System for determining the parameters of a sea vessel by analyzing a photo image

Studenikin D E, Kuku E A

Admiral Ushakov Maritime State University, 93 Lenin Ave., Novorossiysk, 353924, Russian Federation

E-mail: Dm.studenikin@mail.ru

Abstract. In this paper, a method is considered for determining the parameters of a sea vessel, such as its spatial orientation and the distance from the observation point to this vessel from a photographic image to facilitate the task of identifying vessels at night. A measuring system for practical implementation is proposed, consisting of an optical segment, a segment for determining the angle of the vessel by the image, a segment for determining the distance to the object, and a segment for accumulating and filtering data. An algorithm for extracting the parameters of ship lights from a photographic image, their analysis, and the calculation of the quantities required for classification are described. The effectiveness of various classifier architectures for determining the angle of the vessel was experimentally tested, among which the SVM architecture was the most effective. A method for compiling a "depth map" for a static image based on the data of real distances to objects in the daylight image and the coordinates of the corresponding pixels in the same image is described. The method of backpropagation of the error is used for the obtained distances in the corresponding segment of the system based on the existence of the position-distance mapping. The model of the "depth map" constructed based on these data made it possible to obtain a sufficient distance to the object from the photograph.

1. Introduction

The use of computer vision technologies as a method of non-contact measurement of parameters of a sea vessel from photographs has already been considered for cases of ships approaching the berth at distances from the object of photography to the camera not exceeding 500 meters [1]. By the method of analyzing photographic images of sea vessels, it is also possible to solve the problem of identifying traffic vessels in the water area, as discussed in [2]. The application of these methods is possible only for daytime visibility conditions.

In addition, the concept of the functioning of a computer vision system based on neural network technologies has already been proposed, which solves the problem of organizing visual observation directly from the navigation bridge of a sea vessel [3].

Visual observation of sea vessels in the port water area is difficult at night, even in favorable weather conditions. At the same time, it is not possible to control the position of ships other than according to the data of radio navigation systems.

The initial data for the operation of the proposed method are images of a camera, invariably oriented in space, installed in a known position at a suitable elevation above the water level.
Vessels with a mechanical engine underway, according to rule 23 of the COLREGs-72, must, inter alia, carry navigation lights of red and green colors on the port and starboard sides, respectively. Thus, when considering such vessels, the presence of lights of the indicated colors greatly facilitates the task of determining the course of the vessel by analyzing images. This paper deals with a more complex case for sea-going vessels at anchor that light the lights prescribed by regulation 30 COLREG-72 for vessels at anchor.

To solve this problem, in this work, we propose a concept of an optical-measuring system intended for obtaining and subsequent analysis of a photographic image of the water area.

2. System structure
The above system consists of the following functional elements:

1) Optical segment - is a camera mounted on a gyro-stabilized platform. The technical parameters of the camera and the resulting image (the focal length of the lens, the resolution of the final image), as well as the values characterizing the spatial location and orientation of the camera (geographical coordinates, altitude, true bearing of the direction of the main optical axis of the lens and the angle of its inclination relative to the Earth's surface), are known.

2) Segment for determining the view of the vessel, the functioning algorithm of which is described in paragraph 3 of this work.

3) The segment for determining the distance to the object - operates on the principle of extracting from the "Depth Map" a snapshot of the value of the real distance to the object being determined based on information about its location on the image.

4) Segment of storage of distance data - accumulates and stores data on the ratio of real distances with areas of the photographic image for a long period, based on which the “Depth Map” of the photographic image is subsequently filtered and refined.

3. Determination of the view of the vessel.

3.1 Methodology
Determination of the view of the vessel is carried out by the corresponding segment of the proposed optical-measuring system.

The initial images during the study were images of a computer model of the ship, recreated from a real photograph. To achieve plausibility, the images of the computer model were artificially noisy.

The entire array of original images is divided into 12 views of the vessel with a step of 30 °, starting from the direction of the bow of the center plane clockwise.

![Figure 1](image_url)

**Figure 1.** a - The beams show the views of the vessel under consideration (schematic above); b - the image of the vessel from the angle 9.
In night photographs, the objects characterizing the spatial orientation of the ship are the sources of ship lighting - both navigation lights and deck lighting. The first stage of the method considered in the work is to extract several characteristics of the mentioned lights.

The original image is an RGB image. To separate the objects of interest to us from unnecessary information (in this case, almost a black background), the image is converted from a three-dimensional array to a one-dimensional one, where each cell corresponds to the sum of RGB components of the original image. After that, a certain threshold value is set at a cell value below which, the array elements will change the value by 0, higher - by 1. The resulting one-dimensional array contains areas of value 1, corresponding to the location of the lights in the original image.

For each area (figure 1), its center of gravity, its coordinate on the image canvas (row and column), the largest width and height of the figure, and the color of the “center of gravity” pixel are determined.

The original data for the entire sequence of images is stored in a 3D array.

The spatial orientation of the vessel relative to the camera is characterized by the relative position of the overall (key) points of the hull shape in the image. In conditions of night visibility, the indicated points can be approximated by specific lights. According to rule 30 COLREGS-72, power-driven vessels at anchor shall exhibit an all-round white light at the bow of the vessel where it can best be seen. Consequently, on the foreshortenings of the vessel, in which the bow is not obscured from the camera by the superstructure, one of the extreme white lights will be the bow light.

Following the provisions of MSC.1 / Circ.1216, vessels carrying dangerous cargo on board within ports must light an all-round red light at night, which is usually located near the top of the mast. Consequently, the red light in these images will be practically the extreme point of the mast. Vessels less than 50 m in length may exhibit an all-round white light where it can best be seen. Consequently, in the absence of dangerous cargo on board and an all-round red light, the upper dimensional point of the vessel may be an all-round white light. In the absence of the above conditions, deck lighting sources, in most cases, are located on the upper part of the ship's superstructure. In this case, the upper part of the superstructure will also be indicated by the overhead light on the ship image.

Two options are possible regarding the location of the third dimension point of the vessel.

The first is that the stern light of the ship at anchor is visible from the foreshortening of the vessel, which, according to the Rules, should be located at the stern or as close as possible to it at a level lower than the bow light mentioned earlier. Thus, one of the extreme lights in the image will be the stern. The second variant of the arrangement concerns the angles at which the stern light is obscured from the camera by the superstructure. In this case, the side light of the superstructure farthest from the observer is considered the side light, that is, also the light on the right or left in the image.

Based on the foregoing, within the data array corresponding to one image, three “overall” elements of the ship's shape are identified: the light located above all the others (counting from the bottom edge of the image), the leftmost light, and the rightmost light.

Figure 2. The above-mentioned lights are marked with a red frame.
After all, three lights have been identified, the distances in pixels are calculated from the topmost to the left fire, and then, from the top to the right. Since the coordinates of the centers of the lights were previously determined, the ratios of a right-angled triangle can be used to determine the distance:

$$D_{ik} = \sqrt{(x_i - x_k)^2 + (y_i - y_k)^2}$$

(3.1.1)

Where $D_{ik}$ is the distance between pixels $i$ and $k$, $x_i$, $x_k$, are the abscissas of the centers of gravity of lights I and k, $y_i$, $y_k$, are the ordinates of the centers of gravity of lights i and k.

For each image, two distances are obtained in this way. In different images of the same view of the ship model, the distance differs since the images were taken at different distances from the computer model. To normalize the distance D on the images, the following operation was performed.

The smaller of the two distances becomes equal to the ratio of the smaller distance to the larger one, the larger distance is then taken as one. Further, the distance from the top light to the leftmost light is taken to be negative, and the distance from the topmost to the rightmost light is always positive. After that, right-angled triangles are considered, formed by the upper light alternately with the right and left lights. The vertex of the right angle will be the point with the ordinate of the left (right) fire and the abscissa of the upper one. Having calculated the length of the legs, the tangent of the angle at the upper fire is determined, for which the value of the arctangent in radians is then determined. The said angles and distances are indicated in image III, where $D_1$ and $D_2$ are the distances from the upper light to the left and right, respectively; $\alpha$ and $\beta$ are the angles for which the tangent value is calculated.

![Figure 3](image3.png)

Figure 3. Explanation of calculations using the example of the image.

Thus, for each image, we got 4 parameters: 2 relative distances from the top fire to the right and to the left, and 2 sharp corners at the top fire.

3.2 Classification

Various methods of constructing a classification model were tested to solve classification problems.

The support vector machine showed the highest accuracy.

The initial data are presented in the following form:

$$\{\langle c_1 \rangle, \langle x_2, c_2 \rangle ... \langle x_n, c_n \rangle\},$$

(3.2.1)

Where $c_i$ takes the value 1 or -1 depending on which class $x_i$ belongs to ($c_i = \pm 1$), $x_i$ is a p-dimensional real vector $x_i \in \mathbb{R}^p$. Support vector machine is the construction of a p-1 hypersurface in p-dimensional space, which will best divide the data into two groups.

$$f(x) = x^T \beta + b = 0$$

(3.2.2)

Where $\beta \in \mathbb{R}^p \wedge bisarealnumber$.

The optimal solution to the problem is to find the surface that best separates the data. This is achieved by choosing such parameters $b$ and $\beta$ that will minimize $|\beta|$ in such a way that for each data sample $(x_i, c_i)$ the condition
Support vectors will be \( x_j \) on the data boundary for which
\[
y_j f(x_j) = 1
\]
(3.2.4)

It is computationally easier to solve the double quadratic programming problem. To obtain duality, the positive Lagrange multipliers \( \alpha_j \), multiplied by each constraint are subtracted from the objective function:
\[
L_p = \frac{1}{2} \beta' \beta - \sum_j \alpha_j (y_j (x'_j \beta + b) - 1),
\]
(3.2.5)

Where are we looking for stationary points of the function \( L_p \) by arguments \( b \) and \( \beta \). Setting the values of the gradient\( L_p \) to 0 we get:
\[
\beta = \sum_j a_j x_j y_j;
\]
(3.2.6)
\[
0 = \sum_j a_j y_j
\]
(3.2.7)

Substituting into the original equation, we get the dual \( L_D \):
\[
L_D = \frac{1}{2} \beta' \beta - \sum_j \alpha_j a_k y_j x'_j x_k.
\]
(3.2.8)

Which we maximize with respect to \( \alpha_j \geq 0 \). In general, many values of \( \alpha_j \) will be maxima. The nonzero values of \( \alpha_j \) in the solution define the hypersurface, \( k \) is seen in equation 1, which defines \( \beta \) as the sum of \( \alpha_j x_k y_k \). Sample data \( x_j \), corresponding to non-zero \( \alpha_j \) will be a support vector.

The derivative of \( L_D \) with respect to a nonzero \( \alpha_j \) is equal to 0 at the optimum point. This gives:
\[
y_j f(x_j) - 1 = 0
\]
(3.2.9)

This expression gives the value of \( b \) for any \( j \) nonzero element \( \alpha_j \).

In general, the method is used to solve binary classification problems. Since in this work the number of classes is 12, the One-To-Rest approach was used to solve the multiclass classification problem, the essence of which is to split the multiclass classification problem into several binary classification problems. While solving each problem, only samples of one class were separated from all the others; for the entire space of features, several surfaces were built that “separate” a specific class from the rest of the set of objects.

The prediction accuracy reached 92.4%.

4. Determination of the distance.

4.1 Methodology

The functioning of the segment for determining the distance to the object is based on the principle of analyzing the photographic image and calculating certain dependencies between the location of the object in the image and the real distance from the camera to it. Solutions to the problem under consideration were proposed for photographs obtained using stereo cameras in [4, 5, 8]. Proceeding from the assumption that the operation of the coastal optical system will be carried out according to the principle of classical photo and video cameras, the mentioned solution of the problem is inapplicable to the conditions under consideration.

In [6] it was proposed to determine the distance using a depth map compiled using neural network technologies. In [7] the specified photometry problem is solved by calculating and converting the angular values between the rays to the center of the image and the direction to the object.

The functioning of the segment for determining the distance to the object in the photographic image is based on the following algorithm.

Distance measurement is based on the principle of plotting the relationship between the coordinates of an object in the photograph and the actual distance from the camera to it. In other words, the "Depth Map" is being built. Due to perspective distortions, the dependence of the depth map value on the
pixel coordinates will not be linear [9,10]. With enough landmarks with a known distance in the image, it becomes possible to interpolate between the known values of the depth map.

Figure 4. An example of a photograph of the water area.

In red in Figure 4, objects to which the real distance was known in the course of the experiments are selected. To determine the distance to static objects, navigational maps of the port of Novorossiysk were used, to determine the distance to ships - AIS data.

Figure 5. Approximate location of objects and images V.

Figure 6. Image depth map V
In Figure 6 x and y - coordinates of pixels in the image; Depth - the values of the actual distances from the observation point to the designated objects. The surface was constructed using the "Trilinear interpolation" method, as follows:

\[
\begin{align*}
x_d &= \frac{x-x_0}{x_1-x_0}; \quad (4.1.1) \\
y_d &= \frac{y-y_0}{y_1-y_0}; \quad (4.1.2) \\
z_d &= \frac{z-z_0}{z_1-z_0}; \quad (4.1.3)
\end{align*}
\]

Where \(x_d, y_d, z_d\) is the distance between any \(x, y, z\) and the lesser coordinate of the known point; \(x_0\) and \(x_1\) - known values on the axis are less and more than the specified value \(x\), respectively, similarly for \(y_0, y_1, z_0\) and \(z_1\).

Initially, we interpolate along the x-axis (similar to the alignment of the opposite faces of the cube, given by \(C_{0ijk}\) and \(C_{1ijk}\):

\[
\begin{align*}
c_{00} &= c_{000}(1-x_d) + c_{100}x_d; \quad (4.1.4) \\
c_{01} &= c_{001}(1-x_d) + c_{101}x_d; \quad (4.1.5) \\
c_{10} &= c_{010}(1-x_d) + c_{110}x_d; \quad (4.1.6) \\
c_{11} &= c_{011}(1-x_d) + c_{111}x_d; \quad (4.1.7)
\end{align*}
\]

where \(c_{000}\) is the value of the function at the point \((x_0, y_0, z_0)\). Then we interpolate along the y-axis (alignment of the faces specified as \(C_{0ik}\) and \(C_{1ik}\):

\[
\begin{align*}
c_0 &= c_{00}(1-y_d) + c_{10}y_d; \quad (4.1.8) \\
c_1 &= c_{01}(1-y_d) + c_{11}y_d; \quad (4.1.9)
\end{align*}
\]

Then we interpolate along the z-axis:

\[
c = c_{0}(1-z_d) + c_{1}z_d; \quad (4.1.10)
\]

where \(c\) will be the predicted value for the given point \((x, y)\).

Filtration of accumulated data in the distance storage segment assumes the existence of a display

\[
\text{Location}(x, y, A) \rightarrow D \quad (4.1.11)
\]

where \(A\) is the vector of the current interpolation coefficients, \(D\) is the resulting distance.

Further, these coefficients are updated using the well-known backpropagation method.

4.2 Results of the experiment on the operation of the segment for determining the distance to the object.

The predicted accuracy of the distance to the real object differed from the true value by no more than 1 kbt.

Thus, using an optical segment installed on a base statically oriented relative to the water area, and a segment for accumulating and storing data, it is possible, through systematic observations, to compile an accurate map of the "depths" of images taken during the daytime with constant parameters of the direction and focal length of the cameras. The resulting "depth map" can then be used for night images, in which it is often impossible to determine landmarks.

5. Conclusion

The proposed optical-measuring system for processing photographic images, equipped in the vicinity of the water area, subject to the correct and interconnected functioning of all its elements, can significantly increase the navigational safety of the water space. For example, that includes a means of additional control of the movement of vessels within the anchor parking. Isolation of the ship lights of the image area and the subsequent processing of their parameters makes it possible to determine with sufficient accuracy the approximate angle, and subsequently the course of the ship. The image depth map compiled from the results of daytime observations will allow determining the distance to the
vessels in the photograph. These calculated parameters will help in solving the problem of identifying ships in the coastal area at night.

References
[1] Lopatina V and Senchenko V 2020 Methodic of measuring of geometric displacements of a moving object Marine intellectual technologies.50(4) 63
[2] Ferreira J. Branquinho J. Ferreira P. Piedade F 2017 Computer Vision Algorithms Fishing Vessel Monitoring-Identification of Vessel Plate Number 9-17
[3] Studenikin D, Bovan S, Khekert E, and Modina M 2019 The use of neural networks for the organization of visualobservation Marine intellectual technologies 46(3) 92
[4] Fu W, Yan F, Chen K, Ren Z2015 Scene distance measurement method based on light field imaging Applied optics54 6237-6243
[5] Ningthoujam J, Nongmeikapam K2019 Stereo System based Distance Calculation of an Object in Image
[6] Saxena A, Chung SH, Ng A 2008 3-D Depth Reconstruction from a Single Still Image International Journal of Computer Vision7653-69
[7] Fernandes J C, Campos N J 2006 Angle Invariance for Distance Measurements Using a Single Camera, 676 - 680.
[8] Zaslonova V, Golovina A, Popov A N 2020 Creating a ship without a crew within the technological paradigm of the Russian Federation Lecture notes in Networks and systems115 468-474
[9] Popov A N, Kondratiev A I, Smirnov I O 2018 Algorithm for rapid prediction of the degree of collision hazard with ships and surface objects in the electronic navigation zone AGA 2018-19th Annual General Assembly (AGA) of the International Association of Maritime Universities (IAMU), pp. 442-449.
[10] Mironenko A 2013 Methodology of the program movement of the vessel based on the Sigma function IFAC Proceedings Volumes46 (33 Part 1)138-143