

---

# **Geological logging of tunnel surrounding rock based on multi-view geometry and image stitching**

**Zhuang Xie<sup>1</sup>, Zhiheng Zhu<sup>1,\*</sup>, Jinyang Fu<sup>1,3</sup>, Junsheng Yang<sup>1</sup>, Bin Peng<sup>2</sup>**

*1. School of Civil Engineering, Central South University, Changsha 410075, China*

*2. Hunan Electric Power Design Institute Co., Ltd., Changsha 410007, China*

*3. MOE Key laboratory of Engineering Structures of Heavy Haul Railway,  
Changsha 410075, China*

*zzh8207@163.com*

---

**ABSTRACT.** *The geological map of tunnel surrounding rock is essential to the design of dynamic construction and the stability analysis and reinforcement of rock mass. Based on computer vision technology, this paper proposes a fast and flexible method for preparing geological maps of tunnel surrounding rock. By this method, the 3D point cloud of the target tunnel was reconstructed from multiple photos using the multi-view geometry principles; then, the orthographic projection model of the tunnel was determined from the spatial point cloud through 3D surface estimation; after that, each photo on surrounding rock was subjected to geometric correction based on the relative position between cameras and orthographic projection model; finally, the orthographic display maps of the chamber wall and tunnel face were obtained by stitching the corrected photos. The image processing software inspired by this method can automatically generate the geological map on the tunnel surrounding rock in each work cycle based on the set of photos shot freely from multiple angles. Through engineering application, it is proved that the proposed method outperforms the existing tunnel geological logging methods in terms of the flexibility and efficiency of field shooting, as well as the universality and intuitiveness of the automatically generated geological maps on tunnel surrounding rock. The research findings provide an intuitive reference for tunnel construction design and boast profound significance in engineering application.*

**RÉSUMÉ.** *La carte géologique des roches entourant le tunnel est essentielle à la conception d'une construction dynamique ainsi qu'à l'analyse de la stabilité et au renforcement de la masse rocheuse. Basé sur la technologie de vision par ordinateur, cet article propose une méthode rapide et flexible pour préparer des cartes géologiques des roches entourant le tunnel. Par cette méthode, le nuage de points 3D du tunnel cible a été reconstruit à partir de plusieurs photos en utilisant les principes de la géométrie multi-vues; ensuite, le modèle de projection orthographique du tunnel a été déterminé à partir du nuage de points spatial par estimation de surface en 3D; Après cela, chaque photo de la roche environnante a été*

soumise à une correction géométrique basée sur la position relative entre les caméras et le modèle de projection orthographique; enfin, les cartes d'affichage orthographique du mur de la chambre et de la face du tunnel ont été obtenues en assemblant les photos corrigées. Le logiciel de traitement d'images inspiré par cette méthode peut générer automatiquement la carte géologique sur la roche entourant le tunnel à chaque cycle de travail en fonction de l'ensemble de photos prises librement sous plusieurs angles. Grâce à des applications techniques, il est prouvé que la méthode proposée surpassé les méthodes de diagraphie géologique existantes en tunnels en termes de flexibilité et d'efficacité du tournage sur le terrain, ainsi que de l'universalité et de l'intuitivité des cartes géologiques générées automatiquement sur les roches entourant les tunnels. Les résultats de la recherche fournissent une référence intuitive pour la conception de la construction de tunnels et ont une signification profonde dans les applications d'ingénierie.

**KEYWORDS:** tunnel construction, computer vision, photographic geological logging.

**MOTS-CLÉS:** construction de tunnels, vision par ordinateur, diagraphie géologique photographique.

DOI:10.3166/ISI.23.5.43-59 © 2018 Lavoisier

## 1. Introduction

Geological logging, an important part of tunnel excavation, is essential to the design of dynamic construction and the stability analysis and reinforcement of rock mass. Currently, photographic geological logging stands out as the most popular way to record the geological information of the surrounding rock in the tunnel, because it is more efficient and reliable than manual logging and much less expensive than laser scanning logging (Song *et al.*, 2013; Wang *et al.*, 2007). The key to photographic geological logging lies in the construction of geological photomaps, especially those on tunnel walls and faces. As photos with control information, the geological photomaps are mostly shot based on the principle of close-range photogrammetry, laying the basis for subsequent mapping (Li *et al.*, 2014; Yang *et al.*, 2001).

The geological logging using ordinary digital cameras has formed a mature system after the development in the 2000s. For instance, Chen *et al.* (2005) put forward a distortion model and a parameter detection method for common digital cameras. Considering the requirements of geological logging system, Yang *et al.* (2004) proposed a calibration approach for common digital cameras. Yang *et al.* (2003) developed a strategy to measure internal elements and control external elements for cameras. Yan and Gao (2007) established a geological logging photography platform, which realizes the rapid field orientation of digital cameras. Li *et al.* (2004); Yang and Li (2002) explored the measurement and processing algorithms of images taken by ordinary digital cameras. Wang (2007) studied the geometric correction of geological photos on complex chambers, systematically examined the correction algorithm for the external elements of chamber photos, and presented the processing algorithms for the photomaps of chambers in various shapes. Leng (2009) investigated the extraction of geological information from a single, front digital image of tunnel faces, and constructed the TFGIS system. Zhou *et al.* (2008) developed a solution based on infrared images for the dust interference

in New Austrian Tunneling Method (NATM).

In recent years, stereo mapping techniques have been implemented widely in the extraction and logging of the geological information of the surrounding rock. Relying on the ShapeMetriX3D system, Wang *et al.* (2011) explored the rapid recognition of the joint information of tunnel faces using dual-image 3D modelling information. The ShapeMetriX3D system works by the principle of binocular stereo vision. Buyer and Schubert (2017) acquired 3D point clouds via close-range photogrammetry, and computed the spacing of the discontinuities of the rock mass using the ShapeMetriX3D. Li *et al.* (2017) and Zhang *et al.* (2016) carried out digital identification, integrity evaluation and quantitative system research of the surrounding rock structure, with the aid of the CAE Sirovision, which is based on the binocular stereography. However, control points are needed to control field shooting and image pairs are required to control image measurement.

Despite the marked progress, photo-based geological logging still faces many problems. (1) During field shooting, both object space control and image space control are subjected to high field requirements and multiple constraints, which seriously drag down the efficiency of field work and undermine the work quality. (2) During image processing, the geometric correction equations are not universally applicable: the equation should be determined according to the exact section shape and inclination of the tunnel, but the section shape may be highly variable and extremely complex (Wang, 2007). (3) Coupled with the measurement error of the external parameters, the uneven surface of the surrounding rock may lead to different degrees of orthographic projection errors during the geometric correction of photos taken from different angles; In this case, it is inevitable to have seams and misalignment on the display images stitched from the corrected images, which negatively affects the display effect and mapping results (Yan and Gao, 2007).

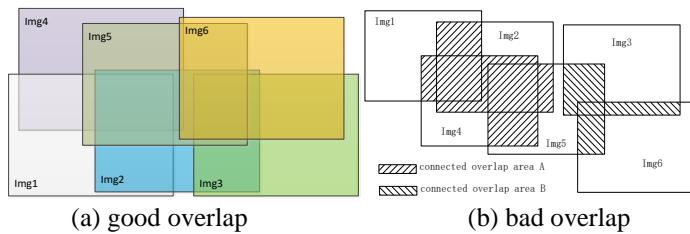
The advancement of computer vision technology has greatly promoted the development of structure from motion (SfM) (Agarwal *et al.*, 2010; Snavely *et al.*, 2007; Wu, 2013) and image stitching (Brown and Lowe, 2007; Szeliski, 2004). The SfM can determine the internal parameters (including distortion), external parameters and feature point positions of the camera by solving large-scale nonlinear optimization problem using multiple overlapping photos, while the image stitching manages to cover the target with a single photo and eliminate the occlusions of a single shot by stitching multiple photos into a large, seamless image. Currently, these two techniques have been applied in the visual inspection of structures like bridges and tunnels (Zhu *et al.*, 2010; Zhu *et al.*, 2016; Heymsfield and Kuss, 2013; Chaiyasarn *et al.*, 2015; Zhu *et al.*, 2014). In this paper, both the SfM and image stitching are adopted to improve the existing method for preparing the geological map of tunnel surrounding rock, aiming to achieve free and efficient field shooting, universal post-processing and good display effect.

## 2. Field shooting of tunnel excavation face

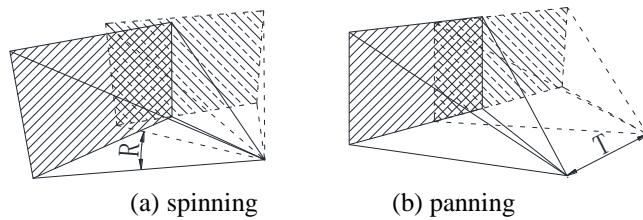
During field shooting, multiple photos on the target area should be taken with a

digital camera, provided that each three consecutive photos share a common overlapping area and the overlapping areas of all photos are fully connected (Figure 1a). Two adjacent photos should be overlapped by greater than 50%, so as to prevent the lack of overlap in Figure 1b and ensure the quality of subsequent 3D reconstruction. To shoot scenes in different positions, the camera should be moved to another place (Figure 2b) instead of being rotated at the original place (Figure 2a). The shooting distance between two photos is positively correlated with the angle of the triangle intersection and the accuracy of the calculated position of spatial points, because it is impossible to derive the spatial position of the corresponding points on two images solely based on the simple rotation transformation relationship in the later stage. To minimize the projection error of the display image, the shooting direction should be as vertical as possible to the exposed surface of the surrounding rock.

In addition, there is no need to do the following things in field shooting: setting a fixed camera station or object space control points, measuring the camera's orientation, determining the inner measurements of the camera in advance, or fixing the lens and focal length. Multiple cameras can be used simultaneously for field shooting. Furthermore, the adoption of image stitching lifts the distance limit in actual shooting, thus avoiding the interference of onsite dusts. Meanwhile, the upper part can be shot with the aid of a trolley to prevent occlusion (Figure 3).



*Figure 1. Schematic plot of pictures overlap pattern<sup>1</sup>*



*Figure 2. Two ways of taking overlapped photograph*

---

1. Overlap part in photos of group (a) are fully connected, and then you can rebuild all the photos if the overlapping region contains enough feature points. Photos of group (b) cannot be totally reconstructed because the overlap part was not fully connected. Reconstruction can only restore 4 photos with connected area A or 3 photos with connected area B.



Figure 3. The situation of photographing worksite

### 3. 3D point cloud reconstruction of tunnel excavation face

The 3D point cloud reconstruction involves the following steps: feature point extraction, image pair registration, corresponding point identification, structure from motion. Specifically, the feature points were extracted using the popular scale-invariant feature transform (SIFT) or speeded up robust features (SURF) operator (Lowe, 2004; Bay *et al.*, 2008). Then, the feature points were registered for each pair of images by approximate nearest neighbour search. After the initial registration, the registered relationship was refined to eliminate mismatch. By the standard 8-point method, the basic matrix between image pairs was calculated through the robust random sampling consensus (RANSAC) (Fischler and Bolles, 1981; Hartley and Zisserman, 2004), leaving the point pairs obeying the epipolar geometric relationship. Registered based on image pairs, the point pairs help to determine the corresponding points on multiple photos. In light of the corresponding points, a series of projection equations was determined by the projection relationship, producing the objective function shown in equation (1). Then, equation (1) was solved by nonlinear optimization, e.g. Levenberg-Marquart algorithm (Nocedal and Wright, 2006), to determine the coordinates and camera parameters of a set of spatial points that minimize the error between the projected value and the observed value of a spatial point on any photo.

$$f(\mathbf{K}^i, \mathbf{D}^i, \mathbf{R}^i, \mathbf{t}^i, \mathbf{X}_j) = \min_{\mathbf{K}^i, \mathbf{D}^i, \mathbf{R}^i, \mathbf{t}^i, \mathbf{P}_j} \sum_{i,j} \left\| \mathbf{x}_j^i - \hat{\mathbf{x}}_j^i \right\|^2 \quad (1)$$

where  $\mathbf{x}_j^i - \hat{\mathbf{x}}_j^i$  is the error between the calculated and actual coordinates of the projection of point  $j$  on photo  $i$ ;  $\mathbf{x}_j^i$  is the actual position of a feature point;  $\hat{\mathbf{x}}_j^i$  is the projected position of the corresponding spatial point, which can be calculated by the following formula (Szeliski, 2010):

$$\tilde{\mathbf{x}}_j^i = \mathbf{K}^i \cdot (\mathbf{R}^i | \mathbf{t}^i) \cdot \mathbf{X}_j \quad (2)$$

$$\hat{\mathbf{x}}_j^i = \text{Dist}_{\kappa_1, \kappa_2}(\tilde{\mathbf{x}}_j^i) = (1 + \kappa_1^i r^2 + \kappa_2^i r^4) \tilde{\mathbf{x}}_j^i \quad (3)$$

$$r = \|\tilde{\mathbf{x}}_j^i - \mathbf{C}^i\| \quad (4)$$

where  $\mathbf{x}_j = [x_j \ y_j \ z_j]$  is a spatial point;  $\mathbf{K}^i = \begin{bmatrix} f_x^i & 0 & c_x^i \\ 0 & f_y^i & c_y^i \\ 0 & 0 & 1 \end{bmatrix}$  are the internal parameters of the camera;  $\mathbf{D}^i = [\kappa_1^i \ \kappa_2^i]$  are the radial distortion parameters of the camera;  $(\mathbf{R}^i \ | \ \mathbf{t}^i)$  are the external parameters of the camera (rotation and translation);  $\mathbf{C}^i = [c_x^i \ c_y^i]$  are the coordinates of the camera's principal point.

For the convergence of the objective function to the global optimum, a good initial value should be identified to solve equation (1) by nonlinear optimization. One of the best identification methods is incremental SfM. By this method, a set of image pairs of a wide baseline was determined to recover the camera parameters and the positions of the observable spatial points; then, the other cameras capable of observing the reconstructed points were added, and their parameters were recovered one by one. The above procedure was repeated till no other camera could be added (Hartley and Zisserman, 2004).

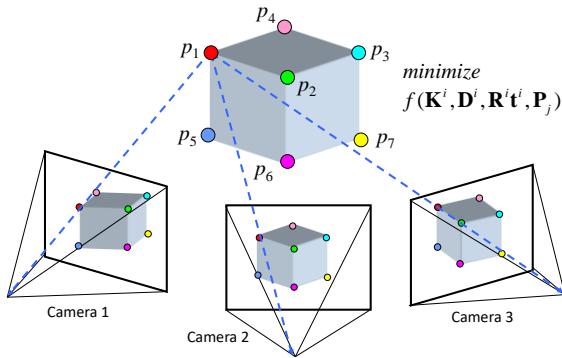


Figure 4. The principle of 3D reconstruction (Snavely et al., 2006)

#### 4. Construction and identification of orthographic projection model of tunnel excavation face

The relationship between the shooting position and the tunnel projection model can be ascertained by constructing a projection model and identifying its position in point cloud of the reconstructed scene. Using this relationship, it is possible to correct a single image by orthographic projection under the following two conditions: First, the model recognition technology must be scale independent, for the point cloud reconstructed by the SfM is a reduced scale simulation of the real situation; Second, the recognition algorithm should be robust against the noises of the projection model caused by the unevenness of the tunnel excavation face.

Inspired by Simon Winkelbach *et al.* (2006), Papazov and Burschka (2011), this paper designs a rapid matching method for similar surfaces based on scale-independent robust surface characteristic operator, develops the RANSAC recognition algorithm from the designed method, and realizes the rapid and robust recognition of the projection model in the reconstructed point cloud.

#### 4.1. Construction of complex orthographic projection model

The sweep method, a popular approach for 3D free-form surface modelling, was adopted to build the surfaces of the projection model. This projection method is universally applicable and adapt to various complex section shapes and inclinations. The specific steps are as follows. First, the designed section information of the chamber was treated as the streamline equation  $S(\eta)$ , while the vertical alignment was treated as the trace equation  $L(\xi)$ :

$$L(\xi) = \begin{cases} L_x(\xi) \\ L_y(\xi) \\ L_z(\xi) \end{cases}, S(\eta) = \begin{cases} S_x(\eta) \\ S_y(\eta) \\ S_z(\eta) \equiv 0 \end{cases} \quad (5)$$

Then, the surface points ( $x$ ,  $y$ ,  $z$ ) constructed by sweeping can be expressed in expanded plane coordinates:

$$f(\xi, \eta) = \begin{cases} f_x(\xi, \eta) = L_x(\xi) + \cos[\arctan \frac{dL_z(\xi)}{dL_x(\xi)}] \cdot S_x(\eta) \\ f_y(\xi, \eta) = L_y(\xi) + S_y(\eta) \\ f_z(\xi, \eta) = L_z(\xi) + \sin[\arctan \frac{dL_z(\xi)}{dL_x(\xi)}] \cdot S_x(\eta) \end{cases} \quad (6)$$

where  $\eta$  and  $\xi$  are the coordinates of the orthographic projection plane; the streamline equation  $S(\eta)$  describes the section shape; the trace equation  $L(\xi)$  describes the horizontal and vertical alignments of the tunnel. On this basis, the author constructed various complex orthographic projection models for the tunnel (Figure 5). The projection plane of the tunnel face can be built directly with two straight lines.

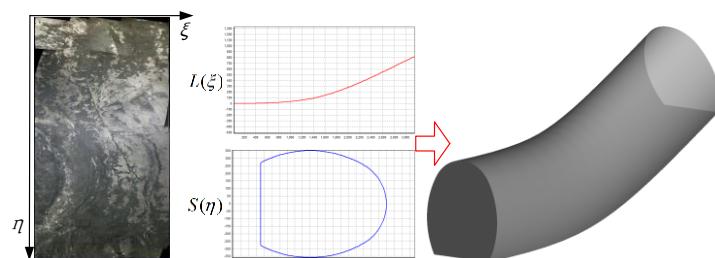


Figure 5. Construction method of projection model for orthographic mapping

#### 4.2. Rapid matching of similar surfaces

For any two directed points  $\mathbf{u} = [\mathbf{p}_u, \mathbf{n}_u], \mathbf{v} = [\mathbf{p}_v, \mathbf{n}_v]$  from the point cloud, a scale-independent surface feature descriptor can be constructed by calculating 3 angles:

$$f(\mathbf{u}, \mathbf{v}) = \begin{bmatrix} \alpha_{uv} \\ \beta_{uv} \\ \delta_{uv} \end{bmatrix} = \begin{bmatrix} \text{Int}[\angle(\mathbf{n}_u, \mathbf{p}_v - \mathbf{p}_u) / \pi * 720] \\ \text{Int}[\angle(\mathbf{n}_v, \mathbf{p}_u - \mathbf{p}_v) / \pi * 720] \\ \text{Int}[\angle(\mathbf{n}_u, \mathbf{n}_v) / \pi * 720] \end{bmatrix} \quad (7)$$

Let  $a, b, c$  and  $d$  be four directed points that respectively belong to model surface  $\mathbf{a}$ , model surface  $\mathbf{c}$ , scene surface  $\mathbf{b}$  and scene surface  $\mathbf{d}$  with  $\mathbf{a}, \mathbf{c} \in \mathbf{S}$  and  $\mathbf{b}, \mathbf{d} \in \mathbf{M}$ . If  $f(\mathbf{a}, \mathbf{c}) = f(\mathbf{b}, \mathbf{d})$ , then there exists a transformation relationship  $tr(\mathbf{a}, \mathbf{c}, \mathbf{b}, \mathbf{d}) = r \cdot \mathbf{T}_{s \rightarrow m}$ , such that:

$$\begin{cases} \mathbf{p}_a = r \cdot \mathbf{T}_{s \rightarrow m} \cdot \mathbf{p}_b \\ \mathbf{p}_c = r \cdot \mathbf{T}_{s \rightarrow m} \cdot \mathbf{p}_d \end{cases} \quad (8)$$

As shown in Figure 6, the scale factor  $r$  and the homography matrix  $\mathbf{T}_{s \rightarrow m}$  can be calculated as:

$$r = \frac{\|\mathbf{p}_b - \mathbf{p}_d\|}{\|\mathbf{p}_a - \mathbf{p}_c\|} \quad (9)$$

$$\mathbf{T}_{s \rightarrow m} = \mathbf{F}(\mathbf{a}, \mathbf{c}, r)^{-1} \cdot \mathbf{F}(\mathbf{b}, \mathbf{d}, 1) \quad (10)$$

$$\begin{aligned} \mathbf{F}(\mathbf{u}, \mathbf{v}, r) &= \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ 0^T & 1 \end{bmatrix} \\ &= \begin{bmatrix} \frac{\mathbf{p}_{uv} \times \mathbf{n}_{uv}}{\|\mathbf{p}_{uv} \times \mathbf{n}_{uv}\|} & \mathbf{p}_{uv} & \frac{\mathbf{p}_{uv} \times \mathbf{n}_{uv} \times \mathbf{p}_{uv}}{\|\mathbf{p}_{uv} \times \mathbf{n}_{uv} \times \mathbf{p}_{uv}\|} & \frac{r \cdot (\mathbf{p}_u + \mathbf{p}_v)}{2} \\ 0^T & 1 & & \end{bmatrix} \end{aligned} \quad (11)$$

$$\mathbf{p}_{uv} = (\mathbf{p}_v - \mathbf{p}_u) / \|\mathbf{p}_v - \mathbf{p}_u\| \quad (12)$$

$$\mathbf{n}_{uv} = \mathbf{n}_u + \mathbf{n}_v \quad (13)$$

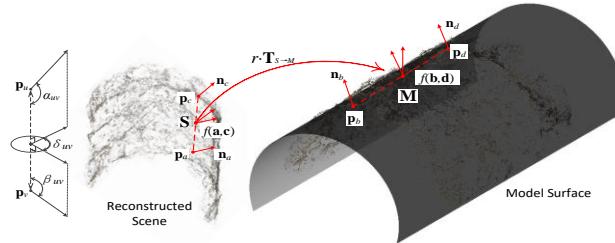


Figure 6. Fast similar surface registration based on scale-invariant robust surface descriptor

#### 4.3. Model recognition

The proportional coefficient was introduced as a surface constraint to solve the similar matching problem between the 3D reconstructed point cloud and the projection model. Meanwhile, the noise interference of the reconstructed point cloud in the model recognition process was eliminated through the development of the RANSAC recognition method.

Let  $\mathbf{M}$  be a model point cloud containing  $n_m$  points, and  $\mathbf{S}$  be a scene point cloud containing  $n_s$  points. Then, the RANSAC recognition algorithm can be explained as follows:

**Step 1:** Randomly select  $\mathbf{a}, \mathbf{c} \in \mathbf{M}$ , calculate the characteristic operator  $f(\mathbf{a}, \mathbf{c})$ , and save the value in the Hash table  $\mathbf{H}$ . If there is a hash collision at  $\mathbf{H}[f(\mathbf{a}, \mathbf{c})]$ , find all the descriptors satisfying  $(\mathbf{b}_j, \mathbf{d}_j) \in \mathbf{S}$  at  $\mathbf{H}[f(\mathbf{a}, \mathbf{c})]$ , calculate the transformation relationship  $tr(\mathbf{a}, \mathbf{c}, \mathbf{b}_j, \mathbf{d}_j)$ , and go to Step 3.

**Step 2:** Randomly select  $\mathbf{b}, \mathbf{d} \in \mathbf{S}$ , calculate the characteristic operator  $f(\mathbf{b}, \mathbf{d})$ , and save the value in the Hash table  $\mathbf{H}$ . If there is a hash collision at  $\mathbf{H}[f(\mathbf{b}, \mathbf{d})]$ , find all the descriptors satisfying  $(\mathbf{a}_i, \mathbf{c}_i) \in \mathbf{M}$  at  $\mathbf{H}[f(\mathbf{b}, \mathbf{d})]$ , calculate the transformation relationship  $tr(\mathbf{a}, \mathbf{c}, \mathbf{b}_j, \mathbf{d}_j)$ , and go to Step 3.

**Step 3:** Evaluate the transformation relationship  $E(tr(\mathbf{a}, \mathbf{c}, \mathbf{b}, \mathbf{d}))$ , increase the number of evaluations  $N_E = N_E + 1$ , record the maximum value of the evaluation function  $E_{best}$  and the corresponding transformation  $tr_{best}$ . The evaluation function is calculated as follows ( $\varepsilon$  is the allowable deviation):

$$E(tr(\mathbf{a}, \mathbf{c}, \mathbf{b}, \mathbf{d})) = \sum_{\mathbf{X} \in \mathbf{M}} c(\mathbf{X}) \quad (14)$$

$$c(\mathbf{X}) = \begin{cases} 1 & dist_S(\mathbf{X}) \leq \varepsilon \\ 0 & dist_S(\mathbf{X}) > \varepsilon \end{cases} \quad (15)$$

$$dist_S(\mathbf{X}) = \min_{\mathbf{Y} \in \mathbf{S}} \|\mathbf{X} - tr(\mathbf{a}, \mathbf{c}, \mathbf{b}, \mathbf{d}) \cdot \mathbf{Y}\| \quad (16)$$

**Step 4:** Determine whether the optimal transformation relationship satisfies the confidence requirement. Terminate the iteration if it is satisfied; otherwise, return to Step 1. The judgment condition is  $N_E > \frac{-\ln(1-p_t)}{p(E_{best})}$ , where  $p_t$  means the confidence of the target is 99.9%, and  $p(E_{best}) = \left( \frac{E_{best}}{n_m} \right)^2$ .

**Step 5:** Perform accurate matching of the optimal transformation  $tr_{best}$  by the iterative closest point (ICP) method (Rusinkiewicz and Levoy, 2002; Besl and McKay, 1992).

## 5. Geometric correction and stitching of photos

### 5.1. Geometric correction of a single photo

As mentioned in Section 3, the transformation relationship between the scene point cloud and the projection model is  $tr(\mathbf{a}, \mathbf{c}, \mathbf{b}, \mathbf{d}) = r \cdot \mathbf{T}_{s \rightarrow m}$ , the camera projection matrix obtained by 3D reconstruction is  $\mathbf{P} = \mathbf{K} \cdot (\mathbf{R} | \mathbf{t})$ , with  $\mathbf{K}$  and  $(\mathbf{R} | \mathbf{t})$  being the internal and external parameters of the camera, respectively. Then, the projection matrix of the camera in the model coordinate system can be expressed as:

$$\mathbf{P}_m = \mathbf{K} \cdot (\mathbf{R} | \mathbf{t}) \cdot r \cdot \mathbf{T}_{s \rightarrow M}^{-1} \quad (17)$$

Let  $(\xi, \eta)$  be the coordinates of a point in the expanded plane coordinate system, and  $(x, y, z)$  be the coordinates of the corresponding point in the model coordinate system. According to the projection model in Section 3.1, the corresponding relationships can be obtained as  $x = f_x(\eta, \xi)$ ,  $y = f_y(\eta, \xi)$  and  $z = f_z(\eta, \xi)$ , with  $f_x, f_y, f_z$  being the functions to convert the projected surface coordinates into the rectangular coordinates. Thus, we have the relationship between the points on the display map and the pixels on the photo:

$$\begin{aligned} \{u, v, 1\}^T &= \{f_u(\eta, \xi), f_v(\eta, \xi), 1\} = \mathbf{f}_{img}(\eta, \xi) = \\ &\text{Dist}_{\kappa_1, \kappa_2}(\mathbf{P}_m \cdot \{f_x(\eta, \xi), f_y(\eta, \xi), f_z(\eta, \xi)\}^T) \end{aligned} \quad (18)$$

where  $\text{Dist}_{\kappa_1, \kappa_2}$  is the radial distortion function. The internal parameter matrix  $\mathbf{K}$ , external parameter matrix  $(\mathbf{R} | \mathbf{t})$  and distortion parameters  $\kappa_1, \kappa_2$  of the camera can be derived through 3D reconstruction.

The display map can be obtained by bilinear interpolation on the original photo according to the following formula:

$$\underset{\eta \in \eta_l \sim \eta_r, \xi \in \xi_l \sim \xi_b, (\eta, \xi) \notin I_m, \alpha \leq \frac{\pi}{12}}{\text{Img}_{flat}(\eta, \xi)} = \text{Img}_{src}(\mathbf{f}_{img}(\eta, \xi)) \quad (19)$$

where  $\eta_l, \eta_r, \xi_l$  and  $\xi_b$  are boundary conditions defining the display range of a single photo after transformation. After acquiring the projection coordinates of the four corner points  $\mathbf{P}_i (i=1 \sim 4)$  on the display map, the optimization function  $\min_{i=1 \sim 4} \|\mathbf{P}_i - \mathbf{f}_{img}(\eta_i, \xi_i)\|$  can be solved by the Levenberg-Marquardt method according to the following formula:

$$\begin{aligned} \eta_l &= \min_{i=1 \sim 4}(\eta_i), \quad \eta_r = \max_{i=1 \sim 4}(\eta_i) \\ \xi_l &= \min_{i=1 \sim 4}(\xi_i), \quad \xi_b = \max_{i=1 \sim 4}(\xi_i) \end{aligned} \quad (20)$$

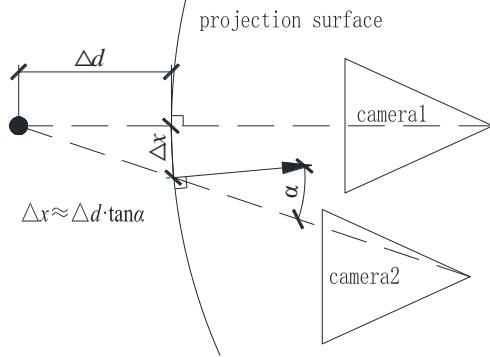


Figure 7. The effect of projection Angle of projection error

Next, the author introduced another two boundary conditions. Specifically, the boundary  $I_m$  indicates the part occluded by other models. To prepare the display map of a tunnel face, the boundary range of the chamber wall of the tunnel should be determined, so that the surface image will not appear in the face image. Otherwise, the subsequent stitching will be negatively affected. The boundary condition  $\alpha$  stands for the projection angle, which can be determined by the angle between the normal vector of the projected model point and the straight line between the point and the camera's optical centre. This boundary condition was designed to reduce the projection error of the uneven surface (Figure 7). Here, the value of  $\alpha$  is set to  $\pi/12$ , i.e. a model point can only be projected by a camera within a  $30^\circ$  cone.

### 5.2. Stitching the panorama of tunnel excavation surface

Project errors are inevitable because the project model cannot recreate the exact surface of the surrounding rock. If the expanded images on the surrounding rock are stitched directly based on the display map coordinates, the resulting photo will carry ghost images and seams (Figure 8a). This calls for one more error compensation before stitching. Here, the objective function is constructed using the relationship between the spatial points determined in 3D reconstruction and their coordinates on the projection plane, and the similar transformation model in the planar stitching model is adopted for error compensation (Szeliski, 2010):

$$f = \min_{s_i, \mathbf{R}_i, \mathbf{t}_i} \sum_{\substack{j=1 \sim n \\ \mathbf{x}_j \in S, c(\mathbf{X}_j)=1}} \left\| \mathbf{x}_j^{ni} - \bar{\mathbf{x}}_j \right\|^2 \quad (21)$$

$$\mathbf{x}_j^{ni} = [s_i \mathbf{R}_i \quad \mathbf{t}_i] \mathbf{x}_j^i \quad (22)$$

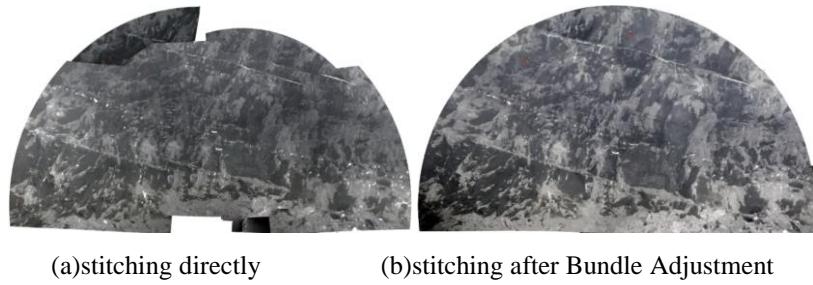
where  $\mathbf{x}_j^i$  are the coordinates of spatial point  $j$  on the expanded plane after being corrected by image  $i$ ;  $s_i$  is the scale factor;  $\mathbf{R}_i$  is the rotation matrix;  $\mathbf{t}_i$  is the

translation vector;  $n$  is the total number of cameras. The objective function should be summed within the inner points  $c(\mathbf{X}_j)=1$  of the scene point cloud that satisfy equation (16) after model recognition and transformation.  $\bar{\mathbf{x}}_j$  is the mean coordinates of spatial point  $j$  on the expanded plane:

$$\bar{\mathbf{x}}_j = \frac{1}{n_{G_j}} \sum_{i \in G_j} \mathbf{x}_j^i \quad (23)$$

where  $G_j$  is the set of cameras that can see spatial point  $j$ ;  $n_{G_j}$  is the number of cameras that can see spatial point  $j$ .

After the error compensation, the panorama was obtained through exposure compensation, the search for the optimal stitching and multi-scale fusion (Figure 8b).

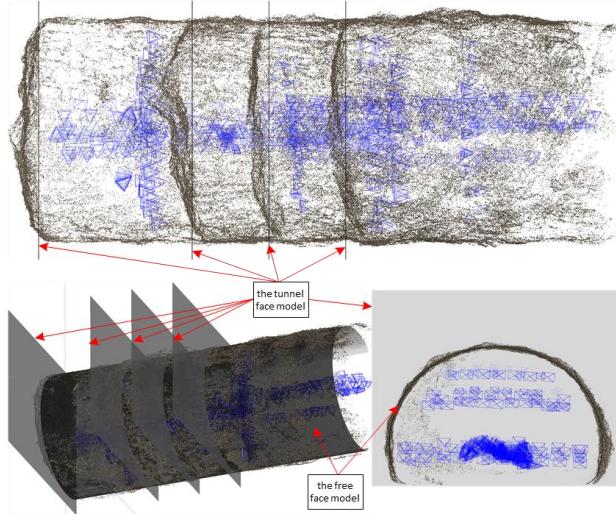


*Figure 8. The effect of Bundle Adjustment before the stitching of orthographic projected photos*

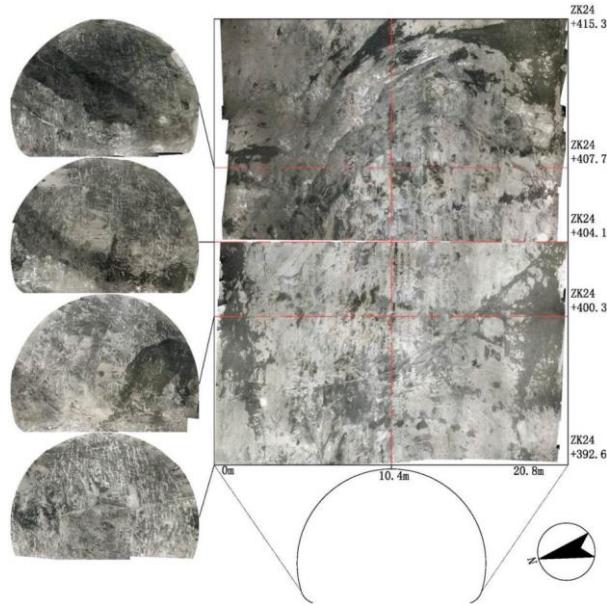
### 5.3. Application

The proposed method was applied for the field shooting of a two-lane expressway tunnel, whose overall axis direction is about 255°. The tunnel was constructed by full-section excavation. The 22.7m-long test section is surrounded by grade III rocks, most of which are hard, medium weathered mudstone. The rock mass is relatively complete with good self-stability. Some pieces of rock may fall off if there is no support. The sidewalls are basically stable. Under concentrated rainfall, the chamber may get wet or have small drips of water. The occurrence of the surrounding rock is 295°∠77°, as shown in the field photos in Figure 3.

During the field shooting, a total of 6 work cycles were captured in 4 sets of photos: the 163 photos in set 1 are about cycles 1~3, the 217 photos in set 2 are about cycles 1~4, the 122 photos in set 3 are about cycles 4~5, the 227 photos in set 4 are about cycles 5~7. These photos were subjected to 3D reconstruction and recognition by the projection model, and the results are presented in Figure 9, which contains five project surfaces. The four rectangular surfaces were adopted to prepare the expanded view of the tunnel face, while the semi-circular surface is the chamber wall model used to prepare the expanded view of the surrounding rock.



*Figure 9. 3D reconstruction result of sample cavern (Blue pyramid indicating the position and direction of cameras when shooting. The upper part is the top view of the reconstructed scene. The lower left part is the oblique view of model estimation result and the lower right part is the front view)*



*Figure 10. Geologic orthographic image of a highway tunnel on zone ZK24+415.3~ZK24+392.6*

The relative position between the tunnel face projection models and the chamber wall projection model was fixed in advance. During the recognition, only the latter was put into use. After recognition, the point cloud of the chamber wall overlapped with the chamber wall projection model, while that of tunnel face did not completely overlap the tunnel face projection models. The mismatch could be eliminated through adjustment of the longitudinal position of each scene. One tunnel surface display map was derived from each set of photos, while the chamber wall display maps were made from sets 2 and 4.

Figure 10 shows the geological map of the entire test section. The map offers a clear view of the changes of the surrounding rock, indicating that the multi-model occlusion method can basically restore the contour of the tunnel face. There is no need for contour extraction and image clipping. The two chamber wall display maps had seams due to excavation disturbance, but the contents of both maps were complete and continuous.

## 6. Conclusions

In this paper, a rapid, flexible construction method is proposed for the geological map of tunnel surrounding rock, through the integration of such computer vision technologies as 3D reconstruction, 3D surface estimation and image stitching. The proposed method was proved efficient and feasible through application in an engineering case. Compared with the mainstream geological logging methods, the proposed strategy enjoys the following advantages:

(1) Great freedom of field shooting: there is no need to calibrate the camera, fix the lens and focal length, measure the shooting position and direction, or determine the object space control points; the shooting distance is unlimited as long as the overlaps are covered; the scenes can be shot freely area by area; if there are heavy dusts in the chamber, the shooting distance can be reduced to avoid blurry images; the shooting position is flexible enough to prevent the occlusion by trolley and other working equipment.

(2) Good universality for image processing: The proposed method applies to tunnels of complex forms, section shapes and inclination angles; the display maps of the chamber wall and tunnel face can be obtained through one shooting due to the multi-model projection and consideration of mutual occlusion relationship.

(3) Excellent image quality: The effect and efficiency of display maps are enhanced by automated image stitching, making it possible to prepare a seamless high-resolution panorama for the geological conditions of the tunnel.

(4) High application value: The proposed method can be implemented at a low cost with ordinary digital cameras.

### Acknowledgement

*This work is supported by Project of National Key R & D Plan (No.2016YFC0802504); Project of National Natural Science Foundation of China*

(No.51608539); Project of China Postdoctoral Science Foundation (No.2016M592451, No. 2017T100610); Science and Technology Project of Guizhou Provincial Transportation Department (No.2018-133-042, No.2018-123-040).

## References

- Agarwal S., Snavely N., Seitz S. M., Szeliski R. (2010). Bundle adjustment in the large. *Computer Vision–ECCV 2010*, pp. 29-42. [https://doi.org/10.1007/978-3-642-15552-9\\_3](https://doi.org/10.1007/978-3-642-15552-9_3)
- Bay H., Ess A., Tuytelaars T., Gool L. V. (2008). Speeded-up robust features (SURF). *Computer Vision and Image Understanding*, Vol. 110, No. 3, pp. 346-359. <http://doi.org/10.1016/j.cviu.2007.09.014>
- Besl P. J., McKay N. D. (1992). A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 14, No. 2, pp. 239-256. <https://doi.org/10.1109/34.121791>
- Brown M., Lowe D. G. (2007). Automatic panoramic image stitching using invariant features. *International Journal of Computer Vision*, Vol. 74, No. 1, pp. 59-73. <https://doi.org/10.1007/s11263-006-0002-3>
- Buyer A, Schubert W. Calculation the spacing of discontinuities from 3D point clouds. *Procedia Engineering*, Vol. 191, pp. 270-278. <http://doi.org/10.1016/j.proeng.2017.05.181>
- Chaiyasarn K., Kim T. K., Viola F., Cipolla R., Soga K., (2015). Distortion-free image mosaicing for tunnel inspection based on robust cylindrical surface estimation through structure from motion. *Journal of Computing in Civil Engineering*, pp. 04015045. [http://dx.doi.org/10.1061/\(ASCE\)CP.1943-5487.0000516](http://dx.doi.org/10.1061/(ASCE)CP.1943-5487.0000516)
- Chen X. X., Li H., Zhang M. Q. (2005). The comparison of two distortion calibration model for general digital camera. *Beijing Surveying and Mapping*, No. 4, pp. 50-54. <http://dx.chinadot.cn/10.3969/j.issn.1007-3000.2005.04.014>
- Fischler M. A., Bolles R. C. (1981). Random sample consensus: A paradigm for model fitting with applicationsto image analysis and automated cartography. *Communications of The ACM*, Vol. 24, No. 6, pp. 381-395.
- Hartley R., Zisserman A. (2004). Multiple view geometry in computer vision. *Cambridge, UK: Cambridge University Press*, pp. 239-259.
- Heymsfield E., Kuss M. L. (2013). Implementing gigapixel technology in highway bridge inspections. *Journal of Performance of Constructed Facilities*, Vol. 29, No. 3, pp. 04014074. [http://dx.doi.org/10.1061/\(ASCE\)CF.1943-5509.0000561](http://dx.doi.org/10.1061/(ASCE)CF.1943-5509.0000561)
- Leng B. (2009). Research of tunne1 face geo1ogy information system based on digital image. *Southwest Jiaotong University*, pp. 25-31. <http://dx.chinadot.cn/10.7666/d.y1689367>
- Li H., Zhang R. C., Yang B., Wu M. F. (2014). Principle and geometric precision of photo geological logging for tunnel. *Journal of Applied Remote Sensing*, Vol. 8, No. 1, pp. 083617. <http://dx.doi.org/10.1117/1.JRS.8.083617>
- Li H., Zhang Y. J., Hua X. S., Yang B. (2004). Geologic logging of digital photos and its basic algorithm. *Editorial Board of Geomatics and Information Science of Wuhan*

- University*, Vol. 29, No. 9, pp. 805-808. <http://dx.chinadoi.cn/10.3969/j.issn.1671-8860.2004.09.013>
- Li S. C., Liu H. L., Li L. P., Shi S. S., Zhang Q. Q., Sun S. Q., Hu J. (2017). A quantitative method for rock structure at working faces of tunnels based on digital images and its application. *Chinese Journal of Geotechnical Engineering*, Vol. 36, No. 1, pp. 1-9. <http://doi.org/10.13722/j.cnki.jrme.2015.1751>
- Lowe D. G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, Vol. 60, No. 2, pp. 91-110. <http://doi.org/10.1023/b:visi.0000029664.99615.94>
- Nocedal J., Wright S. J. (2006). Numerical optimization. *Springer Series in Operations Research and Financial Engineering*, pp. 262-266. <https://doi.org/10.1007/978-0-387-40065-5>
- Papazov C., Burschka D. (2011). An efficient RANSAC for 3D object recognition in noisy and occluded scenes. *computer vision – ACCV 2010*, pp. 135-148. [https://doi.org/10.1007/978-3-642-19315-6\\_11](https://doi.org/10.1007/978-3-642-19315-6_11)
- Rusinkiewicz S., Levoy M. (2002). Efficient variants of the ICP algorithm. *Proceedings Third International Conference on 3-D Digital Imaging and Modeling*, pp. 145. <https://doi.org/10.1109/IM.2001.924423>
- Snavely N., Seitz S. M., Szeliski R. (2006). Photo tourism: Exploring photo collections in 3D. *Acm Transactions on Graphics*, Vol. 25, No. 3, pp. 835-846. <https://doi.org/10.1145/1179352.1141964>
- Snavely N., Seitz S. M., Szeliski R. (2007). Modeling the world from internet photo collections. *International Journal of Computer Vision*, Vol. 80, No. 2, pp. 189-210. <https://doi.org/10.1007/s11263-007-0107-3>
- Song Y., Wang X., Li Y., Li R. (2013). Contrast research on application of 3-dimensional laser scanning technology and digital imaging geology catalog system in acquisition of rock mass information of tunnels. *Tunnel Construction*, Vol. 3, No. 3, pp. 197-202. <http://www.suidaojs.com/EN/10.3973/j.issn.1672-741X.2013.03.005>
- Szeliski R. (2004). Image alignment and stitching: A tutorial, microsoft research. *Faugeras (eds), Handbook of Mathematical Models in Computer Vision*, Vol. 2, No. 11-12, pp. 273-292. <https://doi.org/10.1561/0600000009>
- Szeliski R. (2010). Computer vision: Algorithms and applications. *Springer-Verlag New York, Inc.*, pp. 812. <https://doi.org/10.1007/978-1-84882-935-0>
- Wang M. H. (2007). Key techniques of geologic logging information system for complex tunnel. *HoHai University*, pp. 1-3. <http://dx.chinadoi.cn/10.7666/d.y1128412>
- Wang M. H., Li H., Cang G. H. (2007). Application of ordinary digital cameras in tunnel logging. *Coal Geology & Exploration*, Vol. 4, pp. 15-19. <http://dx.chinadoi.cn/10.3969/j.issn.1001-1986.2007.04.005>
- Wang Y., Wang S. H., Guo M. D., Dong Z. H. (2011). Fast digital identification of joint information of tunnel work face and its stability analysis. *Chinese Journal of Geotechnical Engineering*, No. 11, pp. 1734-1739. [http://manu31.magtech.com.cn/Jwk\\_ytgcxb/CN/Y2011/V33/I11/1734](http://manu31.magtech.com.cn/Jwk_ytgcxb/CN/Y2011/V33/I11/1734)

- Winkelbach S., Molkenstruck S., Wahl F. M. (2006). Low-cost laser range scanner and fast surface registration approach. *Pattern Recognition*, pp. 718-728. [https://doi.org/10.1007/11861898\\_72](https://doi.org/10.1007/11861898_72)
- Wu C. C. (2013). Towards linear-time incremental structure from motion. *2013 International Conference on 3D Vision - 3DV 2013*, pp. 127-134. <https://doi.org/10.1109/3DV.2013.25>
- Yan J. G., Gao G. P. (2007). Application study of geological mapping technique for underground cavern surrounding rock. *Technical Supervision in Water Resources*, No. 3, pp. 27-30. <http://dx.chinadoi.cn/10.3969/j.issn.1008-1305.2007.03.010>
- Yang C. H., Li H., Yang L. (2003). The development of measurable digital camera. *Engineering of Surveying and Mapping*, Vol. 12, No. 2, pp. 34-37. <http://dx.chinadoi.cn/10.3969/j.issn.1006-7949.2003.02.011>
- Yang C. H., Song H. J., Chen C. M. (2001). Digital image processing in photographic geological recording system. *Modern Surveying and Mapping*, Vol. 24, No. 3, pp. 15-18.
- Yang L., Li H. (2002). Digital image projection algorithm and its application in geological catalog. *Modern Surveying and Mapping*, No. 4, pp. 30-33.
- Yang L., Li H., Lv G. N. (2004). A method for digital camera calibration. *Bulletin of Surveying and Mapping*, No. 8, pp. 50-52. <http://dx.chinadoi.cn/10.3969/j.issn.0494-0911.2004.08.018>
- Zhang Y. H., Li L. P., Liu H. L., Yang W. M., Shi S. S. (2016). Digital identification of evaluation of tunnel surrounding rocks discontinuity. *Tunnel Construction*, Vol. 36, No. 12, pp. 1471-1477. <http://www.suidaojs.com/CN/10.3973/j.issn.1672-741X.2016.12.010>
- Zhou C. L., Zhu H. H., Li X. J. (2008). Application of infrared photography and image processing to tunnel construction with new austrian tunneling method. *Chinese Journal of Rock Mechanics and Engineering*, Vol. 27, No. S1, pp. 3166-3172. <http://www.rockmech.org/CN/abstract/abstract24360.shtml>
- Zhu Z. H., Yang J. S., Xiao C., Liu Q. T. (2014). A method for tunnel lining surface photo flattening based on shape functions transform. *Journal of Railway Science and Engineering*, No. 3, pp. 101-106.
- Zhu Z., Fu J. Y., Yang J., Yang J. S. (2016). Panoramic image stitching for arbitrarily shaped tunnel lining inspection. *Computer-Aided Civil and Infrastructure Engineering*, Vol. 31, No. 12, pp. 936-953. <https://doi.org/10.1111/mice.12230>
- Zhu Z., German S., Brilakis I. (2010). Detection of large-scale concrete columns for automated bridge inspection. *Automation in Construction*, Vol. 19, No. 8, pp. 1047-1055. <https://doi.org/10.1016/j.autcon.2010.07.016>

