A Global Photogrammetry-Based Structure from Motion Framework: Application in Oblique Aerial Images

Styliani VERYKOKOU and Charalabos IOANNIDIS, Greece

Key words: Photogrammetry, Oblique Images, Aerial Triangulation, Structure from Motion, Exterior Orientation

SUMMARY

In recent years, oblique aerial images have come back to the foreground, being involved in various photogrammetric applications. Moreover, multi-camera systems have become a well-established technology, providing oblique and vertical aerial images that depict both horizontal and vertical structures of the environment from several perspectives, leading to an increasing market availability of such kind of images. The main prerequisite for their metric exploitation is the knowledge of the camera interior and exterior orientation parameters, which are usually determined through a Structure from Motion (SfM) process that estimates corresponding features between overlapping images and solves the multi-image aerial triangulation problem. This paper presents a complete photogrammetry-based framework that solves the SfM problem, which covers the topics of determining overlapping images, feature extraction, image matching, rejection of outliers, feature tracking and bundle block adjustment. The proposed framework adopts a global SfM workflow that relies on approximate camera exterior orientation parameters, which are almost always available through the data provided by onboard GPS/INS sensors. The proposed SfM methodology is applied in different configurations of oblique aerial images (same perspective oblique images, multi-view oblique images and combined nadir and oblique multi-view images) under a non-ideal aerial triangulation scenario characterized by lack of well-distributed ground control points as well as minimum manual image measurements and the results are outlined in the paper. The exterior orientation parameters computed through the proposed SfM algorithm have better accuracy than the ones achieved through a commercial SfM software package. Thus, the proposed global SfM framework proves to be a good alternative solution to existing SfM methods.

A Global Photogrammetry-Based Structure from Motion Framework: Application in Oblique Aerial Images

Styliani VERYKOKOU and Charalabos IOANNIDIS, Greece

1. INTRODUCTION

Although vertical aerial images have played the leading role in photogrammetric applications for more than a century, in recent years oblique aerial images have gained popularity, mainly because of the fundamental advantages they provide compared to nadir views (Remondino and Gerke, 2015; Verykokou and Ioannidis, 2015), in combination with the progress made in photogrammetric and computer vision algorithms that enable their automatic processing. The fact that they reveal vertical structures (mainly facades), which are not depicted in traditional nadir airborne imagery, constitutes their most fundamental characteristic, that leads to an increasing interest towards oblique images, which is followed by an increasing market availability of such kind of images. As a result, multi-camera systems have become a well-established technology, providing datasets of oblique and vertical aerial images that depict urban environments from several perspectives, being a valuable source of information for numerous photogrammetric applications requiring a digital representation of the world.

The main prerequisite for the metric exploitation of image datasets is the knowledge of their camera exterior orientation parameters, along with camera calibration information (interior orientation parameters) or system calibration in case of a multi-camera system (camera interior orientation parameters and relative orientation of cameras). Several researchers have dealt with the process of estimating the camera exterior orientation of datasets containing oblique aerial imagery through an indirect georeferencing approach. They either used existing software for their research (Jacobsen, 2008; Smith et al., 2008; Fritsch and Rothermel, 2013; Saeedi and Mao, 2014; Rau et al., 2015; Gerke et al., 2016; Jacobsen and Gerke, 2016; Ostrowski, 2016; Passini et al., 2016) or developed their own solution for the whole orientation stage (Gerke and Nyaruhuma, 2009; Wang and Neumann, 2009; Gerke, 2011; Habbecke and Kobbelt, 2012; Saeedi and Mao, 2014; Xiong et al., 2014; Gerke et al., 2016; Sun et al., 2016; Verykokou and Ioannidis, 2016b; Xie et al., 2016; Sun et al., 2017; Verykokou and Ioannidis, 2018b) or for part of it (Wiedemann and Moré, 2012; Rupnik et al., 2013; Rupnik et al., 2015; Moe et al., 2016; Toschi et al., 2017), in combination with existing software. A comprehensive review of orientation procedures for oblique aerial images is provided by Verykokou and Ioannidis (2018a).

The purpose of this paper is the presentation of a Structure from Motion (SfM) framework and its application in challenging datasets of oblique aerial images (same perspective oblique images, multi-view oblique images and combined nadir and oblique multi-view images) with the goal of computing their camera exterior orientation parameters. SfM refers to the process of recovering the camera poses and reconstructing the 3D scene geometry from a 2D image sequence. Its first step is the extraction of features in each image, followed by matching of their descriptors and organization of the extracted correspondences into tracks. Then, an incremental

(Snavely et al., 2006), hierarchical (Farenzena et al., 2009) or global method (Moulon et al., 2013) for exterior orientation estimation along with reconstruction of the sparse scene geometry takes place. Among them, incremental methods register one camera at each iteration, hierarchical methods gradually merge short sequences or partial reconstructions and global methods register all cameras simultaneously. In the case of both incremental and hierarchical SfM methods, intermediate bundle adjustment processes are necessary in order to ensure successful camera pose estimation and sparse 3D point cloud extraction. However, frequent intermediate bundle adjustment procedures significantly increase the computational time; in addition, incremental as well as hierarchical methods may suffer from large drifting error due to gradual incorporation of successive views. Global methods do not face these challenges.

Initial estimates for the camera exterior orientation parameters required by the bundle block adjustment stage may be either computed in the context of a global SfM framework, usually through rotation averaging and translation averaging (Cui, 2017), or they may be available from other sources, e.g., from positioning and orientation information by onboard GPS/INS (global positioning system/inertial navigation system) sensors. In this paper, a complete photogrammetry-based framework that solves the global SfM problem, given initial exterior orientation estimates and camera calibration information along with ground control point (GCP) measurements is presented. The paper covers the topics of determining overlapping images, image matching and outlier removal, feature tracking and bundle block adjustment. The proposed framework is applied in different configurations of oblique aerial images (same perspective oblique images, multi-view oblique images and combined nadir and oblique multi-view images) and the results are compared with those ones achieved via commercial software.

The contributions of this paper are twofold. Firstly, it presents an efficient SfM framework based on photogrammetric algorithms, which may be reproduced by interested readers, as all the necessary details, including programming issues, are discussed. Secondly, it applies the proposed algorithm in challenging datasets of oblique aerial images and contributes in the ongoing research for camera exterior orientation estimation of oblique aerial images towards their metric exploitation. The results of the proposed SfM framework prove to have better accuracy than the ones achieved through well-established commercial SfM software.

2. METHODOLOGY

In this section, the proposed global SfM framework is presented. The input data required for its implementation are: (i) a dataset of overlapping images; (ii) image and ground coordinates of at least three control points; (iii) approximate exterior orientation information, usually computed through GPS/INS information; and (iv) camera interior orientation parameters.

2.1 Determination of overlapping images

In order to reduce the processing time of the image matching stage, the overlapping images are determined in a first step, so that the subsequent search of correspondences takes place solely in corresponding images. If the images are accompanied with metadata from GPS/INS sensors, which define the approximate camera position and orientation, and if the camera focal length

and pixel size are known for each image, along with an average flying height, the overlapping images can be determined based on the estimation of their overlap. If they are not accompanied with such kind of metadata, image-based determination of overlapping images takes place.

2.1.1 GPS/INS-based determination of overlapping images

The proposed workflow is illustrated in Figure 1. It is assumed that georeferencing and orientation metadata from GPS/INS sensors along with focal length and pixel size are encoded in the image headers and that a mean flying height is specified by the user.

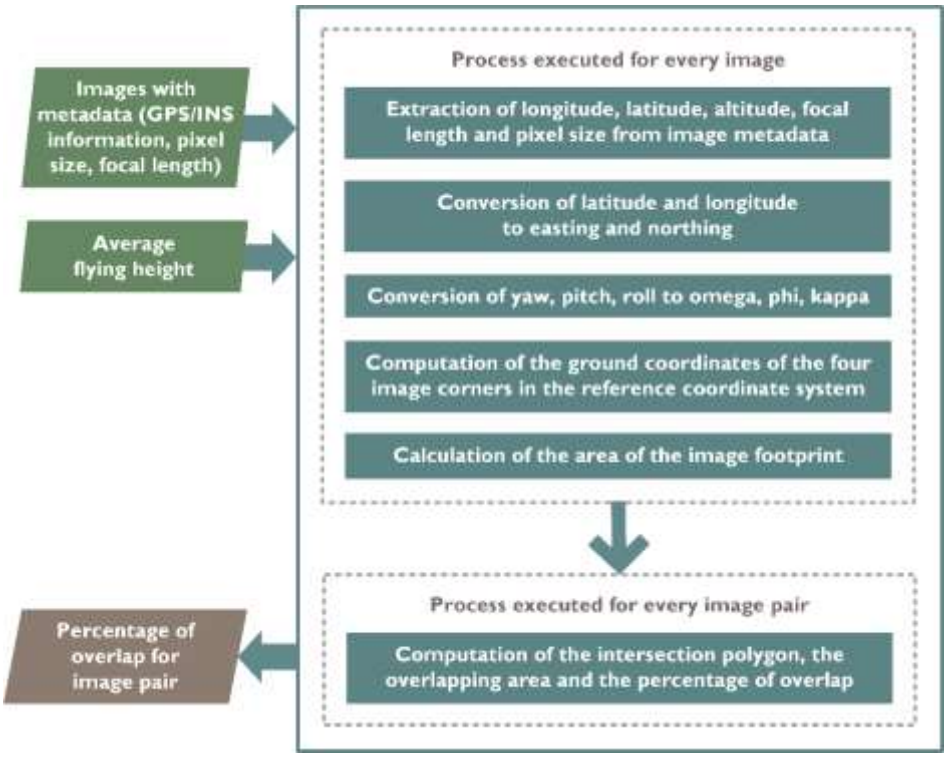

Figure 1. GPS/INS-based determination of overlapping images

The procedure described in the following is implemented for each image separately. The first step is the extraction of the longitude, latitude, altitude, focal length and pixel size from the image metadata. The conversion of latitude and longitude to easting and northing follows. Then, the angular camera orientation, which is usually encoded in terms of yaw (a_y), pitch (a_p) and roll (a_r) is expressed in terms of omega (ω), phi (φ) and kappa (κ). For this reason, the rotation matrix $\boldsymbol{R}$ is firstly calculated according to equations (1), and the angles ω, φ and κ are computed using the elements r_{ij} of the rotation matrix through the set of equations (2).

$$\boldsymbol{R} = \begin{bmatrix} \cos a_y \cos a_p & \cos a_y \sin a_p \sin a_r - \sin a_y \cos r & \cos a_y \sin a_p \cos a_r + \sin a_y \sin a_r \\ \sin a_y \cos a_p & \sin a_y \sin a_p \sin a_r + \cos a_y \cos r & \sin a_y \sin a_p \cos a_r - \cos a_y \sin a_r \\ -\sin a_p & \cos a_p \sin a_r & \cos a_p \cos a_r \end{bmatrix} \quad (1)$$

$$\omega = \tan^{-1}\left(\frac{-r_{32}}{r_{33}}\right) \qquad \varphi = \sin^{-1}\left(r_{31}\right) \qquad \kappa = \tan^{-1}\left(\frac{-r_{21}}{r_{11}}\right) \tag{2}$$

In a next step, the horizontal ground coordinates of the four image corners are computed in the reference coordinate system through the collinearity equations (3), using their coordinates (x, y) in the photogrammetric system, the approximate camera exterior orientation parameters (X_0, Y_0, Z_0, ω, φ, κ), the camera constant c, approximated by the camera focal length, and the mean ground elevation Z, estimated using the average flying height H, as $Z = Z_0 - H$. The coordinates of the image corners are computed in the photogrammetric system using their pixel coordinates and the number of rows and columns along with the pixel size of the imagery.

$$\begin{aligned} X &= X_0 + \left(Z - Z_0\right)\frac{r_{11}\left(x - x_0\right) + r_{21}\left(y - y_0\right) - r_{31}c}{r_{13}\left(x - x_0\right) + r_{23}\left(y - y_0\right) - r_{33}c} \\ Y &= Y_0 + \left(Z - Z_0\right)\frac{r_{12}\left(x - x_0\right) + r_{22}\left(y - y_0\right) - r_{32}c}{r_{13}\left(x - x_0\right) + r_{23}\left(y - y_0\right) - r_{33}c} \end{aligned} \tag{3}$$

Then, the area of the image footprint is calculated using the ground coordinates of the image corners. When the aforementioned workflow has been applied for all images, the intersection polygon between the ground footprints of the images of each possible image pair, their overlapping area and their percentage of overlap are estimated. The percentage of overlap for all image pairs is the output of the whole process.

2.1.2 Image-based determination of overlapping images

The proposed workflow is illustrated in Figure 2. The initial step is the resampling of the images to a sufficiently low resolution. Feature points are then extracted in each image after it has been converted to greyscale, using the speeded-up robust features (SURF) algorithm (Bay et al., 2008). SURF feature points are scale and rotation invariant; skew, anisotropic scaling and perspective effects are also covered to some degree. In addition, the matching of such features is accomplished in less time than the matching of other features with longer descriptor vectors, such as SIFT (scale-invariant feature transform) features (Lowe, 2004).

At the stage of finding correspondences, the feature points extracted from an image are compared to the feature points extracted from all the other images, using the criterion of the minimum Euclidean distance between their descriptor vectors, along with a cross-check test (Verykokou and Ioannidis, 2016a). The correspondences are rejected if the distance between the descriptors of the matched interest points is above a maximum accepted threshold and they are geometrically verified via the RANSAC (random sample consensus) algorithm (Fischler and Bolles, 1981), through computation of the fundamental matrix, using the eight-point algorithm (Hartley, 1997). Whereas the correspondences may still contain some outliers, these ones do not affect the reliability of the resulting information, that is, whether the images overlap. The output of this process is the number of matches between the images of every pair.

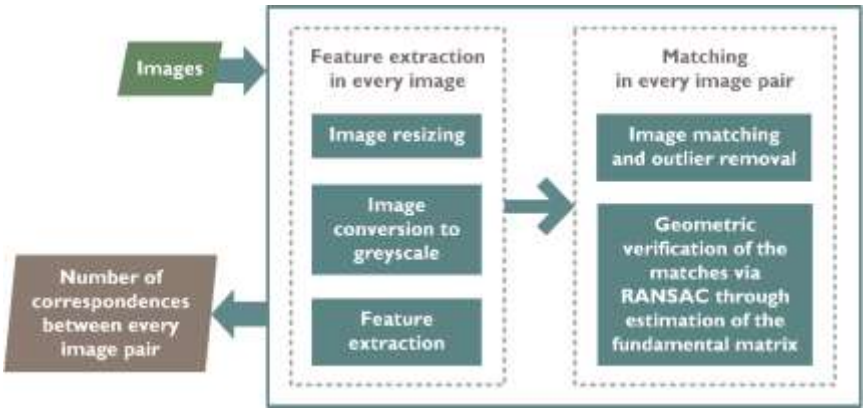

Figure 2. Image-based determination of overlapping images

2.1.3 Graph creation

Regardless of the method applied for determining the overlapping images, an undirected weighted graph that connects the overlapping images is created. Its nodes correspond to the images; an edge connects two nodes that correspond to overlapping images. Two images with a minimum percentage of overlap (if GPS/INS-based determination of overlapping images has taken place) or a minimum number of matches (in the case of the image-based method) are assumed to be overlapping. Each edge of the graph is given a weight, which is the percentage of overlap or the number of correspondences, respectively, between the images that it connects.

2.2 Feature extraction and image matching

Feature points are extracted in all the images, after the reduction of their dimensions, for the scope of accelerating the process, using the SURF algorithm. The resizing factor used here is smaller than the one used during image-based determination of overlapping images. Feature-based matching is applied only to the pairs of overlapping images. A ratio test is implemented (Lowe, 2004) along with a constraint for a maximum value of the Euclidean distance between the descriptors in order to consider two feature points as homologous. The RANSAC algorithm is applied for the removal of outliers via computation of the fundamental matrix. Despite this geometric constraint, some incorrect matches still remain; those feature points which are erroneously considered to match a feature point in another image and happen to lie on the epipolar line of the homologous feature point under consideration are not detected via RANSAC. Thus, a point-to-point constraint is also imposed. Specifically, a homography is fitted to the matches via RANSAC, using a distance threshold for determining the outliers, which is small enough to reject the erroneous matches yet sufficiently large to cope with the homography being approximate and not representing the actual relation between two central projections in cases of non-planar scenes and images not acquired from the same point.

2.3 Feature tracking

The feature tracking algorithm presented by Verykokou and Ioannidis (2018b) is applied for organizing the correspondences into tracks, as required by the aerial triangulation stage. Each

track contains the coordinates of the feature points in different images that correspond to the same 3D point. Feature tracking takes place for each pair of overlapping images after the image matching stage for this pair. The results of the feature tracking process are two matrices, that is, a matrix that stores the image coordinates of the feature points and a matrix that stores information about whether a point is visible in each image; the number of rows of both matrices is equal to the number of images and their number of columns is equal to the number of tracks.

2.4 Aerial triangulation

A bundle block adjustment is implemented to perform the final aerial triangulation step. Approximate exterior orientation parameters are required by the iterative least-squares solution of the bundle block adjustment. Such values are usually available from the onboard sensors (GPS/INS) of the flying platform used for the image acquisition process.

2.4.1 Computation of approximate ground coordinates for the 3D tracks

Approximate values for the 3D coordinates of the corresponding feature points between each pair of overlapping images are computed. In order to obtain more stable tracks for the bundle adjustment process, only tracks that include feature point measurements in at least three images are used. Specifically, a least-squares solution that solves the conventional problem of photogrammetric space intersection through indirect observations is adopted for every pair of corresponding points that belong to a track of at least three feature points, based on the collinearity equations. The initial data required for estimating the 3D coordinates of a point depicted in a pair of images (*img_i*, *img_j*) are (i) its coordinates in both images (x_i, y_i) and (x_j, y_j) expressed in the photogrammetric system; (ii) the interior orientation parameters of the camera that took the images; and (iii) the approximate exterior orientation of the images. The initial values for the 3D coordinates of the point to be triangulated (X_{approx}, Y_{approx}, Z_{approx}) required for the iterative least-squares solution are estimated using the set of equations (4).

$$Z_{approx} = \frac{\left(X_{0j} - X_{0i}\right)c_i'c_j' - x_i'c_j'Z_{0i} + x_j'c_i'Z_{0j}}{-x_i'c_j' + x_j'c_i'}$$
$$X_{approx} = X_{0i} - \frac{x_i'\left(Z_{approx} - Z_{0i}\right)}{c_i'} \qquad Y_{approx} = Y_{0i} - \frac{y_i'\left(Z_{approx} - Z_{0i}\right)}{c_i'} \tag{4}$$

In equations (4), X_{0i}, Y_{0i}, Z_{0i} and X_{0j}, Y_{0j}, Z_{0j} are the projection center coordinates of *img_i* and *img_j* respectively; x_i', y_i', c_i' and x_j', y_j', c_j' are computed via equations (5), where c is the principal distance and $\boldsymbol{R}_i$ is the rotation matrix of *img_i*, calculated using the approximate angular exterior orientation elements (omega, phi and kappa angles) of *img_i*.

$$\begin{bmatrix} x_i' \\ y_i' \\ -c_i' \end{bmatrix} = \boldsymbol{R}_i^T \cdot \begin{bmatrix} x_i \\ y_i \\ -c \end{bmatrix} \qquad \begin{bmatrix} x_j' \\ y_j' \\ -c_j' \end{bmatrix} = \boldsymbol{R}_i^T \cdot \begin{bmatrix} x_j \\ y_j \\ -c \end{bmatrix} \tag{5}$$

2.4.2 Outlier removal

Every 3D feature point is visible in more than two images. Thus, the distance between their ground coordinates, which are computed through photogrammetric intersection using different combinations of overlapping images, are computed. If this distance is above a maximum accepted threshold for at least one combination of image pairs, the entire track is considered to be erroneous and is removed from the initial data. In this way, another check that removes any remaining outliers that have passed through the geometric constraints imposed by the image matching procedure is implemented.

2.4.3 Bundle block adjustment

The ground coordinates of the feature points estimated through space intersection (section 2.4.1), along with the approximate exterior orientation parameters of the images serve as initial values for the bundle block adjustment. Knowledge of the camera focal length, if an auto-calibration process is implemented, or availability of camera calibration information, if the interior orientation is considered to be fixed during the bundle adjustment process, is also required. The mathematical model used is expressed by the collinearity equations. A least-squares solution is implemented that solves the non-linear system through indirect observations.

3. EXPERIMENTS

In this section, the experiments performed for assessing the proposed global SfM framework are described and the results are outlined, being compared with reference data along with results derived through a well-established commercial software package.

3.1 Developed software

A software suite implementing the proposed methodology has been developed in the C++ programming language, making use of some functionalities offered by the OpenCV library for image manipulations, the Eigen library for matrix operations, the GDI+ library for extraction of image metadata, the GeographicLib library for conversion of latitude and longitude to easting and northing and the Boost libraries for graph creation and geometric calculations.

3.2 Test dataset

The images used in the experiments are part of a dataset that consists of 50 multi-perspective oblique and vertical aerial images, which are provided by the ISPRS/EuroSDR initiative "Benchmark on High Density Image Matching for DSM Computation" (Cavegn et al., 2014). The images were taken by a Leica RCD30 Oblique Penta camera system with a 60-Mpixel sensor from a flying height of about 520 m. They were captured in a Maltese-cross configuration with four oblique cameras, the axes of which were tilted at 35° with respect to the vertical, and one nadir camera. The images have a pixel count of 9000×6732 pixels, a pixel size of 6 μm and they are free of distortion. Their calibrated focal length is 53 mm and their ground sample distance varies between 6 and 13 cm. The photogrammetric block consists of one strip. The approximate image overlap in nadir view is 70%. Figure 3 illustrates the ground

footprints of the captured imagery along with the overlap both among the simultaneously acquired multi-view imagery (Figure 3-a) and between the successive images taken from a single camera (Figure 3-b). The images are accompanied with reference exterior orientation parameters computed through a highly accurate triangulation (Cavegn et al., 2014), which are used for assessing the results derived using the proposed aerial triangulation framework.

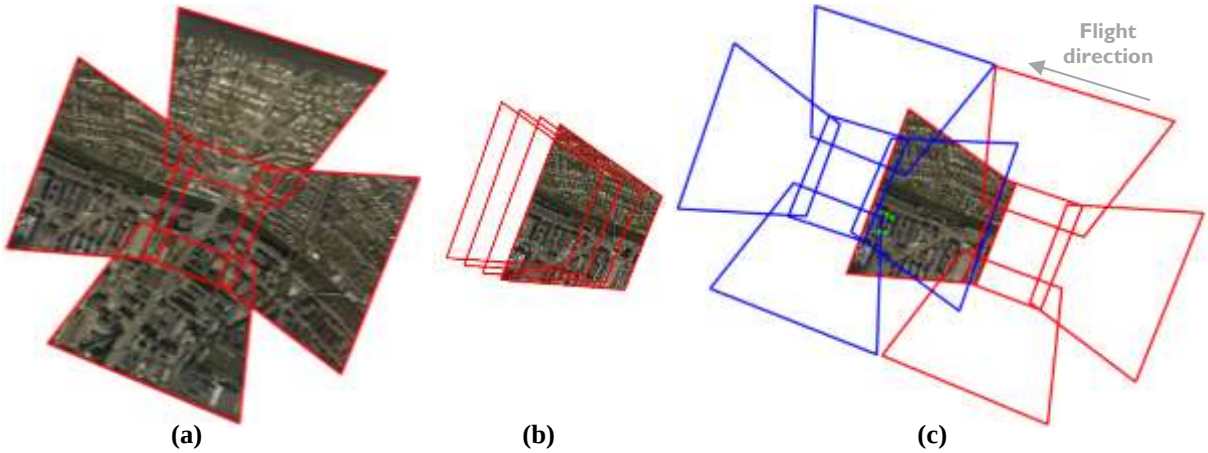

Figure 3. Ground footprints of the images acquired by the Leica RCD30 Oblique Penta camera system, generated using the georeferencing software suite by Verykokou and Ioannidis (2016a); (a): images acquired by the five cameras at one time instance; (b): successive images taken by one oblique camera; (c): images that correspond to the first (shown in red) and the last (shown in blue) acquisition of the strip along with the image used for manual measurements of four coplanar GCPs (shown in green)

Three subsets of the aforementioned dataset encompassing different configurations of oblique aerial images are used in the experiments, as described in the following:

- Scenario 1: Oblique aerial images of the same perspective. This scenario concerns the aerial triangulation of a set of 10 oblique aerial images, captured by the same camera of the multi-view system.
- Scenario 2: Multi-perspective oblique aerial images. This scenario concerns the aerial triangulation of 40 oblique aerial images captured by the four cameras of the multi-view system.
- Scenario 3: Multi-perspective oblique and vertical aerial images. This scenario concerns the aerial triangulation of 40 oblique and 10 vertical aerial images captured in a Maltese cross configuration by the five cameras of the multi-view system.

3.3 Input data

The purpose of the experiments is to investigate the accuracy that can be achieved in a non-ideal scenario characterized by the availability of a small number of coplanar GCPs concentrated in a small area relative to the area of the whole block of images, and image measurements at the minimum possible number of images. For this reason, four coplanar GCPs have been measured in one of the four starting oblique aerial images of the strip (Figure 3-c) and their image coordinates have been automatically estimated in its successive oblique aerial image taken by the same camera, using the method presented by Verykokou and Ioannidis

(2018b). For each one of the three aerial triangulation scenarios, the same four GCP measurements in the same images have been used. The ground coordinates of the GCPs used in the experiments have been estimated by photogrammetric space intersection using the reference exterior orientation parameters. The GPS information that accompanies the imagery along with rough estimates of the camera orientation angles serve as initial values for the exterior orientation parameters of the imagery. The camera calibration information (principal distance along with principal point coordinates and zero distortion) provided by the benchmark are used as fixed interior orientation parameters during the aerial triangulation process.

3.4 Results

In order to estimate the linear and the angular difference between the computed exterior orientation parameters and the reference ones, the following metrics are calculated:

- *DprojCenters* indicates the distance betweeen the computed projection center for an image (X_0, Y_0, Z_0) and the reference one ($X_{0,ref}$, $Y_{0,ref}$, $Z_{0,ref}$); it is calculated by equation (6);
- *Dquaternions* indicates the distance between the unit quaternion that corresponds to the computed Euler angles (ω, φ, κ) for an image ($\boldsymbol{q}$) and the unit quaternion that corresponds to the reference Euler angles for the same image ($\boldsymbol{q}_{ref}$); it is calculated through equation (7), where $\|\cdot\|$ symbolizes the Euclidean norm in $\mathbb{R}^4$. Specifically, the conversion from the angles ω, φ, κ to a quaternion-representation $\boldsymbol{q} = [q_1 \ q_2 \ q_3 \ q_4]^T$ (or equivalently $\boldsymbol{q} = q_1 + \boldsymbol{i}q_2 + \boldsymbol{j}q_3 + \boldsymbol{k}q_4$, where q_1 is the real part of the quaternion, and $\boldsymbol{i}$, $\boldsymbol{j}$ and $\boldsymbol{k}$ are unit vectors pointing along the three spatial axes and forming the imaginary part of the quaternion) is applied, because the latter provides a more efficient framework that permits to obtain a distance metric between 3D rotations (Huynh, 2009; Hartley et al., 2013). Taking into account the fact that $\boldsymbol{q}$ and $-\boldsymbol{q}$ represent the same rotation, the quaternion distance is calculated as the minimum between the Euclidean norm of the difference of quaternions and the Euclidean norm of the sum of quaternions. The formula for calculating a quaternion from the set of rotation angles ω, φ, κ applied sequentially (from ω to κ) is derived by calculating the product of the quaternions that correspond to each one of the rotation angles, as shown in equation (8). This distance gives values in the range $[0, \sqrt{2}]$; it is dimensionless and has been proved to be a metric of the angular distance between 3D rotations (Huynh, 2009; Hartley et al., 2013)

$$D_{projCenters} = \sqrt{\left(X_0 - X_{0,ref}\right)^2 + \left(Y_0 - Y_{0,ref}\right)^2 + \left(Z_0 - Z_{0,ref}\right)^2} \tag{6}$$

$$D_{quaternions} = \min\left\{\left\|\boldsymbol{q} - \boldsymbol{q}_{ref}\right\|, \left\|\boldsymbol{q} + \boldsymbol{q}_{ref}\right\|\right\} \tag{7}$$

$$\left.\begin{aligned} \boldsymbol{q}_{\omega} &= \cos(\omega/2) + \boldsymbol{i}\sin(\omega/2) \\ \boldsymbol{q}_{\varphi} &= \cos(\varphi/2) + \boldsymbol{j}\sin(\varphi/2) \\ \boldsymbol{q}_{\kappa} &= \cos(\kappa/2) + \boldsymbol{k}\sin(\kappa/2) \end{aligned}\right\} \ \boldsymbol{q} = \boldsymbol{q}_{\kappa}\boldsymbol{q}_{\varphi}\boldsymbol{q}_{\omega} \tag{8}$$

Specifically, the average (Avg), maximum (Max) and minimum (Min) values of these metrics among all images for each aerial triangulation scenario along with their standard deviation

(Stdev) are computed. Table 1 illustrates the results derived by the in-house developed software suite that implements the proposed SfM workflow. The results demonstrate a better accuracy in the exterior orientation parameters for the dataset of 10 single-perspective oblique images (scenario 1). The largest differences from the reference data are observed for the dataset of 40 multi-perspective oblique aerial images (scenario 2), as there is very poor overlap between the four oblique aerial images acquired by the multi-view camera system at a single time instance. The dataset of 50 multi-perspective oblique and vertical aerial images (scenario 3) corresponds to smaller differences from the reference data than the oblique-only multi-perspective dataset (scenario 2), as the five images that are acquired by the multi-camera system provide a sufficiently stronger geometry to tie the side oblique aerial images of the strip together.

	Metric	Avg	Max	Min	Stdev
Scenario 1	$D_{projCenters}$ **(m)**	0.529	1.058	0.118	0.286
	$D_{quaternions}$	$2.667 \cdot 10^{-4}$	$4.412 \cdot 10^{-4}$	$1.075 \cdot 10^{-4}$	$1.228 \cdot 10^{-4}$
Scenario 2	$D_{projCenters}$ **(m)**	0.747	1.949	0.246	0.407
	$D_{quaternions}$	$6.782 \cdot 10^{-4}$	$16.175 \cdot 10^{-4}$	$0.000 \cdot 10^{-4}$	$2.884 \cdot 10^{-4}$
Scenario 3	$D_{projCenters}$ **(m)**	0.605	1.719	0.045	0.355
	$D_{quaternions}$	$6.579 \cdot 10^{-4}$	$18.628 \cdot 10^{-4}$	$0.000 \cdot 10^{-4}$	$3.486 \cdot 10^{-4}$

Table 1. Metrics indicating the linear ($D_{projCenters}$) and angular ($D_{quaternions}$) differences between the exterior orientation parameters computed through the developed SfM software and the reference ones for the three aerial triangulation scenarios

For comparison reasons, the exterior orientation parameters of the images of the three aerial triangulation scenarios were computed through existing SfM software packages, which delivered similar aerial triangulation results. For this reason, only results obtained from Agisoft PhotoScan (professional edition, version 1.2.6), which is a well-established commercial SfM software package are shown in Table 2. The same GCP measurements were used and the same camera interior orientation parameters were specified as fixed during the aerial triangulation process, so that the results are comparable to the ones derived through the proposed algorithm. The average value of the distance metric $D_{projCenters}$ is smaller for the case of the proposed approach than the one corresponding to Agisoft PhotoScan for all aerial triangulation scenarios; this improvement reaches the percentage of 20%, 5% and 22% for scenario 1, 2 and 3, respectively. The average quaternion distance $D_{quaternions}$ is also improved for all aerial triangulation scenarios for the case of the proposed approach, compared to Agisoft PhotoScan results; this improvement reaches the percentages of 56% for scenario 1, 1% for scenario 2 and 1% for scenario 3. Similar results were also derived using other existing SfM software solutions.

Hence, the proposed global SfM framework is a robust alternative to existing SfM software solutions. Also, the exterior orientation parameters estimated via the commercial package verify the conclusion derived from the developed software results, according to which better accuracy is achievable in scenario 1 and the worst accuracy corresponds to scenario 2. Finally, bigger systematic error is derived by the commercial software, which is proved by the bigger absolute values of the average differences between the computed and the reference exterior orientation parameters, compared to the results achieved through the developed software solution.

	Metric	Avg	Max	Min	Stdev

Scenario 1	$D_{projCenters}$ **(m)**	0.661	0.957	0.295	0.174
	$D_{quaternions}$	$6.009 \cdot 10^{-4}$	$7.930 \cdot 10^{-4}$	$1.936 \cdot 10^{-4}$	$1.551 \cdot 10^{-4}$
Scenario 2	$D_{projCenters}$ **(m)**	0.786	1.519	0.257	0.223
	$D_{quaternions}$	$6.849 \cdot 10^{-4}$	$16.332 \cdot 10^{-4}$	$0.000 \cdot 10^{-4}$	$3.428 \cdot 10^{-4}$
Scenario 3	$D_{projCenters}$ **(m)**	0.774	1.399	0.263	0.222
	$D_{quaternions}$	$6.641 \cdot 10^{-4}$	$14.744 \cdot 10^{-4}$	$0.000 \cdot 10^{-4}$	$3.150 \cdot 10^{-4}$

Table 2. Metrics indicating the linear ($D_{projCenters}$) and angular ($D_{quaternions}$) differences between the exterior orientation parameters computed through the Agisoft PhotoScan software and the reference ones for the three aerial triangulation scenarios

4. CONCLUSIONS AND FUTURE WORK

This paper presents an efficient SfM framework based on photogrammetric algorithms and demonstrates the results that can be achieved in challenging datasets of oblique aerial imagery in non-ideal aerial triangulation scenarios characterized by lack of well-distributed GCPs and minimum manual image measurements. The exterior orientation parameters computed through the proposed algorithm prove to have better accuracy than the ones achieved through a well-established commercial software package. Thus, the proposed global SfM framework is a good alternative solution to existing SfM methods, in cases of datasets that are accompanied with rough exterior orientation information, e.g., from GPS/INS sensors. Furthermore, the results derived both from the developed solution and the commercial software demonstrate that better accuracy is achieved in cases of combined blocks of multi-view oblique and vertical aerial images acquired through a Maltese-cross configuration in comparison with multi-view oblique-only image blocks with small side overlap between the images.

Future work will focus on the impact of different feature extraction algorithms on the aerial triangulation results of oblique aerial imagery. Furthermore, more research will be carried out to investigate the impact of a weighting strategy for image measurements during the bundle block adjustment of oblique views. Finally, future research will be conducted with the goal of incorporating into the aerial triangulation stage of our proposed global SfM framework iterative bundle adjustment procedures for rejecting any remaining outlier tracks, as proposed by the incremental SfM framework introduced by Verykokou and Ioannidis (2018b).

ACKNOWLEDGEMENTS

The authors would like to acknowledge ISPRS and EuroSDR for image acquisition in the context of the ISPRS/EuroSDR initiative “Benchmark on High Density Image Matching for DSM Computation”. Also, Styliani Verykokou would like to acknowledge the Eugenides Foundation for the financial support through a PhD scholarship.

REFERENCES

Bay, H., Ess, A., Tuytelaars, T., Van Gool, L. 2008. Speeded-up robust features (SURF). *Computer Vision and Image Understanding,* vol. 110, no. 3, pp. 346-359.

Cavegn, S., Haala, N., Nebiker, S., Rothermel, M., Tutzauer, P. 2014. Benchmarking high density image matching for oblique airborne imagery. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences,* vol. XL-3, pp. 45-52.

Cui, Z. 2017. *Global Structure-from-Motion and Its Application* PhD Thesis, Simon Fraser University.

Farenzena, M., Fusiello, A., Gherardi, R. 2009. Structure-and-motion pipeline on a hierarchical cluster tree. *2009 IEEE 12th International Conference on Computer Vision Workshops (ICCV Workshops).* Kyoto, Japan: IEEE.

Fischler, M., Bolles, R. 1981. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM,* vol. 24, no. 6, pp. 381-395.

Fritsch, D., Rothermel, M. 2013. Oblique image data processing - Potential, experiences and recommendations. *In:* Fritsch, D. (ed.) *Photogrammetric Week '13.* Berlin: Wichmann.

Gerke, M. 2011. Using horizontal and vertical building structure to constrain indirect sensor orientation. *ISPRS Journal of Photogrammetry and Remote Sensing,* vol. 66, no. 3, pp. 307-316.

Gerke, M., Nex, F., Remondino, F., Jacobsen, K., Kremer, J., Karel, W., Hu, H., Ostrowski, W. 2016. Orientation of oblique airborne image sets-experiences from the ISPRS/EuroSDR benchmark on multi-platform photogrammetry. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences,* vol. XLI-B1, pp. 185-191.

Gerke, M., Nyaruhuma, A. 2009. Incorporating scene constraints into the triangulation of airborne oblique images. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences,* vol. XXXVIII-1-4-7/W5.

Habbecke, M., Kobbelt, L. 2012. Automatic registration of oblique aerial images with cadastral maps. *In:* Kutulakos, K.N. (ed.) *Trends and Topics in Computer Vision. ECCV 2010. Lecture Notes in Computer Science, vol 6554.* Springer, Berlin, Heidelberg.

Hartley, R., Trumpf, J., Dai, Y., Li, H. 2013. Rotation averaging. *International Journal of Computer Vision,* vol. 101, no. 2, pp. 267-305.

Hartley, R.I. 1997. In defense of the eight-point algorithm. *IEEE Transactions on Pattern Recognition and Machine Intelligence,* vol. 19, no. 6, pp. 580-593.

Huynh, D.Q. 2009. Metrics for 3D Rotations: comparison and analysis. *Journal of Mathematical Imaging and Vision,* vol. 35, no. 2, pp. 155-164.

Jacobsen, K. 2008. Geometry of vertical and oblique image combinations. *In:* Maktav, D. (ed.) *28th Symposium of the European Association of Remote Sensing Laboratories.* Istanbul, Turkey: IOS Press.

Jacobsen, K., Gerke, M. 2016. Sub-camera calibration of a penta-camera. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences,* vol. XL-3/W4, pp. 35-40.

Lowe, D.G. 2004. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision,* vol. 60, no. 2, pp. 91-110.

Moe, K., Toschi, I., Poli, D., Lago, F., Schreiner, C., Legat, K., Remondino, F. 2016. Changing the production pipeline - Use of oblique aerial cameras for mapping purposes. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences,* vol. XLI-B4, pp. 631-637.

Moulon, P., Monasse, P., Marlet, R. 2013. Global fusion of relative motions for robust, accurate and scalable structure from motion. *2013 IEEE International Conference on Computer Vision (ICCV).* Sydney, NSW, Australia.

Ostrowski, W. 2016. Accuracy of measurements in oblique aerial images for urban environment. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences,* vol. XLII-2/W2, pp. 79-85.

Passini, R., Jacobsen, K., Weaver, W., Day, D. 2016. An study on oblique camera systems calibration and the stability of their parameters. *ASPRS 2016 Annual Conference: IGTF-2016 (Imaging and Geospatial Technology Forum).* Fort Worth, Texas, USA.

Rau, J.-Y., Jhan, J.-P., Hsu, Y.-C. 2015. Analysis of oblique aerial images for land cover and point cloud classification in an urban environment. *IEEE Transactions on Geoscience and Remote Sensing,* vol. 53, no. 3, pp. 1304-1318.

Remondino, F., Gerke, M. 2015. Oblique aerial imagery - a review. *In:* Fritsch, D. (ed.) *Photogrammetric Week '15.* Belin: Wichmann/VDE Verlag.

Rupnik, E., Nex, F., Remondino, F. 2013. Automatic orientation of large blocks of oblique images. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences,* vol. XL-1/W1, pp. 299-304.

Rupnik, E., Nex, F., Toschi, I., Remondino, F. 2015. Aerial multi-camera systems: Accuracy and block triangulation issues. *ISPRS Journal of Photogrammetry and Remote Sensing,* vol. 101, pp. 233-246.

Saeedi, P., Mao, M. 2014. Two-edge-corner image features for registration of geospatial images with large view variations. *International Journal of Geosciences,* vol. 5, no. 11, pp. 1324-1344.

Smith, M.J., Kokkas, N., Hamruni, A.M., Critchley, D., Jamieson, A. 2008. Investigation into the orientation of oblique and vertical digital images. *EuroCOW 2008.* Castelldefels, Spain.

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.

Sun, X., Shen, S., Cui, H., Hu, L., Hu, Z. 2017. Geographic, geometrical and semantic reconstruction of urban scene from high resolution oblique aerial images. *IEEE/CAA Journal of Automatica Sinica,* vol. PP, no. 99, pp. 1-13.

Sun, Y., Sun, H., Yan, L., Fan, S., Chen, R. 2016. RBA: Reduced Bundle Adjustment for oblique aerial photogrammetry. *ISPRS Journal of Photogrammetry and Remote Sensing,* vol. 121, pp. 128–142.

Toschi, I., Ramos, M.M., Nocerino, E., Menna, F., Remondino, F., Moe, K., Poli, D., Legat, K., Fassi, F. 2017. Oblique photogrammetry supporting 3D urban reconstruction of complex scenarios. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences,* vol. XLII-1/W1, pp. 519-526.

Verykokou, S., Ioannidis, C. 2015. Metric exploitation of a single low oblique aerial image. *FIG Working Week 2015.* Sofia, Bulgaria.

Verykokou, S., Ioannidis, C. 2016a. Automatic rough georeferencing of multiview oblique and vertical aerial image datasets of urban scenes. *The Photogrammetric Record,* vol. 31, no. 155, pp. 281-303.

Verykokou, S., Ioannidis, C. 2016b. Exterior orientation estimation of oblique aerial imagery using vanishing points. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences,* vol. XLI-B3, pp. 123-130.

Verykokou, S., Ioannidis, C. 2018a. Oblique aerial images: a review focusing on georeferencing procedures. *International Journal of Remote Sensing,* vol. 39, no. 11, pp. 3452-3496.

Verykokou, S., Ioannidis, C. 2018b. A photogrammetry-based structure from motion algorithm using robust iterative bundle adjustment techniques. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences,* vol. IV-4/W6, pp. 73-80.

Wang, L., Neumann, U. 2009. A robust approach for automatic registration of aerial images with untextured aerial LiDAR data. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2009).* Miami, Florida, USA: IEEE.

Wiedemann, A., Moré, J. 2012. Orientation strategies for aerial oblique images. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences,* vol. XXXIX-B1, pp. 185-189.

Xie, L., Hu, H., Wang, J., Zhu, Q., Chen, M. 2016. An asymmetric re-weighting method for the precision combined bundle adjustment of aerial oblique images. *ISPRS Journal of Photogrammetry and Remote Sensing,* vol. 117, pp. 92-107.

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 Geoscience and Remote Sensing Letters,* vol. 11, no. 11, pp. 1976-1980.

BIOGRAPHICAL NOTES

Styliani VERYKOKOU

Styliani Verykokou obtained a diploma in Rural and Surveying Engineering (2013) from the National Technical University of Athens (NTUA), Greece, achieving a grade of 9.29/10. She has been awarded several scholarships and awards by the National Institute of Scholarships of Greece, the NTUA, the Academy of Athens, the Technical Chamber of Greece, the Thomaidion Foundation (Greece) and the Limmat Foundation (Switzerland), as a result of her performance in her studies. Since 2014, she is PhD candidate at the Lab. of Photogrammetry of NTUA and has participated in several national and European research projects. Her PhD topic concerns the metric exploitation of oblique aerial imagery and it has been supported by the Eugenides Foundation scholarship. She has authored several papers in the fields of photogrammetry and computer vision.

Charalabos IOANNIDIS

Dr. Surveying Engineer, Professor at the Lab. of Photogrammetry, School of Rural and Surveying Engineering, National Technical University of Athens, Greece, in the field of Photogrammetry and Cadastre.
1992-1996: Co-chair of Commission VI-WG2 "Computer Assisted Teaching" in ISPRS.
1997-2001: Member of the Directing Council of the Hellenic Mapping and Cadastral Organization and Deputy Project Manager of the Hellenic Cadastre.
2010-2018: Chair of Working Group 3.2 "Technical Aspects of SIM" of FIG Commission 3.
2019-2022: Chair of Working Group 3.2 "Geospatial Big Data: collection, processing, and presentation".
He has given a series of lectures in 15 national and 7 international Seminars, in the fields of Photogrammetry, Geometric Documentation of Cultural Heritage, Terrestrial laser Scanning, Cadastre and GIS. He has been the scientific responsible of 19 research projects and he has participated in more than 35 research projects, in 16 of them as the Principal Investigator. He has written more than 160 reviewed papers in scientific journals, books and proceedings, and another 50 presentations in conferences. He serves as reviewer in various scientific journals, books and international conferences.

CONTACTS

Styliani Verykokou
School of Rural & Surveying Engineering,
National Technical University of Athens
9 Iroon Polytechniou St.
Athens
GREECE
Tel. +302107722687
Email: st.verykokou@gmail.com
Web site: http://users.ntua.gr/sveryk/

Prof. Dr. **Charalabos Ioannidis**
School of Rural & Surveying Engineering,
National Technical University of Athens
9 Iroon Polytechniou St.
Athens
GREECE
Tel. +302107722686
Email: cioannid@survey.ntua.gr
Web site: http://users.ntua.gr/cioannid/