International Core Journal of Engineering Volume 8 Issue 5, 2022
ISSN: 2414-1895 DOI: 10.6919/ICJE.202205_8(5).0001

Spatial Positioning of Overhead Crane Load based on Monocular Vision under Changing Illumination

Xianxian Xu, Weimin Xu, Zhiteng You, Qiang Hu, Xinlei Zhu

Shanghai Maritime University, Key Laboratory of Marine Technology and Control Engineering Ministry of Communications, Shanghai 201306, China

Abstract

A visual hierarchical detection algorithm is proposed in order to solve the problem of tracking drift or failure easily and lower accuracy in the load detection of overhead cranes based on visual detection with illumination change. Firstly, the load position information of the overhead cranes can be obtained by detecting the spherical marker attached on the load, which is insensitive to rotation and tilt. Then, the insensitivity of texture features to light changes is used and an improved mean shift algorithm and circle detection based on sampling of different regions are applied to track continuously and detect spherical markers in real time. Meanwhile, an image processing module is added to eliminate the influence of illumination changes, which can improve the robustness of the algorithm with the changing of illumination. Finally, the real-time swing angle are calculated based on the spatial geometry method. The proposed algorithm is compared and analyzed with the traditional contact method and other visual inspection algorithms. The experimental results show that this method can overcome the shortcomings of traditional encoders of low detection accuracy and dead zone effectively, and meanwhile, the load swing angle and the detection accuracy of the rope length under the change of illumination has been greatly improved and meets the real-time requirements, providing more accurate data for the anti-sway control system of the bridge crane.

Keywords

Overhead Crane; Illumination Change; Real-time Tracking; Image Processing; Load Detection.

1. Introduction

At present, the overhead cranes are widely applied in ports, factories and other occasions as an important means of cargo transportation, and the automation technology of overhead crane has gradually become a hot issue[1]. During the operation of the overhead crane, the overhead crane load will swing due to the movement of the trolley and the influence of external factors such as wind, and the load swing will bring low efficiency and safety impact to the overhead crane transportation and the staff, so it is of great significance to design anti sway control system on the overhead cranes [2].

The key of automatic anti-sway control is to obtain the position information of the bridge crane load in real time [3]. At present, the acquisition of load position information of bridge crane mainly includes contact measurement and non-contact measurement. Contact measurement mainly relies on contact sensors, such as swing angle detectors, photoelectric code discs, etc., but the information detected is single and easy to be affected by electrical noise, which may affect the measurement efficiency directly [4]. The non-contact measurement method based on machine vision principle, which can attain variety of information , be independent of the overhead crane system mechanism and avoid the mechanical conversion in the measurement, has become a hot research direction of overhead crane automation for reducing the measurement error and other advantages [5].

ISSN: 2414-1895 DOI: 10.6919/ICJE.202205_8(5).0001

In [6], two cameras are used to measure the three-dimensional position and swing angle of the load in the way of stereo vision. Two cameras are also applied to identify the marked points on the crane, and the swing angle and rope length of the load were detected in real time combined with the geometric method in [7]. In [8], a smart camera and a laser transmitter are used to measure the load swing angle. In [9], the load space position information is obtained by a camera and a laser rangefinder. These multi-sensor measurement methods have problems such as complex system structure and high cost. Monocular vision measurement methods can effectively solve these problems. At the same time, monocular vision can reduce the processing time of image detection and provide conditions for real-time detection of bridge cranes.

Monocular vision measurement method will lead to detection error in image acquisition and detection stage because of external factors, such as illumination, motion image blur, complex image background and other factors. For overhead cranes working in complex environments, changes in illumination (sunlight and light changes) will cause the detection target to be in an environment with uneven illumination. In order to solve these problems, scholars have proposed many algorithms to enhance the robustness of visual detection of overhead crane load. In [10]a monocular vision measurement method by detecting four marking points fixed on the baffle above the load is proposed. The load eccentric angle, load rotation angle and obstacles were obtained by combining with the geometric relationship of the crane working space. But detecting four marking points may cost too much time to achieve the real-time detection. In [11], the load area is detected by optical flow method, and then the detection area is tracked based on direction code matching. Finally, the three-dimensional position information of crane load is obtained by using a single camera combined with geometric method. In [12], the method of background modeling combined with background subtraction is used to identify the moving load quickly, and the load swing angle measurement algorithm based on cartesian moment is used to calculate the swing angle. But the optical flow method has too much calculation to achieve the real-time detection of the overhead crane, the background modeling method may not deal with the deformation, rotation, and tilt of the image. In order to solve these problems, in [13][14] an improved method based on hough circle principle is proposed to detect the spherical mark above the load, and then calculates the spatial positioning of the load by combining with the spatial coordinate system. However, the hough circle detection has the disadvantages of large amount of calculation and high storage requirements, and cannot accurately detect the target position in case of blur image. In [15] the mean shift algorithm and Kalman filter is proposed to track the moving target continuously, then the minimum area circle method is applied to detect the circle quickly in the region of interest, and the real-time swing angle is calculated by combining the geometric method. In [16] [17] the Bhattacharyya is used to calculate the template and the color histogram of the target to track, and finally the load swing angle is detected by geometric method. However, based on a single color feature, it is easily affected by illumination changes, resulting in tracking failure and unable to achieve accurate load detection.

Since the anti-sway control system of the overhead crane has high requirements on the accuracy and real-time performance of the load position information [18], in order to reduce the image acquisition time, the use of a single-channel-based visual detection algorithm will lead to the detection of the target due to the drastic changes in illumination. Changes in information such as color and edge make it difficult to detect the target accurately [19], so the purpose of this paper is to solve the problem that vision-based load detection of bridge cranes fails or cannot be detected in complex lighting environments accurately. In this paper, aiming at the problem that the tracking drift of the visual detection algorithm in the process of load detection of overhead cranes reduces the detection accuracy due to illumination changes, and the accuracy and real-time requirements of load detection cannot be achieved, a monocular vision-based load swing angle and rope detection method is proposed. In order to reduce the detection time, this method does not process the entire image, but uses the characteristics of spherical markers that are insensitive to rotation and tilt when detected, and uses an improved mean-shift algorithm combined with linear prediction to track spherical markers. In the load detection stage, the image is preprocessed by homomorphic filtering and improved adaptive threshold to

ISSN: 2414-1895 DOI: 10.6919/ICJE.202205_8(5).0001

eliminate the influence of uneven illumination; finally, the circle of interest area is detected and the spatial positioning of the load is performed to obtain the real-time swing angle of the bridge crane load and rope length. The main contributions of this paper are summarized as follows:

1) Aiming at the problem that the traditional target matching based on single color information and the mean-shift algorithm of centroid iteration method for target tracking is susceptible to the influence of illumination changes, this paper uses the characteristics of texture features that are insensitive to illumination changes to design multi-feature adaptive fusion The algorithm is embedded into the MS (Mean Shift) tracking framework for target matching and tracking. At the same time, an adaptive template update mechanism is designed to weight the target template and the candidate template to enhance the robustness of tracking circular markers under illumination changes.

2) Aiming at the defect that the fixed threshold is not sensitive to illumination changes, and the target information cannot be completely described in the process of dynamic illumination changes, the target template and the candidate template in the MS algorithm matching process are subjected to difference operation to obtain a sensitive to illumination change. The value of dynamic and then adaptively threshold the image.

3) Aiming at the real-time problem of bridge crane load detection, a random incremental circle detection algorithm based on differential area sampling is proposed to reduce the detection time. In the sampling stage, the algorithm uses the geometric relationship to limit the sampling area, which greatly reduces the invalid sampling of the random process. At the same time, the feature of minimizing the distance is used to reduce the incremental accumulation time in the circle detection, and improve the efficiency of the circle detection effectively.

2. The Space System of the Bridge Crane

2.1 Monocular Vision System

Figure 1 shows the monocular vision system of the bridge crane. The whole system includes a camera fixed under the trolley, a white spherical marker and a movable black baffle above the load.

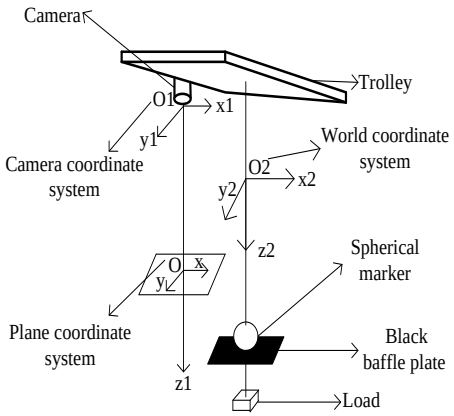


Figure 1. Model diagram of bridge crane visual inspection system

2.2 Spatial Coordinate Transformation

The spatial positioning system consists of three coordinate systems. As shown in Figure 1, camera coordinate system $o_1x_1y_1z_1$, world coordinate system $o_2x_2y_2z_2$ and plane coordinate system oxy. The center of spherical is (x_1, y_1, z_1) in the camera coordinate system, the coordinate which is calculated with space coordinate of the spherical center is (x_2, y_2, z_2) in the world coordinate system , and the pixel coordinate of the center of circle in the plane coordinate system is (x, y). According to the

ISSN: 2414-1895 DOI: 10.6919/ICJE.202205_8(5).0001

transformation principle between camera coordinate system and plane coordinate system in Fig 2, it can be concluded from the similar triangle theorem that:

$$\left(x_1, y_1, z_1\right) = \left(\frac{R}{r} * x, \frac{R}{r} * y, \frac{R}{r} * f * pix\right) \tag{1}$$

Where R is the actual radius of the spherical, r is the radius of the spherical measured by the camera, f is the focal length of the camera, pix and pix is the pixel size.

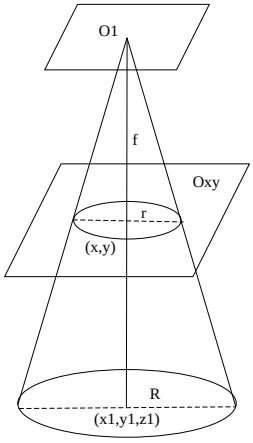


Figure 2. Schematic diagram of coordinate system transformation

According to the relationship of the position between the sling and the camera in Figure 1, it can be seen that the world coordinate system and the camera coordinate system are simple relationship of translation. Therefore, the center of the spherical in the world coordinate system from (1) is:

$$\left(x_2, y_2, z_2\right) = \left(x_1 - l, y_1 - m, z_1 + n\right) \tag{2}$$

Where l is the distance between the origin of the two coordinate systems in the direction, m is the distance between the origin of the two coordinate systems in the direction, n is the distance between the origin of the two coordinate systems in the direction.

3. Measurement Method of Bridge Crane Swing

The use of visual sensors to measure the load position information of bridge cranes is mainly affected by conditions such as too fast motion, blurred images and changes in illumination, and the MS algorithm can be used to deal with the problem of blurred images [15]. Therefore, a hierarchical positioning strategy is designed for illumination changes which the load is tracked first and then detected. Specifically, the improved MS algorithm is used to track the target in real time, and the image processing module is used to eliminate the influence of illumination changes, and the target is accurately detected by combining with the circle detection algorithm based on the difference area.

3.1 Improved MS Tracking Algorithm

The MS algorithm can deal with the problem of image blur effectively [15], but the gray level of the target changes drastically under the illumination change. Therefore, the traditional MS algorithm using a single color feature to match the similarity between the target model and the candidate model cannot accurately track the target, which is easy to cause Track drift. The improved MS tracking

ISSN: 2414-1895 DOI: 10.6919/ICJE.202205_8(5).0001

algorithm utilizes the insensitivity of texture features to illumination changes, and performs linear weighted fusion of color features and texture features. At the same time, an adaptive update strategy for target templates is designed and a linear prediction method is used to initially predict the target position. , the improved MS algorithm effectively enhances the robustness of the algorithm to track under illumination changes.

3.1.1 Traditional MS Algorithm Calculation

The traditional MS algorithm calculates the color features of the target model and the candidate model separately to form a color histogram, and then performs similarity matching based on the color histogram to achieve the purpose of target tracking. Use $q_{u-color}$ to represent the color feature of the target model, $p_{u-color}$ represent the color feature of the candidate model, and use the Bhattacharyya1 coefficient to represent the color feature similarity function $\rho_{color}\left(p, q\right)$ as:

$$\rho_{color}\left(p, q\right) = \sum_{\mu=1}^{m} \sqrt{p_{u-color}\left(f\right) q_{u-color}} \tag{3}$$

Among them, $\mu = 1, , , m$ represents the color histogram index, and f represents the center position coordinates of the candidate target. The tracking process is to maximize the above formula, and the candidate model and the target model are matched with the maximum histogram similarity (for the specific color feature calculation principle, see the literature [20-21]).

3.1.2 Multi-feature Fusion Strategy

LBP (Local Binary Pattern, Local Binary Pattern) is an operator used to describe the local texture features of an image. It has significant advantages such as rotation invariance and grayscale invariance. The extracted features are the local texture features of the image. Its basic calculation method is as follows:

$$LBP\left(x_0, y_0\right) = \sum_{p=1}^{p-1} s\left(i_p - i_0\right) * 2^p \tag{4}$$

Where:

$$s\left(x\right) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \tag{5}$$

The above formula is the original LBP operator calculation method, which defines a 3*3 window, is the pixel coordinates in the window, $\left(x_0, y_0\right)$ is the adjacent pixel value, i_0 is the center pixel value, and $p = 1, 2 \ldots 8$ represents the number of pixels. Taking the central pixel of the window as the threshold, the gray values of the adjacent 8 pixels are compared with it. If the surrounding pixel value is greater than the central pixel value, the position of the pixel is marked as 1, otherwise it is 0. In this way, 8 binary numbers are obtained, and then the binary numbers are converted into decimal numbers, that is, the LBP value of the window is obtained, and this value is used to reflect the texture information of the window, as shown in Figure 3 below $\left(\left(11100011_{10}\right) = 227\right)$.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202205_8(5).0001

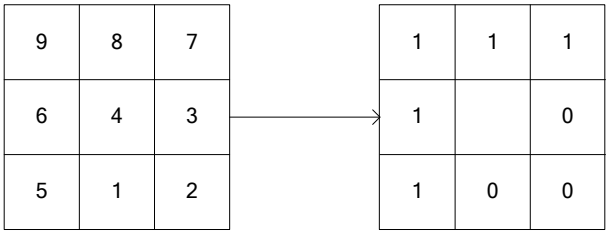


Figure 3. Schematic diagram of LBP calculation

It can be known that the texture information calculated by the basic LBP operator will change due to the change of the position state of the object, and has no rotation invariance from (4). Therefore, combined with the literature [20], the above formula is improved, and the following texture feature calculation method is obtained:

$$LBP_{P,R}^{rot} = \begin{cases} \sum_{p=1}^{p-1} s\left(g_p - g_c\right) & u\left(LBP_{P,R}\right) \le 2 \\ P+1 & U\left(LBP_{P,R}\right) > 2 \end{cases} \tag{6}$$

Among them, P represents the sampling points in the circular area with radius R. $u\left(LBP_{P,R}\right) = \left|s\left(g_{p-1} - g_c\right) - s\left(g_0 - g_c\right)\right| + \sum_{p=1}^{p-1}\left|s\left(g_p - g_c\right) - s\left(g_{p-1} - g_c\right)\right|$ represents the number of transitions of binary numbers (0 and 1) which is divided into two cases: the number of transitions is less than or equal to 2 and the number of transitions is greater than 2. As shown in Figure 4, the number of transitions of 0 and 1 in Figures 4a and 4b are 2 and 4 respectively:

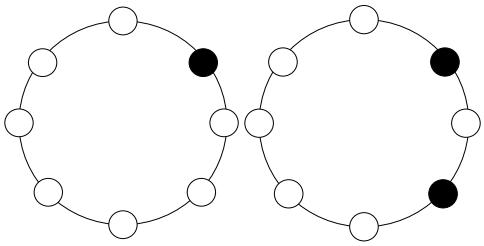


Figure 4. Part of the binary pattern of the domain point $p = 8$

The texture information calculated by the above formula has 9 modes, and the markers 2, 3, 4, 5, and 6 are defined as the main target modes, corresponding to the target areas where the pixel value changes drastically, and 0, 1, 7, and 8 are the secondary modes corresponding to "flat area". Using the main target area to model the texture features of the target while ignoring the influence of the secondary modes can enhance the robustness of the tracking. Therefore, the (6) is improved and the following texture feature modeling method is obtained:

$$LBP_{P,R}^{rot} = \begin{cases} \sum_{p=1}^{p-1} s\left(g_p - g_r - b\right) & u\left(LBP_{P,R}\right) \le 2 and \sum_{p=1}^{p-1} s\left(g_p - g_c - b\right) \in \{2,3,4,5,6\} \\ 0 & otherwise \end{cases} \tag{7}$$

ISSN: 2414-1895 DOI: 10.6919/ICJE.202205_8(5).0001

Among them, the larger the value of $|b|$, the larger the allowable value of pixel fluctuation, and the less influence on the threshold result.

Use the above formula to extract the texture feature of the target to form a texture histogram, and then use the color feature and LBP feature to describe the target, and embed it into the MS tracking framework. Use p_{u-LBP} to represent the texture feature of the target model, q_{u-LBP} to represent the texture feature of the candidate model, and Bhattacharyya2 to represent the similarity function $\rho_{LBP}(p,q)$ as (see the literature [22-24] for the specific calculation method).

$$\rho_{LBP}(p,q) = \sum_{\mu=1}^{n} \sqrt{p_{u-LBP}(f)\, q_{u-LBP}} \tag{8}$$

Where $\mu = 1,\ldots n$ denotes the texture histogram index, f denotes the center coordinate of the candidate target.

Then the improved MS tracking algorithm is an adaptive fusion of color and texture features, and its calculation method is:

$$\rho(p,q) = \alpha\rho_{color}(p,q) + (1-\alpha)\,\rho_{LBP}(p,q) \tag{9}$$

Where:

$$\alpha = \frac{\rho^2_{color}(p,q)}{\rho^2_{color}(p,q) + \rho^2_{LBP}(p,q)} \tag{10}$$

The value of α can adaptively select the update weight and enhance the tracking performance of the algorithm, at the same time reduce the amount of calculation and improve the efficiency of the algorithm.

3.1.3 Adaptive Update Strategy of Target Template

During the tracking process, due to the influence of illumination changes, the gray value of the target changes drastically, resulting in the change of the observation model of the gray target. Therefore, it is necessary to update the apparent model of the target in time to improve the robustness of the algorithm.

The update mechanism of the designed target model is weighted fusion of the current target template and the candidate region template, and the calculation formula of the target model update is:

$$\rho_i = \lambda p_t + (1-\lambda)\, q_c \tag{11}$$

$$\lambda = 1 - \rho(p,q) \tag{12}$$

where $p_t = \{p_{u-color}, p_{u-LBP}\}$, $q_c = \{q_{u-color}, q_{u-LBP}\}$, $\rho(p,q)$ are the similarity functions of feature fusion. The value of λ can be adaptively selected to update the weight in the tracking process

ISSN: 2414-1895 DOI: 10.6919/ICJE.202205_8(5).0001

to better adapt to changes in the external environment. In particular, when $\rho\left(p,q\right) \geq \varepsilon$(ε is a positive number), the update mechanism stops, which can greatly reduce the template update calculation time and improve the algorithm efficiency.

3.1.4 Linear Prediction Method

In the process of target tracking, the tracking algorithm cannot accurately track the target in real time due to factors such as the target moving too fast or occlusion, so a prediction mechanism is introduced. The tracking algorithm predicts the location information of the target by combining the prediction mechanism, which enhances the robustness of the tracking and the efficiency of the algorithm. By analyzing the characteristics of tracking target motion, a linear prediction method based on physical motion characteristics is proposed, which reduces the complexity of the prediction mechanism and improves the running efficiency of the tracking algorithm.

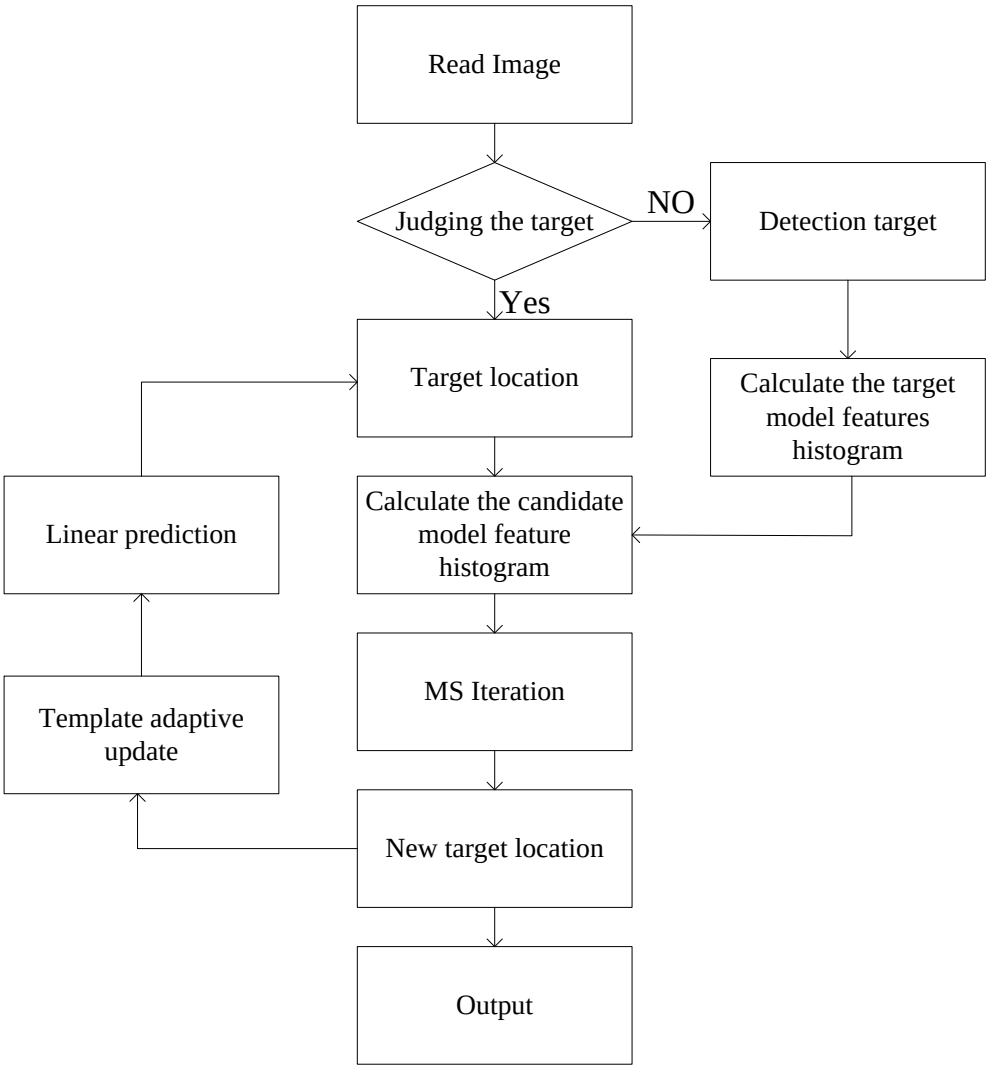


Figure 5. Flow chart of the improved MS algorithm

The linear prediction method is based on the fact that the target movement time between two adjacent frames is short, which can be approximately regarded as a uniform motion, that is $s = v * \Delta t$, s is the distance between two adjacent frames, and Δt is the time interval between two adjacent frames. Therefore, according to the position coordinates of the target at the previous moment and the current position coordinates, the offset of the target coordinates between the two moments is calculated, and then the position of the target in the next frame is predicted according to the offset. Assuming that

ISSN: 2414-1895 DOI: 10.6919/ICJE.202205_8(5).0001

the current k time, the center position coordinate of the target is (x_k, y_k) , and the center position coordinate of the target at k-1 time is (x_{k-1}, y_{k-1}) , then according to the uniform motion, the center coordinate of the next frame (x_{k+1}, y_{k+1}) can be obtained as:

$$(x_{k+1}, y_{k+1}) = (2x_k - x_{k-1}, 2y_k - y_{k-1}) \tag{13}$$

3.2 Detection Module

3.2.1 Dynamic Thresholding

For the problem that the brightness of the image is uneven due to the illumination change, and the fixed threshold processing cannot extract the target contour well, an adaptive threshold processing is proposed.

$$I_1(x, y) = \begin{cases} value & I_2(x, y) \geq M(x, y) + \Delta I_i \\ 0 & otherwise \end{cases} \tag{14}$$

$$\Delta I_i = \frac{k}{W * H} \sum_{x=0}^{W-1} \sum_{y=0}^{H-1} \left| p_t(x, y) - q_c(x, y) \right| \tag{15}$$

where I_i is the pixel value of the image after the threshold, I_2 is the pixel value of the defined n*n window area, $M(x, y)$ is the average value of the pixel values of the n*n window area, ΔI_i is the dynamic threshold and the dynamic threshold reflects the illumination changes in the video frame sequence. When the ambient illumination changes significantly, the pixel difference between the target template and the candidate template is large, and the value of dynamic threshold ΔI_i increases, which in turn increases the judgment threshold of the binarization process, thereby effectively overcoming the problem of illumination changes. When the ambient illumination changes are not obvious, the pixel difference between the target template and the candidate template is small, the value of dynamic threshold ΔI_i becomes smaller, and the threshold is very small in the process of binarization. The threshold range can be dynamically adjusted by adding ΔI_i and k is a dynamic coefficient, which $W * H$ represents the area to be detected in the image, and the value takes 255.

3.2.2 Image Preprocessing

In order to enhance the robustness of detection under illumination changes, an image processing module is added to eliminate the influence of illumination changes and image noise interference. The specific processing steps are as follows:

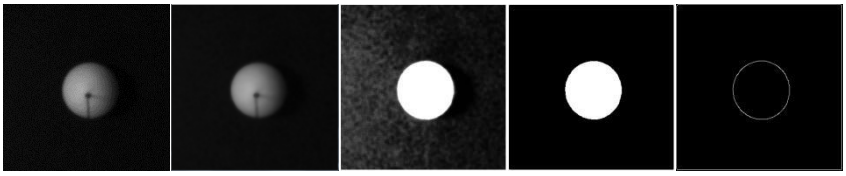
First, define a region of interest to reduce image processing time.

Then, median filtering is performed to remove image noise and preserve details.

Second, the region of interest is processed with the homomorphic filtering [25] algorithm to improve image contrast.

Finally, adaptive thresholding is performed to separate the target from the background.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202205_8(5).0001



(a) (b) (c) (d) (e)

Figure 6. Image processing experiment. (a) original image of region of interest; (b) median filter; (c) homomorphic filter; (d) adaptive threshold; (e) edge detection

3.2.3 Difference Area Circle Detection

Since the mean-shifted tracking frame is manually annotated, it cannot well represent the exact location of the target. In order to achieve accurate positioning of the load, an improved random incremental circle detection method is used to obtain the accurate position information of the target in the tracking frame.

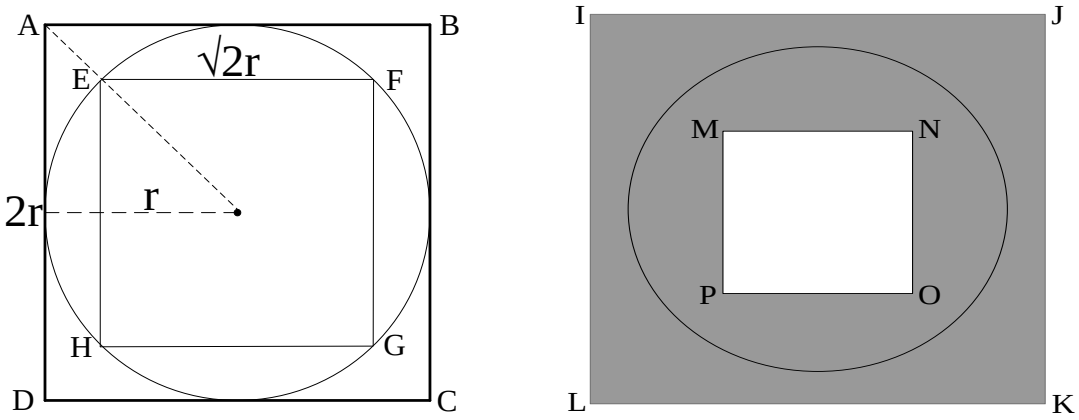


Figure 7. (a) Points on ideal circle (b) Differential sampling area

Since the sampling efficiency is low due to the random sampling of sampling points in the process of judging the candidate circle as a true circle, we improve the sampling efficiency by restricting the area of the candidate circle. Ideally, the points on the boundary of the candidate circle are shown in Fig 7a, in the area sandwiched by the circumscribed square ABCD (with a side length of $2r$, let the radius of the candidate circle be r) and the inscribed square EFGH (with a side length of $\sqrt{2}r$), we perform region limitation by setting a threshold. As shown in Fig 7b, sampling in the gray area sandwiched by the circumscribed square IJKL (with a side length of $2\left(r+\sigma\right)$, σ is the threshold) and the inscribed square MNOP (with a side length of $\sqrt{2}\left(r+\sigma\right)$). The improved circle detection algorithm is as follows:

The first step, in the sampling stage, traverse all the points in the area of interest, find the extreme point on the boundary, denoted as $N_{\min x}$, $N_{\max x}$, $N_{\min y}$, $N_{\max y}$.

The second step is to calculate the value of $N_{\max x}-N_{\min x}$ and $N_{\max y}-N_{\min y}$. If $\left|N_{\max x}-N_{\min x}\right| \geq \left|N_{\max y}-N_{\min y}\right|$, then take $N_{\max x}$ and $N_{\min x}$. Otherwise, take $N_{\max y}$ and $N_{\min y}$ and record the two points as P_1 and P_2.

The third step is to obtain the initial circle with the diameter of P_1P_2, denoted as $C_2\left(c_2, r_2\right)$(C_2 indicates the smallest enclosing circle containing the first i points).

ISSN: 2414-1895 DOI: 10.6919/ICJE.202205_8(5).0001

The fourth step is to add points in turn and set the current point to be P_i. If $\left|P_i - c_2\right| \leq r_2$, then $C_i = C_{i-1}$. Otherwise, take P_1P_i as the diameter to get a $C_i\left(c_i, r_i\right)$.

The fifth step is to find a point that is not in $P_i\left(j < i\right)$ and take P_iP_j as the diameter to get a $C_j\left(c_j, r_j\right)$.

The sixth step is to continue to find a point P_k that is not in C_i and take P_iP_k as the diameter to get a $C_k\left(c_k, r_k\right)$.

Sampling the edge points in the above-mentioned restricted area and denoted as h_k .If $\left|h_k - c_k\right| \leq r_k$, it is judged as a true circle.

In this way, the minimum enclosing circle determined by P_i , P_j , P_k and three points can be obtained, and any point in $C_k\left(c_k, r_k\right)$ satisfies:

$$\left(x - x_0\right)^2 + \left(y - y_0\right)^2 \leq \left(\frac{W}{2}\right)^2 + \left(\frac{H}{2}\right)^2 \tag{16}$$

where $\left(x_0, y_0\right)$ is the center of the region of interest, W and H are the width and height of the region of interest.

3.3 Calculation of Swing Angle and Rope Length

Suppose the position coordinate of the initial ball is $\left(x_0, y_0, z_0\right)$ and the current position coordinate is $\left(x_2\left(t\right), y_2\left(t\right), z_2\left(t\right)\right)$, then the calculation formula of the swing angle and rope length is:

$$d = \sqrt{\left(x_2\left(t\right) - x_0\right)^2 + \left(y_2\left(t\right) - y_0\right)^2} \tag{17}$$

$$\theta = \arctan\frac{d}{z_2\left(t\right)} \tag{18}$$

$$l = \sqrt{\left(x_2\left(t\right) - x_0\right)^2 + \left(y_2\left(t\right) - y_0\right)^2 + \left(z_2\left(t\right) - z_0\right)^2} \tag{19}$$

d Represents the distance that the center of the ball moves at two adjacent moments, θ is the real-time swing angle, $\left(x_2\left(t\right), y_2\left(t\right), z_2\left(t\right)\right)$ can be calculated according to formula 2, and l is the real-time rope length.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202205_8(5).0001

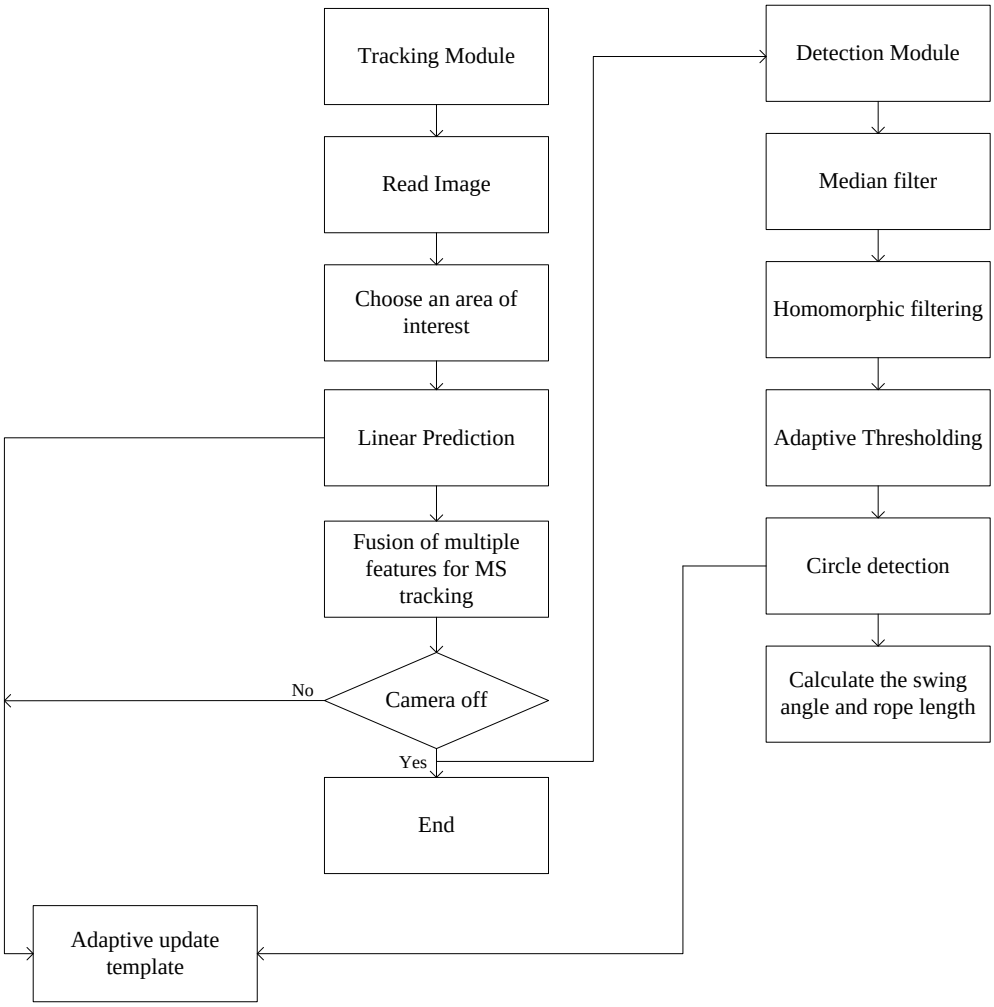


Figure 8. Program flow chart

4. Experiments and Analysis

4.1 Experimental Environment Design

In order to verify the effectiveness of the proposed algorithm, a bridge crane simulation platform is established in this paper, including a cart mechanism, a movable trolley mechanism, a motion control system and a swing angle detection system. The bridge crane simulation platform is shown in Figure 9a. The swing angle detection system includes a load part with a spherical marker with a black background plate, a black and white camera above the load, a detection part of a contact measurement device, and a vision computer. The structure diagram of the swing angle detection system is shown in Figure 9b. At the same time, in order to study the impact of illumination changes on the vision algorithm, we set an adjustable light source above the load and installed ceiling lights on both sides above the bridge suspension simulation (the ceiling light and the bridge suspension platform are illuminated at an angle of about 60 degrees). Control the brightness of the light source and control the top lights on both sides to make the detection target in an experimental environment with uniform brightness changes and oblique light irradiation. In order to verify the validity of the experimental results, the visual inspection is compared with the contact measurement and through the data acquisition card, the data measured by the contact encoder is displayed and saved in real time by the host computer software Delphi7. Meanwhile it is compared with other visual detection algorithms (reference [15] and reference [17]) and a series of experiments verify the effectiveness of the method. The system hardware parameters are shown in Table 1.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202205_8(5).0001

Table 1. Hardware Parameters

Hardware	Parameter	Value
visual computer	CPU	Intel Core i5, clocked at 2.5 GHz
	RAM	4GB
Camera	Model	Manta g125-b/c
	Focal length	5.0mm
	Frame rate	30fps
	Pixel	1296*964pixel
	Pixel size	0.00375mm/pixel
contact measuring device	Encoder	1000line
	Angle sensor	1000line

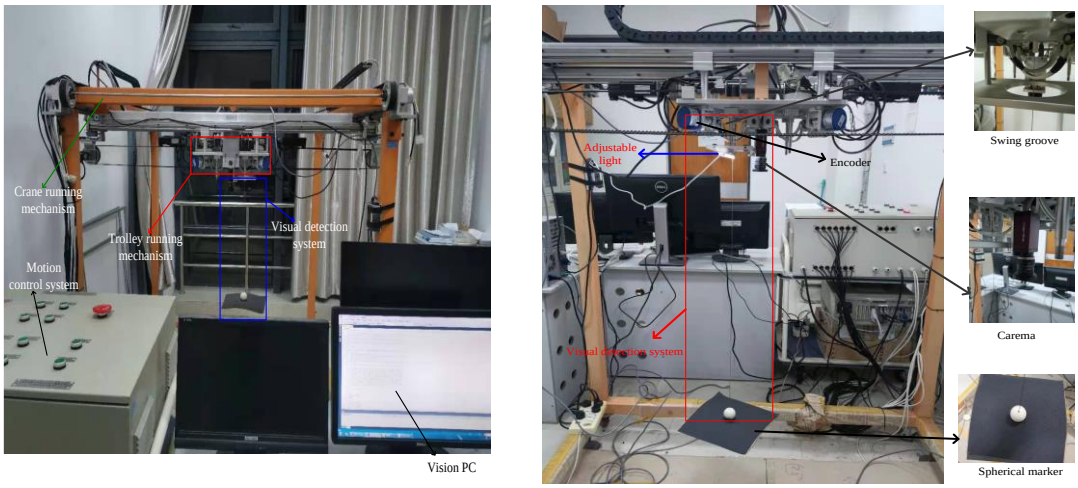


Figure 9. (a) Bridge crane simulation platfo (b) Swing angle detection structure diagram

Image processing is based on a 32-bit Windows7 operating system, using Microsoft Visual Studio 2013combined with OpenCV3.0.0 machine vision library for c++ programming processing.

4.2 Experiment 1: Algorithm Validity Verification

By adjusting the adjustable button of the light source above the load (because there is no relevant measurement tool in the laboratory, no qualitative data can be given) and controlling the switch of the ceiling light on both sides of the simulation platform, the light source can gradually change from bright to dark and in the environment of oblique light irradiation . At the same time, in order to verify the accuracy and rapidity of the algorithm for spherical marker detection under illumination changes, the algorithm proposed in this paper is compared with the algorithm proposed by Huang's (Reference [15]) et al. and Lee's (Reference [17]) et al. comparisons were made.

In the process of processing, three algorithms are used to detect the trackball images in 800 frames of interest collected continuously, and the results are statistically processed. Figure 9a shows the image processing results of the 5th frame, the 300th frame, the 623rd frame, and the 750th frame collected in an environment where the illumination brightness changes uniformly. Figure 9b shows the image processing results of the 10th frame, the 452nd frame, and the 701st frame collected in an environment irradiated by oblique light. At the same time, the accuracy and speed of the three algorithms are compared using the accuracy and average elapsed time metrics. The accuracy index is calculated as follows:

ISSN: 2414-1895 DOI: 10.6919/ICJE.202205_8(5).0001

$$precision = \frac{TP}{TP + FP} \tag{20}$$

Where TP is the number of detected frames and FP is the number of undetected frames.

4.2.1 Experiment on uniform change of illumination brightness

This paper MS(literature 15) BS(literature 17)

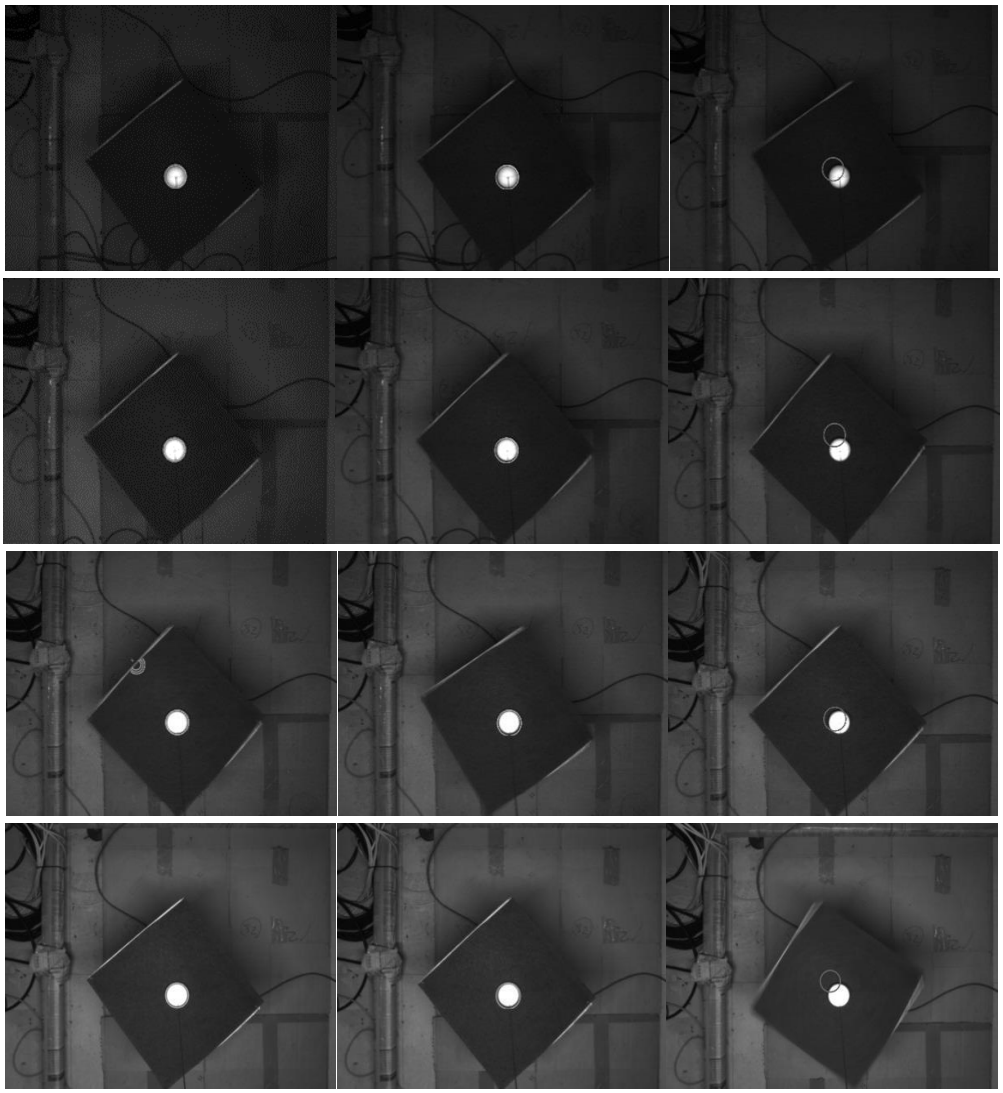
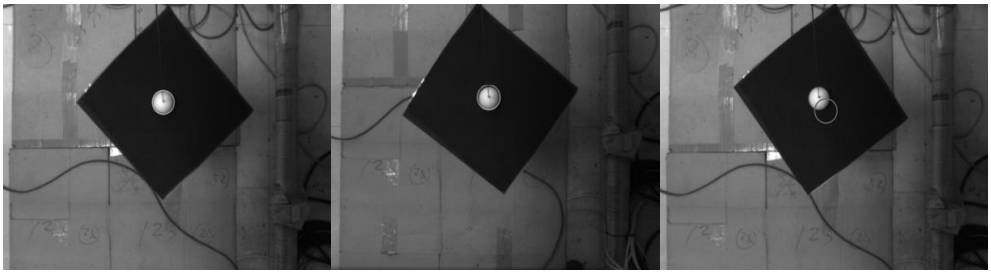


Figure 10a Image detection results of the three algorithms under uniform illumination changes

4.2.2 Oblique light irradiation experiment

This paper MS(literature 15) BS(literature 17)



ISSN: 2414-1895 DOI: 10.6919/ICJE.202205_8(5).0001

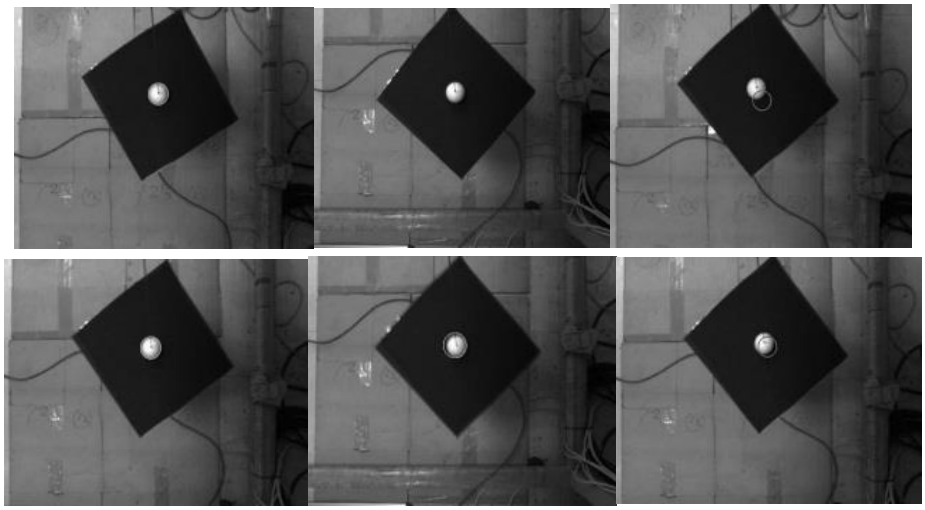


Figure 10b Image detection results of three algorithms under oblique light illumination

Table 2. Average performance of detection algorithms

Lighting environment	precision			Average time		
	This paper	BS	MS	This paper	BS	MS
uniform illumination	0.941	0.832	0.732	0.0323	0.0292	0.1293
oblique light	0.822	0.668	0.603			

From Figure 10a and Figure 10b, it can be seen that the algorithm proposed by Huang and Lee has different degrees of drift in the detection process in the environment where the illumination brightness changes uniformly and the oblique light is illuminated, especially in the oblique light illumination environment, Huang's algorithm sometimes fails to detect due to the shadows influence. It can be seen from Table 2 that in the environment with uniform brightness changes, the detection accuracy of this paper is improved by 13.1% and 28.6% compared to MS and BS respectively. In the oblique light illumination environment, the detection accuracy of this paper is improved by 24.6% and 37.9% compared to MS and BS respectively. In terms of average processing time, this paper reduces the time by 75.2% compared to BS, but is slightly higher than MS. This is because this paper takes a lot of time on the texture and template adaptive update calculation, but in general the detection accuracy and rapidity requirements. In particular, the literature [15] is a self-ablation experiment in this paper, from which the effectiveness of the improved strategy can be seen.

4.3 Experiment 2: Spatial Positioning Experiment and Data Analysis

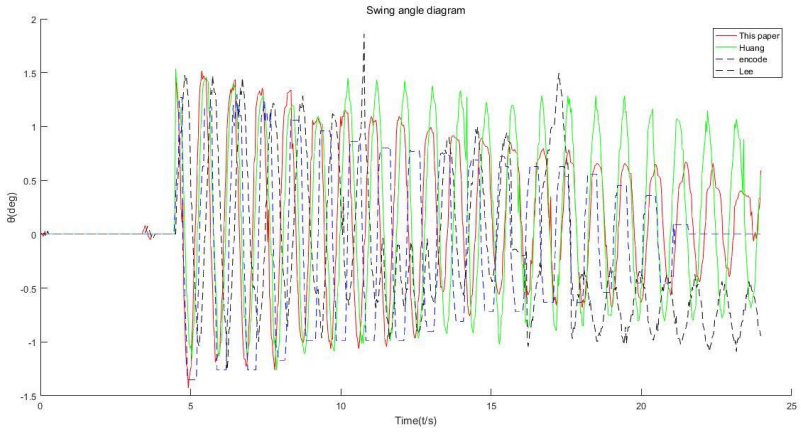


Figure 11.a Comparison of hem angle changes with uniform changes in illumination brightness

ISSN: 2414-1895 DOI: 10.6919/ICJE.202205_8(5).0001

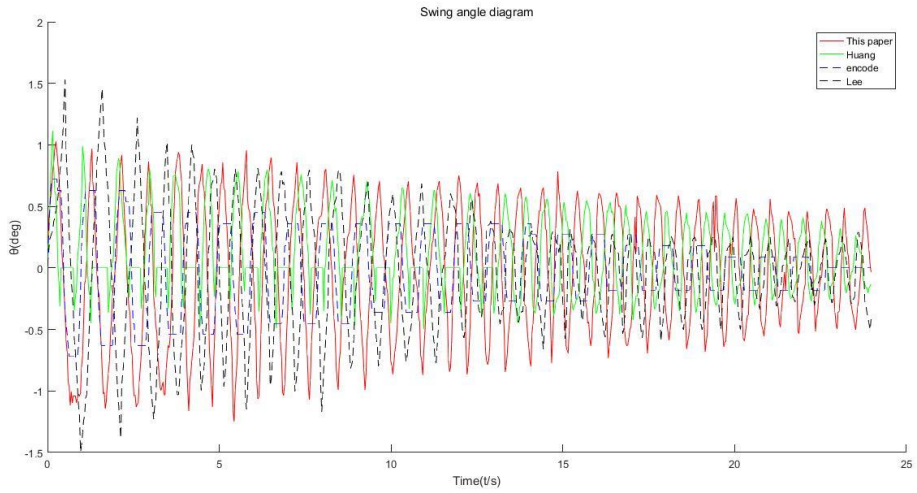


Figure 11.b Comparison of hem angle changes under oblique light irradiation

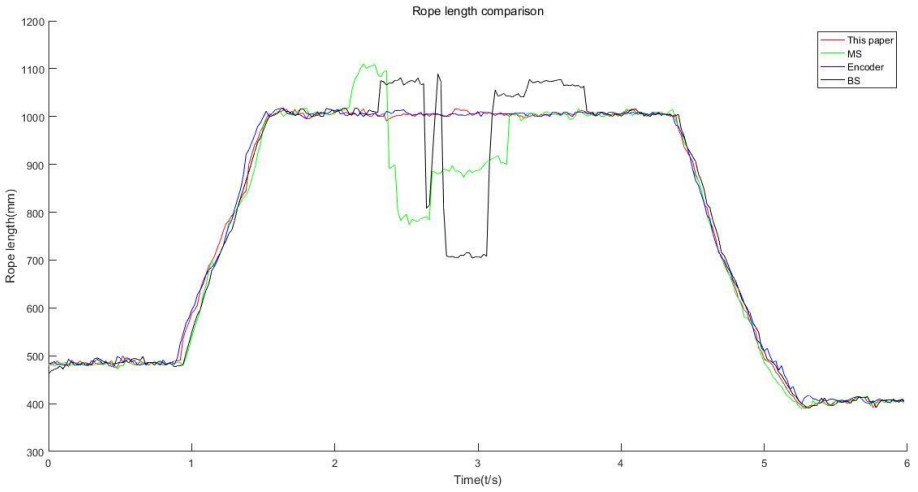


Figure 11.c Comparison of rope length changes under uniform illumination changes

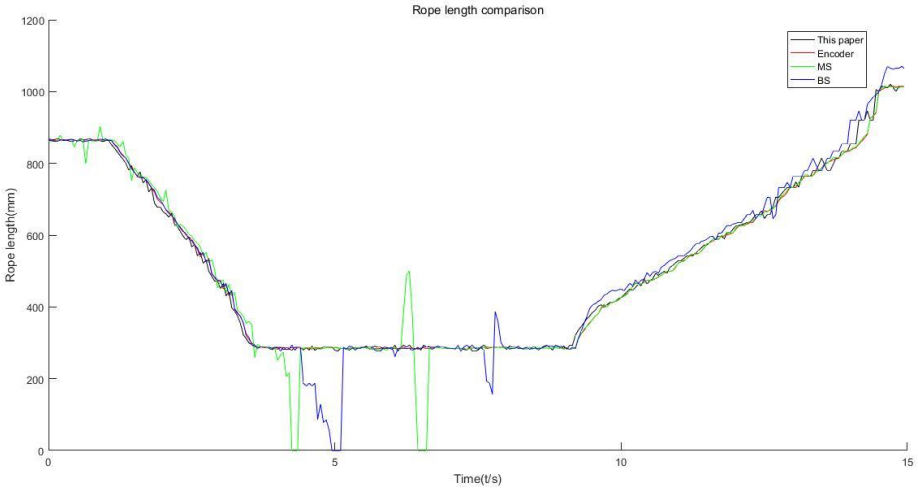


Figure 11.d Comparison of rope length changes under oblique light irradiation

In this experiment, the swing angle information obtained in experiment one is analyzed and compared with the measurement results of the contact encoder. Figures 11a, 11b, 11c and 11d show the measurement results of the four methods under uniform variation of illumination brightness and

ISSN: 2414-1895 DOI: 10.6919/ICJE.202205_8(5).0001

oblique illumination. In the comparison experiment of angle measurement, the angle measured by the touch sensor does not change after reaching a certain angle, while the angle measured by the camera is still changing, which is caused by the limitation of the system hardware. When the touch sensor measures the swing angle, the sling needs to be passed through the sensor's swing groove. In order to make the swing flexible, the width of the swing groove will be larger than the diameter of the sling, so that although the load is swinging, the contact sensor cannot calculate the value. Meanwhile the touch sensor cannot measure the angle when the angle change is less than a certain value due to the dead zone problem of the touch sensor. In the three comparison experiments of visual inspection, under the environment of oblique light illumination, the angle measured by the algorithm proposed by Huang et al. is 0 degrees, because under oblique light illumination, when the ball moves to the shadow, it cannot be detected the target. The swing angle value sometimes increases or decreases suddenly, because the algorithm drifts during the detection process and loses the detection target. In the comparison experiment of rope length measurement, the contact sensor uses the calculation of the circumference of the rotating shaft to count the rope length, which is more accurate. In the comparison of the rope length experiments of the two illumination changes, the rope length calculations of Huang et al. and Lee et al. in Figures 11c and 11d have abrupt changes. This is because the identification cannot be detected in real time due to the influence of the illumination changes during the detection process. For the position of the object circle center, the algorithm proposed by Huang et al. and Lee et al. cannot accurately calculate the rope length in real time under the illumination change, while the rope length calculated by the algorithm proposed in this paper and the calculation results of the sensor are within a certain error range. The method proposed in this paper can well detect the swing angle and rope length information of the bridge crane load, which verifies the robustness of the algorithm in the detection of light changes.

5. Conclusion

In this paper, a visual layered detection strategy is proposed to solve the problem that the tracking drift reduces the detection accuracy due to illumination changes during the load detection of overhead cranes based on visual detection. Firstly, the target is continuously tracked by the improved MS algorithm, and then the collected images are processed and detected in real time to obtain the spatial position of the ball, and finally the change of the load's swing angle is obtained through geometric calculation. Experiments have verified that this method can accurately obtain the spatial information of the load under the illumination change, effectively replace the traditional encoder to measure the load position information, is robust to the illumination change, and can achieve the real-time and accurate detection of the bridge crane load. Require. In the next step, we will consider the influence of complex background and other factors on the detection algorithm and further improve the robustness of the algorithm in the detection of actual working conditions.

Acknowledgments

Fund Project: Supported by the Shanghai Natural Science Foundation of China (13ZR1418800).

References

[1] Lu Fengjiao, Liu Haijiang, You Lei, et al. Overview of anti-sway control algorithms for bridge cranes [J]. Ship and Offshore Engineering, 2020, 036(001): P.1-7.

[2] Abdullahi, Auwalu, M, et al. Control strategies for crane systems: A comprehensive review[J]. Mechanical Systems & Signal Processing, 2017.

[3] L. Anh Tuan, A. Janchiv, G.H. Kim, S.G. Lee, Feedback linearization control of overhead cranes with varying cable length[J]Conf. Control, 2011, pp. 906–911.

[4] M.H. Fatehi, M. Eghtesad, R. Amjadifard. Modelling and control of an overhead crane system with a flexible cable and large swing angle[J] Act. Control. 33 (2014) 395–410.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202205_8(5).0001

[5] KAWAI H, KIM Y B,CHOI Y. Measurement of a container crane spreader under bad weather conditions by image restoration [J]IEEE Transactions on Instrumentation & Measurement,2012,61(1) : 35 - 42.

[6] Yoshida Y,Mason A, Mukhopadhyay S, et al.Gaze-Controlled Stereo Vision to Measure Position and Track a Moving Object: Machine Vision for Crane Control. Sensing Technology vol 8. Springer, Cham.75-93,2014.

[7] Lee L H , Huang C H , Ku S C , et al. Efficient Visual Feedback Method to Control a Three-Dimensional Overhead Crane[J]. IEEE Transactions on Industrial Electronics, 2014, 61(8):4073-4083.

[8] Hyla P. Single camera-based crane sway angle measurement method[C]// International Conference on Methods & Models in Automation & Robotics. IEEE, 2014.

[9] Nara S , Miyamoto D , Takahashi S . Position Measurement of Crane Hook by Vision and Laser[C]// Conference of the IEEE Industrial Electronics Society. IEEE, 2007.

[10] Wu Q , Wang X , Hua L , et al. The real-time vision measurement of multi-information of the bridge crane's workspace and its application[J]. Measurement, 2019, 151:107207.

[11] Takahashi S , Kaneko S . Motion tracking of crane hook based on optical flow and orientation code matching[C] IEEE International Workshop on Advanced Motion Control.2008.

[12] Xu Peng, Fang Yongchun, Chen He. An Overhead Crane Load Swing Angle Measurement Algorithm Based on Background Modeling [J]. Control Engineering, 2019, v.26;No.177(09):11-17.

[13] Luo Yuyang, Xu Weimin, Zhang Mengjie, et al. Spatial positioning method of bridge crane load using monocular vision [J]. Computer Applications, 2016(4):1156-1162.

[14] Zhang Wanpeng, Xu Weimin, Gu Xiutao, Zhang Mingming, Wang Yongshuang. Load space positioning of bridge crane based on improved circle detection algorithm [J]. Sensors and Microsystems, 2020, v.39; No.346(12):51-54+ 57.

[15] Huang J , Xu W , Zhao W , et al. An improved method for swing measurement based on monocular vision to the payload of overhead crane[J]. Transactions of the Institute of Measurement and Control, 2020: 014233122092131.

[16] Chang C Y , Wijaya Lie H. Real-Time Visual Tracking and Measurement to Control Fast Dynamics of Overhead Cranes[J]. IEEE Transactions on Industrial Electronics, 2012, 59(3):1640-1649.

[17] Lee, L. H., Huang,et al. Applying vision feedback to crane controller design. International Journal of Systems Science, 46(2), 294 -302,2015

[18] Ramli L , Mohamed Z , Abdullahi A M , et al. Control strategies for crane systems: A comprehensive review[J]. Mechanical Systems & Signal Processing, 2017, 95(Oct.):1-23.

[19] Zhang Qingpeng, Li Yongqing, Li Guangheng, Wang Yuqing. Research on 3D reconstruction of pressure sensor based on monocular vision [J]. Instrument Technology and Sensors, 2021(08):92-94+105.

[20] H. Xie, X. Fan, M. Sun, Z. Lin, Q. He and L. Sun. Programmable Generation and Motion Control of a Snakelike Magnetic Microrobot Swarm. IEEE/ASME Transactions on Mechatronics, vol. 24, no. 3, pp. 902-912, June 2019.

[21] Liu Tianjian, Qiu Lida, Zhang Ning. An Improved Target Tracking Algorithm Based on Mean Shift [J]. Computer Engineering, 2015, v.41;No.455(09):281-285.

[22] Mujahid O , Ullah Z . High Speed Partial Pattern Classification System using a CAM-based LBP Histogram on FPGA[J]. IEEE embedded systems letters, 2019, PP(99):1-1.

[23] Jie-chun Chen, Jun Wang, Li-ping Zhao, Jin He. Branch-structured detector for fast face detection using asymmetric LBP features[J].Signal, Image and Video Processing,2020,14(prepublish).

[24] Yuan Xiaopei, Chen Xiaofeng, Lian Ming. Multi-feature fusion target detection algorithm based on Haar-like and LBP [J]. Computer Science, 2021, 48(11): 219-225.

[25] Hu Xiujun, Yu Fengqin. A pointer meter reading recognition method based on single-parameter homomorphic filtering and global contrast enhancement [J]. Instrument Technology and Sensors, 2021 (05):42-46+51.