

Video Stitching method for a Surveillance System Deployed at a Fully Mechanized Mining Face

Ben Li

Peking University <https://orcid.org/0000-0001-8216-3599>

Yang Yang

Beijing Longruan Technologies Co. Ltd.

Shanjun Mao (✉ sjmao@pku.edu.cn)

Beijing Longruan Technologies Co. Ltd.

Mei Li

Peking University

Research

Keywords: dynamic image stitching, oblique image correction, fully mechanized mining Thee, video surveillance system, intelligent coal mining

Posted Date: February 9th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-190764/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Video Stitching method for a Surveillance System Deployed at a Fully Mechanized Mining Face

Ben Li¹, Yang Yang², Shanjun Mao^{1,2}, Mei Li¹

Institute of Remote Sensing and Geo-informatic, Peking University, Beijing, 100080)

Beijing LongRuan Technologies Co., Ltd, Beijing, 100080)

Corresponding Author: Shanjun Mao (sjmao@pku.edu.cn)

Abstract: Using video stitching technology, video images with overlapping parts can be stitched into a complete image, with characteristics such as intuitiveness, visualization, and measurable analysis. This technology could be applied in the operation of coal mines for a remote monitoring and control of coal production. However, when the technology is used in coal mines, there are several challenges such as non-uniform illumination, missing scenes, and oblique panorama. In this paper, methods were purposed to solve the above problems: (1) To overcome the non-uniform illumination on a mining face, we applied the wide dynamic range technology to the images from a single camera and histogram matching algorithm on multiple images to reduce the color difference between the images; (2) To overcome the missing scene problem due to the narrow field of view (FOV) of a single camera, the SURF matching and template recognition methods are combined to achieve a stable stitching; (3) To overcome the oblique panorama issue, we applied the vertical correction technology exploiting the posture information of the camera, and then the adjacent images are concatenated. The results of practical experiments show that the proposed methods are suitable for solving the above problems in a fully mechanized mining face. The research provides a new approach for displaying extended scenes of stope faces in the intelligent collieries.

Keywords: *dynamic image stitching; oblique image correction; fully mechanized mining face; video surveillance system; intelligent coal mining*

1 Introduction

A video surveillance system in coal mines can help avoid emergency disasters and is a necessary component of modern coal mines for a safe production (Vujic et al., 2008; Şalap et al., 2009; Zhou, 2011). A real-time monitoring of the underground situation and mining progress can be realized by deploying a video surveillance system on the fully mechanized mining face. Generally, because of the long range of the mining face and the undulating terrain, establishing a video surveillance system requires dozens of cameras to cover the entire face, where each camera covers a limited field of view (FOV) and works independently of the other cameras. It is difficult for surveillants to view dozens of videos simultaneously, making it a challenge to grasp the whole mining situations in time. Clearly, it is valuable to improving the usability of

the video surveillance in a coal mine.

Video stitching has emerged as a new research field in computer vision. This method can be used to stitch two or more digital images with overlapping parts into a panoramic image whose FOV is larger. It has been widely used in many applications such as mine inspection (König & O'Leary, 2020), medical treatment (Zheng et al., 2020), remote sensing (González-García et al., 2020), and 3D reconstruction (Li et al., 2019). However, when it is applied to videos from the surveillance system deployed in coal mines, the stitching technology is associated with the following problems: (1) Underexposure and overexposure due to non-uniform illumination; (2) Missing scenes due to insufficient FOV of the cameras; (3) Obliquity of the panoramic images due to pose change of the cameras.

In underground coal mines, the artificial light is non-uniform so that underexposure and overexposure are common. It is necessary to enhance images taken from coal mine surveillance systems. To this end, algorithms based on illumination adjustment and bi- γ function have been proposed for enhancing non-uniform illumination scenes (Zhi et al., 2017, 2018). Improved algorithms based on the wavelet transformation have contributed to denoising and defogging (Yanqin, 2013; Zhao et al., 2015; Yu et al., 2016). Some researchers used deep networks to solve the low-contrast problem (De-yong & Ze-xun, 2019; Yuehua & Weiqiang 2019). Other methods, such as histogram equalization (Yu et al., 2016), visual characteristics (Hua & Jiang, 2014), dark prime color, and CLANE algorithm (Dongmei & Siqi, 2019), have been applied to enhance coal mine images.

Although post-enhanced images contain more evident features, the stitching of scenes in a fully mechanized mining face is still associated with some problems. Most existing video cameras in fully mechanized mining faces are deployed at the top of the hydraulic supports. The adjacent cameras are too far away to make their coverage overlapped. Therefore, the FOV of a single camera must be expanded to make image stitching possible. Recent studies related to image stitching have focused on reducing the deformation and shortening the running time. Shum et al. (2000) proposed a combination of global and local registration methods to enable unstable handheld shooting methods to quickly obtain panoramic images. Qu et al. (2020) significantly reduced the complexity of computation based on a binary tree and rectified deformed image by estimating overlapped area. The RUF algorithm proposed by Liu et al. (2020) helped improve the accuracy and real-time performance of stitching images. Vishwakarma and Bhuyan (2020) first applied automatic image sorting and reduced the computational complexity of feature detection to improve the efficiency of image stitching. In terms of the real-time performance, the algorithm proposed by Yoon and Lee (2018) can achieve 13 fps for continuous video obtained by a motion camera. Bai et al. (2020) proposed a video stitching method for monitoring systems in coal mines, reaching a frame rate of 26 fps.

Although image stitching can expand the FOV of a single camera, the obtained panoramic images will have a perspective effect in that closer objects appear large and farther objects appear small. Therefore, the extended images captured by adjacent cameras cannot be directly concatenated. The perspective effect is due to oblique shooting, so it can be eliminated by vertical correction. However, the correction of oblique images typically depends on the control points information of the scene (Wang et al., 2011; Wang et al., 2016) or the 3D information of the scene (Habib et al., 2007; Dong et al., 2019). Information of this sort is difficult to obtain in the case of fully mechanized mining faces, making it a challenge to correct the video obliquity.

To realize video stitching of videos taken by existing video surveillance systems deployed at a fully mechanized mining face, methods are proposed in this paper to solve the above problems. In terms of enhancing coal mine images, we performed a wide dynamic adjustment and multi-image histogram matching, to increase the information entropy in a single image and to eliminate hue difference between multiple images under non-uniform illumination. In aspect of image stitching, we used a combination of SURF matching and template recognition to achieve stable stitching. For correcting oblique images, we applied the photogrammetry

principle to complete the vertical correction with the posture information of the camera, enabling the concatenation of adjacent images.

2 Methods

2.1 Principle of image enhancement under non-uniform illumination

2.1.1 Enhancement of images taken by a single camera

The cameras deployed on a fully mechanized mining face are typically located between hydraulic supports. Therefore, their FOV is limited, and the image quality depends heavily on the light source. Due to space and electricity constraints, the distribution of the artificial light sources is nonuniform, resulting in an uneven image quality. The poor image quality can be credited to two reasons. When strong light appears in the scene, images nearby are overexposed. On the other hand, when the illuminance of the scene is extremely low, images are serious underexposed. These issues will lead to loss of image information.

By tuning metering methods, shutter speeds, the size of aperture, the sensitivity, and other parameters of the camera, we can resolve the overexposure or underexposure problems separately. However, simply tuning the above parameters cannot significantly improve the quality of an image which is partially overexposed and underexposed. To this end, we applied a wide dynamic range (WDR) technology to simultaneously perform strong light suppression and low illumination recovery, thus successfully improving the original image quality.

The WDR technology uses multiple exposures of the same scene with different shutter speeds, and then a high-speed digital signal processor combines the multiple images to increase the dynamic range of the output image. For the highlighted part of the scene, the processor will select an image with a high shutter speed; for the shadow part of the scene, the processor will select an image with a low shutter speed. Finally, the selected parts are combined to obtain an output image with a more uniform brightness.

2.1.2 Enhancement of images taken by multiple cameras

The images shot by a single camera cannot provide a large scene of the mining face. Therefore, it is necessary to stitch images from the same camera at continuous poses and in continuous moments. Because of the differences in the light conditions and cameras parameter configuration, the obtained images always differ in aspects of the brightness, contrast, and saturation. Stitching them directly will have unexpected results or even fail. Therefore, these differences should be minimized. We used the histogram matching technology to reduce the differences in the characteristics of the different images and to improve the effect of panoramic images.

The principle of histogram matching is to approximate the probability density function of the source image pixel to that of the target image pixel. Suppose the gray value is continuous in the interval $[0, L - 1]$ in the source image, r is the gray value, and $P_r(r)$ is the probability density function; in the target image, z is the gray value, and $P_z(z)$ is the probability density function. Here, $P_r(r)$ and $P_z(z)$ are known functions, and the conversion relationship between the random variables r and z is completed with the help of an intermediate random variable s . s is a random variable in the interval $[0, 255]$ following a uniform distribution, so the relationships between s and r and z are as follows:

$$s = T(r) = (L - 1) \int_0^r p_r(\omega) d\omega \quad (1)$$

$$s = G(z) = (L - 1) \int_0^z p_z(\omega) d\omega \quad (2)$$

We need to express s using r (expressed in Equation (1)) and then find the inverse function of $G(z)$:

$$z = G^{-1}(s) \quad (3)$$

For digital images, the random variables r , z , and s are discrete, so we discretize the above formulae:

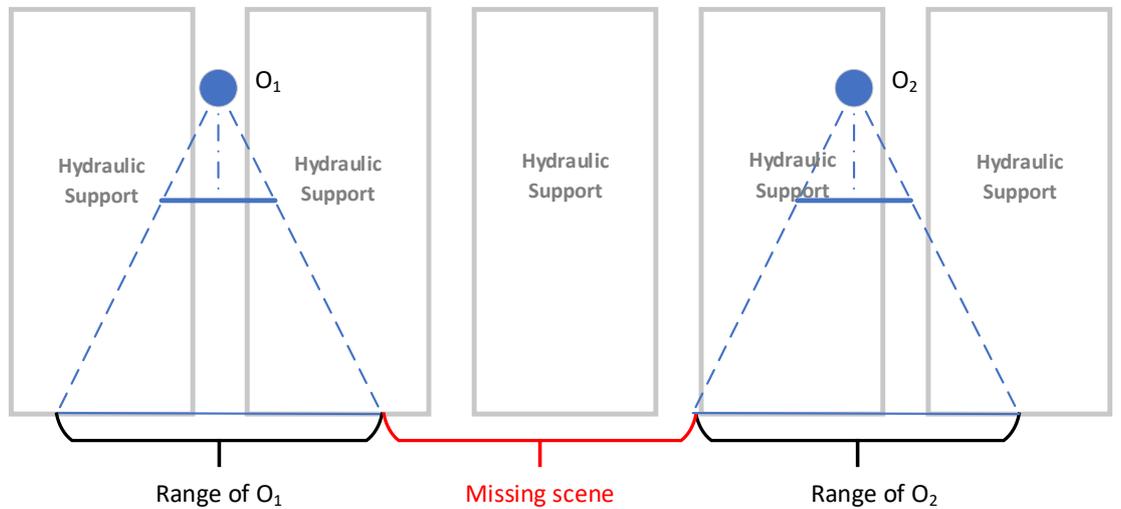
$$s = (L - 1) \sum_{i=0}^r P_r(i), i \in 0, 1, 2, \dots, L - 1 \quad (4)$$

$$z = x, \text{ where } (L - 1) \sum_{i=0}^x P_z(i) \approx s \quad (5)$$

The histogram of the processed image will be similar to that of the target image. In practical experiments, we selected images with better lighting conditions as target images, and the others were histogram matched to the target image, for obtaining the consistency of the brightness, saturation, and contrast of all images.

2.2 Stable stitching by combining SURF matching and template recognition

Because the number of cameras is limited, adjacent cameras are separated by 2–3 hydraulic supports. When shot from the same pose, the images from adjacent cameras cannot be directly stitched (shown in Fig. 1(a)), because no overlapped information exists between the two images. To solve this problem, we successively obtained images shot by the same camera at three different angles (shown in Fig. 1(b)) and then applied the SURF matching and template recognition method to stitch the three images. The three images are stitched to expand the FOV of a single camera, so that the images acquired by adjacent cameras are overlapped to each other, preparing for the next concatenation.



(a)

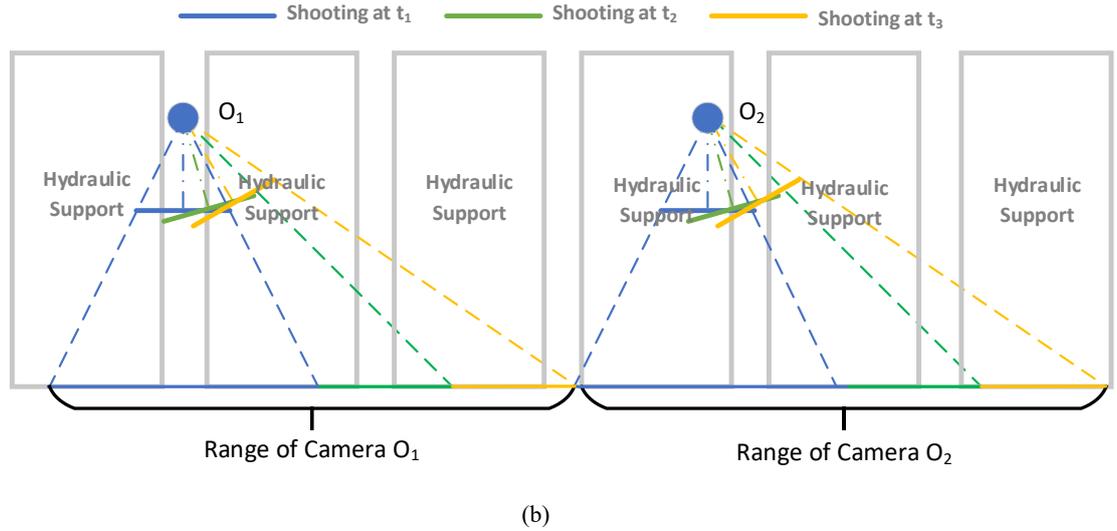


Fig. 1 (a): Single FOV; (b): Triple FOV

The stability is vital important for coal mine video image stitching. A failure of stitching may be credited to the following issues: (1) The image contains a large amount of clastic coal, which leads to fake point features and chaos; (2) The mining face is in operation, and the movement of personnel and equipment can cause point-to-point matching errors. Although stitching methods based on point features matching are relatively mature, affected by the above two issues, stitching coal mine video images requires a more stable method. Methods of this sort are allowed to be inferior to point features matching methods in aspect of the visual effect, but must be more stable than the counterpart. Therefore, we purposed a combination of point features and template recognition method (Fig. 2).

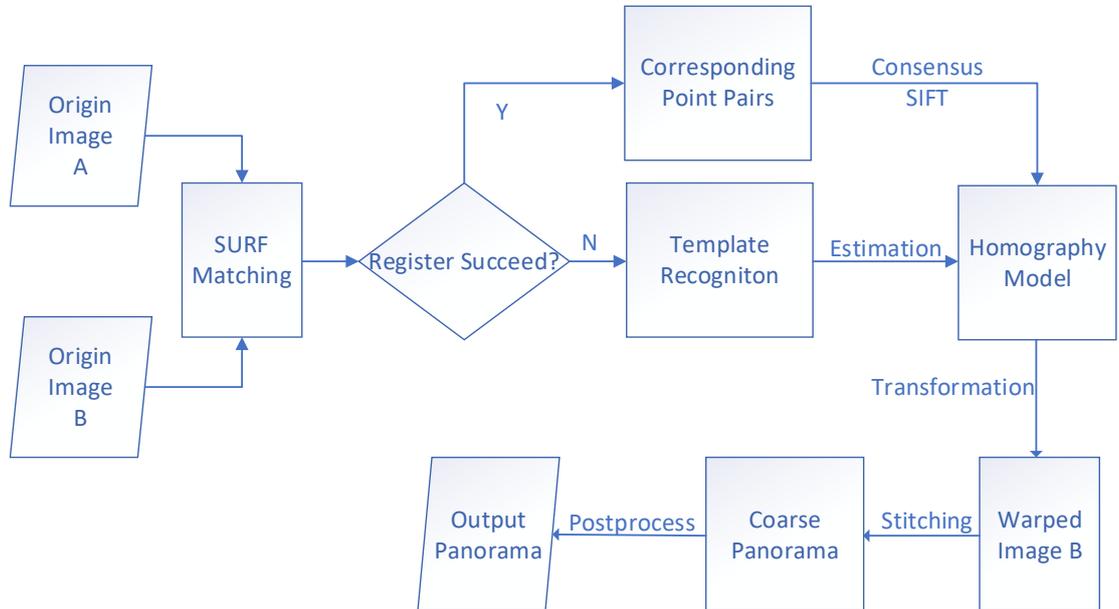


Fig. 2 Pipeline of a more stable stitching method purposed by this article

2.2.1 Stitching based on SURF matching

The speeded up robust feature (SURF) (Bay et al., 2008) was used as point features in this article. In SURF, the detection process is based on the determinant estimation of the Hessian matrix at different scales, where the scale space is generated by Gaussian filtering. For a window template size of $(2k + 1) \times (2k + 1)$, the pixel value after Gaussian filtering can be

represented as follows:

$$G(x, y, \sigma) = \sum_{i=-k}^k \sum_{j=-k}^k w_{ij} I(x + i, y + j) \quad (6)$$

$$w_{ij} = \frac{1}{2\pi\sigma^2} e^{-\frac{i^2+j^2}{2\sigma^2}} \quad (7)$$

In Equations (6) and (7), σ represents the scale parameter of Gaussian filtering. A higher σ will result in a severer blurred image. The calculation formula of the Hessian matrix is:

$$H(x, y) = \begin{bmatrix} \frac{\partial^2 I}{\partial x^2} & \frac{\partial^2 I}{\partial x \partial y} \\ \frac{\partial^2 I}{\partial x \partial y} & \frac{\partial^2 I}{\partial y^2} \end{bmatrix} \quad (8)$$

In Equation (8), x and y are the pixel coordinates, and I represents the corresponded pixel value. The second derivative in Equation (8) is approximated by the second-order difference:

$$\frac{\partial^2 I}{\partial x^2} = I(x + 1, y) + I(x - 1, y) - 2I(x, y) \quad (9)$$

$$\frac{\partial^2 I}{\partial x \partial y} = I(x + 1, y + 1) + I(x - 1, y - 1) - I(x - 1, y + 1) - I(x + 1, y - 1) \quad (10)$$

$$\frac{\partial^2 I}{\partial y^2} = I(x, y + 1) + I(x, y - 1) - 2I(x, y) \quad (11)$$

The Gaussian filtering and Hessian matrix are jointly represented as:

$$H(x, y, \sigma) = \begin{bmatrix} L_{xx}(x, y, \sigma) & L_{xy}(x, y, \sigma) \\ L_{xy}(x, y, \sigma) & L_{yy}(x, y, \sigma) \end{bmatrix} \quad (12)$$

In Equation (12), L_{xx} , L_{xy} , and L_{yy} respectively represent the second derivatives of the image after Gaussian filtering at coordinates (x, y) , and their approximations are the same as that of Equations (9), (10), and (11).

When computing Hessian matrices at different *scales*, a simplified filter is used, which reduces the computing complexity significantly. Different scales are represented by different sizes of the filter window. The minimum window size of the filter should be 9, which is equivalent to the case where Gaussian filter parameter σ equals 1.20. The selection of the possible point features is completed by comparing the Hessian determinant, whose estimation is greater (or lower) than its neighbors are selected. More concretely, a selected feature P has the extreme value among its 27 neighbors (9 of them are at the same scale as P , and 18 of them are at the adjacent scale as P). The estimation of the determinant is:

$$Det(H) \approx L_{xx} * L_{yy} - (0.9L_{xy})^2 \quad (13)$$

After the possible point features are initially located, filters for removing false positives are conducted to obtain the final point features.

The direction invariance is one of the merits of SURF. To enable this trait, the main direction of the feature should be calculated. The Harr wavelet response in the adjacent area of the point feature is used to determine the main direction. In a 60° fan area radius of 6σ with the point feature as the center, the sum of the horizontal and vertical Harr wavelet features of

all points is obtained. As the fan rotating 0.2 radians by step, the value of the corresponded direction is obtained. Finally, the main direction of the point feature is selected, whose value is highest among all directions.

Descriptors are essential for the point features. To calculate the SURF descriptor, firstly, a square area around the point feature height of 20σ is chosen, which can be equally divided into 4×4 sub-areas. Noted that the direction of the square area is consistent with the main direction. Secondly, the Haar wavelets response in each sub-region are counted. The feature vectors of a single subregion consist of the sum of the horizontal response, the sum of the vertical response, the sum of the absolute value of the horizontal response, and the sum of the absolute of the vertical response. Since there are 16 sub-regions in total, the descriptor of a SURF feature point is a 64-dimensional vector.

After obtaining the coordinates and the descriptors of every point features in the two images, we can perform the optimal matching feature pairs by calculating the space distance between descriptors. A correct feature pairs should conform to the homography model:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \quad (14)$$

In Equation (14), $[u, v, 1]^T$ is the pixel coordinate in image A , and $[x, y, 1]^T$ is the corresponded coordinate in image B . Since the homography matrix is homogenous, it has a degree of freedom. To control the degree of freedom, a_{33} is assigned as 1, so the model is simplified as follows:

$$x = \frac{a_{11}u + a_{12}v + a_{13}}{a_{31}u + a_{32}v + 1} \quad (15)$$

$$y = \frac{a_{21}u + a_{22}v + a_{23}}{a_{31}u + a_{32}v + 1} \quad (16)$$

To estimate the parameters in the homography model, at least three pairs of point features are needed. In practice, hundreds of matched pairs are obtained. The random sample consensus (RANSAC) algorithm is always used to exclude erroneous pairs and find an optimal homography model.

After the homography model is obtained, the homography transformation is performed between image A and B , then A is stitched into B to obtain a primary processed result. Generally, there exists little misalignment at the junction of the two images. To improve the stitched result, we performed the post-processing to the primary result for seamlessness (Fig. 2). The pixels around the seams are reassigned by weighed average. As the coordinates move from image A to image B , the weight of image A changes linearly from 1 to 0, while the weight of image B changes from 1 to 0. The pixel value in the result image will transition smoothly from image A to image B , without the unnatural seams at the junction.

2.2.2 Stitching based on template recognition

Compared with the method based on SURF matching, the template recognition method requires some conspicuous template images in the scene. In practical terms, we attached some reflective strips on cable railings on the mining face. Although it is less flexible, this method can improve the stability by increasing the number of templates in the scene. As long as the relative positions and orientations of the templates in the image do not change significantly, the templates can be recognized in a changeable environment. In this case, the template recognition method is superior to the SURF matching method.

The template recognition method is realized by finding the template image T in the source

image I . Suppose the sizes of I and T are $W \times H$ and $w \times h$, respectively, for each feasible pixel coordinate (x, y) in I , we calculate the indicator function:

$$R = \frac{\sum_{i,j} T(i,j) \cdot I(x+i, y+j) - \frac{1}{w \cdot h} \sum_{i,j} T(i,j) \cdot \sum_{i,j} I(x+i, y+j)}{\sqrt{\sum_{i,j} T^2(i,j) - \frac{1}{w \cdot h} (\sum_{i,j} T(i,j))^2} \cdot \sqrt{\sum_{i,j} I^2(x+i, y+j) - \frac{1}{w \cdot h} (\sum_{i,j} I(x+i, y+j))^2}} \quad (17)$$

In Equation (17), $i = 0, 1, \dots, w, j = 0, 1, \dots, h, x = 0, 1, \dots, W - w + 1, y = 0, 1, \dots, H - h + 1$. The R function is the normalized correlation coefficient between the images T and I . We will find the position where the maximum value of R output, which is deemed the position of T in I . After the positions of all templates in the two original images A and B are obtained, the matched pairs are established, and then the homography model can be estimated by Equations (15) and (16). The purposes of template recognition and SURF matching are the same, which aim to estimate the homography model between two images. Afterwards, stitching and post-processing are performed to obtain the final output panorama (as shown in Fig. 2).

2.3 Vertical correction of oblique panorama

When the cameras are turned to three different directions to shoot for eliminating any missing scenes, the angles of view of the cameras are tilted, not vertically downward. So that a so-called perspective effect will deform the images, in which closer objects appear large and farther objects appear small. In this case, although the images shot by adjacent cameras are contiguous in the 3D space (Fig. 1(b)), the same objects will appear to be of different sizes in the two images. Therefore, it is necessary to eliminate the perspective effect in the image as much as possible, which can be done by correcting the oblique images to vertical images.

We consider the camera as a pinhole model, and the oblique image is colored by the orange solid line in Fig. 3. The blue solid line represents the corrected vertical image. After the focal length of the camera is estimated, the corresponded point of each pixel in the oblique image on the vertical image can be obtained.

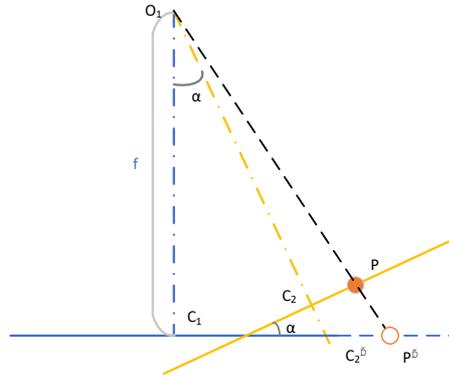


Fig. 3 Correction for oblique image

In Fig. 3, O_1 is the optical center of the camera, C_1 is the origin of the image plane coordinates of the vertical image, C_2 is the origin of the image plane coordinates of the tilted image, C_2' is the corresponded point of C_2 on the vertical image plane, and P is a certain pixel on the tilted image. Suppose the image space coordinates of P in the oblique image are $[x, y, f]^T$, then the image space coordinates of P' are projected onto the vertical image as follows:

$$\begin{bmatrix} x' \\ y' \\ f \end{bmatrix} = \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{bmatrix} \begin{bmatrix} x \\ y \\ f \end{bmatrix} \quad (18)$$

$$\lambda = \frac{f}{y \sin \alpha + f \cos \alpha} \quad (19)$$

In Equations (18) and (19), f represents the focal length of the camera, and α is the angle between the oblique image and the horizontal line (Fig. 3). Through the above transformation, we can reduce the perspective effect, so that the adjacent stitched images both keep the objects the same sizes at contiguous areas.

3 Experiments and results

Practical experiments were carried out on the No. 3301 fully mechanized face of Guotun Coal Mine in Shandong, China. Twenty-three high-resolution cameras were deployed at the fully mechanized mining face to obtain all the images and complete the program testing. In the absence of effective marking points in the underground mining face, we added red and white reflective strips on the cable railings behind the conveyor as artificial signs to enable the method of combining SURF matching and template recognition.

3.1 Framework of the system

To enhance the practicality of our research, we developed a prototype system to operate on the No. 3301 mining face. Considering that multiple cameras work simultaneously, a configurable distributed client program was designed to control multiple cameras for changing their shooting angle, stitching the images, and reducing the perspective effect. In this system, multiple client programs send their corresponded corrected panoramas and necessary information to the message queue managed by RabbitMQ. Subsequently, the stitching program takes multiple images from the same message queue, arranges them in the order of ID as the keyword, and uses the relevant information to process the adjacent panoramic images. The stitching program sends the final results to another message queue linked to a web server for display, which presents large panoramic images in the browser (shown in Fig. 4).

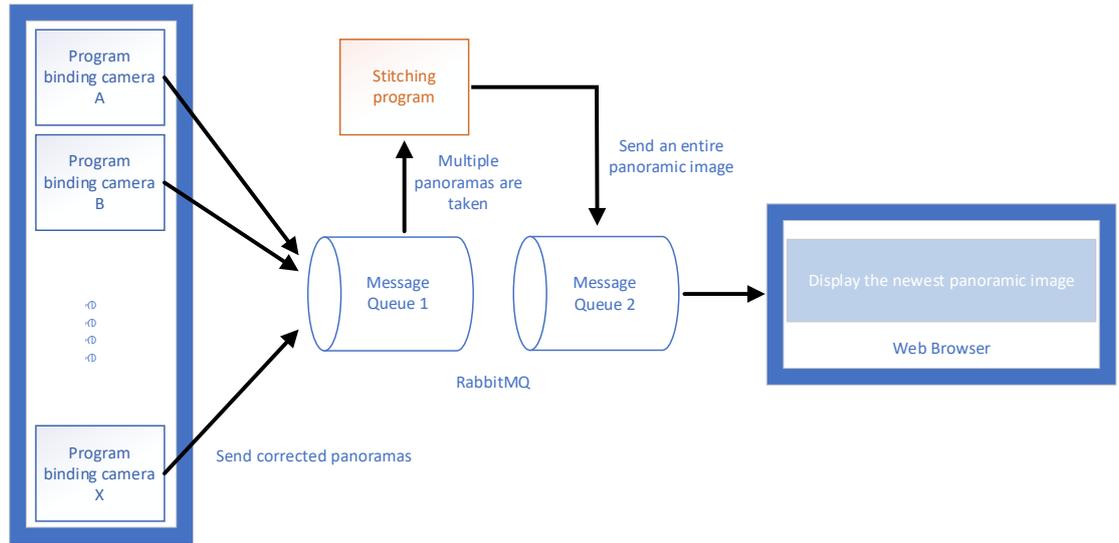


Fig. 4 Three parts of the system framework

RabbitMQ enables a non-blocking communication between multiple camera programs and stitching programs, while ensuring a complete and real-time image sequence. The users don't need to pass through fussy installation or configuration procedures of the system. They just connect to the server, and a complete panorama will be presented in the web browser.

3.2 Results of image enhancement

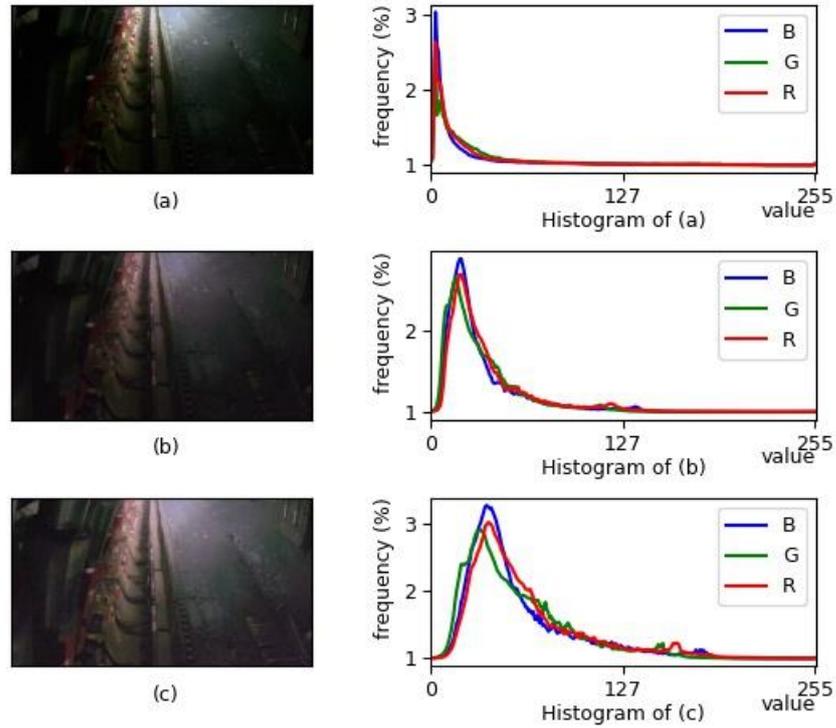


Fig. 5 (a): Wide-range 0; (b): Wide-range 50%; (c): Wide-range 100%

Fig. 5 shows that as the wide dynamic intensity increases, the histogram of the image gradually becomes uniform, and the overall visual effect of the image is gradually improved. The information entropy is used to measure the information contained in the image:

$$H = - \sum_{i=0}^{L-1} P_i \log_2(P_i) \quad (20)$$

In Equation (20), L represents the gray level of the image, which equals 256 in general. P_i represents the frequency of the i -th gray level in the image. Using this equation, we calculated the image entropy respectively of Figs. 6(a), (b), and (c) in the RGB channel. Table 1 lists the results. As the WDR intensity increases, the entropy of the image gradually increases, consistent with the results shown in Fig. 5.

Table 1 Comparison between information entropy of the images shown in Fig. 5

#	WDR intensity	Entropy		
		R	G	B
5(a)	0	6.034	6.256	5.790
5(b)	0.5	6.219	6.154	6.087
5(c)	1	6.862	6.817	6.747

The images taken by different cameras may differ a lot in aspect of the visual effects and the distribution of their histograms due to different camera parameter configurations. We

adopted the method of histogram matching to obtain images that appear as similar as possible.

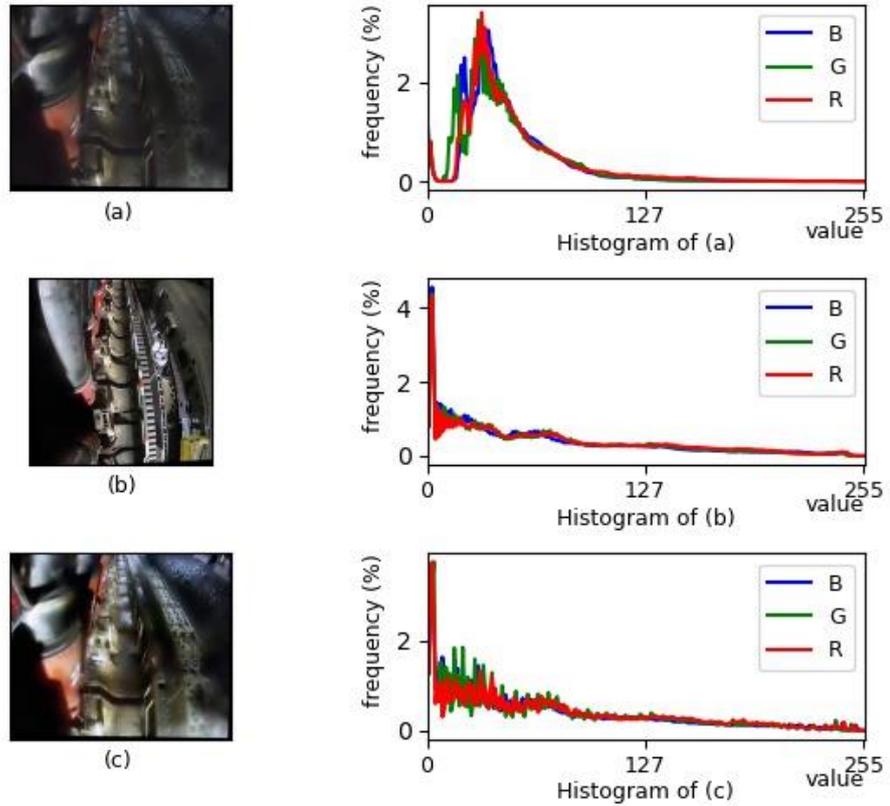


Fig. 6 (a): Origin image; (b): Target image; (c): Image after matching

Comparing Figs. 6(a) and (c), we find that the visual effect of the latter is similar to that shown in Fig. 6(b), with a better illumination. Comparing the histograms of the three images, we find that the histogram matching method does play a role: the histogram distribution shown in Fig. 6(c) is more similar to Fig. 6(b) than to Fig. 6(a). Through the histogram matching, the overall image would be more coordinated in the subsequent stitching procedure.

3.3 Results of stitching images taken from a single camera

The resolution of all the cameras was set to 1920×1080 , and the camera control interfaces were called remotely, so that we could obtain three images continuously with overlapped areas. After stitching three images, a single panoramic image was obtained via the stitching method of combination of the SURF matching and the template recognition.

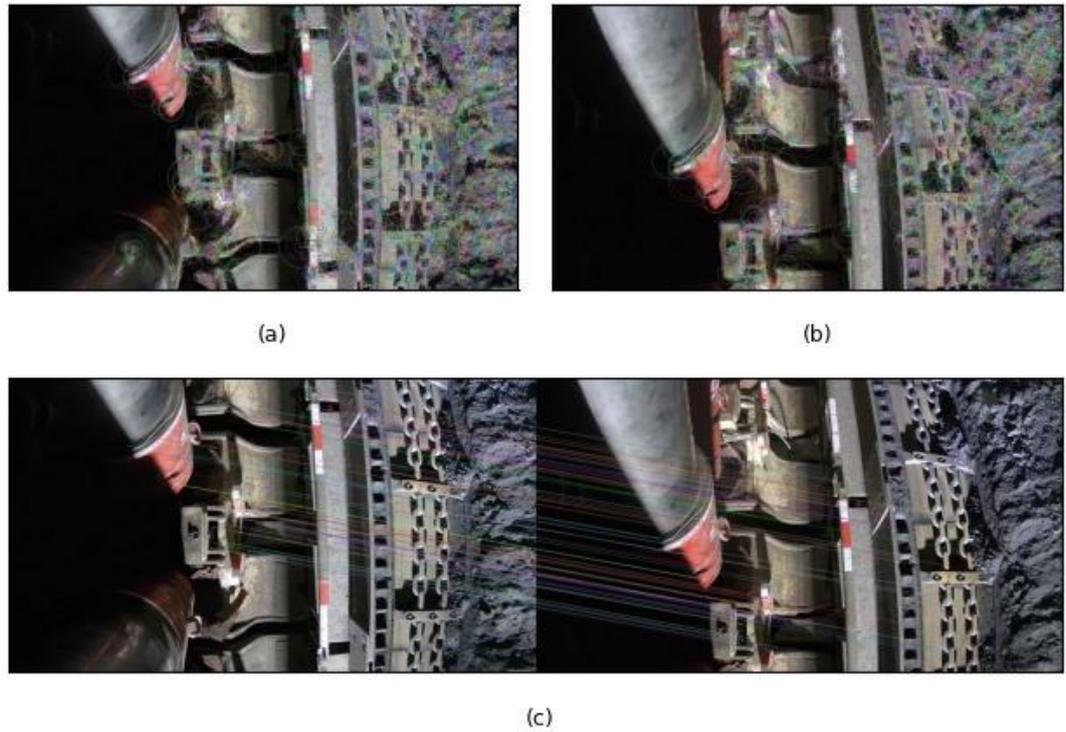


Fig. 7 (a): First origin image and its point features; (b): Second origin image and its point features; (c): The result of SURF matching

The two original images, shown in Figs. 7(a) and (b), have a certain overlapped area, where many corresponded SURF pairs exist. Fig. 7(c) shows the well-matched feature pairs. In Figure 7(c) the connection lines between the matched pairs are oriented toward a same direction, which means that the error of the estimated homography model is relatively small so that the subsequent stitching can be more accurate.

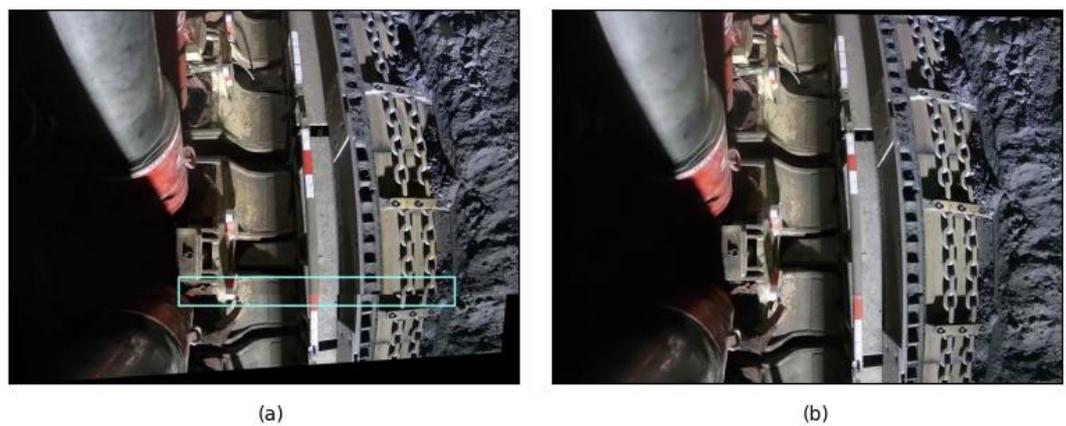


Fig. 8 (a): Primary processed image after homography transformation; (b): image after seamless operation

A coarse and larger result of stitching the images in Figs. 7(a) & (b) covers the ranges of the smaller two. In the region outlined by a cyan rectangle in Fig. 8(a), there exists conspicuous seams. After the seamless procedure, the unnatural seams are hardly visible so that an optimization of the result image is realized (shown in Fig. 8(b)).

The SURF matching method could fail because of the chaos and some unpredictable

changes in the scene, such as during the movement of personnel or coal equipment. In this case, the template recognition method can be an alternative to accomplish the stitching procedure.

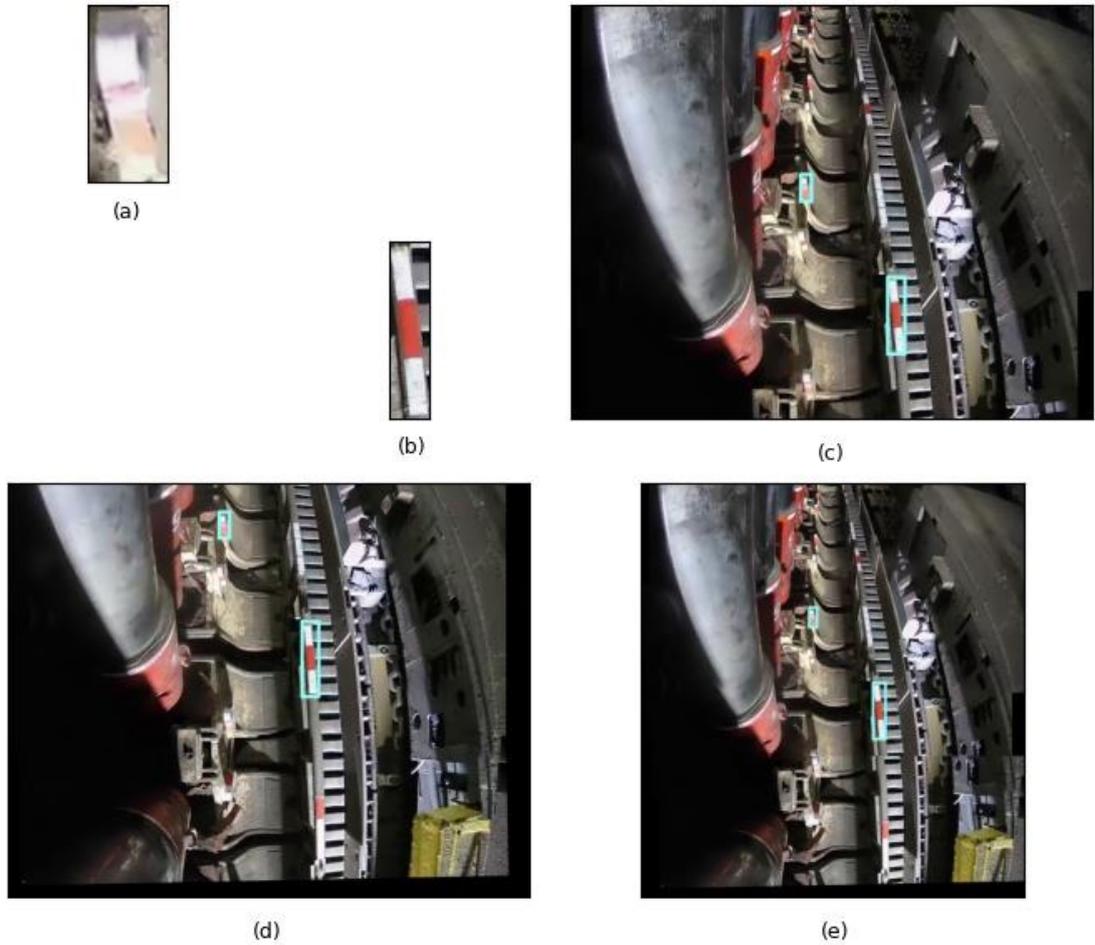


Fig. 9 (a) & (b): Template images; (c) & (d): Images including templates;

(e): The stitching result of the images (c) and (d).

In Figs. 9(c) & (d), the locations of the template images are marked with cyan rectangles. In most cases where the SURF matching fails, the alternative method can accurately locate the templates, so as to perform the subsequent stitching procedure. Fig. 9(e) shows the result of stitching the Figs. 9(c) & (d) via the template recognition method.

3.4 Results of vertical correction for oblique panoramic images

The panoramic image from the three images is deformed by the perspective effect. The distance between the cable trough and the hydraulic support bases decreases gradually along the extension direction, forming a trapezoidal shape (Figs. 10(g)&(h)).

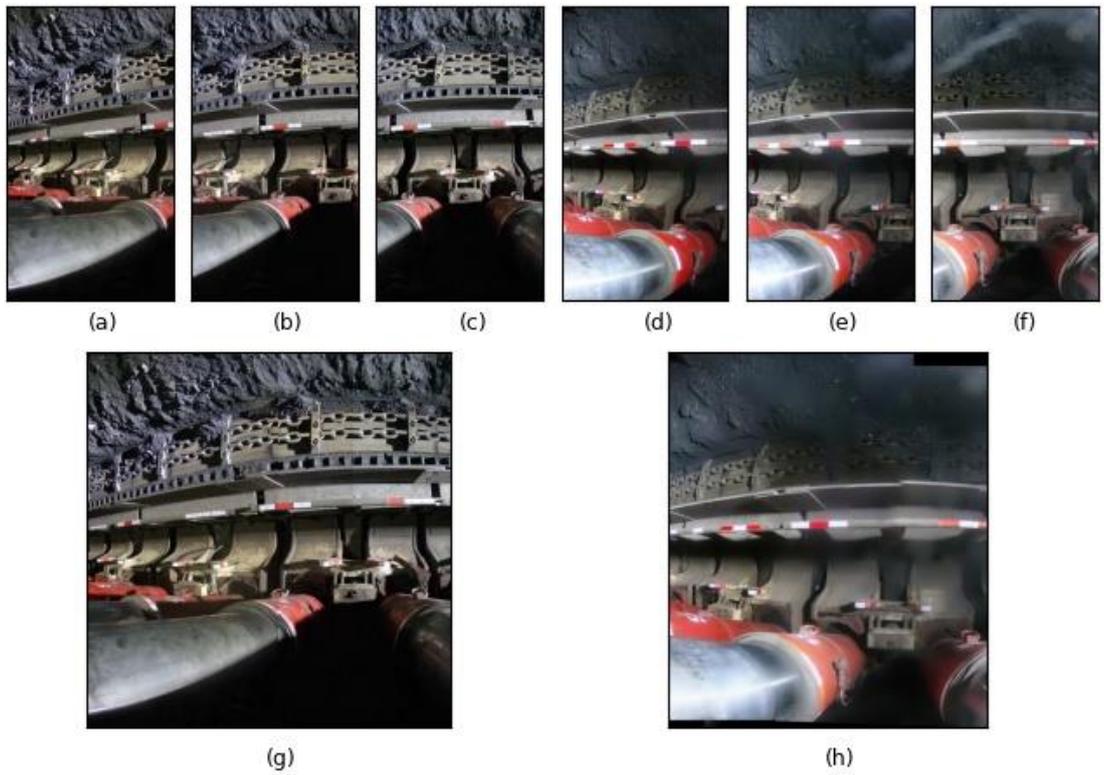


Fig. 10 (a)–(c): Three images from continuous capture; (d)–(f): Three images from continuous capture adjacently; (g): Panorama of the images (a)–(c); (h) Panorama of the images (d)–(f)

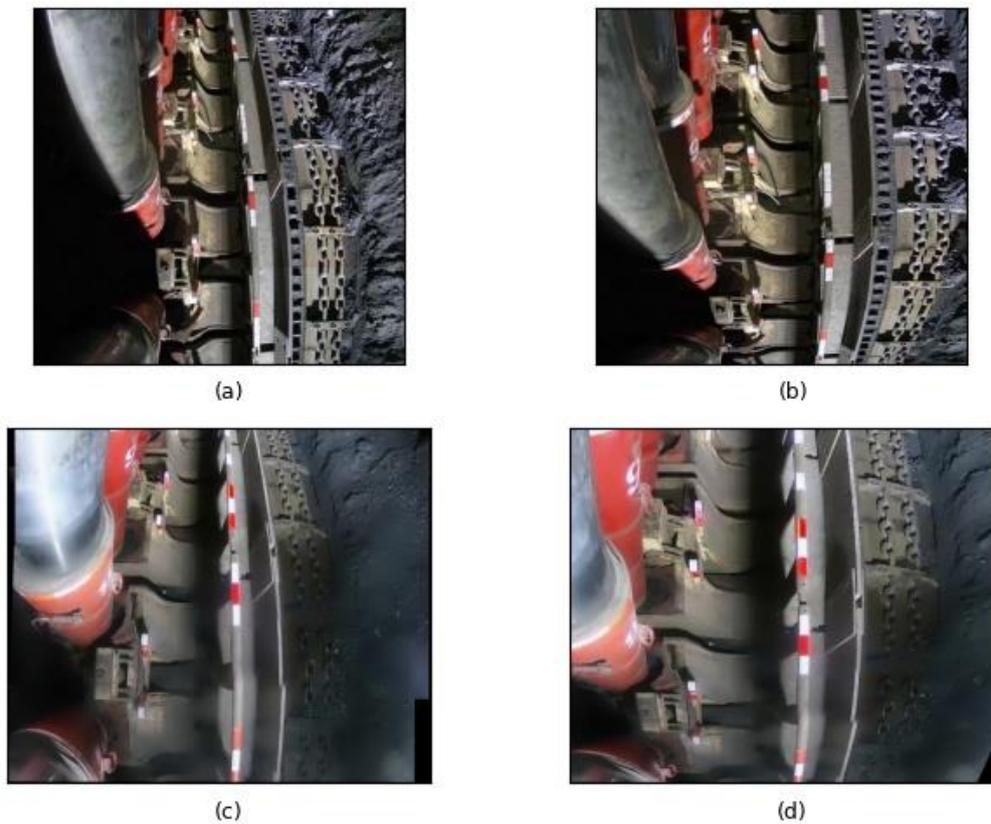


Fig. 11 (a) & (c): panoramic image; (b) & (d): panorama after correction

With the increasing of the distance, the widths of the hydraulic support bases gradually

decrease in the uncorrected images (shown in Fig. 11(a) & (c)), so the images cannot be directly concatenated. After the correction, the result images are displayed as if they are taken from vertical angle (Fig. 11(b) & (d)). In order to concatenate the two contiguous images, we take one of the images as the reference, stretching the other image to keep the width of hydraulic support bases at the same. Finally, the two images are concatenated (Fig. 12).

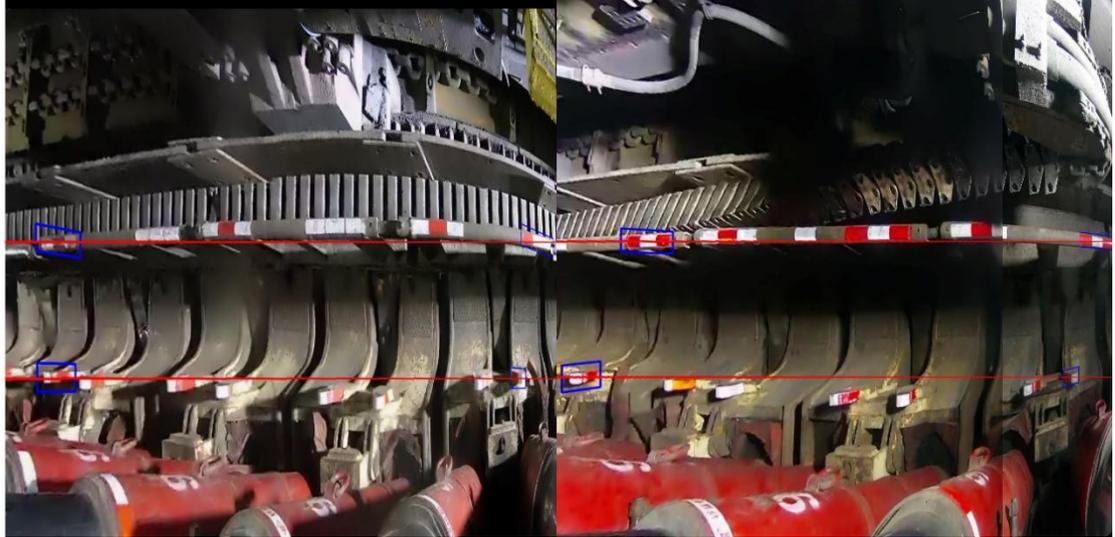


Fig. 12 Result of stitching two panoramic images

Figure 12 shows the final result image derived from two adjacent cameras after a series of operations. Two distributed client programs were used to create such a panoramic image, which took around 3 to 4 seconds. The output image contains at least six sets of hydraulic supports, which is nearly six times the size of the image captured by a single camera, in which we only see one set of supports. The prototype system used the distributed computing and the message queue technologies. In the practical mining environment, the low-speed movement of the shearer could be reflected in the result images. The output panoramic images would be updated continuously, thus providing surveillants with a real-time, large-scale and the most intuitive understanding to the mining environment.

4 Conclusions

Aiming to improve the usability of the video surveillance systems deployed at fully mechanized mining faces, this paper purposed some methods for solving salient problems related to the actual mining environment:

(1) To eliminate the conspicuous differences between images due to non-uniform illumination, the WDR was applied to solve both the overexposure and underexposure problems. The increasing of the information entropy proved our method effective. The differences between the images shot by different cameras was decreased by histogram matching method.

(2) To complete the missing scene of the fully mechanized mining face due to the limited FOV of the camera, a stable stitching method realized by combining SURF matching and template recognition. The range of the result panoramic image was larger than that of any image shot from a single angle of a single camera.

(3) To solve the image obliquity problem after stitching, we used the focal length and the pose of cameras, to eliminate the perspective effect of the single panoramic image. Consequently, the single-panoramic images from the adjacent cameras remain sizes of objects the same. Subsequently the corrected images are concatenated to a non-perspective image with a wider range. The final panoramic images provide an intuitive and dynamic situation of mining

for surveillants.

(4) To prove the effectiveness of the methods we purposed, a prototype system was developed which adopted a distributed configuration to achieve simultaneous correction and stitching of the images. Practical experiments showed that the prototype system was highly robust and usable in actual mining situation.

Acknowledgments The Project was supported by National Key Research and Development Project (2017YFC0804303)

References

- Bai, Z., Li, Y., Chen, X., Yi, T., Wei, W., Wozniak, M., & Damasevicius, R. (2020). Real-Time Video Stitching for Mine Surveillance Using a Hybrid Image Registration Method. *Electronics (Basel)*, 9(9), 1336. doi:10.3390/electronics9091336
- Bay, H., Ess, A., Tuytelaars, T., & Van Gool, L. (2008). Speeded-Up Robust Features (SURF). *Computer vision and image understanding*, 110(3), 346-359. doi:10.1016/j.cviu.2007.09.014
- De-yong, W., & Ze-xun, G. (2019). *Enhancement Algorithms for Images in Coal mines base on Dark Light Denoising Networks*.
- Dong, L., Xu, X., & Xu, Y. (2019). *Multi-point Laser Automatic Oblique Angle Shooting Correction System Based on Image Processing*.
- Dongmei, W., & Siqi, Z. (2019). *Research on image enhancement algorithm of coal mine dust*. Paper presented at the Proceedings - 2018 International Conference on Sensor Networks and Signal Processing, SNSP 2018.
- González-García, J., Swenson, R. L., & Gómez-Espinosa, A. (2020). Real-time kinematics applied at unmanned aerial vehicles positioning for orthophotography in precision agriculture. *Computers and electronics in agriculture*, 177, 105695. doi:10.1016/j.compag.2020.105695
- Habib, A. F., Kim, E.-M., & Kim, C.-J. (2007). New methodologies for true orthophoto generation. *Photogrammetric engineering and remote sensing*, 73(1), 25-36. doi:10.14358/PERS.73.1.25
- Hua, G., & Jiang, D.-H. (2014). A New Method of Image Denoising for Underground Coal Mine Based on the Visual Characteristics. *Journal of applied mathematics*, 2014, 1-7. doi:10.1155/2014/362716
- König, J., & O'Leary, P. (2020). Hyper resolution image mosaics with unbounded vertical field of view. *Computers in industry*, 122, 103281. doi:10.1016/j.compind.2020.103281
- Li, L., Wang, R., Lin, J., Xiao, Z., & Hui, Y. (2019). A novel approach for extraction of ripple mark parameters based on SfM. *Sedimentary geology*, 392, 105523. doi:10.1016/j.sedgeo.2019.105523
- Liu, S. B., Wang, J. H., Yuan, R. Y., Zhao, W. X., Li, L., & Wang, Q. H. (2020). Real-time and ultrahigh accuracy image synthesis algorithm for full field of view imaging system. *Scientific reports*, 10(1), 1-12. doi:10.1038/s41598-020-69353-9
- Qu, Z., Li, J., Bao, K., & Si, Z. (2020). An Unordered Image Stitching Method Based on Binary Tree and Estimated Overlapping Area. *IEEE Transactions on Image Processing*, 29, 6734-6744. doi:10.1109/TIP.2020.2993134
- Şalap, S., Karslıoğlu, M. O., & Demirel, N. (2009). Development of a GIS-based monitoring and management system for underground coal mining safety. *International journal of coal geology*,

80(2), 105-112. doi:10.1016/j.coal.2009.08.008

Shum, H.-Y., & Szeliski, R. (2000). Systems and Experiment Paper: Construction of Panoramic Image Mosaics with Global and Local Alignment. *International Journal of Computer Vision*, 36(2), 101-130. doi:10.1023/A:1008195814169

Vishwakarma, A., & Bhuyan, M. K. (2020). Image mosaicking using improved auto-sorting algorithm and local difference-based harris features. *Multimedia tools and applications*, 79(31-32), 23599-23616. doi:10.1007/s11042-020-09124-w

Vujic, S., Miljanovic, I., & Petrovski, A. (2008). A concept of the establishing the computer supported information-management system at the "Drmno" open pit mine, Coal Basin Kostolac. *Journal of Mining Science*, 44(3), 312-319. doi:10.1007/s10913-008-0022-1

Wang, C., Liu, X., Zhao, X., & Wang, Y. (2016). An effective correction method for seriously oblique remote sensing images based on multi-view simulation and a piecewise model. *Sensors (Basel, Switzerland)*, 16(10), 1725. doi:10.3390/s16101725

Wang, C., Zhang, Y., Wu, Y., & Gu, Y. (2011). Highly accurate geometric correction for seriously oblique aero remote sensing image based on the piecewise polynomial model. *Journal of computational information systems*, 7(2), 342-349.

Yanqin, Z. (2013). Multi-level denoising and enhancement method based on wavelet transform for mine monitoring. *International Journal of Mining Science and Technology*, 23(1), 163-166. doi:10.1016/j.ijmst.2013.03.008

Yoon, J., & Lee, D. (2018). Real-time video stitching using camera path estimation and homography refinement. *Symmetry (Basel)*, 10(1), 4. doi:10.3390/sym10010004

Yu, C., Rui, G., & Li-Jie, D. (2016). *Study of Image Enhancement Algorithms in Coal Mine*.

Yu, C., Xiao-Long, L., & Li-Jie, D. (2016). *An Enhancement Method for Non-uniform Illumination in Coal Mine*.

Yuehua, H., & Weiqiang, F. (2019). Pace recognition method under complex light conditions in coal mines. *Ji guang yu guang dian zi jin zhan*, 56(1). doi:10.3788/LOP56.011003

Zhao, Y., Wen, D., Gao, H., Wang, G., & Wang, Z. (2015). An integrated image clarity method for coal mine safety monitoring. *Journal of information and computational science*, 12(1), 59-64. doi:10.12733/jics20105077

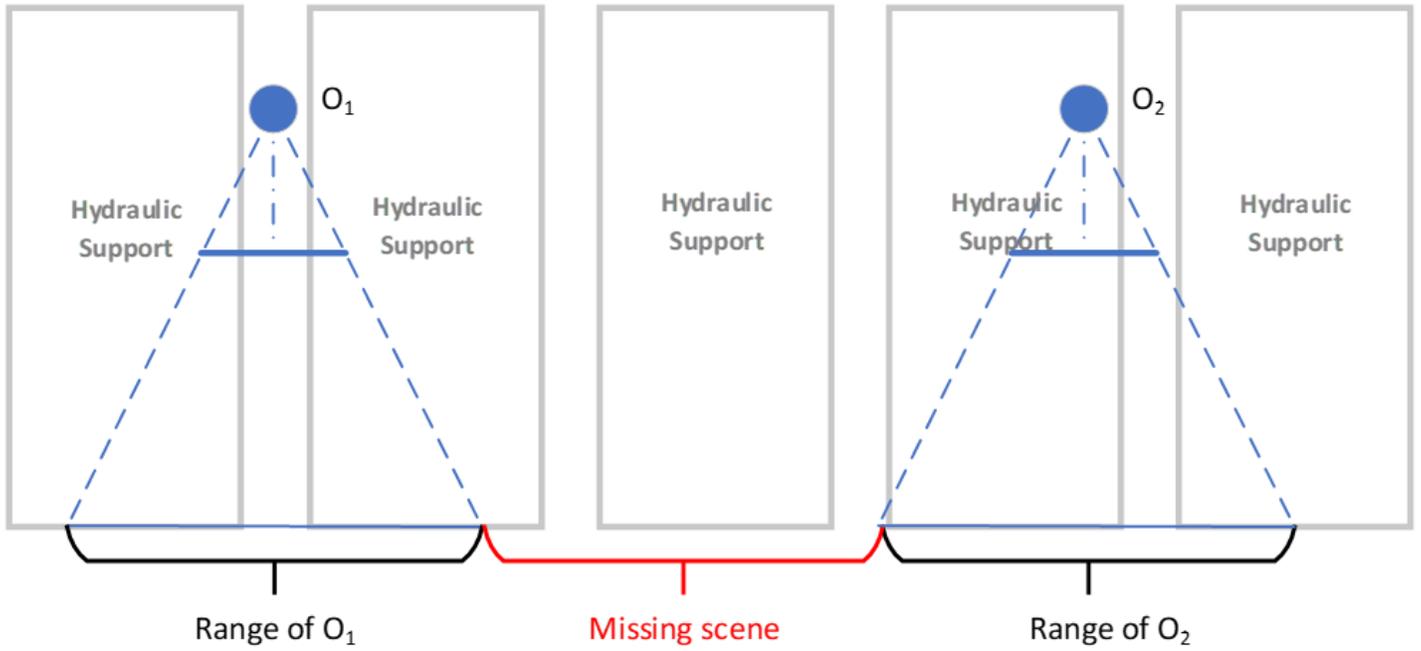
Zheng, S., Zhang, Y., Chen, S., Zhang, Z., Chen, F., Zhang, Z., . . . Wang, L. (2020). A preliminary study of dual-band confocal laser endomicroscopy combined with image mosaic in the diagnosis of liver cancer. *Nanomedicine*, 29, 102250. doi:10.1016/j.nano.2020.102250

Zhi, N., Mao, S., & Li, M. (2017). Enhancement algorithm based on illumination adjustment for non-uniform illuminance video images in coal mine. *Meitan Xuebao/Journal of the China Coal Society*, 42(8), 2190-2197. doi:10.13225/j.cnki.jccs.2017.0048

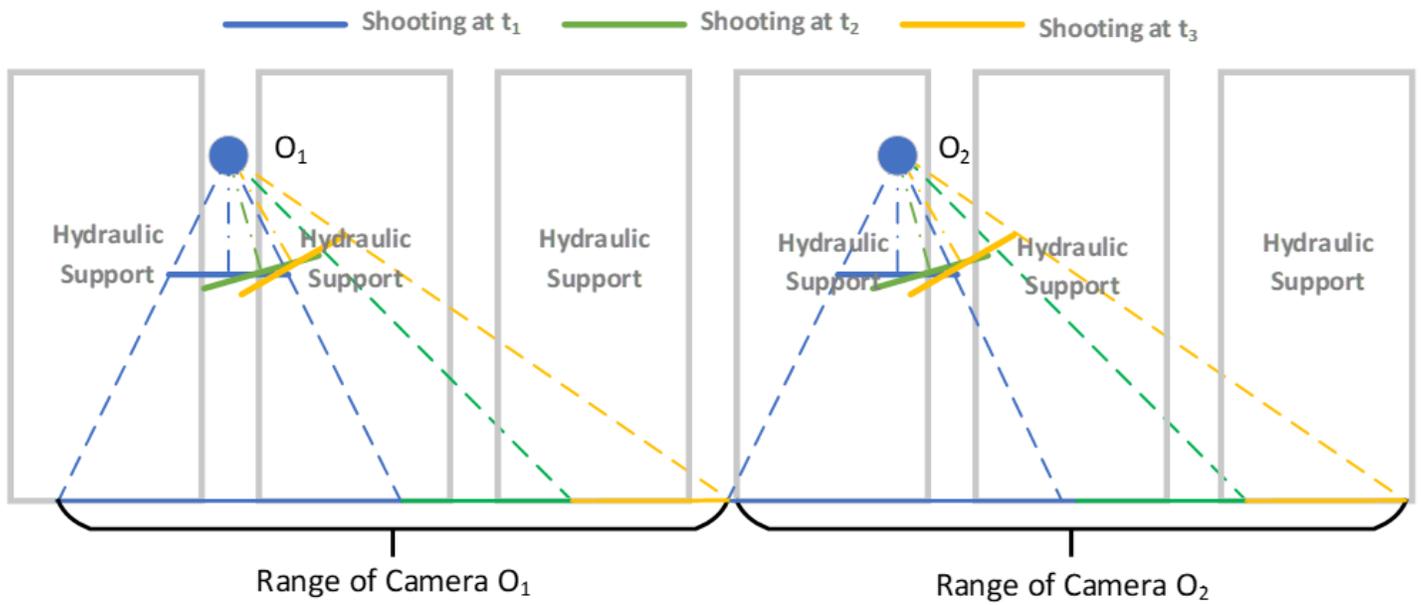
Zhi, N., Mao, S., & Li, M. (2018). An enhancement algorithm for coal mine low illumination images based on bi-gamma function. *Liaoning Gongcheng Jishu Daxue Xuebao (Ziran Kexue Ban)/Journal of Liaoning Technical University (Natural Science Edition)*, 37(1), 191-197. doi:10.11956/j.issn.1008-0562.2018.01.034

Zhou, W. (2011). *Design of video surveillance system based on 3G wireless network in underground coal mine*.

Figures



(a)



(b)

Figure 1

(a): Single FOV; (b): Triple FOV

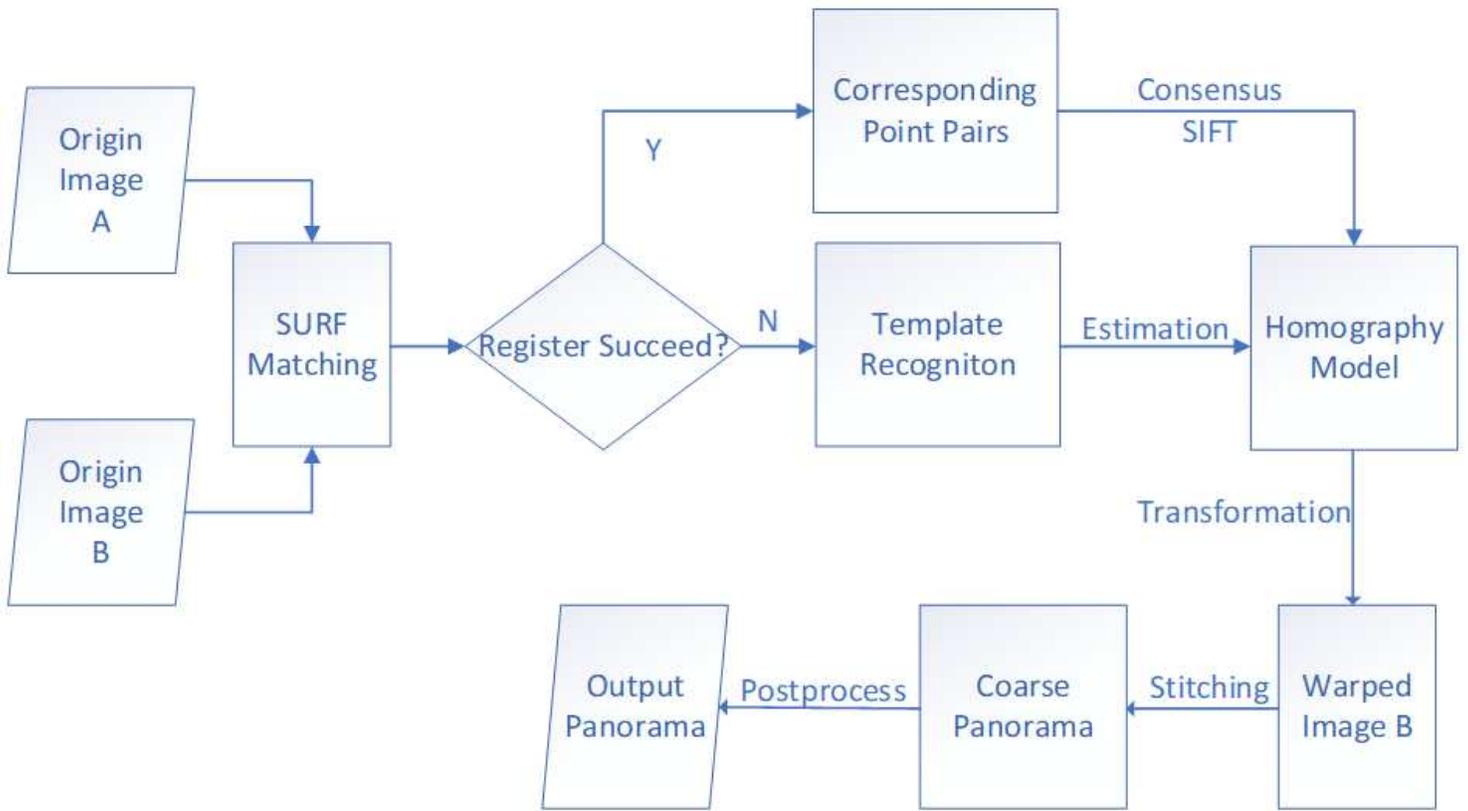


Figure 2

Pipeline of a more stable stitching method purposed by this article

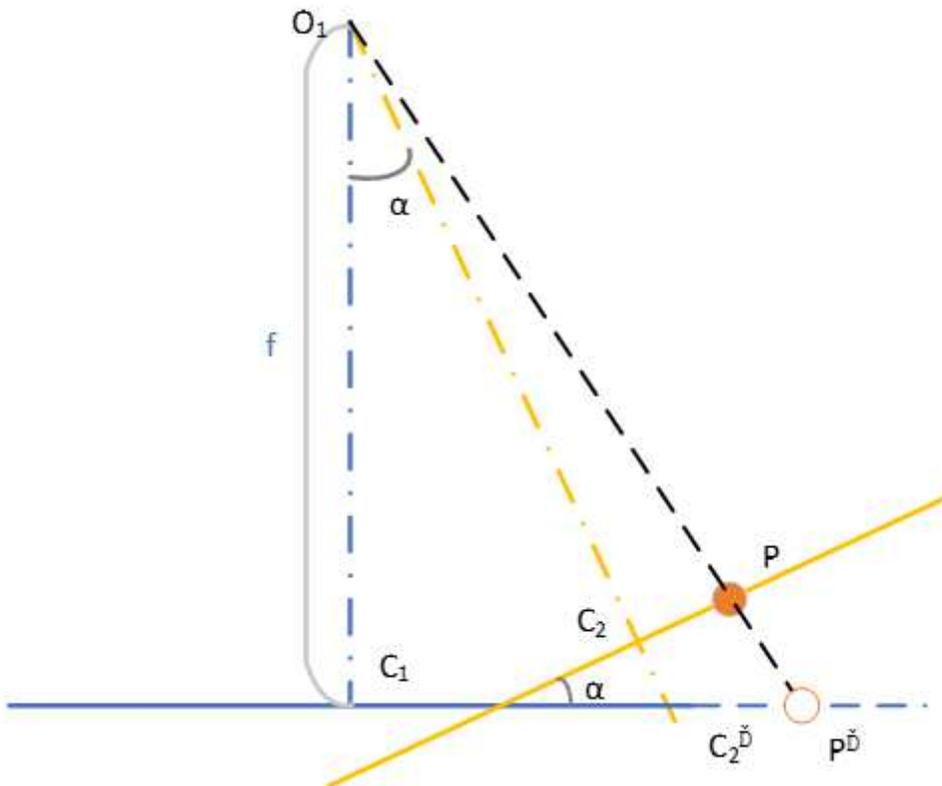


Figure 3

Correction for oblique image

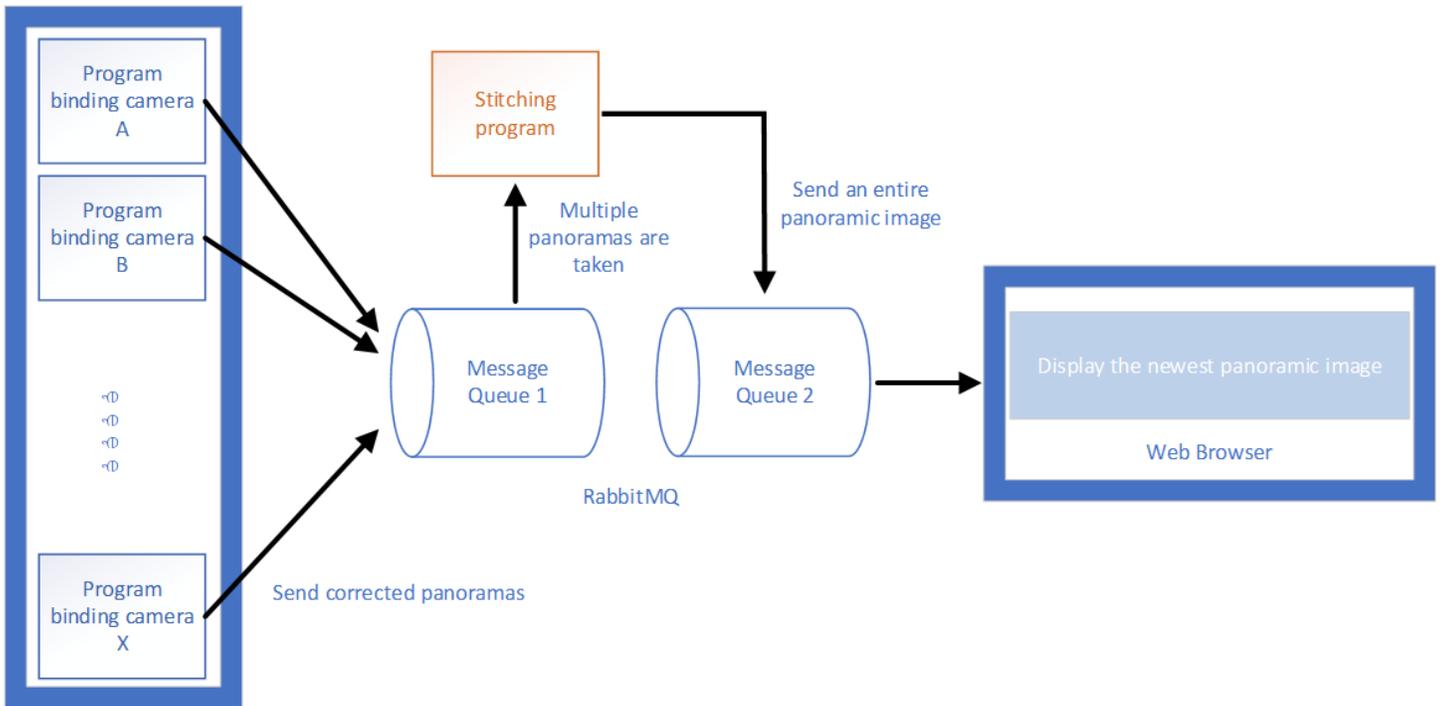
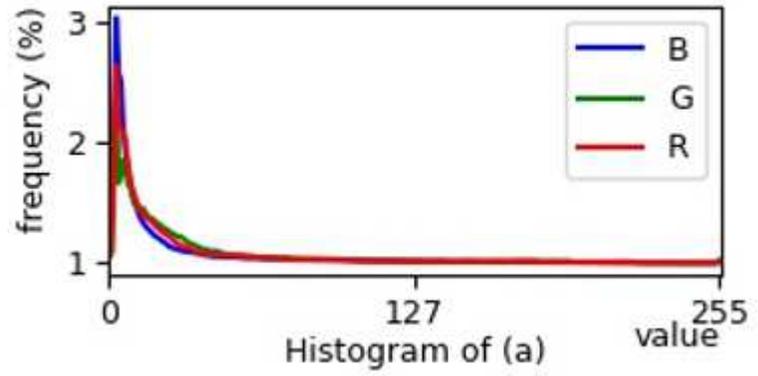


Figure 4

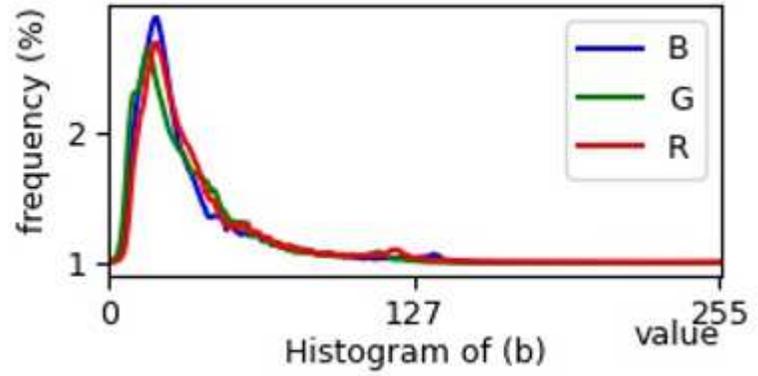
Three parts of the system framework



(a)



(b)



(c)

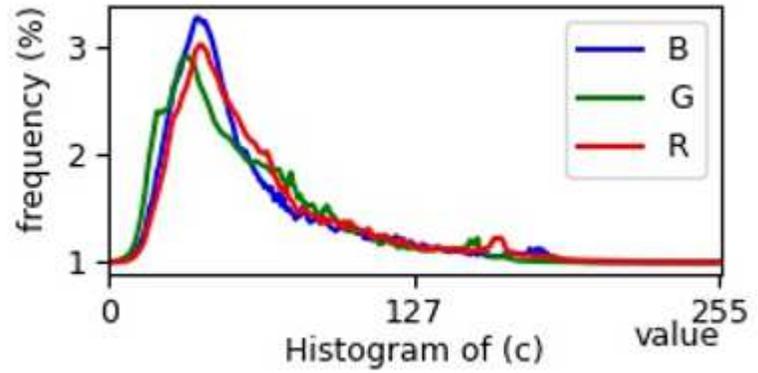
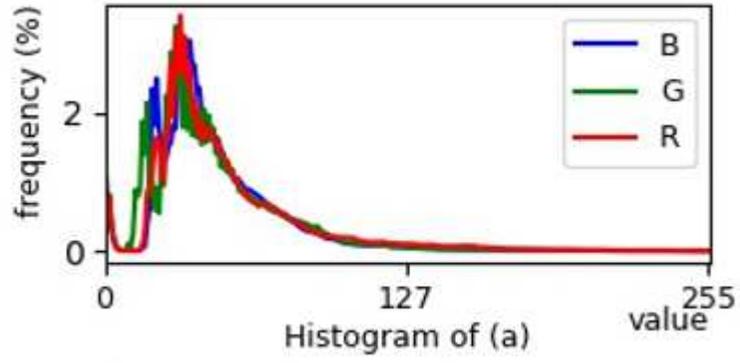


Figure 5

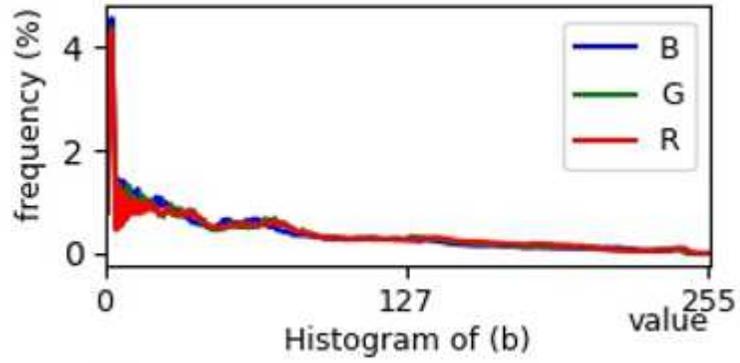
(a): Wide-range 0; (b): Wide-range 50%; (c): Wide-range 100%



(a)



(b)



(c)

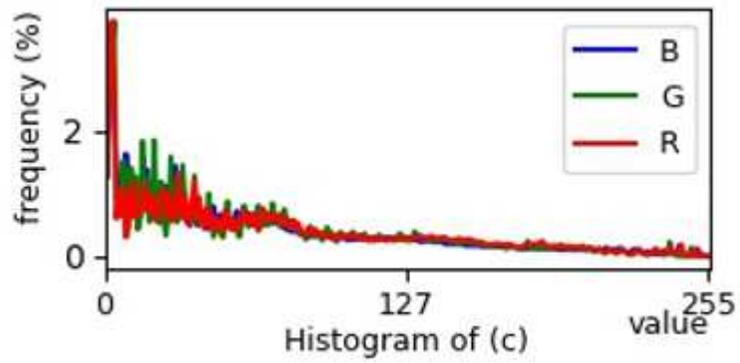


Figure 6

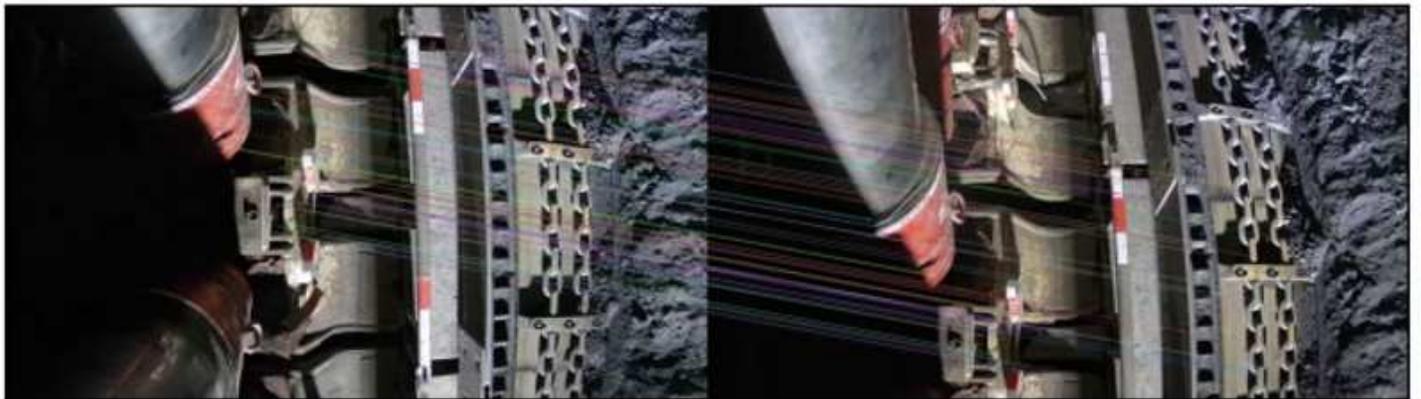
(a): Origin image; (b): Target image; (c): Image after matching



(a)



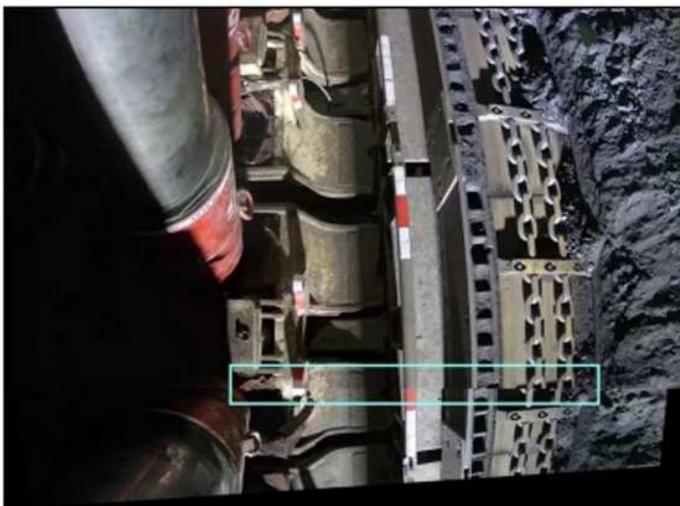
(b)



(c)

Figure 7

(a): First origin image and its point features; (b): Second origin image and its point features; (c): The result of SURF matching



(a)



(b)

Figure 8

(a): Primary processed image after homography transformation; (b): image after seamless operation



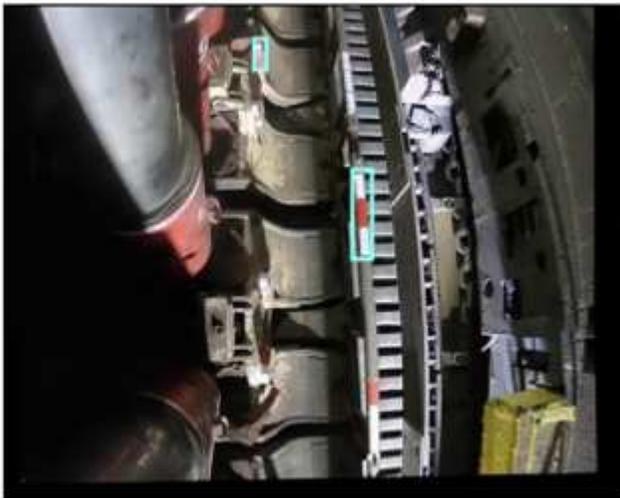
(a)



(b)



(c)



(d)



(e)

Figure 9

(a) & (b): Template images; (c) & (d): Images including templates; (e): The stitching result of the images (c) and (d).

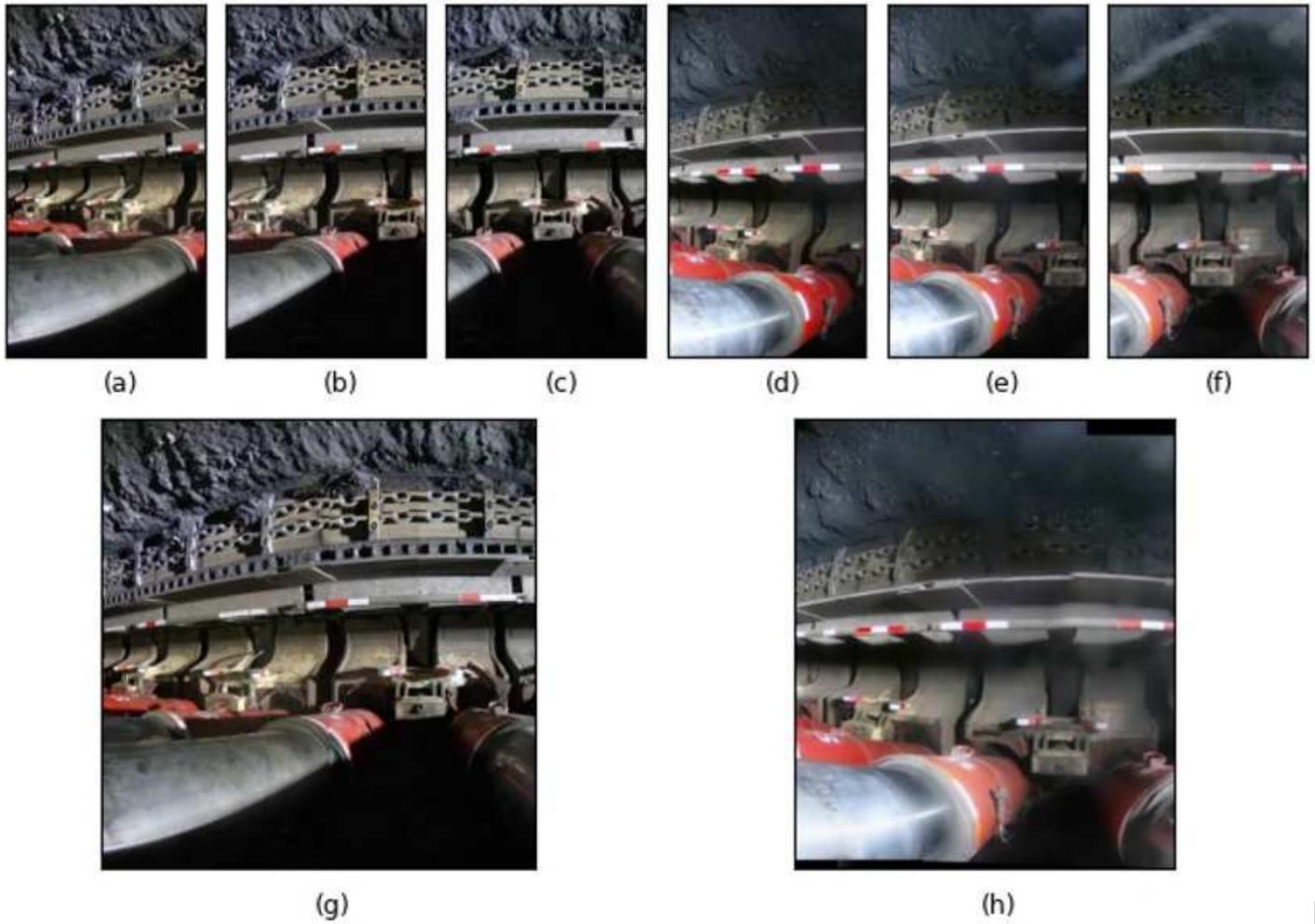


Figure 10

(a)–(c): Three images from continuous capture; (d)–(f): Three images from continuous capture adjacently; (g): Panorama of the images (a)–(c); (h) Panorama of the images (d)–(f)



(a)



(b)



(c)



(d)

Figure 11

(a) & (c): panoramic image; (b) & (d): panorama after correction



Figure 12

Result of stitching two panoramic images