Ship Detection Using Edge-Based Segmentation and Histogram of Oriented Gradient with Ship Size Ratio

Hyukmin Eum¹, Jaeyun Bae¹, Changyong Yoon², and Euntai Kim¹

¹School of Electrical and Electronic Engineering, Yonsei University, Seoul, Korea
²Department of Electrical Engineering, Suwon Science College, Hwaseong, Korea

Abstract

In this paper, a ship detection method is proposed; this method uses edge-based segmentation and histogram of oriented gradient (HOG) with the ship size ratio. The proposed method can prevent a marine collision accident by detecting ships at close range. Furthermore, unlike radar, the method can detect ships that have small size and absorb radio waves because it involves the use of a vision-based system. This system performs three operations. First, the foreground is separated from the background and candidates are detected using Sobel edge detection and morphological operations in the edge-based segmentation part. Second, features are extracted by employing HOG descriptors with the ship size ratio from the detected candidate. Finally, a support vector machine (SVM) verifies whether the candidates are ships. The performance of these methods is demonstrated by comparing their results with the results of other segmentation methods using eight-fold cross validation for the experimental results.

Keywords: Ship detection, Edge-based segmentation, Histogram of oriented gradient, Ship size ratio, Support vector machine

1. Introduction

Collisions between ships have not decreased over time, and the death toll from these accidents has exhibited a general tendency to increase. A shipmate or the harbor controller of a port observes the sea for a long time and performs the duty of preventing marine collisions when dangerous materials are found on the sea. However, it is difficult for a person to detect all of the obstacles on the sea and take action, and there are two hurdles to detection. First, people cannot concentrate for long periods of time. According to research, the maximum amount of time that a person can concentrate on an object for observation is 20 min, and the maximum amount of time that he/she can focus on complex actions such as sports is 60 min. Therefore, a person on a ship sometimes does not recognize upcoming obstacles.

Second, the viewing distance and field of view of humans are limited. In clear weather, a person can observe a distance of about 10 km at sea and his/her field of view is 90° at any point of time. However, the viewing distance and field of view required to detect a specific target are 5 km and 300°, respectively [1]. In addition, obstacle detection by only the sight of a person on a ship is more limited in areas where there are many buildings such as harbors and narrow channels. Therefore, collisions of small ships occurring in harbors and narrow channels can be attributed to the physical limitations of the person who observes the obstacles.
and the sea. Radar is used for overcoming the physical limitations of a person, but it does not detect small ships, wooden ships, or ships at close range [2, 3]. In the international regulations for preventing collisions at sea for a voyage, the sight of a shipmate has a higher priority than radar or the sailing equipment [4]. Because of the two abovementioned reasons, a ship detection method that can solve the problems related to the sight of a person and radar is needed. Currently, sound navigation and ranging, magnetic anomaly detectors, and satellite photographs are being used as aids for ship detection.

However, the above-mentioned devices have low accuracy and detection range, and object detection depends on the knowledge and experience of an expert who has undergone special training. A synthesis aperture radar system that combines the characteristics of radar and a camera has been developed to solve these problems, but it is expensive to install and operate, and is restrictively used only for military purposes since the operation time and locations are limited by the operating characteristics because of aircraft installation [5].

Therefore, we propose a method of vision-based ship detection to overcome the disadvantages of the existing auxiliary equipment for a voyage. The proposed method can be widely used in harbor operations, rescue operations, and marine military activities. The method is comprised of edge-based segmentation, HOG with ship size ratio and ship detection, and offers the advantage of being able to quickly detect ship in real time. Furthermore, detection performance is improved by using HOG with the ship size ratio. The ratio is considered on the basis of the appearance characteristics of a ship.

In this paper, a ship detection system is proposed; this method uses edge-based segmentation and histogram of oriented gradient (HOG) with the ship size ratio to prevent marine collisions. The remainder of this paper is organized as follows: The proposed ship detection system is described in Section 2. In Section 3, the details of proposed method are explained. Experimental results of our system are presented in Section 4. Finally, conclusions and future plans are discussed in Section 5.

2. Description of Entire System

The flow of the entire system is illustrated in Figure 1; this system performs ship detection on the basis of input images obtained using a fixed camera. When an input image is acquired, candidates are detected at the edge-based segmentation stage. In the edge-based segmentation part, the foreground is separated from the background by using a Sobel mask. If a candidate does not exist in the input image, another image is captured by the camera. Otherwise, the features of the candidate region are extracted by using the HOG descriptor and the ship size in the feature extraction part, and the candidate is then verified by using a weighted support vector machine (SVM) in the ship detection part. As a result, ships can be detected in the images through the above process.

3. Proposed Method

3.1 Edge-Based Segmentation

It is important that the target be separated from the rest of the image to detect a specific object in an image because of the detection speed in real time. In other words, the background is subtracted to find the object area in the image and the desired object is then verified in this area. Among the various existing segmentation methods (differential image, watershed technique, saliency map, etc.), we use edge-based segmentation because edge information adequately separates ships from the background unlike other information (color, texture, etc.) [6-9]. This can be attributed to the fact that ship information is sometimes similar to the background, and hulls (body of the ship) can be composed of different textures. Therefore, the candidates are detected by using the information of an edge, which is a boundary between the background and the ship.
The process of edge-based segmentation is shown in Figure 2. The edge information may be defined as a localized discontinuity of pixel values that exceed a threshold. In other words, the edge is the difference between pixel values that can be observed in the image, and includes useful information. Since calculations are easy and this information can be obtained quickly, it can be used in real time for segmentation.

Among a variety of mask types (Prewitt, Robert, Laplacian of Gaussian [LoG], etc.) to detect an edge, we use a Sobel mask because this mask does not lose information and subtracts the background unlike other masks, as shown in the experimental results [8, 10, 11]. The edge image is obtained by applying horizontal and vertical Sobel masks (1) based on a grayscale image, as shown in Figure 2(c).

\[
\begin{bmatrix}
-1 & 0 & 1 \\
-2 & 0 & 2 \\
-1 & 0 & 1
\end{bmatrix} \quad \begin{bmatrix}
-1 & -2 & -1 \\
0 & 0 & 0 \\
-1 & -2 & -1
\end{bmatrix}
\]

The candidate is determined by the amount of the detected edge information after morphology operations (dilation, closing, and opening) that can reduce noise are applied to the edge image, as shown in Figure 2(d)–(f) [12].

### 3.2 Histogram of Oriented Gradient with Ship Size Ratio

After candidate detection, the features of the candidate are extracted using an HOG descriptor with the ship size ratio. This descriptor is suitable for ship detection because the method is described by the distribution of the oriented gradients and is known for frequently using object detection [13, 14]. Moreover, we propose the use of HOG features with the ship size ratio to improve the performance of the ship detection.

The HOG features are created by calculating the number of oriented gradients in the local area of an image and are generated by the process shown in Figure 3. The detected candidate is divided into areas called cells that are connected to each other, and the respective histograms are created by using the magnitude and the orientation of the gradient in each of these cells. The feature vector is obtained by collecting the histograms, and the method enables the system to maintain consistent performance throughout the entire cell normalization process.

The process of HOG feature extraction can be divided into the following three steps: the first step calculates the gradient to apply a one-dimensional discrete differential mask for the horizontal and the vertical orientations as follows:

\[
D_X = \begin{bmatrix}
-1 & 0 & 1 \\
\end{bmatrix} \quad D_Y = \begin{bmatrix}
-1 & 0 & 1
\end{bmatrix}^T
\]

\[
I_X = I * D_X \quad I_Y = I * D_Y
\]

The magnitude and the orientation are calculated using (4).
The second step calculates the histogram for the divided cells. The orientation of the gradient obtained in the previous step belongs to the corresponding band on the basis of the number of pre-defined bins. The band distribution is \(0^\circ–180^\circ\) or \(0^\circ–360^\circ\) depending on the sign of the gradient. The third step locally normalizes the magnitude of the gradient, and the normalization method is divided into three types, as expressed in (5).

\[
\begin{align*}
L_1 - \text{norm} : f &= \frac{v}{\|v\|_1 + e} \\
L_1 - \text{sqrt} : f &= \frac{v}{\sqrt{\|v\|_1 + e}} \\
L_2 - \text{norm} : f &= \frac{v}{\|v\|_2 + e^2}
\end{align*}
\]  

where \(v\) denotes the normalized vector, \(\|v\|_k\) represents the k-norm, and \(e\) indicates a constant to evaluate descriptors densely, but it does not affect the results over a large range. In this study, we use \(L_1 - \text{norm}\).

Moreover, we consider the ship size ratio to improve the performance of the ship detection on the basis of the appearance characteristics of a ship, as shown in Figure 4. The ratio of the full length and the maximum width for the ship does not exceed \(1:3\) [15]. The proportion of the full length and the overall height is \(1:1\), and the ratio of the full length and the upper works is \(1:10\) [16].

Therefore, if the aspect ratio of a candidate is more than \(1:10\), the candidate is most likely not a ship. The aspect ratio is added to the feature vector with the HOG features. For example, if the number of cells and bins in the HOG descriptor is \(4 \times 2\) and \(9\), respectively, a 36-dimensional vector is created, and the ship size ratio is added to the 37th dimension. When the feature of the ship size ratio is added, the deviation of the ratio information is larger than that of the HOG feature values. Therefore, we reduce this ratio to an appropriate value by using (6).

\[
SSR_{\text{new}} = \frac{SSR_{\text{origin}}}{C_X + C_Y + B}
\]

where \(SSR_{\text{new}}\) denotes the new ratio and \(SSR_{\text{origin}}\) represents the original ratio. \(C_X\), \(C_Y\), and \(B\) indicate the cell sizes of the x and y axes, and the number of bins, respectively. Thus, we extract features by using an HOG descriptor and the ship size ratio for ship detection.

### 3.3 Ship Detection

A ship is detected by using the feature vector obtained through the above procedure, and the ship detection method uses an SVM as the supervised learning method because of its fast rate of object detection [17-19]. The SVM calculates the hyperplane that can maximize the margin between classes and is divided into hard and soft margin techniques. We use a weighted nu-SVM which is the soft margin technique and considers the penalty and the characteristics of the distribution for feature vector. (7) presents the optimization problem of the nu-SVM.

\[
\min \left\{ \frac{1}{2} \|\omega\|^2 - \nu \rho + \frac{1}{m} \sum_i \xi_i \right\} \\
(\text{s.t.}) \ y_i (\langle \omega, x \rangle) + b \geq \rho - \xi_i, \ \xi_i \geq 0, \ \rho \geq 0
\]

The nu-SVM is advantageous for the selection of an effective error penalty [20]. Furthermore, it is means that weight in weighted nu-SVM is assigned to feature vector. In training step, the weighted nu-SVM is learned by the feature vector for training data set. After detecting candidates through the edge-based segmentation in an input image, the weighted nu-SVM verifies whether the detected candidates are ships in testing step.

### 4. Experimental Results

#### 4.1 Experimental Environment

In this study, we use a video recorded directly; its resolution is \(320 \times 240\). The specifications of the computer used in the experiment are Intel Core 2 Quad 2.4-GHz processor and 4-GB RAM. Approximately 11,000 image frames are used for evaluating the proposed system and include eight types of ships with different sizes and shapes for a variety of experiments as shown in Figure 5.
4.2 Segmentation Results

The performance of the proposed system is verified by comparing the obtained segmentation results with the results of other segmentation methods, namely differential image method, saliency, watershed technique, thresholding technique, Prewitt mask, Robert mask, and LoG mask, as shown in Figure 6 [6-8].

The results show that the differential image method cannot detect an anchored ship because of its use of the difference between the previous and the current images. The saliency map often detects only a portion of the ship and occasionally does not find the ship part as a candidate. The watershed technique frequently divides a ship into a number of candidates or detects a candidate together with the ship and ground areas. If the background is similar to a hull, the thresholding technique cannot detect a candidate for the ship part. The edge masks exhibit no significant performance difference, but the Sobel mask sometimes shows a small edge inside the ship and exhibits less wake noise or illumination noise than the other masks. The confusion matrix is defined to calculate the evaluation metrics as shown in Table 1.
Table 1. Confusion matrix for performance evaluation

| Division | Predict |
|----------|---------|
|          | Positive | Negative |
| Actual   | Positive | False negative (FN) |
|          | True positive (TP) | |
| Negative | False positive (FP) | True negative (TN) |

Table 2. Candidate detection performance

| Segmentation method            | Precision | Error rate |
|--------------------------------|-----------|------------|
| Differential image             | 0.14908   | 0.75093    |
| Saliency map                   | 0.39233   | 0.46613    |
| Watershed technique            | 0.09180   | 0.90792    |
| Thresholding technique         | 0.29161   | 0.55719    |
| Prewitt mask                   | 0.36807   | 0.57577    |
| Robert mask                    | 0.36372   | 0.57981    |
| Laplacian of Gaussian mask     | 0.38947   | 0.55073    |
| Sobel mask                     | 0.40460   | 0.53835    |

The various evaluation metrics can be calculated by Table 1.

\[
\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN}
\]

\[
\text{Recall} = \frac{TP}{TP + FN}
\]

\[
\text{Precision} = \frac{TP}{TP + FP}
\]

\[
\text{Error rate} = \frac{FP}{TP + FP + TN + FN}
\]

\[
\text{False negative rate} = \frac{FN}{TP + FN}
\]

\[
\text{False positive rate} = \frac{FP}{FP + FN}
\]

\[
\text{True negative rate} = \frac{TN}{TP + FP}
\]

The results for all the considered segmentation methods are shown in Table 2.

The candidates using edge masks are generally consistent with the ship area, and the precision is higher than that obtained in the case of the other methods; however, the LoG mask is sensitive to wake and illumination. On the other hand, the other methods have a problem in that they frequently detect only a portion of the ship. Therefore, we use a Sobel mask, which exhibits the best performance.

4.3 Feature Extraction and Ship Detection Results

The detection performance of the proposed system is verified by using eight-fold cross validation. First, we have to select the parameter values (the number of cells and bins) for feature extraction that yield the best performance, and change the number of cells to $2 \times 2$, $3 \times 3$, and $4 \times 4$ and the number of bins to 7, 9, and 11. The best accuracy is confirmed in the case of $4 \times 4$ cells and 11 bins. Then, we evaluate the proposed method using C-SVM with a linear kernel as shown in Table 3.

The proposed system demonstrates a slightly better performance through these results, but we want to evaluate the C-SVM here. Therefore, we compare the performances of various types of SVMs and SVMs with different kernels, as shown in Figure 7 [17, 20-22].

As shown in Figure 7, the weighted nu-SVM using the RBF kernel demonstrates the best performance; its accuracy is 0.936 [23]. Furthermore, evaluation metrics such as recall and precision are higher than those of the other types.

5. Conclusion

In this paper, we propose methods for detecting ships using edge-based segmentation and HOG descriptors with the ship size ratio. By using a Sobel mask, the foreground is separated from the background in the input images and candidates are detected in the segmentation part. The features are extracted by
employing an HOG descriptor with the ship size ratio, and the candidates are then verified by applying a weighted SVM.

Figure 7. Accuracy results for various types of SVMs and SVMs with different kernels. SVM, support vector machine; RBF, radial basis function.

Among various types of edge masks, the segmentation method using a Sobel mask shows that the edge of a ship adequately appears and the edge of light reflecting from the water appears the least. Further, it is not appropriate to use segmentation methods based on saliency maps and multi-difference images in candidate detection for detecting ships because the performance differences according to the image characteristics are noticeable. In the feature extraction part, an HOG descriptor with the ship size ratio exhibits a slightly better performance than a common HOG descriptor, and the weighted nu-SVM shows the best performance in the ship detection part.

The proposed system exhibits good performance and results in the experiment. However, the problems of an irregular decrease in the detection rate and a long operating time in the case of high-resolution images may occasionally occur. Therefore, it is necessary to solve these problems and thus improve the method. Furthermore, ship detection requires the use of images taken under a variety of weather conditions by a camera installed on an actual ship to automatically avoid collisions.

Conflict of Interest

No potential conflict of interest relevant to this article was reported.

References

[1] D. Cornish and D. Dukette, The Essential 20: Twenty Components of an Excellent Health Care Team. Pittsburgh, PA: RoseDog Books, 2009.

[2] D. Y. Kim, G. K. Park, and H. Y. Kim, “A study on the ship information fusion with AIS and ARPA radar using by blackboard system,” Journal of Korean Institute of Intelligent Systems, vol. 24, no. 1, pp. 16-21, 2014, http://dx.doi.org/10.5391/JKIIS.2014.24.1.016

[3] H. Lee, E. K. Kim, and S. Kim, “A study on fuzzy logic based clustering method for radar data analysis,” Journal of Korean Institute of Intelligent Systems, vol. 25, no. 3, pp. 217-222, 2015, http://dx.doi.org/10.5391/JKIIS.2015.25.3.217

[4] Lloyd’s Register Rulefinder, “COLREGS: International Regulations for Preventing Collisions at Sea,” Available http://www.mar.ist.utl.pt/mvventure/Projecto-Navios-I/IMO-Conventions%20(copies)/COLREG-1972.pdf

[5] H. J. Kim and J. H. Choi, “The phase error correction scheme using the iterative signal bandwidth estimation in SAR imaging system,” in Proceedings of the Korean Institute of Intelligent Systems Conference, Daegu, Korea, 2000, pp. 463-464.

[6] F. C. Monteiro and A. Campilho, “Watershed framework to region-based image segmentation,” presented at the 2008 19th International Conference on Pattern Recognition, Tampa, FL, December 8–11, 2008, pp. 1-4. http://dx.doi.org/10.1109/icpr.2008.4761587

[7] G. Fritz, C. Seifert, L. Paletta, and H. Bischof, “Attentive object detection using an information theoretic saliency measure,” in Attention and Performance in Computational Vision, L. Paletta, J. K. Tsotsos, E. Rome, and G. Humphreys, Eds. Berlin: Springer-Verlag, 2005, pp. 29-41.

[8] A. McAndrew, Introduction to Digital Image Processing with Matlab, 1st ed. Boston: Thompson Course Technology, 2004.
[9] Y. H. Baek and S. R. Moon, “Color edge detection using variable template operator,” International Journal of Fuzzy Logic and Intelligent Systems, vol. 6, no. 2, pp. 116-120, 2006. http://dx.doi.org/10.5391/ijfis.2006.6.2.116

[10] R. Milanese, S. Gil, and T. Pun, “Attentive mechanisms for dynamic and static scene analysis,” Optical Engineering, vol. 34, no. 8, pp. 2428-2434, 1995. http://dx.doi.org/10.1117/12.205668

[11] S. T. Seo, K. Sivakumar, and S. H. Kwon, “Dempster-Shafer's evidence theory-based edge detection,” International Journal of Fuzzy Logic and Intelligent Systems, vol. 11, no. 1, pp. 19-24, 2011. http://dx.doi.org/10.5391/ijfis.2011.11.1.019

[12] R. M. Haralick, S. R. Sternberg, and X. Zhuang, “Image analysis using mathematical morphology,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 9, no. 4, pp. 532-550, 1987. http://dx.doi.org/10.1109/TPAMI.1987.4767941

[13] N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” in Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2005, San Diego, CA, 2005, pp. 886-893. http://dx.doi.org/10.1109/CVPR.2005.177

[14] H. Eum, C. Yoon, H. Lee, and M. Park, “Continuous human action recognition using depth-MHI-HOG and a spotter model,” Sensors, vol. 15, no. 3, pp. 5197-5227, 2015. http://dx.doi.org/10.3390/s150305197

[15] C. G. Soares and T. Moan, “Model uncertainty in the long-term distribution of wave-induced bending moments for fatigue design of ship structures,” Marine Structures, vol. 4, no. 4, pp. 295-315, 1991. http://dx.doi.org/10.1016/0951-8339(91)90008-y

[16] P. F. Chen and C. H. Huang, “An inverse hull design problem in optimizing the desired wake of ship,” Journal of Ship Research, vol. 46, no. 2, pp. 138-147, 2002.

[17] H. Byun and S. W. Lee, “A survey on pattern recognition applications of support vector machines,” International Journal of Pattern Recognition and Artificial Intelligence, vol. 17, no. 3, pp. 459-486, 2003. http://dx.doi.org/10.1142/S0218001403002460

[18] H. Eum, J. Lee, C. Yoon, and M. Park, “Human action recognition for night vision using temporal templates with infrared thermal camera,” in Proceedings of 2013 10th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI), Jeju, Korea, 2013, pp. 617-621. http://dx.doi.org/10.1109/URAI.2013.6677407

[19] C. C. Chang and C. J. Lin, “LIBSVM: a library for support vector machines,” ACM Transactions on Intelligent Systems and Technology, vol. 2, no. 3, pp. 1-27, 2011. http://dx.doi.org/10.1145/1961189.1961199

[20] F. Perez-Cruz, J. Weston, D. J. L. Herrmann, and B. Scholkopf, “Extension of the nu-SVM range for classification,” in Advances in Learning Theory: Methods, Models and Applications, J. Suykens, G. Horvath, S. Basu, C. Micchelli, and J. Vandewalle, Eds. Amsterdam: IOS Press, 2003, pp. 179-196.

[21] Y. Chen, X. S. Zhou, and T. S. Huang, “One-class SVM for learning in image retrieval,” in Proceedings of 2001 International Conference on Image Processing, Thessaloniki, Greece, 2001, pp. 34-37. http://dx.doi.org/10.1109/icip.2001.958946

[22] J. Park, J. Kim, H. Lee, and D. Park, “One-class support vector learning and linear matrix inequalities,” International Journal of Fuzzy Logic and Intelligent Systems, vol. 3, no. 1, pp. 100-104, 2003. http://dx.doi.org/10.5391/ijfis.2003.3.1.100

[23] D. M. Powers, “Evaluation: from precision, recall and F-measure to ROC, informedness, markedness & correlation,” Journal of Machine Learning Technologies, vol. 2, no. 1, pp. 37-63. 2011.

Hyukmin Eum received his B.S. degree in Electronic Engineering from Pai Chai University, Daejeon, Korea. He is a graduate student of the combined master’s and doctoral degree programs at Yonsei University. His current research interests include intelligent vehicle system, computer vision, pattern recognition, machine learning and fuzzy application systems.

E-mail: hmeum@yonsei.ac.kr

Jaeyun Bae received his B.S. degree in International Relationship from Republic of Korea Naval Academy, Changwon, Korea. He is a graduate student of the master’s degree programs at Yonsei University. His current research interests include intelligent vision system, computer vision, pattern recognition, machine learning and fuzzy application systems.

www.ijfis.org
Changyong Yoon received B.S., M.S., and Ph.D. degrees in Electrical and Electronic Engineering from Yonsei University, Seoul, Korea, in 1997, 1999, and 2010, respectively. He was a senior research engineer in LG Electronics, Inc., and LG-Nortel, and he developed system software for the DVR and WCDMA from 1999 to 2006. From 2010 to February 2012, he was a chief research engineer in LG Display and developed the circuit and algorithms in touch systems. Since 2012, he has been a Professor in the Department of Electrical Engineering, Suwon Science College, Korea. His main research interests include intelligent transportation systems, pattern recognition, robot vision, and fuzzy application systems.

E-mail: cyyoon@ssc.ac.kr

Euntai Kim received his B.S., M.S., and Ph.D. degrees in Electronic Engineering from Yonsei University, Seoul, Korea, in 1992, 1994, and 1999, respectively. He was a full-time lecturer with the Department of Control and Instrumentation Engineering, Hankyong National University, Anseong, Korea from 1999 to 2002. Since 2002, he has been with the faculty of the School of Electrical and Electronic Engineering, Yonsei University, where he is currently a Professor. He was a visiting researcher at Berkeley Initiative in Soft Computing, UC at Berkeley, Berkeley, CA, USA. His current research interests include computational intelligence and statistical machine learning and their application to intelligent robot, vehicle, and machine vision.

E-mail: etkim@yonsei.ac.kr