kappa_1(n_1) - \kappa_2(n_2)| \quad (2)$$

The possibility to match contours in the case of local elastic deformation, implies that the correspondance between contour C_1 and contour C_2 must not be bijective. Then, to one point of contour C_1 can coincide several points of contour C_2 . Though this property does not seem to be of major importance in the semi-rigid framework of mobile robotics, we can presuppose a better efficiency of the elastic matching in the case of perspective deformation and in even other cases.

g) Multi-Resolution Wavelet Decomposition.

The multi-resolution is performed on a periodized version of the contour $C(x(n), y(n))$. More precisely, we perform a pyramidal one-dimensional discrete W-MRA on each signal $x(n)$ and $y(n)$, with an assumption of signal periodicity to avoid border effects. The multi-resolution wavelet decomposition is now a classical scheme, in which the action of a linear projector, a convolution product, on a function f , provides its decomposition on the basis of scale functions $\varphi_{j,m}(x)$, and j and m stand for the scale and position parameters. The coefficients of this decomposition, the A_m^j , constitute the approximation, or "coarse view", at scale j of the initial function f . This W-MRA decomposition is completed by the projection of f on the basis of wavelet functions $\psi_{j,m}(x)$ which define the detail coefficients and provide, for each scale, detail subspaces W_j that are orthogonal complement of their homologous approximation subspaces V_j at scale $j-1$.

$\varphi(x)$ defines the "cascade" of approximation closed subspaces $\{V_j\}_{j \in \mathbb{Z}}$, which will represent the multiple resolutions of the contour C . The projection of contour $C(x,y)$ on the basis of scale functions $\varphi_{j,n}$, and for all scales, can be written :

$$A[C(x, y)] = \sum_j \sum_n \underbrace{\langle C(x, y), \varphi_{j,n} \rangle}_{a_j^n} \varphi_{j,n} \quad (3)$$

where the a_j^n are the approximation coefficients.

Our purpose is to compute coarser versions of the same contour, so we will consider here only the projections of both series $C_1(n_1)$ and $C_2(n_2)$ on the scale function $\varphi_{j,n}(x)$, and not on the wavelet function $\psi_{j,m}$. The projection are made with separable scaling functions along the x and y directions :

$$C_1 \left\{ \begin{array}{l} A[x_1(n_1)] = \sum_j \sum_{n_1} \underbrace{\langle x_1(n_1), \varphi_{j,n_1} \rangle}_{a_j^{n_1}(x_1)} \varphi_{j,n_1} \\ A[y_1(n_1)] = \sum_j \sum_{n_1} \underbrace{\langle y_1(n_1), \varphi_{j,n_1} \rangle}_{a_j^{n_1}(y_1)} \varphi_{j,n_1} \end{array} \right. C_2 \left\{ \begin{array}{l} A[x_2(n_2)] = \sum_j \sum_{n_2} \underbrace{\langle x_2(n_2), \varphi_{j,n_2} \rangle}_{a_j^{n_2}(x_2)} \varphi_{j,n_2} \\ A[y_2(n_2)] = \sum_j \sum_{n_2} \underbrace{\langle y_2(n_2), \varphi_{j,n_2} \rangle}_{a_j^{n_2}(y_2)} \varphi_{j,n_2} \end{array} \right. \quad (4)$$

The difference with a Fourier transform is that each approximation coefficient owns, with the parameter n , a localization information on the local contour regularity, the amplitude, or "strength", of the coefficient.

The decomposition depth used for this W-MRA have also to be determined. This parameter will be discussed later on.

h) Contour matching

Dynamic programming is used for contour matching. The algorithm starts at the lowest resolution, with a cost which is the difference between curvatures for each point of the contours. A matching constraint is to stay close from the matching at former resolution, so we start with no constraint at the first, low resolution, level. A supplementary cost can be provided in the vertical and horizontal transitions.

i) Inverse WT

We start the matching process at the coarsest resolution. Each step towards the maximum resolution is then reconstructed from the former step by an inverse wavelet transform.

4. TEST EXAMPLE

To demonstrate the ability of shape recognition, we first have performed tests on two different car shapes (see figs. 1 and 2). Car contours are drawn with a bitmap editor on an image background of 256x256 pixels. The contours, interpolated by the algorithm, have a length of 512 points. Then the matching process between both contours is started. It is clear, on figs. 7 and 8, that car B is mapped on the top and car A on the bottom.

The result of the multiresolution matching (figs. 5,6,7,8) is an interesting correspondence of similar parts of the cars (front and rear bumpers, wheels). Perceptible defaults in the matching process (see the test example in⁴) have been solved by using images with a higher resolution of 256x256 points.

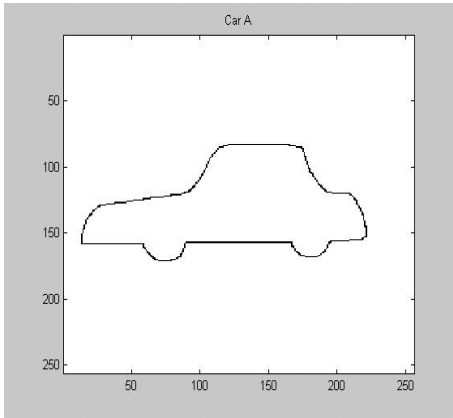


Figure 1. Contour of the "A" car.

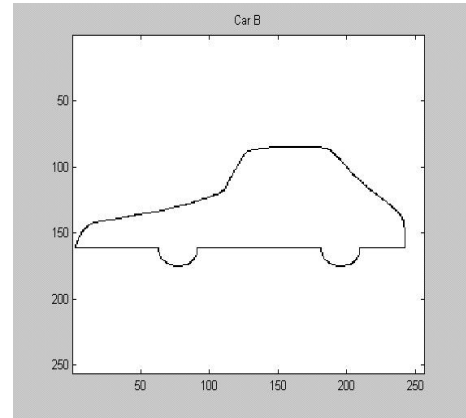


Figure 2. Contour of the "B" car.

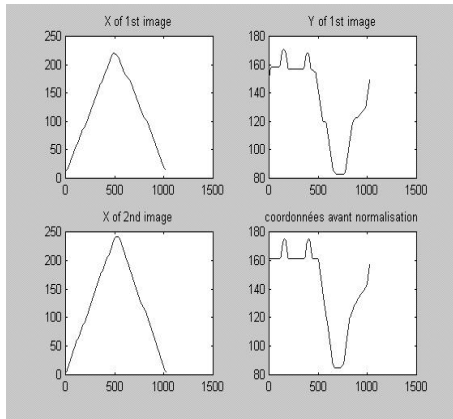


Figure 3. X and Y Contour coordinates series.

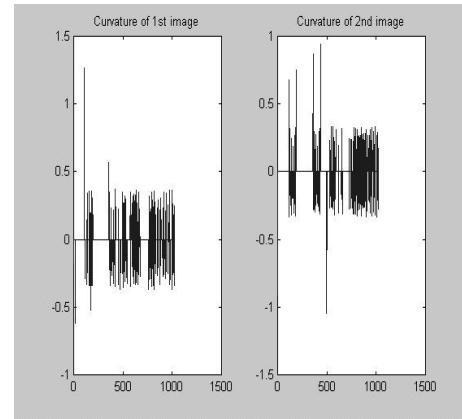


Figure 4. Curvature value versus contour position.

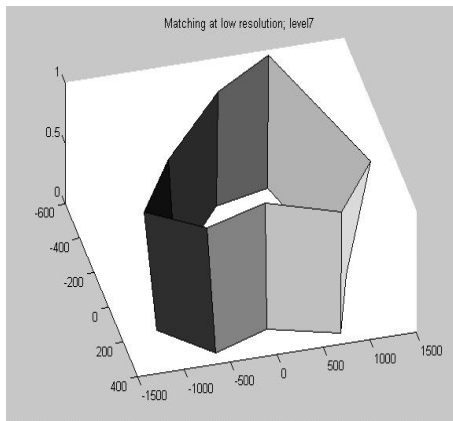


Figure 5. Contour matching at the coarsest resolution (level 7)

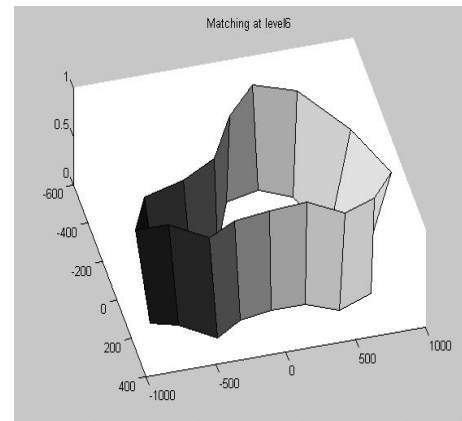


Figure 6. Contour matching at level 6

5. INVARIANCE TO TRANSFORMATIONS

5.1. Translation

The referencing of coordinates to the object centroid provides invariance to translation.

5.2. Rotation

The choice of curvature as matching parameter for shape recognition provides invariance to rotation. We have performed tests with one of the contours rotated more than 10 degrees. Above 50° the matching exhibits wrong couples.

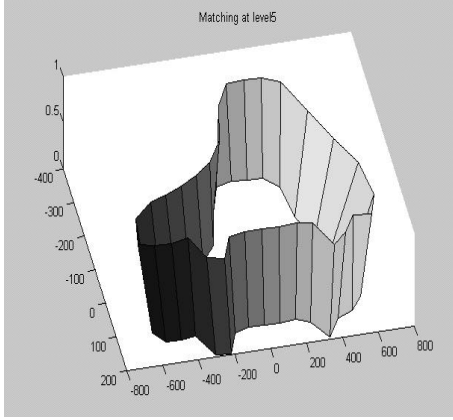


Figure 7. Contour matching at level 5

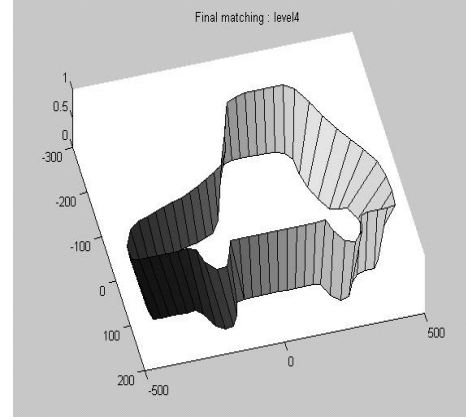


Figure 8. Contour matching at level 4

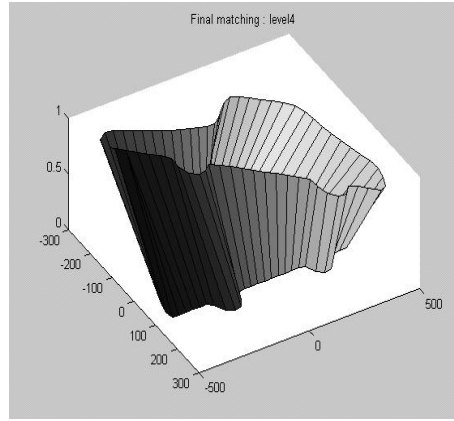


Figure 9. Matching with car B rotated 30 degrees

5.3. Scaling

Scaling does not affect the matching process if both objects are still defined by the same number of points.

5.4. Perspective and other small elastic deformations

Elastic deformations like perspective or other very local deformations due to noise or weather conditions, do not affect the matching process which is mainly based on high curvature points. The recognition parameter computed in the following section (6) will only be affected by such deformation. In the case of perspective deformation, one idea is to compute the recognition parameter on a part of the shape with a tight constraint, and on the rest of it with a weak constraint. Nevertheless, the curvature value along the shape is in general supposedly not very much altered by the elastic deformations. So the recognition constraint should only be slightly relaxed in this case.

6. MATCHING DECISION

The curvature, used as a cost function in the matching process by dynamic programming, is used again as a recognition parameter in the decision process. We compute a final recognition parameter " r_F " by summing the absolute value of the difference between the curvatures for each point of the A and B contours (see eq. 5 below). The decision for

shape matching then takes place by comparing the r_F parameter with a maximum threshold r_{Max} . When the shapes are identical, $r_F = 0$.

$$r_F = \sum_{n=1}^N |\kappa_1(n) - \kappa_2(n)| \quad (5)$$

A second scheme is to compute a global decision parameter " r_G ". This one is computed over a range of scales " j ", from the low to the highest resolution, in order to verify shape matching for all scales. We then write :

$$r_G = \sum_j \sum_{n=1}^N |\kappa_1(n) - \kappa_2(n)| \quad (6)$$

The third scheme is a point-by-point, individual, recognition parameter r_I which behaves as a maximum threshold for the curvature distance computed at each contour point and over all scales.

$$|\kappa_1(n) - \kappa_2(n)| < r_I, \quad \forall(j, n) \quad (7)$$

In presence of noise, the maximum level of resolution in the W-MRA has to be limited in order to prevent the curvature, i.e. the recognition process, to be perturbed by noise.

Another scheme is an iterative recognition process. The recognition at a highest resolution takes place only if the recognition is made at the former resolution, thus enabling an "accelerated" matching rejection. On the contrary, the recognition can be made even with unmatching at low resolution, but with correct matching at high resolution if the decision criterions are rather "high frequency" ones.

Experimental results : We give here results for the recognition parameter r_F . The image definition has been increased to 512x512, for low resolution contours were not able to provide accurate enough curvature values. In this case the contour length is 1024 points.

The r_F parameter has been computed between the shape of a "C" car and locally deformed versions of this shape. It has also been computed between the C car shape and a totally different shape, a rectangle . The values obtained below show that this method is relevant for shape recognition :
For a very slightly different car shape, $r_F = 0,0101$. For a car shape modified on the front and the rear part (Fig. 10), $r_F = 0,0557$. For a rectangle, $r_F = 0,0774$. The choice of a maximum threshold r_{Max} can now be done on the basis of this first "learning" and could be taken between 0,5 and 0,6.

7. CONCLUSION

Shape recognition based on a 1D characterization criterion, the contour curvature, has been investigated by means of multiresolution wavelet descriptors. Robustness to local transformations is given by a coarse to fine wavelet multiresolution matching scheme. Robustness to global transformations is given : for the symetries by an initial correlation between contours, for the rotation by the use of a 1D curvature criterion and for the translation by the use of cartesian coordinates referenced to the object centroid. We have also demonstrated the possibility of an automated shape recognition by the computation of recognition decision parameters.

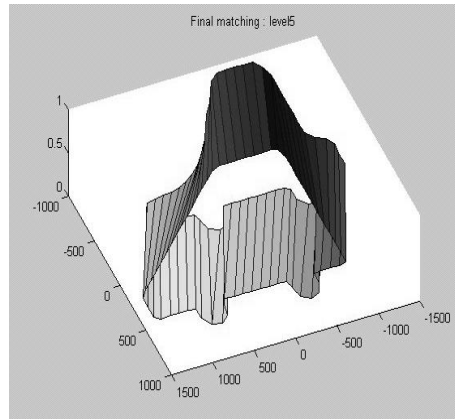


Figure 10. Recognition decision between the C car and a modified version : $r_F = 0,0557$

REFERENCES

1. P. Abry, E. Chassande-Mottin and P. Flandrin, "Algorithmes rapides pour la décomposition en ondelettes continue - Application à l'implantation de la réallocation du scalogramme". *Proceedings of the GRETSI*, 1995.
2. J.P. Antoine, D. Barache, R.M. Cesar Jr and L. da F. Costa, "Multiscale Shape Analysis using the Continuous Wavelet Transform", *Proceedings of the IEEE International Conference on Image Processing (ICIP-96)*, Lausanne, Delogne ed., Vol. 1, pp. 291-294, 1996.
3. R. Basri, L. Costa, D. Geiger and D. Jacobs, "Determining the similarity of deformable shapes". *IEEE Workshop on Physics-based modeling in computer vision*, p.135-143, 1995.
4. P. Brault and H. Mounier, "Wavelet Multi-Resolution Transform Applied to Shape Recognition Based on a Curvature Criterion", *ICISP01 Int'l Conference on Image and Signal Processing, Agadir*, Vol. 1, pp. 250-258, 2001.
5. T.D. Bui and G.Y. Chen, "Multiresolution Moment-Wavelet Fourier Descriptor for 2D Pattern Recognition", *Proceedings of the SPIE*, vol. 3078, pp. 552-557, 1997.
6. T.D. Bui and G.Y. Chen, "Invariant Fourier-wavelet descriptor for pattern recognition". *Pattern recognition* Vol.32, 1083, 1999.
7. G.W. Cook and E.J. Delp, "Multiresolution sequential edge linking". Purdue University. *IEEE Washington Int'l Conference on Image Processing*, Oct 1995.
8. W.T. Freeman and E.H. Adelson, "The design and use of steerable filters". *IEEE Pattern Anal. Machine Intell.*, vol 13, pp. 891-906, 1991.
9. D. Gatica-Perez and F. Garcia-Ulgade, "Compact representation of planar curves based on a wavelet shape descriptor for multimedia applications", *Vision interface 99*, Trois-Rivieres, Canada, 1999.
10. D. Geiger, A. Gupta, L.A. Costa and J. Vlontzos, "Dynamic programming for detecting, tracking and matching deformable contours". *IEEE PAMI* 17(3), march 1995.
11. A. K. Jain, Y. Zhong, S. Lakshmanan, "Object matching using deformable templates". *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 18, no. 3, pp. 267-278, 1996.
12. T-L. Liu, D. Geiger and R.V. Kohn, "Representation and self-similarity of shapes". *Int'l conference on computer vision (ICCV)*, Bombay, 1998.
13. S. Mallat and S. Zhong, "Characterization of signals from multiscale edges". *IEEE transactions on pattern analysis and machine intelligence*, Vol. 14, N°7, July 1992.
14. S. Mallat, *A wavelet tour of signal processing*, Academic Press, 1998.
15. S. Pittner and S.V. Kamarthi, "Feature extraction from wavelet coefficients for pattern recognition tasks". *IEEE PAMI*, Vol. 21, N°1, Jan. 1999.
16. A. Quddus, F. A. Cheikh and M. Gabbouj, "Content-based Object Retrieval Using Maximum Curvature Points In Contour Images", *Proceedings of SPIE: Storage and Retrieval for Media Databases 2000*, volume 3972, M. M. Yeung and B.-L. Yeo and C. A. Bouman Ed., p. 98-105, 2000.

17. N. Saito and G. Beyklin, "Multiresolution representations using the auto-correlation functions of compactly supported wavelets", *IEEE Trans. Inform. Theory*, Jan. 1992.
18. B. Serra and M. Berthod, "Subpixel contour matching using continuous dynamic programming". *Proc. IEEE Comput. Soc. Conf. Comput. Vision and Pattern Recogn.*, pp 202-207, 1994.
19. M.J. Shensa, "The Discrete Wavelet Transform: Wedding the "à trous" and Mallat algorithms". *IEEE Trans. on Signal Proc.*, vol. 40 n°10, October 1992.
20. E.P. Simoncelli, W.T. Freeman, E.H. Adelson and D.J. Heeger, "Shiftable Multiscale Transforms" . *IEEE Trans. Inform. Theory*, vol 38, pp. 587-607, Mar. 1992.
21. Y.Y. Tang, Ma Hong, Liu Jiming and Xi Dihua, "Multiresolution analysis in extraction of reference lines from documents with gray level background", *IEEE PAMI*, vol 19, N°8, Aug. 1997.
22. Q.M. Tieng and W.W. Boles, "Wavelet-based affine invariant representation : a tool for recognizing planar objects in 3D space", *IEEE PAMI*, Vol.19, N°8, Aug. 1997.
23. Q.M. Tieng and W.W. Boles, "Recognition of 2D object contours using the wavelet transform zero-crossing representation", *IEEE PAMI*, Vol.19, N°8, Aug. 1997.
24. A. Vapillon, B. Collin and A. Montanvert, "Appariement de contours 2D par analyse multirésolution hiérarchique de la déformation", *proceedings of GRETSI'97*, Grenoble, France, Sept. 1997.
25. S.C. Zhu and A.L. Yuille, "FORMS : a flexible object recognition and modeling system", *Int'l journal of computer vision*, Vol 20, N°3, dec. 1996.