Automatic classification of ships from infrared (FLIR) images

Paul Withagen[a], Klamer Schutte[a], Albert Vossepoel[b], and Marcel Breuers[a]

[a]Electro-Optical Systems, TNO Physics and Electronics Laboratory, PO Box 96864, NL-2509 JG The Hague, The Netherlands
[b]Pattern Recognition Group, Delft University of Technology, Faculty of Applied Sciences, Lorentzweg 1, NL-2628 CJ Delft, The Netherlands

ABSTRACT

The aim of the research presented in this paper is to find out whether automatic classification of ships from Forward Looking InfraRed (FLIR) images is feasible in maritime patrol aircraft. An image processing system has been developed for this task. It includes iterative shading correction and a top hat filter for the detection of the ship. It uses a segmentation algorithm based on the gray value distribution of the waves and the Hough transform to locate the waterline of the ship.

A model has been developed to relate the size of the ship and the angle between waterline and horizon in image coordinates, to the real-life size and aspect angle of the ship. The model uses the camera elevation and distance to the ship. A data set was used consisting of two civil ships and four different frigates under different aspect angles and distances. From each of these ship images, 32 features were calculated, among which are the apparent size, the location of the hot spot and of the superstructures of the ship, and moment invariant functions.

All features were used in feature selection processing using both the Mahalanobis and nearest neighbor (NN) criteria in forward, backward, and branch & bound feature selection procedures, to find the most significant features.

Classification has been performed using a k-NN, a linear, and a quadratic classifier. In particular, using the 1-NN classifier, good results were achieved using a two-step classification algorithm.

Keywords: FLIR, target, classification, automatic target recognition, infrared

1. OVERVIEW OF LITERATURE

Several methods have been described in literature to perform the automatic classification of ships from FLIR images. Note that most papers are 10 to 20 years old, which is a long time for an image processing and pattern recognition application.

Depending on the data used, which differs from real FLIR images [4, 8] to recorded paintings [14], simulated images [11], and visual-light images [6], most methods need some pre-processing before the objects can be detected. Detection itself is performed manually [4] or automatically, depending on the quality of the images, and the number of objects in a single image.

Together with detection, a Region Of Interest (ROI) is declared and the ship is segmented. This has been done by methods varying from a simple threshold [6, 8, 9] to two-class relaxation algorithms [4] and edge detection [8], depending on the difference in temperature between object and background in the images.

The calculated features can be divided into three groups. Features can be calculated using the gray value distribution, by using moments and moment invariant functions [11, 14]. Also, features can be calculated using the shape of the silhouette, like location and size of the superstructures [4, 6, 8, 9]. Beside this, the location of the hot spot, being the funnel for most ships, can be used as a feature [3].

The classification of ships has been done by two classes of methods, namely k-Nearest Neighbor (k-NN) classification [11, 14] and using a binary decision tree [6, 8].

The performances of all methods are hard to compare because of the large variety in data and number of classes used. Results of 70-93% [4] and 63-90% [8] correct classifications were obtained using real data and over eight classes.

Part of the SPIE Conference on Signal Processing, Sensor Fusion, and Target Recognition VIII
Orlando, Florida • April 1999 SPIE Vol. 3720 • 0277-786X/99/$10.00

2. IMAGES USED

Although hundreds of tapes with FLIR images were available for this project, it was hard to get sufficient images. This was because the data had to meet the demands below to be usable:

- The ships in the images had to be known.
- The height of the camera had to be known.
- The distance of the ships had to be known or calculable.
- Of each class, images with several aspect angles were needed.

As a result, most of the images could not be used.

The images used were recorded with a Barr & Stroud Ltd. IR18 camera on Umatic tape. The images were digitized using an 8-bit frame grabber to 720 by 576 pixels.

The camera has a field of view of five degrees horizontal and three degrees vertical. It operates in the 8-12μm region and has a CCIR video output.

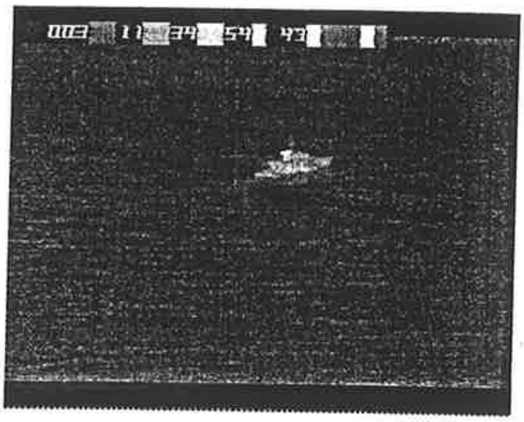

Figure 1 Image directly from the frame grabber.

Images were used from six ships: two civil ships and four different frigates. On these images, classification between civil and frigate (2-class) and between all different ships (6-class) has been performed. A training set has been chosen by taking approximately eight images with different aspect angles from each class. From all remaining images of each class, eight images were randomly chosen to be used as test set, see Table 1. The images were recorded with a camera on a stable platform, at an altitude of 100m. Most images contained ships at ranges between 4 and 7 km.

Table 1 Number of images in each of the data sets.

# images ship	training set	test set	random test set
Rossetti	6	11	8
coaster	2	9	8
Euro	8	15	8
Elli	13	15	8
Coventry	15	16	8
Bowen	10	16	8
Total:	54	82	64

3. IMAGE PROCESSING

To be able to classify ships, that is, to do pattern recognition, features are needed for each object. However, the calculation of features is not possible on raw images (see Figure 1). A small ROI is needed for this purpose.

To obtain this ROI, first shading is removed by fitting a quadratic surface to the image and subtracting it from the image. The ship can be detected using a top-hat transform (see [13]). This way the funnel, being the primary hot-spot of the ship, is located in the image.

The segmentation used is quite simple. It is based on the statistical distribution of the waves. The gray values from waves have an almost Gaussian distribution, with a larger number of points at the higher gray values because of the existence of whitecaps, which appear warmer then the surrounding. By using a bi-threshold at μ-2σ and μ+3σ, where μ and σ are the mean and standard deviation of the image after the shading removal, most ship-pixels are found. Pixels which were not found are those situated between hot and cold parts of the ship. A partial solution used in this project is to fill these gaps with a closing operation. Only segmented ships at locations where nearby a funnel was detected are taken into account (see figure 2a).

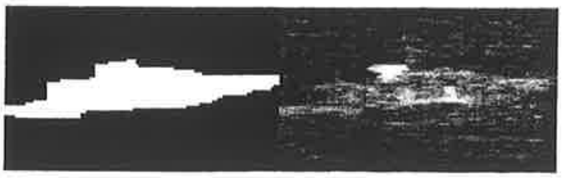

Figure 2a Images after segmentation. Left the binary segmented image and right the gray value image.

Figure 2b Images after skewing. The features are calculated from these images.

A Hough transform [12] is applied to a gradient filtered version of the gray scale ROI to find the waterline of the ship, which appears relatively bright in the images. The quantization of the Hough-space is done automatically [12]. Finally, the ship is skewed based on the waterline orientation, to have it horizontal for further processing (see Figure 2b).

4. CALCULATION OF REAL DIMENSIONS

From the size of the ship in the image, the true size of the ship can be calculated. To do this, first the distance of the ship has to be known. Using the fact that the height of the camera was fixed and known, the distance d of the ships can be calculated with $d \approx \phi R - \sqrt{\phi^2 R^2 - 2hR}$, with R the radius of the earth, h the height of the camera, and ϕ the angle between the horizon and the ship. For this, the horizon had to be in the image, which was the case in all images.

In an operational system this distance can be retrieved from flying height and the elevation of the camera, or by direct distance measurements using radar or laser range finder. The first method has the advantage that a passive method is used, making the plane harder to detect.

Now that the distance to the ship is known, the real aspect angle can be calculated from the elevation of the camera and the measured aspect angle in the image α_i by:

$$\alpha_r = \arctan \frac{\tan \alpha_i}{\beta}$$ (see Figures 3a and 3b)

Using the real aspect angle α_r, the elevation of the camera β and the distance of the ship d, the real dimensions of the ship can be calculated by

$$l_{ship} = \frac{n}{N \cos \alpha_r} 2r \tan \frac{\theta_h}{2}$$ for the length, and $$h_{ship} = \frac{m}{M \cos \beta} 2r \tan \frac{\theta_v}{2}$$ for the height. In these formulas, n and m are the horizontal number of pixels over the length and the vertical number of pixels over the height of the ship in the image after shading correction, N and M are the total number of pixels horizontally and vertically and θ_h and θ_v are the horizontal and vertical field of view of the camera.

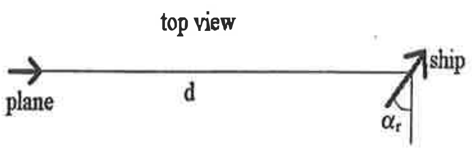

Figure 3a Definition of the aspect angle α_r.

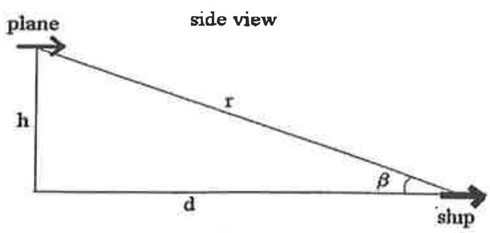

Figure 3b Definition of the camera elevation β

5. FEATURES

The real height and length of the ships are used as features, but for high classification performance, more features are necessary. All other features are calculated from the binary silhouette and/or the gray value ROI, both after skewing so that the waterline is horizontal.

The features used can be divided in the following groups:
1. size parameters
2. location and height of superstructures
3. moment invariant functions (both of binary and gray value images)
4. hot spot & cold spot
5. centroid of binary image

Of each group, a short description is given below.

size parameters The size of the ship (its length and height) is measured in the image (in pixels) together with the aspect angle in the image. From these values, the real length and height (in meters) of the ship can be calculated.

location and height of superstructures The x- and y-locations of the highest four superstructures (mast, cabin, funnel, etc.) as seen in the binary image are used as features. The locations are given relative to the total length and height of the ship. For the image scaled to 10 by 5 pixels, the highest two superstructures are also used.

moment invariant functions (both binary and gray value) The use of moments has the disadvantage that the moment value changes as the size of the image changes, or the image is rotated or translated. Using the second and third-order central moments [1], seven scale, rotation, and translation invariant moment functions can be derived [2]. These seven moment invariant functions are calculated for both the binary and the gray value image.

hot spot & cold spot The x- and y-location of the largest region of hot pixels (hottest 5%) and cold pixels (coldest 5%) within the ship are used as features. Again, the locations are relative to the total length and height of the ship.

centroid of binary image After taking the distance transform of the binary image, the point with the highest value is called the centroid of the image. The x- and y-locations of this point, relative to the length and height of the ship, are used as features.

6. FEATURE SELECTION AND CLASSIFICATION

Now, a total of 32 features is available for each ship. This amount of features is too large. A classifier needs a small number of good features to distinguish between the classes. Especially with the small number of training images available for each class, the use of all features would result in adaptation to details of the training data, rather than to extract the overall trends, which are needed to classify new images correctly.

Before feature selection can be done, we normalize the features. This is performed, for all features, by subtracting the mean and dividing by the standard deviation of all training data. For the testing data, the means and standard deviations of the training data are used.

To find out which features perform best, a measure as to what is good or bad performance is needed. In this research, two measures were used, namely, the nearest neighbor measure and the Mahalanobis measure (see [1]). We used forward, backward and branch & bound feature selection to find the best performing set of features.

Three classification methods are considered, which will be described below.

6.1 Nearest neighbor classifier

The k-Nearest Neighbor (k-NN) method searches in the feature-space for the k nearest objects. The object is assigned to the same class as the majority of these k objects. In the training-phase, the k-NN algorithm used automatically adjusted the number of neighbors to use. With all experiments performed, the optimal k turned out to be one.

6.2 Linear classifier

A linear discriminant function D is a function of the form $D_A = a_{Ak}x_{Ak} + a_{Ak-1}x_{Ak-1} + \ldots + a_{A0}x_{A0}$, in which x_{Ak} is feature k of class A and a_{Ak} is a parameter adjusted by training. This function is derived for each class. The object will be assigned to the class for which the discriminant function is the largest.

As a linear classifier, the minimum leased squares linear classifier (mlslc) has been used, which is based on the Fisher's discriminant [1] and has two modes of operation:

- Single: Between each class and the combined set of other classes a single linear classifier is computed.
- Multiple: For each of the i classes a combined linear classifier is computed separating it from the other i-1 classes. This increases the computing time by about a factor i.

In both modes the error during training is minimized in least squares sense.

6.3 Quadratic classifier

A non-linear classifier has a discriminant function which is not linear in the features. To construct such a function, a larger number of training objects is needed because the number of degrees of freedom is larger. The normal densities based quadratic classifier (nqc) has been used [1]. This classifier assumes normally distributed classes and finds the quadratic (n-dimensional) surface which optimizes classification performance.

7. RESULTS

To choose the optimal feature sets, all three classifiers are now used on all feature sets generated by feature selection with all feature selection methods and both criteria. From these classification results, the best-performing feature sets are chosen for each classifier. The classifiers with the smallest number of classification errors are listed in Table 2. From these results, it can easily be concluded that the 1-NN classification method gives the best results.

Of each of the feature sets used above, the features are listed below.

1. civil versus frigate, 1-NN, three features
   - height of the ship
   - first moment function on binary image
   - fifth moment function on binary image
2. civil versus frigate, mlslc, four features
   - height of the ship
   - y-coordinate of the center of hot spot
   - first moment function on binary image
   - sixth moment function on binary image
3. civil versus frigate, nqc, four features
   - height of the ship
   - y-coordinate of the center of the hot spot
   - first moment function on binary image
   - sixth moment function on binary image
4. 6-class, 1-NN or mlslc, eleven features
   - height of the ship
   - y-coordinate of the third superstructure
   - x-coordinate of the center of the hot spot
   - y-coordinate of the center of the hot spot
   - first moment function on gray value image
   - second moment function on gray value image
   - third moment function on gray value image
   - fourth moment function on gray value image
   - first moment function on binary image
   - fourth moment function on binary image
   - x-coordinate of the centroid of the binary image
5. 6-class, nqc, one feature
   - height of the ship

Because the classification between civil ships and frigates is almost perfect, there should be no wrong classifications from a civil ship to a frigate and vice versa, when looking at the confusion matrix after 6-class classification. So there should be no classifications in the gray regions of Table 3. However, there is a significant number of erroneous classifications in these regions, indicating that classification might profit from a two-step classification scheme.

First, classification is performed between civil ships and frigates. Then, two separate classifications are performed, one between the two civil ships and the second between the four frigates. The results of the separate classifications are shown in Table 4.

Table 2 Classification results using the best-performing feature sets. The code for the used feature selection method is made of the used criterion (Nearest Neighbor or Mahalanobis) and the used feature selection method (Forward or Backward).

	2-class, civil versus frigate			**6-class**		
classification method	1-NN	mlslc	nqc	1-NN	mlslc	nqc
feature selection method	NN-b	M-f	M-b	NN-f	NN-b	M-b
#features	3	4	4	11	11	1
error	2.1%	2.1%	6.3%	10.4%	25.0%	54.2%

Table 3 Confusion matrix after 6-class classification using the 1-NN classifier. The number of test images for each class is eight.

to:	civil-1	civil-1	frigate-1	frigate-2	frigate-3	frigate-4
from:						
civil-1	100%	0%	0%	0%	0%	0%
civil-2	0%	62.5%	12.5%	0%	0%	25%
frigate-1	0%	0%	87.5%	0%	12.5%	0%
frigate-2	0%	0%	12.5%	75%	12.5%	0%
frigate-3	0%	0%	12.5%	0%	75%	12.5%
frigate-4	12.5%	0%	0%	0%	0%	87.5%

Table 4 Results of two-step classification. The code for the used feature selection method is made of the used criterion (Nearest Neighbor or Mahalanobis) and the used feature selection method (Forward, Backward).

	civil ships	**frigates**		
classification method	1-NN	1-NN	1-NN	1-NN
feature selection method	NN-b	NN-b	NN-f	NN-f
#features	1	11	8	7
error	0%	3.1%	6.2%	9.4%

When classification between civil ships and frigates is done using three features and 1-NN, and classification between frigates is done using eleven features and 1-NN, the total classification error becomes 4.2%. One error for the 2-class classification, no errors for the classification of civil ships and one error for the frigates (3.1% of 4x8 ships), so $100\% \times \left(\frac{2 errors}{6 classes \times 8 ships}\right) = 4.2\%$. This is better than classifying all classes simultaneous, which gave an error of 10.4%.

The features, used in the two-step classification are:

1. civil versus frigates, 1-NN, three features
   - height of the ship
   - first moment function on binary image
   - fifth moment function on binary image
2. civil ships, 1-NN, one feature of the following two
   - height of the ship
   - x-coordinate of the centroid of the binary image
3. frigates, 1-NN, eleven features
   - height of the ship
   - length of the ship
   - x-coordinate of the first superstructure
   - y-coordinate of the second superstructure
   - y-coordinate of the third superstructure
   - x-coordinate of the hot spot
   - second moment function on gray value image
   - second moment function on binary image
   - fifth moment function on binary image
   - x-coordinate of the centroid of the binary image
   - y-coordinate of the centroid of the binary image

8. CONCLUSIONS AND DISCUSSION

The goal of the research described in this paper was to see whether a system could be made for the automatic classification of ships in FLIR images. The conclusion can be drawn that this is possible, although further research is needed before an operational system is ready. Below, some further conclusions will be drawn and recommendations for further study will be given.

8.1 Conclusions

A system has been made which can automatically detect a ship, calculate its real dimensions and extract features from it. With these features, it is able to distinguish between the different ships and to distinguish between the civil ships and the frigates. This shows that an operational automatic classifier could well be possible.

Some further conclusions can be drawn from this research:

- For the feature selection, it is necessary to use both forward and backward selection, with both the Mahalanobis and the nearest neighbor criteria, as no single method produces optimal performance for all classification problems.
- It proved necessary to use different kinds of features together. It is not possible to recommend one specific class of features.
- The 1-nearest neighbor classifier performs best.

8.2 Further research

Before an operational system is ready, there are several problems which will have to be solved.

- Other classes of ships should be evaluated, and classes consisting of more than one ship should be used.
- It should be evaluated how many images will be necessary to describe a class in the training set. Both the number of different ships in each class, and the number of aspect angles per ship have to be known.
- More features should be evaluated. Also features should be found to distinguish between other classes.
- The generation of a training set should be made easier. This might be done by using features which can be calculated from existing ship-databases, like Jane's Fighting Ships [10].
- Other classification-methods should be evaluated. For example neural nets and decision trees that make use of expert-knowledge.
- There is a problem when ships are seen from the front or back. These images look very different and the detection of the waterline will probably go wrong. These images could be detected by looking at the length-to-height ratio of the ship, which will become very small. Then the images could be rejected or a database could be available with front-views from all classes.

9. REFERENCES

1. C.H. Chen, L.F. Pau, and P.S.P. Wang, "Handbook of pattern recognition and computer vision", World Scientific, Singapore, 1993
2. M.K. Hu, "Visual pattern recognition by moment invariants", *IRE Transactions on Information Theory*, Vol. IT-8, pp. 179-187, 1962
3. K. Kanzaki, E. K. Wong, and M. Kabrisky, "The use of morphology and the low Fourier frequencies in hierarchical fuzzy target segmentation and classification", *SPIE Proc. Conf. on Architecture, Hardware, and Forward-Looking Infrared Issues in Automatic Object Recognition*, Vol. 1957, pp. 95-106
4. D.N. Kato, R.D. Holben, A.S. Politopoulos, and B.H. Yin, "Ship classification and aimpoint maintenance", *SPIE Proc. Conf. on Infrared Systems and Components, Los Angeles*, Vol. 890, pp. 174-181, 1988
5. R.A. McLaughlin and M.A. Alder, "The Hough transform versus the UpWrite", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 20, pp. 396-400, 1998
6. R.W. McLaren and H.-Y. Lin, "A knowledge-based approach to ship identification", *SPIE Proc. Conf. on Applications of Artificial Intelligence III, Orlando*, Vol. 635, pp. 257-267, 1986
7. V.F. Leavers, "Shape detection in computer vision using the Hough transform", Spring Verlag, Berlin, 1992
8. F.C. Luce and W.B. Schaming, "Automatic classification of ship targets", *RCA Engineer*, Vol. 31, pp. 18-23, 1986
9. M.S. Schmalz and F.M. Caimi, "Deformation-tolerant statistical correctors for enhancement of ship silhouette recognition", *SPIE Proc. Conf. on Infrared Image Processing and Enhancement, Orlando*, Vol. 781, pp. 148-158, 1987
10. R. Sharpe, "Jane's fighting ships 1997-1998", Jane's Information Group, London, 1997
11. C.J.S. deSilva, G. Lee, and R. Johnson, "All-aspect ship recognition in infrared images", *IEEE Proc. Conf. on electronic technology directions to the year 2000, Adelaide*, pp. 194-198, 1995
12. T.M. Veen and F.A.C. Groen, "Discretization errors in the Hough transform", *Pattern Recognition*, Vol 14, pp. 137-145, 1981
13. P.J. Withagen, "Automatic classification of ships from infrared (FLIR) images", TNO report FEL-98-S297, The Hague, 1998
14. B. Zvolanek and I.J. Kessler, "Autonomous ship classification from infrared images", *IEEE Proc. Conf. on Electronics and Aerospace System Conventions, San Francisco*, pp. 76-80, 1980