

## SHIP CLASSIFICATION FROM MULTISPECTRAL VIDEOS

FRÉDÉRIQUE ROBERT-INACIO<sup>✉,1,2</sup>, GHISLAIN OUDINET<sup>2</sup> AND FRANÇOIS-MARIE COLONNA<sup>2,3</sup>

<sup>1</sup>CNRS Institut Matériaux Microélectronique et Nanosciences de Provence, (IM2NP UMR 7334), France;

<sup>2</sup>Institut Supérieur de l'Electronique et du Numérique de Toulon, Place Pompidou, 83000 Toulon, France;

<sup>3</sup>CNRS Laboratoire des Sciences de l'Information et des Systèmes (LSIS UMR 7296), France

e-mail: [frederique.robert,ghislain.oudinet,francois-marie.colonna]@isen.fr

(Received February 1, 2012; revised May 12, 2012; accepted May 30, 2012)

### ABSTRACT

Surveillance of a seaport can be achieved by different means: radar, sonar, cameras, radio communications and so on. Such a surveillance aims, on the one hand, to manage cargo and tanker traffic, and, on the other hand, to prevent terrorist attacks in sensitive areas. In this paper an application to video-surveillance of a seaport entrance is presented, and more particularly, the different steps enabling to classify mobile shapes. This classification is based on a parameter measuring the similarity degree between the shape under study and a set of reference shapes. The classification result describes the considered mobile in terms of shape and speed, as speed is determined by target tracking.

Keywords: pattern recognition, ship classification, similarity parameter, video surveillance.

### INTRODUCTION

Video-surveillance is of increasing importance in our every-day life. For example, crime prevention and prevalence have been initialized during the 1970s and 1980s by experiments aiming to increase security in banks. The United Kingdom is one of the European countries which is the most under surveillance, and London more specifically is the best-surveyed town. Video-surveillance has been extended to streets, parking lots, airports (Yi and Marshall, 2000; Besada *et al.*, 2005) and finally to every location dealing with crowds or traffic (Reulke *et al.*, 2007). Being faced with such an amount of video data induces to find automatised processing procedures in order to get intelligent cameras. Human behavior is now analyzed in order to determine if people are behaving peculiarly. More particularly human motion analysis is of a great help to distinguish normal behaviors from others (peculiar, even threatening) (Moeslung and Granum, 2001; Wang and Suter, 2006). It can also be combined with face recognition for people identification (Kale and Roychowdhury, 2004; Zhou and Bhanu, 2008). However this behavior analysis has been extended to any moving object in order to detect intrusions in sensitive areas.

In the case of seaport surveillance (Zhu *et al.*, 2010) different features can be clues to behavior analysis. For example, a tanker with a too high speed must be suspected. As well a non-identified small boat approaching rapidly the coast must be intercepted before something could be thrown overboard or before

somebody could go ashore. That is why every mobile must be classified in terms of shape and speed in order to primarily analyze its behavior.

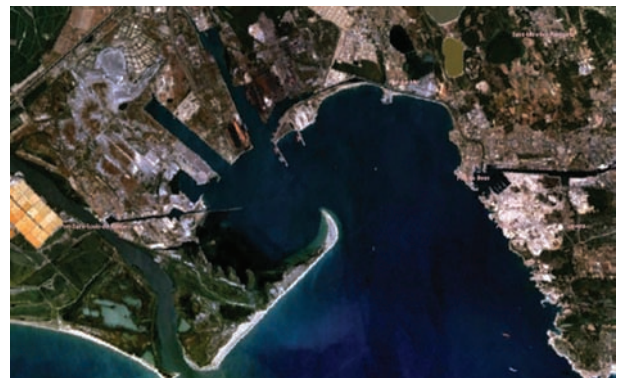


Fig. 1. *Seaport entrance.*

In this paper targets are first detected on the sea surface before being classified. Shape classification of 3D objects from video images raises the problem of bias induced by 2D projection. Actually the pattern recognition step must integrate some parameters such as the camera axis and position, the object trajectory if it moves, and so on. In this way, several transformations must be carried out on images in order to correct the mobile silhouette before the classification step. These corrections can be achieved in 2D or in 3D and followed by a projection.

Several data inputs must be known to supplement the videos: camera position and axis, mobile course, and so on. By combining this supplementary data the

transformations are defined and then applied to the mobile in order to get a profile as close as possible to the real one. Once the profile is computed the classification step begins by a contour detection and vectorization in order to deal only with vector objects. Then the profile under study is compared to several reference profiles in order to determine the closest one in terms of shape and size. Finally, examples are given in order to illustrate the whole process.

## APPLICATION CONTEXT

The surveillance system is set at the entrance of a trading seaport (Fig. 1). Shipping traffic includes cargo vessels and tankers as well as pleasure or fishing boats. A pylon (Fig. 2) supports the optronic sensor systems described below and is located at the end of the dike (Fig. 3). The whole system also includes a radar and a sonar.



Fig. 2. Pylon and optronic sensors.

The software system includes a module carrying out shape classification from infrared and daylight videos sent by two different sensor systems (Fig. 4). On the one hand, the surveillance system is composed of six fixed cameras (three for infrared and three for videos in visible wavelengths) each with a field of view of about  $20^\circ$ , providing a panoramic field of view of about  $60^\circ$ , of the seaport entrance in order to detect object presence on the sea, and more particularly those of moving objects. On the other hand, the tracking system includes two servo-controlled cameras (one for infrared and one for videos in visible range) with a

narrower angle (about  $5^\circ$ ). This second system is able to focus on particular objects and can track them not only in the area under survey but also outside it, as the tracking system can rotate through  $360^\circ$ . Shape classification is achieved at different levels (level 1: coarse classification, level 2: accurate classification, level 3: identification) and aims to determine what kind of mobile is captured on videos in order to extract features such as ship category (tanker, speedboat and so on), ship state (for example: sailing dinghy with shortened sail or not), etc. These features are then simultaneously displayed on the shape classification module monitor (Fig. 4) and sent to the Shore Center via a local network. The whole system provides a computer-aided tool helping the operator to make his decision (Fig. 5: displayed information).

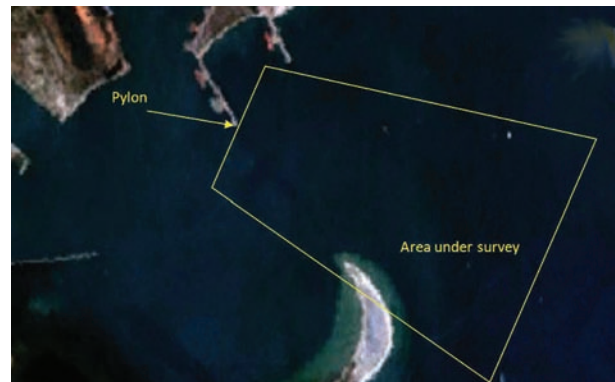


Fig. 3. Seaport entrance and video-surveillance system location.

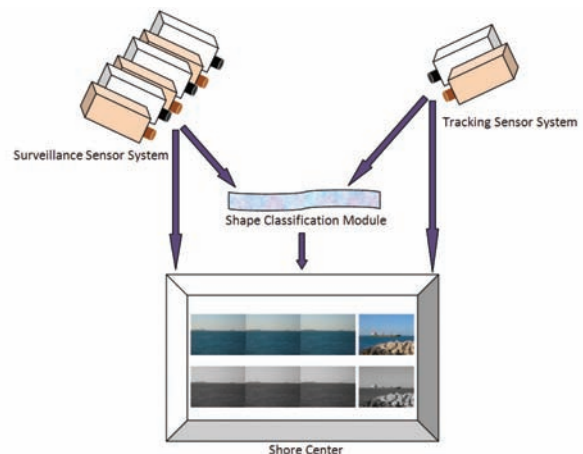


Fig. 4. Surveillance systems: sensor systems, shape classification system and shore center.

The main goal of this study is to secure a seaport by providing a video surveillance in order to protect it from terrorist attacks. The attacks under consideration are restricted to attacks coming from the sea and from objects moving at the surface.

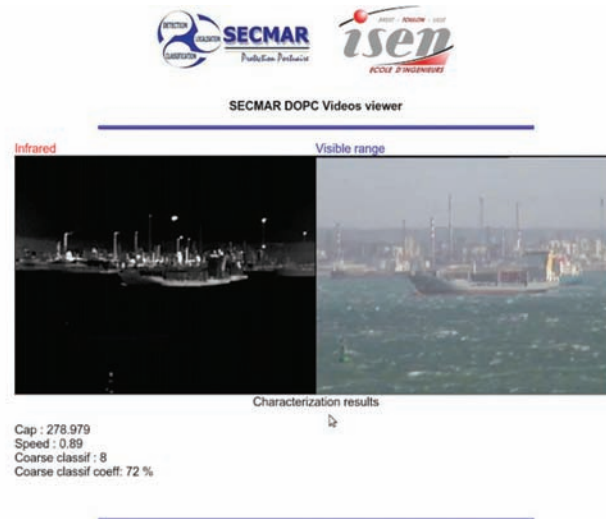


Fig. 5. Classification results displayed on the module monitor.



Fig. 6. Full process outline.

Mobiles to be identified range from swimmers to tankers, including windsurfers, jet skis, speedboats, sailing boats and other kinds of ships. In this way, the mobiles to be detected can be of various kinds regarding their shapes, dimensions or speeds. The main features of these mobiles concern shape, size speed and trajectory. The analysis of the video stream will consists of three different stages: detection and tracking, trajectory analysis and mobile identification. The trajectory analysis will allow to determine if the mobile is threatening for the seaport and the mobile

identification will be based on pattern recognition and shape classification from geometrical features.

## ALGORITHM DESCRIPTION

Fig. 6 describes the full process from detection to decision whereas Fig. 7 focuses on the classification process.

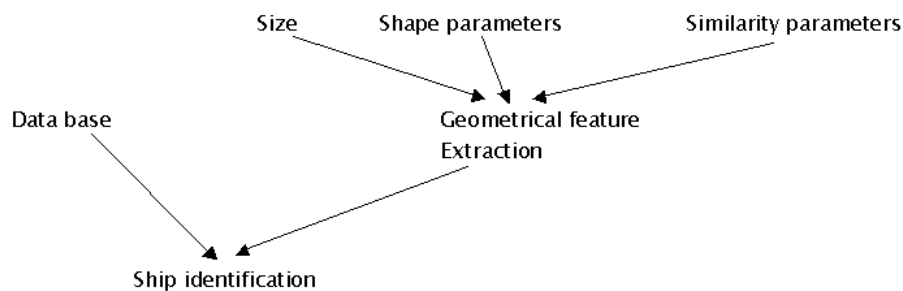


Fig. 7. Classification process outline.

## SILHOUETTE DETECTION

The first step in the image processing stage is to detect moving objects on multispectral videos (Manolakis *et al.*, 2003; Robert-Inacio *et al.*, 2007). A background image for both visible and infrared videos is periodically refreshed while acquiring videos (Karimi-Ashtiani and Kuo, 2006; Yu *et al.*, 2008). Each new image is then compared to this background image in order to extract moving objects. The detection is carried out when the same object is captured on several consecutive images. Thus, waves and ripples can be rightfully identified as noise. In order to compare the current image with the background image the two images are divided into several square areas where data is gathered by averaging.

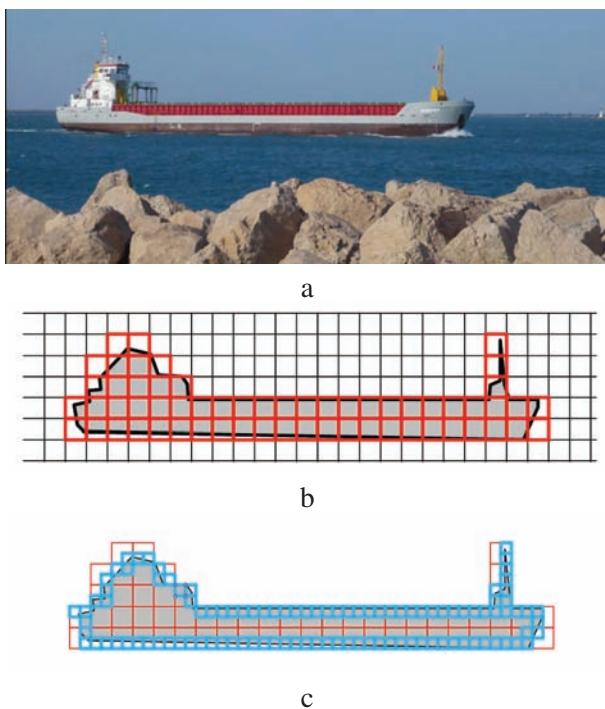


Fig. 8. *Silhouette primary detection: a) original image, b) coarse detection, c) accurate detection.*

Then the intersection of the two coarsely sampled images is computed and moving objects are represented by the complementary set of this intersection. In this way a silhouette of each moving object is given by a binary image.

The silhouette detection can be more or less accurate depending on the square area size. This size must be chosen large enough at first in order not to detect too much noise. Afterwards, the silhouette contour can be refined by iterating the detection process at lower scales only on bordering squares (Fig. 8).

Detection is carried out in two stages. The primary detection is based on motion detection. It is achieved only on infrared videos whereas the secondary detection also uses videos in visible range in order to refine results of primary detection.

### Motion detection

Targets are localized by motion detection. Target detection is achieved by considering a background image periodically refreshed throughout the process (Hall *et al.*, 2005). The time interval between two background images is set to 3 minutes but can be chosen by the operator. The video stream is set to 10 frames per second, in this way mobiles are detected by difference between the image under study and the background image.

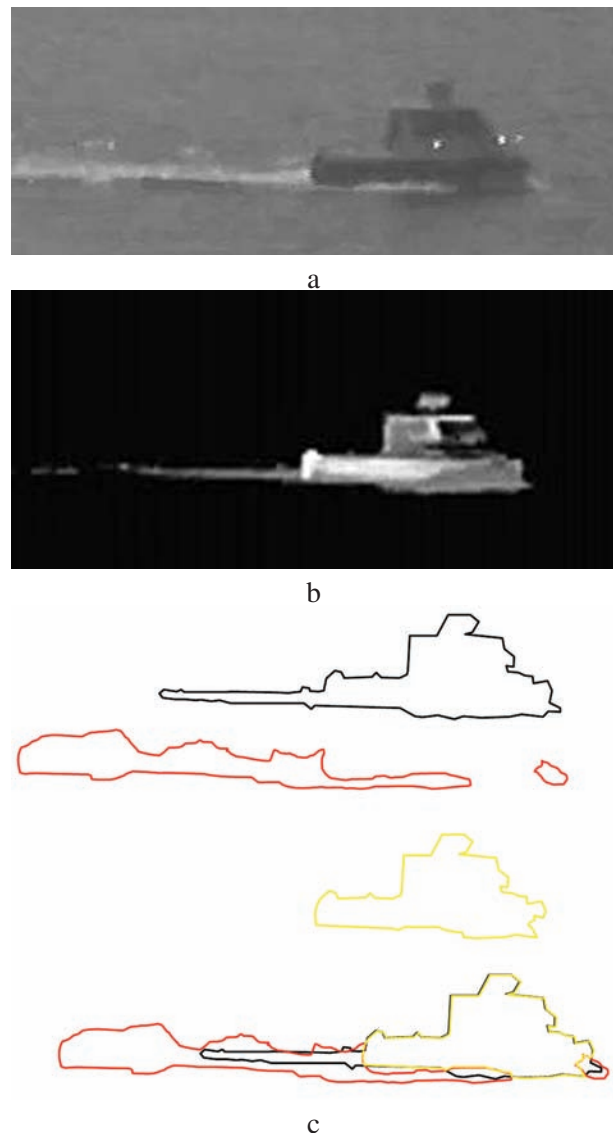


Fig. 9. *Secondary detection: a) image in visible range, b) infrared image, c) object segmentation.*

This method allows to take into account persistent changes in the background. For example, if a new ship arrives and stands in the camera frame for a long time (anchoring), its shape is integrated to the background after a given time.

In order to determine if a part of the image has been modified, the image under study (Fig. 9) is divided into several arbitrary elementary areas in which an average value is computed. Then, the average value is compared to the corresponding one on the background image. If the difference of the two average values is greater than an error tolerance value the elementary area is assimilated to a target. Colored squares in Fig. 10a correspond to modified elementary areas. Afterwards elementary areas are gathered according to their location. In other words if two or more elementary areas are neighbors, they are considered as a unique target (Fig. 10b).

Nevertheless, a major problem to solve is wave motion. Actually, even if the sea is calm, wave motion induces the detection of several irrelevant targets when using a camera acquiring color data. In order to avoid this drawback infrared data is used as the water temperature does not vary significantly between the top and the bottom of a given wave.

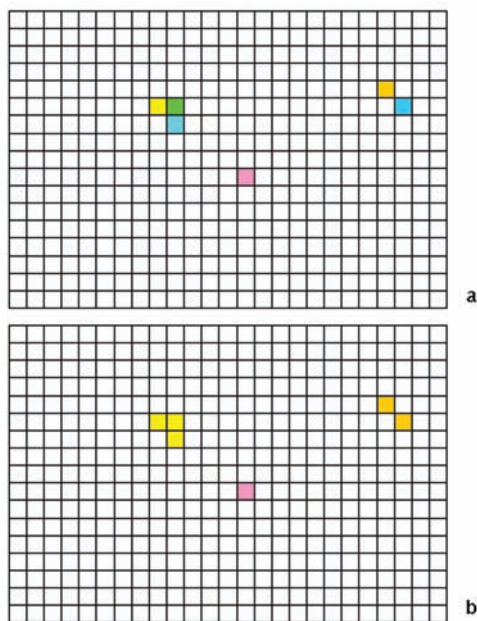


Fig. 10. Target detection process: a) primary target detection, b) elementary area gathering.

Target tracking is achieved in two steps. For a given image, the first step consists in comparing detected targets to those of the previous image. Then if their location is coherent in terms of speed, the targets are not rejected as a false alarm. Note that the rejection is effective after several images.

Afterwards, the following step consists in elaborating the history of the target position along the sequence (Fig. 11). This step is preliminary to the trajectory determination and analysis. Such an analysis should provide a helpful tool to decide if a target is threatening or not. For example, if a target is very fast and sailing directly towards the seaport, the trajectory analysis should classify it as a probable attack and the operator must be warned in order to confirm the diagnosis.

But there remain problems to solve, such as target tracking with multiple targets overlapping each others or targets disappearing from the image and reappearing a short time after.

### Silhouette refining

The silhouette refining stage is achieved by combining infrared and visible videos. For example the stem wave and the wake wave are included in the object detection on infrared images. But this part can be easily separated from the real object shape by considering that the waves appear in white, and so as a homogeneous region in terms of color, on images in visible wavelengths, and as a region of different value on infrared images. Actually, waves are of darker values if the object is hot or lighter if it is cold, depending on the weather. The secondary detection is then carried out on both visible and infrared images, by using a watershed algorithm (Beucher and Meyer, 1993; Soille, 1999) restricted to the area corresponding to the moving object in infrared. Fig. 9 shows the image in visible range (a), the corresponding infrared image (b) and the object segmentation (c). In Fig. 9c the boat outline extracted from the IR image is drawn in black, the wave outline extracted from the image in visible wavelengths is drawn in red and finally the resulting boat outline is drawn in yellow. A superposition of the three outlines is also shown. As well, images in visible wavelengths will be further used in order to classify naval vessels as they usually appear in gray tones.

### CONTOUR EXTRACTION

From the binary image previously computed to get the silhouette, contours are extracted by detecting inner bordering points and then contours are vectorized on the one hand to arrange bordering points clockwise and on the other hand to deal with less data. Furthermore, geometrical transformations are more accurate when carried out on vectorized objects. For example, a homothetic contour can be exactly computed even if the scale ratio is not an integer.

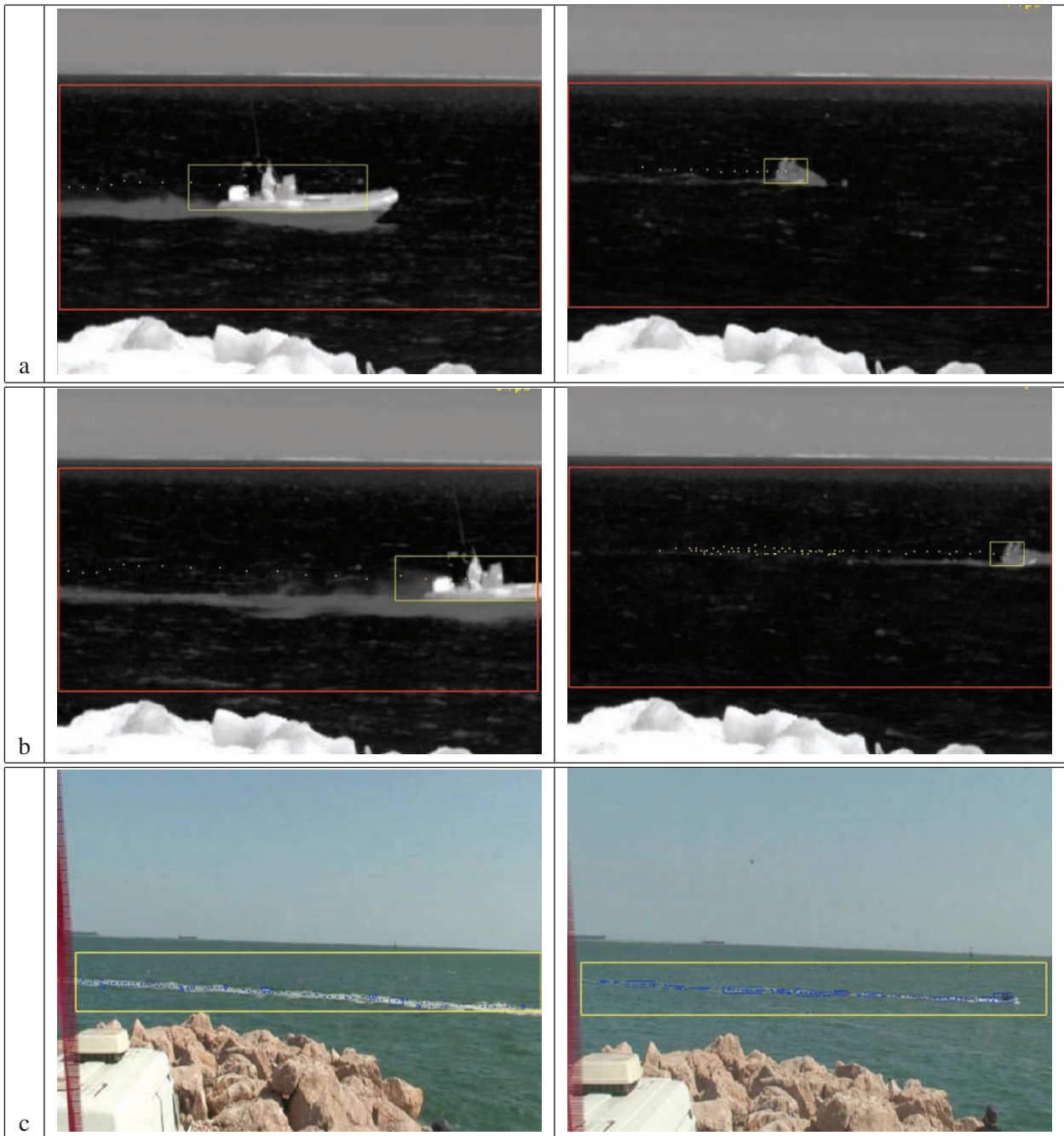


Fig. 11. Acquisition sequences for a) IR data (mid time), b) IR data (full time), c) color data (full time).

### PROFILE CORRECTION

In this application several object features must be taken into account. For example the object course is combined to the camera axis at each instant  $t$  (in other words, for each image of the current video) in order to compute the angle between the camera axis and the object direction (Fig. 12). Such data (object course, object speed, object position) are computed during the tracking stage and camera axis is provided by sensor (XML message). In this way, a rotation is applied in 3D to the object in order to get its profile (Fig. 13).

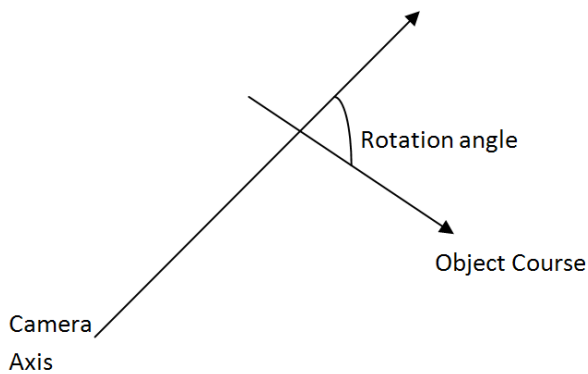


Fig. 12. Boat angle according to the ship course and the camera axis.

Obviously, such a transformation is biased as it is not possible to get hidden information from 2D videos. Hidden parts are then reconstructed by interpolation. In other words, a geometric transformation combining 3D rotation and projection is applied to the binary silhouette. This is not a significant drawback as shape classification only deals with binary data.

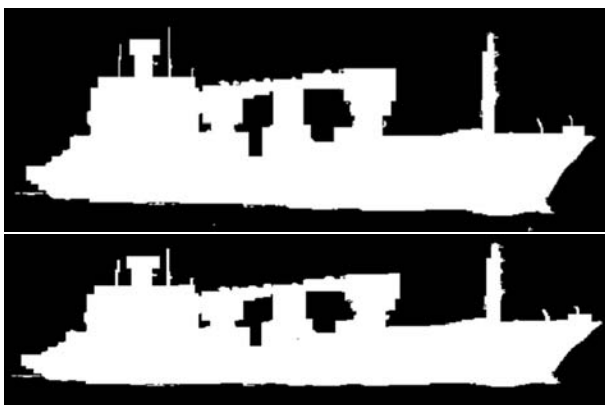


Fig. 13. Rotation in 3D.

Secondly, speedboats require a supplementary correction. The higher the speed of the boat, the more the bow rises. A rotation in 2D can correct this error and the stern part is reconstructed while the bow part disappears (Fig. 14). The boat angle is computed by using the method of least squares. Lowest points of the silhouette are considered as input data. This gives

features of the least square fitting line and the angle between this line and the horizontal line is the 2D rotation angle (boat angle).

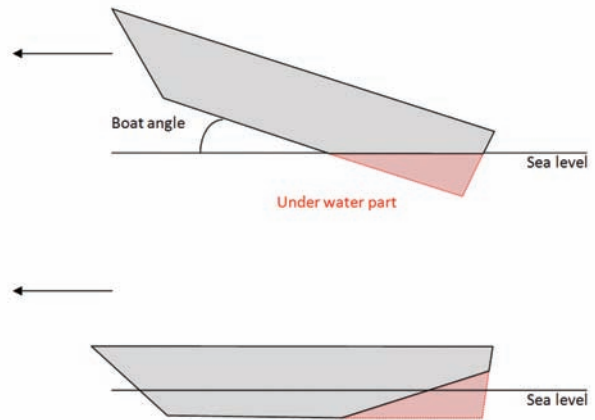


Fig. 14. Boat angle according to the ship course and the camera axis (rotation in 2D).

### SHAPE CLASSIFICATION

Different methods can be used in order to achieve ship silhouette classification: k nearest neighbor (Luo and Folleco, 2006), shape-driven level-set segmentation (Tao *et al.*, 2009), etc. In our case, shape classification is based on a similarity parameter (Robert, 1998) that is a generalization of the circularity parameter defined by the well-known formula, for any compact set  $X$  of  $\mathbb{R}^2$ , as follows:

$$P_c(X) = \frac{r}{R}, \tag{1}$$

where  $r$  is the radius of the inscribed disk into  $X$  and  $R$  the one of the circumscribed disk to  $X$  (Fig. 15).

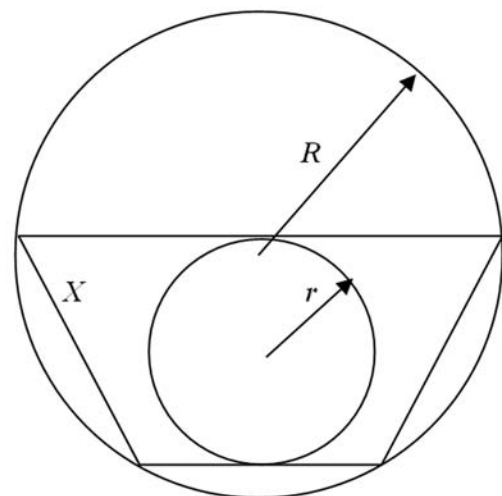


Fig. 15. Estimation of the circularity degree of a convex set  $X$ .

We can note that this circularity parameter is equal to 1, when  $X$  is a disk, as the inscribed disk merges with the circumscribed disk and so  $r = R$ .

The similarity parameter  $P$  is defined for any pair of convex shapes  $(X, Y)$  by considering two scale ratios. The first one gives the smallest homothetic set of  $X$  containing  $Y$ , and the second one, the smallest homothetic set of  $Y$  containing  $X$ . Thus, let us define the following function  $S_X(Y)$ :

$$S_X : \mathbb{K} \longrightarrow \mathbb{R}^+ \\ Y \mapsto S_X(Y), \quad (2)$$

where

$$S_X(Y) = \inf\{k > 0; Y \subset_t k.X\}. \quad (3)$$

$\subset_t$  means "included in, regardless of any translation". And then, a preliminary definition of the similarity parameter can be:

$$P_1(X, Y) = \frac{S_Y(X)}{S_X(Y)}. \quad (4)$$

Unfortunately, the parameter  $P_1$  is very sensitive to scaling. In order to solve this problem, an appropriate solution seems to be the introduction of a weight defined by the ratio of the surface areas of the two sets  $X$  and  $Y$ . The second definition follows from that:

$$P(X, Y) = \frac{S_Y(X)}{S_X(Y)} \cdot \frac{\mu(X)}{\mu(Y)}, \quad (5)$$

where  $\mu$  is the surface area measure.

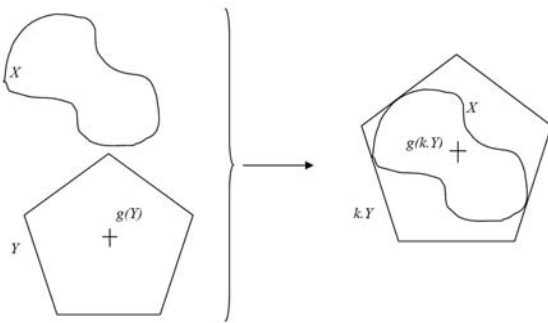


Fig. 16. *Circumscribed convex  $k.Y$  to a compact set  $X$ .*

The similarity parameter has the following properties:

- (i) If  $X \subset_t Y$  then  $P(X, Y)$  belongs to  $]0, 1]$
- (ii)  $P$  is invariant by translation
- (iii)  $P$  is invariant by scaling
- (iv) If  $X$  and  $Y$  are of the same shape regardless of a positive scale ratio, then  $P(X, Y) = P(Y, X) = 1$
- (v)  $P(X, Y) = P(Y, X)^{-1}$

In conclusion, the similarity parameter  $P$  is invariant under translation and scaling, and the closer its value to 1, the closer in shape  $X$  and  $Y$ .

In order to implement the similarity parameter, we must create fast algorithms determining the features of circumscribed convex shapes. In this way, the scale ratios required for the similarity degree estimation will be easily evaluated. Assuming that we have at our disposal an algorithm called *CircumRatio* $(X, Y)$ , determining the scale ratio to apply to a convex set  $X$  to circumscribe it to a convex set  $Y$ , the algorithm estimating the similarity degree between two shapes is the following:

```

for all pair of shapes  $(X, Y)$  do
     $k = \text{CircumRatio}(X, Y)$ 
     $k' = \text{CircumRatio}(Y, X)$ 
    Compute the two surface areas of  $X$  and  $Y$ 
    Compute the similarity degree of  $X$  and  $Y$ 
end for
    
```

First of all, let us consider the circumscribed disk algorithm that allows to design the minimal disk containing a given planar object  $X$ . The extension of this algorithm to convex sets (Serra, 1988) is the algorithm used to evaluate the scale ratios, and then, the similarity parameter. Fig. 16 shows the circumscribed convex to a compact. Some classical algorithms exist for the convex hull computation (Avis *et al.*, 1997; Barber *et al.*, 1996). They enable to take into account the corresponding convex hull  $C(X)$  to a compact set  $X$  instead of  $X$  itself.

Finally shape classification is achieved by comparing the object under study to a set of reference shapes regarding the similarity parameter. The best-scoring reference object is the most similar one. That means that objects under study are gathered into families according to each reference shape. The similarity parameter value gives a quantitative estimation of the closeness between two shapes.

## SHIP DATABASE

Several particular ship silhouettes have been extracted from images in order to obtain the reference set of shapes. These particular shapes are significant of a given kind of moving object. Each reference shape is known by its attributes: its vectorized contour, the extremal point set of its convex hull (Fig. 17), its size range and its orientation. The size range is necessary to determine if the real size of the ship under study fits or not and then if it is worth comparing this ship to the corresponding reference shape. The orientation tells if the bow is on the left or on the right. Obviously



only profiles are stored as reference shapes but the orientation must be known in order not to duplicate shapes.

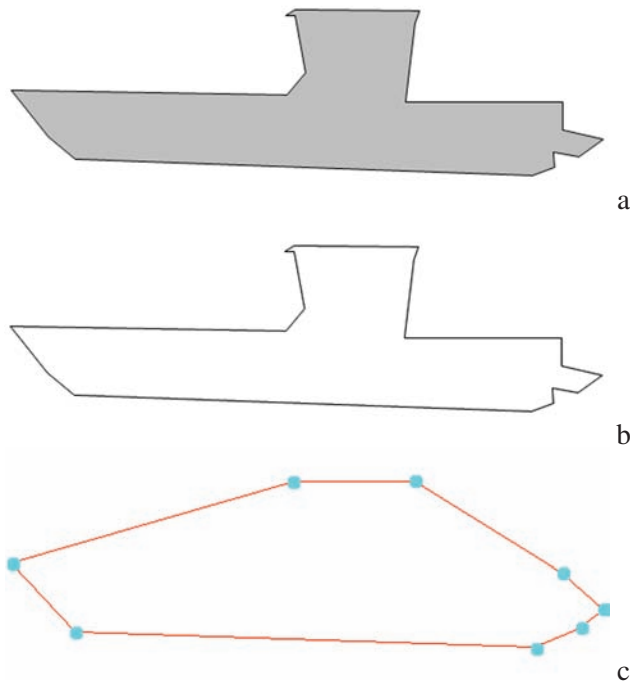


Fig. 17. a) Ship silhouette, b) contour and c) convex hull and extremal points in cyan.

Table 1. Types of moving objects and classification code.

| Classification Code | Moving Object Type |
|---------------------|--------------------|
| 0                   | Unknown            |
| 1                   | Kayak              |
| 2                   | Jet ski            |
| 3                   | Inflatable boat    |
| 4                   | Speedboat          |
| 5                   | Tug                |
| 6                   | Sailing dinghy     |
| 7                   | Naval vessel       |
| 8                   | Cargo liner        |
| 9                   | Tanker             |

Table 1 gives the list of reference shapes considered for the classification stage. A silhouette of each kind of moving objects is stored in order to be compared to the shape under study if necessary. The shape under study is then compared to each reference shape by using the similarity parameter. This shape is then associated with the best-scoring one, if this score is higher than a given threshold (usually set to 20%). This ensures that the shape under study is similar enough to the best-scoring reference shape. If the best score is less than the threshold value, then the classification code is set to *unknown*. The

reference shape database can be extended to more ship kinds. However a too large database can bring about an increasing computation time, preventing real-time processing.

## EXPERIMENTAL RESULTS

Experimental results are presented on static images on the one hand, and on videos on the other hand. The processing is the same as images are extracted from videos for the characterization step.

### EXPERIMENTAL RESULTS ON IMAGES



Fig. 18. Original images of reference shapes.

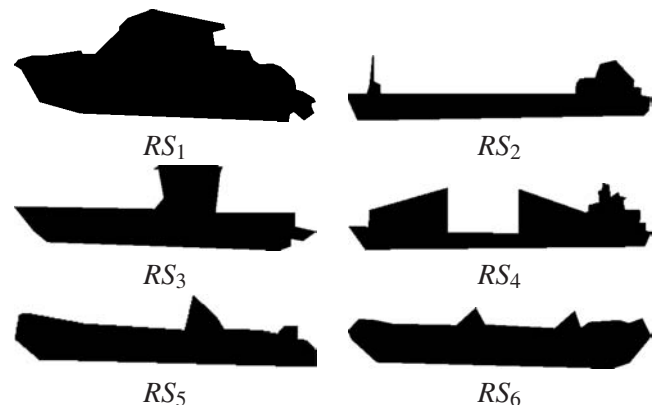


Fig. 19. Reference shapes.

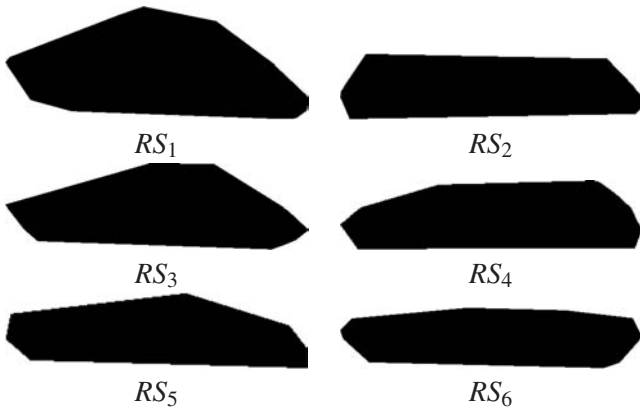


Fig. 20. Convex hulls of shapes of Fig. 19.

Table 2. Reference shapes *RS*.

| <i>RS</i> | Moving Object     |
|-----------|-------------------|
| 1         | Speedboat (big)   |
| 2         | Tanker            |
| 3         | Speedboat (small) |
| 4         | Cargo liner       |
| 5         | Inflatable boat   |
| 6         | Kayak             |

The previously described similarity parameter is suitable for this application as it is really efficient for classifying objects of same orientation (Robert, 1998). And when the two rotation corrections are achieved all ships are seen from the side. The only supplementary feature that is required is the orientation knowledge. If the orientation of the shape under study does not match that of the reference shape an axial symmetry map must be applied to the reference shape using to a vertical axis.

Shapes are extracted from images of Fig. 18. Table 3 gives the similarity parameter values when comparing shapes of Fig. 19 to each others. Shapes are described in Table 2 and Fig. 20 shows their convex hulls.

Table 3. Similarity parameter values for shapes of Fig. 19.

| <i>SP</i>              | <i>RS</i> <sub>1</sub> | <i>RS</i> <sub>2</sub> | <i>RS</i> <sub>3</sub> | <i>RS</i> <sub>4</sub> | <i>RS</i> <sub>5</sub> | <i>RS</i> <sub>6</sub> |
|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|
| <i>RS</i> <sub>1</sub> | 1.00                   | 0.24                   | 0.49                   | 0.11                   | 0.05                   | 0.38                   |
| <i>RS</i> <sub>2</sub> | 0.24                   | 1.00                   | 0.47                   | 0.45                   | 0.19                   | 0.63                   |
| <i>RS</i> <sub>3</sub> | 0.49                   | 0.47                   | 1.00                   | 0.22                   | 0.09                   | 0.39                   |
| <i>RS</i> <sub>4</sub> | 0.11                   | 0.45                   | 0.22                   | 1.00                   | 0.44                   | 0.28                   |
| <i>RS</i> <sub>5</sub> | 0.05                   | 0.19                   | 0.09                   | 0.44                   | 1.00                   | 0.12                   |
| <i>RS</i> <sub>6</sub> | 0.38                   | 0.63                   | 0.39                   | 0.28                   | 0.12                   | 1.00                   |

## EXPERIMENTAL RESULTS ON VIDEOS

Sensors acquire video data either in a static mode or in a servo-controlled mode. In the first case, the camera axis and the zoom ratio are the same for the whole video, whereas they can evolve in the second case. These two values depend on the tracking stage results and are provided to the classification module by XML messages.

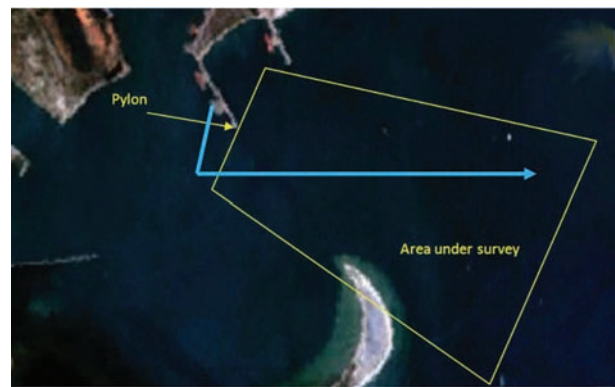


Fig. 21. In blue, speedboat trajectory for video of Fig. 23.

Fig. 21 gives an example of mobile trajectory. Fig. 22 shows images extracted from a video acquired in the static mode. Fig. 23 represents a set of images obtained from a video acquired in the servo-controlled mode. In this case, the camera can zoom in or out and the camera points to the mobile in order to keep it approximately at the image center. The camera moves by saccades and its axis is refreshed when the mobile is about to exit the area under survey of the current sensor. The two videos are sampled at one frame every five seconds but processing is achieved at a frequency of five frames per second.

On the video of Fig. 22, the inflatable boat moves back and forth on images 1 to 15 and then it follows a curved trajectory to exit the sensor field (images 16 to 20).

On the video of Fig. 23, the speedboat moves away from shore (Fig. 21). The camera axis and the zoom factor are the same for images 1 to 3. They are updated 13 times: for images 4, 5 to 6, 7 to 9, 10, 11, 12, 13, 14 to 15, 16, 17, 18, 19, and 20.

Table 4. Results for video of Fig. 22.

| Image | Classification code | Classification score |
|-------|---------------------|----------------------|
| 1     | 3                   | 0.92                 |
| 2     | 3                   | 0.91                 |
| 3     | 3                   | 0.90                 |
| 4     | 3                   | 0.92                 |
| 5     | 3                   | 0.89                 |
| 6     | 3                   | 0.91                 |
| 7     | 3                   | 0.91                 |
| 8     | 3                   | 0.90                 |
| 9     | 3                   | 0.91                 |
| 10    | 3                   | 0.90                 |
| 11    | 3                   | 0.91                 |
| 12    | 3                   | 0.91                 |
| 13    | 3                   | 0.90                 |
| 14    | 3                   | 0.92                 |
| 15    | 3                   | 0.90                 |
| 16    | 3                   | 0.91                 |
| 17    | 3                   | 0.57                 |
| 18    | 0                   | 0.12                 |
| 19    | 3                   | 0.44                 |
| 20    | –                   | –                    |

Table 5. Results for video of Fig. 23.

| Image | Classification code | Classification score |
|-------|---------------------|----------------------|
| 1     | 4                   | 0.57                 |
| 2     | 4                   | 0.52                 |
| 3     | 4                   | 0.50                 |
| 4     | 4                   | 0.51                 |
| 5     | 4                   | 0.49                 |
| 6     | 4                   | 0.51                 |
| 7     | 4                   | 0.46                 |
| 8     | 4                   | 0.40                 |
| 9     | 4                   | 0.41                 |
| 10    | 4                   | 0.38                 |
| 11    | 4                   | 0.41                 |
| 12    | 4                   | 0.41                 |
| 13    | 4                   | 0.40                 |
| 14    | 4                   | 0.42                 |
| 15    | 4                   | 0.43                 |
| 16    | 4                   | 0.53                 |
| 17    | 4                   | 0.72                 |
| 18    | 4                   | 0.82                 |
| 19    | 4                   | 0.88                 |
| 20    | 4                   | 0.89                 |

Table 4 gives classification results for video of Fig. 22. While the inflatable boat is seen by profile the classification score is around 0.90. For image 18 the classification code is set to 3 *a posteriori* though the silhouette is not really significant of an inflatable

boat. Actually a history of previous values is also used to find the new value, combined to the mobile position evolution. The speedboat in the background is not considered for the study as it does not move in a sufficiently significant way.

Table 5 gives classification results for video of Fig. 23. The classification score highly depends on the influence of the geometric transformation enabling to correct the 3D rotation. The highest scores are obtained when the speedboat moves in profile.

Images of Fig. 23 are of different quality. Their features depend on the sensor orientation relative to the sun.

## STATISTICS

The characterization module receives XML messages from eight different sensors: three IR sensors in static mode, three visible sensors in static mode, one IR sensor in servo-controlled mode and one visible sensor in servo-controlled mode. Sensors work by pair, an IR sensor corresponding to a visible sensor. After the characterization stage, the characterization module sends XML messages to the shore center as soon as the computation is achieved (Fig. 4).

Table 6. Statistics on three examples of acquisition in a static mode.

|   |         |         |         |
|---|---------|---------|---------|
| <b>Video length</b>                                 | 85683 s | 86323 s | 78736 s |
| <b>Target number</b>                                | 694     | 330     | 185     |
| <b>Output XML message number</b>                    | 173144  | 53819   | 40045   |
| <b>Average message number per target</b>            | 249.48  | 163.08  | 216.45  |
| <b>Average classification score</b>                 | 0.7461  | 0.7293  | 0.8505  |
| <b>Average message number per second</b>            | 2.02    | 0.62    | 0.51    |
| <b>Average message number per target per second</b> | 0.0029  | 0.0019  | 0.0027  |

Table 6 gives statistics when videos are acquired in a static mode. In these conditions, several targets can be tracked simultaneously. The output XML message number represents the number of achieved classifications, as a message is sent whenever a characterization is completed. The message frequency (average message numbers: per second, per target, per target and per second) depends on shipping traffic

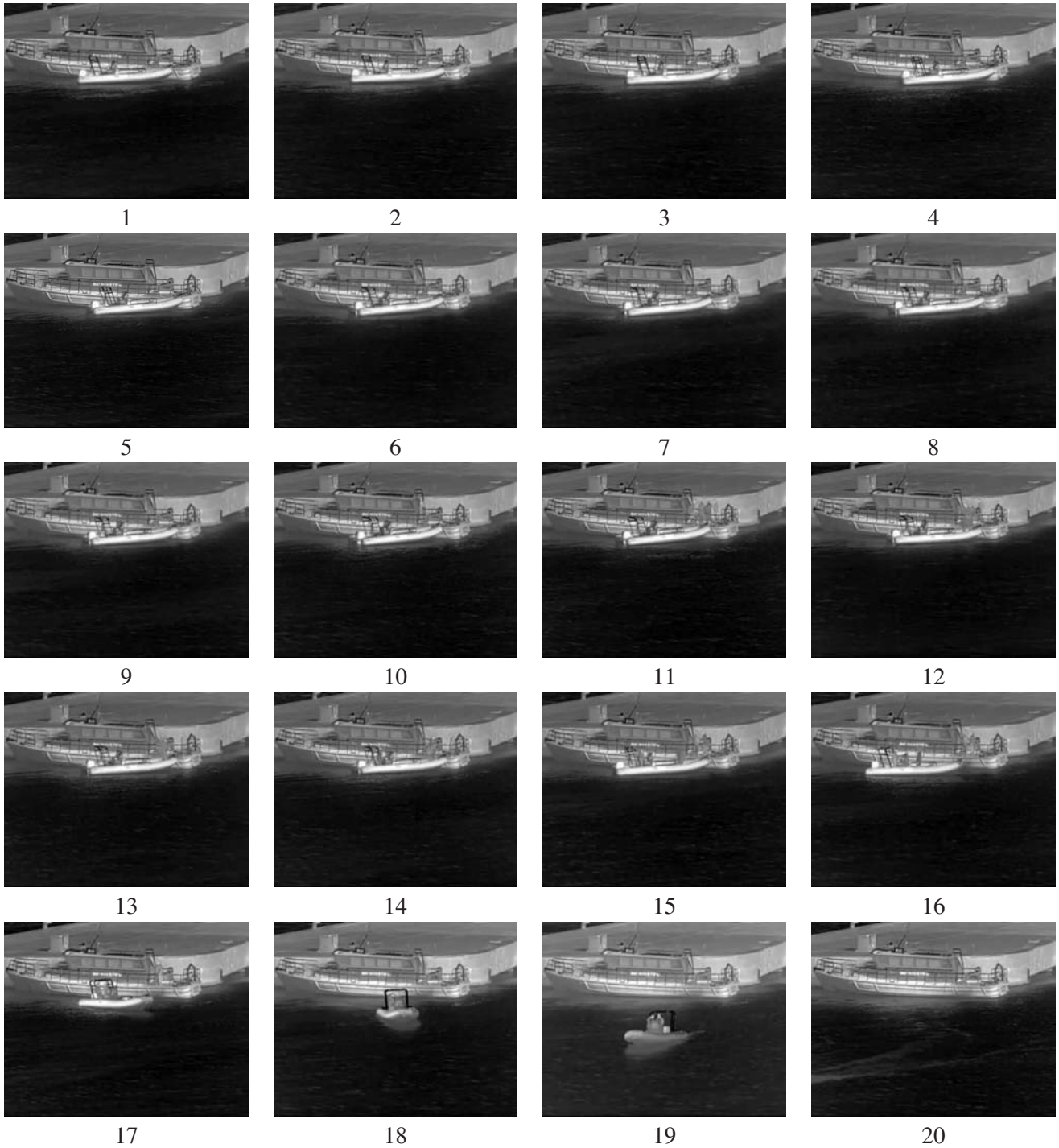


Fig. 22. *Images from an IR video acquired in a static mode.*

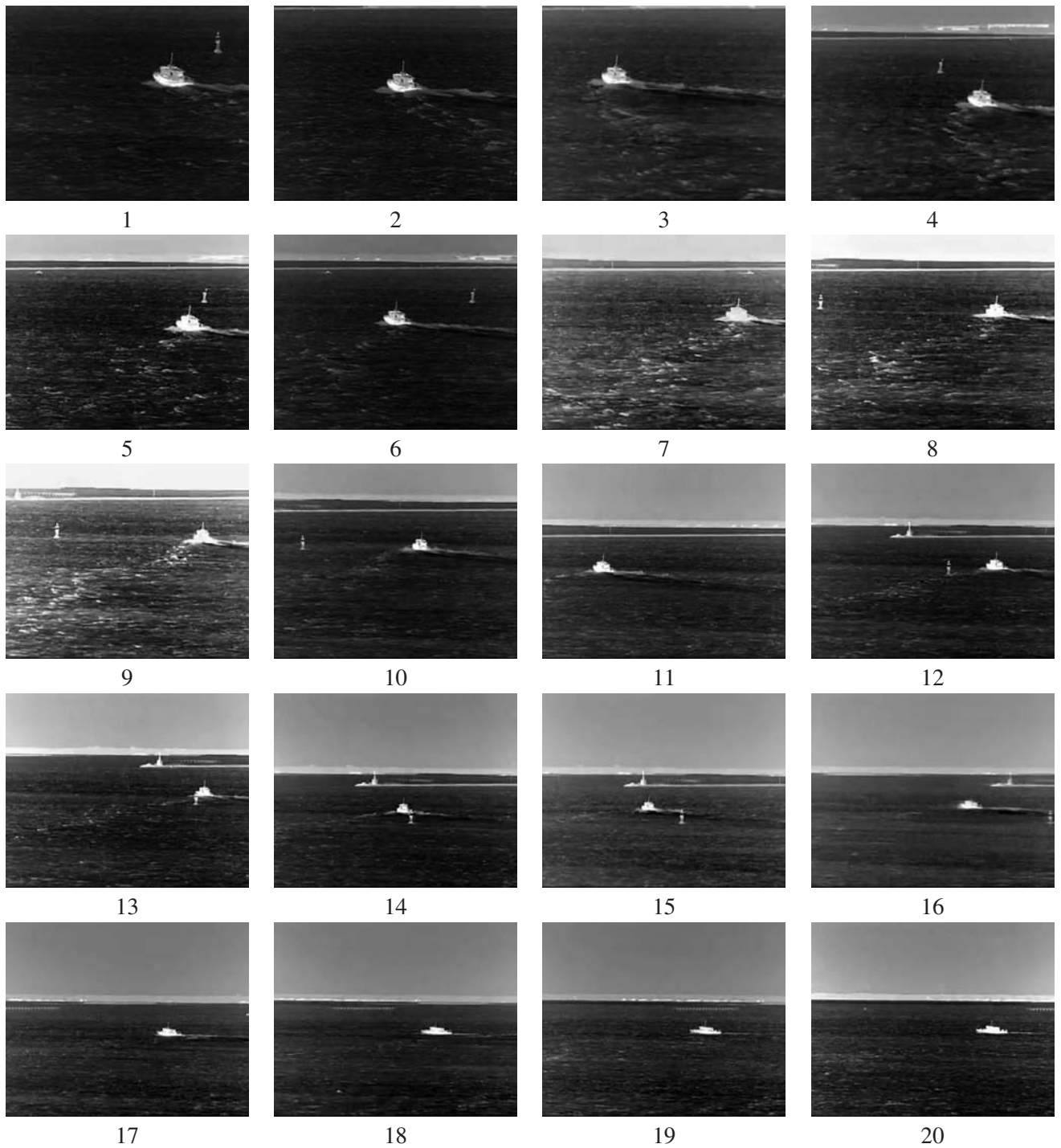


Fig. 23. Images from an IR video acquired in a servo-controlled mode.

and is not really significant of the system efficiency. The average classification score is high and proves that classification has been achieved in appropriate conditions.

Table 7 gives statistics when videos are acquired in a servo-controlled mode. In these conditions, only one target is tracked at a time. The average message number per second varies from 1.00 to 2.70. This means that one to three classifications can be completed in one second. The minimal values can be explained by the fact that classification cannot be achieved if images are too blurred. This can occur when the sensor features (camera axis and zoom ratio) are updated too often: the sensor moves and images are blurred. This problem appears with too fast mobiles moving too close to the sensors.

Table 7. *Statistics on six examples of acquisition in a servo-controlled mode.*

| Video length | Output XML message number | Average message number per second |
|--------------|---------------------------|-----------------------------------|
| 2856 s       | 6096                      | 2.13                              |
| 23 s         | 62                        | 2.70                              |
| 35 s         | 35                        | 1.00                              |
| 15 s         | 37                        | 2.47                              |
| 19 s         | 34                        | 1.79                              |
| 491 s        | 617                       | 1.26                              |

## CONCLUSION

Seaport surveillance is a very sensitive task to achieve as it deals with a lot of parameters such as tanker or cargo traffic as well as small boats such as fishing vessels or sailing boats or even inflatable boats or jet skis. Automatic video processing implies fast algorithms. Ours must be able to classify moving objects in terms of shape and analyze their behaviors in order to detect threatening mobiles. In this paper a full classification process has been described associating each moving object to a reference shape. This reference shape is the best-scoring one according to a similarity parameter comparing shapes and their convex hulls. This parameter takes into account shapes in their whole and is invariant in scale and translation, but not in rotation. That is one of the reasons why the object outlines are preprocessed in order to deal only with profile. The other reason is that profiles are the most significant sides for ships, and then the most discriminant. But the whole classification system still has to be evaluated on a large sample of moving objects to estimate more accurately the similarity parameter efficiency. Furthermore the video

database must be enlarged to all weather conditions to test the segmentation algorithm performance. In this way the classification system will be more robust. Finally target features must be correlated with data from other sensors such as radar (Vasile and Marino, 2005) or sonar.

## ACKNOWLEDGEMENTS

This research is supported by the SECMAR project, funded by the Pôle de Compétitivité (cluster) Mer of the Région Provence-Alpes-Côte-d'Azur, France, and Direction Générale de l'Armement, and managed by Thales Under Water Systems, Sophia Antipolis, France. Many thanks to CS Communications & Systèmes, Toulon, France, and Bertin Technologies, Aix-les-Milles, France, for providing videos.

Images of Fig. 1, 3 and 21 are extracted with World Wind 1.4.0.0, software developed by the National Aeronautics and Space Administration (<http://worldwind.arc.nasa.gov/>).

## REFERENCES

- Avis D, Bremner D, Seidel R (1997). How good are convex hull algorithms. *Comp Geom Theory Appl* 7:265–302.
- Barber C, Dobkin D, Huhdanpaa H (1996). The quickhull algorithm for convex hulls. *ACM Trans Math Softw* 22(4):469–83.
- Besada J, Garcia J, Portillo J, Molina J, Varona A, Gonzalez G (2005). Airport surface surveillance based on video images. *IEEE Trans Aero Elec Syst* 41(3):1075–82.
- Beucher S, Meyer F (1993). The morphological approach to segmentation: The watershed transformation. In Dougherty R, ed. *Mathematical Morphology in Image Processing*. New York: Marcel Dekker. 433–81.
- Hall D, Nascimento J, Ribeiro P, Andrade E, Moreno P, Pesnel S, List T, Emonet R, Fisher R, Santos Victor J, Crowley J (2005). Comparison of target detection algorithms using adaptive background models. In: *Proc 2nd IEEE Int Works Visual Surveil Perf Eval of Track Surveil*. Beijing, China. October 15–16. 113–20.
- Kale A, Roychowdhury K (2004). Fusion of gait and face for human identification. In: *Proc IEEE Int Conf Acoust Speech Signal Process*. Montreal, Canada. May 17–21. 5:901–4.
- Karimi-Ashtiani S, Kuo C (2006). Automatic real-time moving target detection from infrared video. In: *Proc Int Conf Intel Inform Hiding Multimedia Signal Process*. Pasadena, USA. December 18–20. 435–40.
- Luo Q, Folleco A (2006). Classification of ships in surveillance video. In: *Proc IEEE Int Conf Inform*

- Reuse Integr. Waikoloa Village, USA. September 16–18. 432–7.
- Manolakis D, Marden D, Shaw G (2003). Hyperspectral image processing for automatic target detection applications. *Lincoln Lab J* 14(1):79–116.
- Moeslung T, Granum E (2001). A survey of computer vision-based human motion capture. *Comput Vision Image Und* 81:231–68.
- Reulke R, Bauer S, Doring T, Meysel F (2007). Traffic surveillance using multi-camera detection and multi-target tracking. In: *Proc Image Vision Comput New Zealand*. Hamilton, New Zealand. December 5–7. 175–80.
- Robert F (1998). Shape studies based on the circumscribed disk algorithm. In: *Proc IEEE-IMACS Conf Comput Eng Systems Appl*. Hammamet, Tunisia. April 1–4. 4:821–6.
- Robert-Inacio F, Raybaud A, Clément E (2007). Multispectral target detection and tracking for seaport video surveillance. In: *Proc Image Vision Comput New Zealand*. Hamilton, New Zealand. December 5–7. 169–74.
- Serra J (1988). *Image analysis and mathematical morphology*. Vol.2: Theoretical advances. London: Academic Press.
- Soille P (1999) *Morphological image analysis*. Berlin: Springer Verlag.
- Tao C, Tan Y, Cai H, Tian J (2009). Ship detection and classification in high-resolution remote sensing imagery using shape-driven segmentation method. In: *Proc 6th Int Symp Multispectr Image Process Pattern Recogn*. Yichang, China. October 30–November 1. *Proc SPIE* 7495:74954N.
- Vasile A, Marino R (2005). Pose-independent automatic target detection and recognition using 3D laser radar imagery. *Lincoln Lab J* 15(1):61–78.
- Wang L, Suter D (2006). Analyzing human movements from silhouettes using manifold learning. In: *Proc IEEE Int Conf Video Signal Based Surveil*. Sydney, Australia. November 22–24. 7.
- Yi W, Marshall S (2000). A novel approach for automatic aircraft detection, In: *Proc 10th Eur Signal Proces Conf*. Tampere, Finland. September 4–8. 2117–20.
- Yu W, Yu X, Zhang P, Zhou J (2008). A new framework of moving target detection and tracking for UAV video application. *Int Arch Photogram Remote Sens Spatial Inform Sci* 37:609–613.
- Zhou X, Bhanu B (2008). Feature fusion of side face and gait for video-based human identification. *Pattern Recogn* 41(3):778–95.
- Zhu C, Zhou H, Wang R, Guo J (2010). A novel hierarchical method of ship detection from spaceborne optical image based on shape and texture features. *IEEE Trans Geosci Remote Sens* 48(9):3446–56.