Chapter 16
Research on Video Mosaic Technology Based on Fully Mechanized Face of Coal Mine

Weihu Zhang, Zhihui Tao, and Xu Li

Abstract In the process of coal mine safety production, video surveillance is one of the important technical means to ensure production safety. However, due to the complicated environment of the coal mine itself, video surveillance has some defects, and the means are single, which cannot effectively complete the effective monitoring of the production environment. Therefore, a video monitoring method that meets the actual needs and is suitable for the complex environment of coal mines is proposed. In this paper, in view of the fact that the coal mine fully mechanized mining face needs the whole roadway visual field to monitor the safety of the whole roadway, the whole roadway monitoring can be completed by studying the panoramic video splicing technology, which can meet the production conditions and facilitate the safe mining activities of coal mines. The experimental results show that the method used in this paper has good robustness and timeliness and can be used for panoramic video stitching in coal mines.

16.1 Introduction

At present, video stitching algorithm has some development but it is not applicable to the actual needs of fully mechanized coal face [1]. The video splicing technology based on ORB feature proposed by Huang [2] effectively satisfies the video splicing requirements and reduces the number of feature extractions by dividing the ROI area of interest to improve the splicing rate. However, the coal mine underground environment is complex, and the dust is large which requires more real-time splicing effect. In 2017, Cheng et al. [3] proposed to stitch images through the SURF algorithm and the improved sampling consensus algorithm to solve the problem of ghosting and stitching. In 2017, Feng and Peng [4] and others proposed to improve the SURF algorithm descriptor to optimize the panoramic video splicing technology. The video

W. Zhang · Z. Tao (✉)
Xi'an University of Science and Technology, Xi'an, China
e-mail: txtao123@163.com

X. Li
Xi'an Hezhiyu Information Technology Corporation, Xi'an, China

S.-L. Peng et al. (eds.), *Sensor Networks and Signal Processing*, Smart Innovation, Systems and Technologies 176, https://doi.org/10.1007/978-981-15-4917-5_16

splicing effect is better but it is not applicable in coal mines. In 2018, Guan et al. [5] proposed improving the SURF algorithm for dimensionality reduction and dividing the region of interest by ROI to improve the efficiency of downhole video stitching. In 2018, Chen and Zhang [6] proposed the importance of panoramic video in coal mining and proposed a feasible method which provided a very constructive opinion for the application of video stitching in coal mines. In view of the current theoretical stage for coal mine video splicing technology in practical applications, there is a problem that the delay is large and more video frames need to be discarded. A method of quickly splicing video is proposed which can meet the actual needs and save more video. Frame information to avoid the loss of real video information in coal mine roadways [7]. And in the real environment of coal mine, the dust is large, the lighting conditions are poor, etc., the corresponding preprocessing technology is adopted, the video frame is denoised and preprocessed, and the video frame information is optimized to provide the basis for subsequent video stitching.

16.2 Introduction to Video Stitching Principle

16.2.1 Technical Route and Block Diagram

In this paper, the video splicing technology collects the video through the coal mine roadway camera, performs video preprocessing, and denoises the video information to provide more feature points extraction for the next video splicing. Then, the Moravec operator was used to extract feature points from different video frames, and FLANN method was used for coarse matching. The RANSAC algorithm performs the pure matching and the model parameter estimation on the extracted feature points, performs affine transformation [8], and then performs video frame fusion and output image. If a scene changes or timeout occurs during video stitching, the matching point is recalculated. If there is no change, the stitching is performed according to the previous affine transform coefficient. This method effectively improves the program running efficiency. The results show that the method has good real-time performance and can meet the actual production needs. The schematic is shown in Fig. 16.1.

16.2.2 Video Stitching Algorithm Flow Introduction

Moravec Operator

This paper detects feature points of video frames by Moravec [9] operator. This operator defines feature points as points with low "autocorrelation". The algorithm detects each pixel of the image, uses a neighborhood around the pixel as a patch, and detects the correlation of the patch with other patches around it. This correlation is

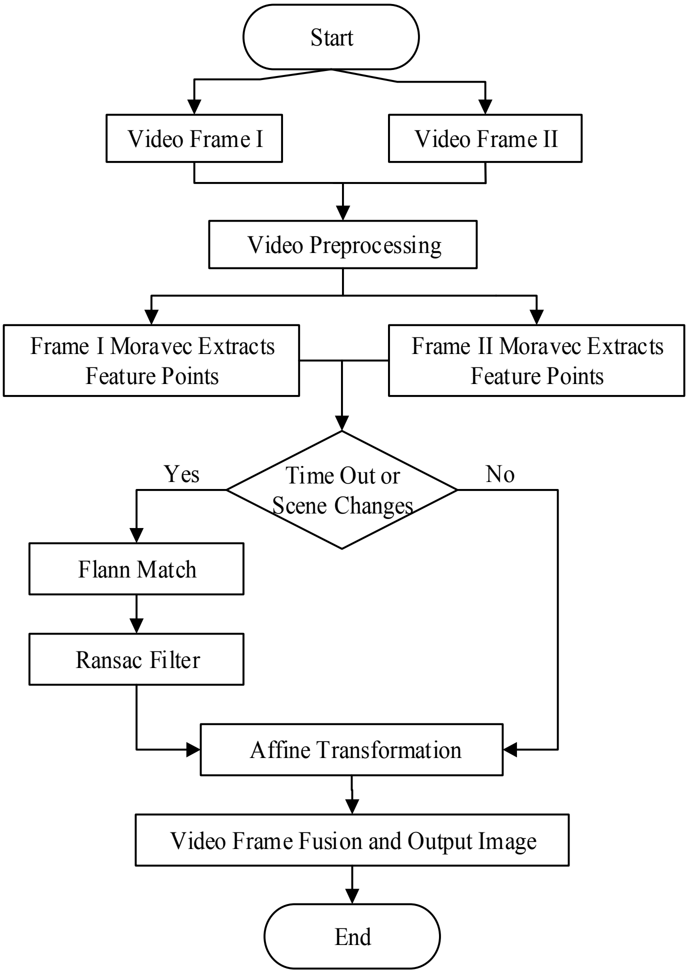


Fig. 16.1 Video mosaic flowchart

measured by the sum of squared differences (SSD) between the two patches, and the smaller the SSD value, the higher the similarity is.

When the Moravec operator calculates feature points, the following occurs: Pixels are in the smooth image area, the surrounding patches are very similar, and the SSD will be small; if pixels are on the edges, the surrounding patches will be in the direction orthogonal to the edges. There are big differences, which are similar in the direction parallel to the edges. SSD is small in one direction, and there are great differences in other directions. Pixels are feature points that change in all directions, and all patches around them are not very similar. In all directions, the SSD will be large. The algorithm steps are as follows.

In the $n * n$ window centered on the image pixel (x, y), the grayscale variance V in the four directions (horizontal, vertical, diagonal and anti-diagonal) of the image, the average grayscale variation, is calculated using the following formula [9]:

$$\begin{cases} V_h = \sum_{i=-k}^{k-1} \left(I_{x+i,y} - I_{x+i+1,y}\right)^2 \\ V_v = \sum_{i=-k}^{k-1} \left(I_{x,y+i} - I_{x,y+i+1}\right)^2 \\ V_d = \sum_{i=-k}^{k-1} \left(I_{x+i,y+i} - I_{x+i+1,y+i+1}\right)^2 \\ V_o = \sum_{i=-k}^{k-1} \left(I_{x+i,y-i} - I_{x+i+1,y-i-1}\right)^2 \end{cases} \tag{16.1}$$

where i represents the value of the change and accumulates k, $k = [n/2]$ represents the number of pixels to be calculated in the window, $[n/2]$ represents rounding, $I(x + i, y)$ represents the pixel value of the corresponding position of the image, and the SSD in each direction is calculated by such a method, and the minimum value is selected as the feature point response value.

According to the set threshold of the actual image, let the window traverse the image and use the feature point response value larger than the threshold as the candidate corner point.

Local non-maximum value suppression selects feature points. In a window of a certain size, the candidate feature points selected in the second part are removed from the point where the response value is not the maximum value, and the reserved feature points are the feature points of the region.

FLANN Match

The result of the feature points matching will result in a correspondence list of the two feature sets. The first set of features is called the training set, and the second set is called the query set. FLANN [10] trains a matcher to improve the speed of the match before calling the match function. The training phase is to optimize execution performance. The training set class will build an index tree of feature sets. Match each feature point of the query set with the training set matcher to find the best match; that is, match the trainer one by one from the features of the query set, that is, each feature set of the query set will have one the best match.

RANSAC Extracts Feature Points

The random sampling consensus algorithm uses the idea of iteration to estimate the correct mathematical model parameters from a set of normal data and anomalous data.

RANSAC [11] achieves its goal by iteratively selecting a random set of data in the data. The selected subset is assumed to be intra-office point and verifies with the following method:

1. Randomly select several data contained in the data and set it as an intra-office point;
2. Calculate the model suitable for the intra-site point;
3. Put the data that is not selected in the first step into the model calculated in the second step to determine whether it is an intra-office point;
4. Record the number of points within the office;
5. Loop iteration, repeat the above steps until you find the model with the most in-office points, so that the coarse matching feature points are filtered.

Parameter determination is as follows [11].

In the sampling consensus algorithm, it is necessary to determine the parameter t (which is used to determine whether the data is adapted to the threshold of the model), d (determine whether the model is suitable for the number of data in the data set), and k (the number of iterations of the algorithm). Where t and d are determined by specific experimental conditions, and k can be deducted by theory.

Assuming that the probability that each point is an intra-point is a, then a is equal to the number of intra-points divided by the total amount of data, but it is not actually known what the w is. The nth power of a indicates that the selected point of each of n times is the probability of the intra-point, and $1 - a^n$ indicates the probability that at least one of the n points is not the intra-point, indicating that a bad one is the calculated model. We use p to indicate a model with a good selection and evaluation.

$$1 - p = \left(1 - a^n\right)^k \tag{16.2}$$

Number of iterations:

$$k = \frac{\log(1 - p)}{\log(1 - a^n)} \tag{16.3}$$

Image Fusion

In this paper, the video frame is fused and outputted by using the fading-integrated fusion method. The method is to linearly assign weights to the pixels of the two images in the overlapping region [12].

$$\text{frame} = \text{frame1} * d + \text{frame2} * (1 - d) \tag{16.4}$$

The frame represents the pixel point of the fused video frame, and frame1 and frame2 are the pixel points of video frame to be splicing. d is the distance from a pixel point in the overlapping region to the boundary. In this method, the pixel points of the corresponding position video frame were fused into one pixel point, and different video frames were merged into one video frame with a larger observation field.

16.3 Result Analysis

To verify whether video splicing can be performed on scenes with poor lighting in real time, we deal with Figs. 16.2, 16.3, and 16.4 through the Moravec corner points and Figs. 16.5, 16.6, and16.7 through the SIFT algorithm.

Through program simulation verification, we can get the difference between the number of extracted feature points and the extraction speed of Moravec corner detection operator and SIFT extraction algorithm (Tables 16.1 and 16.2).

According to the results obtained, we can see that the SIFT algorithm is better than Moravec in extracting effect, but the time required is longer than the corner detection time. Considering that in practical applications, we need more in the case of meeting the demand. Focus on real time, so choose the corner detection algorithm.

Fig. 16.2 Moravec test

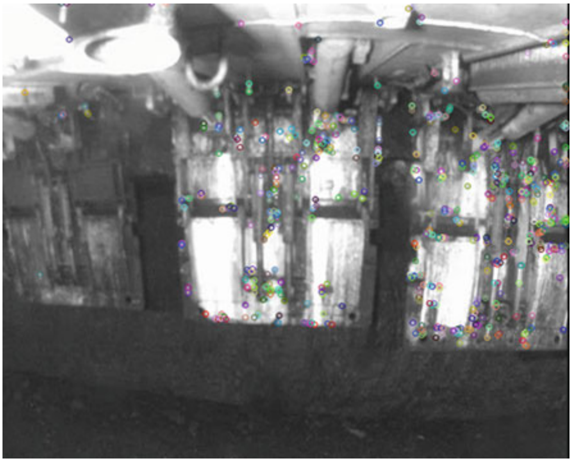
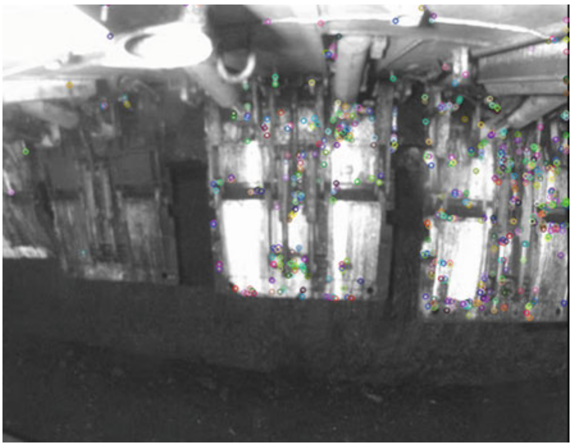


Fig. 16.3 Moravec test



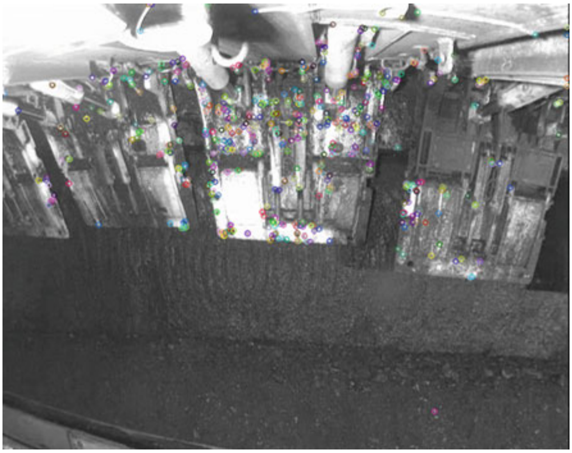


Fig. 16.4 Moravec test

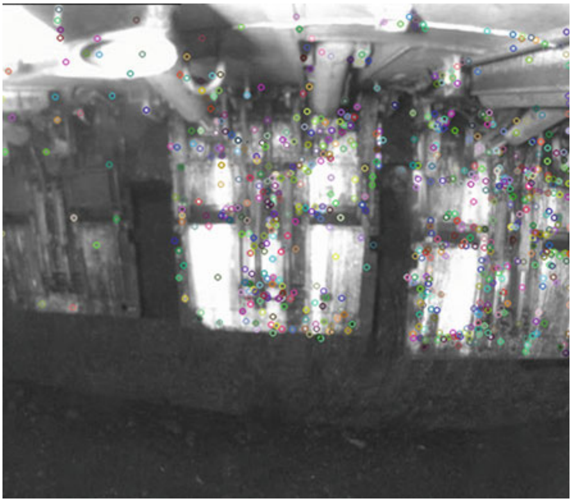


Fig. 16.5 SIFT test

After matching by the FLANN algorithm, the feature matching result filtered by the RANSAC random sampling consensus algorithm. Select two images for splicing as an example, as Fig. 16.8 shows.

Since the video cameras of the intelligent working surface may have angular differences, the viewing angle correction is performed during the preprocessing process of the entire algorithm, thereby ensuring that the video images are in the same plane and ensuring the stitching effect [13]. Figures 16.9 and 16.10 indicate the stitching effect.

It can be seen from Figs. 16.9 to 16.10 that the algorithm effectively meets the actual requirements. The algorithm extracts feature points from the same part of different video frames and then stitches different video frames together by matching

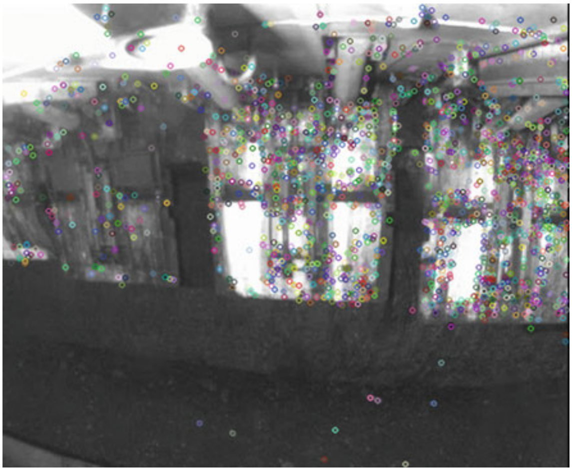


Fig. 16.6 SIFT test

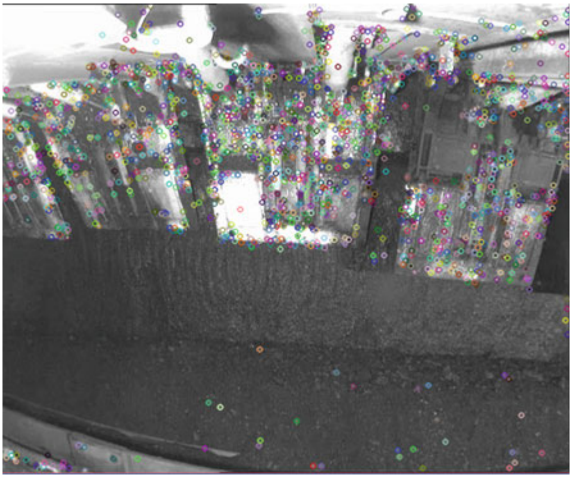


Fig. 16.7 SIFT test

Table 16.1 Moravec test results

	Figure 16.2	Figure 16.3	Figure 16.4
Number of feature points	500	512	618
Speed (ms)	338.021	318.394	319.074

Table 16.2 SIFT test results

	Figure 16.5	Figure 16.6	Figure 16.7
Number of feature points	800	1772	2000
Speed (ms)	753.978	927.605	1052.39

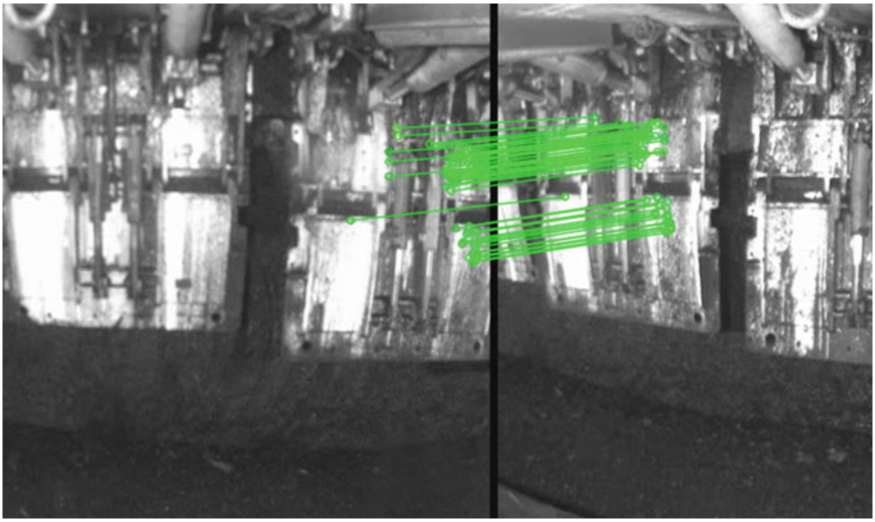


Fig. 16.8 Feature point extraction matching

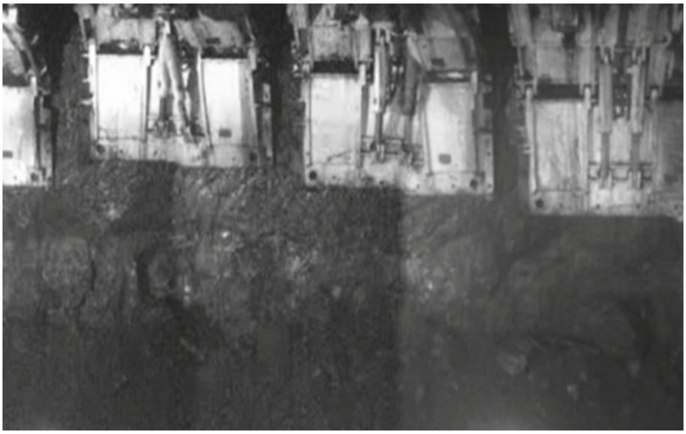


Fig. 16.9 Stitching result 1

feature points. The results show that the splicing work can be performed on the complex coal mining face, so as to save more video frames and avoid the loss of real video information of coal mine roadway. Moreover, by preprocessing the image, the poor illumination condition can be solved and the roadway in the coal mine can be monitored better.

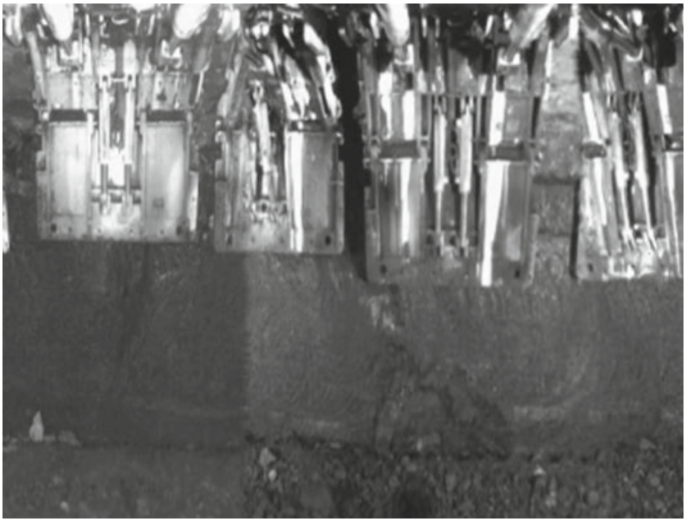


Fig. 16.10 Stitching result 2

16.4 Conclusion

Under the current experimental conditions, compared with the SIFT algorithm, the real-time performance is worse than the Moravec operator, the algorithm proposed in this paper can effectively realize the stitching effect, complete the design and meet the real-time performance, and whether to use the previously retained transform matrix coefficients to perform video in real time by judging whether the scene changes or recalculating the affine transformation matrix within a certain time. But we can also see that the stitching effect is not very good, so next we need to optimize the algorithm under the existing conditions, so that the stitching effect is more impressive.

References

1. Yuan, C.: Development trend of application of coal mine monitoring and control system. Electr. Technol. Softw. Eng. **8**, 100–101 (2019)
2. Huang, J.: Improved video real-time stitching technology based on ORB feature. Electr. Autom. **46**(3), 212–215 (2017)
3. Cheng, D., Li, H., Huang, X.: Video Mosaic algorithm based on improved random sampling consensus algorithm. Work. Condition Autom. **43**(8), 50–55 (2017)
4. Feng, Z., Peng, M.: Research on real-time panoramic video stitching technology. Softw. Guide **16**(8), 193–195 (2017)
5. Guan, Z., Gu, J., Zhao, G.: The downhole video mosaic algorithm based on improved accelerated robust feature. Autom. Work. Conditions **44**(11), 69–74 (2018)
6. Chen, L., Zhang, D.: Application of panoramic video splicing technology in coal mining. Sci. Technol. Econ. Inf. **26**(11), 10–10 (2018)

7. Zhao, X.: Analysis on the problems and counter measures in the application of coal mine safety monitoring system. Eng. Technol. Res. **4**(5), 28–29 (2019)
8. Yu, S.: Classification and application of planar affine transformation. J. Ningbo University (Sci. Technol.) **4**(1), 69–72 (2010)
9. Yu, F., Zhu, D., Hu, L., et al.: Research on feature extraction algorithm of aerial photography image. Softw. Guide **18**(8), 66–70 (2019)
10. Wang, J., Zhou, Z.: Research on feature extraction based on SIFT image and FLANN matching algorithm. Comput. Measur. Control **26**(2), 175–178 (2018)
11. Li, J., Zhang, F., Cui, H.: A homography matrix estimation method based on improved Ransac algorithm. Softw. Guide (2019)
12. Shi, M.H.: Summary of the development of image fusion technology. Comput. Er **9**, 27–29 (2019)
13. Xiang, H., Bai, X.: Photo image correction technology scheme. Technol. Econ. Guide **7**, 33–34 (2016)