Vehicle Positioning and Tracking Algorithm in the Visual Positioning System in Ship Cabin

Ke-Yang SUN\textsuperscript{1}, Chuan-Xu YAN\textsuperscript{2}, Li-Hua GE\textsuperscript{2}, Hong QIAN\textsuperscript{2}, Hui ZHAO\textsuperscript{1} and Wei TAO\textsuperscript{1,a}

\textsuperscript{1}Department of Instrument Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China
\textsuperscript{2}Chinese Shipping and Marine Engineering Design and Research Institute, Shanghai 201114, China

Abstract. A large cargo ship is widely used for vehicle transporting, and it is necessary to position vehicles in the cabin. For the characteristic of large space, low height of ship cabin and large vehicle, building a multi-camera positioning system for vehicles is the optimal one. In this paper, a high accuracy vehicle positioning algorithm is proposed to locate multiple moving vehicles in the positioning system. Firstly, the conversion relationship between the pixel coordinate and the world coordinate is calculated. Zhang’s calibration method is adopted to correct the lens distortion, the images are stitched to get the complete cabin image and the height of vehicle body is calibrated. Secondly, a method for vehicle positioning and tracking is proposed. Background subtraction method is adopted to detect the preliminary vehicle position, and Spatio-Temporal Context (STC) algorithm is adopted to track the mark on the vehicle body. The experimental results show that in 2.56m*0.62m space of the simulated ship cabin, for three vehicle models with size of 0.45m*0.14m*0.15m, the horizontal positioning accuracy can reach 2 cm each, less than 1%. With the proposed algorithm, multiple moving vehicles positioning with high accuracy and high precision in ship cabin can be reached.

1 Introduction

A large cargo ship is widely used for vehicle transporting. In the cabin of a cargo ship, multiple vehicles are usually moving independently to the specified location. However, many accidents happen during the period. In order to ensure the safety of vehicles in ship cabin, the real time position of vehicles should be monitored and calculated.

The ship cabin has the characteristics of long length, poor width and low height. Besides, the size of vehicle is large and the ratio of vehicle’s top height to cabin’s height is almost 1:2. It is difficult to position a vehicle in this application scenario.

Two methods are commonly adopted in vehicle positioning, which are radio frequency identification (RFID) and GPS technology.

RFID is widely used in the traffic positioning system at the crossroads. This technology can identify high-speed moving items, have fast multiple label reading or writing functions, and have strong intelligence. However, there are also shortcomings in RFID, which cannot get the precise location of each target, the cost is high.

Shladover and Tan (2007) \cite{1} proposed a system to position vehicles and avoid collision using GPS technology and wireless communication technologies. The system can reach a positioning accuracy of approximately 50–100 cm in order to support consistent and reliable warnings. However, the accuracy values of the system rely on the data of GPS, which demand good sky exposure. The cabin is a closed structure, and GPS may not provide data with enough accuracy. Besides, the positioning data provided by a GPS positioning system is the location of the receiver, which cannot provide the contour of the vehicle.

For this reason, building a multi-camera visual positioning system for vehicle positioning is the optimal one. To solve the problem of long length and poor width of the ship cabin, the cabin is divided into several partitions in the horizontal direction in the system. Each partition is vertically equipped with one camera, whose view field can cover the partition. Each camera captures image continuously to obtain a image sequence.

In this paper, a vehicle positioning algorithm is proposed to locate multiple moving vehicles in the visual positioning system in ship cabin. The algorithm can deal with the ratio of vehicle’s top height to cabin’s height, and reach high accuracy and high precision.

2 System Layouts

The visual positioning system in ship cabin is consisted of three parts. Part one is the vehicle positioning module, which is the key module of the system. Part two is the license plate recognition module. Part three is the displaying module. Modules of the system are shown in Figure 1.
2.1 License Plate Recognition Module and Displaying Module

The license plate recognition module continuously gets images from the entrance camera and exit camera. If the vehicle is at the entrance or exit of the ship cabin, the module will get the possible character areas and recognize the possible digital combinations in the areas. If the digit combination recognized is consistent with the pre-stored license plate number in the database, the digit combination will be recorded. After that, if the same digit combination is recorded successively twice, we consider that the digit combination is a vehicle license plate number. This license plate number will be sent to the vehicle positioning module, and the displaying module. Besides, the particular license plate number will be registered or cancelled.

The structure of the license plate recognition module is shown in Figure 2.

![Figure 2. Structure of license plate recognition module.](image)

The displaying module is the module for displaying results of the system.

For one thing, this module gets the license plate number of each entering vehicle and then shows it immediately. For another, this module gets the positioning information from the vehicle positioning module, and combines the identified license plate information and then displays it.

2.2 Vehicle Positioning Module

The vehicle positioning module is the key module of the system. This module gets the license plate number of entering vehicles, and correspond it to the entering vehicle. Then the vehicle positioning module continuously gets the position of the vehicle in cabin ship and sends it to the displaying module.

As the ship cabin has the characteristics of long length and poor width, it is necessary to partition the space of the cabin. The length and width ratio of the cabin is about 4:1. In order to ensure the location accuracy of the partition, 4 cameras are used, and each camera handles 1/4 of the field of cabin.

Considering the effects of camera and lens distortion performance, each camera is equipped with a wide angle lens. In this way, we can ensure that the total field of view of all the cameras covers the cabin space, thus ensuring the full observation of the vehicles in the cabin. The field of view of the adjacent cameras can be partially overlapped, and the field of view can be stitched to improve the stitching accuracy.

For one thing, because the ratio of vehicle’s top height to cabin’s height is about 1:2, the proportion of vehicle height to the height of the cabin is large. If the height of a camera is low, it will cause serious occlusion. Therefore, the camera is arranged along the straight line to the top of the cabin. For another, because of the large size of the vehicle, in order to ensure the positioning accuracy, all the cameras in the positioning module are the same. The cameras are placed vertically down, hanging at the same height. This can avoid the occlusion caused by the complex environment in the cabin, at the same time, to a certain extent, reduce the occlusion between vehicles, and improve the accuracy of the positioning.

The structure of the vehicle positioning module is shown in Figure 3.

![Figure 3. Structure of vehicle positioning module.](image)

3 Vehicle Positioning and Tracking Algorithm in Vehicle Positioning Module

In this paper, a high accuracy vehicle positioning algorithm is proposed to locate multiple moving vehicles in the visual positioning system. Firstly, the conversion relationship between the pixel coordinate and the world coordinate is calculated. Secondly, a method for vehicle
positioning in pixel coordinates is proposed. After that, the position of a vehicle can be calculated.

3.1 Calculating the Conversion Relationship between the Pixel Coordinate and the World Coordinate

To calculate the conversion relationship between the pixel coordinate and the world coordinate, each camera lens in the vehicle positioning module needs to be corrected. Then the images captured by the cameras are stitched, and the height of the vehicle body is calibrated.

The flow of this part in the algorithm is shown in Figure 4.

Figure 4. Flow of calculating the conversion relationship in the positioning algorithm.

3.1.1 Distortion Correction of a Single Camera

As each camera is equipped with a wide angle lens, there is a serious distortion in the lens. In order to correct the lens distortion, Zhang’s calibration method is adopted.

Zhang’s calibration method is a camera calibration method proposed by Professor Zhang Zhengyou in 1998[2]. It is based on a moving plane template. This method is a method between the traditional calibration and self-calibration. The advantages of Zhang’s calibration method overcome the shortcomings of the two methods and have the advantages of the two methods.

Zhang’s calibration method can be used to obtain the camera intrinsic parameter matrix and external parameter matrix. The image captured by the camera can be corrected by the internal and external parameters to get the image of relatively small distortion. The calibrated precision can reach 1 pixel.

3.1.2 Image Stitching and Height Calibration

After correction of each camera distortion, we can get images with lower distortion captured by the cameras. Then, we stitch the images captured by each camera to get a complete cabin image. After that, we need to get the conversion relationship between the pixel coordinate in the complete image and the world coordinate of the cabin.

By calibrating a particular plane in the cabin, this relationship can be obtained effectively. As the ratio of vehicle’s top height to cabin’s height is about 1:2, if the ground is the datum plane, there will be great error. Because the surface of the vehicle body is roughly parallel to the ground, and the height of the vehicle body of different vehicles is roughly equal, it is a reasonable choice to take the horizontal plane of the vehicle body as a datum plane.

We suppose the body height is h, and we assume that the camera is installed vertically downward, and the lens center and CCD center in the same plumb line. The world coordinates of each installed camera can be obtained by measurement, in turn (X1, Y1), (X2, Y2), (X3, Y3), (X4, Y4). In this way, this is the same as the world coordinates of the image center of each camera.

On the horizontal plane of the height of h, we place the fixed length and width rectangular markers in two lines parallel to the direction of the ship, in the designated position of the field of view of each camera. We measure the length, the width, and the world coordinates of the four corners of the rectangular marker. Then we get the image coordinates of four angles from the image.

The view of the overlook is shown in Figure 5[3].

Figure 5. Schematic diagram of the overlook of the cabin.

After that, we can calculate the conversion relationship between the pixel coordinate and the world coordinate of the markers, so as to calculate the conversion relationship of any point on the image. Considering the small error of the camera installation height on the vertical, the stitched image of the complete cabin may not be completely matched. Therefore, we must first calculate the mapping of coordinates to each camera separately, and then combine them.

In the case of camera 2, we assume that the actual length of the measured rectangle is M, and the width is N. At the same time, we assume that the image coordinate of the center of image 2 (captured by camera 2) is (l, w), the length of the rectangle marker on the image is m₀, and the width is n₀. So, as the marker is parallel to the cabin, in the X direction, the ratio of the world coordinate to the image coordinate is m₀/m₀ while in the Y direction it is n₀/n₀. The conversion relationship between the pixel coordinate in image 2 and the world coordinate is shown in Figure 6.
Then, we take the image 2 corner A as (0, 0) point in the pixel coordinates, and the world coordinate of A is:

\[ X_A = X_2 - l * \frac{M}{m_0} \]  
(1)

\[ Y_A = Y_2 - w * \frac{N}{n_0} \]  
(2)

We suppose that any point P in the image 2 is \((x_p, y_p)\) in the pixel coordinate, and then its world coordinate is:

\[ X_p = X_2 - l * \frac{M}{m_0} + x_p * \frac{M}{m_b} \]  
(3)

\[ Y_p = Y_2 - w * \frac{N}{n_0} + y_p * \frac{N}{n_b} \]  
(4)

In this way, we can calculate the conversion relationship between the pixels coordinates of the images captured by each camera in 1, 2, 3, 4, and the world coordinates. Then we can easily calculate the conversion relationship between the pixels coordinates of the vehicle height plane and the world coordinate XY in the complete cabin image after stitching, so as to complete the height calibration.

### 3.2 Positioning Method

#### 3.2.1 Flow Chart of the Positioning Method

Background subtraction method is used for judging when the vehicle enters, as well as the subsequent rough location, and the situations STC algorithm cannot identify the vehicle. Besides, the STC algorithm is used for accurate positioning of vehicles in the cabin.

The flow chart of the Positioning method is shown in Figure 7.

The first step is to obtain the registered license number of the vehicle from the license plate recognition module. In the second step, the target vehicle information in cabin is continuously obtained through the background subtraction method. The third step is to determine whether the vehicle is detected, and if no vehicle is detected, the second step is returned to the background subtraction method detection.

In the fourth step, if the vehicle is detected, then the method determines whether the STC positioning the mark is successful, and if the positioning is successful, the position of the mark is obtained. Fifth, if the location fails, the second step is returned to the background subtraction method detection.

#### 3.2.2 Background Subtraction Method

Background subtraction method is a general method for motion segmentation for static scene [4, 5]. It will get the gray image of the target area through difference operation between the current frame and background image. The moving region is extracted by thresholding of gray scale, and to avoid environmental illumination changes, background image updates according to the current frame.

Gauss background modelling is a background that based on statistical information of pixel samples [6]. It is based on the probability and statistics information of each pixel in the time domain (mean or variance) to build the color distribution model of each pixel, and then achieve the purpose of background modeling. After the
background modelling is completed, the target pixels are judged by statistical difference to achieve the detection of the foreground target, and the background model is updated with target pixels to achieve the modelling and fitting of the dynamic background.

A mixed Gauss model is established for each pixel in a video sequence frame. In this model, the greater weight represents the background and the smaller weight represents the foreground. If the new pixel can match the Gauss model corresponding to the background, the new pixel is considered as the background. If the matching is a smaller Gauss model, or a Gauss model that doesn’t match, then the pixel is regarded as the foreground.

\[
P(x_j) = \sum_{i=1}^{K} \omega_{j,i}^{(i)} \eta(x_j;x_i^{(i)};\mu_{j,i}^{(i)},\sigma_{j,i}^{(i)})\]

(5)

For a pixel in the new video frame, it matches the mixed Gauss model. If there is a successful Gauss model, the weight is changed according to the following formula [7].

\[
\mu_{i,t} = (1 - \rho) \cdot \mu_{i,t-1} + \rho \cdot X_t
\]

(6)

\[
\sum_{i,j,t} = (1 - \rho) \cdot \sum_{i,t-1} + \rho \cdot \text{diag}([X_t - \mu_{i,t}]^T (X_t - \mu_{i,t})]
\]

(7)

\[
\rho = \frac{\alpha}{\omega_{j,i}^{(i)}}
\]

(8)

Because the color difference between the cabin and the vehicle is great, the background subtraction method using the Gauss mixed model can be used to extract the vehicle target with high efficiency.

However, since the size of a vehicle is large, the positioning accuracy of the background subtraction method is not high. This method is suitable for the preliminary positioning of the vehicle and the supplementary positioning when the vehicle target is lost.

### 3.2.3 STC Tracking

Accurate positioning needs clear target positioning information. However, because of the large size of the vehicle, and different vehicles have no unified position identification information, it is an effective method to place marks that can be accurately tracked on the vehicle body.

The spatio-temporal context algorithm is based on the Bayes algorithm framework. It establishes spatial-temporal correspondence between the target and its surrounding area. STC uses the low order characteristics of the image to do statistical modelling. Through the calculation of the map, we find the maximum likelihood probability region, that is, the location of the next tracking result.

Supposing the location of the target is known in the current frame, we set it as \(x^*\), the image features of the current frame can be represented as follows [8]:

\[
X^C = \{c(z) = (I(z), z) \mid z \in \Omega_c(x^*)\}
\]

(9)

The model of the spatial context by learning is:

\[
h^S(x) = F^{-1}\left(\frac{F(be \alpha)}{F(I(x)\alpha(x - x^*))}\right)
\]

(10)

Then local predestination area of \(t+1\) frame based on \(t\) frame target location can be found:

\[
x_{t+1}^* = \arg \max_{x \in \Pi(x^*)} e_{t+1}(x)
\]

(11)

By calculating the position of the maximum likelihood probability in the confidence graph in the \(t+1\) frame, the current frame is tracked.

High accuracy positioning and tracking can be achieved by using STC algorithm. Even if there is a problem of mark missing, we can use the location information of the vehicle's last frame to create a STC rectangle box based on the vehicle center obtained by background subtraction method, and use this rectangle box as the initial location to continue tracking as a transition. When the STC algorithm recaptures the mark, it can get more accurate position.

### 4 Experimental Results and Analysis

#### 4.1 Experimental Environment

In the experimental environment, the simulated ship cabin is 2.56m*0.62m, while a vehicle model is the size of 0.45m*0.14m*0.15m whose body height is 10cm. The rectangular marker for height calibration is 42cm*30cm, and the mark for STC tracking is the rectangle of 2cm*1cm in red. Four cameras are hanging at the height of 0.28m, vertically down. The distance between two adjacent cameras is 64cm and the pixel of each camera is 1292*964.
4.2 Results and analysis

According to the experimental results, three vehicles can be positioned at the same time. To obtain the positioning accuracy, a vehicle was placed in 50 random locations in the simulated ship cabin, and we measured the world coordinates of the location. 47 vehicle positions were successfully positioned by the algorithm. We calculated the positioning results given by the algorithm and compared the results with the measurement results.

The results are shown in figure 9.

![Figure 9. Experimental results.](image)

The average value of the measurement results are x=127.50cm and y=31.24cm, while the average value of the positioning results given by the algorithm is X=127.06cm and y=31.04cm. The results fit well. Meanwhile, the variance of the absolute value of the measurement error is 1.6cm². The average error of the measurement results is 1.95cm in the 2.56m*0.62m area, which means that the horizontal positioning accuracy is less than 1%.

Therefore, the Vehicle Positioning and Tracking algorithm is of high accuracy and high precision, and it is with quite satisfactory results.

Conclusion

The ship cabin has the characteristics of long length, poor width and low height. Besides, the size of vehicle is large and the ratio of vehicle’s top height to cabin’s height is about 1:2. It is difficult to position a vehicle in this application scenario through the existing methods.

To solve the problem, a vehicle positioning algorithm is proposed to locate multiple moving vehicles in the visual positioning system in ship cabin in this paper.

Firstly, through the camera distortion correction, image mosaic and height calibration, we get the conversion relationship between pixel coordinates and the world coordinates of the plane of the vehicle body height in the whole ship cabin.

Secondly, the background subtraction method and the STC algorithm are combined to accurately position the vehicles in the cabinet, and the high robustness of the vehicle positioning in the cabin is realized.

Experimental results show that the algorithm proposed in this paper is feasible, and it has high positioning accuracy and precision.

References