# WAVELET TRANSFORM FOR PARTIAL SHAPE RECOGNITION USING SUB-MATRIX MATCHING

El-hadi Zahzah

*Laboratoire de Mathématiques Appliquées, Avenue M Crépeau La Rochelle 17042, France*

Abstract:     In this paper, we propose a method for 2D partial shape recognition under affine transform using the discrete dyadic wavelet transform invariant to translation well known as *Stationary Wavelet Transform or SWT*. The method we propose here is about partial shape matching and is based firstly on contour representation using the wavelet transform. A technique of sub matrix matching is then used to match partial shapes. The representation is based on three steps, the contour is first parameterized by enclosed area, the affine invariant feature is then calculated to finally determine the natural axis which enable to fix the starting point. The knowledge of the orientation of the natural axis enables to adjust the starting point on the contour between the query and the models in a given database. Furthermore, the method can selects a subset of useful invariant features for the matching step. A sub-matrix matching algorithm developed by (Saber et al., 2005)is then used to determine correspondences for evaluation of partial similarity between an example template and a candidate object region. The method is tested on a database of 5000 fish species, and the results are very satisfactory.

## 1 INTRODUCTION

Object recognition is a main problem in computer vision. Literature is abundant in this field. These works belong to two main categories: methods based on contour and methods based on region. The criterion used to classify a method to one category or another, is to see if the descriptor is calculated on the contour or on the region. For a good overview of the various representation, description and recognition techniques see (Zhang and Lu, 2004). Although methods based on region seem to be more general comparing to methods based on contour, in many applications they require more data and are more time consuming. Our work limits to the first case, taking into account only the contour object. In (Mallat, 1989), the author describes a mathematical model to calculate and interpret the concept of *multi-resolution (multi-scale)* representation. Mallat shows that information can be extracted from two successive resolutions and he then defines a new and complete representation well known as *wavelet representation*. This representation is widely used in computer vision and signal processing. In (Mallat, 1991), the author studies also the completeness, the stability and the application to the recognition of the model based on *zero-crossings*

representation. The discrete wavelet transform invariant to translation introduced by Mallat in (Mallat, 1991), called later in (Misiti et al., 2003), *Stationary Wavelet Transform (SWT)* which is the non decimated version of the classical *dyadic discrete wavelet transform* (*DWT*). The *SWT* is often used for 2D shape recognition. Tieng et al. use this transform to deduce the contour representation by the *zero-crossings* in (Tieng and Boles, 1997a) and by the *extremums* in (Tieng and Boles, 1997b), but these methods need post processing to remove the false zero-crossings or false extremum. In (Kimcheng and El-hadi, 2004), we proposed another approach for 2D shape recognition under affine transform. The method is based on the parameterization by the *enclosed area*. We develop a technic to align the starting point between the model of the database and the query. The general scheme of this process is illustrated in the figure (1). The different steps are detailed in the following sections. In this paper we add to our representation a sub-matrix matching algorithm developed by E. Saber et al in (Saber et al., 2005). This algorithm is proposed to determine correspondances for evaluation of partial similarity between an example template and a candidate object region. The method is translation,
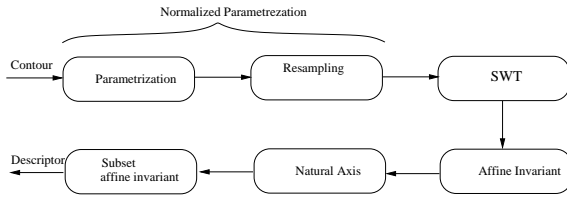
Figure 1: General scheme of the different stages for the descriptor construction.

rotation, scale and reflection invariant. Applications of the proposed partial matching technique include recognition of partially occluded objects in images.

## 2 PARAMETERIZATION

2D shape recognition under affine transform uses generally two types of parameters as detailed in (Khalil and Bayoumi, 2001), (Khalil and Bayoumi, 2002) and (Tieng and Boles, 1997b): The parameterization by the *affine length* and by the *enclosed area*. The first type needs the derivative of high order and then are very sensitive to quantization and to noise. We adopt the second type of parameterization,using the enclosed area defined by:

$$s(t) = \frac{1}{2} \int_a^t |x(u)y'(u) - x'(u)y(u)| du \qquad (1)$$

in the discrete case this formula may rewritten as:

$$s(k) = \begin{cases} 0 \text{ if } k=1, \\ \frac{1}{2} \sum_{t=1}^{k-1} |x(t)y(t+1) - x(t+1)y(t)| \\ \text{otherwise.} \end{cases} \qquad (2)$$

this parameter is based on the affine transform property which assess that is *"under affine transform, the objects areas change in the same proportion"*.

## 3 THE AFFINE INVARIANT

In this section, we show how to obtain the affine invariant from the coefficients of the *SWT*. Others methods to obtain the affine invariant can also be use as those proposed as in (Tieng and Boles, 1997b) and (Khalil and Bayoumi, 2001). Let's assume that the contour $\tilde{\Gamma}(\tilde{x}_i, \tilde{y}_i)$ is the affine transform of a given contour $\Gamma(x_i, y_i)$, ie.

$$\tilde{x}_i = a_1 x_i + a_2 y_i + t_1 \ ; \ \tilde{y}_i = b_1 x_i + b_2 y_i + t_2 \qquad (3)$$

where $1 \le i \le N$, $a_j$, $b_j$, $1 \le j \le 2$ are the coefficients of the affine transform with the determinant

of the affine matrix transform $a_1 b_2 - a_2 b_1$ is non zero and $t_1$, $t_2$ represent respectively the translation in $x$ and $y$ axis of the contour. As the origin is previously shifted to the centroid, we get $t_1 = t_2 = 0$. Consequently, if the *SWT* is applied for both previous equations one obtains

$$\begin{pmatrix} A^l \tilde{x}_i & W^l A^l \tilde{x}_i \\ A^l \tilde{y}_i & W^l \tilde{y}_i \end{pmatrix} = \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} \cdot \begin{pmatrix} A^l x_i & W^l x_i \\ A^l y_i & W^l y_i \end{pmatrix} \qquad (4)$$

where $A^l x_i$ and $W^l x_i$ for $1 \le i \le N$ represent respectively the *ith* approximation and the detail coefficient of the decomposition of $x = (x_i)_{1 \le i \le N}$ by *SWT* at level $l$. If we note

$$M^l(i) = A^l x_i W^l y_i - A^l y_i W^l x_i \qquad (5)$$

we obtain

$$\tilde{M}^l(i) = (a_1 b_2 - a_2 b_1) M^l(i) \qquad (6)$$

then $M$ is a relative invariant. $M^l(i)$ is normalized by $M^k(j)$ to obtain the absolute invariant $I$.

$$I^l(i) = M^l(i)/M^k(j) \qquad (7)$$

To reduce the noise effect on the invariant $I$, the level $k$ is selected if the magnitude sum (in absolute value) of equation (5) is the greatest value. The value of $j$ corresponds to the position where the magnitude (in absolute value) of equation (5) is maximale as in (Tieng and Boles, 1997b).

$$k = \underset{1 \le l \le n}{argmax} \sum_{j=1}^N [M^l(j)]^2 \ ; \ j = \underset{1 \le i \le N}{argmax} |M^k(i)|$$

## 4 THE NATURAL AXIS

In section 3, One assumes that the starting point on the contour $\tilde{\Gamma}$ and $\Gamma$ are the same. In practice, this assumption is not true since one cannot has $\tilde{I}^l(i) = I^l(i)$. However, if the contour $\tilde{\Gamma}$ is the affine transform of the contour $\Gamma$, then, one can shows that exist an integer number $\tau$ such as :

$$\tilde{I}^l(i + \tau) = I^l(i), \quad 1 \le i \le N \qquad (8)$$

Actually, $\tau$ is the required shift to adjust the starting point of the contour $\tilde{\Gamma}$ and $\Gamma$. In practice we cannot know *a priori* the value of $\tau$, because we do not know whether yes or not the query is an affine transform of the model, and how to match points. The value of $\tau$ can be estimated by computing the correlation between $I$ et $\tilde{I}$, ie.:

$$\tau = \underset{1 \le j \le N}{argmax} \sum_{1=1}^N \tilde{I}^l(i+j) I^l(i)) \qquad (9)$$

However, this calculus is relatively heavy. Another way to obtain this value is to use the *Natural Axis* as in (Shen et al., 1999). The technic principle in a general framework is as follow: One starts from an ordered set $S = \{s_i, \ 1 \leq i \leq N\}$ (in our case $S$ is the set of values $I^l(i)$), one associates to $S$ the necklace of radius one and $N$ pearls. To each pearl correspond a weight $s_i$. The distance between two successive pearls is equal to $2\pi/N$. The natural axis of the set $S$ of the $N$ pearls necklace is represented by a vector with an origin in the center of the necklace, and the extremity is a point with coordinates $(X_{natural}, Y_{natural})$ defined by:

$$X_{natural} = \sum_{i=1}^{N} s_i \cos(\frac{2\pi}{N}(i-1));$$

$$Y_{natural} = \sum_{i=1}^{N} s_i \sin(\frac{2\pi}{N}(i-1)) \qquad (10)$$

## 5 THE CONTOUR DESCRIPTOR

In the above section 3, we showed how to obtain the affine invariant $I^l(i)$ from the *SWT* coefficients and we also how to obtain the correspondence between $(\tilde{I}^l(i))_{1 \leq i \leq N}$ and $(I^l(j))_{1 \leq j \leq N}$ using the natural axis orientation. The *SWT* transform is a non decimated version of the *DWT*, so it is redundant as the $I^l$ invariant is. To avoid this redundancy and to keep the invariant values useful for the contour description, we remove from the invariant $I^l$ all values which not correspond to the invariant obtained from the *DWT* on the contour $(x_i, y_i)_{1 \leq i \leq N}$. In (Misiti et al., 2003), the author showed that giving a *SWT* set coefficients, it is possible to get all the *DWT* $\varepsilon$-decimated for any sequence $\varepsilon = [\varepsilon_1 \varepsilon_2 ... \varepsilon_n]$ with $\varepsilon_j = 0$ or 1 for all $1 \leq j \leq n$. In a same way, it is possible to obtain the descriptors $\varepsilon$-decimated of the invariant $I^l$ by the following relation: $D^l = (D^l(j))_{1 \leq j \leq 2^{n-l}} =$
$(I^l(ind), I^l(ind + 2^l), ..., I^l(ind + (2^{n-l} - 1)2^l))$ with $ind = 1 + \sum_{i=1}^{l} \varepsilon_i 2^{i-1}$. Note that there is $N = 2^n$ possible descriptors. Furthermore, to make our algorithm more efficient, only a subset of descriptors with scale from $K$ to $L$, $K \leq L$, is used. The selection of these scales are automatically performed, using the level histogram, such the vector magnitude $(I^l)$ $1 \leq l \leq n$ of equation (7) is maximal.

## 6 THE PARTIAL-SHAPE MATCHING

The method described above is not adapted to open curves. In real applications and depending of the image quality, the contours are generally open even using effective segmentation methods. The opening may also be due to objects occlusion. Much efforts has been devoted of finding effective methods for recognition of partially occluded objects. The disparity matrix to perform similarity matching of occluded objects modeled by line segments has been used by many authors as in (Price, 1984) and (Bhanu and Ming, 1987). Khalil and Bayoumi in (Khalil and Bayoumi, 2002) proposed to use maxima lines of the continuous wavelet transform and recognize occluded objects by identifying singularities on their boundaries. In this paper, we use the method of partial shape matching based on features extracted by the wavelet transform. Actually, the shape descriptor used is the affine wavelet descriptor, and its coefficients are used in the sub-matrix matching algorithm proposed by E. Saber *et al* in (Saber et al., 2005). In the following section, we recall this method.

### 6.1 Distance Matrix

The distance Matrix represents the proximity of feature points within each example template or potential object region in order to determine complete or partial correspondances between two sets of feature points. Let $(Xi, Yi)$, $i = 1, 2, ..., n$, be feature points for a potential object region or example template contour; the distance matrix $D$ for the contour is defined as

$$D = \begin{pmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ d_{21} & d_{22} & \cdots & d_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ d_{n1} & d_{n1} & \cdots & d_{nn} \end{pmatrix} \qquad (11)$$

where $d_{kl} = \sqrt{(X_k - X_l)^2 + (Y_k - Y_l)^2}$, $k, l = 1, 2, ..., n$ denotes the distance between the feature points $k$ and $l$ along the contour. The distance matrix is a symmetric matrix, and is invariant to translation and rotation by definition, since it only depends on distances between feature points; (2) invariant to isometric scale variation of a contour, i.e. zoom or contraction, since that corresponds to scaling all distances by a constant factor; (3) reflection of a contour is equivalent to reordering of feature points in the clockwise or counter clockwise direction essentially inverting the initial ordering, and (4) if two contours partially match, their distance matrices have matching sub-matrices. The distance matrix depends on

which feature point is selected as the reference (start-ing) point for enumeration of points along the contour. For example, if $(X_k, Y_k)$ were selected as the reference point, the new distance matrix becomes.

$$\tilde{D} = \begin{pmatrix} d_{kk} & d_{k(k+1)} & \cdots & d_{k1} & \cdots & d_{k(k-1)} \\ d_{(k+1)k} & d_{(k+1)(k+1)} & \cdots & d_{(k+1)1} & \cdots & d_{(k+1)(k-1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ d_{1k} & d_{1(k+1)} & \cdots & d_{11} & \cdots & d_{1(k-1)} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \cdots \\ d_{(k-1)k} & d_{(k-1)(k+1)} & \cdots & d_{(k-1)1} & \cdots & d_{(k-1)(k-1)} \end{pmatrix} \quad (12)$$

Note that $D$ and $\tilde{D}$ are related by a circular shift, that is $D$ can be obtained by shifting $\tilde{D}$ by $(k-1)$ rows up and $(k-1)$ columns to the left. This is important because starting points for enumerating feature points on a potential object region and an example template may differ. However, this does not pose a difficulty in finding partial matches between the two. Sub-matrix matching suppose $(X_j^Q, Y_j^Q), j = 1, 2, 3, ..., m$, and $(X_i^R, Y_i^R), i = 1, 2, 3, ..., n$, denote the feature points for a given example template and candidate image region respectively, the corresponding distance matrices for the template $D_{(m \times m)}^Q$ and the image region $D_{(n \times n)}^R$ as defined in (Saber et al., 2005)

## 7 EXPERIMENTATION

To validate our algorithm with a consistent database, we created 3 databases $DB_1$ $DB_2$ and $DB_3$ from an original database $DB_0$ containing 500 objects (marine species). All these objects are nine time rotated with angle $\theta = \frac{k\pi}{9}$ $k = 1, 2, .., 9$, followed a stretching with a factor $s$ according to the transform: $\begin{bmatrix} 1 & s \\ 0 & 1 \end{bmatrix}$ we fix $s = 1$ for the first database, $s = 2$ and $s = 3$ for the second and the third database. Finally, each database contains 5000 marine species contours (500 original contours and 4500 contours obtained with the 9 rotations followed by a stretching). In our experimentation, we have re-sampled the contour in $N = 2^9 = 512$ points and $\varepsilon_i$ is fixed to 0 for each $i \in [1, 9]$ and $K = L = 6$ (see section 5). Note that for the first step the contour is closed, and for the partial shape matching, parts of the shape of all the database objects were randomly removed to obtain open contours. Each model of the database is represented by a vector $D^6$ of dimension $2^{9-6} = 8$. For the matching step, the Haar's wavelet and the city-block distance were used. In this paper, and to illustrate the results obtained by the proposed method, we use four contours of the figure (2). P1 is the original contour fish. P2 is P1 rotated with $\pi/4$ followed by a stretch of $s = 0.5$. P3 is
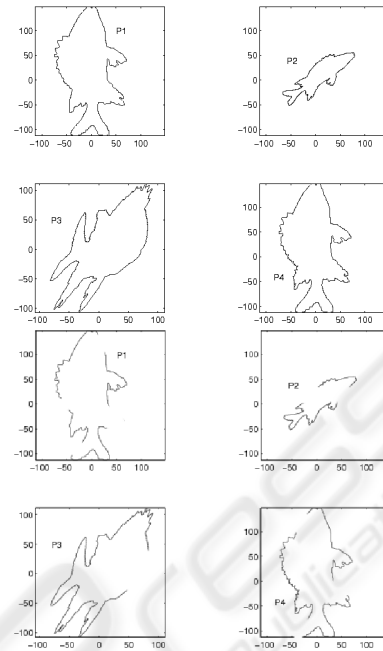


Figure 2: (a) P1 Original contour fish. P2 is P1 with rotation of $\pi/4$ followed by a stretch of $s = 0.5$. P3 is P1 rotated with $\pi/2$ followed by a stretch of $s = 1.5$ and a scaling of 0.5 and 1.5 on the $X$ and $Y$ axis. On $P4$ a uniform noise is added locally on the contour . (b) The same figures as in the previous figure, with open contour, the cutting contour is performed randomly.

P1 rotated with $\pi/2$ followed by a stretch of $s = 1.5$ and a scaling of 0.5 and 1.5 on the $X$ and $Y$ axis. On $P4$ a uniform noise is added locally on the contour (The noise is not clearly visible due to scaling of the figure in printing).

## 8 CONCLUSIONS

In this paper, we develop a method of 2D shape recognition under affine transform using the discrete dyadic wavelet transform invariant to translation *(Stationary Wavelet Transform or SWT)*. We have also showed how to choose the starting point on the contour using the orientation of the natural axis. This orientation enables the elimination of the redundant values on the affine invariant which is stable for a large distortion. We the generalized the method to partial shape using the the sub-matrix matching algorithm developed by (Saber et al., 2005) to determine correspondances for evaluation of partial similarity between an example template and a candidate object region. The method is tested on a database of 5000 fish species, and the results are very probate. We are actually plan to generalize this method to partial shape matching of 3D
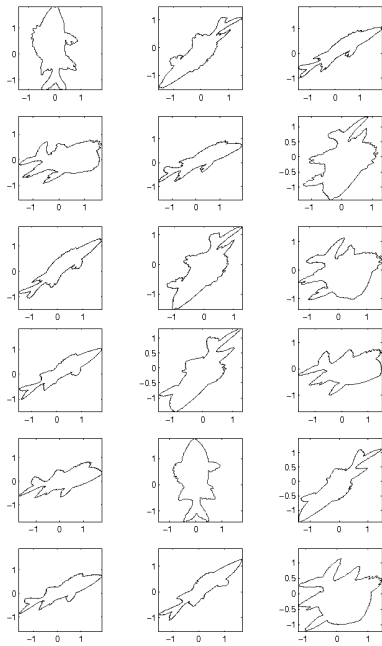
Figure 3: (a)The results obtained for queries P1, P2 and P3, using any of the 3 database with partial shape. (b) Results of the query *P*4 with partial shape.
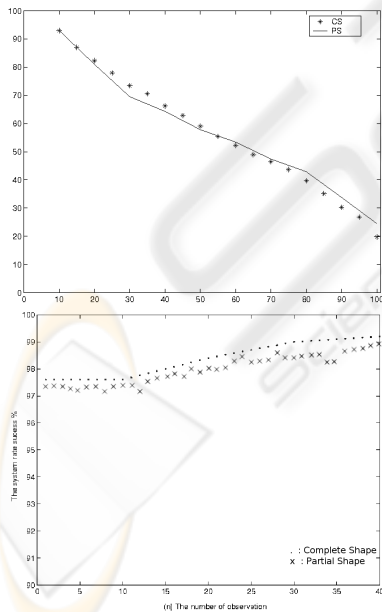


Figure 4: (a) The Recall/Precision evaluation between the Complete Shape(CS) and Partial Shape (PS), performed on 500 queries. (b) The global success rate in % according to the number of observation of the system performed on 500 queries.

objects.

## REFERENCES

Bhanu, B. and Ming, J. (1987). Recognition of occluded objects, a cluster-structure algorithm. *Pattern Recognition*, 20 (2):199–211.

Khalil, M. and Bayoumi, M. (2001). A dyadic wavelet affine invariant function for 2d shape recognition. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 23(10):1152–1163.

Khalil, M. and Bayoumi, M. (2002). Affine invariant for objet recognition using the wavelet transform. *Pattern Recognition Letters*, 23:57–72.

Kimcheng, K. and El-hadi, Z. (2004). 2D Affine-Based Recognition Using Discrete Wavelet. *International Conference on Computer Vision and Graphics, September 22-24, 2004 Warsaw, Poland*.

Mallat, S. (1989). A theory for multiresolution signal decomposition : the wavelet representation. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 11:674–693.

Mallat, S. (1991). Zero-crossings of a wavelet transform. *IEEE Transactions on Information Theory*, 37:1019–1033.

Misiti, M., Misiti, Y., Oppenheim, G., and Poggi, J.-M. (2003). *Les ondelettes et leurs applications*. hermes Sciences.

Price, K. (1984). Matching closed contours. *International Conference on Pattern Recognition, Jerusalem*.

Saber, E., Yaowu, X., and Tekalp, A. (2005). Partial shape recognition by sub-matrix matching for partial matching guided image labeling. *Pattern Recognition*, 38:1560–1573.

Shen, D., him Wong, W., and Ip, H. H. (1999). Affine-invariant image retrieval by correspondence matching of shapes. *Image and Vision Computing*, 17:489–499.

Tieng, Q. M. and Boles, W. W. (1997a). Recognition of 2d object contours using the wavelet transform zero-crossing representation. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 19(8):910–916.

Tieng, Q. M. and Boles, W. W. (1997b). Wavelet-based affine invariant representation: A tool for recognizing planar objects in 3d space. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 19(8):846–857.

Zhang, D. and Lu, G. (2004). Review of shape representation and description techniques. *Pattern Recognition*, 37(1):1–19.