

Improved Method for Positioning Crane Grab Boom Corner Points using Hough Transform and K-means Clustering

Min Wang¹, Longkun Wan², Chengli Zhao², and Zhangyan Zhao²

¹CCCC

²Wuhan University of Technology

April 29, 2023

Abstract

To ensure that the crane can smoothly calibrate and align the lifting rod with the beam body lifting hole, it is necessary to use image processing technology to locate and detect the corner coordinates of the crane's lifting rod. Traditional corner detection methods are not suitable for this scene. This article proposes a new idea for corner positioning, which locates corner coordinates through the intersection of straight lines. Firstly, using the R and G channels of the RGB color space to construct a grayscale difference map is beneficial for Otsu's threshold segmentation; Secondly, this article proposes an optimal adaptive threshold determination method to filter the number of votes in the clustering results, eliminate interfering straight lines, and improve the clustering centroid calculation method based on the weight calculation formula of different voting proportion, replacing the original clustering centroid as the basis for line fitting; Finally, calculate the corner coordinates of the crane's grab boom based on the straight line fitting results, and compare the recognition accuracy under different lighting conditions. This method is significantly superior to traditional corner detection methods, providing a method basis for solving the algorithm accuracy and robustness problems of port cranes under multiple environmental variables.

Improved Method for Positioning Crane Grab Boom Corner Points using Hough Transform and K-means Clustering

Min Wang ¹, Longkun Wan ², Chengli Zhao ²,Zhangyan Zhao^{2*}

1 CCCC Second Harbor Engineering Company LTD,Wuhan 430040,China

2 School of Transportation and Logistics Engineering, Wuhan University of Technology, Wuhan 430063,China

Correspondence:zzy63277@163.com

Abstract: In the process of automatic grabbing of bridge segment beams, in order to ensure that the crane can smoothly calibrate and align the lifting rod with the beam body lifting hole, it is necessary to use image processing technology to locate and detect the corner coordinates of the crane's grabbing lifting rod. When applying traditional corner detection methods to this scene, there are challenges such as low detection accuracy and unsuitability. This article proposes a new idea for corner positioning, which locates corner coordinates through the intersection of straight lines. This method is divided into three steps: first, use the R and G channels of the RGB color space to construct a grayscale difference map, so that the grayscale histogram of the foreground and background presents a bimodal feature, which is conducive to Otsu's threshold segmentation. And use the open close operation to denoise the small impurities in the Canny edge detection results; Secondly, this article proposes the optimal adaptive threshold determination method to filter the number of votes in the clustering results, eliminate interfering straight lines, and then improve the clustering centroid calculation method by using weight calculation formulas based on different proportion of votes, replacing the original clustering centroid as the basis for line fitting; Finally, calculate

the corner coordinates of the crane's grab boom based on the straight line fitting results, and compare the recognition accuracy under different lighting conditions. The experimental results show that when there are many interfering edge points in the edge detection result graph, compared to other line detection algorithms, the detection error of our algorithm is smaller and has strong robustness. The calculated corner coordinate accuracy is pixel level. The algorithm in this article has the best detection performance under strong complementary light conditions. The average detection error within 0-2 pixels accounts for 97.1% and a recognition accuracy of 98.6%. The recognition success rate under different lighting conditions is higher than 92.9%. This method is significantly superior to traditional linear detection methods and meets the needs of automatic gripping of the boom. It has certain engineering application value and provides a method basis for solving the algorithm accuracy and robustness problems of port cranes under multiple environmental variables.

Keywords: Crane grabbing boom \ Hough transform \ K-means clustering \ Line fitting \ Corner detection

0 INTRODUCTION

During the lifting operation of port cranes, it is necessary to position the grab boom at the designated position, accurately control the angle and direction of the boom, and enable it to complete the task of grasping and transporting objects. If the position or angle deviation of the crane grasping the boom is too large, it can lead to low efficiency, inaccurate operation, and major safety accidents caused by operational errors. Therefore, by using image processing technology to detect the corner coordinates of the crane's grab boom, the position and direction of the crane's grab boom can be monitored in real-time, and whether the boom is at the correct angle and position can be determined, thereby ensuring efficient and accurate alignment between the crane's grab boom and the segment beam body lifting hole. At the same time, the detection of the corner coordinates of the crane's grab boom also lays the foundation for the autonomous control and intelligent application of the crane[1-2]. By detecting and analyzing the coordinates of corner points, the crane can achieve automated control and intelligent decision-making, thereby further improving the operational efficiency and safety of the crane.

Corner detection algorithms aim to detect all possible planar coordinate information of corners in an image. During the process of extracting useful information from an image through feature extraction, edge detection, and other methods, it is inevitable to lose some useful information or introduce unnecessary interference. Corner detection algorithms usually start from the global perspective and detect all possible corner coordinates in the input image information. Therefore, this series of corner detection algorithms cannot directly provide corner coordinates for certain specific positions. Domestic and foreign scholars have conducted extensive research on the direction of diagonal point detection. Harris corner detection[3], Shi Tomasi corner detection[4] based on Harris corner detection, Fast corner detection[5-6], Sift feature detection algorithm[7-9], etc. These corner detection algorithms are suitable for matching in massive feature databases. The above literature mainly focuses on detecting all corners in the image and cannot be used for corner localization at certain specified positions.

Semantic segmentation [10-14] based on deep learning can segment images at the pixel level, achieving precise and accurate segmentation of images, and distinguishing different objects, object boundaries, and backgrounds in the image. Semantic segmentation can automatically segment the crane's grab boom, but this series of semantic segmentation methods often make it difficult to accurately recognize and segment different objects when dealing with complex scenes and small targets. At the same time, semantic segmentation methods based on deep learning require processing a large amount of data and require a large amount of computational resources, which increases costs. Before applying this series of semantic segmentation methods, it is usually necessary to annotate and train a large number of high-quality images. The time cost of collecting images and annotating data is extremely high. If the quality of annotation and image acquisition is not good, it will lead to low accuracy of the trained model and even complete segmentation failure. Compared to semantic segmentation, Otsu algorithm is suitable for most images and can quickly and

accurately classify images into foreground and background categories. It does not require prior information and has good robustness to noise. Ashish [15] et al. introduced an optimal multi-level 3D Otsu image thresholding technique and proposed a 1-D-Otsu thresholding method based on the CFA cuttlefish algorithm to reduce noise and weak edge effects, optimizing the traditional Otsu algorithm for color image segmentation. Jiqing Chen [16] et al. proposed a navigation extraction method for greenhouse cucumber harvesting robots using predicted point Hough transform. A new grayscale factor was used for image segmentation, and finally the predicted point Hough transform was used to fit the navigation path. The calculation time of this method was reduced by 35.20ms compared to traditional Hough transform, but the grayscale factor in this method is prone to image oversegmentation. Ziwen Chen [17] et al. proposed a vegetable crop extraction method based on automatic Hough transform accumulation threshold. The image was segmented using a component independent of light in the Lab color space, and the feature points of crop rows were extracted using the dual threshold segmentation vertical projection method. Finally, the cluster analysis in the accumulator was clustered into the same number of classes as the number of crop rows using the k-means clustering method. This method provides a certain basis for solving the robustness and adaptability problems of algorithms under multiple environmental variables, but the accuracy of line fitting needs to be improved.

The application scenarios and issues presented in the above literature for precise calibration and alignment of the crane’s grab boom and segmental beam body lifting holes. This article provides a new approach to corner positioning, which involves calculating the intersection coordinates of the fitted straight line to locate the three corner coordinates of the two sides of the crane’s grab boom. Use these three corner coordinates to fit the plane and determine the position and direction of the crane’s grab boom in space.

Firstly, in the image preprocessing stage, a grayscale difference map is constructed through the R and G channels of the RGB color space. The resulting difference grayscale map avoids the problem of oversegmentation and undersegmentation of the target object, making the grayscale histogram of the foreground and background appear bimodal, which is conducive to Otsu’s threshold segmentation. And use the open close operation to denoise the small impurities in the Canny edge detection results. Secondly, in the edge line detection and fitting process of the crane grabbing the boom, this paper proposes an optimal adaptive threshold determination method to screen the number of votes in the clustering results, eliminate interfering straight lines, and then improve the clustering centroid calculation method by using weight calculation formulas based on different proportion of votes, replacing the original clustering centroid as the basis for line fitting. Finally, in the corner detection and plane fitting process, the coordinates of the three corner points of the crane’s grab boom are calculated, and the plane information is determined using the three corner point coordinates. The research results of this article provide a methodological basis for solving the algorithm accuracy and robustness problems of port cranes under multiple environmental variables.

1 IMAGE PREPROCESSING

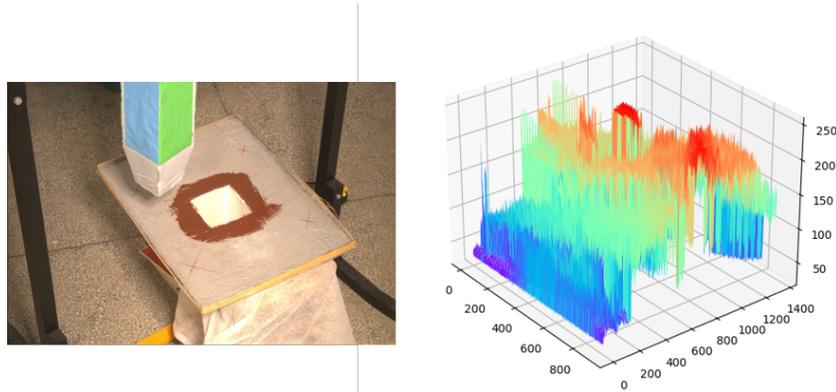
1.1 Color Space Selection

The main goal of crane grab boom extraction is to separate the crane grab boom from the background in a complex environment. The position of the industrial camera constantly changes with the movement of the crane grab boom, and its background mainly includes foreign objects such as segment beam lifting holes, segment beam planes, stones on the road surface, local buildings on the construction site, and other mechanical equipment. Due to significant differences in color between the crane’s grab boom and the surrounding environmental characteristics. Therefore, this article uses color features as the basis for image segmentation. However, considering that the on-site construction environment is outdoor and the working environment is complex, for example, it is necessary to work under lighting conditions such as rainy, cloudy, and night, which are very unstable. The resulting uneven lighting, shadows, weak lighting, and strong lighting can make image segmentation unstable. Using the HSV color model, although it can easily segment the target color, it requires high stability of lighting and cannot meet the requirements of outdoor construction. This paper

adopts RGB color space, which will make a specific color different under different lighting conditions. The RGB color space contains 256 levels of red, green, and blue, which can represent very subtle color difference, thus ensuring that the edge details of the target will not be lost during image segmentation. Secondly, after separating the three channels in the RGB color space, the values of the image matrix can be mapped to the 0-255 range, and the calculation between each channel can be simplified by adding, subtracting, multiplying, etc.

1.2 Construct differential grayscale images

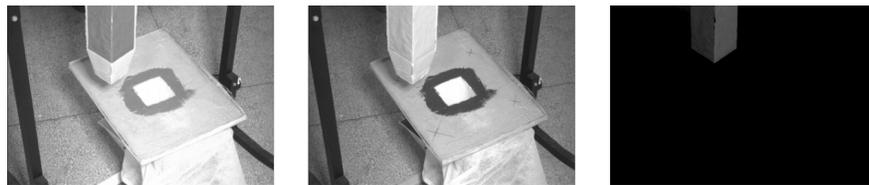
This article adopts the method of constructing differential grayscale images to prepare for image segmentation. In the original grayscale image, the grayscale image is constructed based on the different weights of each pixel's R, G, and B, which cannot reflect the differences between each component, as shown in Figure 1.



(a) (b)

Fig 1. Grab boom of crane and its three-dimensional gray distribution

Due to the blue and green colors on the two sides of the crane grabbing boom, the R, G, and B channels of the original image are separated, and the corresponding grayscale values are displayed in each channel. Using the R and G channels to construct a grayscale image, as shown in Figure 2, the grayscale values of the two sides of the boom in Figure 2a in the R channel are smaller than those in Figure 2b in the G channel. This is because in RGB color models, the color components G that are usually blue and green are greater than R. Therefore, it is reflected in the grayscale images constructed by the R and G channels that the brightness of the boom in the G channel is significantly higher than that of the boom in the R channel. From this, a differential grayscale image 2c is constructed. The three-dimensional gray distribution diagram is shown in Figure 3.



(a) (b) (c)

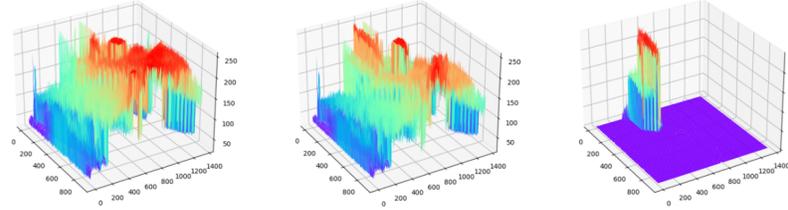


Figure 2. R channel grayscale map, G channel grayscale map, and differential grayscale map

(a) (b) (c)

Figure 3. R Channel 3D Gray Distribution Map, G Channel 3D Gray Distribution Map, Difference 3D Gray Distribution Map

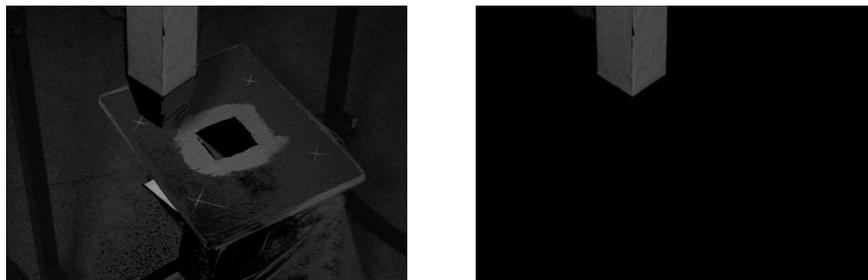
Compared with Figure 3b, Figure 3a shows that the distribution of background pixel values in the R and G channels is roughly the same, but there are small differences. For example, the hanging hole red paint in Figure 3b is significantly darker in brightness than in 3a. When the G channel is subtracted from the R channel, the result in this area is negative. To ensure that image details are not lost, the absolute value of the difference is usually taken as the result and stored in the corresponding pixel coordinates. However, based on prior information, it is known that in the RGB color model, there are multiple mixed colors in an image. The R component in the red, brown, and yellow series colors is usually greater than the G component. This does not highlight the gray level difference in the area to be segmented, which can cause the gray level peak of the interfering pixel area to be similar to the target hanger, and even the average gray level value of the interfering pixel area to be higher than the target hanger, Although retaining more edge information in the image, it also introduces a large amount of interference.

The reason why is not used in this article is that we do not want to construct the grayscale difference too clearly. When there are dark blue or green areas in the background, the constructed grayscale difference formula enlarges the background clutter, resulting in the problem of image oversegmentation.

This article adopts the following formula to construct a differential grayscale image:

1.3 Image segmentation and morphological processing

This article uses the Otsu algorithm to perform threshold segmentation on the constructed differential grayscale image, completing the separation of crane grab boom and background. The Otsu algorithm solves a binary classification problem, which satisfies the requirement of maximum inter class variance and minimum intra class variance. The Otsu algorithm is more suitable for grayscale histograms with bimodal features. Namely, construct a grayscale difference map to meet the bimodal features as much as possible. If the absolute value of the difference is taken to construct a grayscale image, as shown in Figure 4a.



(a) (b)

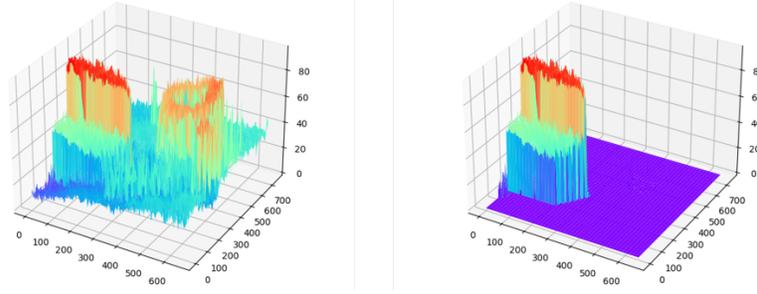
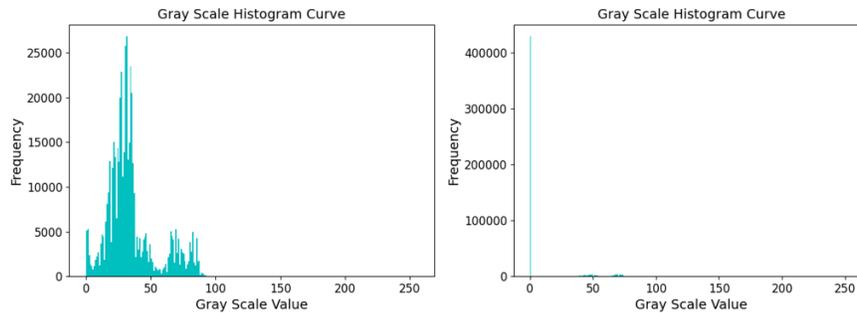


Figure 4. Absolute Value Difference Grayscale Map and Difference Grayscale Map

(a) (b)

Figure 5. Three dimensional gray distribution of absolute value difference gray map, and three dimensional gray distribution of difference gray map

It is obvious that there is more image information in 4a, but most of them are interference pixels. Its two-dimensional gray-scale histogram shows multi peak characteristics, and its three-dimensional gray-scale distribution diagram 5a shows highly consistent with the information in Figure 4a. The method proposed in this article is used to construct a differential grayscale image. Figure 4b shows that the grayscale values in the area where the crane grasps the boom are brighter than the surrounding environment, and are highlighted to achieve the purpose of constructing differences. Figure 6b 2D gray histogram shows that the gray value of more than 90% of the pixels in the whole image is 0, with obvious peaks and troughs, showing a bimodal distribution.



(a) (b)

Figure 6: Grayscale Histogram of Absolute Value Difference Grayscale Map, Grayscale Histogram of Difference Grayscale Map

Use the Otsu algorithm to segment two grayscale images. Calculate the maximum grayscale level K according to formula (2), which is the threshold of Otsu.

In the formula, when the grayscale level is set as the threshold, the probability of a pixel being classified as A is ω_A , μ_A is the grayscale value of the entire image, and σ_A^2 is the variance.

The two-dimensional grayscale histogram in Figure 7a presents a multimodal distribution, and the Otsu algorithm cannot find the optimal threshold in the multimodal distribution. Therefore, many backgrounds are segmented, resulting in poor segmentation performance. Figure 7b is a binary image for constructing a differential gray-scale image, highlighting the gray-scale characteristics of the target area, so that the grayscale distribution of the target and the background presents a double peak feature, and the segmentation

effect is good. To remove the influence of paint peeling on the wall surface of the crane's grab boom, this article adopts morphological opening operation to eliminate white impurities in the black area, and then performs morphological closing operation to eliminate black pixels in the white area, as shown in Figure 7c.



(a) (b) (c)

Figure 7. Binary image of absolute difference grayscale image, binary image of difference grayscale image, morphological denoising image

In preparation for the subsequent Hough transform, Canny edge detection was performed on Figure 7c, resulting in a clear edge detection image with low noise and accurate edges, as shown in Figure 8.

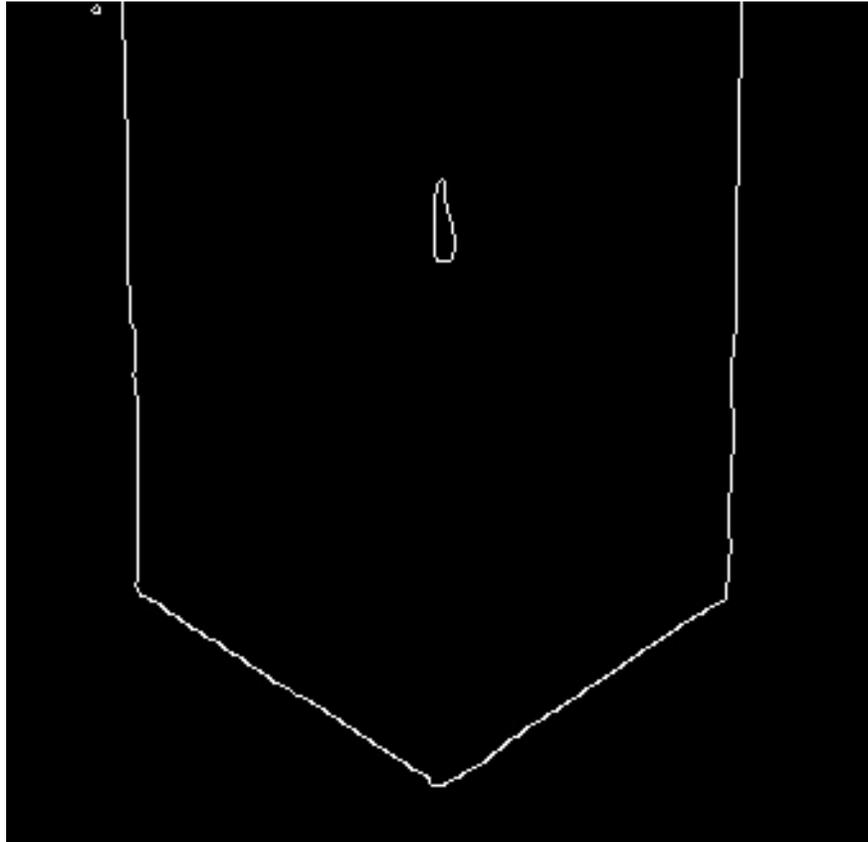


Figure 8. Edge detection diagram

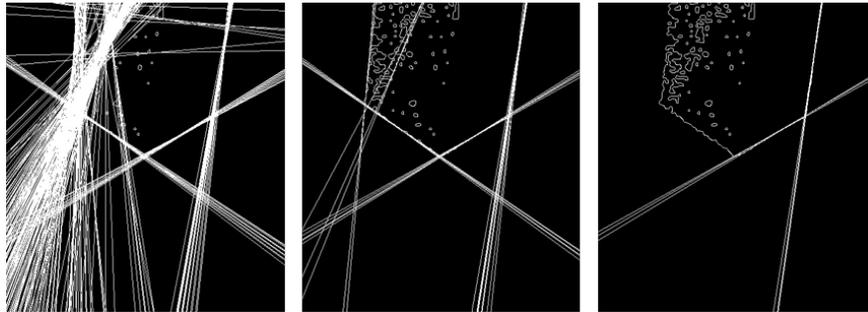
2 Edge line detection and fitting of crane grab boom

2.1 Edge Line Detection of Crane Boom Based on Hough Transform

In an image, pixels with linear features need to be detected. The commonly used methods are divided into two categories. The first type of method predicts the distribution of pixel points through linear regression, with the least squares method[18] and Ransac [19] line fitting being the most representative. The least squares method is only applicable to a group of pixels with straight line characteristics. Serious deviation of outliers will directly affect the accuracy of linear regression. When used to detect straight lines in Figure 5, only one line with serious deviation can be obtained. Although Ransac line fitting has strong anti-interference ability for outliers, it is only suitable for detecting a group of pixels with line features. The second type of method can perform line detection on any pixel in an image with line features from a global perspective. Hough transform, probabilistic Hough transform, LSD, and other methods can detect global line features, among which Hough transform is the most classic and can be used to detect any shape that can be expressed using mathematical formulas. The principle is to transform points on a specific graph into a parameter space, and obtain a maximum solution based on the vote accumulator in the parameter space. This solution corresponds to the parameter of the desired geometric shape. The transformation principle is shown in Figure 9.

Figure 9. Hough transform principle diagram

According to the principle of Hough transform, the more obvious the line features are, that is, the more pixels in an edge detection image are located on the same line, the higher the linearity of the pixel arrangement, and the more detectable they can be. But usually, due to the continuous movement of the crane's grab boom, the industrial camera also moves, resulting in constantly changing construction backgrounds on site and the impact of different lighting conditions. It is inevitable to cause some small interfering pixels in the differential grayscale image. Select edge detection images detected in harsh working environments for explanation. As shown in Figure 10. The Hough transform detects straight lines when the corresponding voting thresholds in Figure 10 are 30, 40, and 60, respectively.



(a) (b) (c)

Figure 10. Line detection results under different Hough transform thresholds

From Figure 10a, it can be seen that when the voting threshold is 30, the Hough transform detects multiple straight lines in the figure. That is, the smaller the voting threshold setting, the shorter the line segment features that can be detected, It is obvious that this will detect the interfering pixels in the edge image as straight lines. In Figure 10b, when the threshold of vote count is 40, the number of interfering pixels detected in the Hough transform detection results significantly decreases. Therefore, it can be concluded that increasing the threshold of vote count will improve the accuracy of line detection. But when the threshold of the number of votes is 60, the Hough transform only detects two edge lines of the crane grabbing the boom, missing the other two edge lines. This indicates that the edge line detection of the crane's grab boom cannot

be achieved solely by increasing the threshold setting of the number of votes. In the edge detection image shown in Figure 10, Due to factors such as uneven lighting ,the linearity of the pixel arrangement of the four edge lines of Crane grabbing boom will undergo a sudden change , There may be slight differences in the parameters of the lines detected on the same edge line. And these subtle differences may directly affect the accuracy of line fitting. Therefore, it is difficult to determine which straight line to use as the edge straight line for the crane to grab the boom, Meanwhile, the purpose of conducting line detection in this article is to calculate the coordinates of the three corner points of the crane’s grab boom through the intersection of lines, Therefore, it is necessary to fit the detected straight lines and simplify them to four.

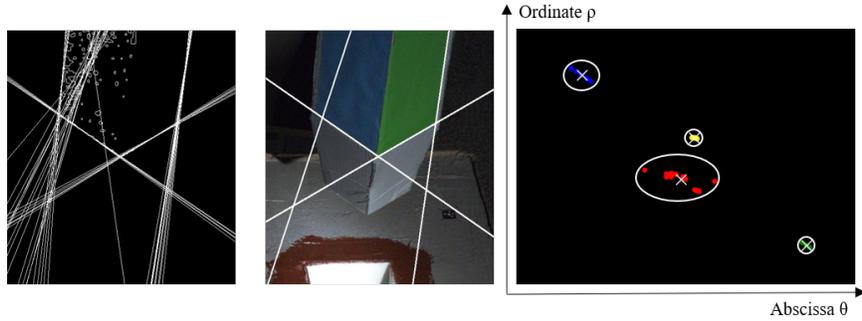
2.2 Clustering of Line Parameter Points Based on k-means

The coordinate points mapped to the Hough space correspond to a straight line in the image coordinate system, Crane grabbing boom has 4 edge lines. Therefore, the clustering method[20-21] is adopted to cluster all the line parameter points detected by the Hough transform into the same number of classes as the number of edges of Crane grabbing boom . Compared to other clustering algorithms, the k-means algorithm has advantages such as simple algorithm idea, fast convergence speed, better clustering effect, and the main parameter that needs to be adjusted is only the number of clusters K. When other clustering algorithms are applied to edge maps with significant interference, they will cluster the coordinates of the line parameter points detected by Hough transform into an uncertain number of clusters. Intuitively manifested as fitting 3, 4, 5, or even more straight lines, From an engineering perspective, it is not possible to guarantee the stability of the coordinates of the three corner points used to obtain the crane’s grasping boom. The k-means algorithm needs to specify the number of clustering clusters K in advance to ensure that the lines detected by the Hough transform can be clustered into four categories, intuitively manifested as fitting four corresponding lines. Therefore, this article uses the k-means algorithm to cluster linear parameter points.

K-means algorithm is a unsupervised learning clustering method. According to the attributes of N data objects, it is divided into K clusters, and the center point of each cluster is calculated. Make the clustering results meet the requirements of high similarity among samples within different clusters and low similarity among samples between different clusters. It uses the Euclidean distance and the sum of squared errors criterion as the evaluation criteria.

Specifically, for a given K clusters,the center point of each cluster is , all data points in the -th cluster are in set , and represents the Nth data point in the -th cluster.The sum of squared errors of the -th cluster and the sum of squared errors are shown below,the goal of the k-means algorithm is to find a partitioning method,the process of recalculating the cluster center point and reallocating data points to the cluster through iterative methods to minimize the total sum of squared errors (SSE).

Perform k-means clustering analysis on the parameter coordinates of straight lines mapped to the Hough space, as shown in Figure 11. The edge detection results shown in Figure 11a contain a large number of interfering edge points, which have weak linear relationships. Under low threshold conditions, pixels with weak linear relationships are easily detected as straight lines. The voting threshold in Figure 11 is 40, Normalize the coordinates of parameter points with votes higher than 40 in the Hough space voter and perform k-means clustering analysis. The clustering results are affected by the interference coordinate points in the Hough space. Figure 11c shows that the clustering center of the red cluster is affected by the interference data points. The intuitive manifestation is that the fitted straight line produces a significant deviation, as shown in Figure 11b.

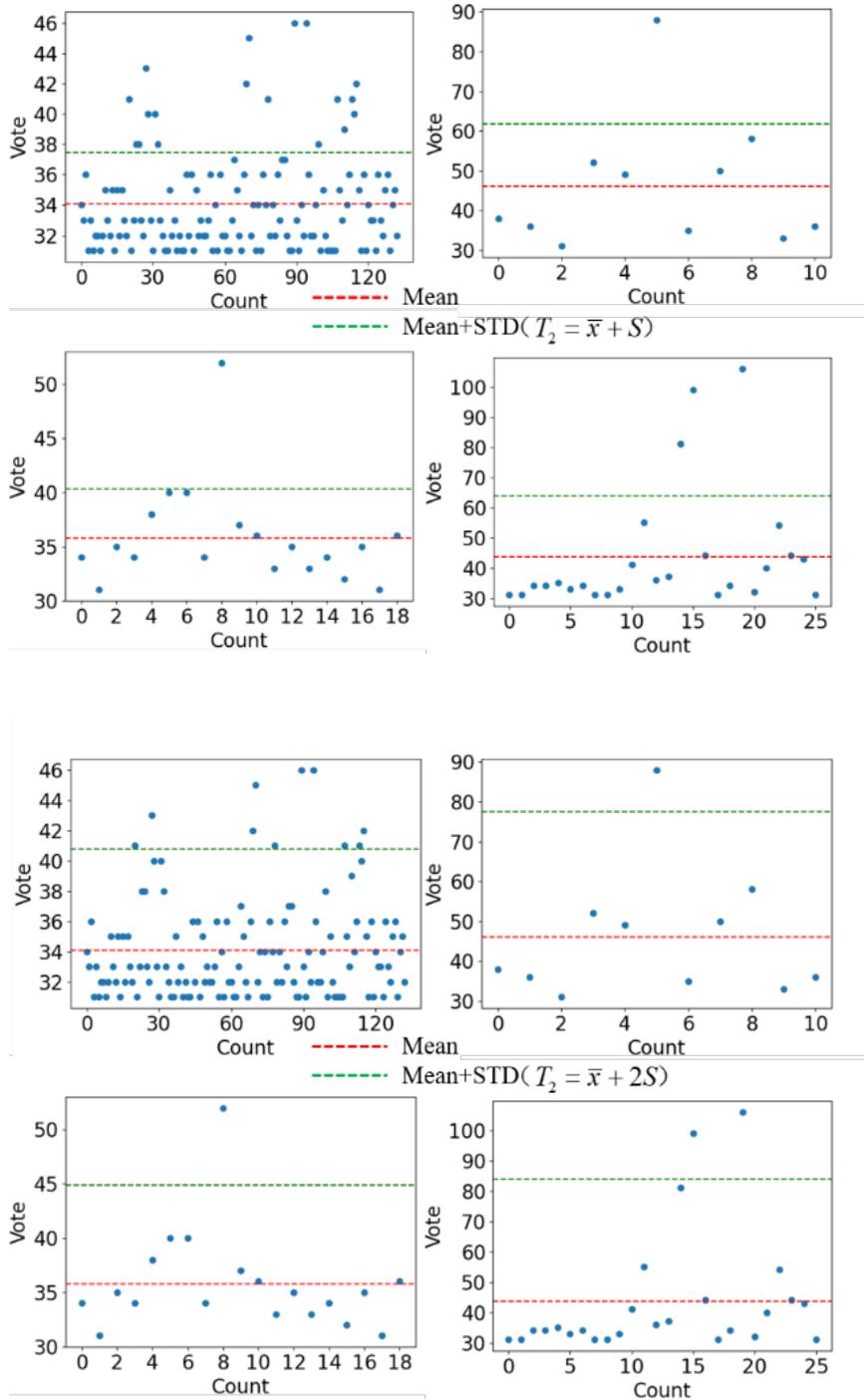


(a) (b) (c)

Figure 11. Parameter Coordinate Clustering Analysis of Lines in Hough Space

2.3 Determine the optimal adaptive threshold by counting the number of votes

To ensure that there is no missed detection during Hough transform, it is necessary to detect straight lines at low voting thresholds. However, the detection results at low voting thresholds are often affected by interfering pixel points, resulting in the detection of many interfering straight lines. Under this premise, the straight line fitting results obtained by k-means clustering often have significant errors, and in extreme cases, the detection method may even fail. Based on the above situation, this article proposes a method for determining the optimal adaptive threshold. After the clustering analysis is completed, this method calculates the distribution of voting numbers, degree of dispersion, and the relationship between the mean and standard deviation of the calculated voting numbers of coordinate points in the parameter space. By observing the distribution of these data points, the optimal adaptive threshold is set. Specifically, the low threshold ensures that Hough transform does not miss detection, and k-means clustering is performed on the data points in the parameter space detected by the low threshold, ensuring that there are four straight lines in the fitting result. Calculate the mean and standard deviation of the number of straight line votes in the i -th cluster clustering results, and set. The value of is generally determined based on experimental conditions and is generally within the range of $[0,2]$. Using the piecewise function to place the element in the i -th cluster that satisfies into a new set. After setting an adaptive threshold for each cluster result, 4 new clustering results will be generated, The number of votes in these four new clustering results is all high, which means that the interference lines detected under low thresholds have been filtered out. Figure 12a shows that when $m=1$, there are many lines in the first cluster that are lower than the green dashed line, all of which are interference lines, Figure 12b shows the distribution of votes when $m=2$. Based on the experimental analysis, this article takes $m=1$. From the experimental results, it can be seen that the method proposed in this article for determining adaptive thresholds based on the number of votes is to identify the lines with high voting numbers in each cluster, preparing for the improved clustering centroid calculation and line fitting method proposed in Section 2.4.



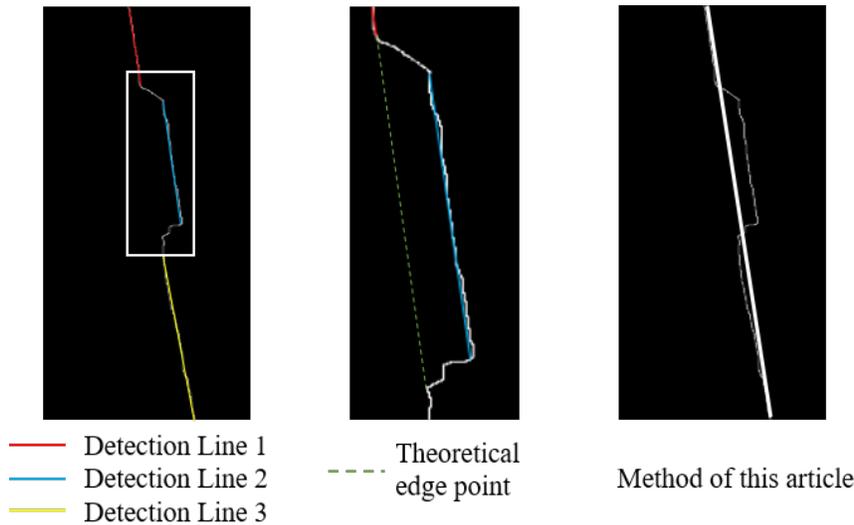
(a) (b)

Figure 12. Distribution of Linear Voting Numbers under Different Values of m

2.4 Improved clustering centroid calculation method

The clustering centroid for k-means will eventually move to the average of all samples within the cluster, which is greatly affected by errors. This article proposes an improved clustering centroid calculation method to replace the original clustering centroid of k-means as the basis for line fitting. Take the weight of the votes of each line into the calculation formula. Figure 13 shows the local image of line detection. When the edge image is affected by uncontrollable factors such as lighting, a small portion of correct edge pixels are lost, but the overall linear relationship does not change much. Therefore, this type of straight line still has some reference value for line fitting and cannot be directly removed. Using the optimal adaptive threshold method proposed in 2.3, find a series of lines with high voting numbers in each cluster. In Figure 13a, in detection lines 1, 2, and 3, detection line 2 shows the detection results of missing edge pixels, but its linear relationship is basically consistent with other detection lines, with slight errors. Therefore, based on the weight of the number of straight line votes, a formula is proposed:

The improved clustering centroid calculation method proposed in this article takes into account the weight proportion of each line, and the calculation results are theoretically more in line with the linear relationship of real edge points. By using the method of determining adaptive thresholds based on the number of votes counted, lines with low votes are deleted, improving the reliability of the lines to be fitted in each cluster. And the improved clustering centroid calculation method further brings the straight line fitting results closer to the real edge in theory. The detection results are shown in Figure 13c.

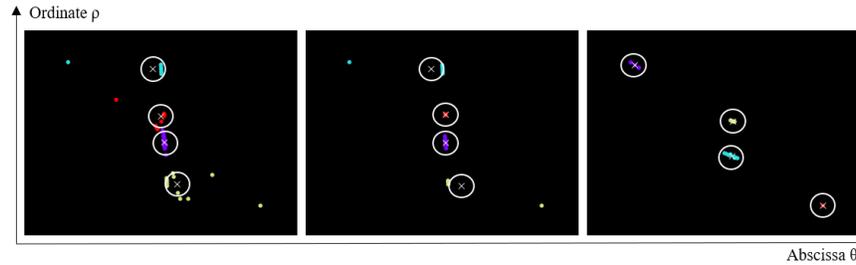


(a) (b) (c)

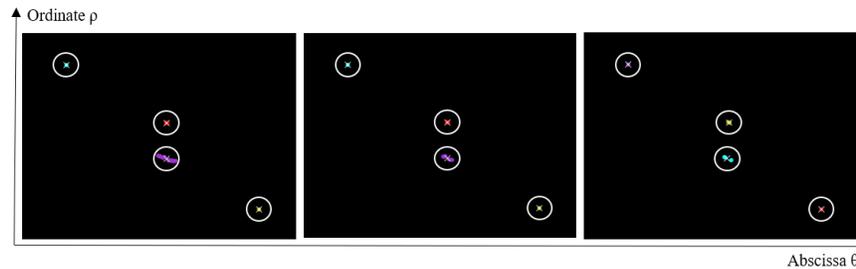
Figure 13. Partial view of line detection

This method is compared with the original k-means clustering results, as shown in Figure 14. Figure 14a shows the clustering results of the original k-means algorithm in the Hough space for the fitted lines at thresholds of 25, 30, and 35. Use circles and oblique crosses to represent the centroids of k-means clustering, and data points in different clusters are represented in different colors. In Figure 14, when is 25, due to the presence of many interfering line parameter points in the Hough space, the clustering results are severely affected by interference, and the interfering points are mistakenly clustered into one category; When is 30, due to the increase in threshold, some of the interference line parameter points in the Hough space have been deleted, but there is still some interference, resulting in inaccurate clustering results; When is 35, the interfering line parameter points are further deleted, and the line parameter points within each cluster happen to be accurate lines, so the clustering results are good. Using the adaptive threshold proposed in

this article combined with an improved clustering centroid calculation method, In Figure 14b, the method of determining the optimal adaptive threshold based on statistical voting numbers adaptively sets according to the distribution of voting numbers, removes the interfering line parameter points, and then recalculates the clustering centroid according to the weight proportion of voting numbers in the clustering results. When the Hough transform threshold is 25, 30, and 35, clustering analysis can be completed correctly. Overall, the clustering robustness and accuracy of the method proposed in this article are superior to the original k-means clustering. The calculation flowchart is shown in Figure 15.



(a)



(b)

Figure 14. Cluster analysis of two algorithms under different thresholds

Figure 15. Calculation flowchart

3 EXPERIMENTAL RESULTS AND ANALYSIS

3.1 Image acquisition scheme for crane grab boom

The automatic grabbing device for segment beams is connected to the automatic grabbing lifting tool through the crane's crown block. The bottom two sides of the lifting bracket are equipped with automatic grabbing lifting rods, which are used to insert the segment beam lifting hole to complete the grabbing task. After it is fully locked, the position of the crown block and the posture of the lifting tool are adjusted through the control system, and the segment beam can be grabbed and placed in a designated area. The layout of the industrial camera is shown in Figure 16. The camera is installed at the bottom of the hanger at a 30 ° angle to the vertical direction. The industrial camera follows the movement of the hanger and maintains its relative position with the automatic grabbing boom, ensuring that the bottom of the automatic grabbing boom is always in the center of the industrial camera's field of view.

Under working conditions, the overhead crane pulls the lifting tool to a certain distance directly above the section beam to be grasped, adjusts the posture to make it basically consistent with the posture of the section beam to be grasped. The industrial camera automatically takes photos and sends the captured image of the suspension rod back to the local computer through the image transmission module. The method described in this article is used for image processing, and the captured photos are also transmitted to the crane driver’s cab for reference by the operator. Using a local computer to calculate the intersection coordinates of straight lines, provide image point coordinates for photogrammetric 3D space calculation, and locate the 3D spatial position of the lifting tool. Output the parameters to the control system, which drives the adjustment mechanism of the suspension rod based on the parameters to achieve precise alignment between the suspension rod and the lifting hole. Finally, the lifting tool is lowered to achieve the purpose of automatic grasping of the segmental beam.

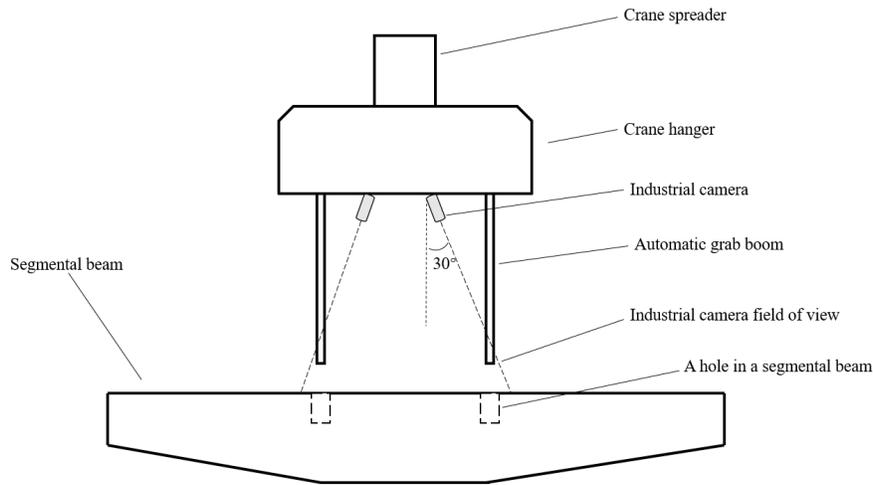


Figure 16. Schematic diagram of industrial camera layout

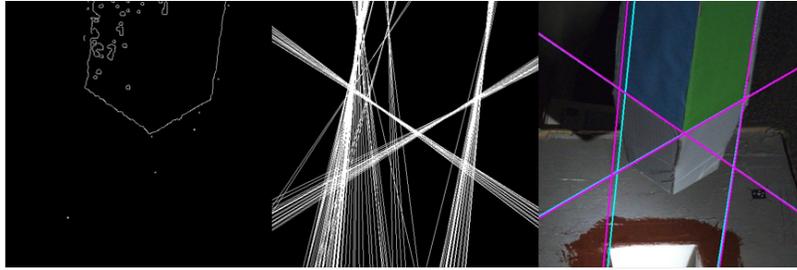
3.2 Evaluation indicators for testing results

In this paper, the average distance between the intersection coordinates between the detection lines and the corresponding corner coordinates of the suspender is used to evaluate the error size of the line detection results, and determine whether the line detection results are within the error range. Manually mark A, B, and C as the suspender corner points, the intersection point between line and line is a, the intersection point between line and is b, and the intersection point between line and is c. Aa, Bb, and Cc are error distances between corresponding coordinate points. Finally, the average error distance of Aa, Bb, and Cc is used to measure the detection result. The schematic diagram of evaluation criteria is shown in Figure 17:

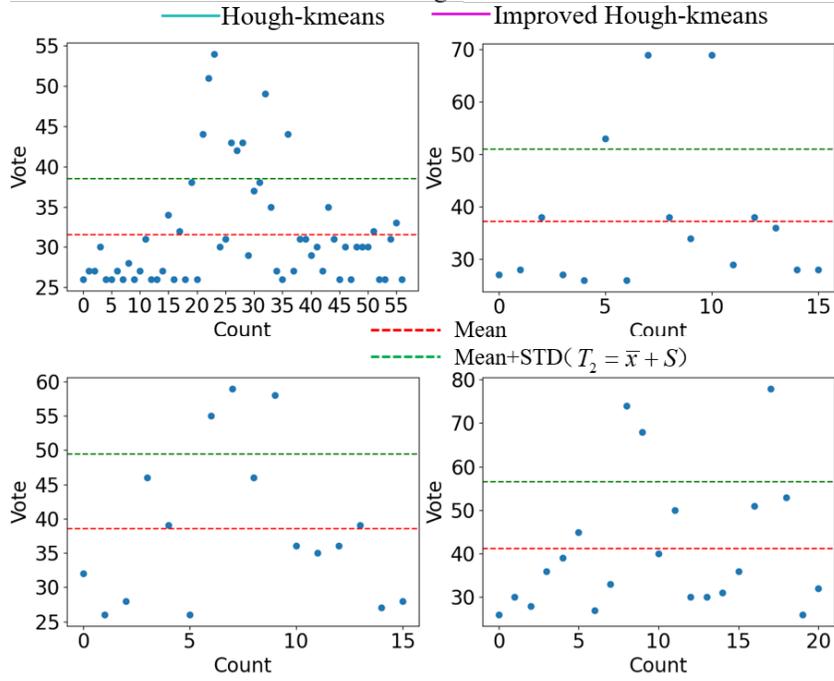
Figure 17. Schematic Diagram of Evaluation Criteria

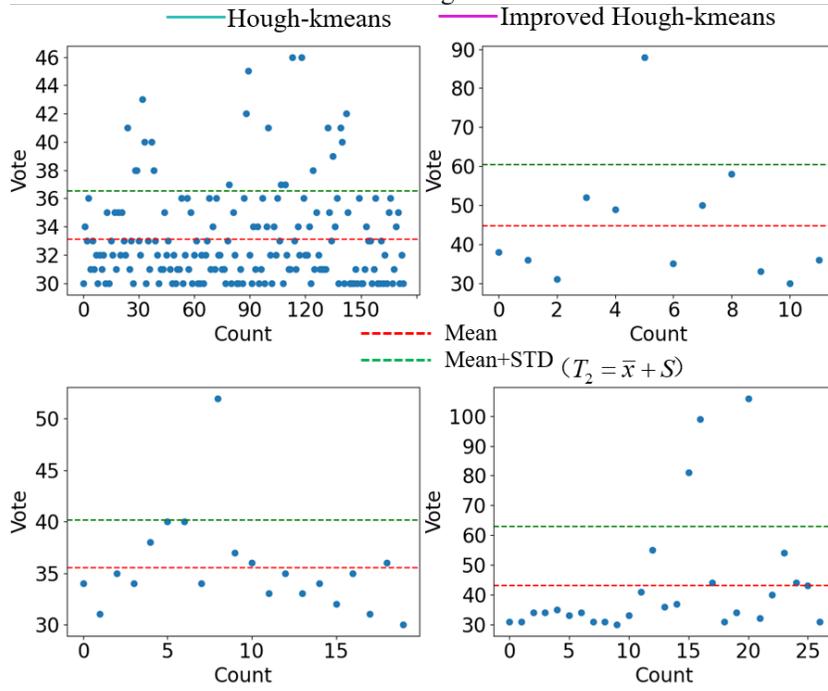
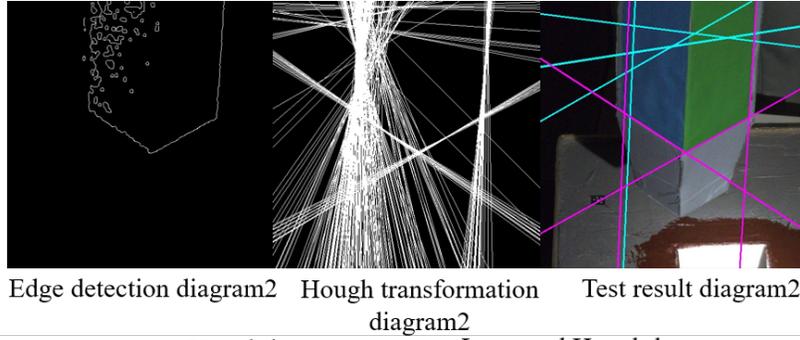
3.2 Analysis of results under different levels of interference

This article lists four datasets with different interferences, and uses the method proposed in this article and the original k-means clustering method to detect four datasets. Each example includes an edge detection graph, a Hough transform graph, the detection results of the two algorithms, and a scatter plot of the straight line voting numbers of the clustering results. To facilitate the display of the detection effect, the detection image is now magnified by 10 times, and the test example is shown in Figure 18.

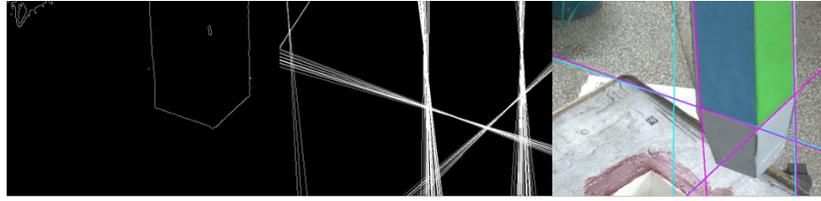


Edge detection diagram1 Hough transformation diagram1 Test result diagram1

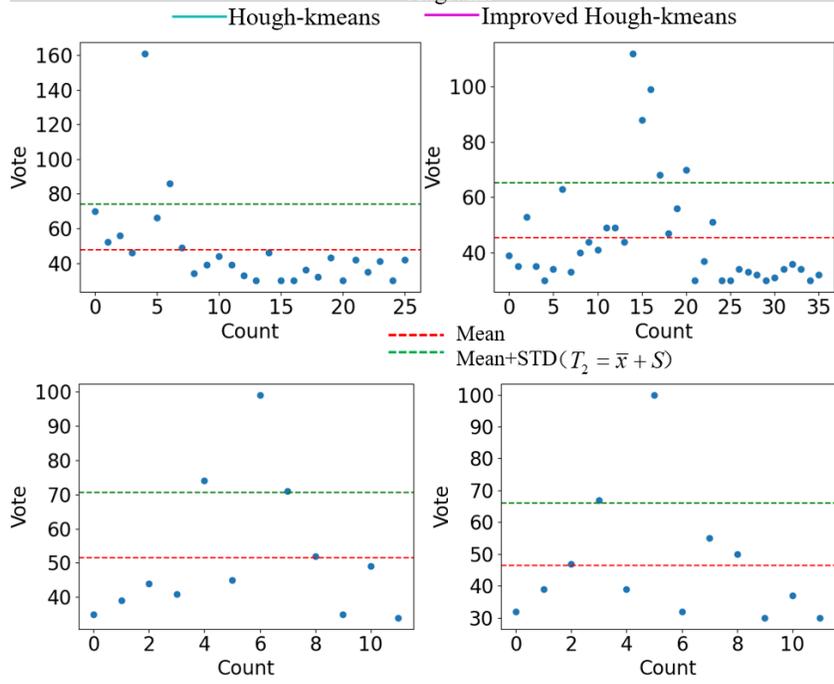


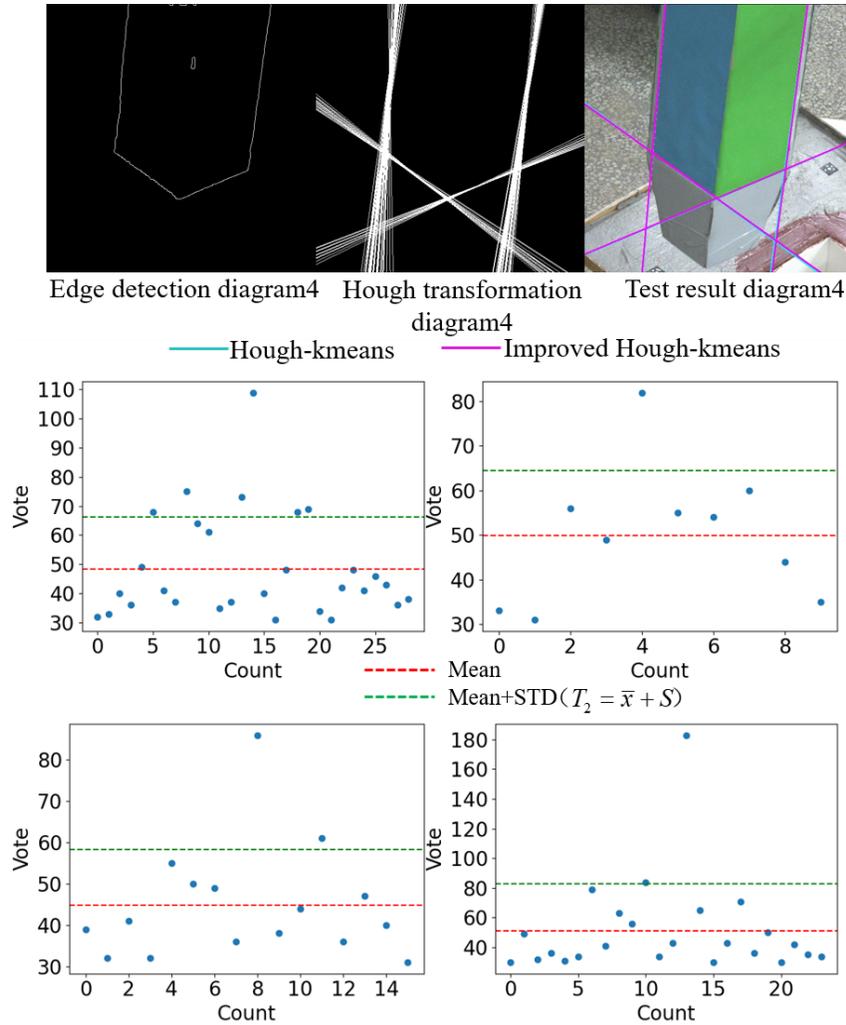


Example (a) Example (b)



Edge detection diagram3 Hough transformation diagram3 Test result diagram3





Example (c) Example (d)

Figure 18. Test Example Diagram

In example (a), the Hough transform threshold T is taken as 30. There are many interfering pixels on the left side of the edge detection image, and there are many non real edge detection lines in the lines detected by the Hough transform. The original k-means method directly clusters the Hough transform results and takes the cluster centroid as the basis for line fitting. In example (a), the method in this article takes $m=1$, which is a standard deviation. After clustering is completed, the threshold is adaptively calculated. The method in this article preserves reliable straight lines, and then recalculates the clustering centroid to replace the original k-means clustering centroid. From the detection results graph, it can be intuitively seen that the fitting line of this method is more accurate in areas with more interference than the fitting line of the original k-means method.

In example (b), there are more interfering pixels on the left side of the edge detection image, making the edge image more noisy and visible to the human eye. Due to the further increase in interfering pixels, the error of the original k-means clustering method further increases, resulting in complete detection failure. Under such interference conditions, the method in this article can still detect the straight lines we need.

In example (c), the edge detection result graph shows that there is interference edge at a distance far from

the crane to grab the boom. Similarly, the original k-means detection result produces a significant error.

In example (d), there are no other interfering pixels on the four edges of the crane grabbing the boom in the edge detection image. In such detection situations, the detection performance of our method is basically consistent with the original k-means clustering method.

These four sets of examples demonstrate that for edge detection images with different levels of interference, the detection performance of our method is superior to the original k-means clustering method, with higher accuracy and robustness.

3.3 Error result analysis

The following verifies the error analysis of the algorithm in this paper and the original k-means fitting method for straight lines when the Hough transform threshold is set to 25, 30, 35, and 40, respectively. This article selects four noisy edge detection images for verification, as shown in Figure 19. In 19 (a), it can also be seen that for the same edge detection image, both algorithms show a decreasing trend in error as the Hough transform threshold continues to increase. But when , the error of the original k-means algorithm in fitting straight lines is significantly greater than the error of the algorithm in this paper. When , the errors of the two algorithms are close. Even under low threshold conditions of Hough transform, the algorithm proposed in this paper has a smaller error and obvious advantages, enabling it to achieve very small errors even on noisy edge detection images. And all four images show this pattern.

To verify the effectiveness of the method proposed in this article under different lighting conditions, some straight line detection results were taken as display, as shown in Figure 20. The images in group 20 (a) show the straight line detection results under outdoor natural lighting conditions. The shadow of the visible part has little impact on the detection results, with a slight deviation. It can detect the straight line of the suspension rod contour normally. Group 20 (b) images show the results of line detection under indoor lighting conditions, which can effectively complete the line detection task. Group 20 (c) images show the results of line detection under dark fill conditions, which can detect corresponding lines. Due to the dim light, there may be detection failures. 20 (d) sets of images show the results of line detection under bright fill light conditions, with sufficient fill light, which can effectively complete the task of line detection. The images in group 20 (e) show the results of line detection under strong supplementary lighting conditions. Under this condition, the light is strong, and compared to slightly darker supplementary lighting conditions, it significantly highlights the target object, making edge detection more accurate, and the overall detection effect is the best.

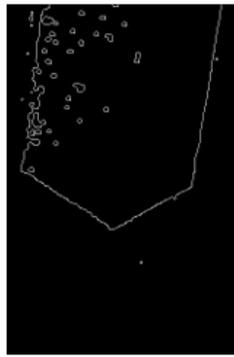
Identify the crane grab boom under different light source conditions, and test 70 images under each light source condition. The recognition accuracy results of the algorithm in this article are shown in Table 1. The average error range is within the interval $[0,10]$, indicating successful recognition. The average error range is within the interval $(10, [?])$, indicating recognition failure. Under dark light conditions, the average detection error within the range of $[0,2]$ accounts for 90.0%, and the recognition success rate is 92.9%. Under strong supplementary light conditions, the recognition success rate is the highest, with an average detection error of 97.1% within 0-2 pixels and a recognition accuracy of 98.6%.

Hough-kmeans algorithm data error table

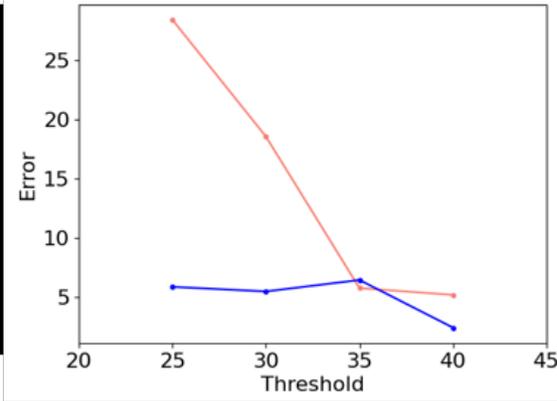
	Aa	Bb	Cc	Mean value
T1=25	34.104	50.397	0.812	28.438
T1=30	21.482	32.109	2.1	18.564
T1=35	13.158	2.689	1.433	5.76
T1=40	9.831	3.483	2.247	5.187

Improved Hough-kmeans algorithm data error table

	Aa	Bb	Cc	Mean value
T1=25	12.842	3.891	0.868	5.867
T1=30	11.076	4.493	0.869	5.479
T1=35	14.238	3.111	1.985	6.445
T1=40	2.893	2.231	2.048	2.42



Edge detection diagram1



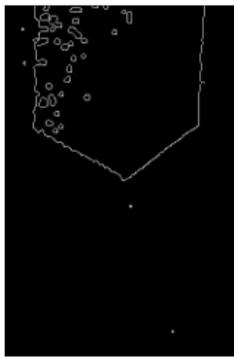
Error Comparison Line Chart 1

Hough-kmeans algorithm data error table

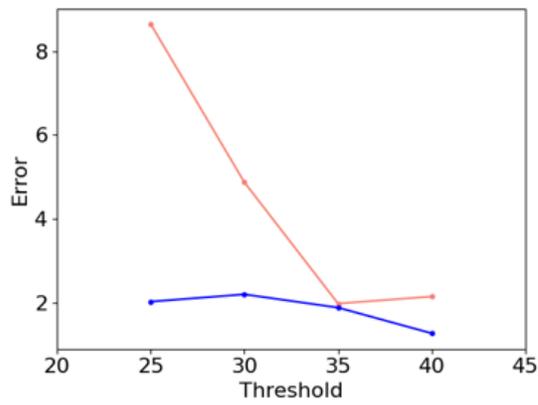
	Aa	Bb	Cc	Mean value
T1=25	21.701	2.459	1.781	8.647
T1=30	11.012	2.438	1.178	4.876
T1=35	2.035	2.449	1.473	1.985
T1=40	2.817	1.911	1.728	2.152

Improved Hough-kmeans algorithm data error table

	Aa	Bb	Cc	Mean value
T1=25	2.399	2.108	1.585	2.031
T1=30	2.919	2.108	1.585	2.204
T1=35	1.972	2.108	1.585	1.888
T1=40	2.181	0.469	1.176	1.275



Edge detection diagram2



Error Comparison Line Chart 2

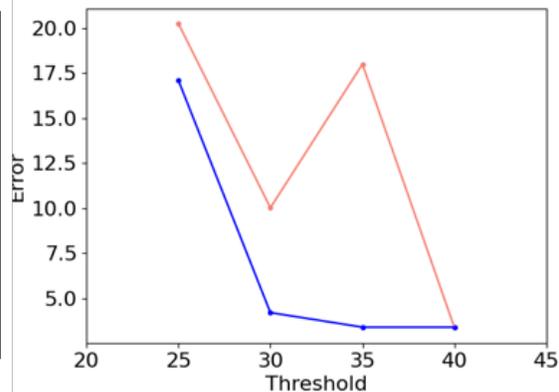
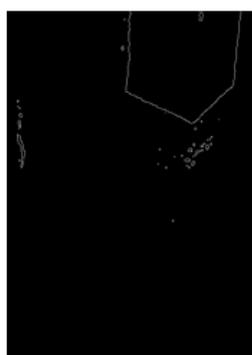
(a) (b)

Hough-kmeans algorithm data error table

	Aa	Bb	Cc	Mean value
T1=25	56.233	3.12	1.283	20.241
T1=30	25.657	3.42	0.995	10.024
T1=35	48.932	3.972	1.056	17.987
T1=40	4.996	4.005	1.118	3.373

Improved Hough-kmeans algorithm data error table

	Aa	Bb	Cc	Mean value
T1=25	45.217	4.197	1.869	17.094
T1=30	5.255	4.726	2.648	4.209
T1=35	3.107	4.726	2.375	3.403
T1=40	3.107	4.726	2.375	3.403



Edge detection diagram3

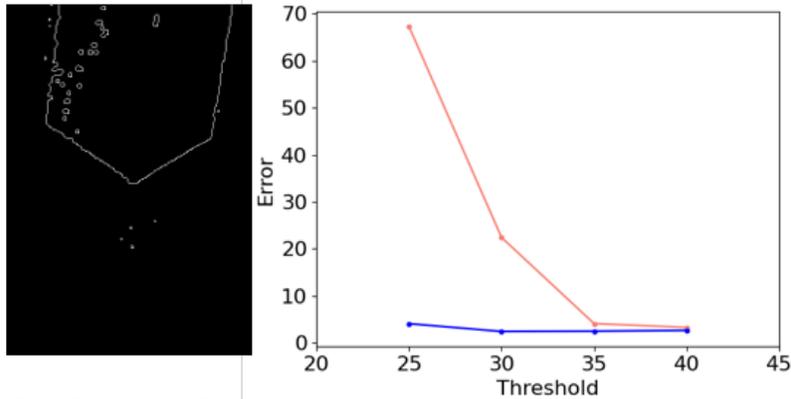
Error Comparison Line Chart 3

Hough-kmeans algorithm data error table

	Aa	Bb	Cc	Mean value
T1=25	46.456	87.779	67.656	67.297
T1=30	14.405	26.585	26.308	22.433
T1=35	6.333	4.114	1.903	4.117
T1=40	2.312	5.479	1.667	3.293

Improved Hough-kmeans algorithm data error table

	Aa	Bb	Cc	Mean value
T1=25	6.286	5.393	0.608	4.095
T1=30	0.869	5.752	0.728	2.441
T1=35	1.175	5.52	0.79	2.495
T1=40	1.107	5.707	1.193	2.668



Edge detection diagram4 Error Comparison Line Chart 4

(c) (d)

Figure 19. Error Results of Hough-kmeans and Improved Hough-kmeans Algorithms.

Comparing the method proposed in this article with the probabilistic Hough line detection method and the LSD line detection method, it can be seen in Figure 21a that the probability Hough line detection has a high false detection and missed detection rate, and some unrelated background lines have been detected, with a large number of lines. The LSD line detection method in Figure 21b also suffers from false detections, with the detected lines being intermittent and a large number of interfering lines in the background being detected. As shown in Figure 21c, the original k-means line fitting algorithm has the advantages of low false detection and missed detection rates. However, when the Hough transform threshold is low, the detection error increases or even fails. The 21d line fitting algorithm proposed in this article only displays the four lines we need after clustering, and still has good detection performance and high robustness even in the presence of significant interference in the edge image.

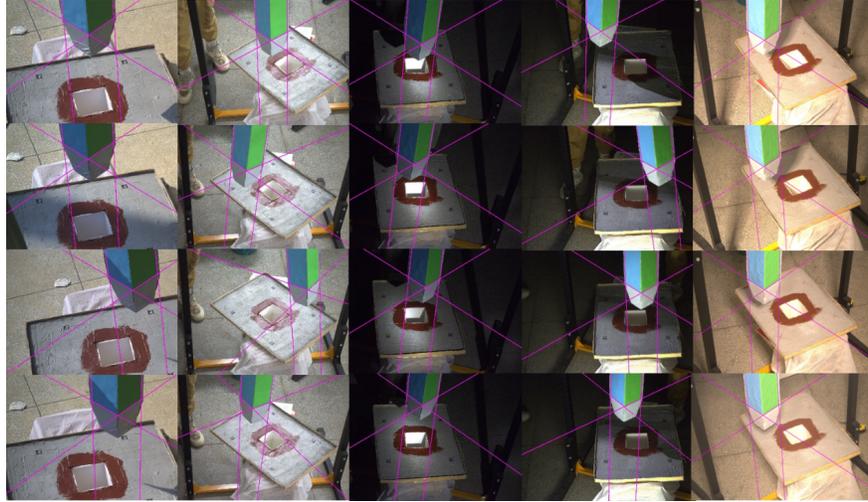
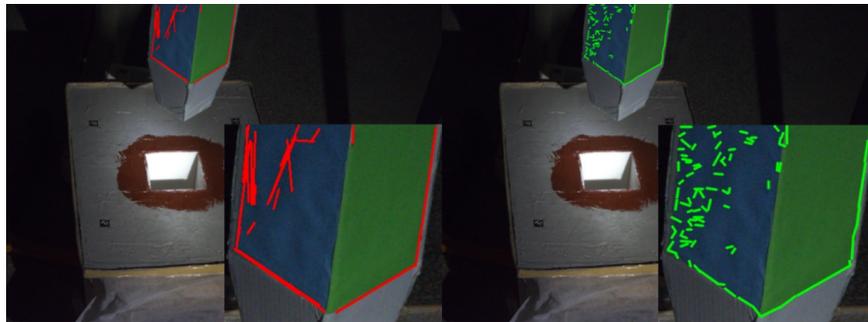


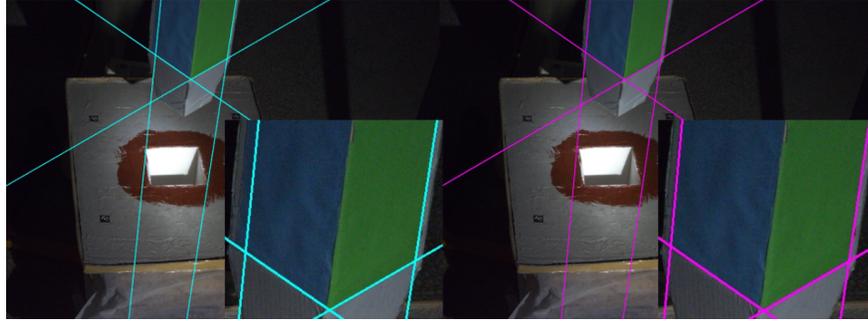
Figure 20. Test results under different lighting conditions

Table 1 Algorithm recognition accuracy under different light source conditions

Different light sources	Average error range(pixel)	Ratio of error[0,2]	Ratio of error[0,10]				
	[0,2]	(2,4]	(4,6]	(6,10]	(10,[?])		
Natural	66	1	0	1	2	94.3	97.1
Indoor	67	1	1	0	1	95.7	98.6
Dim	63	0	1	1	5	90.0	92.9
Bright	67	0	1	0	2	95.7	97.1
Intense	68	1	0	0	1	97.1	98.6



(a) (b)



(c) (d)

Figure 21. Comparison between traditional linear detection methods and this method

4 CONCLUSION

1) To solve the problem of difficult detection of the corner points of the crane's grab boom during the alignment process between the boom and the segmental beam body lifting hole. This article proposes a method for constructing grayscale difference maps, which presents a bimodal feature between the region to be segmented and the background structure, and it more suitable for binary processing using the Otsu algorithm. Accurately extract the edge image of the crane grabbing boom under sufficient lighting conditions and without obstructions. The image preprocessing method in this article solves the problem of over segmentation and under segmentation that other image segmentation methods are prone to, and can meet the segmentation requirements under different lighting conditions. The method proposed in this article has the advantage of fast computation speed, and compared to semantic segmentation, it does not require a large amount of time to produce datasets to train neural networks.

2) The threshold setting of Hough transform requires artificial selection. Due to the changing weather, lighting conditions, construction environment and other external conditions, it is necessary to ensure the reliability of application and engineering application. Combining the Hough transform with k-means clustering, the Hough line detection threshold is set low, simplifying lines with the same features into one, providing a new approach for corner detection.

3) The method proposed in this article for determining the optimal adaptive threshold by counting the number of votes removes the error lines in the low threshold of the Hough transform and ensures that there is no missed detection; Replace the original cluster centroid with an improved clustering centroid calculation method as the basis for line fitting, Improved the accuracy of line fitting under the same lighting conditions and its robustness under different uniform lighting conditions. The algorithm in this article has the best detection performance under strong complementary light conditions. The average detection error within 0-2 pixels accounts for 97.1% and a recognition accuracy of 98.6%. The recognition success rate under different lighting conditions is higher than 92.9%. This method is significantly superior to traditional linear detection methods and meets the needs of automatic gripping of the boom.

Conflicts of Interest: Authors have no conflict of interest relevant to this article

REFERENCE

1. Wu,QX,et,al:The real-time vision measurement of multi-information of the bridge crane's workspace and its application. MEASUREMENT.151,107207(2020)

- 2.Gou,HL,et,al: Automated Selection and Localization of Mobile Cranes in Construction Planning. BUILDINGS.12,05(2022)
- 3.Harris,C.,&Stephens,M,:A combined corner and edge detector. In Alvey vision conference.15,50,pp.10-5244(1988)
4. Shi, J., & Tomasi, C.: Good features to track. In Proceedings of the IEEE conference on computer vision and pattern recognition.pp.593-600(1994)
- 5.Wang,YF,ea,al: A Target Corner Detection Algorithm Based on the Fusion of FAST and Harris. MATHEMATICAL PROBLEMS IN ENGINEERING.2022,4611508(2022)
- 6.Wang,MZ,et,al: Efficient corner detection based on corner enhancement filters. DIGITAL SIGNAL PROCESSING.122,103364(2022)
- 7.Zhang,WN,et,al: Combination of SIFT and Canny Edge Detection for Registration Between SAR and Optical Images. IEEE ACCESSIEEE GEOSCIENCE AND REMOTE SENSING LETTERS.19,4007205(2022)
- 8.Liu,YY,et,al: Farmland Aerial Images Fast-Stitching Method and Application Based on Improved SIFT Algorithm. IEEE ACCESS.10,pp.95411-95424(2022)
9. Hossein-Nejad, Z,et,al: Adaptive RANSAC and extended region-growing algorithm for object recognition over remote-sensing images. MULTIMEDIA TOOLS AND APPLICATIONS.81,22,pp.31685-31708(2022)
- 10.Zhao,Y,et,al: Multi-directional feature refinement network for real-time semantic segmentation in urban street scenes. IET COMPUTER VISION.(2023)
- 11.Jia,DY,et,al: Multi-stream densely connected network for semantic segmentation. IET COMPUTER VISION.16,02,pp.180-191(2021)
- 12.Dong,YS,et,al: Refinement Co-supervision network for real-time semantic segmentation. IET COMPUTER VISION.(2023)
- 13.Zhou,H,et,al: Edge gradient feature and long distance dependency for image semantic segmentation. IET COMPUTER VISION.13,01,pp.53-60(2019)
- 14.Yi,ZK,et,al: Scene-Aware Deep Networks for Semantic Segmantation of Images. IEEE ACCESS.7,pp.69184-69193(2019)
- 15.Bhandari,AK,et,al: Cuttlefish Algorithm-Based Multilevel 3-D Otsu Function for Color Image Segmentation. IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT.69,05,1,pp.1871-1880(2020)
- 16.Chen,JQ,et,al: Navigation path extraction for greenhouse cucumber-picking robots using the prediction-point Hough transform. COMPUTERS AND ELECTRONICS IN AGRICULTURE.180,105911(2021)
- 17.Ziwen,Chen,et,al: Vegetable crop row extraction method based on accumulation threshold of Hough Transformation.TRANSACTION OF THE CHINESE SOCIETY OF AGRICULTURAL ENGINEERING.35,22,PP.314-322(2019)
- 18.Nagai,S,et,al,: Acceleration of surface roughness evaluation using RANSAC and least squares method for Running-in wear process analysis of plateau surface. MEASUREMENT.203,111912(2022)
- 19.Zhou,YJ,et,al,: Accurate coupled lines fitting in an errors-in-variables framework. SURVEY REVIEW.5.,362,pp.386-396(2018)
- 20.Lu,L,et,al,: Moving shadow detection via binocular vision and colour clustering. IET COMPUTER VISION.14,08,pp.665-673(2020)
- 21.Zhou,SB,et,al,: Automatic grayscale image segmentation based on Affinity Propagation clustering. PATTERN ANALYSIS AND APPLICATIONS.23,01,pp.331-348(2020)