

Mine Classification using a Hybrid Set of Descriptors

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-This paper is concerned with the problem of recognition of objects laying on the sea-bed. A high resolution sonar provides high-quality acoustic images of the sea-bed, allowing the classification of objects from their cast shadow. After the segmentation step, a set of features is extracted from the shadow. We propose an approach based on a hybrid set of descriptors, combining features of different origins. We first compute topological parameters: the extent and the elongation. In addition to these classical features, affine moment invariants seem suitable for sonar images. Indeed, under weak perspective conditions, the perspective transformation is well approximated by an affine transformation. A four-dimensional vector is then computed characterizing the shadow. The method has been tested on simulated sonar images.

I. INTRODUCTION

Pattern recognition from imagery may be accomplished by identifying an unknown object as a member of a set of well-known objects. Precisely the classification process includes two distinct and successive phases of feature extraction and feature classification [1]. In the context of mine identification, the question is to select a set of appropriate features with acceptable recognition accuracy. We focus on techniques based on binary images (segmented shadows) of preprocessed 2-dimensionnal intensity images (sonar images). Several important issues may be considered to improve shadow identification. As the underwater environment induces several noises and distortions on the shadow, robustness of features to noise is required [2]. Moreover, objects viewed under different acoustic and geometric conditions have a changing appearance. If invariance under translation, rotation and scaling of the shadow often gives satisfaction, the introduction of invariants under general affine transformations is necessary for complex objects. The feature extraction step is a process by which the initial binary images obtained from gray-level images are transformed into a pattern feature vector. Several kinds of features are used for recognition, especially geometrical features and statistical features [3] [4] [5] [6] [7]. Choosing four features between them, a hybrid set of descriptors may characterize the 2D shape i.e. the cast shadow.

The paper is organized as follows. An overview of the topological parameters is given in Section II. Section III introduces the concept of affine transformation and affine moment invariants. Section IV specifies the preliminary steps. Performances of the proposed features are exhibited in Section V. Finally, the conclusion of our study is given in Section VI.

II. TOPOLOGICAL PARAMETERS

Among the geometrical characteristics including area, parameter or dimensions along a given direction, topological parameters characterize global shadows. Three of them are widely used : the compactness, the extent and the elongation. The two last are based on second order moments known as moments of inertia and are independent of scale and orientation. The compactness is more sensitive to noise

because it depends on the perimeter of the shadow. Indeed, it is more robust to take into account the information content of the whole plane shape than of its boundary.

Let's define the *extent* and the *elongation*. Moments and functions of moments capture global information about the shadow. An introduction to moment theory and properties can be found in appendix. Moment values of the distribution $I(i,j)$ that is binary and contiguous, i.e. pixels of a segmented image, may be easily explained in terms of geometric characteristics like area, center of mass or principal axes [8]. In addition, the extent and the elongation are computed from second order central moments [3]:

$$\left\{ \begin{array}{l} \text{extent} = (\mu_{20} + \mu_{02}) / \mu_{00}^2 \\ \text{elongation} = \sqrt{4\mu_{11}^2 + (\mu_{20} - \mu_{02})^2} / (\mu_{20} + \mu_{02}) \end{array} \right.$$

where μ_{pq} stands for the central moment of order p+q (see Appendix).

III. STATISTICAL PARAMETERS

Defining the general affine transformation $\begin{cases} u = a_0 + a_1x + a_2y \\ v = b_0 + b_1x + b_2y \end{cases}$,

Flusser et al. decomposed it into six one-parameter transformations [9]:

$$\begin{array}{l} \underline{\text{Translation}} \quad \begin{cases} u = x + \alpha \\ v = y \end{cases} \quad \text{and} \quad \begin{cases} u = x \\ v = y + \beta \end{cases} \\ \underline{\text{Scaling}} \quad \begin{cases} u = \omega x \\ v = \omega y \end{cases} \\ \underline{\text{One-axis scaling}} \quad \begin{cases} u = \delta x \\ v = y \end{cases} \\ \underline{\text{Skew transformation}} \quad \begin{cases} u = x + ty \\ v = y \end{cases} \quad \text{and} \quad \begin{cases} u = x \\ v = t'x + y \end{cases} \end{array}$$

The idea is to extract moment invariants, i.e. algebraic functions of moments (see Appendix) that are invariant under the six elementary transformations of the shadow.

Beside the well-known property of invariance under *translation* of the central moments, they derived a function F invariant under each of the transformations above. F is a polynomial of central moments which is constrained in order to ensure the invariance under the three other transformations. For *scaling*, a normalization of members of the polynomial F is computed. Then the sum of p th indexes of each member must be equal to the sum of q th indexes to ensure invariance under *one-axis scaling*. Lastly the derivative of F with respect to t and t' must be equal to zero concerning the *skew transformation*.

Finally F may be written as :

$$F = \left(\sum_i k_i \mu_{p_1(i)q_1(i)} \dots \mu_{p_{c(i)}q_{c(i)}} \right) / \mu_{00}^{z(i)}$$

with $c(i)$ the number of central moments for the i^{th} member,

$$z(i) = \left(\sum_{j=1}^{c(i)} (p_j(i) + q_j(i)) \right) / 2 + c(i)$$

$$\text{and } \sum_{j=1}^{c(i)} p_j(i) = \sum_{j=1}^{c(i)} q_j(i).$$

Flusser et al. derived a complete set of four polynomials of second and third-order moments [9]. According to some preliminary tests we kept two of them I_1 and I_2 defined as :

$$\begin{cases} I_1 = (\mu_{20}\mu_{02} - \mu_{11}^2) / \mu_{00}^4 \\ I_2 = (\mu_{30}\mu_{03}^2 - 6\mu_{30}\mu_{21}\mu_{12}\mu_{03} + \\ 4\mu_{30}\mu_{12}^3 + 4\mu_{21}^3\mu_{03} - 3\mu_{21}^2\mu_{12}^2) / \mu_{00}^{10}. \end{cases}$$

IV. PRELIMINARY STEPS

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. 2 and Fig. 3.

A. Segmentation step

Segmentation consists in partitioning the image into homogeneous regions. In our case, objects are classified from their cast shadow. Each image presents an echo (more or less important) corresponding to the detected object and the cast shadow on the sea-bed. Giving the label zero for pixels belonging to the shadow and the label one elsewhere we obtain binary from graylevel image. Based on the knowledge of the image formation related to the spatial sampling rate (along-range and across-range), a filter may be synthesized allowing to make the histogram of the filtered image bimodal [10]. This permits to perform image thresholding.

B. Noise reduction

Irregularities of the outer boundary of the shadow may have undesired effects on the recognition system. While preserving the global information of the shadow, we aim to smooth 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 new shadow is smoother than the original one.

C. Image normalization

To improve robustness of topological features, an image normalization is performed. It has to provide a new image as it would be seen under a grazing angle of 45 degrees preserving shadow ratios. The normalized length given this angle is computed as shown in Fig. 1. Moreover on account of the sonar parameters, image resolution is generally different along the two dimensions. To form the image each pixel has to be approximately square to prevent from disproportions.

L_i : length of the shadow on the sea-bed,
 R_i : slant range from sonar to object,
 θ_i : grazing angle,
 h : height of the object,
 H_i : altitude of the sonar,

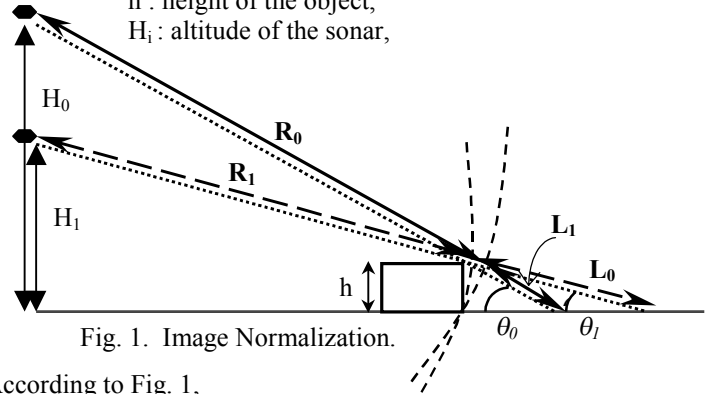


Fig. 1. Image Normalization.

According to Fig. 1,

$$\sin \theta_0 = \frac{H_0}{R_0 + L_0} \quad \text{and} \quad \sin \theta_1 = \frac{H_1}{R_1 + L_1},$$

$$\frac{H_0}{h} = \frac{R_0 + L_0}{L_0} \quad \text{and} \quad \frac{H_1}{h} = \frac{R_1 + L_1}{L_1},$$

$$\text{then } \frac{\sin \theta_0}{\sin \theta_1} = \frac{H_0}{H_1} \frac{R_1 + L_1}{R_0 + L_0} = \frac{L_1}{L_0},$$

$$\text{the normalized distance is } L_0 = L_1 \frac{1}{\sin \theta_0} \frac{H_1}{R_1 + L_1}.$$

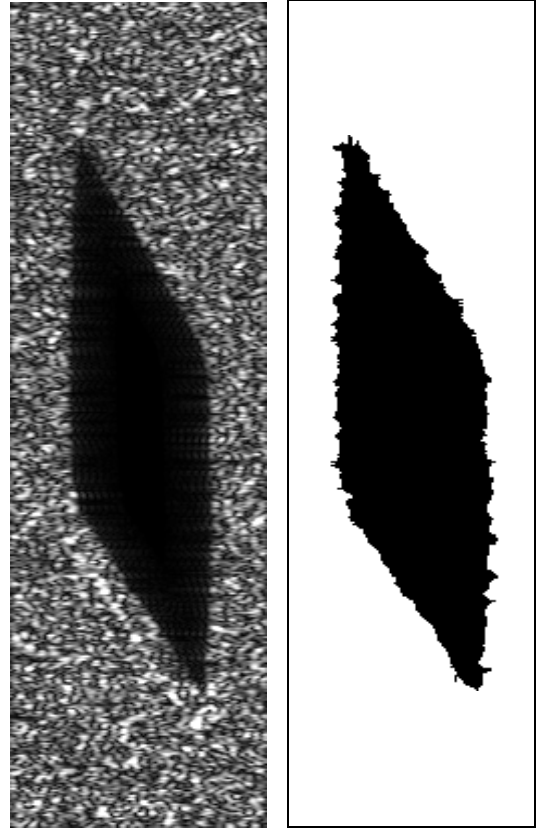


Fig. 2. Sonar image and binary image.

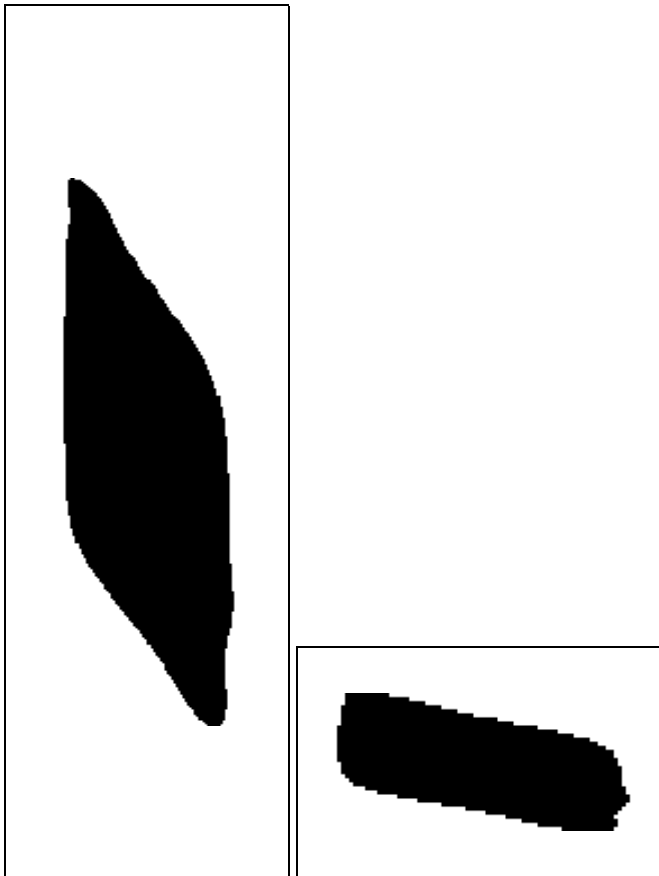


Fig. 3. Smoothed binary image and final normalized image.

V. EXPERIMENTS

A. Training Set

Sonar data are simulated in order to cover a wide set of configurations in terms of types of mines and appearances from different points of view. On one hand, moored mines generally look like a sphere coupled to an anchor box for laying. On the other hand, ground mines are usually cylindrical in shape. Nevertheless while cylindrical shaped objects reflect fairly definite sonar shadows, some ground mines have specific shape which improves stealth capabilities resulting in a difficult identification [11]. In our experiments, four classes have then been considered: cylinders, spheres, and two stealthy mines which look like truncated cones (the Manta and Sigeel mines). Simulating a circular path of the sonar around each mine with a shot every 10 degrees, we simulate sequences of 10 points of view for each mine except for the cylinder which requires a entire semicircular path because of its two axes of symmetry. Indeed poor symmetrical properties make shadows different under different points of view. Finally 300 images have been generated for each class, so the training set is made of 1200 patterns. Fig. 4 shows some shadows (taken under an approximate grazing angle of 10 degrees) to be recognized, namely a sphere, a cylinder, a Manta mine and a Sigeel mine. For an easier visualisation, Fig. 5 gives the corresponding

normalized images (see paragraph IV.C.) which are more representative of the proportions of the detected objects. Remember that this normalization really occurs after the segmentation step (see paragraph IV.).

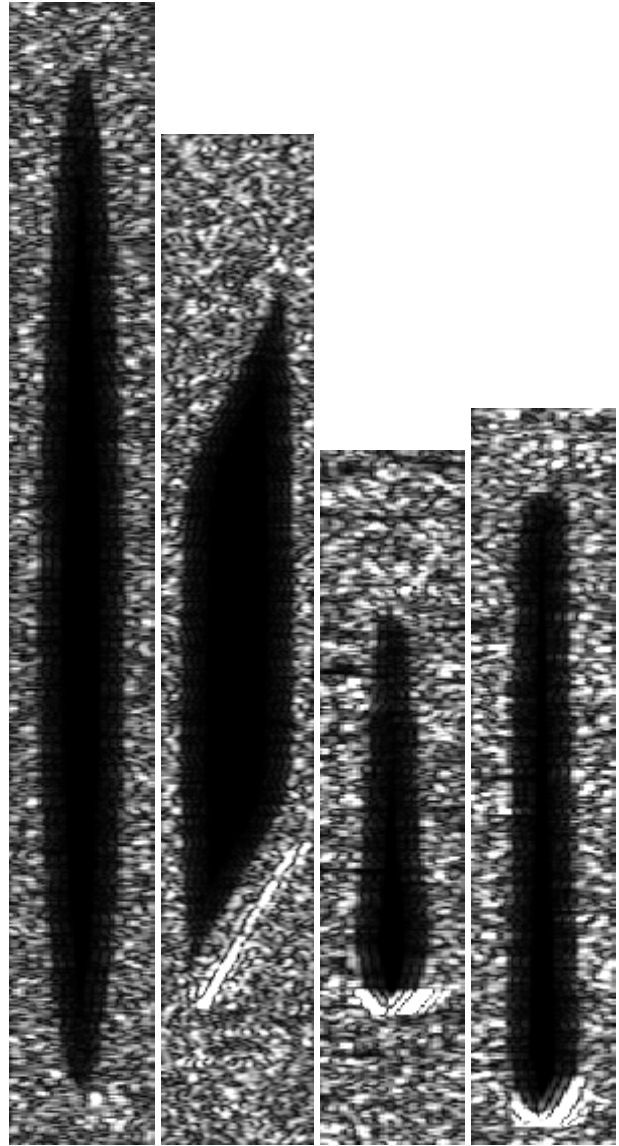
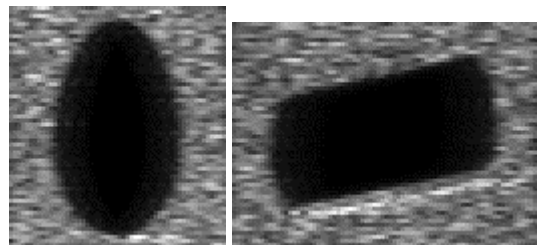


Fig. 4. Examples of shadows: sphere, cylinder, Manta mine and Sigeel mine.



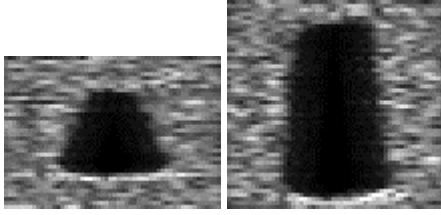


Fig. 5. The corresponding normalized shadows.

B. Experimental Results

a) Features analysis

Features must not be redundant but, on the contrary, must vary according to the objects, in other words to be differently distributed. In Fig. 6 the histograms of the four classes for each feature are plotted (the interval $[\min, \max]$ of a given feature is divided by 50). Considering a topological parameter and a statistical parameter, classes are not ordered in the same way along the interval of the values. This provides a good property for classification.

In order to show that both kind of features are complementary, we compare the performances of three pairs of them: (I_1, I_2) , $(\text{extent}, \text{elongation})$ and a hybrid pair $(I_1, \text{elongation})$. Their performances are compared according to the following criterion: the class k , $k \in \{1, \dots, 4\}$, is assigned to the pattern i if

$$k = \text{Arg min}_{k \in \{1, \dots, 4\}} \left\{ \sum_j \frac{(a_j(i) - \bar{a}_j(k))^2}{\sigma_j^2(k)} \right\}, \text{ where } \bar{a}_j(k) \text{ is the}$$

mean value of the feature a_j for the class k , and $\sigma_j^2(k)$ the associated variance. For each index k , we have the correspondence below:

- $k=1$ is the class of cylinders,
- $k=2$ is the class of spheres,
- $k=3$ is the class of Manta mines,
- $k=4$ is the class of Sigeel mines,

and for index j ,

- $j=1$ stands for the feature I_1 ,
- $j=2$ stands for the feature I_2 ,
- $j=3$ stands for the feature *extent*,
- $j=4$ stands for the feature *elongation*.

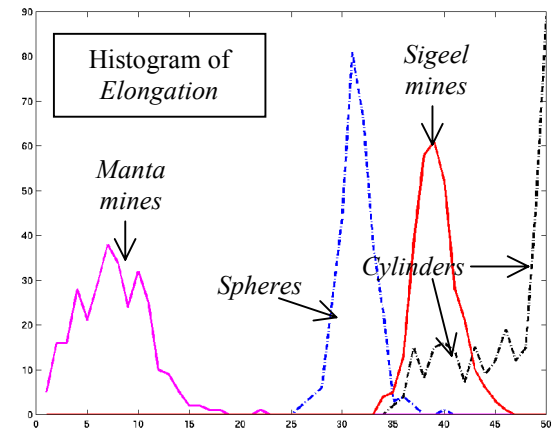
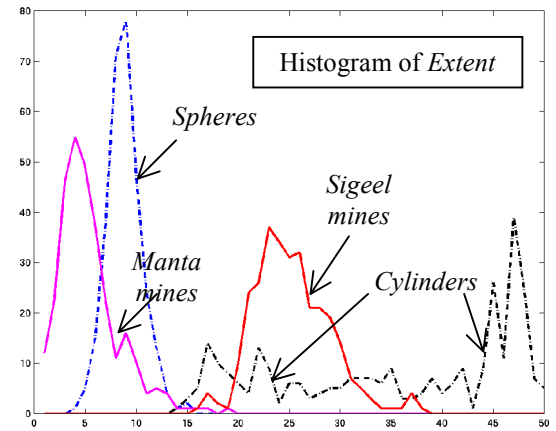
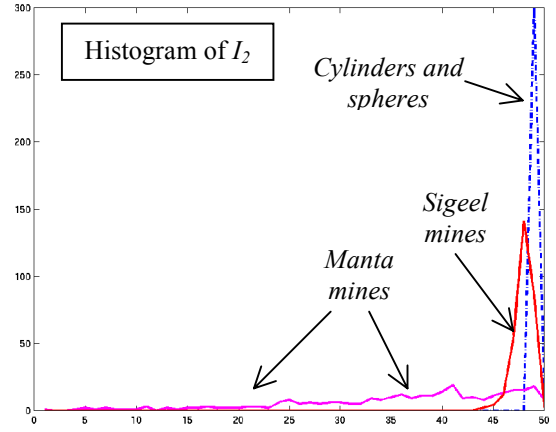
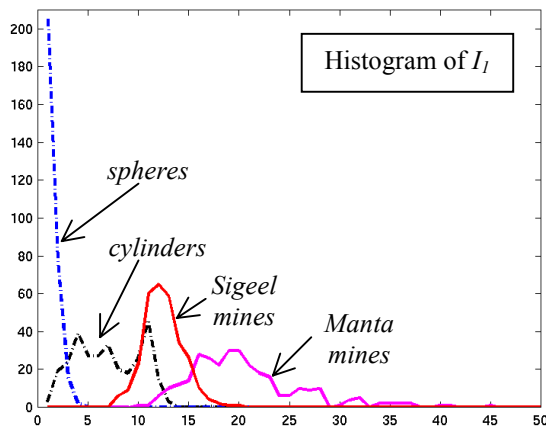


Fig. 6. Histograms.

The 4×4 pairs of statistic values $(\bar{a}_j(k), \sigma_j^2(k))_{(j,k) \in \{1, \dots, 4\} \times \{1, \dots, 4\}}$ are calculated from a training set of 600 measures while classification is performed on the 600 remaining measures. Table I gives the percentage of classification results. It shows that the hybrid pair is the most effective. For information, performance of the whole set of features is given in the last column.

Table I
Performances in percents.

| | I_1, I_2 | extent, elongation | I_1 , elongation | I_1, I_2 , extent, elongation |
|--------------|-------------|-----------------------|-----------------------|------------------------------------|
| cylinder | 89.3 | 80.0 | 100 | 100 |
| sphere | 89.3 | 97.3 | 98.7 | 98.7 |
| Manta mine | 99.3 | 100 | 100 | 100 |
| Sigeel mine | 82.7 | 76.7 | 88.0 | 100 |
| total | 90.2 | 88.5 | 96.7 | 99.7 |

b) Principal Component Analysis

In order to reduce the problem size while preserving the maximum of variance, the technique of Principal Component Analysis (PCA) has been applied. It consists in finding the principal axes of inertia through the feature space and projecting data along these axes. From the set of the 4 feature measures, we compute the corresponding centered and reduced measures allowing feature measures using different units to be compared and combined. For a given measure $a_j(k)$, we compute the following operation:

$$a_j^{new}(k) = (a_j^{old}(k) - \bar{a}_j) / \sigma_{a_j}$$

where \bar{a}_j is the mean and σ_{a_j} is the standard deviation for that feature. In the remainder, we call *individual* a given shadow described by its four features $(a_j)_{j \in \{1..4\}}$.

A covariance matrix $1/1200 \times CC^T$ of size 4×4 is then constructed from the set of the new measures C of size 4×1200 (T stands for the transpose operation). The principal axes through the feature space are the eigenvectors of this matrix associated to its eigenvalues. We can then visualize the best clustering in the feature space projecting the feature measures on the base of the eigenvectors U .

Clearly, given the matrix of feature measures

$$C = \begin{pmatrix} (I_1)_1 & \dots & (I_1)_{1200} \\ (I_2)_1 & \dots & (I_2)_{1200} \\ (extent)_1 & \dots & (extent)_{1200} \\ (elongation)_1 & \dots & (elongation)_{1200} \end{pmatrix},$$

← 1200 measures →

the eigenvalues of the corresponding covariance matrix are $\lambda_1=2.60077$, $\lambda_2=0.989262$, $\lambda_3=0.376325$ et $\lambda_4=0.033642$, and U is written as

$$U = \begin{pmatrix} V_{11} & V_{21} & V_{31} & V_{41} \\ V_{12} & V_{22} & V_{32} & V_{42} \\ V_{13} & V_{23} & V_{33} & V_{43} \\ V_{14} & V_{24} & V_{34} & V_{44} \end{pmatrix} = \begin{bmatrix} -0.42 & 0.67 & 0.52 & 0.32 \\ 0.51 & -0.27 & 0.81 & -0.10 \\ 0.46 & 0.67 & -0.14 & -0.57 \\ 0.60 & 0.19 & -0.23 & 0.75 \end{bmatrix}$$

where $V_i = (V_{ij})_{j \in \{1..4\}}$ stands for the i^{th} eigenvector.

As the measures are centered and reduced, the total inertia I of the cluster of the whole individuals, i.e. the sum of the eigenvalues, is equal to the number of features, i.e. 4. Given the ratio $(\lambda_1 + \lambda_2)/I = 0.8975$, i.e. the part of the inertia related to

the two first principal axes, we can sum up the measures by the only two first components and reduce the size of the feature space from four to only two dimensions.

Doing the projection $F = U^T C$, we visualize the 2D-subspace $f_1 = f(f_2)$ on Fig. 7, where f_1 and f_2 are the two first lines of F and are related to the coordinates of the individuals in the principal plane. On this figure, farther an individual is from the origin of the axes (i.e. the mean principal component), easier it is to be classified. Actually, the principal plane is the best projection of the initial 4D-space when the four features $(a_j)_{j=1..4}$ are combined to compute two new synthetic features, i.e. the principal components f_1 and f_2 . To give an interpretation of the positions of the individuals in this plane, one can use the correlation circle of unit radius in Fig. 8. On this graph, coordinates of each point stand for the correlation of each feature a_j with the two principal components f_1 and f_2 . For our centered and reduced measures, correlation is equal to $r(f_i, a_j) = V_{ij} \times \sqrt{\lambda_i}$ [12]. The correlation circle creates a link between the two spaces by means of correlations.

The part of the inertia related to the two first principal axes allows us to extract some informations from the principal plane given the correlation circle. In other words, considering the mean features $(\bar{a}_j)_{j=1..4}$ and $(\bar{f}_j)_{j=1,2}$, we can write that

$\bar{A}^R = (\bar{a}_j^R(k))$ is an approximation of $\bar{A} = (\bar{a}_j(k))$ where

$$\bar{A}^R = \begin{pmatrix} \bar{a}_1^R(1) & \bar{a}_2^R(1) & \bar{a}_3^R(1) & \bar{a}_4^R(1) \\ \bar{a}_1^R(2) & \bar{a}_2^R(2) & \bar{a}_3^R(2) & \bar{a}_4^R(2) \\ \bar{a}_1^R(3) & \bar{a}_2^R(3) & \bar{a}_3^R(3) & \bar{a}_4^R(3) \\ \bar{a}_1^R(4) & \bar{a}_2^R(4) & \bar{a}_3^R(4) & \bar{a}_4^R(4) \end{pmatrix} = \begin{pmatrix} \bar{f}_1(1) & \bar{f}_2(1) \\ \bar{f}_1(2) & \bar{f}_2(2) \\ \bar{f}_1(3) & \bar{f}_2(3) \\ \bar{f}_1(4) & \bar{f}_2(4) \end{pmatrix} \times \begin{bmatrix} V_{11} & V_{12} & V_{13} & V_{14} \\ V_{21} & V_{22} & V_{23} & V_{24} \end{bmatrix}.$$

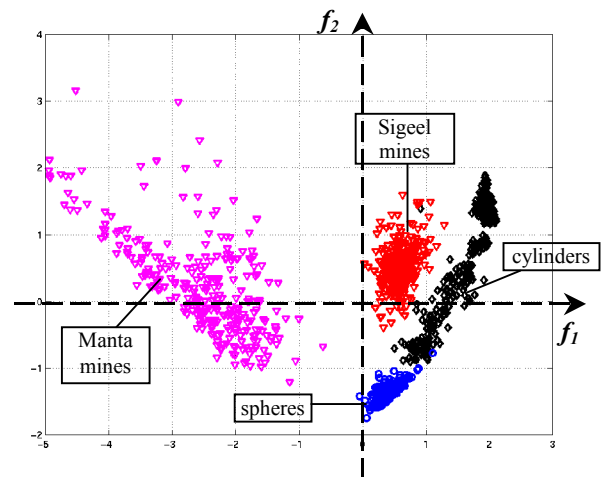


Fig. 7. Partition of the subspace.

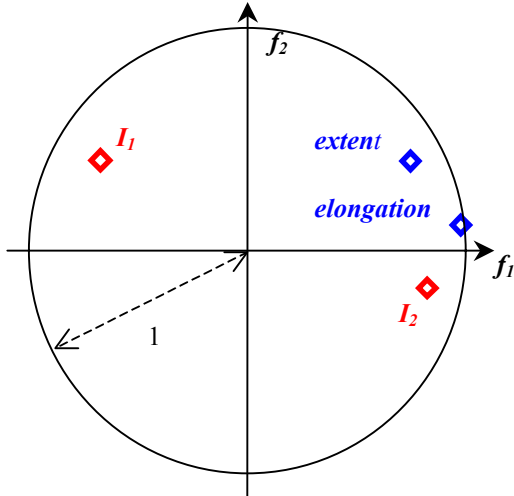


Fig. 8. Correlation circle.

 Table II
 Approximated (and exact).mean features.

| | \bar{a}_1^R (\bar{a}_1) | \bar{a}_2^R (\bar{a}_2) | \bar{a}_3^R (\bar{a}_3) | \bar{a}_4^R (\bar{a}_4) |
|---------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|
| <i>cylinders</i> | -0.25 (-0.42) | 0.64 (0.46) | 1.11 (1.22) | 1.04 (1.00) |
| <i>spheres</i> | -1.08 (-1.14) | 0.57 (0.46) | -0.73 (-0.72) | -0.02 (0.03) |
| <i>Manta mines</i> | 1.23 (1.30) | -1.38 (-1.23) | -1.00 (-0.97) | -1.48 (-1.59) |
| <i>Sigeel mines</i> | 0.10 (0.27) | 0.16 (0.30) | 0.61 (0.47) | 0.45 (0.56) |

Table II gives the approximated mean features $(\bar{a}_j^R)_{j=1\dots 4}$ and the exact ones $(\bar{a}_j)_{j=1\dots 4}$ in parenthesis. Bold (resp. bold and italic) data is related to the largest positive (resp. negative) values for a given feature. These values justify the positions of the clusters related to the four classes in the principal plane. Note particularly that the cluster of Manta mines deviates from the origin because of a large positive value of I_1 and large negative values of the other features that is consistent with their positions on the correlation circle.

 Table III
 Results in percents.

| | Classifier 1 | | Classifier 2 | |
|--------------|------------------------------|-------------|------------------------------|-------------|
| | I_1 , <i>elongation</i> | f_1, f_2 | I_1 , <i>elongation</i> | f_1, f_2 |
| cylinder | 100 | 100 | 81.3 | 98.7 |
| sphere | 98.7 | 97.3 | 96.7 | 100 |
| Manta mine | 100 | 100 | 100 | 100 |
| Sigeel mine | 88.0 | 88.7 | 90.0 | 100 |
| total | 96.7 | 96.5 | 92.0 | 99.7 |

c) Results

Two different classifiers have been used for experiments: the classifier defined in paragraph V.B.a. (classifier 1), and the well-known K-nearest neighbor with K=1 (classifier 2). The results (percentage of classification) are summarized in Table III when each shadow is only characterized by two features. Improvement is clearly observed after PCA using the second classifier.

VI. CONCLUSION

It is generally difficult to make comparison of different recognition systems, even for the same problem, since different test sets are used to evaluate performance. Moreover simulated sonar images refer to a specific environment joint with a particular sonar. Nevertheless given a set of four existing mines simulated in a specific case of mine warfare using current sonar specifications we succeed in pointing out the advantage of using a hybrid set of features.

In addition to the widely used topological parameters, we were interested into moments of shadow. Some particular functions of moments have the useful property of affine invariance. So they are well-suited for cast shadows viewed under weak perspective conditions. Thanks to their different origins, these four features appeared complementary for classifying both classical and stealthy mines. Performances included in this paper are encouraging especially doing a principal component analysis. In that case promising results were obtained in a restricted 2-dimensionnal space.

APPENDIX

A two-dimensional density distribution function $I(i,j)$ is provided from gray-level image giving the value one for pixels belonging to the shadow and the value zero elsewhere. Moments may be used to characterize a shadow and to extract properties from the spatial distribution of mass. In this appendix, we only focus on two dimensional moments for a $M \times N$ discretized image.

- Moments : definition and properties

The two-dimensional moment m_{pq} of order $p+q$ is defined as $m_{pq} = \sum_{i=0}^M \sum_{j=0}^N i^p j^q I(i,j)$.

The low-order moments represent geometric properties of the 2D distribution illustrating simple shape characteristics of the object:

Area - zeroth order moment represents the total mass of the image $m_{00} = \sum_{i=0}^M \sum_{j=0}^N I(i,j)$

Center of mass - the two first order moments locate the center of mass of the object whose coordinates are $i_g = m_{10}/m_{00}$ and $j_g = m_{01}/m_{00}$

Size and orientation of the shadow - second order moments are known as moments of inertia. Useful object

features may be determine as principal axes, approximation of the original image to an ellipse.

- **Central moments**

Let us make (i_g, j_g) the position of the center of mass of the shadow. Central moments μ_{pq} of order $p+q$ are computed as

$$\mu_{pq} = \sum_{i=0}^M \sum_{j=0}^N (i - i_g)^p (j - j_g)^q I(i, j).$$

REFERENCES

- [1] J. Wood, "Invariant pattern recognition: a review", *Pattern Recognition*, Vol. 29, No. 1, pp1-17, 1996.
- [2] D. Jan and J. Minot, "Les traitements d'image en sonar lateral", *L'onde électrique*, Vol.69, No. 3, pp.13-19, Mai-Juin 1989.
- [3] D. Jan, "Traitement d'images numériques appliqué aux sonars à haute résolution", *11^{ème} colloque du GRETSI*, Nice, 1-5 juin 1987.
- [4] J.C. Delvigne, "Shadow classification using neural networks", *Undersea Defence Technology*, London, 1992.
- [5] D.M. Lane and J.P. Stoner, "Automatic interpretation of sonar imagery using qualitative feature matching", *IEEE Journal of Oceanic Engineering*, Vol. 19, No. 3, pp.391-405, July 1994.
- [6] S.A. Dudani, K.J. Breeding and R.B. McGhee, "Aircraft identification by moment invariants", *IEEE Transactions on Computer*, Vol.C-26, No. 1, pp.39-46, January 1977.
- [7] P. Pejnovic, Lj. Buturovic and Z. Stojiljkovic, "Object recognition by invariants", *International Conference on Pattern Recognition*, La Hague, September 1992.
- [8] R.G. Prokop and A.P. Reeves, "A survey of moment-based techniques for unoccluded object representation and recognition," *CVGIP: Graphical Models and Image Processing*, vol. 54, No. 2, pp. 438-460, September 1992.
- [9] J. Flusser and T. Suk, "Pattern recognition by affine moment invariants", *Pattern Recognition*, Vol.26, No.1, pp.167-174, 1993.
- [10] I. Quidu, J.Ph. Malkasse, G. Burel and P. Vilbé, "Spécification d'un filtre optimal pour le seuillage d'une image sonar", *Proceedings of Forum Jeunes Chercheurs du GDR ISIS*, pp.30-31, 16-17 Mai 2000.
- [11] *Jane's Underwater Warfare Systems*, Ed. By Anthony J. Watts, Tenth Edition 1998-1999.
- [12] J.M. Bourroche and G. Saporta, *L'analyse des données*, Collection «Que sais-je ?», Presses Universitaires de France, 1998.