

Mine Classification based on raw sonar data: an approach combining Fourier Descriptors, Statistical Models and Genetic Algorithms

I. Quidu*, J. Ph. Malkasse*, G. Burel**, P. Vilbé**

(*) Thomson Marconi Sonar, Route de Sainte Anne du Portzic, 29601 BREST cédex, France

(**) L.E.S.T. - UMR CNRS 6616, 6 avenue Le Gorgeu, BP 809, 29285 BREST cédex, France

Abstract- In the context of mine warfare, detected mines can be classified from their cast shadow. A standard solution is to perform image segmentation first (we obtain binary from graylevel image giving the label zero for pixels belonging to the shadow and the label one elsewhere), and then to perform a classification based on features extracted from the 2D-shape of the segmented shadow. Consequently, if a mistake happens during the process, it will be propagated through the following steps. In this paper, to avoid such drawbacks, we propose a novel approach where a dynamic segmentation scheme is fully classification-oriented. Actually, classification is performed directly from the raw image data. The approach is based on the combination of *deformable models*, *genetic algorithms*, and *statistical image models*.

I. INTRODUCTION

Deformable contours are flexible models able to fit the data. Two techniques have been widely used: snakes and deformable templates. *Snakes* are active contour models guided by internal constraint forces and image forces [1]. These free-shape models can be deformed in order to match salient image features without prior information on the geometry of the shape. They do not require image preprocessing. On the opposite, *deformable templates* describe the shape by using a deformation of a basic template [2]. The image must be segmented first, in order to obtain an explicit contour description.

Our approach takes place between these two techniques. Searching the best solution through a large space of potential solutions, a particular optimization process has to be applied. Genetic algorithms are such balanced methods which exploit the best solutions while exploring the search space. For our purpose, a set of templates (i.e. *individuals*) are deformed in order to maximize an energy function based on the statistical properties of the observed image. These individuals are contours alike snakes which manage to fit the observed image. Alterations among individuals lead to a solution characterized by the Fourier decomposition of its contour.

The paper is organized as follows. In Section II, we review briefly the principle of the genetic algorithms. Section III describes the implementation. In Section IV, experimental results obtained on both simulated and real sonar images are provided. Finally the conclusion of our study is given in Section V.

II. GENETIC ALGORITHMS

Genetic Algorithms (GA) are stochastic search methods that mimic the natural biological evolution. They operate on a *population* of potential *individuals* applying the principle of survival of the fittest to produce better and better

approximations of a solution [3] [4]. GAs are executed over a sequence of iterations on a set of coded individuals, the population, with three basic operators : *selection/reproduction*, *crossover*, and *mutation*. In each generation a probabilistic selection is performed based upon the individual's fitness such that the best individuals have an increased chance of being selected to reproduce in the next generation. Genetic operators are applied on these *parent chromosomes* and new chromosomes (*offspring*) are generated.

III. DESCRIPTION OF THE PROPOSED ALGORITHM

The genetic algorithm toolbox implemented in Matlab by Houck et al. gave us some useful functions adaptable to the given problem [5]. Their Matlab toolbox, named GAOT, Genetic Algorithms for Optimization Toolbox, provides a group of related functions with easy extensibility and modularity.

Given the raw image data, the population converges to a contour well-fitted to the image. Each individual is characterized by specific Fourier descriptors related to its shadow contour. This contour splits the observed image in two homogeneous regions whose statistical properties are used to evaluate the individual's fitness. Over a sequence of iterations, GA generates better sets of individuals thanks to selection and reproduction (using genetic operators) of the best individuals.

A. Individual encoding

Shadows are described by two sets of Fourier coefficients, corresponding to the Fourier transformations of the coordinates of the contour pixels. Given the set of N uniformly spaced pixels $(x_i, y_i)_{i=0 \dots N-1}$ extracted from the closed contour we compute the double set of Fourier descriptors hereafter:

$$\begin{cases} X_k = \frac{1}{N} \sum_{i=0}^{N-1} x_i \exp\left(-j2\pi \frac{ki}{N}\right) \\ Y_k = \frac{1}{N} \sum_{i=0}^{N-1} y_i \exp\left(-j2\pi \frac{ki}{N}\right) \end{cases}$$

In order to avoid the effect of an arbitrary starting point and to provide an unique description for a given contour, these coefficients are normalized as proposed by K. Arbter [6].

Given $C_k = \frac{1}{N} \sum_{i=0}^{N-1} (x_i + jy_i) \exp\left(-j2\pi \frac{ki}{N}\right) = X_k + jY_k$ we really compute the following coefficients :

$$\begin{cases} \tilde{X}_k = X_k \frac{C_1^{-k}}{|C_1|^{-k}} = X_k \frac{(X_1 + jY_1)^{-k}}{|X_1 + jY_1|^{-k}} \\ \tilde{Y}_k = Y_k \frac{C_1^{-k}}{|C_1|^{-k}} = Y_k \frac{(X_1 + jY_1)^{-k}}{|X_1 + jY_1|^{-k}} \end{cases}$$

Chromosomes are made of units called *genes* (actually the features related to the set of Fourier descriptors) arranged in a definite succession. Here each chromosome $\theta_{(K)}$ presents the real and imaginary parts of the two sets of K positive descriptors (because $\tilde{X}_{-k} = \tilde{X}_k^*$ and $\tilde{Y}_{-k} = \tilde{Y}_k^*$):

$Re(\tilde{X}_1)$				$Re(\tilde{X}_K)$				$Im(\tilde{X}_1)$				$Im(\tilde{X}_K)$
gene 1				gene K				gene K+1				gene 2K
$Re(\tilde{Y}_1)$				$Re(\tilde{Y}_K)$				$Im(\tilde{Y}_1)$				$Im(\tilde{Y}_K)$
gene 2K+1				gene 3K				gene 3K+1				gene 4K

B. Population

Initially, a given number Np of prototypes is provided for each of the five classes considered : cylinders, spheres, and three stealthy mines, the Manta and Sigeel mines which look like truncated cones and the Rockan mine with its low and peculiar profile. This population is supposed to cover a large set of possible situations especially under any point of view. Fig. 1 shows some of the initial individuals.

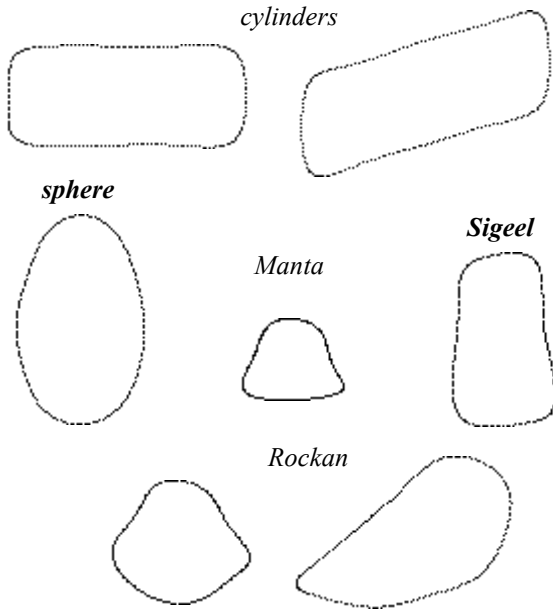


Fig. 1. Some initial individuals.

C. Evaluation function

The evaluation or fitness function is chosen in a way such that highly fitted strings (or chromosomes) are preferred. During the selection process (see III.D.1.), individuals whose chromosomes have high fitness values are chosen with higher probability. Given the observed image, fitness of a given contour can be evaluated using the statistical framework below.

We aim at finding the boundary between two specific homogeneous regions of the observed image, namely the reverberation and the shadow of the detected object. The conditional probability function of each pixel depends only on whether it belongs to the inside or outside region of the

contour: i.e., all pixels inside (*resp.* outside) have a common distribution characterized by a parameter vector ϕ_{in} (*resp.* ϕ_{out}), with $\phi = [\phi_{in}, \phi_{out}]$ [7]. The likelihood function to maximize is written as :

$$P(I/\theta_{(K)}, \phi) = \left(\prod_{(i,j) \in R_{in}} p(I(i,j)/\phi_{in}) \right) \left(\prod_{(i,j) \in R_{out}} p(I(i,j)/\phi_{out}) \right),$$

where $I(i,j)$ denotes the intensity of pixel whose coordinates in the image are (i,j) , R_{in} (*resp.* R_{out}) region of pixels belonging to the shadow (*resp.* both the reverberation and the echo). Components of the vector $\theta_{(K)}$ (whose length is $2 \times 2K$) are the K Fourier descriptors characterizing the individual (see III.A.). The corresponding contour whose coordinates are $(x_i, y_i)_{i=1 \dots N}$ can be rebuilt with:

$$\begin{cases} x_i = \sum_{k=-K}^K \tilde{X}_k w^{ki} \\ y_i = \sum_{k=-K}^K \tilde{Y}_k w^{ki} \end{cases} \quad \text{where } w = \exp(j2\pi / N)$$

Given $\theta_{(K)}$, the explicit contour representation is obtained with a degree of smoothness defined by the order K . This contour splits the image in two regions: the inside (shadow) and the outside (reverberation and echo).

The goal is to estimate $\theta_{(K)}$ and ϕ from the observed image maximizing the likelihood function. For sonar images, pixels generally have a Rayleigh distribution. For each region, it follows:

$$p(I(i,j)/\Phi_{region}) = \frac{I_{region}(i,j) - d_{region}}{\alpha_{region}^2} \exp \left\{ -\frac{[I_{region}(i,j) - d_{region}]^2}{2\alpha_{region}^2} \right\}$$

where the parameters $\Phi_{region} = [d_{region}, \alpha_{region}^2]$ have to be estimated.

Appendix A gives the following estimates:

$$\begin{cases} \hat{d}_{region} = \bar{I}_{region}(i,j) - \sqrt{\frac{\pi \sigma_{region}^2}{4 - \pi}} \\ \hat{\alpha}_{region}^2 = \frac{\sum_{(i,j) \in region} [I_{region}(i,j) - \hat{d}_{region}]^2}{2 \cdot N_{region}} \end{cases}$$

where

N_{region} is the number of pixels belonging to the region,

$\bar{I}_{region}(i,j)$ is the mean of the intensities of these pixels,

σ_{region}^2 the variance of these pixels.

The maximum likelihood estimation of $\theta_{(K)}$ and ϕ consists in maximizing the log-likelihood function $\ln(P(I/\theta_{(K)}, \phi))$,

when $\tilde{I}(i,j) = I(i,j) - \hat{d} > 0$. It amounts then to maximize the following evaluation function:

$$E = \sum_{I(i,j) \in R_{in}} \ln \tilde{I}(i,j) - N_{R_{in}} \ln \alpha_{R_{in}}^2 + \sum_{I(i,j) \in R_{out}} \ln \tilde{I}(i,j) - N_{R_{out}} \ln \alpha_{R_{out}}^2$$

The evaluation requires the mean point $(\tilde{X}_0, \tilde{Y}_0)$ around which the contour has to be rebuilt from the others descriptors $(\tilde{X}_k, \tilde{Y}_k)_{k=1...K}$. Practically this point is found in the vicinity of an initial point (chosen by the operator for the given observed image) in such a way that it maximizes the evaluation function E.

D. Basic operators

1) Selection/reproduction procedure

During an iteration, a fixed number of individuals is maintained. Each individual is evaluated thanks to the above evaluation function to give some measure of its fitness. Then a new population is formed by selecting the best individuals. There are several schemes for the selection process [8]. Based on ranking methods, *normalized geometric ranking* only requires the evaluation function to map the individuals to a partially ordered set [5]. A probability of selection P_i is assigned to individual i when all individuals are sorted:

$$P_i = q(1-q)^{r_i-1}$$

where

- q = the probability of selecting the best individual
- r_i = the rank of the individual i , where 1 is the best
- Np = the population size
- $q' = q/(1-(1-q)^{Np})$

Choosing $q=0.15$ with $Np=95$, about 1/3 of the individuals have a probability P_i higher than 0.001. It appears as a good compromise between a lack of diversity (for $q<0.15$) and large computational times (for $q>0.15$).

2) Genetic operators

Genetic operators are applied on parent chromosomes: new chromosomes, i.e. *offspring*, are generated. Altering the composition of children, a certain diversity is preserved. Dealing with particular genes namely real and imaginary parts of Fourier descriptors, we have implemented some suited genetic operators making good use of Fourier descriptors properties. In order to keep constant the number of individuals during each iteration, *crossover* takes two individuals and produces two new children while *mutation* alters one individual to produce a single new child. Note that c_i (*resp.* p_i) will refer to a child (*resp.* a parent):

- crossover

A crossover operation recombines genetic material of two parent chromosomes to produce offspring for the next generation. We implement four different operations (*best*(p_1, p_2) means "the best of (p_1, p_2) in terms of fitness"):

1. *Random exchanges occur between the descriptors of the parents with a probability equal to 0.2*

2. $c1 = (p1 + p2)/2$
 $c2 = \text{best}(p1, p2)$
3. $c1 = (p1 + 0.5 \times p2)/1.5$
 $c2 = (0.5 \times p1 + p2)/1.5$

4. *same operation as 2. but only on the coefficients X_1 and Y_1 (related to the principal axes)*

The three last crossovers occur with a probability equal to 0.2/3. On the whole, the crossover probability is fixed as 0.4.

- mutation

Based on the general idea of mutation, the first operation happens one time per iteration and consists in additionning a white gaussian noise on one of the descriptors. The second operation occurs more frequently with a probability equal to 0.2 and operates on the whole chromosome. It consists in doing an affine transformation of the Fourier descriptors (see appendix B). Child's contour appears like the affine transformation of the parent's as if it was seen through a different point of view. Six affine transformations are possible using a 2x2 matrix A:

- scaling $A = \begin{bmatrix} 1+\alpha & 0 \\ 0 & 1+\alpha \end{bmatrix}$, with the random parameter $\alpha \in [-0.4; +0.4] \setminus \{0\}$

- elongation along the rows of the image $A = \begin{bmatrix} 1+\alpha & 0 \\ 0 & 1 \end{bmatrix}$, with the random parameter $\alpha \in [-0.4; +0.4] \setminus \{0\}$

- elongation along the columns $A = \begin{bmatrix} 1 & 0 \\ 0 & 1+\alpha \end{bmatrix}$, with the random parameter $\alpha \in [-0.4; +0.4] \setminus \{0\}$

- rotation $A = \begin{bmatrix} \cos\left(\frac{\pi}{40} \times \alpha\right) & -\sin\left(\frac{\pi}{40} \times \alpha\right) \\ \sin\left(\frac{\pi}{40} \times \alpha\right) & \cos\left(\frac{\pi}{40} \times \alpha\right) \end{bmatrix}$, with the random parameter $\alpha \in [-2; +2] \setminus \{0\}$

- skew transformation on the rows $A = \begin{bmatrix} 1 & \alpha \\ 0 & 1 \end{bmatrix}$, with the random parameter $\alpha \in [-0.1; +0.1] \setminus \{0\}$

- skew transformation on the columns $A = \begin{bmatrix} 1 & 0 \\ \alpha & 1 \end{bmatrix}$, with the random parameter $\alpha \in [-0.1; +0.1] \setminus \{0\}$.

After a genetic operation, each child is characterized by its normalized descriptors (normalization has to be ensured after any genetic operation) and its fitness value.

E. Termination criterion

The genetic algorithm is terminated when one of the following conditions is met:

- no improvement in the best individual happens during three consecutive generations,
- the specified maximum number of generations taken as 20 is reached.

F. Classification step

Classification can be performed by comparing the Fourier descriptors of the winner with the Fourier descriptors of the initial prototypes. Thanks to the uniqueness of the normalized Fourier descriptors, we compute a classification criterion assigning the fittest individual to the same class as its nearest neighbour according to a minimal distance.

A distance d_j can be computed from genes of the fittest individual and those of the j^{th} individual of the initial population such that $d_j = \left(\sum_{i=1}^{4K} \left[\frac{(g_i^j - \tilde{g}_i)}{p_i} \right]^2 / 4K \right)^{\frac{1}{2}}$ where

p_i is the weight related to the maximal value allowed for the gene g_i

$(\tilde{g}_i)_{i=1 \dots 4K}$ are the estimated genes and $(g_i^j)_{i=1 \dots 4K}$ are the genes of the j^{th} individual's chromosome of the initial population.

Each index j refers to a specific prototype mine whose class is known. The index $j = \text{Arg} \min_{j \in \{1 \dots Np\}} d_j$ (Np is the size of the population) is then related to the nearest prototype. Notice that individuals constituting the initial population is arranged as follows: 19 cylindrical mines, 19 spherical mines, 19 Manta mines, 19 Siegel mines and 19 Rockan mines.

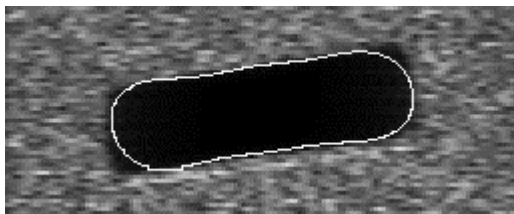
Remember that genes are related to the real and the imaginary parts of the normalized Fourier descriptors according to the individual representation (see III.A.).

IV. EXPERIMENTAL RESULTS

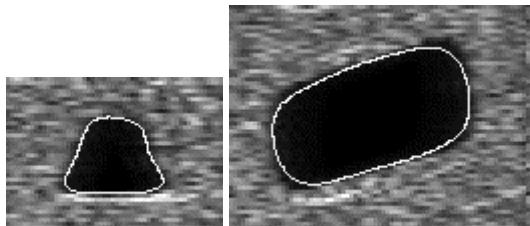
To improve the robustness of the optimization, an image normalization is performed to provide a new image as it would be seen through a grazing angle of 45 degrees preserving shape ratios.

For each real and simulated example, we give the contour rebuilt from the estimated descriptors and the best distance d_j as defined above.

❖ Simulated sonar images :

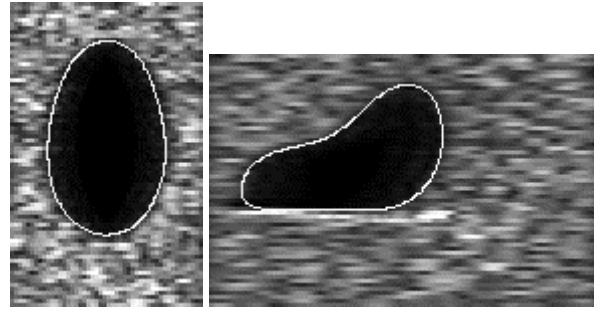


$d_4 = 0.133 \Rightarrow$ cylinder



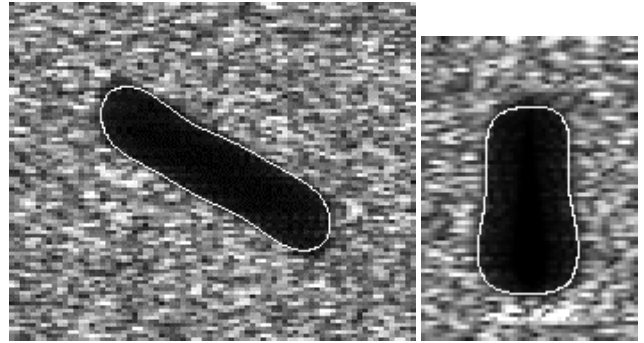
$d_{56} = 0.018 \Rightarrow$ manta

$d_5 = 0.073 \Rightarrow$ cylinder



$d_{38} = 0.011 \Rightarrow$ sphere

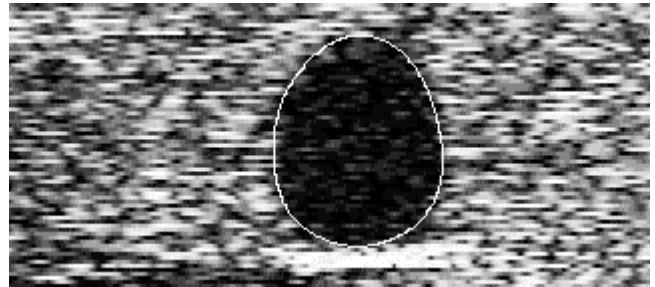
$d_{84} = 0.029 \Rightarrow$ rockan



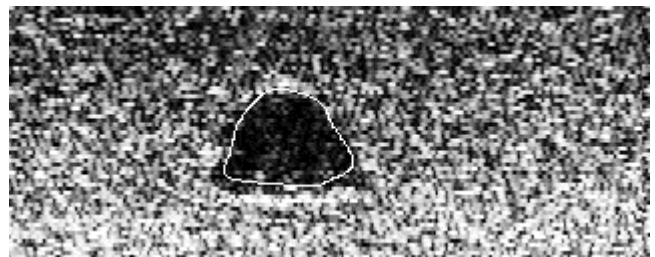
$d_8 = 0.160 \Rightarrow$ cylinder

$d_{58} = 0.059 \Rightarrow$ siegel

❖ Synthetic aperture sonar images [9]:



$d_{37} = 0.073 \Rightarrow$ sphere



$d_{55} = 0.228 \Rightarrow$ manta

V. CONCLUSION

In the context of mine warfare, shapes of possible mines are well-known [10]. In terms of shadow recognition, one need only to consider a limited set of cases depending on the geometry of mines. Genetic algorithms are such techniques for searching through a space (population) of potential individuals. In this paper, we characterize each individual by

a normalized double Fourier transformation of the coordinates of the contour pixels. Given a statistical criterion, namely the evaluation function, individuals converge towards an optimal one. From raw image data, a Maximum Likelihood approach allows to rate individuals in terms of their fitness. Furthermore specific genetic operators take advantage of the double Fourier descriptors: while crossovers give offspring similar to their parents, new offspring appears thanks to mutations preserving diversity. Finally, the fittest individual is classified comparing its descriptors and those of the initial population.

Our approach stands out against classical classification processing. Dealing with raw data, we discard some undesirable steps. Indeed, without sequential processing, a punctual perturbation will have less repercussion on the final result. Instead of doing image segmentation firstly, feature extraction secondly and performing classification from these features lastly, each individual acts as a potential solution identifiable thanks to the Fourier decomposition of its contour. In this way, the whole of the contours move in order to match the shadow resulting in a dynamic classification process.

APPENDIX A

Probability density function $\mathfrak{R}(x, \Phi)$ of Rayleigh distribution is defined as

$$\mathfrak{R}(x, \Phi) = p(x/\Phi) = \frac{x-d}{\alpha^2} \exp\left\{-\frac{[x-d]^2}{2\alpha^2}\right\}$$

whose parameters are $\mathfrak{R}(x, \Phi)$. d stands for the shift from the origin and α is the scale parameter.

If $x=(x_1, x_2, \dots, x_N)$ is a realization of the variate $X=(X_1, X_2, \dots, X_N)$ considering N random variables (pixels of the observed image), independent of each other and obeying a Rayleigh law $\mathfrak{R}(x, \Phi)$, formally the maximum likelihood estimate $\hat{\Phi}$ is

$$\hat{\Phi} = \arg \max_{\Phi} \{\ln P_{X/\Phi}(x/\Phi)\}$$

Thanks to independence,

$$\begin{aligned} \ln P_{X/\Phi}(x/\Phi) &= \ln \left\{ \prod_{i=1}^N \left[\left(\frac{x_i-d}{\alpha^2} \right) \exp\left(-\frac{(x_i-d)^2}{2\alpha^2}\right) \right] \right\} \\ &= \sum_{i=1}^N \ln(x_i-d) - N \ln \alpha^2 - \frac{1}{2\alpha^2} \sum_{i=1}^N (x_i-d)^2 \end{aligned}$$

from which we obtain the estimate $\hat{\alpha}^2 = \frac{\sum_{i=1}^N \tilde{x}_i^2}{2N}$ with $\tilde{x}_i = x_i - d > 0, \forall i$.

To find the shift d , one can use the statistical properties of the centered Rayleigh law ($d=0$) that is to say the mean μ and the variance σ^2 defined as :

$$\mu = \alpha \sqrt{\frac{\pi}{2}} \quad \text{and} \quad \sigma^2 = \alpha^2 \left(2 - \frac{\pi}{2}\right)$$

For the general Rayleigh law ($d \neq 0$) the mean μ_d depends on d such as $\mu_d = \mu + d$.

Consequently for pixels X_i obeying to the same law, d corresponds to the difference between the mean of these pixels' intensities and a term depending on their variance :

$$d = \mu_d - \sqrt{\frac{\pi \sigma_d^2}{4 - \pi}} \quad \text{with} \quad \begin{cases} \mu_d = \sum_{i=1}^N x_i / N \\ \sigma_d^2 = \sum_{i=1}^N (x_i - \mu_d)^2 / N \end{cases}$$

APPENDIX B

An affine transformation of the contour is achieved doing this operation on the double set of descriptors. Indeed, given the following affine transformation of the descriptors (except translation):

$$\begin{bmatrix} U_k \\ V_k \end{bmatrix} = A \times \begin{bmatrix} X_k \\ Y_k \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \times \begin{bmatrix} X_k \\ Y_k \end{bmatrix}, \quad \forall k=1, \dots, K$$

Computing the coordinate values of the N corresponding contour pixels from $\begin{cases} U_k = aX_k + bY_k \\ V_k = cX_k + dY_k \end{cases}$ we have, with

$$w = \exp(j2\pi / N),$$

$$\begin{cases} u_i = \sum_{\substack{k=-K \\ k \neq 0}}^K U_k w^{ki} = \sum_{\substack{k=-K \\ k \neq 0}}^K (aX_k + bY_k) w^{ki} = ax_i + by_i \\ v_i = \sum_{\substack{k=-K \\ k \neq 0}}^K V_k w^{ki} = \sum_{\substack{k=-K \\ k \neq 0}}^K (cX_k + dY_k) w^{ki} = cx_i + dy_i \end{cases}$$

$$\text{so: } \begin{bmatrix} u_i \\ v_i \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \times \begin{bmatrix} x_i \\ y_i \end{bmatrix} = A \times \begin{bmatrix} x_i \\ y_i \end{bmatrix}$$

what proves that the corresponding contour undergoes the same affine transformation.

REFERENCES

- [1] M. Kass, A. Witkin and D. Terzopoulos, "Snakes: active contour models," *International Journal of Computer Vision*, pp. 321-331, 1988.
- [2] 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, March 1996.
- [3] D.E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, Reading, MA, 1989.

- [4] Z. Michalewicz, *Genetic Algorithms + Data Structures = Evolution Programs*, Springer-Verlag, Berlin, 1992.
- [5] C.R. Houck, J.A. Joines, M.G. Kay, "A genetic algorithm for function optimization : a Matlab implementation," *NCSU-IE Technical Report 95-09*, 1995.
- [6] K. Arbter, "Affine-invariant Fourier descriptors," in *From Pixels to Features*, J.C. Simon (ed.). Amsterdam, The Netherlands: Elsevier Science, 1989.
- [7] M.A.T. Figueiredo, J.M.N. Leitão, and A.K. Jain, "Adaptative parametrically deformable contours," *Proceedings of Computer Vision and Pattern Recognition*, 1997.
- [8] D.E. Goldberg and K. Deb, "A comparative analysis of selection schemes used in genetic algorithms," *Foundations of Genetic Algorithms*, Ed. G.J.E. Rawlins, 1991.
- [9] D. Billon and F. Fohanno, "Theoretical performance and experimental results for synthetic aperture sonar self-calibration", in *OCEANS'98 MTS/IEEE*, pp. 965-970, September 1998.
- [10] *Jane's Underwater Warfare Systems*, Ed. By Anthony J. Watts, Tenth Edition 1998-1999.