Development of Track Estimation System for Floating Object Surveillance

NGUYEN Minh Duc*, Monotobu IMASATO**, Yoichi FUTAKI*** and Takayuki ASANUMA****

Abstract

Floating object detection and tracking has always been an important task for ensuring marine traffic safety and other surveillance purposes. To enable on scene observation, NMRI has developed a hybrid observing system. The system consists of a Laser Camera, a Night-vision Camera and an Infra-red (IR) Camera. By processing the collected IR images, this study aims at developing a floating objects track estimating system to support observation and provide warnings. The system should be able to automatically collect IR camera images and detect floating objects from those images, transferring the object positions to the sea surface and estimate the objects tracks from their positions on consecutive camera images.

Keywords: Surveillance, IR image, Floating Object Detection, Track Estimation

1. Introduction

Floating object detecting and tracking has always been an important task for ensuring marine traffic safety, as well as in search and rescue missions. Thanks to the technology advancement, different techniques (e.g. radar, AIS, satellite based observation etc.) can be (or have been) used for this purpose. Each observing method has its own advantages and disadvantages and is therefore applied in appropriate fields. In an effort to support real time observation on scene, especially for small objects that are not detectable by ship radar, NMRI has developed a hybrid observing system⁽¹⁾. The system consists of a Laser Camera, a Night-vision Camera and an Infra-red (IR) Camera. These 3 different cameras are situated in a camera box located on a stabilizer. This stabilizer has the function of maintaining camera box in a horizontal plane while the ship, on which the system is installed, is moving with 6 degrees of freedom. The stabilizer is controlled automatically by a computer and can be rotated around the own ship heading. This enables the system to provide real time images of sea surface around the ship (camera) position.

Using the collected IR images, this study aims at developing a floating object track estimating system to support observation and provide warnings. For this target, the system must be able to solve the following tasks simultaneously:

- Collecting IR camera images and detecting floating objects from the images.
- Transferring the object positions from the image coordinate to the ship coordinate systems and then to absolute positions on the sea surface.
- Predicting the object track from its positions extracted from consecutive camera images.
- Providing warnings if the tracked floating

object is entering a Guard Area.

The system, in the software form, is installed in a computer connected to the IR camera. To get the camera position, communication is established between this computer and a GPS receiver. The computer is also connected to a Stabilizer Control computer to allow the access to camera direction data (pitch, roll, and yaw) as shown in Fig. 1.

Fig.1 NMRI Hybrid Observing System

2. Program Outline and Position Calculating Algorithm

2.1 Floating Object Tracking Program Outline

The system outline is illustrated by the flow chart in Fig. 2. Going down the flow chart, IR camera images are captured, using an image capturing board (matrox). Capturing interval can be decided by the user. In the study, this interval is set to 1 or 2 seconds, taking into account the existing interval of waves and floating object movement. Field of view of the IR Camera is 21.7 by 16.4 degrees. It uses 8-13μm wavelength, with minimum detectable

*Student Member: Tokyo University of Marine Science and Technology nmduchhh@yahoo.com

**Member: National Maritime Research Institute

*** NTT Communication Corporation, ****Japan Oil, Gas and Metals National Corporation

temperature difference of 0.08C, to produce 640 by 480 pixels images (see Ref.1 for more details).

Then, own ship position and course are extracted from the GPS receiver logs, using a RS-232C serial communication link. The pan data, which is necessary for determining camera direction, is derived from the Stabilizer-control computer through a local network cable. This dataset contains ship roll, pitch and stabilizer roll, pitch and yaw.

Next, the sea horizon is searched and floating objects are detected from the image. This is the main task of the program and will be discussed in later sections.

As mentioned above, to transform object positions from the image to the sea surface coordinate system, camera bearing must be known. In this step, own ship direction and camera pan data collected from previous steps are used. The transforming algorithm will be discussed in more detail in the next section.

Fig.2 Floating Object Tracking Program Outline

In the following steps, object tracks are predicted from its consecutive positions and the result is displayed to the user.

2.2 Object Position Calculating Algorithm

Equation (1) is used for camera bearing calculation. Starting with a unit vector e pointing north in a north east down (NED) coordinate system, vector e' is derived by rotating e through the ship yaw, pitch and roll angles.

In the same manner, the camera axis (illustrated by vector e'') is the result of rotating e' around axe

of ship body fixed coordinate system by the angles equivalent to the stabilizer yaw, pitch and roll respectively. This is realized by matrix multiplications.

$$e = [1 \ 0 \ 0]$$

$$e' = [e'_n \ e'_e \ e'_d] = R_{Ship} \times e$$

$$e'' = [e''_n \ e''_e \ e''_d] = R_{Stabilizer} \times e' \quad (1)$$

$$R_K = R_z^{K-yaw} \times R_y^{K-pitch} \times R_x^{K-roll}$$

K denotes [Ship] or [Stabilizer]

$$Cam_Bearing = \arctan(e''_e / e''_n)$$

Here, the rotation matrices (R_{Ship} , $R_{Stabilizer}$) are calculated from their conventional rotation matrices around the z-axis (yaw, R_z), y-axis (pitch, R_y), and x-axis (roll, R_x). In this study, NED (Fig. 3a) is an Earth-Fixed coordinate system centered at the camera position. $R_{Stabilizer}^T$ (i.e. transpose of $R_{Stabilizer}$) is, in fact, the matrix for transforming the camera axis to the ship longitudinal axis (e'), while R_{Ship}^T is for transforming the latter (axis) to the north-direction in NED.

Using (1), camera bearing can be derived from e'' north and east components in the NED coordinate.

Fig.3 NED and Camera Coordinate Systems

Transformation of object position on the image to its equivalent relative position is carried out by the following equation

$$\text{Object relative Position : } [X \ Y]$$

$$X = H / \tan(\alpha) \quad (2)$$

$$Y = X \times \tan(\beta)$$

where:

- X , Y : longitudinal and traverse distance, relative to camera position and bearing.
- H : Camera height from the sea surface.
- α : vertical angle between true horizon direction and the direction from the camera to the object.
- β : horizontal angle between camera axis and direction from camera to the object.

The two parameters α , β can be inferred from the image as illustrated in (Fig. 4), using camera vertical and horizontal view angles.

It can be seen from (2) that the relative position (X , Y) is most sensitive to error in α . Therefore, it is important to determine this parameter accurately.

Object position on the sea surface can be deduced from its relative position by

$$\begin{aligned} N_{Obj} &= N_{Ship} + X \cos(CB) + Y \cos(CB + 90) \\ E_{Obj} &= E_{Ship} + X \sin(CB) + Y \sin(CB + 90) \end{aligned} \quad (3)$$

where:

- N_{Ship} , E_{Ship} : Own Ship (Camera) position in an earth fixed coordinate system.
- N_{Obj} , E_{Obj} : Object position in the same system
- CB : Camera true bearing.

Fig.4 Camera Direction and Sea Horizon

Then (2), (3), together with the camera bearing defined in (1) make it possible to transform object position, from a point on the image to its equivalent sea surface position.

3. Horizon Line Detection

As stated above, detecting the sea horizon (and α accordingly) is an important task to ensure accuracy. The sea horizon is, in fact, an edge between the sea water and the sky. Therefore, an edge detection technique should be applied⁽²⁾⁽³⁾.

All the detecting techniques begin with the calculation of the image intensity gradient. As the sea horizon does not deviate largely from the image horizontal direction, just the horizon component (G_X) of the gradient is concerned. In this study, G_X is determined using convolving matrices defined as the followings:

$$\begin{aligned} G_X^1 &= \begin{bmatrix} 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 \end{bmatrix} \\ G_X^2 &= \begin{bmatrix} 0 & 0 & 1 & 1 & -1 & -1 & 0 & 0 \end{bmatrix} \\ G_X^3 &= \begin{bmatrix} 0 & 1 & 1 & 1 & -1 & -1 & -1 & 0 \end{bmatrix} \\ G_X^4 &= \begin{bmatrix} 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \end{bmatrix} \end{aligned} \quad (4)$$

$$Grad_i = Image * G_X^i \quad i = 1 \text{ to } 4$$

$$Grad = \prod_{i=1}^4 (Grad_i)$$

*: convolution operation

The image is convolved with 4 kernel matrices (G_X^{1-4}) respectively to give 4 gradient components. These components denote intensity variation for different frequencies, from high ($Grad_1$) to low ($Grad_4$). Then, the gradient is defined as the product of these 4 gradient components. The idea lying behind this tactic is that the intensity gradients are significant for all components at the sea horizon, but not for all of them at, e.g. wave edges.

To illustrate the technique applied, comparison has been made between the gradient calculated by the technique and the one determined using common Sobel operator with convolving matrix G_X

$$G_X = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \quad ImageGradX = Image * G_X \quad (5)$$

2 images in Fig.5 are captured by IR camera for different weather conditions.

Fig.5 Sea Horizon

In Fig.5a, water and air temperature difference is large, resulting in a clear edge between the sea water and the sky above. The gradient of this image is shown in Fig.6a, for Sobel operator (upper) and the operator used in this study (lower) for a vertical line drawn in Fig. 5. It can be seen that the lower figure provides a clear and significant peak for the point on the sea horizon. This peak is well over other wave edges and is easy to be detected. For Sobel operator, the gradient peak values are more similar for sea horizon and wave edges. Therefore, it is difficult to decide a threshold value for the sea horizon and wave edges separation.

Difference between the 2 approaches is even more obvious for the second case (Fig.5b). In this case, image was captured when the temperature difference between sea water and air is small, resulting in a vague horizon line, if it is detectable.

Fig.6a Gradient Comparison

In this example, it is almost impossible to recognize the horizon line peak if Sobel operator is applied (Fig.6b, upper). It is due the image nature that produces unwanted horizontal lines and edges of these lines are even more significant than the sea horizon edge when temperature difference is too low. On the other hand, the operator used in this study still gives a significant peak for the horizon line, as the result of accumulation of different frequencies intensity variations.

Fig.6b Gradient Comparison

These 2 images were captured in Jan. 2010. The average sea water temperature for this month was 15°C. The weather was clear for Fig.5a (27th, around 16:00), with air temperature (sea level) of 11.3°C. On 28th, at around 10:00 am (Fig.5b), it was rainy (0.5mm) and temperature was from 15.1 to 15.2°C⁽⁴⁾. However, quality of IR camera images depends on a variety of other conditions (e.g. cloud conditions, humidity etc.) which is hard to determine. Therefore, weather condition is not further mentioned in this paper (refer to Ref.1 for

more information in the performance of IR camera on different conditions).

Then, sea horizon detecting procedure is carried out as shown in the flow chart in Fig.7.

Fig.7 Sea Horizon Detection Flow Chart

After gradient calculation, edges which are not significant should be suppressed. The aim of this step is to remove wave edges to eliminate possibility of false detection. It is difficult to decide a single gradient Threshold Value for non significant edges removing for all weather conditions. Therefore, a better way to do is to keep just some most significant edges, especially those around the estimated position of the sea horizon.

Due to noise and wave effects, edges may be corrupted that causes discontinuities. To solve the problem, broken edge parts nearby and of similar tendency should be connected. The false connection (parts are not of the same edge) will be detected in the later step and, in almost all cases, will not cause false horizon line detection.

After connecting, edges are fitted with lines, taking into account the fact that with camera height of 10 – 15 meters, the sea horizon is very similar to a straight line.

From those fitting lines, the best fitting line will be selected. The best fit line is the one with maximum edge points on it and does not deviate largely from its estimated position. The estimated position can be inferred from ship attitude (roll, pitch) and camera stabilizer pan data (stabilizer roll, pitch and yaw). As these data have an accuracy of 0.5 degrees, the estimated sea horizon should not be used directly for object position calculation, but it gives a good approximation of where to search for the sea horizon.

To be taken as the sea horizon, the best fit line must satisfy the following decisive conditions:

- Number of edge points on this line must be larger than a threshold value.
 - Deviation from estimated sea horizon (roll, pitch differences) must be less than a threshold.
- The total deviation can be evaluated by (6):

$$\delta = \sqrt{\delta_{roll}^2 + c \times \delta_{pitch}^2} \quad (6)$$

Where c (> 1) is an adjusting factor used to put more weight on pitch deviation.

If fitting conditions are satisfied, the line is considered to be sea horizon and will be used for later calculation. However, there are still cases where the sea horizon is almost invisible such as those in Fig.8. For this case, use of estimated sea horizon is the only possible solution. Object position accuracy, accordingly, should be considered with care. In this case, detection procedure returns a fail.

Fig.8 Horizon Line Detection False Case

The horizon detection algorithm, if successful, can ensure the accuracy of horizontal direction (direction to true horizon, after adding correction for horizon dip, Fig.4) to be better than 0.06 degrees (2 pixels). Total effect of this error and other inaccuracy (the camera height, e.g.) is illustrated in Fig.14 and Fig.15.

4. Floating Object Detection

Constant false alarm rate (CFAR) method is used in this study for object detection. Using this method, the existence of an object is detected by the intensity (or brightness) difference between the object pixels and the surrounding background pixels, including noise, clutter and interference⁽⁵⁾⁽⁶⁾. If intensity difference between a pixel and its surrounding pixels is above a threshold, the pixel is considered to be an object pixel. If this threshold is too low, then more objects will be detected at the expense of increased numbers of false alarms. Conversely, if the threshold is too high, then fewer objects will be detected, but the number of false alarms will also be low. The method is used for cases in which it is difficult to decide object existence just from its intensity peak. For example, for the sea surface images, pixel intensity varies as a function of distance to camera, swell and wave pattern thus it is impossible to decide a single value for the peak intensity to decide whether it is an

object or simply the wave crest.

For floating object on the sea surface, its edge is normally clearer than wave edges. This can be taken into consideration in the object detection program. In Fig.9, an IR camera image in wavy sea condition is taken as example. Pixel intensity value variation for different directions (horizontal, vertical, $\pm 45^\circ$ upward) are shown in the graphs in this figure. In these graphs, a peak can be easily seen for a buoy at approximately 150 meter distance from camera position. Wave peaks, however, are also quite significant, in comparison with the object (the buoy).

Object detection procedure is carried out in the following 2 steps

Fig.9a Pixel Intensity Variation

Fig.9b Pixel Intensity Variation

Step 1: Candidate object determination (local peaks detection). To find the local peaks, local average intensity is first calculated for each region of the image (each horizon line or image row). The background average is then determined to be the average intensity of pixels with intensity value less than a threshold. This threshold is chosen to be a value between the local average and maximum intensity of the row.

After the local background average has been calculated, local peaks are defined to be those image pixels having intensity larger than local background average a certain amount.

Step 2: Repeat local intensity difference calculation with variable size frames.

In this step, local background average is calculated again for the area surrounding candidate object. Area size is chosen not too small to reduce noise effect in intensity variation. A guard region is also used (as in Fig.8) around the candidate object. Then, object pixels value will be compared again with this average local background and to be erased if intensity difference is less than a threshold. The process is repeated for the 4 main direction (horizontal, vertical, $\pm 45^\circ$). Therefore, wave peaks, of which edges are not significant, will be erased gradually.

Fig.10 Object Detection

In Fig.11, consecutive IR images are shown, and the objects detected in these images are marked by suitable rectangles.

Fig.11 Object Detection Result

The wave crests, however, may appear to be very similar to small objects. Objects and wave crests distinction, in this case, can just be done by scanning their existence through several images continuously.

5. Floating Object Track Prediction

The object track is predicted from its positions. The aim of track prediction is to check whether objects are floating into the Guard Area, which is an area behind our Own Ship. Then, it is necessary to gather the image objects of the same target on consecutive camera images.

In this study, relation between image objects (see Fig.12a) is evaluated by a relating value. The value takes into account similarities in the objects size (S_{rel}), distances to camera (D_{rel}) and bearings. The relation value is calculated by (7). Components of (7) are defined in the following equations. If the relation value of 2 image objects is less than a threshold and less than relating value with any other image objects, the 2 image objects will be considered to be objects on different images of a single target.

Fig.12 Object Track Prediction Principle

$$f(obj1, obj2) = S_{rel} \times D_{rel} \times B_{rel} \quad (7)$$

where

$$S_{rel} = LargeObjectSize / SmallObjectSize$$

$$D_{rel} = DistBetweenObjects / Limit_Dist$$

$$B_{rel} = (Delta_Bearing + 0.5) / Limit_Bearing$$

$$f(obj1, obj2) < Threshold_relation$$

$$\Rightarrow \text{image object of the Same Target}$$

In Fig.13, the objects detected in Fig.11 are plotted on the sea surface. Using (7), these objects are found to be of a single target.

Fig.13 Object Position and Track

However, due to errors in position determining algorithm (the sea horizon detection error, e.g.), the object positions are fluctuating about its track. Then, in this study, object track is predicted by least square method (LMS). LMS is used with the assumption that target movement is constant. This assumption is appropriate because target can not change its speed and course much in a short period of time (less than 1 minute).

In this LMS algorithm, target latest position (X_0 , Y_0) and its 2 velocity components (V_x , V_y) will be determined so as to minimize the total square error (denoted by J). For illustration, predicted track was calculated and drawn in Fig. 13.

$$X_i = X_0 - i \times \Delta X$$

$$Y_i = Y_0 - i \times \Delta Y$$

$$J = \sum (X_i^* - X_i)^2 + \sum (Y_i^* - Y_i)^2 \quad \text{or}$$

$$J = \sum (X_i^* - X_0 - i \times \Delta X)^2 + \sum (Y_i^* - Y_0 - i \times \Delta Y)^2$$

$$i = 0 \text{ to } \text{NumberOfPoints}$$

X^*, Y^* : Observed target position

X, Y : best fit target position

$\Delta X, \Delta Y$: best fit x, y speed

Due to the effects of different error sources (sea horizon detection error, object water line detection, camera height etc.), object position calculated by the program has some extent of uncertainty that depends largely on the relative distance to camera position. This dependence is illustrated in Fig.14 and 15. Objects material may also contribute some error to the accuracy due to the error in water line detection and should be studied further.

Fig. 14 shows the variation of distance from camera to a non moving object (an anchoring ship) at different distances. An increase in variation with increasing distance can be easily seen. For a target at 1100m, the deviation is up to 20m. The experiment was carried out on Oct. 28th 2009

(17:00 to 18:00). The weather was fine with air temperature of approximately 21.0°C and average sea temperature to be 22°C⁽⁴⁾.

To check the position calculation algorithm accuracy, distance to the object deduced from IR image is compared with its equivalent Lidar measurement and result is shown in Fig. 2. In this figure, a fluctuation of the former about the latter is noticed. It is due to the fact that camera height correction due to ship motion (rolling, pitching and heaving) is not applied. The experiment (20:00, Mar. 17th 2010) was in heavy weather (sea state), with air and average sea temperature to be 0.3°C and 4°C respectively⁽⁴⁾.

Fig.14 Tracked Distance Variation

Fig.15 Distance Comparison with Lidar Measurement

The track prediction, accordingly, pertains to some error. It depends on, among others, the distance to the camera and number of observation used. In our experiment with an unmoving object (buoy or small boat), speed error for distance of app. 1200m is 1.5 (m/s) (if 20 seconds of observation is used). This can be reduced by increasing the number of observations (50 seconds, e.g.) at the cost of more calculation needed. Further data should be taken to determine this fitting uncertainty, especially the course prediction.

6. Conclusions:

In this paper, possibility of using IR Camera for marine traffic observation and development of a floating object track estimating system has been studied. It is shown that, IR camera, if properly used, is a very effective tool for this purpose. The system is able to detect objects, predict their track and give warnings if the objects are floating into the Guard Area. For track predicting purpose, system is reliable for targets at less than 2000m. The further the target is, the less accurately object position can be estimated. Effectiveness of the system is, however, seriously reduced in bad weather condition.

Acknowledgement

This study was carried out as the JOGMEC (Japan Oil, Gas and Metals National Corporation) Public Offering Type R&D Project, "Study of Cooperative Navigation Support System aimed for Safe and Effective Operation of the Seismic Vessel using High-speed and Large-capacity network from Ship to Land -Study of the all-day all-weather marine surveillance technology-". The authors are thankful to whom it may concerned.

Further thanks should go to Professor Kohei OHTSU (TUMSAT) and crew members of ships for the cooperated experiments.

Reference

- (1) M. Sasano, J. Kayano, Y. Futaki and K. Maeda: Development of All-Day, All-Weather Hybrid Marine Surveillance System, Journal of Japan Institute of Navigation, Sep. 2010
- (2) E. Nadernejad: Edge Detection Techniques, Evaluation and Comparisons, Applied Mathematical Sciences, Vol. 2, 2008, pp. 1507-1520
- (3) J. Ponce: "Lecture 26, Edge Detection II", <http://www-cvr.ai.uiuc.edu/~ponce/fall04/lect26.ppt>
- (4) Japan Meteorological Agency: Weather, Climate & Earthquake Information, <http://www.jma.go.jp/jma/menu/report.html>
- (5) H. You and G. Jian: A New CFAR Detector with Greatest Option, Journal of Electronics, Vol. 14, 1997, pp. 125-132
- (6) T. Y. Liu, K.T. Lo et al.: A New Cut Detection Algorithm with Constant False-Alarm Ratio for Video Segmentation, Journal of Visual Communication and Image Representation, 2004, pp. 132-144

Question and answer

FURUSHO Masao (Kobe University):

What is the reason for you to treat with luminance but not illuminance?

NGUYEN Minh Duc:

As temperature range is not so large, to limit the volume of data transferred, grayscale images are

captured and pixel intensity value (0-255) is used for processing in this study. It is actually the "brightness" of the image. "Relative Luminance" (0 for pure black and 255 for white) might be an inappropriate word to use.

FURUSHO Masao (Kobe University):

Your research results might be significant for applying in SAR (Search and Rescue) at sea. How is the possibility of applying this system to the SAR issue?

NGUYEN Minh Duc:

Applying in SAR is really a potential use for the observing system developed in this study. Unfortunately, we have not yet been able to capture images of man at sea. However, experiments have shown that birds could be seen clearly in the IR images, so we strongly believe that human detection at sea is quite possible due to the temperature difference between man and sea water. We would like to gather data and study the cases in more details in the future.