

Research on Machine Vision Image Mosaic Algorithm for Multi-workpiece Cutting Platform



Long Xue, Yingyu Cao, Junfen Huang, Jiqiang Huang, Zhengyu Zhang, Lanlan Bai, and Jianting Xi

Abstract The cutter head parts of the tunnel boring machine (TBM) are of different sizes, and the cutting platform is of large size. The panoramic view of the cutting platform that meets the accuracy requirements through image stitching is a prerequisite for realizing the visual guidance of autonomous cutting. An image stitching algorithm is studied to reduce the complexity of the algorithm on the basis of ensuring the stitching precision. The ORB algorithm is selected to implement feature detection, and a series of strategies are adopted on this basis. The ROI region is extracted, and only the images in the region of interest are matched. Then, the feature detection, description and matching are performed by the ORB feature detection algorithm and Hamming distance. The feature points are purified by the RANSAC algorithm, and the transformation matrix H is calculated. Finally, the image fusion is carried out. The experimental results show that the ROI-based ORB image stitching algorithm satisfies the accuracy requirements of subsequent visual guidance.

Keywords Machine vision · Image stitching · ORB algorithm

1 Introduction

The groove processing of cutter head parts of shield machine belongs to the typical small batch discretization mode [1]. The structure of this parts is complex and diverse, and the forms of groove are also different. At present, manual groove and semi-automatic groove are the main two forms [2], which means the quality of groove depends on the proficiency of the workers, and always the production efficiency is very low. Therefore, it is urgent to develop an automatic groove cutting system for the cutter head parts of the robot shield machine, so that the workers can place the workpiece to be cut on the robot working platform at will, and the robot will sense

L. Xue · Y. Cao (✉) · J. Huang · J. Huang · Z. Zhang · L. Bai · J. Xi
School of Mechanical Engineering, Beijing Institute of Petrochemical Technology, Beijing 102617, China
e-mail: caoyingyu@bipt.edu.cn

© Springer Nature Singapore Pte Ltd. 2020
S. Chen et al. (eds.), *Transactions on Intelligent Welding Manufacturing*,
Transactions on Intelligent Welding Manufacturing,
https://doi.org/10.1007/978-981-15-6922-7_7

the workpiece position and plan the cutting path independently, so as to improve the production efficiency and groove processing accuracy.

Robots usually work in the static and structured environment, that is, to keep their position fixed and to complete tasks by running programs written offline in advance [3]. However, in the autonomous cutting system, robots need to meet the requirements of autonomous planning: Autonomous planning means that robots can realize intelligent perception of the external environment, identify, locate the workpiece and plan the path [4]. Now, visual guidance is a more mainstream way of perceiving the environment, which uses two-dimensional or three-dimensional camera to obtain the working environment image to realize the environmental perception. Because the size of cutter head parts of shield machine varies, the smallest one is tens of centimeters, the largest one will be five or six meters, so it is difficult to use a single camera to obtain a panorama of cutting work with required accuracy. Moreover, because the camera needs to be mounted on the mobile mechanism, the acquired image will be deformed due to mechanical jitter. Therefore, it is a prerequisite for independent planning of visual guidance to generate panorama of cutting platform with required accuracy through image splicing algorithm.

In this paper, oriented fast and rotated brief (ORB) algorithm is proposed to splice panoramic images of cutting platform. ORB algorithm is a matching algorithm proposed by Roblee E et al. in 2011 by the IEEE International Conference on compute vision. Its operation speed is greatly improved compared with scale-invariant feature transform (SIFT) and speed-up robot features (SUFT). It is suitable for the industrial environment with high requirements for splicing time and accuracy. Based on the original algorithm, the matching is limited to region of interest (ROI), which reduces the occurrence of matching time and mismatch.

2 ORB Image Mosaic Based on ROI

The experiment was carried out under the condition of workshop. The shoot target of the experiment was a 600 mm × 800 mm working platform. Da Heng GT2450 camera and Computer M0824-MPW2 industrial lens were used. The computer configuration was Intel I5 590 processor with 8G RAM. The algorithm compiling environment was Visual Studio 2015, and OpenCV image processing algorithm library was called. The camera was about 400 mm away from the workpiece, and the shooting field size was 232.7 mm × 172.9 mm. The camera was installed on the six-axis flange of KUKA mechanical arm. The camera moved in one direction to shoot, and the camera position was adjusted to ensure that the coincidence area between the two adjacent images was about 60%.

ORB is a kind of local invariant feature descriptors, which has certain scale invariance to matrix transformation of image such as translation, rotation and scaling. The algorithm flowchart is shown in Fig. 1. The core idea of the algorithm is to use the improved FAST (oriented FAST) algorithm to detect corner points on the image pyramid and generate the main direction, so that the feature points have certain scale

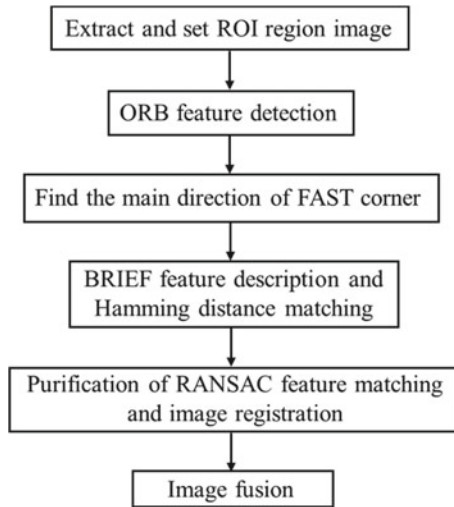


Fig. 1 Algorithm flow of ORB image mosaic based on ROI

invariance [5]; then use the binary robot independent element features (BRIEF) to describe the features, so that the descriptors have certain scale invariance. Finally, through Random Sample Consensus (RANSAC) algorithm, the feature points are purified to find the matching feature points; according to the matching feature points, image registration and image fusion are carried out to realize image splicing.

2.1 Improved Image Matching of ROI

Feature detection and description is a time-consuming part of image mosaic and during image registration; only the feature points in the coincidence areas of different images, which is, the region of interest, are needed. Therefore, if the whole image is matched, the useless feature points outside the overlapped region will consume a lot of time and easily produce mismatches. In this paper, ROI is used to limit image matching to feature points in the region of interest [6], which not only reduces the time of ORB feature extraction, but also reduces the time of matching and mismatching. As shown in Fig. 2, the left and right images are image sequences collected by camera, the coincidence area is the middle part of the two images, the left image's region of interest is the right part, the rest is the uninterested region; similarly, the right image's region of interest is the left part, and the rest is the uninterested region. Generally, the coincidence area of the video frame collected by the camera is 30–60%, In this paper, the camera is installed on the manipulator, and the manipulator will be moved to ensure the coincidence area is about 60%, and 50% of the whole image is considered as ROI area.

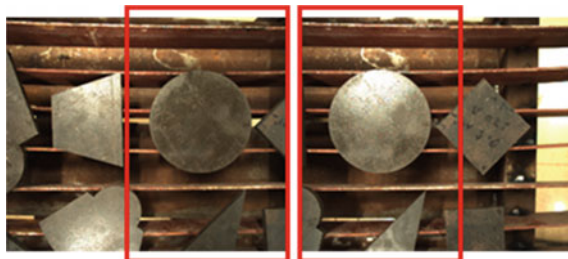


Fig. 2 ROI region selection

2.2 ORB Feature Detection

ORB algorithm uses FAST corner detection to extract feature points. The basic idea of FAST algorithm is that there are enough pixels in the circle neighborhood of a certain radius around the pixel point and the point is in the different area. In gray image, the algorithm can be explained as there are enough pixels whose gray value differs from the point's by more than a certain constant, and then, the pixel point can be considered as a corner point. As shown in Fig. 3, select any pixel point P , and set the pixel value of the point as I_p ; compare the pixel point with the surrounding 16 pixels. If the circle composed of 16 pixels with P as the center has N consecutive pixels, the pixel value is either greater than $I_p + T$ or smaller than $I_p - T$; then, it can be considered as a corner point, where N is generally 9 or 12, and T is the selected threshold.

However, FAST corner detection only compares the gray value of the pixel. Although it has the advantages of simple calculation and fast speed, the feature points detected by FAST do not have scale invariance or rotation invariance. Therefore, it is necessary to calculate the main direction of the feature points so that the algorithm has certain rotation invariance and increases the robustness of the algorithm.

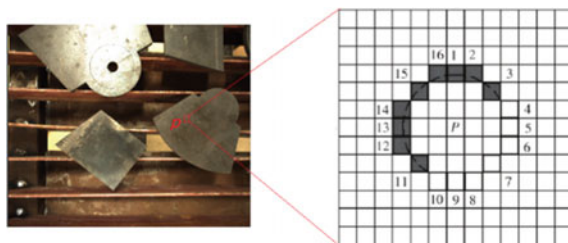


Fig. 3 FAST corner detection

2.3 Finding the Main Direction of Feature Points

In order to make the descriptors obtained by ORB algorithm have rotation invariance, it is necessary to assign a main direction to each feature point. The main direction of FAST corner is calculated by Intensity Centroid [7], which takes the feature point as the center, calculates the position of the center of mass in its neighborhood S , and then constructs a vector with the feature point as the starting point and the center of mass as the end point. The direction of the vector is the main direction of the feature point. The definition is shown in Formula (1):

$$m_{p*q} = \sum_{(x,y) \in S} x^p y^q f(x, y) \tag{1}$$

where p and q are the order of pixel moments in the neighborhood; x and y are the corresponding row and column values; $f(x, y)$ is the gray value of the image, and then, the location of the region's center of mass is:

$$C = \left(\frac{m_{1,0}}{m_{0,0}}, \frac{m_{0,1}}{m_{0,0}} \right) \tag{2}$$

Thus, the direction of the feature points is:

$$\theta = \text{atan}(m_{01}, m_{10}) = \arctan\left(\frac{m_{01}}{m_{10}}\right) = \arctan\left(\frac{\sum_{x*y} yI(x, y)}{\sum_{x*y} xI(x, y)}\right) \tag{3}$$

In order to improve the rotation invariance of the method, it is necessary to ensure that x and y are in the circular region with radius r , which is $x, y \in [-R, R]$, R is equal to the neighborhood radius. After calculating the main direction, the feature points need to be described to prepare for feature point matching.

2.4 Feature Point Description

In ORB, the BRIEF descriptors are used to describe the detected feature points, which solves the primary defect that the BRIEF descriptors do not have rotation invariance.

BRIEF descriptors, which are simple and fast in calculation, contain the main thought that image's neighborhood can provide a relatively small amount of intensity contrast to express.

Define the criterion τ of $S \times S$ size image's neighborhood P as:

$$\tau(p : x, y) = \begin{cases} 1, & p(x) < p(y) \\ 0, & \text{else} \end{cases} \tag{4}$$

where $p(x)$ is the gray value of the image's neighborhood P at $x = (u, v)$ τ after smoothing.

Select n_d sets of (x, y) position pairs, define the binary criterion uniquely, and the BRIEF descriptor is the binary bit string of n_d dimension:

$$f_{nd} = \sum_{1 \leq i \leq 2^{i-1}} \tau(p : x_i, y_i) \quad (5)$$

n_d can be 128, 256, 512, etc. Selecting different values requires a trade-off between speed, storage efficiency and recognition rate. By weighing speed, storage efficiency and recognition rate, n_d in this experiment was set to be 256.

The criterion of image neighborhood only considers a single pixel, which makes its sensitivity toward noise. In order to solve this problem, each test point uses 5×5 sub-windows in the 31×31 -pixel domain, in which the selection of sub-windows follows the Gaussian distribution, and then uses the integral image to accelerate the calculation.

BRIEF descriptors do not have direction or rotation invariance. The solution of ORB is trying to add a direction to BRIEF. For any n pair (x_i, y_i) , define a $2n$ matrix:

$$S = \begin{pmatrix} x_1, x_2, \dots, x_n \\ y_1, y_2, \dots, y_n \end{pmatrix} \quad (6)$$

Using the neighborhood direction θ and the corresponding rotation matrix R_θ , we construct a corrected version of S : $S_\theta : S_\theta = R_\theta S$, then steered BRIEF descriptors will be: $g_n(p, \theta) = f_{nd}(p) | (x_i, y_i) \in S_\theta$.

After obtaining the steered BRIEF, a greedy search is performed to find 256-pixel block pairs with the lowest correlation from all possible pixel block pairs, i.e., the final rotated BRIEF.

2.5 Feature Matching Purification and Image Registration

The feature point detection algorithm described above is used to detect and mark the feature points. The result of feature point matching using Hamming distance is shown in Fig. 4. The connecting points inside figure are the matching feature point pairs in the two images. There are inevitably some mismatches between the feature point pair detected and matched using BRIEF descriptor and Hamming distance. In this paper, RANSAC algorithm is used, and the RANSAC algorithm is a robust parameter estimation algorithm proposed by Fischler and Bons [8].

Image registration is to solve the projection transformation between two images according to the coincidence area of them, i.e., plane homography matrix, so that the images to be spliced under different coordinate systems can be aligned in image space.

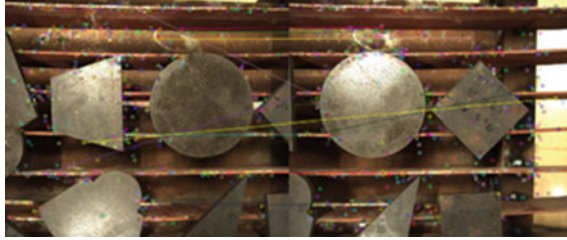


Fig. 4 Show match points

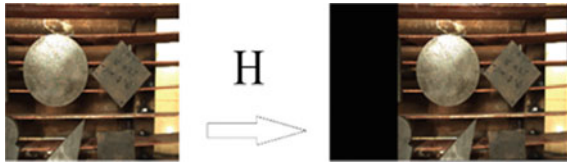


Fig. 5 Image registration

Using matrix H to define the plane homography matrix of two images, $M(x, y)$ and $m(x, y)$ are a set of correctly matched feature point pairs, so the corresponding relationship is:

$$\begin{bmatrix} x'_i \\ y'_i \\ 1 \end{bmatrix} = H \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = \begin{bmatrix} h_0 & h_1 & h_2 \\ h_3 & h_4 & h_5 \\ h_6 & h_7 & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} \tag{7}$$

According to Formula (6), eight degrees of freedom of H matrix can be solved by four pairs of correct feature points, which means, H matrix can be calculated, and then coordinate transformation can be carried out for the image to be registered. Image registration will be realized by applying H matrix transformation to the corresponding image. The result of applying H matrix change to the image to be matched in the above is shown in Fig. 5, and the left picture in the figure is the one before registration, while the pictures right side are after registration.

2.6 Image Fusion

The images collected by different cameras will inevitably have different light intensity, angles, etc. Thus, the simple application of transformation matrix H for transformation and splicing will produce seams, and the worse images will also have defects such as ghosting, distortion and blur. Then, the image fusion part is set to solve this problem and smooth the seams of image splicing. Generally speaking, the

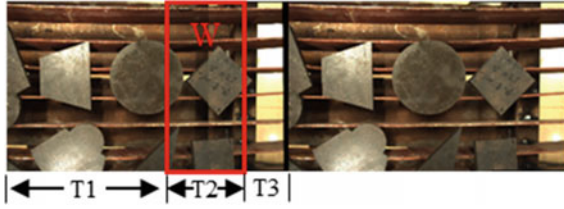


Fig. 6 Before and after image fusion

methods of image fusion include average value method, position weighting method, weighted method with threshold value and Laplace fusion method. Because the camera-moving mode involved in this paper is only translation, position weighting algorithm is adopted [9]. The algorithm complexity is small, and the image fusion effect is good under this condition.

Assuming that $I_1(x, y)$, $I_2(x, y)$ are two images with coincidence areas to be spliced, and the pixels of the fused image are $f(x, y)$, the fusion formula is:

$$f(x, y) = \begin{cases} I_1(x, y) & (x, y) \in T_1 \\ d_1 I_1(x, y) + d_2 I_2(x, y) & (x, y) \in T_2 \\ I_2(x, y) & (x, y) \in T_3 \end{cases} \quad (8)$$

$$\begin{cases} d_1 + d_2 = 1 \\ d_2 = \frac{w_i}{W} \end{cases} \quad (9)$$

d_1, d_2 are the weight values used in the fusion process, w_i is the transverse distance between the current pixel point and the left edge of the coincidence area, W is the total width of the coincidence area of the two images, in this experiment, W was set as 200. T_1 is the value in the left image, T_2 is the value in the coincidence area, and T_3 is the value in the right image. The image is smoothly transitioned by weighting. As shown in Fig. 6, the left image is before image fusion, and the right one is after image fusion. Before image fusion, the stitching trace can be seen clearly. After image fusion using position weighting algorithm, the stitching trace basically disappears that means the success of image fusion.

3 Experimental Results and Analysis

Based on the parameters of the test platform and the requirement of a certain coincidence area between the two images, $3 \times 2 = 6$ images were selected for splicing. Before the images were taken, the optical axis of the camera was adjusted to be perpendicular to the workbench, and the distance that the workbench need to move before and after the two taken adjacent images was determined according to the field

size of the camera and the size that the adjacent images need to be overlapped; then adjusted the camera and lens manually. After the clear image with good contrast could be taken, placed the calibration plate above the workpiece, took the calibration plate image and completed the camera calibration. Finally, the upper left corner of the worktable started to take pictures of the workpieces in blocks. According to the image splicing sequence shown in Fig. 7, the splicing of panoramic images of the large worktable for six images was completed.

Figure 8 shows the six images to be spliced, and Fig. 9 shows the final splicing result. The splicing result is ideal, and the whole splicing process took 3.2 s in debug mode.

As shown in Fig. 9, image correction was carried out for the spliced image. The external and internal participants of the calibrated camera were used to convert the image coordinate system to the world coordinate system. Five lines $L1-L5$ at the junction of the spliced image are taken for accuracy verification. The results were as in Table 1.

According to the calculation results, the error between the length of the line segment in the coincidence area of the mosaic image and the actual length is less than 2 mm, which meets the subsequent needs of the autonomous cutting visual guidance processing.

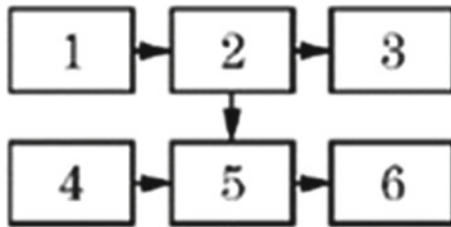


Fig. 7 Splicing method

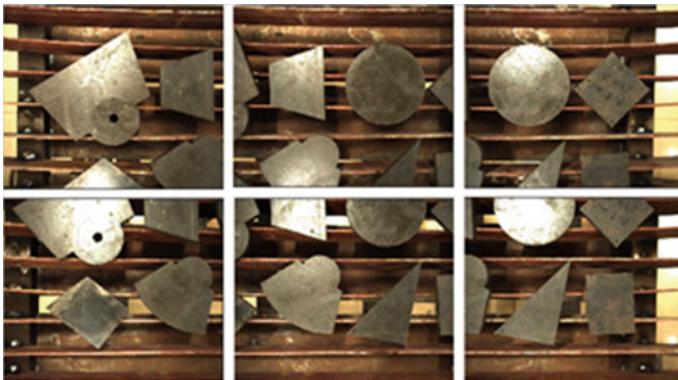


Fig. 8 Six original images to be spliced

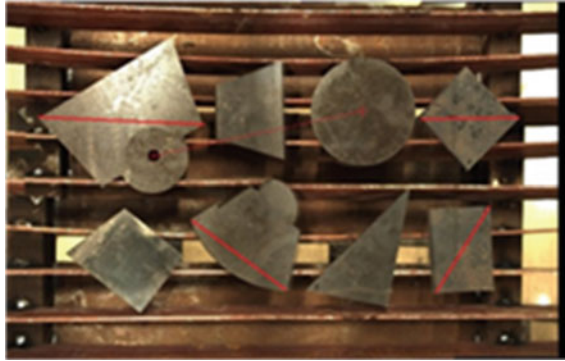


Fig. 9 Splicing result

Table 1 Image mosaic accuracy verification (unit: mm)

Line number	Actual size	Calculated size	Error
<i>L1</i>	112.50	111.64	0.86
<i>L2</i>	117.51	118.68	1.17
<i>L3</i>	253.52	255.19	1.67
<i>L4</i>	137.90	139.29	1.39
<i>L5</i>	212.20	213.89	1.69

4 Conclusion

An image mosaic algorithm for large-scale workpiece cutting platform is proposed, which uses ROI to limit the image-matching feature points in the region of interest; use ORB algorithm to extract feature points, determine the main direction of feature points, and describe the detected feature points; use RANSAC algorithm to combine the plane homography matrix estimation method to eliminate the unmatched points to optimize the matching point sets, and carry out image registration; Finally, the position weighting algorithm is used to complete image fusion and realize image splicing. The experimental results show that the accuracy of panorama obtained by splicing meets the requirements of autonomous cutting vision guidance.

Acknowledgements Thanks for National Key Research and Development Program of China (2017YFB1303300) and Project jointly supported by Beijing Natural Science Foundation and Beijing Education Commission (KZ201810017022).

References

1. Liu XL, Chen S (2010) Current status and prospect of shield machine automatic control technology. *J Mech Eng* 46(20):152–160
2. Wang TM (2007) Fully promote Chinese robot technology. *Robot Technol Appl* 2:17–23
3. Xu F (2007) Industrial robot industry status and development. *Robot Technol Appl* 05:20–30
4. Hong Z (2007) Image processing and analyzing-the core of machine vision. *J Appl Opt* 28(1):121–124
5. Zhang YS, Zhou ZR (2013) Remote registration image automatic registration method based on improved ORB algorithm. *Land Resour Remote Sens* 25(3):20–24
6. Shou ZY, Ning OY, Zhang HJ (2013) Real-time video stitching based on SURF and dynamic ROI. *Comput Eng Des* 34(3):998–1003
7. Rosin PL (1999) Measuring corner properties. *Comput Vis Image Underst* 73(2):291–307
8. Li XH, Xie CM, Jia YZ (2013) Fast target detection algorithm based on ORB features. *J Electron Meas Instrum* 27(5):455–460
9. Fei L, Wang WX, Wang XL (2015) Real-time video stitching method based on improved SURF. *Comput Technol Dev* 25(3):32–35