Mine Classification based on a Multiview Characterisation

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

[(*)] Thomson Marconi Sonar, Route de Sainte Anne du Portzic, 29601 BREST cédex, France
isabelle.quidu@tms.thomson-csf.com, jean-philippe.malkasse@tms.thomson-csf.com
[(**)] L.E.S.T. - FRE CNRS 2269, 6 avenue Le Gorgeu, BP 809, 29285 BREST cédex, France
Pierre.Vilbe@univ-brest.fr, Gilles.Burel@univ-brest.fr

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, and then to extract a set of features from the shape allowing classification in a final step. In this paper, we extend this procedure to a sequence of images obtained along a part of a circular trajectory of the sonar.
For a given ground mine except mine with radial symmetry, cast shadow appearance generally depends on the point of view. Consequently, different features values can describe the same object. Whereas this often entails misclassification when a single view is used, we propose to use feature values computed over a sequence of images, especially its evolution, to characterise objects from multiple views. Our supervised classification scheme is based on the correlation, for each feature, between the sequence of values obtained from the unknown object and typical values related to each class.

I. Introduction

A high resolution sonar provides high-quality acoustic images of the sea-bed, allowing the classification of objects from the shadow they cast. After the segmentation step, a set of features is extracted from the shadow (1) (2) (3). Using a single view, pattern recognition can be accomplished by identifying the unknown object as a member of a set of well-known objects. Such a supervised classification can then be performed by using several classical techniques such as k-nearest-neighbour classifier for example (4). But even if we use features which are invariant under appropriate transformations, conditions are not always propitious. By taking more than a single view, ambiguities can be removed while making classification become more robust. A solution consists in fusing information from different views of the object (5). Actually, fusion techniques enable to represent imprecise or uncertain data by means of particular measures called degrees of belief. The decision is taken after the combination of these pieces of information. But as far as images acquired according to a multiview strategy are concerned, an powerful kind of information can be judiciously used, i.e. the sequential evolution of feature values. This interesting information is unfortunately lost by use of fusion techniques which consider each sensor or sources independently. Under these considerations, we propose to characterise the detected object by the evolution of feature values computed over the sequence of images provided that sonar images are acquired under a precise trajectory.
This paper is organised as follows. Main steps of shadow shapes characterisation are described in the second section. Then section 3 derives the multiview from a given single-view characterisation of an object. The multiview classification algorithm is detailed is section 4. Section 5 shows the experimental results.

II. Characterisation of shadow shapes in sonar imagery

Before evaluating the features, image data are preprocessed in order to obtain a binary image and to improve the robustness of the features. An example is given in Fig. 1.

II.1. Segmentation step

Segmentation consists in partitionning the image into homogeneous regions. In our case, objects are classified from their cast shadow. Each image is made up of both the echo and the cast shadow caused by the detected object and the sea-bed reverberation area. Giving the label zero for pixels belonging to the shadow and the label one elsewhere we obtain binary from greylevel image. Sonar image's grey-level histogram is generally unimodal and then threshold selection by the mode method is impossible. To make it bimodal, a specific spatial filtering can be applied in order to minimise pixels variance. In the resulting histogram, peaks represent shadow and reverberation regions (differing in average grey-level) while the single valley represents shadow edge. The optimal filter designed for sonar images does not look like a classic low-pass filter but strongly depends on the principle of the sonar image formation in terms of size and coefficient values. For details on this technique, please refer to (6).

II.2. Image normalisation

To improve robustness of topological features, an image normalisation is performed. It has to provide a new image as it would be seen under a grazing angle of 45 degrees preserving shadow ratios. Moreover on account of the sonar parameters, image resolution is generally different along the two dimensions. Consequently, each pixel is made approximately square to prevent from disproportions.

II.3. Noise reduction

Irregularities of the boundary of the shadow may have undesired effects on the recognition system. While preserving the global shape of the shadow, we aim at smoothing the boundary. The shadow's closed boundary can be represented by a periodic function of the contour coordinates. Computing Fourier descriptors and removing the high frequencies, the resulting shadow is smoother than the original one.

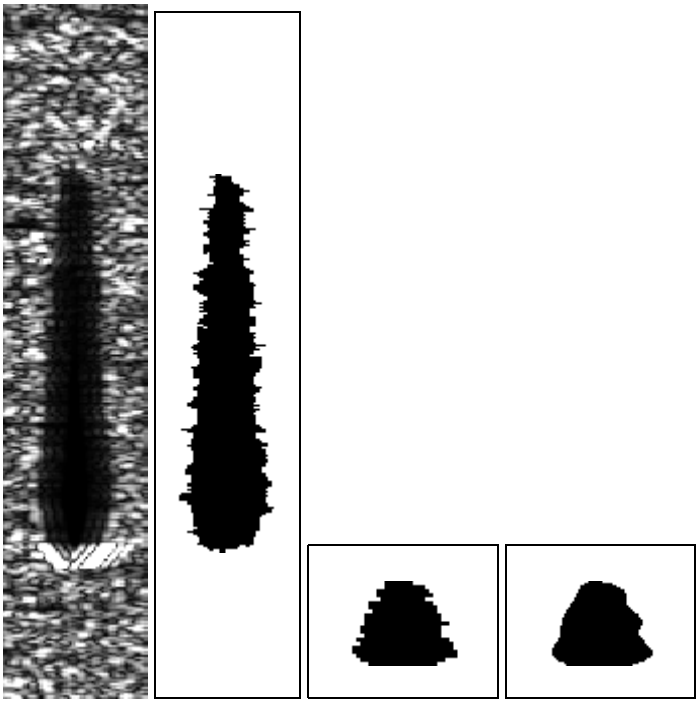

Fig. 1. The initial sonar image, the corresponding binary, normalised binary and final smoothed normalised binary images

II.4. Features extraction

In the context of mine classification, the question is to select a set of L appropriate features with acceptable recognition accuracy. Several kinds of features are used for recognition, especially geometrical features and statistical features. In addition to the widely used topological parameters as in Jan(7), we were interested into moments of shapes. Some particular functions of moments introduced by Flusser and Suk(8) have the useful property of affine invariance. Choosing only three features, a hybrid set of descriptors may characterise the 2D shape i.e. the cast shadow. Components of the three-dimensional characteristic vector $(f_i)_{i=\{1\ldots3\}}$ are then related to two topological parameters and a moment invariant. This last feature characterises shapes deprived of any affine transformation. It is interesting in case of complex mines such as the Rockan mine for which any shadow shape cannot be referred to affine transformations of some reference one. The other features emphasize major transformations of shapes. The first one is the length (taken along the range axis) to the width (taken along the azimuth axis) ratio and the second one the elongation parameter computed from second order central moments (7).

III. Towards a multiview characterisation

III.1. Problem description

To ensure ship safety against modern mines, some minehunting vessels will now carry an offboard system such as a ROV (Remotely Operated Vehicule) mounted sonar. For this purpose, Thomson Marconi Sonar promotes the Propelled Variable Depth Sonar (PVDS) which can easily turn around an object to be classified and hence provides short range multiple views of the object from different points of view. In our application,

a sequence of N sonar images I_n is simulated as if the sonar turned around each mine with a shot every 10 degrees. The object is then characterised by N characteristic vectors $V_n = \left(f_1^n \quad f_2^n \quad f_3^n\right)^T$ sequentially stored.

III.2. Features evolution

In our experiments, five classes (K=5) have then been considered: cylinders (3 sizes), spheres (3 sizes), and three stealthy mines, i.e. two which look like truncated cones (the Manta and Sigeel mines) and another one with sloping angled faces and low profile (the Rockan mine). By storing features values along the circular trajectory of the sonar, these objects are easily distinguishable. Moreover, geometrical properties and especially symmetries are clearly displayed for a cylindrical mine on Fig.2 around which the sonar turned round.

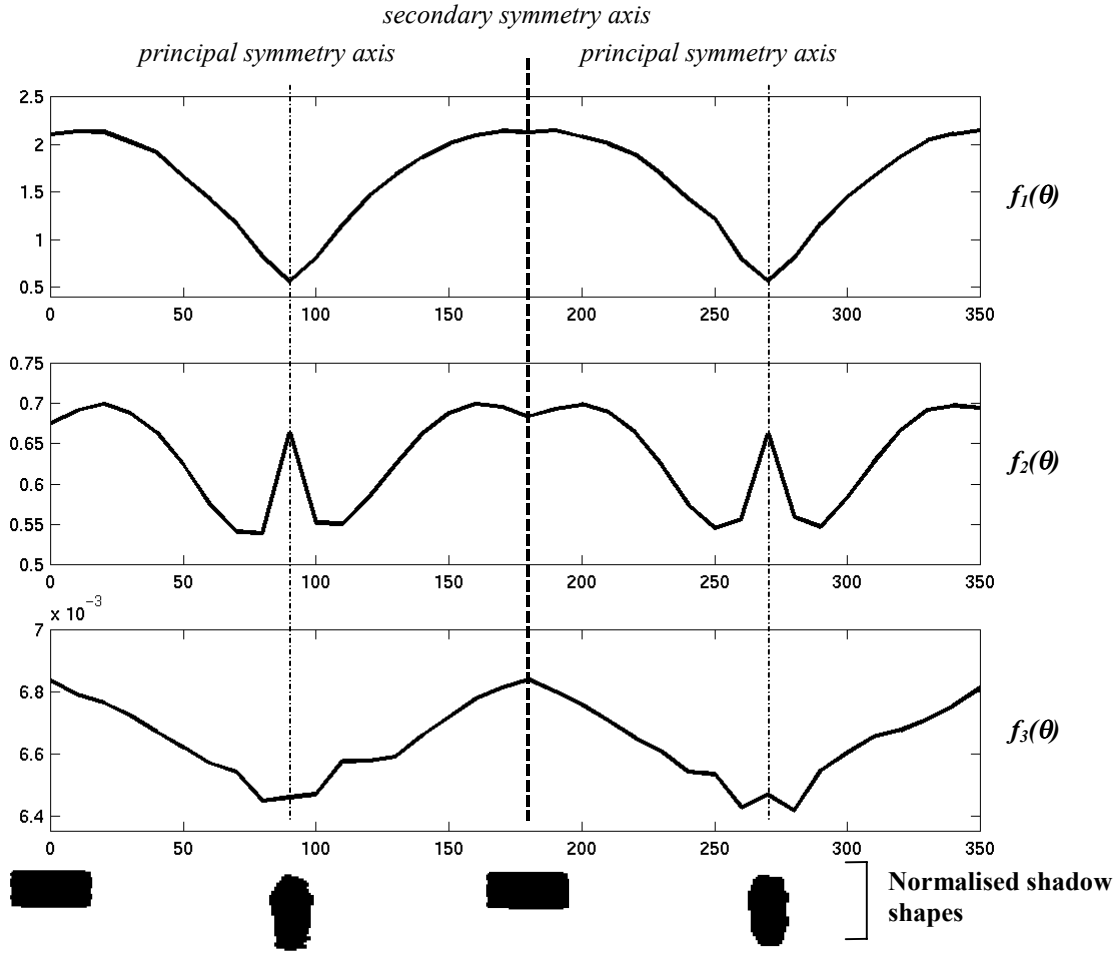

Fig.2. Geometrical properties

Given the geometric properties of the objects we have to classify, viewpoints from 0 to 180 degrees are sufficient to completely discriminate the different types of mines.

IV. Mutiview classification algorithm

Under the previous observations, we propose to compare experimental sequential features values to theoretical values related to each class and computed over a training set of 15 sequences of 19 images for each class. As far as objects with radial symmetry axis are concerned, features values are obviously constant. Fig.3 displays evolutions of these features for the five classes.

Our algorithm is based on correlation operations designed to measure the similarity between two features evolution curves.
φ is the correlation function defined as :

$$\varphi_{xy}(k) = \frac{2}{M+N}\sum_{l=0}^{M-1}\mathbf{x}(l)\mathbf{y}(l+k)$$

where **x** and **y** are two vectors of length *M* and *N*. This expression returns a vector of length *M+N-1*. φ is maximal for $k=k_{max}$ such that the similarity is the most important.

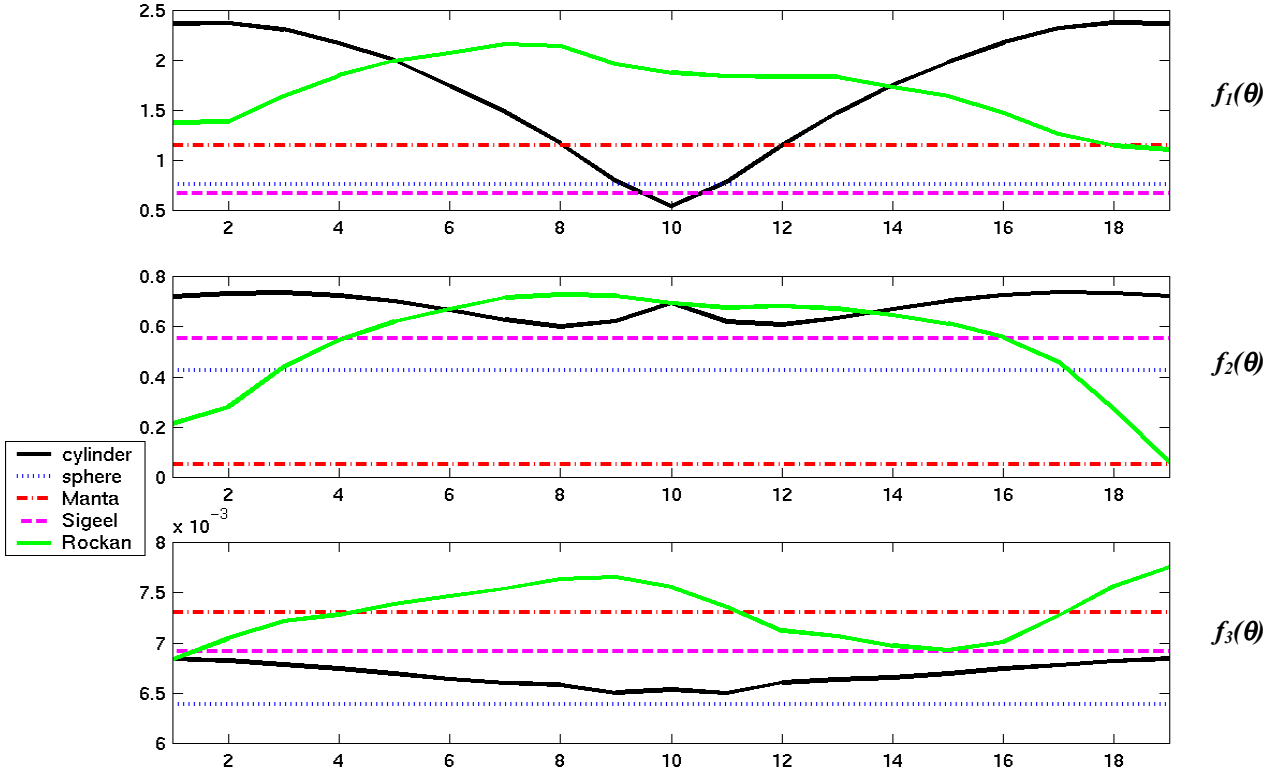

Fig.3. Examples of feature evolution curves.

The correlation coefficient ρ is used to measure the similarity between two functions x and y:

$$\rho_{xy}(k) = \frac{\varphi_{xy}(k)}{\sqrt{\varphi_{xx}(0)\varphi_{yy}(0)}} \leq 1$$

$\rho_{xy}(0)$ is equal to one if x and y are identical.

Of course, if we deal constant curves, ρ provides no interesting information. For instance, if $x(l) = x_0, \forall l$, $\rho_{xy}=\rho_y$ does not depend on x. In this particular case, a simpler comparison will be used between x_0 and the average value $\bar{y} = \frac{1}{N}\sum_{l=1}^{N} y(l)$.

The diagram of the Fig.4 give the main steps we develop hereafter.

In the following, Sf_i (respectively Sf_i^j) stands for the experimental (respectively theoretical) sequence related to the ith feature f_i.

Step 1. **Comparison** of sequence vectors $\mathbf{Sf_i}$ and $\mathbf{Sf_i^j}$ by means of cross-correlation functions $\varphi_{Sf_i^j Sf_i}$.

Selection of k_{max} such that $k_{max} = Arg\max_k \left[\sum_{i=1}^{L}\varphi_{Sf_i^j Sf_i}(k)\Big/L\right]$

Step 2. Computation of **correlation coefficients** $\rho_{Sf_i^j Sf_i}(k_{max}) = m(i,j)$

The resulting L×K matrix $\boldsymbol{M} = (m(i,j))_{(i,j)\in\{1\ldots L\}\times\{1\ldots K\}}$ contains all the maximal correlation coefficients for each feature between the experimental sequence and the theoretical one.

Step 3. **Test** about the symmetry of the observed object.
Theoretical sequential feature values of objects with radial symmetry are constant because their cast shadow appearance does not depend on the point of view. Consequently, for three classes namely sphere, Manta and Sigeel mine, correlation coefficient values are the same and only depend on experimental values describing the observed object. For these particular objects, correlation coefficient values are not discriminant.
For this reason, we have to do the following test before the final classification step. If at least half the maximal correlation coefficient values per feature do not depend on the class C_j, i.e. $\max_j m(i,j) = \rho_{Sf_i^j Sf_i} = \rho_{Sf_i}$, then we can assume that the observed object has radial symmetry.

Step 4. **Decision** from all the features (L=3 features)

- *First case:* radial symmetry object (sphere, Manta or Sigeel mine)

  Comparison occurs between average features of the observed object and constant theoretical ones (given that we only consider classes of objects with radial symmetry)

$$j = Arg\min_j \left[\frac{1}{L} \sum_{i=1}^{L} \left(\bar{f}_i - \bar{f}_i^j \right)^2 \right] \text{ where } \bar{f}_i^j = f_i^j(l), \forall l$$

- *Second case :* complex object (cylinder or Rockan mine)

  Correlation coefficient values are discriminant values and we can write

$$j = Arg\min_j \left[\frac{1}{L} \sum_{i=1}^{L} \left(1 - m^2(i,j) \right) \right]$$

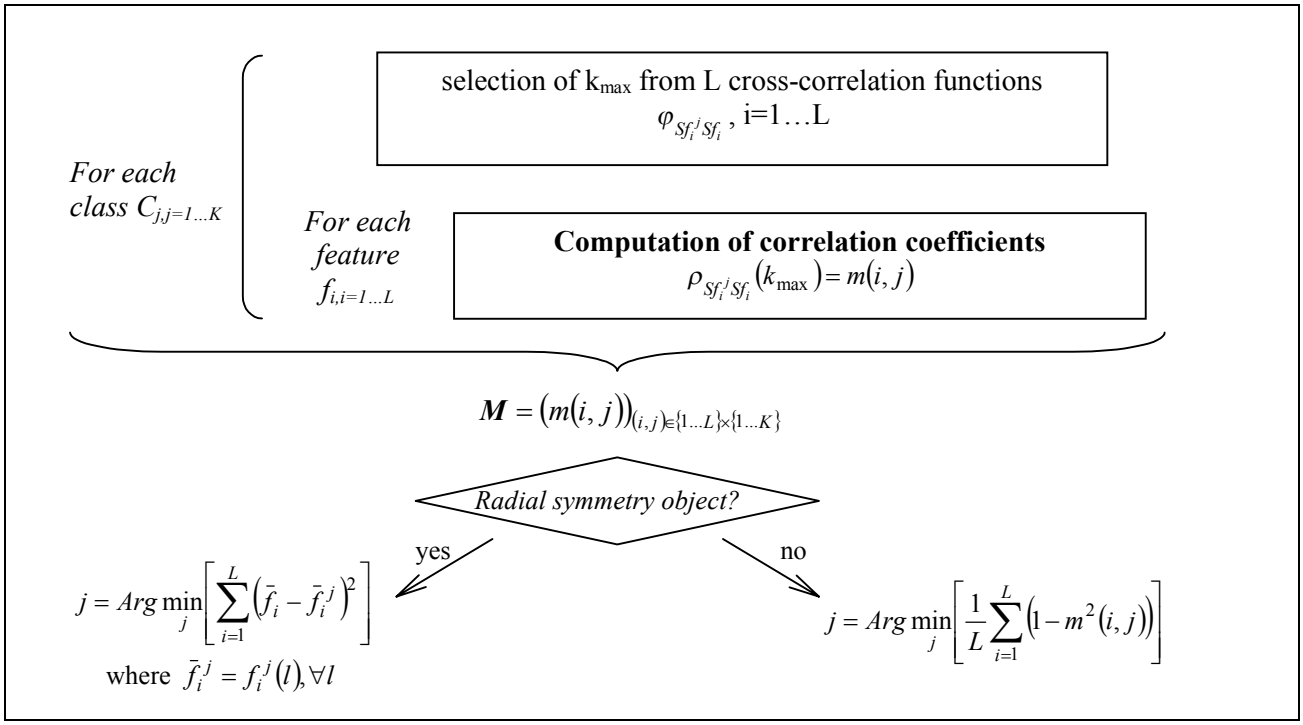

Fig.4. Multiview classification algorithm

V. Experimental results

This algorithm has been tested on nine different mines over 30 sequences per mine: three different cylinders, three different spheres, the Manta, the Sigeel and the Rockan mine (see paragraph III.2). Given that we make a supervised classification, the database is first divided to provide two sets: while learning examples whose classes are known enable us to define theoretical (average) features values, remaining examples (i.e. 15 sequences per mine) are classified using the proposed method. Results are summarised for each feature in confusion matrices obtained for two different length N of the experimental sequential feature values (that is the number of views within the covered sector). Fig.5 shows the evolution of the average recognition rate versus N increasing from 3 views to M=19 views.

It appears that complex mines (cylinder and Rockan mine) are the most sensitive to the number of views N. Nevertheless we observe a rate of good classification more than 90% if N is over 9.

Table I: confusion matrix when M=N=19 (respectively N=9, i.e. less than a quarter of the sequence) corresponding to an average rate of good classification of 99.4% (respectively 86.2%)

	C1	C2	C3	C4	C5
cylinders	**100 (68)**	0 (31)	0 (0.5)	0 (0)	0 (0.5)
spheres	0 (0)	**100 (100)**	0 (0)	0 (0)	0 (0)
Manta mines	2 (2)	0 (0)	**98 (88)**	0 (0)	0 (12)
Sigeel mines	1 (0)	0 (0)	0 (0)	**99 (100)**	0 (0)
Rockan mines	0 (24)	0 (0)	0 (1)	0 (0)	**100 (75)**

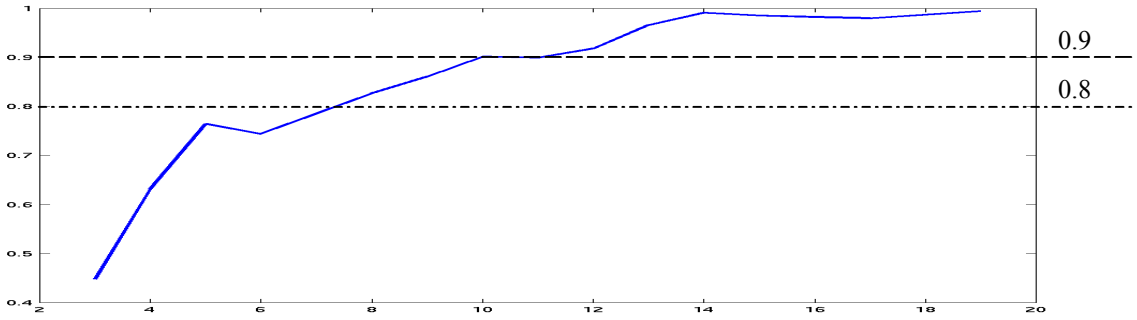

Fig.5. Rates of good classification versus N from 3 to 19

VI. Conclusion

In this paper, a multiview-based recognition method was described. Indeed, the shadow cast by a given object (objects with radial symmetry excepted) has different appearances with respect to the sonar observation viewpoint. As a consequence, different feature values computed from pixels belonging to the shadow shape can characterise the same object. Provided that we know the successive positions of the sonar for each image, sequential feature values can be stored in order to be compared to theoretical values. This comparison between the experimental and the theoretical curves is easily done by means of correlation operators. Our algorithm has been successfully applied to simulated sequences of sonar images for five classes of mines with different symmetry properties, and a set of three features chosen in such a way that they appear discriminant over a sequence of viewpoints. As future works, it would be interesting to test this flexible technique with other features and various objects.

References

(1) J. Wood; Invariant Pattern Recognition; 1996; Pattern Recognition, vol. 29, no. 1, pp.1-17.
(2) H. Kauppinen, T. Seppänen, M. Pietikäinen; An Experimental Comparison of Autoregressive and Fourier-based Descriptors in 2D Shape Classification; 1995; IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 17, no. 2, pp.201-207.
(3) A.P. Reeves, R.J. Prokop, S.E. Andrews and F.P. Kuhl; Three-Dimensional Shape Analysis Using Moments and Fourier Descriptors; 1988; IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 10, no. 6, pp.937-943.
(4) R.O. Duda and P.E. Hart; Pattern Classification and Scene Analysis; New York: John Wiley & Sons.
(5) S. Daniel, B. Solaiman and E.P. Maillard; Object recognition on the sea-bottom using possibility theory and sonar data fusion; 1998; Proceedings of the International Conference on Multisource-Multisensor Information Fusion, Las vegas, USA, Vol.1, pp. 396-402.
(6) I. Quidu, J.P. Malkasse, G. Burel and P. Vilbé; A 2-D Filter Specification for Sonar Image Thresholding; July 30 – August 3, 2001; to appear in Advanced Concepts for Intelligent Vision Systems (ACIVS'2001) conference, Baden-Baden, Germany.
(7) D. Jan ; Traitement d'images numériques appliqué aux sonars à haute résolution ; 1-5 juin 1987 ; 11ème colloque du GRETSI, Nice, France.
(8) J. Flusser and T. Suk ; Pattern recognition by affine moment invariants ; 1993 ; Pattern Recognition, Vol.26, No.1, pp.167-174.