Aerial multi-camera systems: Accuracy and block triangulation issues

Ewelina Rupnik*, Francesco Nex, Isabella Toschi, Fabio Remondino

3D Optical Metrology Unit (3DOM), Bruno Kessler Foundation (FBK), Trento, Italy

A R T I C L E   I N F O

Article history:
Received 28 August 2014
Received in revised form 22 December 2014
Accepted 23 December 2014

Keywords:
Oblique images
Simulation
Flight planning
Orientation
Bundle adjustment
Accuracy
Automation

A B S T R A C T

Oblique photography has reached its maturity and has now been adopted for several applications. The number and variety of multi-camera oblique platforms available on the market is continuously growing. So far, few attempts have been made to study the influence of the additional cameras on the behaviour of the image block and comprehensive revisions to existing flight patterns are yet to be formulated. This paper looks into the precision and accuracy of 3D points triangulated from diverse multi-camera oblique platforms. Its coverage is divided into simulated and real case studies. Within the simulations, different imaging platform parameters and flight patterns are varied, reflecting both current market offerings and common flight practices. Attention is paid to the aspect of completeness in terms of dense matching algorithms and 3D city modelling – the most promising application of such systems. The experimental part demonstrates the behaviour of two oblique imaging platforms in real-world conditions. A number of Ground Control Point (GCP) configurations are adopted in order to point out the sensitivity of tested imaging networks and arising block deformations. To stress the contribution of slanted views, all scenarios are compared against a scenario in which exclusively nadir images are used for evaluation.

© 2014 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.

1. Introduction

Oblique aerial images, alone or in combination with vertical source thanks to their intuitive nature. Documented use of oblique imagery dates back to the nineteenth century. Gaspard Félix Tournaire acquired the first aerial photograph in 1858 by raising a balloon over the village of Petit Bicêtre in France (Colwell, 1997). Two years later, the oldest surviving oblique aerial image was taken by J.W. Black and S. King in Boston, USA, from a balloon. In 1887 a balloon again was employed by a German forester to acquire aerial photos over forests in order to identify and measure stands of trees (Colwell, 1997). The same year also saw the birth of aerial photography from a kite, thanks to a British meteorologist, E.D. Archibald (Colwell, 1997). At about the same time in France, the Tissandier brothers conducted other experiments of kite and balloon aerial photography. Thereafter this practice of acquiring images from kites moved across the Atlantic and advanced rapidly. In 1906 G.R. Lawrence, using between nine and seventeen large kites to lift a huge camera, took some oblique aerial photographs of San Francisco (USA) after an enormous earthquake in the area.

Thanks to the rapid progress in military photography from airplanes, kites were then abandoned in favour of powered flight, which gained importance for military reconnaissance during World War I. Simultaneously, single and multiple lens cameras were produced for oblique-only or combined configurations in the UK, Germany, France, Italy, Switzerland and USA (Manual of Photogrammetry, 1952). In the 1930s for instance, the U.S. Geological Survey and the U.S. Army used a Fairchild T-3A five-lens film camera for mapping, surveillance and reconnaissance purposes. Finally, during World War II, oblique aerial photography was employed for reconnaissance purposes before and after bombing missions (Nocerino et al., 2012), thus further stimulating the rapid development of new cameras, lenses, films and camera mounting systems (Aber et al., 2010).

The concept of fitting several imaging sensors into a unique camera housing re-emerged for civil applications with the introduction of digital imaging technology (Petrie, 2010). However, the true revival of oblique imaging systems for geospatial applications occurred in 2000 when Pictometry International introduced their five-lens camera system that incorporated vertical and slant views (Patent Ser. No. 60/425,275, filed November 8, 2002). Pictometry imagery has also, more recently, been disseminated through the Bird’s Eye view function of current Bing Maps. Over the past decade, the geospatial industry has enthusiastically taken up oblique imaging technology. New businesses have been born
results, what has been largely overlooked until now is any revision of traditionally adopted flight patterns. Open questions include: (i) How do different overlap scenarios influence the precision of computed object points? (ii) What is the right balance between aerial survey precision and productivity? (iii) How should a flight be planned in terms of overlap and camera configuration parameters (e.g. tilt) to achieve most complete results? (iv) What, after all, is the influence of the oblique views on the image block triangulation from the point of view of obtainable accuracy?

In the following sections of the paper, the authors address the above problems related to oblique multi-camera systems in terms of image triangulation (self-calibration is not the primary scope of the paper). Section 2 reports on existing flight planning routines and simulation studies of various block configurations. The influences of varying sensor size, tilt angle and overlap (in nadir and oblique cameras) are considered regarding the obtainable accuracy, with attention being given to the issue of completeness of point cloud retrieval in urban environments. Section 3 briefly describes the image orientation task related to oblique imagery. Lastly, Section 4 reports the results of two real-case studies. The behaviour of image blocks is evaluated for different Ground Control Point (GCP) configurations, different image overlaps, and GNSS assisted image triangulation.

2. Flight planning – simulated case studies

Flight planning refers to the initial determination of flight geometry, given the area of interest and the required end-product, and thus the desired accuracy. In the case of traditional nadir imagery, the parameters that are optimized are the flying height, the camera with its optics and the overlap pattern. Depending upon the type of photogrammetric application, adopted lenses can vary from wide to narrow angles. As a general rule, long focal lengths are used when large height variations are present (cities, mountainous areas), whereas short focal lengths are preferred in rural areas, for overview flights or when better height accuracy is targeted. Extensive studies on optimization of flight pattern scenarios were carried out during the second half of the 20th century (Förstner, 1985; Ackermann, 1992; Kraus, 1997, 2007). At that time, digital photogrammetry was still an unrealised concept and manual stereoscopic observations were commonplace. Hence, the attention remained around vertical imaging while combined vertical and oblique photography was conceived merely for pictorial representation. The prevailing rule-of-thumb suggested that flying scenarios with 60% forward overlap would ensure a 3-fold coverage for points down each side of the image, providing for strip formation. Furthermore this also afforded good stereoscopic viewing with a good base length for 3D measurement. At the same time, the 20–40% side overlap provided a sufficiently strong geometry to tie the neighbouring strips of photography together. Reduced processing times and compilation costs were a top priority; therefore the fewer images covering the area of interest the better the economy of a project.

The costs of mapping with photogrammetry started to significantly reduce as the technology matured and the digital era was embarked upon. This brought about automated image measurement, which raised the compilation efficiency and reduced human operator interaction. Automation gave rise to a new tendency to fly denser patterns, in particular in the in-flight direction. Also, thinking in terms of single stereo models gave way to thinking in terms of multi-image bundles of rays (McGlone, 2004; Aber et al., 2010).

Within this context of “digital evolution”, oblique multi-camera systems started again to be developed and employed for different application fields and market niches. They began to be thought of as instruments for cartographic mapping, rather than only for...
visualization purposes or reconnaissance, as was the case in the past. The imaging geometry of oblique camera systems is analogous to close-range applications, with all their characteristics, e.g. greatly varying scale within the images, illumination changes, multiple viewing directions and wide baselines and hence large perspective differences between the views. To fully exploit the oblique cameras on-board, be it for texture mapping or image dense matching, the flight planning must thus more cautiously adapt to the situation on the ground, and the overlaps must be separately calculated for nadir, backward/forward, and side cameras. Apart from the flying height and the camera optics, the effective overlaps are conditioned on the camera tilt angles. As will be shown later, only a just combination of all the parameters will produce images that can serve for precise, and complete 3D scene reconstruction (Fritsch and Rothermel, 2013; Rupnik et al., 2013).

In the following paragraphs, the authors evaluate the quality of 3D point triangulation under several imaging configurations. Similar investigations using multi-view nadir and stereo oblique geometry have been presented in Förstner (1998) and Gerke (2009), respectively.

2.1. Simulation procedure

The purpose of simulation is to provide *a priori* variances of triangulated points from the inversion of the normal equation matrix of the bundle adjustment, without performing a real measurement. It is a way of optimizing object point accuracy when minimum information about the scene is available, a process known within close-range photogrammetry as network design. The design problem (Fraser, 1984, 1996; Mason, 1995) typically divides into Zero-order (the datum definition problem), First-order (the network configuration problem), Second-order (the weight problem) and Third-order (the densification problem) design. In the following we consider only First-order design, i.e. the optimization of observation configurations by (i) simulating a flight, (ii) generating image observations by back-projecting 3D points to all cameras/images and (iii) performing a bundle adjustment to retrieve the standard deviations from the diagonal elements of the symmetric covariance matrix. Object points, camera positions and orientations are regarded as free parameters. The estimated precision of object points is based on the following standard errors – image coordinate 1 pixel, camera position 0.1 m, camera orientation 0.001°. Collinearity equations constitute the basis for the 3D restitution.

The simulated scenarios are imaged with a *maltese-cross* system and vary according to the following parameters (Table 1 and Fig. 1): sensor size, forward and side overlap, camera focal length and field of view. The flying height is adapted in order to keep the Ground Sampling Distance (GSD) computed in the nadir images constant. The choice and values of particular parameters reflect the current market offerings and the most common flight practices. To distinguish between the sensors, they are referred to as small, medium and large although the size of the latter would normally be classified as medium. The unconstrained scenarios reflect the situation where observations from all cameras are accepted, while in the constrained scenarios only observations in the nadir cameras and oblique cameras of same look direction are considered, as discussed in Section 3. In order to account for possible occlusions that influence point observability, normally prevailing in urban environments, a complex of buildings (Fig. 4a, street width of 5 m and 10 m, building height 20 m) was embedded in the simulated scene. Every point back-projected to a camera is verified whether obstructing with the scene or not. Further constraint is imposed on the oblique cameras, specifying that object points can be observed only from cameras that look the same direction (see also Section 3). This tolerance is full-fledged given the large perspective distortion between slanted views that hamper the matching of homologous points (Rupnik et al., 2013).

2.2. Precision, productivity, completeness

2.2.1. Precision

According to the data contained in Table 1, presented case studies involve the *maltese-cross* system with small (A), medium (B) and large format (C) cameras. Results presented in Fig. 3 show the relationship between object point precision (computed from the covariance matrix) and factors such as overlap, camera tilt angle, and imaging ray redundancy. Due to the fact that the scale of the nadir images is kept constant for all case studies, the figures of precision do not vary significantly between the cases.

Within the different imaging configurations, the overlap and tilt angle of the oblique cameras, and the imposed constraint, largely influence the results. It can be noted (Fig. 3) that the increased tilt angle contributes to:

- a larger point redundancy (observations),
- improved precision, i.e. a better standard deviation of 3D points,
- a homogenous standard deviation in all three dimensions.

However, imposition of the look direction constraint causes:

- degradation in precision along the Z axis through a decrease in both the number of observations and the intersection angles,
- inhomogeneous planimetric precision (especially noticeable for large tilt angles) due to the different intersecting angles in the XZ and YZ planes.

Compared with the classical scenario of exclusively nadir images (Fig. 3), it can be observed that while the planimetric precision is only slightly improved, the standard deviation in the Z direction (height) is 2–3 or 4–5 times better, depending whether the look direction constraint is enabled.

2.2.2. Productivity

Increasing overlap between subsequent acquisitions and strips is costly in the image collection time and post-processing but causes substantial gain in precision, as proves Fig. 2. The figure displays the precision differences between the overlap 60/40 and 80/
Throughout all the scenarios, flying with an 80/60 overlap pattern, rather than 60/40, delivers standard deviations better by ca. 50% along X, Y, Z directions. On the other hand, when comparing the standard deviations between 80/60 and 80/80 overlap, a precision increase of only about 20–30% is observed. To give more meaning to the figures, they should be “weighted” by the additional number of acquisitions needed to densify the data collection, especially that the number grows exponentially (Fig. 2d). Confronting the two, the rationality suggests that the 80% along- and 60% across-track patterns would be the golden mean as the return in precision for the cost of extra labour is substantial. Having said this, it must
be underlined that dense matching in oblique datasets is feasible from images of the same perspective. This is analogous to the automatic tie point extraction step where SIFT-like features are matched across views of similar look direction (see Section 3). It implies that both across- and along-track overlaps are equivalent and favourably equal to 80%.

With the GSD size equal for all cases, the variable sensor size affects only the survey productivity. In Fig. 3 the redundancy figures of the C camera system are in the majority of cases 20–30% smaller than in A and B. Notably, the measures of precisions do not drop accordingly (see also Fig. 2).

2.2.3. Completeness
Besides the precision and productivity aspects, sensor size and tilt angle are also important from the standpoint of the completeness of the reconstructed scene. Increasing the value of both parameters enlarges the camera’s field of view (FOV) and provides for a bigger GSD, respectively (assuming fixed focal length). But, larger FOVs contribute to more severe occlusions – an undesirable artefact in image-based 3D city modelling (Haala and Kada, 2010). Because of this, as a part of the project planning, attention should be paid to careful adaptation of the camera platform parameters to the situation in the field, i.e. to the topography of the city to be flown over.

Fig. 4b–e exemplifies the effect of the tilt angle on the visibility of two selected façades. Both façades are oriented parallel to the flight direction, and the width of the street in front of them are 10 m (building E10) and 5 m (building E5). The simulated example was derived from the flight with 80/60 overlap, surveyed with the camera system of type A. Due to the occlusion effects, the tilt angle of 45° introduces inhomogeneous distribution of redundancy along the façades. The 30° tilt angle results in fewer but uniform redundancies. The width of the street and the building height play a major role in planning a successful urban survey campaign. The taller the city architecture the lower the camera incidence angle should be. Therefore, a compromise ought to be found between the camera tilt setting, given focal length, sensor size, overlap and the geometry of the surveyed area.

3. Image block orientation
The processing workflow for oblique imagery can be divided in sequential steps (Rupnik et al., 2014), namely image orientation, dense matching and point cloud filtering, feature extraction, orthoimage generation and monoplotting (Murtiyoso et al., 2014). While this follows the traditional photogrammetric processing pipeline, the algorithms for the orientation and matching phases need to be slightly adjusted in order to cope with oblique image datasets. The orientation task is recapped below. Dense matching and feature extraction issues are discussed in Nex et al. (2013).

In the image orientation process, the interior and exterior parameters are nowadays often known a priori, as retrieved with a prior calibration procedure or measured directly with on-board sensors (GNSS/IMU), respectively. Nonetheless, these parameters are generally regarded as approximate if one has metric and automatic applications in mind. An adjustment in a least squares sense is therefore mandatory, requiring linearization of the equations and good approximate values of unknown parameters. On the basis of image data (homologous points), these parameters are (i) retrieved with direct methods (e.g. resection, perspective-n-point problem) individually for each image, (ii) concatenated to a common reference system with the help of the fundamental/essential matrix or trifocal tensor, and eventually (iii) refined inside the bundle adjustment. The homologous points...
are found automatically across multiple aerial views with standard area-based image matching techniques (such as Normalized Cross Correlation and Least Squares Matching) or feature-based methods (SIFT, SURF, etc.) (Apollonio et al., 2014) coupled with robust estimators to remove possible wrong correspondences (RANSAC, Least Median of the Squares, etc.).

Oblique imagery is richer in content compared to nadir imagery. Objects (building, roads, etc.) are recorded at different scales, and object features are back-project under different projective transformations. Also, due to different viewing directions grave occlusions are exhibited. Such configurations, are characteristic of unordered terrestrial image networks (Barazzetti et al., 2010; Pierrot-Deselligny and Clery, 2011) and the generation of putative correspondences between images is therefore prone to a greater number of mismatches. Indeed, the literature reports that traditional photogrammetric software is optimized for nadir image acquisition and generally performs poorly for other imaging geometries (Jacobsen, 2008; Gerke and Nyaruhuma, 2009).

In the developed processing pipeline, the computation of approximate orientation parameters is guided with the help of a connectivity graph (Rupnik et al., 2013). The connectivity between images refers to a graph with nodes and edges being representations of images and their relationships. Two images are linked with an edge if and only if they are spatially compatible. GNSS/IMU data
beneficially provides the input information for establishing the image associations. Traversing across the connectivity graph helps in speeding up the extraction of image correspondences and in reducing the number of possible outliers. There are three conditions (constraints) to be fulfilled for an image pair to be regarded as compatible:

### Table 2

Dataset specifications (N and O stand for nadir and oblique). The given overlap is calculated on the nadir images.

<table>
<thead>
<tr>
<th>Data</th>
<th>Camera</th>
<th>Sensor size (mm)</th>
<th>Focal N/O (mm)</th>
<th>FOV along/ across N (°)</th>
<th>Image scale</th>
<th>GSD N (m)</th>
<th># images N/O</th>
<th>Y/X overlap (%)</th>
<th>Area (km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Milan</td>
<td>Midas 5</td>
<td>36 × 24</td>
<td>80 100</td>
<td>17/25</td>
<td>-15,000</td>
<td>-0.10</td>
<td>125/375</td>
<td>70/30</td>
<td>3 × 5</td>
</tr>
<tr>
<td>Graz</td>
<td>Vexcel Osprey I</td>
<td>70 × 45 (nadir): 23.5 × 36 (left/right) 71.5 × 23.5 (for-/backward)</td>
<td>51 48/70</td>
<td>-21,000</td>
<td>-0.12</td>
<td>20/160</td>
<td>75/65</td>
<td>3 × 1.5</td>
<td></td>
</tr>
</tbody>
</table>

### Table 3

A summary of all performed tests. ID refers to a name given to a particular test. The columns indicate the adjustment parameters used as constants, unknowns, observed unknowns, and observations. FX means fixed IOR, FR refers to self-calibrating bundle adjustment (i.e. free IOR).

<table>
<thead>
<tr>
<th>ID</th>
<th>Input images</th>
<th>Constants (C) unknowns (U) observations (O)</th>
<th>Exterior orientation EOR</th>
<th>Interior orientation IOR + AP</th>
<th>GCPs, observed in nadir</th>
<th>GCPs, observed in oblique</th>
<th>Tie points</th>
</tr>
</thead>
<tbody>
<tr>
<td>1FX</td>
<td>Nadir</td>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>O</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2FX</td>
<td>Nadir + oblique</td>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>O</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3FX</td>
<td>Nadir + oblique</td>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>O</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1FR</td>
<td>Nadir</td>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>O</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2FR</td>
<td>Nadir + oblique</td>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>O</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3FR</td>
<td>Nadir + oblique</td>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>O</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a GNSS observations are used inside the bundle block adjustment as observed unknowns only in the GNSS-assisted scenario.
b Tie points in image space treated as observed and in object space as unknown.

Fig. 7. Overlaps in forward/backward and side oblique cameras for Midas 5 (left) and UltraCam Osprey I (right) systems corresponding to 70/30 and 75/65 overlaps in nadir views. The overlap in the oblique views is given as a ratio of the complete length of the median of a trapezoid (forward/backward camera in cross-track overlap and side camera in along-track overlap) or the axis of symmetry (forward/backward camera in along-track overlap and side camera in cross-track overlap), to its length between subsequent acquisitions (red line). It is computed in along- (horizontal red line) and cross-track (vertical red line). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
Table 4
The computed a posteriori standard error of unit weight $\sigma_u$ (pixel), for the different configurations of ground control information reported in Figs. 5 and 6. Note that the sigma values are the same for both, with and w/o the GNSS support.

<table>
<thead>
<tr>
<th>Data</th>
<th>GCPs config</th>
<th>1FX</th>
<th>2FX</th>
<th>3FX</th>
<th>1FR</th>
<th>2FR</th>
<th>3FR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Milan</td>
<td>0–4</td>
<td>0.49</td>
<td>0.47</td>
<td>0.55</td>
<td>0.47</td>
<td>0.45</td>
<td>0.52</td>
</tr>
<tr>
<td>Graz</td>
<td>1–4</td>
<td>0.38</td>
<td>0.36</td>
<td>0.39</td>
<td>0.31</td>
<td>0.29</td>
<td>0.31</td>
</tr>
</tbody>
</table>

a Their ground footprints coincide by a given percent.
b Look directions of the images are similar or one is nadir-looking (similarity can be then violated).
c The number of extracted homologous points for the pair is above a given threshold.

given a graph complying with conditions a and b, feature extraction is allowed between images with at least two edges. Next, the edges of the graph are enriched with another attribute, i.e. the number of extracted tie points, and all images are concatenated. Concatenating images for the initial solution is understood to (i) find triplets of images within the connectivity graph to be used for computing approximate orientations, and (ii) providing all the triplets a sequence in the concatenation order. A bundle adjustment controlled in this way minimizes the risk of divergence by maximizing the similarity of images within particular triplets, hence maximizing the ratio of good to bad matches and ensuring the block’s cohesion.

As a last step, the control information in form of GNSS/IMU data and/or GCPs is included in the adjustment computations. With this, the data is transformed to the coordinate system of the control information. The quality of the resulting orientation is strongly coupled to the spatial distribution of the control information, its “observability” in the images as well as the importance given in the adjustment. It should be remarked that if the surveyed area surpasses a certain size, the Earth curvature effects must be considered.

4. Real case studies

4.1. Datasets

The first dataset (Milan) comprised 550 images and was acquired with a Midas 5 multi-camera system (one nadir and four oblique looking cameras tilted at 45°). The image block covered an area of ca. 3 km by 5 km that was flown with 70/30 (Y/X) overlap related to nadir images (Fig. 5). In order to evaluate the quality of the block orientation results, 25 GCPs were GNSS-surveyed with a mean positional accuracy less than 5 cm.

The second dataset (Graz) comprised 180 images and was acquired using an UltraCam Osprey I system (one nadir camera with six oblique-looking cameras tilted at 45° along the four cardinal directions – see Gruber and Wolfgang (2014)). The area covered ca. 3 km by 1.5 km and was flown with an overlap of 75/65 (Y/X) related to nadir images (Fig. 6). In a similar manner to the Milan case study, a metric evaluation of the achieved results was performed using ground truth information in form of GCPs and check points (CP) GNSS-surveyed to a mean accuracy of 5 cm.

In both datasets the approximate exterior orientation parameters (EOR) were available as GNSS/IMU observations. The dataset specifications are reported in Table 2. The structure of the flights resemble the comprehensive block specification of Honkavaara et al. (2006a), without crossing flight lines. Corresponding overlaps computed for oblique images are presented in Fig. 7.

4.2. Tests

Two pieces of control information (GCPs and GNSS observations) were incorporated in the computations, and regarded as soft constraints, i.e. treated as unknowns and observations (so-called observed unknowns), therefore assigned corrections within the bundle adjustment. The blocks were oriented using different configurations of GCPs (yellow circles in Figs. 5 and 6) and independent CPs (red markers in Figs. 5 and 6). For both datasets, configuration 1 shows a good GCP distribution along the borders of the block, whereas configurations 2, 3 and 4 are examples of unfavourable control point distributions. Although the latter may be seen as unrealistic cases, they are shown for the sake of completeness and to stress the role of ground control information in the bundle solution. Thanks to the large number of points surveyed in the Milan test area, an additional configuration was added (configuration 0). It comprises a good GCPs distribution across the whole area of interest. A summary of the performed tests is provided in Tabs. 3 and 5, where abbreviated descriptions of the different processing scenarios are listed as well.

The initial values of internal orientation parameters (IOR) and additional parameters (AP) were taken from a calibration certificate provided by the company executing the flight. These input data can be either held fixed within the bundle adjustment process or included as unknowns in a self-calibration procedure. In our tests, the scene contains significant height differences – leading to variable scale across images, hence both possibilities were tested. A version of the Fraser calibration model (Fraser, 1997), involving three coefficients of radial distortion and two coefficients of decentering distortion, was used in the experiments. While for middle-format cameras this could not be sufficient and more sophisticated approaches are available (Cramer, 2009), the camera calibration issue was not the primary scope of this article.

In the following, two measures of adjustment quality are reported, the sigma naught ($\sigma_0$) and the root mean square values (RMS), defined as the square root of the mean squared difference between nominal values and corresponding adjusted observations. Moreover, to give a more global impression of the possible deformations within the image block, the resulting sparse point clouds are compared to a reference point cloud, which was chosen to be the one exhibiting the optimal measures of accuracy against CPs. Network 3FR in GCPs configurations 0 (Milan) and 1 (Graz), i.e. self-calibrating condition and good distribution of GCPs was selected. The imaging geometry corresponds to the constrained case of Section 3, i.e. tie point observations were allowed from all nadir images and those oblique images with the same viewing directions.

The purpose of the performed tests is to investigate how the accuracy of triangulated 3D points is affected by the following factors: (i) introduction of oblique imagery in the image block, (ii) self-calibration, (iii) spatial distribution of GCPs and (iv) GNSS support. All tests were conducted using MicMac, an open-source software suite for bundle block adjustment and dense image matching (Pierrot-Deseilligny and Paparoditis, 2006; Pierrot-Deseilligny and Clery, 2011).

4.3. Bundle adjustment results

MicMac solves the orientation task within a bundle adjustment by means of weighted least mean squares. Classically, the imaging function $f$ (see Eq. (1)), i.e. the functional model of the adjustment, is defined in terms of the following parameters: rotations $R$, perspective centers $C$, camera interior parameters $l$, and object coordinates of the tie points $P$. The optimal values of the parameters are

\[ f(R, C, l, P) = 0 \]

For interpretation of color in Figs. 5 and 6, the reader is referred to the web version of this article.
found by minimizing an objective function (Eq. (3)) which is the sum of squared discrepancies multiplied by a weighting function $\rho(v)$, see Eq. (4). The weight function is composed of three terms: $\frac{1}{\sigma_{\text{obs}}^2}$ that weights observations given their a priori accuracy, $\frac{1}{\sqrt{1+\beta^2}}$ that limits the influence of a particular group of observations on the solution, and $\frac{1}{\sqrt{1+(\frac{l}{\sigma^2})^2}}$ that “penalizes” observations with large residuals making it a robust estimation. The normalizing factor $\mu$ is dataset-dependent and should be set according to the estimated observation accuracy. As a general rule, its value will change from iteration to iteration. The normalizing factor should be high when unknowns are poorly approximated and recede along the iterations. If a residual exceeds the user-defined $B$ parameter, the weight takes on the value zero, hence eliminating the influence of that particular observation in the adjustment. An in-depth explanation of the adopted weighting scheme can be found in (Pierrot-Deseilligny, 2014).

$\sum \rho(v) v^2 \rightarrow \text{minimum}$

$\rho(v) = \frac{l}{\sigma_{\text{obs}}} \frac{N_{\text{max}}}{N_{\text{max}} + N_{\text{obs}}} \frac{1}{\sqrt{1+(\frac{l}{\sigma^2})^2}}$
pling distance can be larger than that for the nadir looking camera. Furthermore, although freeing camera parameters can help to compensate for systematic lens- and sensor-related errors, the improvement in terms of a reduced $\sigma_0$ is minor. Finally, from scenario to scenario, $\sigma_0$ remains nearly constant and a plausible explanation for this is that even if the GCPs introduce “strain” into the block, resulting in locally larger residuals, the robust weight function (Eq. (4)) diminishes their effect on the sigma naught value.

Fig. 8. Discrepancy vectors at GCPs (red triangles) and CPs (open circles). Above: Milan dataset, GCP configurations 0 and 3, with GNSS. Below: Graz dataset, GCP configurations 1 and 3, with GNSS. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
4.3.1. Influence of oblique views

It is evident from the RMS values in Table 5 that for the Milan dataset, observations in oblique views improve object point accuracy in height (compare $1FR$ and $1FX$ with $2FR$/$3FR$ and $2FX$/$3FX$), which is consistent with expectations from the simulations presented in Section 2. For the Graz dataset the advantage of oblique views is less obvious but nevertheless visible especially in the Z direction. What is symptomatic is the fact that in GCP configurations 3 and 4, i.e. the unfavourable point distributions, nadir images ($1FR$) deliver either better or the same Z-accuracy results as $2FR$ or $3FR$ with the oblique views included. The exclusively nadir block seems to be more stable in the event of deficient ground control information, which is opposite to the results in the Milan dataset. The different behaviour could be due to (i) different parameters of the nadir camera models, i.e. large sensor and shorter focal length of the Osprey I system provide for better imaging geometry than the small sensor and long focal length of Midas 5 (see Table 2), and (ii) different overlap of the two image blocks (75/65 versus 70/30 related to nadir cameras).

4.3.2. Influence of self-calibration

For the Milan dataset, the bundle solutions without GNSS generally show a gain in accuracy when IO parameters are treated as unknowns (self-calibration). In the GNSS-assisted solutions this evidence is less pronounced and the RMS errors at independent CPs are comparable, suggesting that (i) the available camera calibration is valid and self-calibration is superfluous, and (ii) the newly adjusted camera calibration parameters in the GCP-only cases do not correct for systematic errors stemming from physical deficiencies in the additional parameters model but, due to projective coupling with the EOR, do compensate for the suboptimal estimates of camera rotations and positions (see Section 4.4 for more evidence). For the Graz dataset, freeing self-calibration parameters in the bundle adjustment improved results in the Z coordinates for all GCPs configurations (compare all FX versus FR). The provided
camera calibration, unlike that of the Milan dataset, was not sufficient to achieve the most optimal result and self-calibration afforded superior results. As seen in the discrepancy plots in Fig. 8, the vectors in Milan dataset do not change between the cases of fixed or freed. Their lengths in the Graz dataset, however, clearly recede if camera calibration parameters are re-adjusted.

4.3.3. Influence of GCPs distribution and GNSS observations

The GCP distribution plays a crucial role in determining CP accuracy. So long as the distribution of control information is even and appropriately dense, the constraint on CPs is satisfactory (GCP configuration 0 and 1 in Table 5). As soon as the GCP configuration gets sparser, extrapolation effects emerge and are reflected in larger RMS error values (GCPs configuration 2, 3 and 4 in Table 5 and Figs. 8 and 10). The inclusion of GNSS constraints largely mitigates the problem (GCPs configuration 2, 3 and 4 in Table 5, values in brackets). Under poor GCP distributions, the Z coordinate suffers the most, and this is the case for both datasets. In GCP configuration 3 the blocks “wobble” in counter directions (compare the RMS in X and Y in Table 5). Nonetheless, although the Milan dataset appears highly sensitive to GCP variations, the sensitivity of the Graz network is minor, especially when GNSS constraints are included. Again, the imaging geometry and extent of image overlap are the most probable factors influencing this sensitivity. In the Milan case study, switching on the GNSS constraints in one of the “good” scenarios (3FR, GCPs configuration 1) leads to larger coordinate discrepancies at CPs (0.20 m against 0.13 m). This behaviour could be a sign of “tension” between ground control information and perspective centers. As a remedy, less optimistic weights for the GNSS data should be used within the bundle adjustment.

4.4. Evaluation of derived sparse point clouds

No ground truth in the form of DSM data was available for either the Milan or Graz networks. Possible deformations and the level of noise in the derived 3D object coordinates (i.e. sparse point cloud) were evaluated for the cases of with and without oblique images, GNSS/IMU data and when GCP configurations changed. Conclusions were drawn from comparisons in Figs. 9 and 10.

In Fig. 9, in order to ascertain the influence of the inclusion of oblique images, scenarios 1FX and 1FR were compared with scenarios 3FX and 3FR. To free the analyses from the impact of GCP locations inasmuch as possible, results obtained in GCPs configuration 0 (Milan) and 1 (Graz) were used. As a side-product, also the influence of self-calibration on the triangulated 3D points was devised from the differences 1FX with 1FR and 3FX with 3FR.

In Fig. 10, in order to determine the influence of the GCPs distribution, the oblique image blocks 3FR and 3FX, in all GCP configurations, were compared against a reference result. The reference is unique for all comparisons – 3FR in GCP configuration 0 in Milan, and GCP configuration 1 in Graz. Left: Milan, right: Graz datasets. Note that the legends should be decoded separately for each dataset.
4.4.1. Influence of oblique views and self-calibration

The sparse point differences confirm and complement the conclusions of the previous subsection. Adopting oblique cameras reduces the magnitude of noise in 3D object space and homogenises the distribution of discrepancies within the block. Blocks containing the full set of 5 or 7 imaging cones are less sensitive to the GCPs arrangement and shorten the growth of discrepancies outside the region containing the GCPs. Bundle adjustment with self-calibration brings about further profits with regard to noise and deformations (see the patterns in 1FX – 1FR and 3FX – 3FR). The “dooming” effect that is characteristic for un-modelled radial distortion error in 1FX – 1FR of the Graz dataset (Fig. 9, bottom), provides further evidence of an outdated camera calibration and justifies the freeing of additional parameters in the bundle adjustment. James and Robson (2014) found that such systematic errors can be mitigated if oblique images are acquired in the block. Based on real and simulated studies, they showed that in the presence of radial distortion, with an assumed error-free camera model, (i) a nadir block will produce the “dome” in the DEM (see 1FX – 1FR) and (ii) an oblique block will decrease the magnitude of discrepancies (no “dome” effect in 3FX – 3FR). It can be validly claimed that under conditions of self-calibration, the systematic vertical DEM error is eliminated, for all practical purposes, which explains the small discrepancies obtained in 1FR – 3FR. Similar conclusions are reported in Wackrow and Chandler (2011), Nocerino et al. (2014). Besides the radial distortion, another cause of the systematic block deformations in Figs. 9 and 10 is possibly the un-modelled out-of-plane distortion, i.e. unflatness. Fraser (1997) warns that systematic errors caused by the sensor unflatness are critical in that the image space perturbations might be concealed while their effect on photogrammetric triangulation precision is severe. The induced radial displacement depends on the value of the incidence angle thus wide-angle lenses are especially subject to manifest the errors. The standard 8-parameter and the extended 10-parameter models are insufficient to compensate for these misalignments. Yet, they are present in today’s photogrammetric systems. In theory higher degree polynomial functions could remove all inaccuracies and thus more suitable for the employed sensors.

Development of new mathematical models is therefore encouraged (Honkavaara et al., 2006b; Tang et al., 2012).

4.4.2. Influence of GCPs distribution and GNSS observations

In both datasets the “strain” that the GCPs induce on the entire block is striking in all the block adjustment solutions that do not include GNSS constraints. The discrepancies are kept small within the general region enclosed by GCPs (marked with red dots in Fig. 9) whereas extrapolation errors increased away from this region. When GNSS constraints are introduced into the bundle adjustment, the magnitude of these discrepancies is substantially reduced, as is the extrapolation error. GNSS observational constraints stabilize the bundle adjustment solution, due to projection centers behaving like control points. The visible elongated deformation pattern (Fig. 9, top) arising in the comparison of the nadir image blocks (1FX – 1FR) with GNSS constraints was most likely due to dependencies between camera calibration parameters and exterior orientation parameters.

5. Conclusions

This paper has presented an analysis of oblique multi-camera systems employed for mapping applications. The simulated and real case studies demonstrate that increasing the overlap, and thus point redundancy, improves 3D precision and accuracy. The traditional 60/40 flight pattern applied to an oblique image block produces substantial enhancement in precision over the corresponding block with nadir images only. The precision in height determination is increased some two to five times compared to the nadir image case. While increasing the overlap to an 80/60 scenario can bring about further improvements, the change in precision between 80/60, 80/80 and 90/90 is less dramatic. To grasp the actual value of the increased overlap, the gain in precision ought to be confronted with the extra labour imposed by densified data collection. It has been shown that the latter grows exponentially. On the other hand, having in mind the application of 3D city modelling from imagery and knowing that the content of urban scenes is complex, with visibility restricted due to occlusions, the overlap should be kept large. It is advised that as a part of the project planning camera system and flight parameters are adapted to the topography of the city. The building height to street width ratio is an important factor in the planning phase.

If productivity is considered as the number of images needed to conduct a survey over an area, evidently larger sensors and greater tilt angles will outperform smaller sensors and tilt angles (provided that there is no change in focal lengths). Nonetheless, it has been shown that although the 45° tilt angle gives best results from the perspective of object point precision, it is more prone to occlusions. The empirical part of the paper evaluated two datasets, one from a Midas 5 camera system with small sensors (Milan dataset) and the other recorded with an UltraCam Osprey I with large sensor sizes (Graz dataset). The UltraCam Osprey I dataset outperforms the Midas 5 dataset in accuracy as a consequence of (i) larger along- and cross-strip overlap, (ii) different imaging geometry that stems from different camera internal parameters, i.e. large sensor and short focal length versus small sensor and long focal length, and (iii) the camera being designed for metrically accurate mapping purposes rather than comprising integrated cameras designed more for the consumer market. The advantage of the UltraCam system is noticeable when considering CP accuracy, which was in the same range as the GCPs measurement accuracy. Furthermore, the magnitude of the noise in the sparse point cloud comparisons was smaller. The tests conducted, however, do not constitute a basis for direct comparison of the two oblique camera systems due to (i) the different flight patterns, (ii) different extent of the regions covered, and (iii) it was not the aim of this paper to comment upon which of the two systems might be superior.

Both evaluations indicate that inclusion of oblique imagery into the bundle adjustment brings a significant gain in height accuracy, as well as improved stability of the image block when (i) the overlap is small and (ii) no GNSS or well distributed GCPs are available. Likewise, the height accuracy is improved in the self-calibrating bundle adjustment. The characteristic deformation patterns arising in point cloud comparisons revealed the remaining sensor-related errors. The applied 8-parameter camera model failed to correct all remaining systematic errors. For this reason, future work will involve the study of other additional parameter models able to remove all inaccuracies and thus more suitable for the employed sensors.

Acknowledgements

This work is partially funded by the RAPIDMAP project, a European Union (EU) funded project in the International Cooperation Activities under the Capacities Programme of the EU 7th Framework Programme for Research and Technology Development (CONCERT-Japan).

The authors are thankful to Blom-CGR S.p.A. (Parma, Italy) and Microsoft UltraCam (Graz, Austria) for providing the oblique imagery datasets processed in the presented work.
References


Karbo, N., Simmons, G., 2007. Aerial imagery from different angles. Prof. Surveyor Mag. 27, 8–11.


