AUTOMATIC DETECTION OF SHIPS IN RADAR IMAGE SEQUENCES USING NEURAL NETWORKS

Giuliano BENELLI[(1)], Andrea GARZELLI[(2)], Alessandro MECOCCI[(1)]

(1) Dipartimento di Elettronica, Università di Pavia, Via Abbiategrasso 209, 27100 Pavia, Italy
(2) Dipartimento di Ingegneria Elettronica, Università di Firenze, Via S.Marta 3, 50139 Firenze, Italy

RÉSUMÉ

Dans cette communication on propose un système pour le contrôle automatique du traffic maritime à proximité des ports. Le système utilise des réseaux de neurones pour le traitement de séquences d'image formées par un système radar opérant dans la bande X. L'utilisation de réseaux de neurones accroît la précision par rapport aux techniques traditionelles. On indique les résultats fournis par le système proposé et par quelques techniques classiques de traitement des images. Le problème de la détection de l'angle de dérive a été étudié pour prévenir les situations de collision.

ABSTRACT

In this paper a system for the automatic control of ship-traffic in the access area of a seaport is presented. The system employs Artificial Neural Network techniques applied to sequences of X-band real-aperture radar images to achieve system reliability and measurement accuracy. A comparative study between neural and classical approaches is presented. A particular attention has been devoted to the detection of drift-angles for automatic collision avoidance.

1. INTRODUCTION

In ship traffic control, the analysis of the trajectory of a ship with respect to the position of the ship-prow is fundamental to prevent drifting situations and collisions.

In automatic collision avoidance systems using imaging radar sensors, real-time processing and high accuracy in feature location are requested. In this paper a neural system for the automatic control of ship traffic in the access area of a seaport is proposed. The system employs Artificial Neural Networks (ANN's), thus improving system reliability and measurement accuracy and reducing the computational complexity with respect to classical image processing techniques.

A real-aperture radar system can provide significant image sequences of the area near a seaport. In this study, images formed by a radar simulator have been processed, in order to test precision and reliability of the ship location and tracking process. The characteristics of the simulated radar system are summarized in Table 1.

The paper is structured as follows: in Section 2 the image segmentation phase is described. It is performed in two steps. First, a class membership probability is assigned to each pixel of the input image by using second order gray-level statistics obtained from the cooccurrence matrix. These probabilities then form the input of a multi-layer perceptron which performs the final supervised classification.

In Section 3 the feature-extraction phase is explained. Ship-prows are accurately located in order to estimate the motion directions of ships and compute possible drift-angles. The automatic location of ship-prows is performed in two steps, both employing multi-layer perceptrons.

Section 4 describes the phase of ship-tracking, which refines the ship-orientation estimates and provides a technique for motion prediction. The experimental results and a comparative study between the proposed feature-extraction neural system and some classical shape-analysis techniques are reported in Section 5.

2. SEGMENTATION

The image segmentation phase is performed in two steps. First, a probabilistic segmentation of the image is performed, based on the location of the intensities of each pixel and its neighbor within the cooccurrence matrix [1]. Since on-diagonal peaks characterize regions and off-diagonal peaks characterize boundaries, the cooccurrence matrix can be utilized as a feature space. Let $S_\Delta(i,j)$ denote the cooccurrence matrix corresponding to a displacement Δ

$$S_\Delta(i,j) = \sum_x \sum_{x'} \delta(i;I(x))\ \delta(j;I(x+\Delta))\ \delta(x';x+\Delta) \qquad (1)$$

where I(x) is the intensity of pixel x, the summations are extended to the whole image and $\delta(i;j)$ denotes the

Kronecker delta function δ_{ij}.

In order to define a segmentation algorithm, it is possible to relate the cooccurrence matrix of an image to the intensity and region probabilities of a set of images. The objective is to express the matrix in terms of the conditional probability for a pair of pixels to have observed intensities, given that their region membership is known. $S_\Delta(i,j)$ will contain a number of symmetric maxima R_a (a = 1,..,k) on the leading diagonal together with corresponding off-diagonal asymmetric maxima B_{ab} (a,b = 1,..,k; $a \neq b$). The classification assigns both region membership and whether the pixel is interior or boundary relative to the cooccurrence direction. It is expressed as a probability $p_x^\Delta(a,\epsilon)$, where a = 1,..,k denotes region and ϵ = 0,1 interior or boundary respectively. For example, the probability of x belonging to region R_a is

$$p_x^\Delta(a;0) = p_\Delta(ia;ja) / \sum_{a=1}^{k} p_\Delta(ia;ja) \quad (2)$$

where

$$p_\Delta(ia;ja) = p_\Delta(i|a;j|a)\, p_\Delta(a;a) \quad (3)$$

The conditional probability in (3) can be approximated by a Gaussian distribution

$$p_\Delta(i|a;j|a) = M_{aa}^{-1} \exp\left[-\frac{(i-i_a)^2 + (j-i_a)^2}{2\sigma_a^2} - \frac{(i-i_a)(j-i_a)}{2\sigma_{aa}^2} \right] \quad (4)$$

where M_{aa} is a normalization factor, the values i_a (a = 1,..,k) are given by the maxima on the main diagonal of $S_\Delta(i,j)$ and σ_a and σ_{aa} may be determined directly from the half-widths of the corresponding maxima in $S_\Delta(i,i)$.

The second phase uses the probability map generated by the first phase as the input of a multilayer perceptron which produces the final pixel classification.

Purely statistical classifiers consider only gray-level information, completely disregarding spatial relationships between pixels. This fact produces incorrect and noisy segmentations. Conversely, the ANN classifier improves spatial consistency and also adds expert knowledge by means of manually segmented images [2].

Training of the perceptron (with three layers) is achieved by Back Error Propagation. The input of the network is formed by the class membership probabilities of pixels over a given window W; the output is interpreted as a class assignment of the central pixel. W has been chosen of size 7x7, only two classes - ship and sea - have to be classified, hence the first layer is formed by 49 neurons and the output layer contains only one neuron. The hidden layer has been limited to five neurons.

3. FEATURE EXTRACTION

In the feature-extraction phase, ship prows are accurately located in order to estimate the motion directions of ships and compute possible drift-angles. The automatic location of ship-prows is performed in two steps. The first step generates search windows for prows from coarse-resolution images. The second step uses those search windows to locate prow features inside high-resolution images [3]. Both steps employ multilayer perceptrons. The multiresolution approach for prow detection is based on a pyramidal decomposition of the original images. This approach allows to reject a large number of false alarms that would be detected in a purely high resolution system and strongly reduces the computational complexity.

The purpose of the MLP in the coarse-resolution stage is to approximately locate the ship prows from the previously segmented images to form raw search regions. The image shrinking operation reduces each original image of the sequence by a factor of 2 or more for each direction, by a convenient multistep process. Each coarse-resolution image is raster scanned by a 5x5 window forming the 25 pattern vector presented to the MLP. Output values above a threshold imply a search region pixel for prows and are set to 1; those below the threshold imply a rejected pixel and are set to 0.

The total area covered by the search regions should be small enough (about 10% of the total image area or less) to maintain a reasonable level of computational complexity and to reduce mistakes in the high-resolution stage.

The fine-resolution stage takes the search regions generated by the coarse resolution stage and, using a raster scanning sliding window, searches within them for ship prow regions. As in the coarse-resolution stage, an MLP feature detector is used for the location process.

The orientation of a ship is then computed as the angle formed by the ship centroid and the ship prow previously detected. Starting from this angle, it is possible to estimate the angle of yaw, formed by the ship-centroid velocity vector and the ship-prow orientation.

The radar resolution affects the estimation of the prow position and consequently the computation of the ship orientation. In real-aperture mapping radars the resolution in the cross-range direction is proportional to the antenna beamwidth. For a rectangular antenna generating a fan beam, the approximate expression for half-power antenna beamwidth is given by

$$\eta = 51\, \lambda / L \quad (5)$$

where η is expressed in degrees, λ is the system wavelength and L is the horizontal dimension of the antenna. The azimuth resolution is obviously limited by the maximum antenna dimension, while the range resolution can be easily improved by reducing the duration of the transmitted pulse. The radar simulator used in this work simulates a radar system with η = 0.4 degrees and range resolution ΔR = 6.25m. If the distance d from the antenna increases, the

Table 1 - Main characteristics of the simulated radar system

Transmission frequency	10 GHz	Antenna beam (azimuth)	0.4°
Pulse Repetition Frequency	1600 Hz	Antenna beam (elevation)	20°
IF band	24 MHz	Antenna rotation	0.25 cps
Transmitter peak power	45 kW	Quote of radar site	50 m

range error introduced in the radar map can be neglected with respect to the cross-range position error. Thus, for d > 4÷6 km, we can assume that the image I_η of a ship in the radar map is a modified version of the ship image I_0 formed by a "maximum resolution" radar. In particular, I_η can be considered as the result of a stretching of I_0 along the azimuth direction. This shape distortion is irrelevant for the centroid position estimation, because of its symmetry properties. Conversely, it shifts the ship contour points by an amount that is proportional to the antenna beamwidth. A calibration algorithm has been implemented. It shifts the computed prow-position toward the scanning direction (increasing azimuth) if $180° < \alpha < 360°$ or toward the opposite direction (decreasing azimuth) if $0° < \alpha < 180°$ (Fig. 1). The amount of this shift is $\eta_r R$, where η_r is η expressed in radians. We assume that the side-lobes of the antenna produce a distortion in the ship image whose effect is eliminated by the image segmentation, while the distortion due to the main-lobe width is not affected by the binarization phase. Such considerations has been proved by some experimental tests regarding the computation of the length of a ship.

For all the images of the radar sequence, the coordinates of prows after calibration and of ship-centroids are memorized together with the angle of orientation of every ship.

4. SHIP TRACKING

Working on a whole sequence of images, the positions of all ships are tracked. First a matching operation is performed in order to pair each point from the set of the ship centroids at a given time with a point from the same set at the following time. The main requirement of this operation is to minimize the sum of the distances between the points in these pairs. This process is iterative and it is effective even if some ships appear in the scene or leave it [4]. The repetition of the coupling process for all the images of the sequence provides the direction of motion of every ship approaching the port. It should be noted that the system characteristics make the matching phase easy to perform. Indeed, ships move very slowly, almost insignificantly, during a scan period.

The subsequent phase refines the ship-position estimates and provides a technique for motion prediction. The α-ß filter, which is derived from the Kalman filter theory has been used [5]. The simple α-ß tracking filter is widely used in track-while-scan (TWS) radar applications. It is generally preferred to Kalman filter, because of its easier implementation and better flexibility. However, α-ß filter is less effective than Kalman filter in non-stationary environments, as during track initiations and maneuvering target tracking.

In its easiest form with constant parameters, the α-ß filter is defined by the following equations at time instant k:

$$
\begin{aligned}
i(k) &= x_m(k) - x_p(k) \\
x_s(k) &= x_p(k) + \alpha\, i(k) \\
v_s(k) &= v_s(k-1) + \frac{\beta}{T} i(k) \\
x_p(k+1) &= x_s(k) + T\, v_s(k)
\end{aligned}
\qquad (6)
$$

where $x_m(k)$, $x_p(k)$ and $x_s(k)$ are the measured, predicted and smoothed positions, respectively, $v_s(k)$ the smoothed velocity estimation, i(k) the innovation and T the sampling interval.

The optimal predictor (in the mean-square sense) minimizes the power of innovations. Therefore the optimal parameters α and ß depend on the statistics of the input signal which, in turn, depend on the assumed model for the target dynamics.

We assume that ships move along a straight line with constant velocity (case of non-maneuvering targets) or follow a circular trajectory with constant angular velocity (case of maneuvering targets).

In [5] a strategy to adaptively change the parameters α and ß under the above assumptions is presented. The objective is to minimize the error variance function. This adaptation method belongs to the stochastic gradient family, because it updates the parameters α and ß by an amount proportional to the negative gradient of the instantaneous quadratic innovation. This technique has been applied to solve the ship tracking problem.

5. EXPERIMENTAL RESULTS

Table 2 shows the performance of the proposed neural segmentation and of a threshold segmentation using an optimum binarization threshold T, which is computed from the statistics of the original image. Note that a lower final classification error is achieved by the neural system, especially using a 7x7 mask.

Table 2 - Comparison of classifier performance

Classifier	Accuracy (%)
Threshold segmentation	88.3
Neural (mask 3x3)	87.7
Neural (mask 5x5)	93.1
Neural (mask 7x7)	96.8

Table 3 shows some results about the automatic ship-course computation in the case of yaw = 0, for different processing techniques. The deviation of the prow-orientation from the actual value of the coarse-angle thus denotes a processing error. For aspect angles α differing from 90°, the estimation of the ship-prow orientation is affected by an error which depends on α and on the distance R of the ship from the radar site. Course-angles are computed by a classical image processing technique (CP) which locates the prow positions from the centroidal map of each ship contour and by the proposed neural system (NN). Subscript C denotes the application of the calibration algorithm described in Section 3. Each course-angle represents an average on an observation time of 40 s (10 scans).

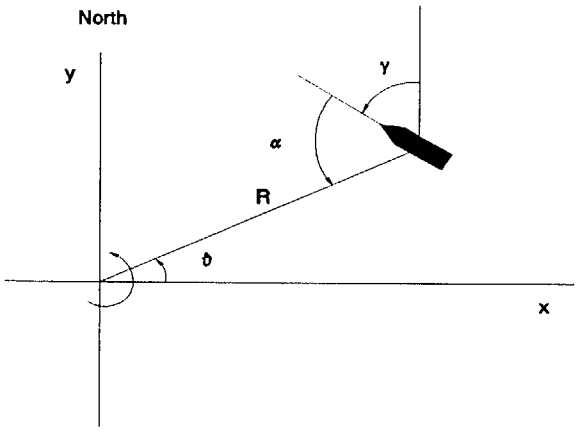

Figure 1 - Geometrical notation: α is the aspect angle, γ measures the prow orientation.

The experimental data show that the proposed centroid procedure does not introduce relevant errors if compared with the sensor uncertainty (6.25 m in range and 4÷6 m for every km of distance from the antenna in cross-range). This condition is strictly verified if the ship dimension and the distance from the radar site allow a shape-preserving representation of the target (ships longer than 80-100 m located at a distance lower than 6 km from the radar antenna have generated images correctly processed). Moreover, prow positions are generally computed with no error in the image plane, showing an extremely accurate behavior of the process of prow detection.

Table 3 - Performance of the ship-prow location process for different values of aspect-angles, for a classical (CP) and a neural (NN) technique. Subscript C denote the application of the calibration algorithm.

	Course angle (deg) (* = actual)				
Aspect angle (deg)	*	CP	NN	CP_C	NN_C
90	0	-1.0	-1.0	0.8	0.8
70	20	17.3	18.4	20.7	20.2
60	30	23.5	24.3	30.8	30.6
50	40	28.6	30.1	40.1	41.1
40	50	37.7	38.2	48.5	51.6
30	60	42.3	45.7	56.2	58.4

This work has been supported by *Alenia S.p.A.*

REFERENCES

[1] J.F. Haddon, J.F. Boyce, 'Image Segmentation by Unifying Region and Boundary Information', *IEEE PAMI*-12, 1990, pp. 929-948

[2] D.L. Toulson, J.F. Boyce, 'Segmentation of MR images using neural nets', *Image and Vision Computing*, 10, 1992, pp. 324-328

[3] J.M. Vincent, J.B.Waite, D.J.Myers, 'Automatic location of visual features by a system of multilayered perceptrons', *IEE Proc.-F*, 139, 1992, pp. 405-412

[4] J. Wiklund, G. Granlund, 'Tracking of multiple moving objects', *Proc. Int. Workshop of Time-Varying Image Processing and Moving Object Recognition*, Firenze, 1986, pp. 241-250

[5] M.G. Otero, J.M. Páez Borrallo, 'Adaptive algorithms for alpha-beta radar tracking', *Proc. Int. Conf. 'Digital Signal Processing-91'*, Elsevier Science Publisher, pp. 535-540