ISPRS Journal of Photogrammetry and Remote Sensing 117 (2016) 92–107

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

An asymmetric re-weighting method for the precision combined bundle adjustment of aerial oblique images

Linfu Xie [a,d], Han Hu [b,c,*], Jingxue Wang [b], Qing Zhu [b,d], Min Chen [b]

[a] State Key Laboratory of Information Engineering in Surveying Mapping and Remote Sensing, Wuhan University, P.O. Box C310, 129 Luoyu Road, Wuhan, Hubei, PR China
[b] Faculty of Geosciences and Environmental Engineering, Southwest Jiaotong University, Chengdu, Sichuan, PR China
[c] Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hum Hong, Kowloon, Hong Kong
[d] Collaborative Innovation Center of Geospatial Technology, 129 Luoyu Road, Wuhan, Hubei, PR China

ARTICLE INFO

Article history:
Received 30 September 2015
Received in revised form 8 January 2016
Accepted 21 March 2016

Keywords:
Oblique images
Bundle adjustment
Asymmetric re-weighting
Robust estimation

ABSTRACT

Combined bundle adjustment is a fundamental step in the processing of massive oblique images. Traditional bundle adjustment designed for nadir images gives identical weights to different parts of image point observations made from different directions, due to the assumption that the errors in the observations follow the same Gaussian distribution. However, because of their large tilt angles, aerial oblique images have trapezoidal footprints on the ground, and their areas correspond to conspicuously different ground sample distances. The errors in different observations no longer conform to the above assumption, which leads to suboptimal bundle adjustment accuracy and restricts subsequent 3D applications. To model the distribution of the errors correctly for the combined bundle adjustment of oblique images, this paper proposes an asymmetric re-weighting method. The scale of each pixel is used to determine a re-weighting factor, and each pixel is subsequently projected onto the ground to identify another anisotropic re-weighting factor using the shape of its quadrangle. Next, these two factors are integrated into the combined bundle adjustment using asymmetric weights for the image point observations; greater weights are assigned to observations with fine resolutions, and those with coarse resolutions are penalized. This paper analyzes urban and rural images captured by three different five-angle camera systems, from both proprietary datasets and the ISPRS/EuroSDR benchmark. The results reveal that the proposed method outperforms the traditional method in both back-projected and triangulated precision by approximately 5–10% in most cases. Furthermore, the misalignments of point clouds generated by the different cameras are significantly alleviated after combined bundle adjustment.
© 2016 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.

1. Introduction

As interest in urban 3D reconstruction increases (Rottensteiner et al., 2014), aerial oblique images have become objects of public attention due to their ability to display a building facade from different directions and their re-invention by Pictometry (Petrie, 2009). Since then, similar penta-view oblique camera systems, including Pictometry, MIDAS, IGI, Leica RCD30, Icaros, and SWDC-5, have prospered (Rupnik et al., 2014). For photogrammetric practitioners, these novel platforms have numerous applications, which range from building verification (Nyaruhuma et al., 2012a, 2012b), damage assessment (Gerke and Kerle, 2011), and public mapping services to creating textures for models of buildings (Kada, 2009).

In most of their previous applications, oblique images were only used as "pretty images". But recent advances in dense image matching (DIM) have triggered interest in exploiting their potential for 3D applications (Fritsch et al., 2012; Fritsch and Rothermel, 2013; Xiong et al., 2014; Rupnik et al., 2015), in which a combined bundle adjustment (BA) of all the images is necessary. However, making a combined BA is non-trivial compared to the sophisticated solutions available for nadir images. The difficulties are due to the irregular configurations of the oblique images, which have consequences for feature matching, as indicated in previous study (Hu et al., 2015), and for determining the *a priori* uncertainties of the observations in a bundle block.

BA is regarded as the gold standard for solving 3D reconstruction problems using image correspondences (Snavely et al., 2008), and the errors in observations are commonly assumed to

* Corresponding author at: Faculty of Geosciences and Environmental Engineering, Southwest Jiaotong University, Chengdu, Sichuan, PR China. Tel.: +86 13437119857.
E-mail address: huhan8807@gmail.com (H. Hu).

http://dx.doi.org/10.1016/j.isprsjprs.2016.03.017
0924-2716/© 2016 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.

have normal distributions, which is known as the Gaussian assumption (Baselga, 2007). Traditional bundle adjustment, which was designed for nadir images with rectangular footprints and similar scales, usually applies identical and symmetrical weights to the different parts of point observations made from different directions, in accordance with the assumption that all of the point observations are of the same quality. However, due to the peculiar characteristics of their tilted viewpoints (Höhle, 2008; Gerke, 2009), oblique images have trapezoidal footprints on the ground, and different parts of them have conspicuously different ground sample distances (GSDs), even when the terrain is flat. Hence, the accuracies of matching points in different regions of oblique images are no longer identical. In other words, the error distribution for all the observations does not conform to the same Gaussian distribution; in fact, the error becomes a function of the image's scale and GSD. Therefore, weighting the points in an image asymmetrically according to the image's scale and the area of each pixel's projection onto the ground is more rigorous.

Aiming at correctly modeling this kind of systematic inhomogeneity, in this paper, we present an asymmetric re-weighting method for the combined BA of aerial oblique images. Two factors are considered in the determination of the *a priori* weight of each image observation: (1) a normalized scale factor represents the scale of each pixel, and (2) a relative anisotropic GSD represents the different GSDs of a single pixel for different directions. The integration of these factors assigns greater weights to observations with fine resolutions and penalizes those with coarse resolutions. Initial exterior orientation (EO) parameters, which are derived from the onboard global navigation satellite system (GNSS) and pre-calibrated platform parameters, are exploited to initialize the re-weighting factors and update them according to the renewed EO parameters during each iteration of the BA.

Section 2 briefly reviews the existing combined BA methods for multi-camera systems. Then, in Section 3, the asymmetric re-weighting method is introduced in detail. The performance of the proposed method is evaluated in Section 4 using four blocks of images acquired by three different multi-camera systems, includes a proprietary dataset in China and two ISPRS/EuroSDR benchmark datasets. Finally, the conclusions are presented in Section 5.

2. Related work

In the photogrammetry and computer vision communities, BA is always formulated as a nonlinear least squares problem that assumes there is underlying Gaussian noise when it is viewed as an inference method in a probabilistic model (Triggs et al., 2000; Zach, 2014). In BA, a variety of observations and scene constraints can be incorporated, including image correspondences (points, lines, curves, and surfaces), ground control points (GCPs), GPS/IMU measurements (Ip et al., 2007), right angles (Gerke, 2011), and parallel lines (Zhang et al., 2011), to refine the camera parameters accurately and efficiently (Triggs et al., 2000). The observations have different scales and dimensions, and the standard practical method for combining them into a single BA is to weight their contributions to the block appropriately using their *a priori* uncertainties. For example, the uncertainties of points matched automatically is between 0.1 and 0.5 pixels (Gruen, 2012), while manually measured image points are assumed to have an uncertainties of 0.5 pixels. For GCPs, the uncertainties of the object space depends on the surveying technique; it ranges from 0.05 m to 0.5 m (Yuan et al., 2009; Gülch, 2012). For onboard GPS/IMU measurements, the positioning uncertainties is between 0.05 m and 0.3 m; the uncertainties of the attitudes, roll and pitch could reach 0.005°, and the uncertainty of the heading is 0.008° (Ip et al., 2007). The weights of constraints between points (horizontal and vertical lines, right angles) should be greater than those of the other observations (Gerke, 2011).

The primary difference between BA for oblique images and nadir images is in the pre-calibrated platform parameters, which represent the rotational and translational relationships between nadir and oblique cameras. For creating a combined BA, two models are widely used: orientation with and without platform parameters (Rupnik et al., 2013). When a BA is implemented without platform parameters, the framework becomes similar that of traditional image orientation. When a BA is processed with fixed platform parameters, one retrieves the EOs of oblique images from a combination of the EOs of the corresponding nadir images and the platform parameters (Tommaselli et al., 2010). This process stabilizes the image block and decreases the number of unknowns (Wiedemann and Moré, 2012). However, due to mechanical reliability and possible problems with synchronization and re-installation, others argue that the pre-calibrated platform parameters should be considered unknowns and be adjusted during the BA (Jacobsen, 2009). In this way, the observations of the initial platform parameters that were used for calibration in the laboratory environment should be weighted according to their calibration accuracies.

The weighting methods described above generally adopt the assumption that the observation errors have a Gaussian distribution; however, for outliers, this assumption does not hold. In applications of aerial triangulation, the majority of outliers come from mismatches of tie points. They can be modeled using the so-called mean-shift model and the variance inflation model (Cen et al., 2003), which lead to two fundamental techniques for handling them: (1) gross error detection and (2) robust estimation (Huber, 1971; Klein and Förstner, 1984). Due to the large number of observations and unknown parameters in aerial photogrammetry engineering, robust estimation is very efficient for practical applications, and the iteratively reweighted least squares method has become the standard strategy for exploiting robust estimation in a nonlinear least squares framework (Zach, 2014). As summarized in (Baselga, 2007), using an equivalent weight function, one can rescale *a priori* weights and obtain an updated minimized objective function for the next iteration. This allows robust estimators (also called loss functions) such as the L_1-norm (Appa and Smith, 1973) and Huber (Huber, 1971) estimators (see Fig. 1) to be embedded as equivalent weight functions, making the parameter estimation less sensitive to outliers. All of these robust estimators down weight outliers to make them less influential.

As described above, previous studies have generally categorized image correspondences into sets of inliers and outliers. It is expected that outliers are eliminated and that inliers have the same *a priori* uncertainty. Therefore, tie point observations receive identical weight in both directions and regardless of their locations during the combined BA. However, this is not the case for oblique images. Due to their large tilt angles, the image scale changes significantly in an oblique view, and the projection area of each pixel is trapezoidal instead of rectangular (Wiedemann and Moré, 2012). Therefore, the uncertainties of the different locations of corresponding points in each image are probably not the same. For digital images, pixel is the smallest unit to record information. When a certain point in object space has been captured in images with different GSDs, the differences in loss of information will result in different levels of random uncertainty. Image tie points that appear in locations with coarse resolutions are not as accurate as those that appear in locations with fine resolutions. Since their expectations remain zero, their error distributions still satisfies the zero-mean Gaussian assumptions, but their variances are larger. To correctly model the error distribution of a bundle block combining nadir images and oblique images, in this paper, we propose an asymmetric re-weighting method that gives each correspondence different

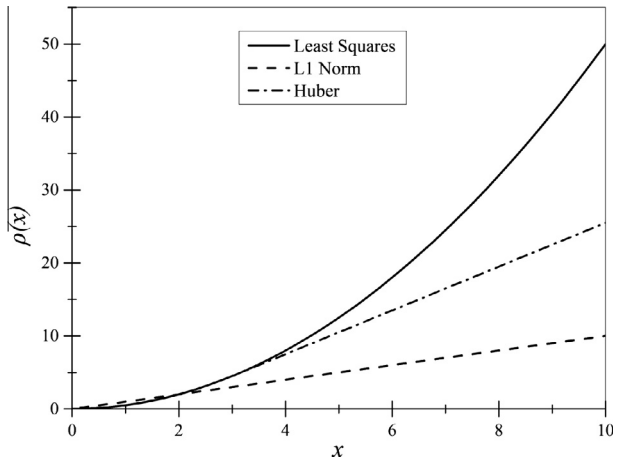

Fig. 1. The cost functions of different robust estimators: x denotes the residual of an observation, and $\rho(x)$ denotes its contribution to the cost.

weights in the x and y directions; these weights depend on the scale of the image and the shape of each pixel's footprint on the ground.

3. Methods

3.1. Principles of asymmetric re-weighting

The combined BA model for multi-camera systems shown in Fig. 2 is developed. Given the initial EO parameters for all of the nadir images and the platform parameters, the EO parameters of all of the oblique images are initialized within a certain precision. After the automatic generation of tie points between overlapping images (Hu et al., 2015), a BA based on the following co-linearity equation (Wang, 1990) is used to refine the EO parameters of all of the images:

$$\begin{cases} x - x_0 + \Delta x = -f\,\frac{m_{11}(X-X_S)+m_{12}(Y-Y_S)+m_{13}(Z-Z_S)}{m_{31}(X-X_S)+m_{32}(Y-Y_S)+m_{33}(Z-Z_S)} \\ y - y_0 + \Delta y = -f\,\frac{m_{21}(X-X_S)+m_{22}(Y-Y_S)+m_{23}(Z-Z_S)}{m_{31}(X-X_S)+m_{32}(Y-Y_S)+m_{33}(Z-Z_S)} \end{cases}, \tag{1}$$

where f, x_0, and y_0 are the image's interior orientation (IO) parameters, Δx and Δy represent image distortions, x and y are image observations, X, Y, and Z are the object coordinates of the tie point, X_S, Y_S, and Z_S are the translated EO parameters, and m_{11} through m_{33} are the elements of the rotation matrix for the EO parameters. Once Eq. (1) has been linearized, the observation equation for the multi-camera system can be represented in matrix form as shown in Eq. (2),

$$V = AX - L, P \tag{2}$$

where V contains the residual vectors of all of the observations, A contains the partial derivates, L contains the observations, and P is the weighting matrix for the observations.

The basic assumption in weighting image observations in a combined BA is that the precision should be directly related with the matching precision and the image quality. When estimating image correspondences between nadir and oblique images, image patches should be rectified locally or globally, in order to relieve the significantly geometrical deformations. In this procedure, more pixels will be used to determine the correspondence if the GSD is smaller and thus a more stable match. Furthermore, in our practical experiences, due to complex lighting conditions and other effects, the images are blurred in the far end of the oblique images that are of larger GSD, so inferior matching qualities are expected for them. Because the actual image matching quality is hard, if not possible, to be quantitatively measured, an intuitive method is to weigh the image observations with the GSD.

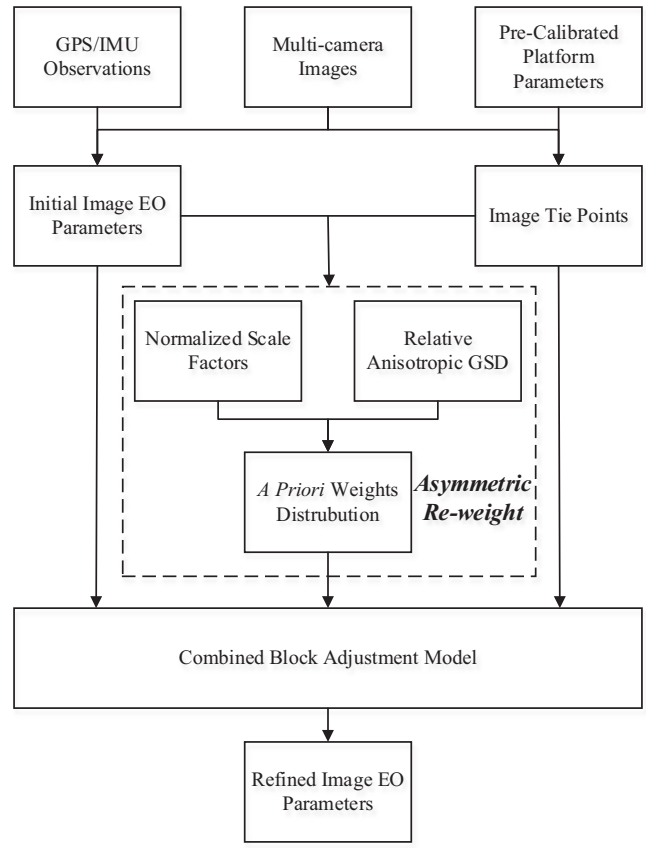

Fig. 2. The framework for the combined bundle adjustment of aerial oblique images using the asymmetric re-weighting method.

Although the image measurement accuracy is hard to be quantitatively measured, in order to give an intuitive understanding of the fall-off on the accuracy of image observations with respect to the GSD, we conduct a Monte Carlo simulation with the following simplified set-up. We assume there are one nadir and oblique image each, which covers the same planar area. Thus a one on one homography relationship can be estimated (Hu et al., 2015) between the two images denoted as ($\boldsymbol{p}_o$, $\boldsymbol{p}_n$), which are considered as ground truth match. We then add Gaussian noises of one pixel to the oblique images and projected the noised coordinates to the nadir image. The GSD of $\boldsymbol{p}_o$ is estimated and binned into intervals of 0.05 m. Fig. 3 clearly shows that the average and standard deviation of the projected error are related with the GSD.

In this paper, the *a priori* weights of each image observation in the x and y directions are computed and injected into the standard BA as shown in Fig. 2. The *a priori* weights (p_x and p_y) are determined using the scale and the shape of the pixel in object space; they are not symmetric with respect to the two axes of the image coordinate system. As the cost of an observation is still in direct proportion to the square of its residual, the stochastical model is maintained while the variances of different image observations in the Gaussian model are modified. Furthermore, a different statistical model is actually used to handle outliers, which is the Huber loss function (Huber, 1971). Specifically, the weights are defined using the two factors shown in Eq. (3): the normalized scale factor (F_S) and the relative anisotropic GSD (F_R).

$$p_x = \frac{1}{F_S}, p_y = \frac{1}{F_S F_R} \tag{3}$$

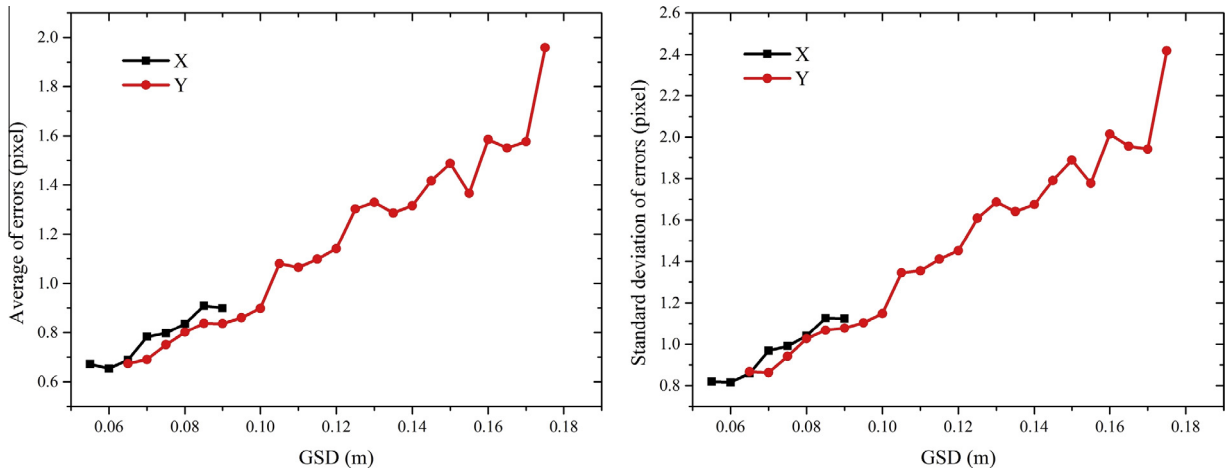

Fig. 3. Results of the Monte Carlo simulation. The figures show the average and standard deviation of the projected errors with respect to different GSD at the same level of noises on the oblique images.

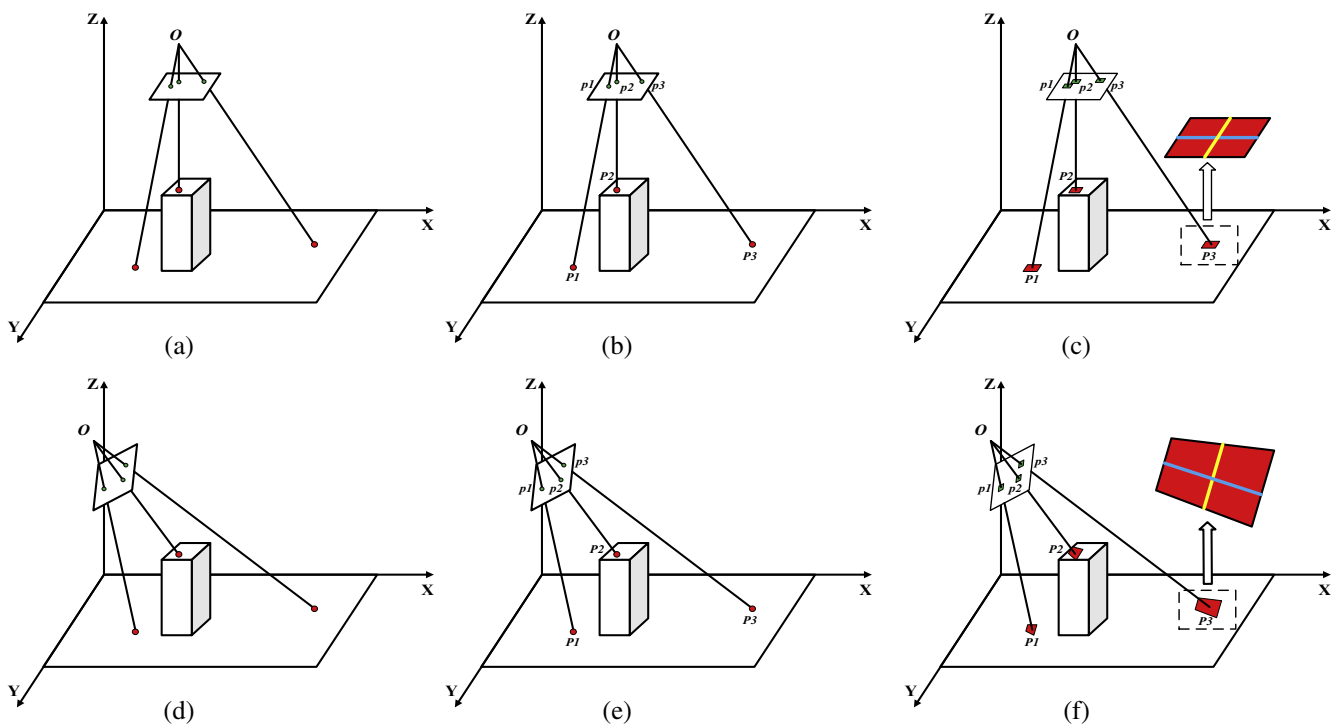

Fig. 4. Schematic diagrams of the normalized scale factor and relative anisotropic GSD for vertical and oblique images. The left schematics show the terrain and the camera's POS, the middle schematics depict the differences in the image scales, and the right schematics depict the footprints of single pixels. The first row depicts the vertical view and the second row depicts an oblique view.

The normalized scale factor portrays the variation of the scale of matched image point without any input from the trapezoidal shape of the pixel's footprint. As shown in Fig. 4 the scale number is the ratio of distance between the ground point and the center of the projection (*O-Pi*) to the distance between the image point and the center of the projection (*O-pi*), which is the reciprocal of the image scale. It is determined using the 3D position of the tie point and camera's IO and EO parameters. A larger number corresponds to a lower image resolution. To normalize the factor, the above scale number is divided by that of the nadir image at the average flight altitude. As shown in Fig. 4, the normalized scale factor is influenced by two major aspects: the variation of the terrain and the attitude of the aerial vehicle. For nadir images with nearly negligible yaw and roll angles, F_S is primarily determined by variations in the terrain. For oblique images, F_S is also affected by the tilt angle, as shown in Fig. 4(e).

The relative anisotropic GSD expresses the difference in the GSDs within a single pixel and is influenced by the trapezoidal shape of the pixel's footprint. It reveals the relative quality of the image in one pixel along different axes. In Fig. 4(c) and (f), the yellow[1] and blue lines represent the GSD in the *x* and *y* directions, respectively. Because the area of one pixel is sufficiently small and only the relative ratio is of interest, the four corners of a pixel are only projected onto a horizontal plane to determine the GSDs in the two directions. As shown in Fig. 4(c) and (f), the differences in GSDs are significantly larger in oblique images compared to those in nadir images. If the image is tilted along the *y*-axis of the image, then F_R will be larger than one and thus penalize the contributions

[1] For interpretation of color in Fig. 4, the reader is referred to the web version of this article.

of y coordinates of the image measurements. On the other hand, if the image is tilted along the x-axis, F_R will be smaller than one, thus award larger weights accordingly.

The synthesis of these two factors reflects the GSDs of each image observation along the two orthogonal image coordinates and determines their *a priori* weights. Neither of them are dispensable. On one hand, the normalized scale factor evaluates the image's quality on the basis of its scale but ignores the anisotropy in each pixel. On the other hand, the relative anisotropic GSD focuses on the trapezoidal shape of the pixel's footprint caused by its three rotation angles regardless of projection center position and variations in the terrain.

Once the *a priori* weights have been determined, the initial weight matrix, P in Eq. (2), of image observations is reset by rescaling the equivalent weight function (Baselga, 2007). In this way, the covariance matrix appropriately models the statistical behavior of the tie points observed in the oblique images. The residuals of the observed tie points with coarser resolutions have Gaussian distributions with higher variances. In the following section, a method of calculating these two factors is introduced in detail and their wide variety in oblique images is discussed.

3.2. The normalized scale factor

As described in (Höhle, 2008), for an oblique image with a focal length f and tilt angle θ along one axis, as shown in Fig. 5, the scale of an image point can be determined using Eq. (4),

$$m = \frac{O\text{-}PT}{O\text{-}pt} \approx \frac{h/\cos(\theta+\beta)}{f/\cos\beta} \tag{4}$$

where m is the *scale number* of the image point, h is the flight altitude, β is the angle between the line from the projection center to the principal point and the line from the projection center to the image point. The lengths of $O\text{-}pt$ and $O\text{-}PT$ are approximately equal to $f/\cos\beta$ and $h/\cos(\theta+\beta)$, respectively. The normalized scale factor is defined in Eq. (5),

$$F_s = m/m_0 \tag{5}$$

where m_0 is the scale of the corresponding nadir image obtained using Eq. (4) with $\theta = 0$. In practice, the rough 3D position of the target point (PT) in object space and the initial EO (O) parameters are used to calculate scale as shown in Eq. (6).

$$m = \frac{\sqrt{(X-X_S)^2+(Y-Y_S)^2+(Z-Z_S)^2}}{\sqrt{f^2+(x-x_0)^2+(y-y_0)^2}}. \tag{6}$$

For a given camera in a given flight, h and f are constant. When $h/f = 1$, the relationship between m, θ and β is apparent. For nadir images, the image scale remains constant. For most off-the-shelf

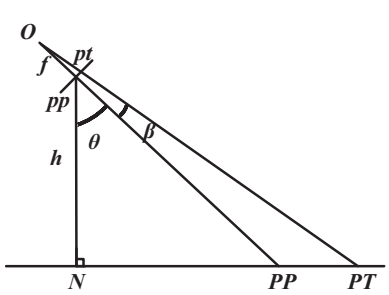

Fig. 5. A sketch of an oblique view. O is the projection center, pp is the principal point of the image, PP is the intersection of $O\text{-}pp$ and the ground, pt is a target point in the image, PT is the ground point corresponding to pt, h is the altitude of the flight, f is the focal length, θ is the tilt angle, β is the angle between the principal optical axis and the ray from the target point to the center of the projection.

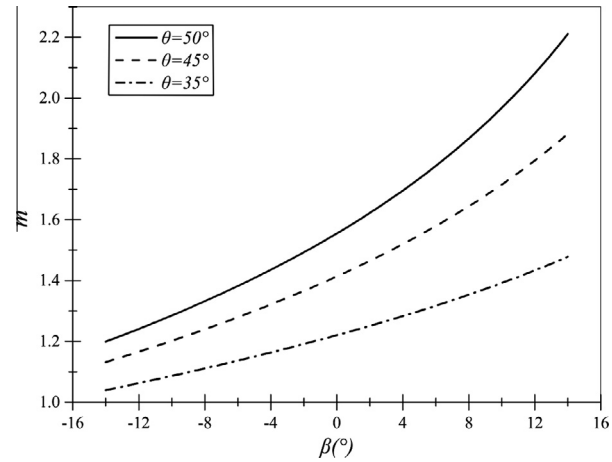

Fig. 6. The functional relationship between the image scale and the angle between the principal optical axis and the ray from the target point to the projection center (β in Fig. 5) for a given tilt angle (θ in Fig. 5). A tilt angle of 35° represents the default configuration of the Leica RCD30 Oblique Penta system, one of 45° represents the SWDC-5 multi-camera system, and one of 50° represents the oblique images provided by Pictometry.

aerial oblique camera systems, the tilt angle is between 35° and 50° (Gerke, 2009; Leica-Geosystems, 2014). From Fig. 6, it is obvious that the scale of an oblique image increases distinctly from the back to the front of the image. The larger the tilt angle is, the greater the difference is. Due to flight conditions, oblique images may not be taken at the anticipated tilt angle. Instead, the tilt angle may become as much as 60° (Grenzdörffer et al., 2008). We can infer that, for off-the-shelf commercial multi-camera systems in real flights, the ratio between the maximum and minimum scale may be as large as two.

3.3. The relative anisotropic GSD

Fig. 7 shows the results of simulations of the footprints of nadir images and oblique images performed by projecting each pixel's four corners onto flat ground. For the nadir images, the shape of each pixel's ground projection is almost rectangular and its GSDs in the x and y directions are similar, even in the presence of significant yaw angles, as shown in Fig. 7(a) and (c). However, for oblique images, this property does not exist. In the simplest case, when the oblique image tilts along one axis only, the ground projection of each of its pixels is always trapezoidal (Fig. 7(b)). Although we assume tilting in y-axis in this figure, however the formulation is general and can be applied for any flight configurations. Focusing on a single pixel, it is obvious that the GSD in the x direction is not equal to that in y direction, and the ratios between them are different due to differences in the pixel coordinates. In an actual flight, an oblique image may tilt along more than one axis. The simulation results shown in Fig. 7(d) indicate that this type of GSD anisotropy changes asymmetrically in the presence of small rotations around the other two axes.

For simplicity, the GSDs of one pixel in two directions are approximately represented by the average of the lengths of two sides of its footprint, as shown in Eq. (7),

$$\begin{cases} GSD_x = (a+c)/2 \\ GSD_y = (b+d)/2 \end{cases} \tag{7}$$

where GSD_x and GSD_y are the GSDs in the x and y directions, respectively, and the yellow and blue lines, respectively, in Fig. 4, and a, b, c, and d are the lengths of the sides of the quadrangle shown in Fig. 7(d). When the ground is perfectly flat, the ground coordinates

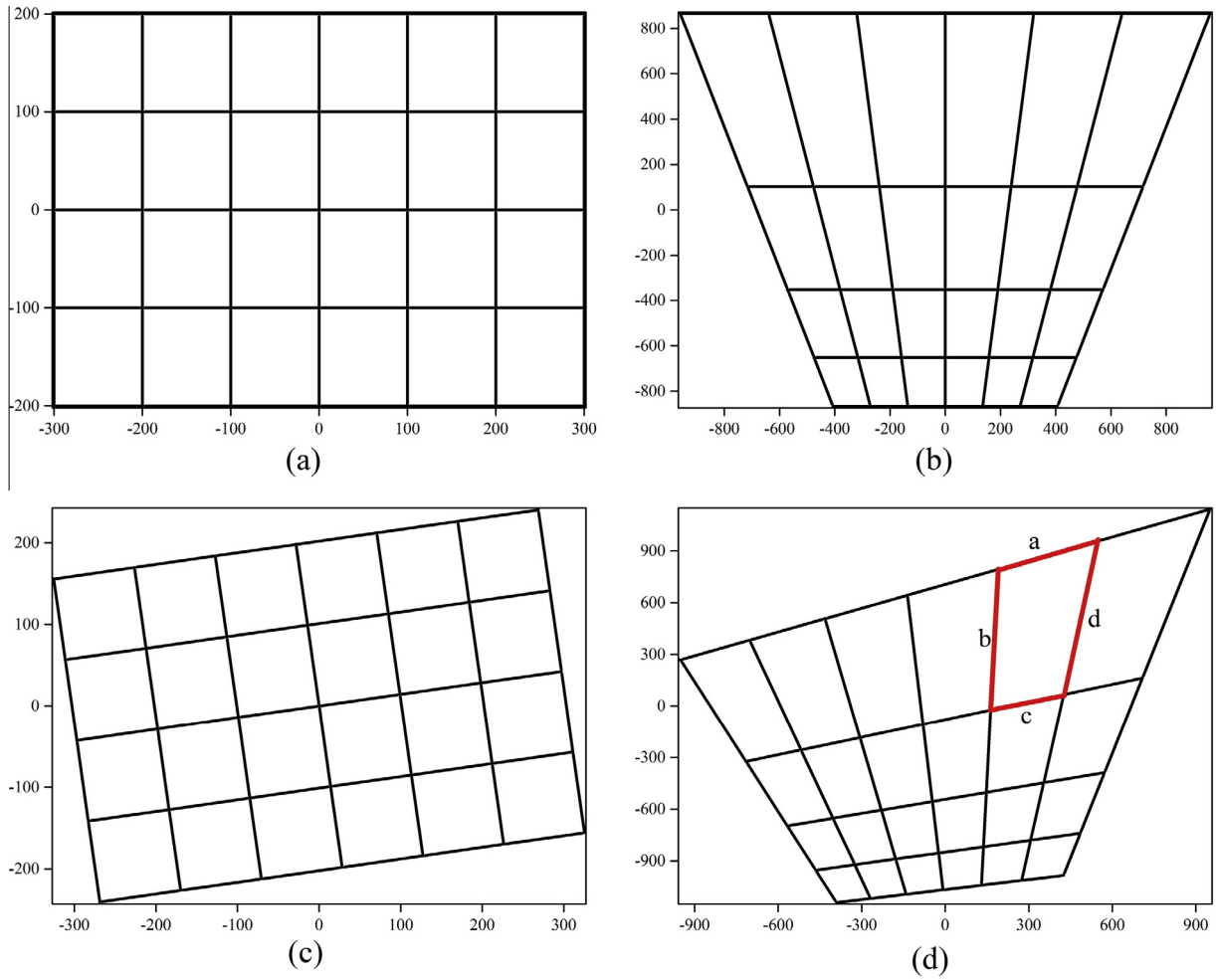

Fig. 7. Simulated footprints of nadir and oblique images with and without unanticipated rotation. In the simulation, the image's resolution is 6 × 4 pixels, the physical pixel size is 10 × 10 mm, the focal length is 100 mm and the flight altitude is 1000 m. The footprints of nadir images are on the left, and those of oblique images are on the right. The simulated tilt angle is 60°, and the unanticipated rotation is 15° around the *Z*-axis.

of the pixel's four corners are determined using the pre-calibrated IO and initial EO parameters, and the side lengths are calculated. Because of Eqs. (7) and (8) is used to obtain the relative anisotropic GSD,

$$F_R = \frac{GSD_y}{GSD_x} \tag{8}$$

In Fig. 8, the distributions of the relative anisotropic GSDs for vertical and oblique views are compared. In the anticipated vertical configuration shown in Fig. 8(a), the value of the vertical F_R for all of the pixels is 1.0. In an actual flight, which always has some amount of pitch, yaw and roll, the relative anisotropic GSDs remain in the vicinity of 1.0, and for the extreme case in which the pitch angle is as large as 15°, see Fig. 8(b) and (c). However, the situations of the oblique images are quite different, and the relative anisotropic GSD is different for different pixels. Under the ideal conditions shown in Fig. 8(d), the value of F_R ranges from 1.15 to 1.86 and changes significantly in one image. The same phenomenon occurred in the situation shown in Fig. 8(e), which includes small rotations around the other two axes. For the extreme case shown in Fig. 8(f), the largest relative anisotropic GSD is 3.17. Through this comparison, we learn that the relative anisotropic GSD of a vertical image remains almost unchanged at 1.0 regardless of the pixel's location and any unexpected rotation, and that for oblique images, the GSD not only is inconsistent between pixels in different regions of one image but also changes significantly due to the attitude at which the image is captured. Although, the ratio F_R is influenced by the unflatness in the area of the single pixel, quantitative gauging of the slope inside a pixel is hard, if not impossible. Furthermore, due to occlusion and other factors, most of the tie points lie on the roof of the building or the flat grounds, so in practical consideration, a flat assumption is plausible approach.

4. Experimental results and analysis

4.1. Description of the test data

To compare the performance of the proposed BA method with that of the traditional method, four blocks of images acquired by three different multi-camera systems are employed (see Table 1), which cover Jinyang, Zurich and Dortmund and are described in Sections 4.2–4.4, respectively.

The first and second image blocks are captured in Jinyang, Guizhou Province, China by a pre-calibrated SWDC-5 multi-camera system, which comprises one nadir camera and four oblique cameras. The first block contains 140 images from four strips, and each strip includes seven camera stations and covers an urban area of 1.3 km by 1.7 km. The second block includes 135 images captured in three strips with nine camera stations

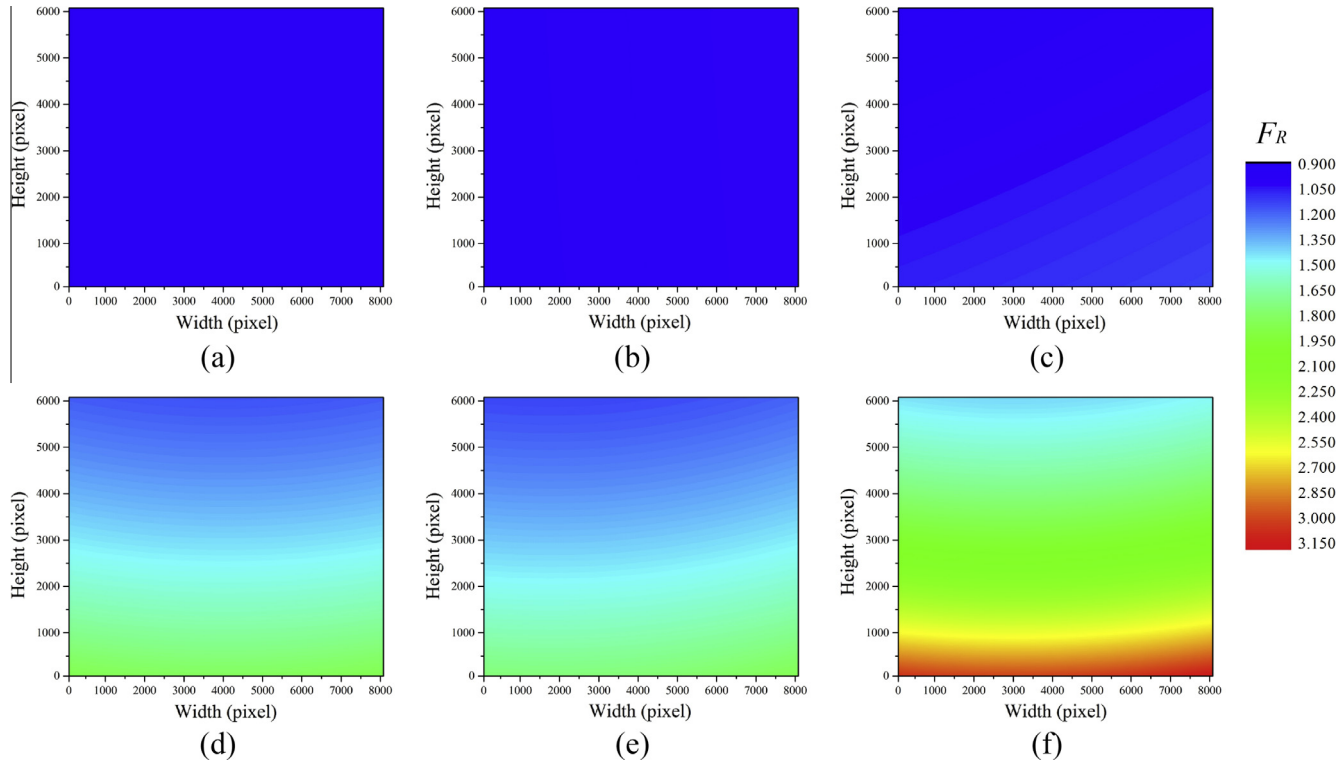

Fig. 8. Simulated relative anisotropic GSDs for vertical and oblique images. The first row represents vertical images; the three rotations, (pitch, roll, heading), are (0°, 0°, 0°), (0°, 5°, −5°), and (15°, 5°, −5°). The second row represents oblique images; the three rotations are (45°, 0°, 0°), (45°, 5°, −5°), and (60°, 5°, −5°).

Table 1
Primary parameters of the experimental image data used in this paper. *N* stands for nadir camera, *O* stands for oblique camera, *F* stands for flight direction, *C* stands for cross flight direction.

Datasets	Jinyang	Zurich	Dortmund
Sensor	SWDC-5	RCD 30	IGI PentaCam
Focal length (*N*/*O*) (mm)	50/80	53/53	50/80
Image resolution (*pixel*)	8176 × 6132	9000 × 6732	8176 × 6132
Pixel format (mm)	0.006	0.006	0.006
Nadir overlap (*F*/*C*)	60%/40%	70%/50%	60%/60%
Flight altitude (m)	750	520	900
Nadir GSD (cm)	8	6	10

per strip and covers a rural area of 1.6 km by 1.2 km. As shown in Fig. 9, the SWDC-5 system adopts the typical "Maltese Cross" configuration (Petrie, 2009), with the plane symbol indicating the actual flight direction. The image resolution is 8176 × 6132 pixels and the physical pixel size is 0.006 × 0.006 mm. The average flight altitude is 750 m above the ground. The principal distances of the nadir and oblique cameras are approximately 50 mm and 80 mm, respectively. The along-track overlap is 60%, and the cross-track overlap is approximately 40%. With this configuration, the average GSDs for the nadir and oblique images are approximately equal at 0.08 m. The nadir images are adjusted using an integrated sensor orientation (ISO) process (Heipke et al., 2002) with the assistance of GPS measurements, and the practical accuracies of the EO parameters obtained using the ISO process for the nadir images are 0.030 m and 0.070 m in the vertical and horizontal directions, respectively, at independent checkpoints. Therefore, the EO parameters for the nadir images are fixed during the combined BA to provide absolute references.

The third image block is delivered by the EuroSDR/ISPRS oblique airborne benchmark dataset on dense image matching, covering an area of about 1.75 km^2 in Zurich, Switzerland (Cavegn et al., 2014). The Leica RCD30 Oblique Penta is employed, which equipped with Leica NAG-D 50 mm lenses and mounted with tilt angles of 35°. The image resolution is 9000 × 6732 pixels and the physical pixel size is 0.006 × 0.006 mm. The average flight altitude is about 520 m above the ground. The along-track overlap is 70%, and the cross-track overlap is approximately 50%. As reported by the benchmark sponsor, the 135 images captured by 5 cameras in 3 strips have already been well oriented, that for GCPs the RMSE in object space is 0.032 m, 0.026 m and 0.091 m in *X*, *Y*, *Z* direction respectively while for tie points the RMSE of 3D accuracy is about 0.195 m. Therefore, we add some Gaussian noises to EO parameters of all the oblique images and use the nadir images to provide absolute references during the combined BA as in Jinyang area.

The fourth image block is provided by the ISPRS benchmark for multi-platform photogrammetry, surveyed over the city of Dortmund, Germany (Nex et al., 2015). The released image block has 314 images in 7 strips with truncated camera position, which are captured by the PentaCam IGI multi-camera system (tilt angles are 45°) with 60% along-track and 60% across-track overlaps. The focal length of the nadir and oblique cameras are approximately 50 mm and 80 mm, respectively. The image resolution is 8176 × 6132 pixels and the physical pixel size is 0.006 × 0.006 mm. Besides not well distributed GCPs and control points (CPs) are provided, so we also estimate the EO parameters for the nadir image rather than fixing them during the combined BA as described above.

For all the four test datasets, we start from a set of image correspondence points and first perform a BA with the Huber robust estimator (Huber, 1971) to reject mismatches. Then, the standard and proposed BA methods are compared with correspondence after outlier detection. It should be noted that the only difference between the asymmetrical approach and the classical one is the preprocessing step in initializing the weights for image observations. When comparing against to other parts of the BA, the computational cost is almost negligible. For the Jinyang and Zurich

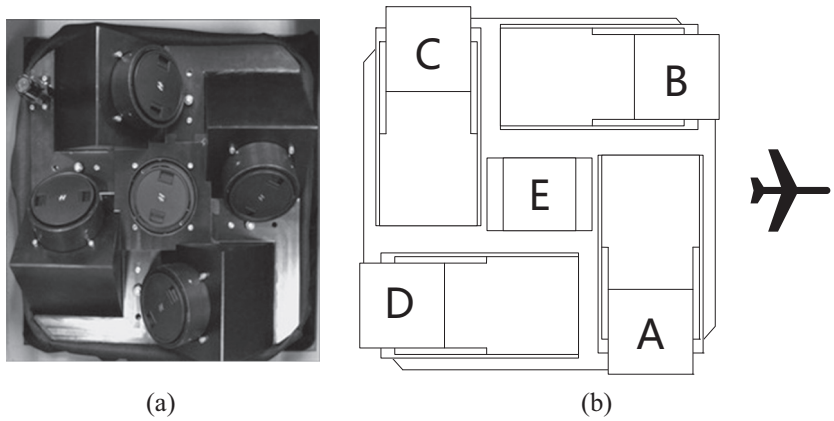

Fig. 9. Camera configuration of the SWDC-5 multi-camera system. A–E refer to *CamA–CamE*. The aircraft symbol indicates the flight direction. (a) Top view of the multi-camera system and (b) schematic bottom view of the cameras with flight direction.

Table 2
The theoretical BA accuracy in the Jinyang urban area.

Camera	σ_0 (*pixels*)	*X* (mm)	*Y* (mm)	*Z* (mm)	φ (1/1000°)	φ (1/1000°)	φ (1/1000°)
CamA	0.25	7.12	8.52	12.52	0.88	0.68	0.51
CamB		6.92	6.77	12.77	0.68	0.51	0.46
CamC		5.67	7.18	8.58	0.87	0.66	0.51
CamD		8.82	8.24	15.76	0.71	0.51	0.53

Table 3
The theoretical BA accuracy in the Jinyang rural area.

Camera	σ_0 (*pixels*)	*X* (mm)	*Y* (mm)	*Z* (mm)	φ (1/1000°)	φ (1/1000°)	φ (1/1000°)
CamA	0.27	7.92	10.01	13.25	0.95	0.80	0.61
CamB		6.75	5.28	7.89	0.61	0.62	0.53
CamC		4.92	6.15	7.53	0.72	0.69	0.45
CamD		7.83	6.35	8.73	0.66	0.67	0.60

datasets, because the EO parameters for the nadir images are fixed, a set of quad-overlap tie points are used as checkpoints for further evaluation, which are observed in two nadir images and two oblique images captured with the same camera. At each checkpoint, we triangulate 3D points using the nadir and oblique observations separately and check their discrepancies. In addition, the 3D points from the nadir images are back-projected onto the oblique images, the discrepancies in the image coordinates are evaluated to reveal the BA result in image space. Then, for the entire datasets, we generate point clouds using SGM (Rothermel et al., 2012) and compare the profiles at several typical locations. For the Dortmund dataset, because GCPs and CPs are also provided, we follow the evaluation protocol of the benchmark requirements (Nex et al., 2015), which include evaluation of the image re-projection errors and CPs errors for the time being.

4.2. Result analysis of the Jinyang datasets

4.2.1. Analysis of the results of the bundle adjustment

The theoretical accuracies of the two BA blocks calculated using the method described by Yuan (Yuan, 2008) are shown in Tables 2 and 3. As described in our previous study (Hu et al., 2015), *CamA* through *CamD* are the four oblique cameras, and *CamE* is the nadir camera. In both of the blocks, the theoretical accuracies of the EO parameters are satisfactory. The RMSEs of the positioning parameters are on the order of centimeters, and those of the rotation parameters are less than 1/1000°. From Tables 4–7, we have compared the different BA strategies, including (1) "rigid", only estimate the EO parameters for the nadir images and obtain those for the oblique images using installation parameters; (2) "classical", estimate EO parameters for all the images using identical weights; (3) "asymmetrical", estimate EO parameters for all the images using asymmetrical weights.

Table 4
Back-projection errors of the ground control points in the Jinyang urban area after classical and asymmetrical re-weighting BAs. Rigid: using calibrated installation parameter for oblique EO parameters; Classical: using the classical BA method; Asymmetrical: using the asymmetric re-weighting BA method; Improvement: relative accuracy improvement of the proposed method.

Camera	Method	*x* RMSE (*pixels*)	*y* RMSE (*pixels*)	RMSE (*pixels*)	Maximum (*pixels*)
CamA	Rigid	7.53	2.69	7.99	12.77
	Classical	0.38	0.50	0.63	1.68
	Asymmetrical	0.36	0.48	0.60	1.60
	Improvement	***5.15%***	***4.26%***	***4.58%***	***4.90%***
CamB	Rigid	7.59	6.10	9.74	14.95
	Classical	0.34	0.64	0.72	2.01
	Asymmetrical	0.31	0.59	0.66	2.00
	Improvement	***8.78%***	***7.63%***	***7.88%***	***0.42%***
CamC	Rigid	11.21	5.93	12.69	16.39
	Classical	0.42	0.75	0.86	2.96
	Asymmetrical	0.40	0.75	0.85	2.87
	Improvement	***4.93%***	***0.26%***	***1.37%***	***2.94%***
CamD	Rigid	6.64	5.18	8.42	15.85
	Classical	0.38	0.69	0.79	2.98
	Asymmetrical	0.35	0.63	0.72	2.75
	Improvement	***9.37%***	***9.04%***	***9.12%***	***7.56%***
Average improvement		***7.06%***	***5.30%***	***5.74%***	***3.96%***

Table 5
Errors in the 3D positions of the ground control points in the Jinyang urban area after classical and asymmetrical re-weighting BAs. The second row of each method denotes the relative accuracy with respect to the flight altitude. Rigid: using calibrated installation parameter for oblique EO parameters; Classical: using the classical BA method; Asymmetrical: using the asymmetric re-weighting BA method; Improvement: relative accuracy improvement of the proposed method.

Camera	Method	*XY* RMSE (mm)	*Z* RMSE (mm)	RMSE (mm)	Maximum (mm)
CamA	Rigid	875.35	976.24	1311.21	3326.29
		1.17E−03	1.30E−03	1.75E−03	4.44E-03
	Classical	53.43	62.26	82.05	257.71
		7.12E−05	8.30E−05	1.09E−04	3.44E-04
	Asymmetrical	50.92	61.71	80.01	239.29
		6.79E−05	8.23E−05	1.07E−04	3.19E-04
	Improvement	***4.70%***	***0.89%***	***2.49%***	***7.15%***
CamB	Rigid	497.60	468.85	683.68	2067.72
		6.63E−04	6.25E−04	9.12E−04	2.76E-03
	Classical	74.56	90.99	117.64	297.55
		9.94E−05	1.21E−04	1.57E−04	3.97E-04
	Asymmetrical	63.95	77.06	100.14	242.54
		8.53E−05	1.03E−04	1.34E−04	3.23E-04
	Improvement	***14.23%***	***15.31%***	***14.87%***	***18.49%***
CamC	Rigid	1432.05	256.55	1454.85	2053.39
		1.91E−03	3.42E−04	1.94E−03	2.74E-03
	Classical	65.73	64.73	92.25	220.50
		8.76E−05	8.63E−05	1.23E−04	2.94E-04
	Asymmetrical	63.81	57.32	85.77	208.11
		8.51E−05	7.64E−05	1.14E−04	2.77E-04
	Improvement	***2.93%***	***11.45%***	***7.03%***	***5.62%***
CamD	Rigid	866.35	1001.49	1324.21	2717.22
		1.16E−03	1.34E−03	1.77E−03	3.62E-03
	Classical	87.44	104.23	136.05	349.81
		1.17E−04	1.39E−04	1.81E−04	4.66E-04
	Asymmetrical	73.77	76.89	106.56	270.07
		9.84E−05	1.03E−04	1.42E−04	3.60E-04
	Improvement	***15.64%***	***26.23%***	***21.68%***	***22.80%***
Average improvement		***9.38%***	***13.47%***	***11.52%***	***13.51%***

Table 6
Back-projection errors of the ground control points in the Jinyang rural area after classical and asymmetrical re-weighting BAs. Rigid: using calibrated installation parameter for oblique EO parameters; Classical: using the classical BA method; Asymmetrical: using the asymmetric re-weighting BA method; Improvement: relative accuracy improvement of the proposed method.

Camera	Method	*x* RMSE (*pixels*)	*y* RMSE (*pixels*)	RMSE (*pixels*)	Maximum (*pixels*)
CamA	Rigid	7.53	6.01	9.63	14.87
	Classical	0.37	0.78	0.86	2.36
	Asymmetrical	0.33	0.71	0.78	2.15
	Improvement	***9.00%***	***8.94%***	***8.95%***	***9.20%***
CamB	Rigid	5.13	5.80	7.75	11.74
	Classical	0.32	0.54	0.62	1.63
	Asymmetrical	0.30	0.52	0.60	1.55
	Improvement	***4.97%***	***2.84%***	***3.39%***	***4.88%***
CamC	Rigid	10.64	3.27	11.13	16.19
	Classical	0.32	0.74	0.80	2.80
	Asymmetrical	0.31	0.68	0.75	2.78
	Improvement	***4.67%***	***6.95%***	***6.58%***	***0.67%***
CamD	Rigid	6.61	4.59	8.05	15.64
	Classical	0.46	0.63	0.78	2.33
	Asymmetrical	0.42	0.61	0.74	2.07
	Improvement	***7.33%***	***3.31%***	***4.68%***	***11.06%***
Average improvement		***6.49%***	***5.51%***	***5.90%***	***6.45%***

In the urban test area, the RMSE of unit weight in image space is 0.25 pixels. From Table 4, it is clear that for all of the oblique cameras, the standard and asymmetric re-weighting BAs both improve the internal consistency of the block, and the image back-projection errors are significantly less after the BA. The proposed method is better than the standard one; the improvement in the ratio in image space ranges from 1.37% to 9.12% for the different cameras; its average is 5.74%. Because the 3D positions of the checkpoints are determined using the forward intersections of the two nadir observations, which results in a less accurate height, the RMSEs in image space at the checkpoints are greater than the estimated RMSE of unit weight (sigma naught). The 3D position errors at the checkpoints for each method are listed in Table 5. The first row lists the absolute value of the RMSE, and the second row lists their relative accuracies (the RMSE divided by the relative flight altitude). According to Table 5, after an asymmetric re-weighting BA, the horizontal RMSE in object space is less than 0.08 m and the vertical RMSE is less than 0.11 m for all of the oblique cameras. The accuracy of the empirical point determination in terms of the flight altitude (*FA*) after an asymmetric BA is better than 1.5×10^{-4} *FA*, which is superior to that of the standard method by an average of 11.52%. A comparison of their maximum residuals showed more improvement, which implies that the proposed asymmetric re-weighting method is more balanced than the standard method.

Table 7
Errors in the 3D positions of the ground control points in the Jinyang rural area using classical and asymmetrical re-weighting BAs. The second row of each method denotes the relative accuracy with respect to the flight altitude. Rigid: using calibrated installation parameter for oblique EO parameters; Classical: using the classical BA method; Asymmetrical: using the asymmetric re-weighting BA method; Improvement: relative accuracy improvement of the proposed method.

Camera	Method	*XY* RMSE (mm)	*Z* RMSE (mm)	RMSE (mm)	Maximum (mm)
CamA	Rigid	641.56	724.89	968.02	1383.20
		8.55E−04	9.67E−04	1.29E−03	1.84E-03
	Classical	64.41	94.21	114.12	307.84
		8.59E−05	1.26E−04	1.52E−04	4.10E-04
	Asymmetrical	63.19	80.49	102.34	280.73
		8.43E−05	1.07E−04	1.36E−04	3.74E-04
	Improvement	***1.88%***	***14.55%***	***10.32%***	***8.81%***
CamB	Rigid	692.86	856.21	1101.43	2472.64
		9.24E−04	1.14E−03	1.47E−03	3.30E-03
	Classical	67.32	75.10	100.85	255.88
		8.98E−05	1.00E−04	1.34E−04	3.41E-04
	Asymmetrical	66.55	73.92	99.47	248.41
		8.87E−05	9.86E−05	1.33E−04	3.31E-04
	Improvement	***1.14%***	***1.57%***	***1.38%***	***2.92%***
CamC	Rigid	1304.54	322.20	1343.74	1744.33
		1.74E−03	4.30E−04	1.79E−03	2.33E-03
	Classical	56.76	77.06	95.71	205.58
		7.57E−05	1.03E−04	1.28E−04	2.74E-04
	Asymmetrical	52.31	62.41	81.44	202.07
		6.98E−05	8.32E−05	1.09E−04	2.69E-04
	Improvement	***7.84%***	***19.01%***	***14.92%***	***1.71%***
CamD	Rigid	1218.18	1296.11	1778.72	3214.14
		1.62E−03	1.73E−03	2.37E−03	4.29E-03
	Classical	103.02	112.58	152.60	354.91
		1.37E−04	1.50E−04	2.03E−04	4.73E-04
	Asymmetrical	88.21	98.81	132.45	321.33
		1.18E−04	1.32E−04	1.77E−04	4.28E-04
	Improvement	***14.37%***	***12.24%***	***13.20%***	***9.46%***
Average improvement		***6.31%***	***11.84%***	***9.95%***	***5.72%***

In the rural empirical test area, the RMSE of unit weight in image space is 0.27 pixels. The results are shown in Tables 6 and 7. The proposed method performs better than the standard method, as it did in the urban area. In image space, the proposed method reduces back-projection errors by 5.90%, and in object space, the average ratio is improved by 9.95%. In this area, the horizontal RMSE after an asymmetric re-weighting BA is less than 0.11 m and the vertical RMSE is less than 0.14 m for all four oblique cameras, and the empirical point determination accuracy is better than 1.8×10^{-4} *FA*.

Note that, in the urban area, the improvements for *CamB* and *CamD* are greater than those for *CamA* and *CamC*, while the improvements for *CamA*, *CamB* and *CamD* are more remarkable in the rural area. Except for intrinsic qualities of the cameras, this type of difference between different cameras is a product of the asymmetrical weight distribution. Tie points on the top of high buildings have greater weights, for example, at a flight altitude of 750 m, and the F_S of object points on the roof of a building 100 m tall is distinctly greater than that of object points on the ground. In addition, the vehicle's attitude influences the photographs captured at each station because of its effect on the angles of the different cameras, which are different and even opposite; for example, enlarging the tilt angle of the front camera requires minimizing that of the back camera. After all, even in the worst situation, the proposed method did not provide a worse result than the standard method. In both the areas, improvement in object space is larger than in image space, which imply a better 3D measurability for oblique images achieved by the proposed method, which is crucial to subsequent productions, such as DIM and reconstruction.

Figs. 10 and 11 show the discrepancy vectors at the checkpoints for all four oblique cameras in the urban area and the rural area, respectively. Note that there are no significant systematic errors in either block. The results of the BA are better in the urban area than in the rural area due to the varying quality of the tie point matching. Generally, tie points in the urban area overlap more. There is a stream flowing across the rural test area shown in Fig. 13 that results in fewer matched points. In addition, in the rural area, object points located at the crowns of trees may not remain static due to the influence of the wind. In both of the areas, the vast majority of planar and vertical discrepancies at checkpoints are less than 0.2 m, which shows that the proposed method provides good results.

4.2.2. Analysis using dense matching point clouds

One major application of oblique images is DIM-based 3D mapping. To evaluate the feasibility of BA using the proposed method, the DIM point clouds generated from the oriented oblique images and nadir images are examined. Only adjacent photos captured by cameras pointed in the same direction are selected to perform DIM. We evaluate profiles of point clouds derived from all five orientations (one vertical and four oblique) to reveal potential systematic errors. The width of each profile is set to 0.08 m, which is the average GSD of the nadir cameras. In each area, four groups of profiles

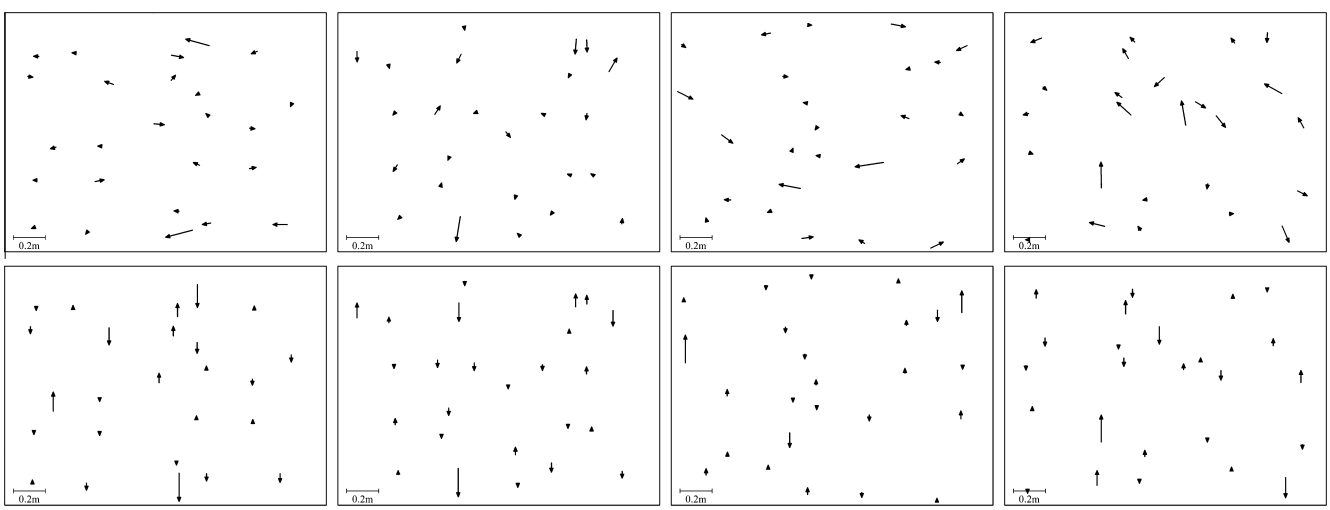

Fig. 10. Discrepancy vectors for the GCPs in the urban area after bundle adjustment. From left to right are discrepancy plots for the four oblique cameras, *CamA* through *CamD*. The first row represents the area discrepancies and the second row represents the vertical discrepancies.

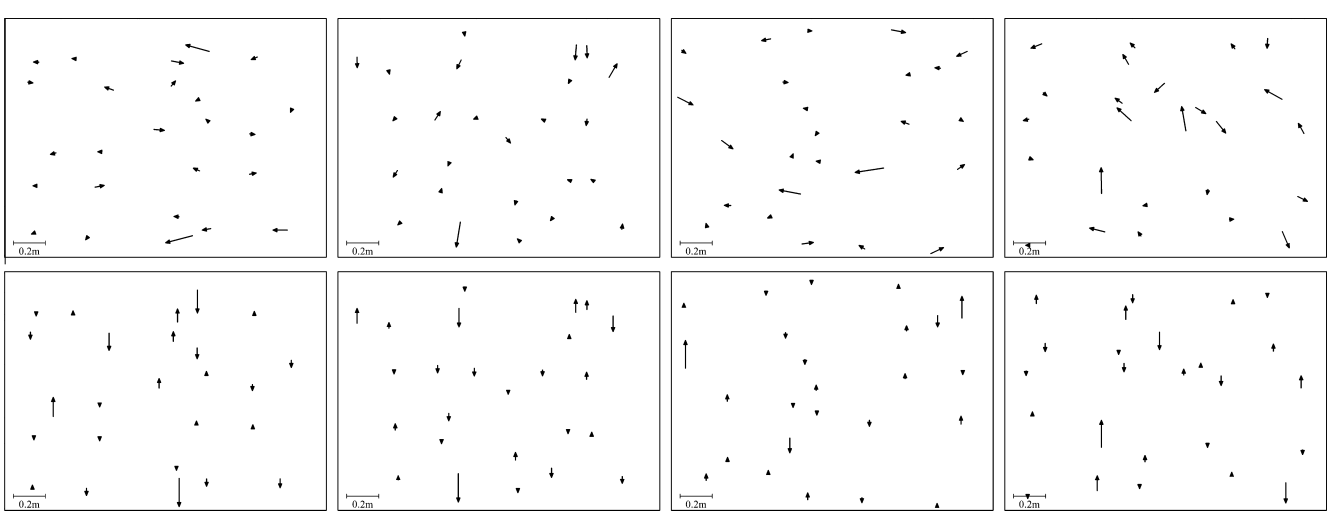

Fig. 11. Discrepancy vectors for the GCPs in the rural area after bundle adjustment. From left to right are discrepancy plots for the four oblique cameras, *CamA* through *CamD*. The first row represents the area discrepancies and the second row represents the vertical discrepancies.

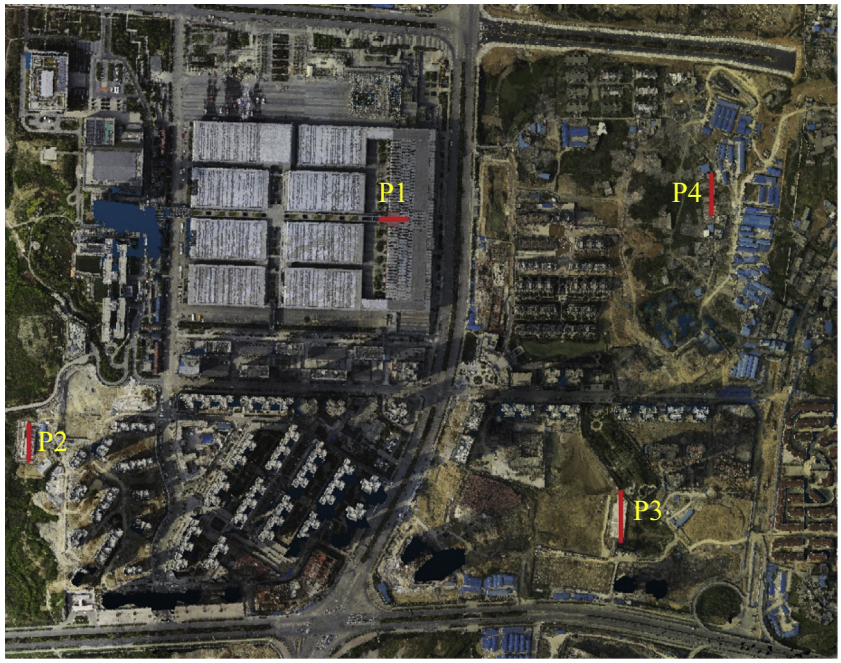

Fig. 12. A map of the profiles selected for the DIM point cloud comparisons made in an urban area. The red lines indicate the four selected regions (P1 to P4). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

are selected for the comparison; distribution maps of the selected profiles are shown in Figs. 12 and 13.

Comparison diagrams for the urban and rural areas are shown in Figs. 14 and 15, respectively. Comparing Fig. 14(a)–(d) and (e)–(h), we find that before the combined BA, the roofs of buildings have different altitudes according to different cameras, and the misalignment is up to 3 m. Using the DIM point clouds observed from five directions, we can infer that the ambiguity of the surface patches poses a substantial obstacle to reconstructing the buildings with a certain geometric accuracy. After performing a combined BA using the proposed method, the discrepancies between the DIM point clouds observed from different directions are eliminated in all four profiles, and the horizontal structures of the buildings are clear, which is favorable for 3D applications such as building extraction and reconstruction. In Fig. 15(a)–(d), there are systematic unconformities in the ground surface as viewed from different cameras' DIM point clouds; this situation improved significantly after the combined adjustment using the proposed method; see Fig. 15(e)–(h). In both of the test areas, the internal consistencies of the DIM point clouds generated using different cameras are generally promoted due to the high accuracy achieved by combined BA using the proposed method.

4.3. Result analysis of the Zurich dataset

4.3.1. Analysis of the results of the bundle adjustment

The theoretical accuracies of this block after BA are shown in Table 8. In this block, the theoretical accuracies of the EO parameters are as well as those of the first two blocks. In Tables 9 and 10, we have compared the different BA strategies using the same strategy as Tables 4 and 5. In the Zurich test area, the RMSE of unit weight in image space is 0.42 pixels. In Tables 9 and 10, we can notice improvements in both image space and object space that achieved by the proposed method compared to the standard one. The improve ratio in image space ranges from 4.19% to 7.27% with an average of about 5% while that in object space ranges from 2.65% to 11.70% with an average of about 6%. Compared to the first two blocks, the improvement ratio in this block is lesser. This is probably because the configuration tilt angles of oblique images in this block are 35°, which are smaller than those in the first two blocks. According to Table 10, after an asymmetric re-weighting BA, the horizontal RMSE in object space is less than 0.16 m and the vertical RMSE is less than 0.18 m for all of the oblique cameras, which have shown slight improvement than the classical method, but also less significant than the accuracies in the first two block. There are two reasons for this result: first, the orientation accuracies of nadir EO parameters are not as good as those of the Jinyang datasets; second, since the Zurich datasets are designed for dense image matching, the margin of some oblique images have been masked, which lead to fewer and not well distributed auto-matched points. However, under the situations above, the proposed asymmetric re-weighting method still outperform the classical BA method.

Fig. 16 shows the discrepancy vectors at the checkpoints for all four oblique cameras in the Zurich area, respectively. Similar to the first two blocks no significant systematic errors in this block, and the majority of planar and vertical discrepancies at checkpoints are less than 0.2 m.

4.3.2. Analysis using dense matching point clouds

Similar to the Jinyang datasets, four groups of profiles have been selected in the Zurich area to evaluate the DIM point clouds generated after asymmetric re-weighting BA. The width of each profile is set to 0.06 m, which equals to the average GSD of the nadir camera. The distribution map of the selected profiles is shown in Fig. 17. Comparison diagrams for the Zurich areas are shown in Fig. 18. Comparing Fig. 18(a)–(d) and (e)–(h), it is obvious to see that before the combined BA, there exit distinct deviations between DIM point clouds from different looking directions, which are generated by the Gaussian noises. The position of eaves in object space have apparent uncertainty especially in Z direction. After conducting a combined BA using the proposed method, the discrepancies between different DIM point clouds have been largely relieved, the plane structure could be clearly recognized in the profiles which imply the satisfactory 3D accuracies achieved by the combined BA. The experiment results in this area are consistent with the situations in the Jinyang datasets, thus the validity of the proposed BA method is proven again.

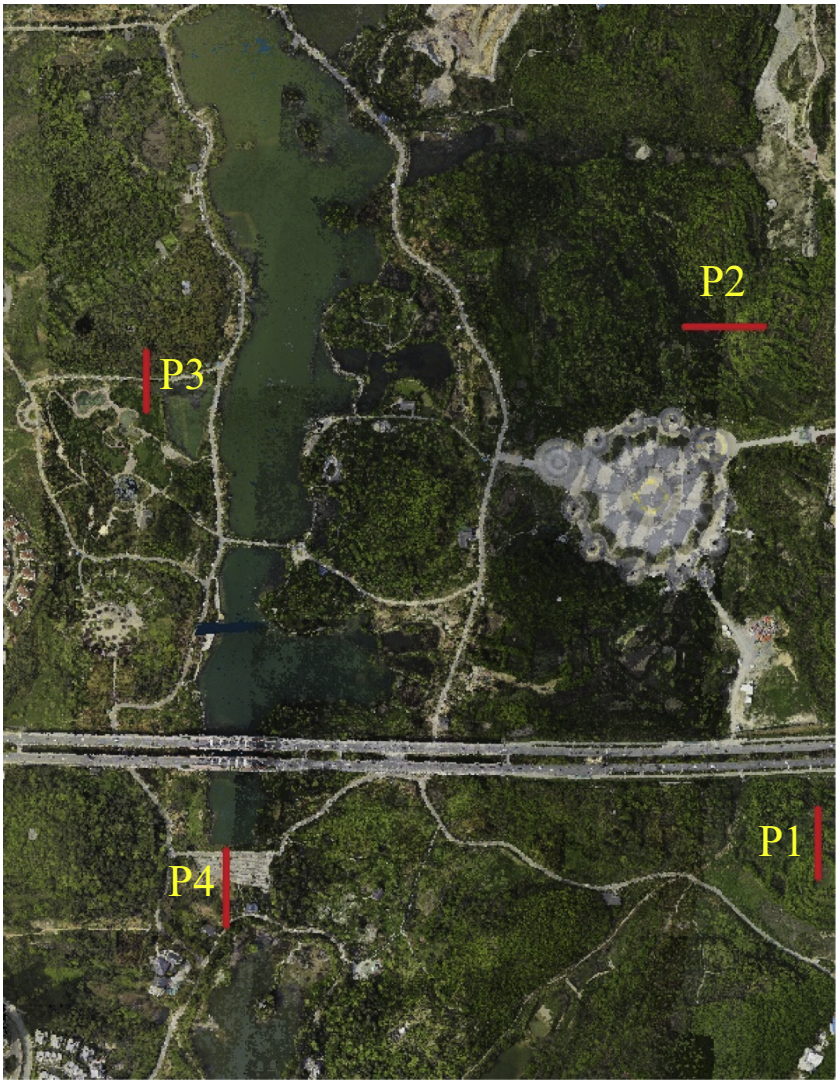

Fig. 13. A map of the profiles selected for the DIM point cloud comparisons made in a rural area. The red lines indicate the four selected regions (P1 to P4). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

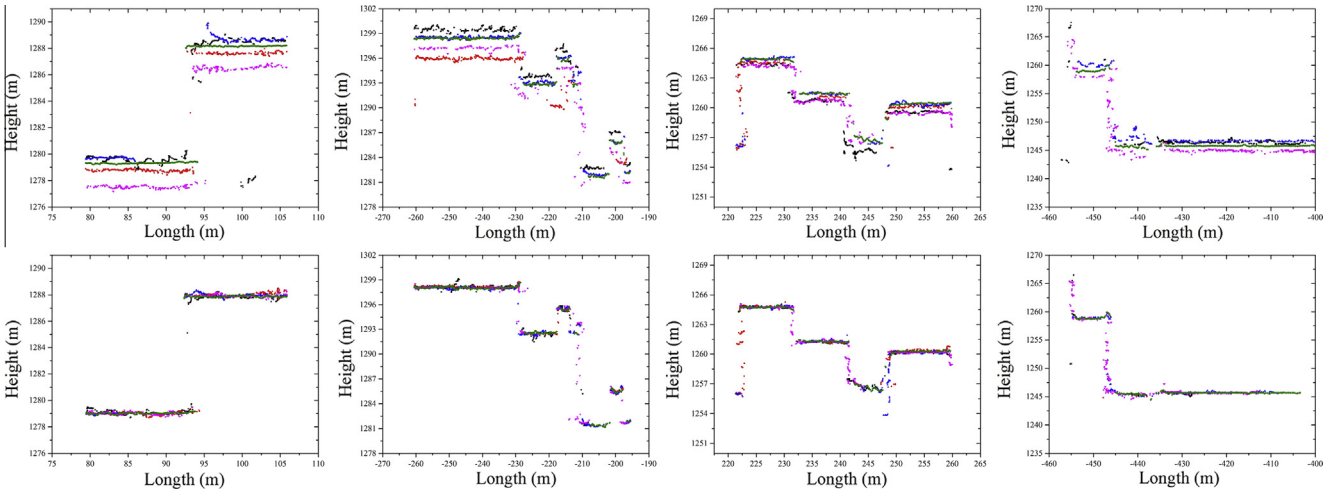

Fig. 14. Comparisons of the DIM point cloud profiles for the five cameras in the urban area. From left to right are the four test areas marked in Fig. 12 as P1 through P4. All of the points' coordinates are subtracted from the coordinates of the center block. The upper row shows the profiles before bundle adjustment, and the lower row show the profiles after bundle adjustment using the asymmetric re-weighting method. In all eight diagrams, black represents *CamA*, red represents *CamB*, blue represents *CamC*, pink represents *CamD*, and green represents *CamE*. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

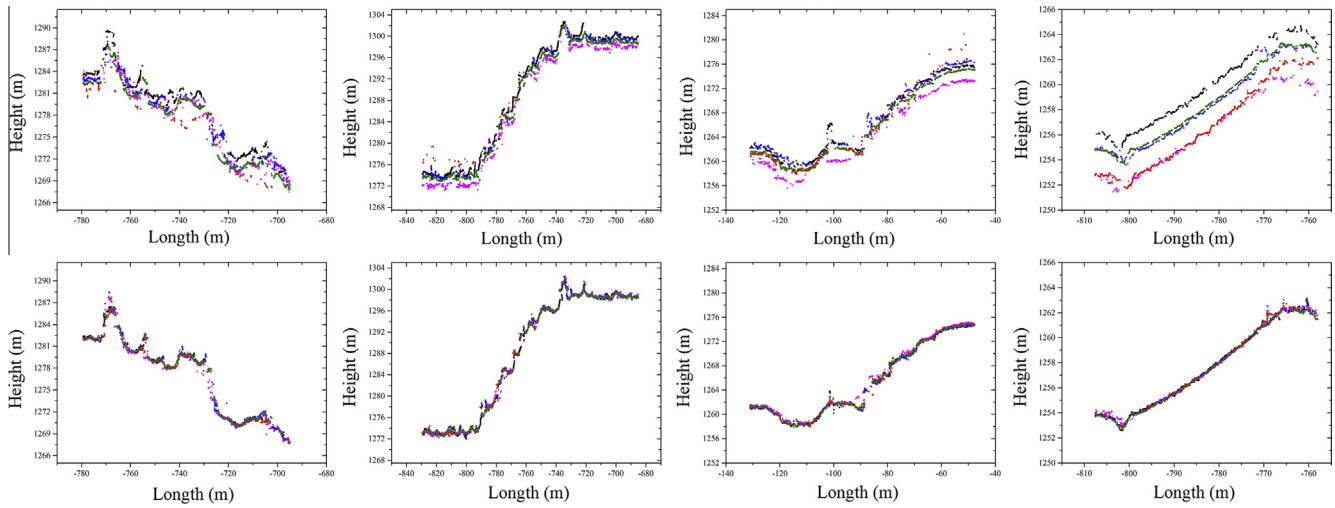

Fig. 15. Comparisons of the DIM point cloud profiles for the five cameras in the rural area. From left to right are the four test areas marked in Fig. 13 as P1 through P4. All of the points' coordinates are subtracted from the coordinates of the center block. The upper row shows the profiles before bundle adjustment, and the lower row shows the profiles after bundle adjustment using the asymmetric re-weighting method. In all eight diagrams, black represents *CamA*, red represents *CamB*, blue represents *CamC*, pink represents *CamD*, and green represents *CamE*. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 8
The theoretical BA accuracy in the Zurich area.

Camera	σ_0 (*pixels*)	*X* (mm)	*Y* (mm)	*Z* (mm)	φ (1/1000°)	φ (1/1000°)	φ (1/1000°)
CamA	0.42	7.35	8.06	9.94	0.85	0.69	0.55
CamB		8.10	7.84	8.43	0.65	0.52	0.58
CamC		5. 25	6.76	8.62	0.66	0.66	0.52
CamD		8.49	7.90	9.47	0.61	0.58	0.61

Table 9
Back-projection errors of the ground control points in the Zurich area after classical and asymmetrical re-weighting BAs. No adjustment: using the EO released with additional Gaussian noise; Classical: using the classical BA method; Asymmetrical: using the asymmetric re-weighting BA method; Improvement: relative accuracy improvement of the proposed method.

Camera	Method	*x* RMSE (*pixels*)	*y* RMSE (*pixels*)	RMSE (*pixels*)	Maximum (*pixels*)
CamA	No adjustment	0.94	0.94	1.33	2.81
	Classical	0.59	0.62	0.86	2.35
	Asymmetrical	0.57	0.56	0.79	2.28
	Improvement	***3.78%***	***10.51%***	***7.27%***	***2.89%***
CamB	No adjustment	0.86	1.11	1.41	3.04
	Classical	0.76	0.81	1.11	2.66
	Asymmetrical	0.70	0.78	1.05	2.66
	Improvement	***7.41%***	***3.08%***	***5.08%***	***0.14%***
CamC	No adjustment	0.75	0.86	1.14	2.57
	Classical	0.61	0.64	0.88	2.30
	Asymmetrical	0.58	0.62	0.85	1.98
	Improvement	***6.11%***	***2.43%***	***4.19%***	***13.93%***
CamD	No adjustment	0.92	1.02	1.37	2.92
	Classical	0.72	0.82	1.09	2.67
	Asymmetrical	0.71	0.74	1.03	2.53
	Improvement	***1.70%***	***9.36%***	***5.93%***	***5.09%***
Average improvement		***4.75%***	***6.35%***	***5.62%***	***5.51%***

Table 10
Errors in the 3D positions of the ground control points in the Zurich area using classical and asymmetrical re-weighting BAs. The second row of each method denotes the relative accuracy with respect to the flight altitude. No adjustment: using the EO released with additional Gaussian noise; Classical: using the classical BA method; Asymmetrical: using the asymmetric re-weighting BA method; Improvement: relative accuracy improvement of the proposed method.

Camera	Method	*XY* RMSE (mm)	*Z* RMSE (mm)	RMSE (mm)	Maximum (mm)
CamA	No adjustment	113.39	93.42	220.94	448.14
		2.18E−04	1.80E−04	4.25E−04	8.62E-04
	Classical	81.65	143.84	165.39	387.13
		1.57E−04	2.77E−04	3.18E−04	7.44E-04
	Asymmetrical	76.12	125.85	147.09	354.36
		1.46E−04	2.42E−04	2.83E−04	6.81E-04
	Improvement	***6.77%***	***12.50%***	***11.07%***	***8.47%***
CamB	No adjustment	150.13	249.89	291.52	592.24
		2.89E−04	4.81E−04	5.61E−04	1.14E-03
	Classical	101.99	166.29	195.07	416.81
		1.96E−04	3.20E−04	3.75E−04	8.02E-04
	Asymmetrical	97.28	152.71	181.06	411.32
		1.87E−04	2.94E−04	3.48E−04	7.91E-04
	Improvement	***4.62%***	***8.17%***	***7.19%***	***1.32%***
CamC	No adjustment	106.07	196.75	223.51	543.87
		2.04E−04	3.78E−04	4.30E−04	1.05E-03
	Classical	78.47	139.37	159.94	426.42
		1.51E−04	2.68E−04	3.08E−04	8.20E-04
	Asymmetrical	74.42	136.76	155.69	419.23
		1.43E−04	2.63E−04	2.99E−04	8.06E-04
	Improvement	***5.16%***	***1.87%***	***2.65%***	***1.69%***
CamD	No adjustment	135.22	239.62	275.14	895.85
		2.60E−04	4.61E−04	5.29E−04	1.72E-03
	Classical	97.90	156.58	184.67	431.36
		1.88E−04	3.01E−04	3.55E−04	8.30E-04
	Asymmetrical	93.07	151.99	178.22	398.16
		1.79E−04	2.92E−04	3.43E−04	7.66E-04
	Improvement	***4.94%***	***2.93%***	***3.49%***	***7.70%***
Average improvement		***5.37%***	***6.37%***	***6.10%***	***4.79%***

4.4. Result analysis of the Dortmund dataset

In this block, we conduct the evaluation protocol by the benchmark (Nex et al., 2015). After asymmetric re-weighting BA, the RMSE of unit weight in image space is 0.52 *pixels*. In order to

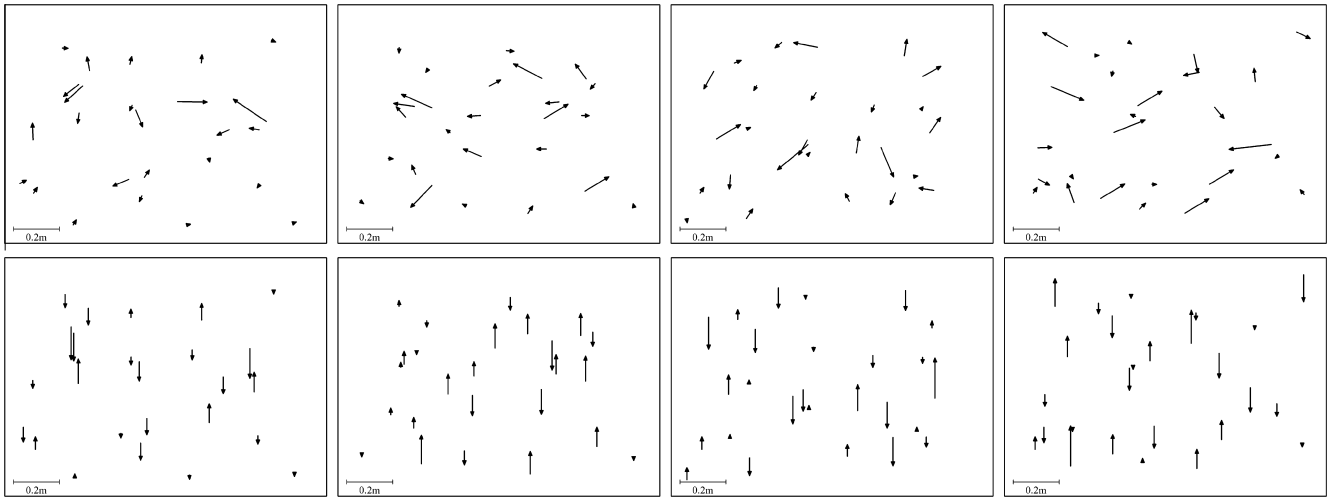

Fig. 16. Discrepancy vectors for the GCPs in the Zurich area after bundle adjustment. From left to right are discrepancy plots for the four oblique cameras, *CamA* through *CamD*. The first row represents the area discrepancies and the second row represents the vertical discrepancies.

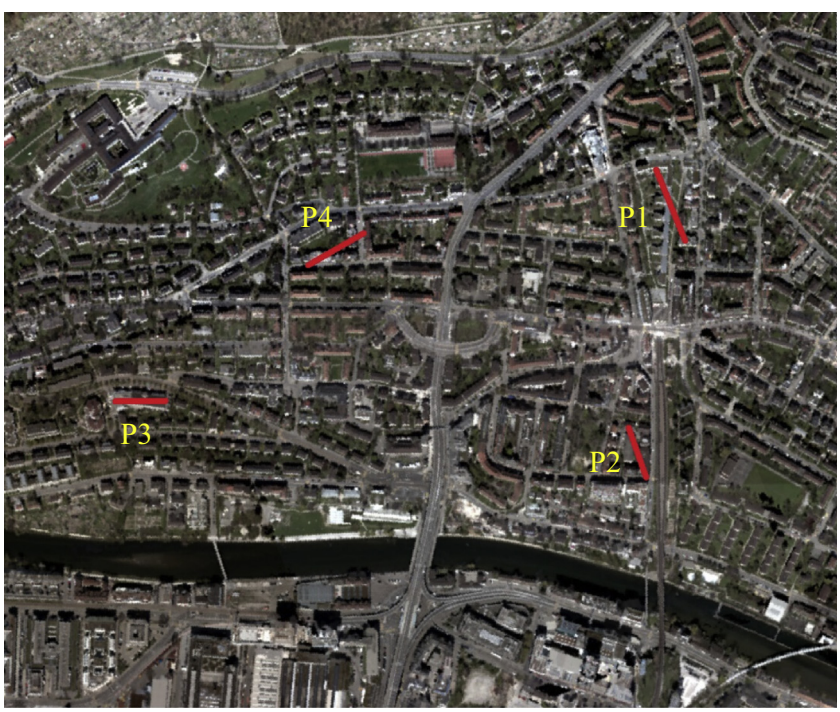

Fig. 17. A map of the profiles selected for the DIM point cloud comparisons made in the Zurich area (Cavegn et al., 2014). The red lines indicate the four selected regions (P1 to P4). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

explore the potential systematic errors related to different cameras, we grouped the re-projection errors by the five different cameras. As shown in Table 11, when grouped by cameras, the average re-projection error ranges from 0.276 *pixels* to 0.325 *pixels* while the standard deviations have a min value of 0.399 *pixels* and a max value of 0.455 *pixels*. The maximum re-projection error is beyond the value of 2.6 *pixels* for all the five cameras. The differences between the four tilt configuration cameras are unconspicuous which indicates a balanced result achieved by the proposed BA. The average and standard deviation of the nadir camera is slightly larger than that of oblique ones, the reason is that the nadir camera has a shorter focal length.

We have submitted our BA result to the benchmark sponsor and the statistical information (Table 12) at checkpoints and the residual plot (Fig. 19) have been return. As shown in Table 12, the RMSE in X and Y direction is 0.069 m and 0.124 m respectively, stating a good planar accuracy achieved by the proposed BA method. Notice that the RMSE in Z direction is 0.260 m while the mean of discrepancies in height is about −0.207 m, which imply that there possibly exist some systematic errors in the vertical direction. Consider that the GCPs were intentional not well distributed, the 3D RMSE after our combined BA is still above 0.3 m, demonstrating the high reliability of the proposed method.

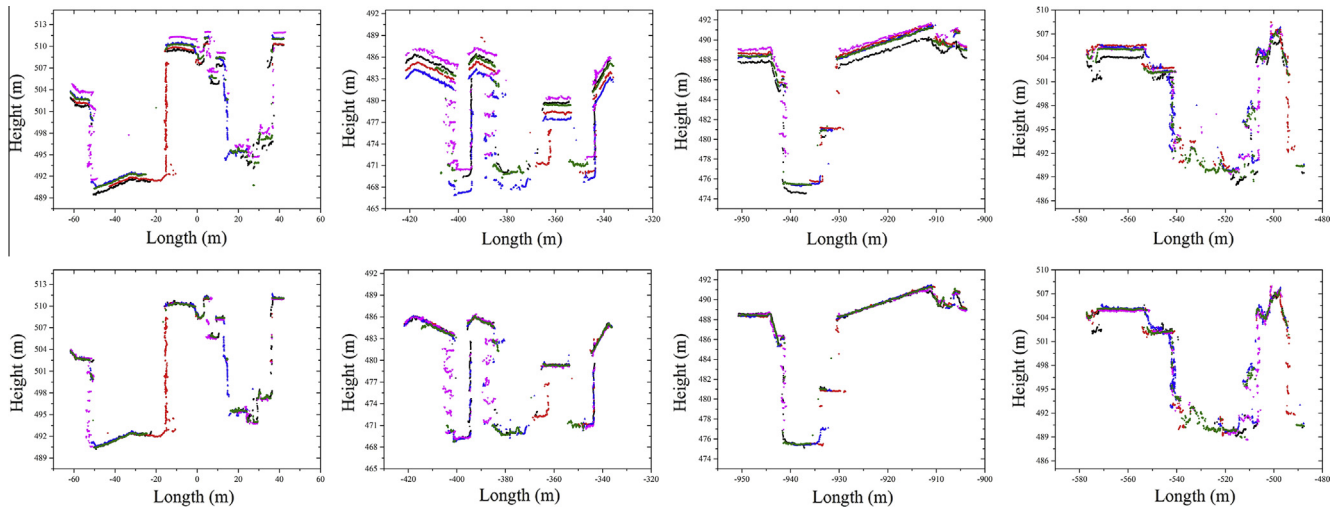

Fig. 18. Comparisons of the DIM point cloud profiles for the five cameras in the Zurich area. From left to right are the four test areas marked in Fig. 17 as P1 through P4. All of the points' coordinates are subtracted from the coordinates of the center block. The upper row shows the profiles before bundle adjustment, and the lower row shows the profiles after bundle adjustment using the asymmetric re-weighting method. In all eight diagrams, black represents *CamA*, red represents *CamB*, blue represents *CamC*, pink represents *CamD*, and green represents *CamE*. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 11
Statistics of re-projection error per camera of the Dortmund dataset.

Camera	Average (*pixel*)	Std dev (*pixel*)	Maximum (*pixel*)
Front	0.276	0.399	2.306
Right	0.286	0.410	2.515
Back	0.317	0.438	2.332
Left	0.305	0.429	2.354
Nadir	0.325	0.455	2.466

Table 12
Mean, standard deviation and RMSE at checkpoints of the Dortmund dataset.

	X (m)	*Y* (m)	*Z* (m)	*XYZ* (m)
Mean	−0.026	0.109	−0.207	0.236
Std dev	0.065	0.059	0.161	0.183
RMSE	0.069	0.124	0.26	0.296

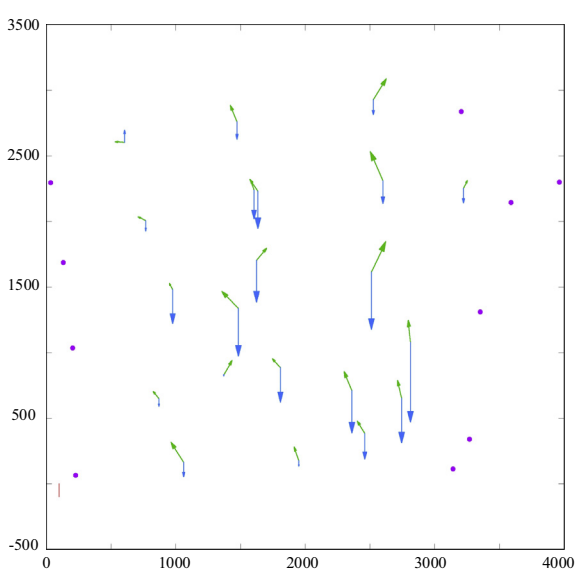

Fig. 19. Residual plot at checkpoints of the Dortmund dataset. The purple dots represent GCPs; the green arrows represent horizontal residual vectors while the blue arrows represent vertical residual vectors. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

5. Conclusions

An asymmetric re-weighting method for precise combined BA is presented for more reliable processing of oblique aerial images. This method determines an *a priori* weight for an image observation based on its scale and the amount of anisotropy in its GSDs, which varies markedly between different oblique views. Using four blocks of images acquired by three different multi-camera systems which include both urban and rural scenes, it is shown that the proposed method's performance is superior to that of the standard method in terms of practical BA accuracy. By analyzing the residual distributions at checkpoints and comparing the profiles of different cameras' DIM point clouds, we testify that our method results in high-quality adjustments, laying a good foundation for various applications of oblique images. Following the ISPRS benchmark for multi-platform photogrammetry, the satisfying results demonstrated that the proposed method is capable of handling large airborne multi-view image blocks.

One major limitation of the proposed work is that we need a relatively accurate initial information of the entire scene. Although for aerial oblique images, GNSS/IMU systems could be obtained easily for these information, this limitation has impeded the extension of the proposed method to terrestrial photogrammetry.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (Nos. 41471320 and 41501421) and the Technology Plan Projects of Sichuan (No. 2014SZ0106). The authors would like to acknowledge the provision of the datasets by ISPRS and EuroSDR, released in conjunction with the ISPRS scientific initiative 2014 and 2015, lead by ISPRS ICWG I/Vb.

References

Appa, G., Smith, C., 1973. On L1 and Chebyshev estimation. Math. Program. 5 (1), 73–87.
Baselga, S., 2007. Global optimization solution of robust estimation. J. Surv. Eng. 133 (3), 123–128.
Cavegn, S., Haala, N., Nebiker, S., Rothermel, M., Tutzauer, P., 2014. Benchmarking high density image matching for oblique airborne imagery. Int. Arch. Photogramm., Remote Sens. Spatial Inform. Sci., Zürich 3, 45–52.
Cen, M., Li, Z., Ding, X., Zhuo, J., 2003. Gross error diagnostics before least squares adjustment of observations. J. Geodesy 77 (9), 503–513.

Fritsch, D., Rothermel, M., 2013. Oblique image data processing – potential, experience and recommendation. Photogrammetric Week 2013, Stuttgart, Germany, pp. 73–88.
Fritsch, D., Kremer, J., Grimm, A., 2012. Towards all-in-one photogrammetry. GIM Int. 26 (4), 18–23.
Gerke, M., 2009. Dense matching in high resolution oblique airborne images. Int. Arch. Photogramm., Remote Sens. Spatial Inform. Sci. 38 (3/W4), 77–82.
Gerke, M., Kerle, N., 2011. Automatic structural seismic damage assessment with airborne oblique pictometry imagery. Photogramm. Eng. Remote Sens. 77 (9), 885–898.
Gerke, M., 2011. Using horizontal and vertical building structure to constrain indirect sensor orientation. ISPRS J. Photogramm. Remote Sens. 66 (3), 307–316.
Grenzdörffer, G.J., Guretzki, M., Friedlander, I., 2008. Photogrammetric image acquisition and image analysis of oblique imagery. Photogramm. Rec. 23 (124), 372–386.
Gruen, A., 2012. Development and status of image matching in photogrammetry. Photogramm. Rec. 27 (137), 36–57.
Gülch, E., 2012. Photogrammetric measurements in fixed wing UAV imagery. Int. Arch. Photogramm., Remote Sens. Spatial Inform. Sci. 39 (B1), 381–386.
Heipke, C., Jacobsen, K., Wegmann, H., 2002. Analysis of the results of the OEEPE test "Integrated Sensor Orientation". OEEPE Official Publication, 43, pp. 31–49.
Höhle, J., 2008. Photogrammetric measurements in oblique aerial images. Photogrammetrie, Fernerkundung, Geoinform. 2008 (1), 7–14.
Hu, H., Zhu, Q., Du, Z., Zhang, Y., Ding, Y., 2015. Reliable spatial relationship constrained feature point matching of oblique aerial images. Photogramm. Eng. Remote Sens. 81 (1), 49–58.
Huber, P.J., 1971. Robust statistics. International Encyclopedia of Statistical Science, pp. 1248–1251.
Ip, A., El-Sheimy, N., Mostafa, M., 2007. Performance analysis of integrated sensor orientation. Photogramm. Eng. Remote Sens. 73 (1), 1–9.
Jacobsen, K., 2009. Geometry of vertical and oblique image combinations. In: Remote Sensing for a Changing Europe: Proceedings of the 28th Symposium of the European Association of Remote Sensing Laboratories, Istanbul, Turkey, 2–5 June 2008. Millpress Science Pub., p. 16.
Kada, M., 2009. The 3D Berlin Project, Photogrammetric Week. Wichmann Verlag, Heidelberg, pp. 331–340.
Klein, H., Förstner, W., 1984. Realization of automatic error detection in the block adjustment program PAT-M43 using robust estimators. Int. Arch. Photogramm. Remote Sens. 25, 234–245.
Leica-Geosystems, 2014. Leica RCD30 Oblique Specifications <http://www.leica-geosystems.com/downloads123/zz/airborne/RCD30_Oblique/Flyer/Leica_RCD30_Oblique_FLY_en.pdf> (accessed 24 September, 2015).
Nex, F. et al., 2015. ISPRS benchmark for multi-platform photogrammetry. ISPRS Ann. Photogramm., Remote Sens. Spatial Inform. Sci. II-3/W4, 135–142.
Nyaruhuma, A.P., Gerke, M., Vosselman, G., 2012a. Verification of 3D building models using mutual information in airborne oblique images. ISPRS Ann. Photogramm., Remote Sens. Spatial Inform. Sci. 1, 275–280.
Nyaruhuma, A.P., Gerke, M., Vosselman, G., Mtalo, E.G., 2012b. Verification of 2D building outlines using oblique airborne images. ISPRS J. Photogramm. Remote Sens. 71, 62–75.
Petrie, G., 2009. Systematic oblique aerial photography using multiple digital cameras. Photogramm. Eng. Remote Sens. 75 (2), 102–107.
Rothermel, M., Wenzel, K., Fritsch, D., Haala, N., 2012. Sure: photogrammetric surface reconstruction from imagery. In: Proceedings LC3D Workshop, Berlin, pp. 1–9.
Rottensteiner, F. et al., 2014. Results of the ISPRS benchmark on urban object detection and 3D building reconstruction. ISPRS J. Photogramm. Remote Sens. 93, 256–271.
Rupnik, E., Nex, F., Remondino, F., 2014. Oblique multi-camera systems – orientation and dense matching issues. Int. Arch. Photogramm., Remote Sens. Spatial Inform. Sci. 38 (W1), 107–114.
Rupnik, E., Nex, F., Toschi, I., Remondino, F., 2015. Aerial multi-camera systems: accuracy and block triangulation issues. ISPRS J. Photogramm. Remote Sens. 101, 233–246.
Rupnik, E., Nex, F., Remondinob, F., 2013. Automatic orientation of large blocks of oblique images. Int. Arch. Photogramm., Remote Sens. Spatial Inform. Sci. 40-1 (W1), 299–304.
Snavely, N., Seitz, S.M., Szeliski, R., 2008. Modeling the world from internet photo collections. Int. J. Comput. Vis. 80 (2), 189–210.
Tommaselli, A., Galo, M., Marcato Jr., J., Ruy, R.S., Lopes, R.F., 2010. Registration and fusion of multiple images acquired with medium format cameras. In: Proceedings of the Canadian Geomatics Conference 2010 and Symposium of Commission I, Calgary, AB, Canada, 15–18, June.
Triggs, B., McLauchlan, P.F., Hartley, R.I., Fitzgibbon, A.W., 2000. Bundle Adjustment—A Modern Synthesis. Springer, Berlin Heidelberg, pp. 153–177.
Wang, Z., 1990. Principle of Photogrammetry: With Remote Sensing. Press of Wuhan Technical University of Surveying and Mapping, Publishing House of Surveying and Mapping, Beijing, China.
Wiedemann, A., Moré, J., 2012. Orientation strategies for aerial oblique images. Int. Arch. Photogramm., Remote Sens. Spatial Inform. Sci. 39 (B1), 185–189.
Xiong, X., Zhang, Y., Zhu, J., Zheng, M., 2014. Camera pose determination and 3-D measurement from monocular oblique images with horizontal right angle constraints. IEEE Geosci. Remote Sens. Lett. 11 (11), 1976–1980.
Yuan, X., 2008. A novel method of systematic error compensation for a position and orientation system. Prog. Nat. Sci. 18 (8), 953–963.
Yuan, X., Fu, J., Sun, H., Toth, C., 2009. The application of GPS precise point positioning technology in aerial triangulation. ISPRS J. Photogramm. Remote Sens. 64 (6), 541–550.
Zach, C., 2014. Robust bundle adjustment revisited. Computer Vision – ECCV 2014. Springer International Publishing, pp. 772–787.
Zhang, Y., Hu, B., Zhang, J., 2011. Relative orientation based on multi-features. ISPRS J. Photogramm. Remote Sens. 66 (5), 700–707.