# Experiences and achievements in automated image sequence orientation for close-range photogrammetric projects

Luigi Barazzetti [a], Gianfranco Forlani [b], Fabio Remondino [c], Riccardo Roncella [b], Marco Scaioni [a]

[a] Politecnico di Milano, B.E.S.T. Dept., P.za Leonardo da Vinci 32, Milan, Italy
(luigi.barazzetti, marco.scaioni)@polimi.it

[b] Università degli Studi di Parma, D.I.C.A.T. e A. Dept., Viale Usberti 181/A, 43124 Parma, Italy
(riccardo.roncella, gianfranco.forlani)@unipr.it

[c] Bruno Kessler Foundation (FBK), 3D Optical Metrology Unit, via Sommarive 18, Trento, Italy
remondino@fbk.eu

## ABSTRACT

Automatic image orientation of close-range image blocks is becoming a task of increasing importance in the practice of photogrammetry. Although image orientation procedures based on interactive tie point measurements do not require any preferential block structure, the use of structured sequences can help to accomplish this task in an automated way. Automatic orientation of image sequences has been widely investigated in the Computer Vision community. Here the method is generally named "Structure from Motion" (SfM), or "Structure and Motion". These refer to the simultaneous estimation of the image orientation parameters and 3D object points of a scene from a set of image correspondences. Such approaches, that generally disregard camera calibration data, do not ensure an accurate 3D reconstruction, which is a requirement for photogrammetric projects. The major contribution of SfM is therefore viewed in the photogrammetric community as a powerful tool to automatically provide a dense set of tie points as well as initial parameters for a final rigorous bundle adjustment. The paper, after a brief overview of automatic procedures for close-range image sequence orientation, will show some characteristic examples. Although powerful and reliable image orientation solutions are nowadays available at research level, there are certain questions that are still open. Thus the paper will also report some open issues, like the geometric characteristics of the sequences, scene's texture and shape, ground constraints (control points and/or free-network adjustment), feature matching techniques, outlier rejection and bundle adjustment models.

**Keywords:** Automation, orientation, close-range photogrammetry, image matching, calibration

## 1. INTRODUCTION

Automatic orientation of targetless close-range image is still a major research topic in Computer Vision (CV) and photogrammetry, although with somehow different goals (visual vs. metric reconstruction) and/or assumptions (e.g. uncalibrated camera vs. known calibration). By far the largest contribution to the progress in the field comes from CV, with the developments in automated projective reconstruction [1-3], the introduction of image descriptors [4-6] and robust estimators [7-9] and the realization of free tools for image-based rendering, internet-related 3D applications or location-based services (e.g. ARC3D, Photosynth, etc.) They all contributed to a strengthening of the mathematical, statistical and geometrical foundations of techniques for image orientation [10].

In a strictly photogrammetric environment, automation in targetless image orientation (and surface reconstruction) followed the achievements in geometric Computer Vision and was primarily driven by the increasing use of Terrestrial Laser Scanning (TLS), in order to preserve or regain market share in close-range surveying and 3D modeling applications.

At the end of the 90's computer vision researchers started to estimate automatically and simultaneously camera interior and exterior parameters together with 3D object coordinates from long image sequences acquired with uncalibrated

Videometrics, Range Imaging, and Applications XI, edited by Fabio Remondino, Mark R. Shortis,
Proc. of SPIE Vol. 8085, 80850F · © 2011 SPIE · CCC code: 0277-786X/11/$18 · doi: 10.1117/12.890116

video-cameras [1]. The key of the success was the very short baseline which endorsed the automated extraction of image correspondences for a successive global image registration in a projective and then Euclidean space [2]. Hence the name "Structure from Motion" with the main aim being the 3D reconstruction of the scene recorded moving the camera by on operator. The photogrammetric community has adopted the short-baseline concept only recently, mainly to increase the level of automation in the tie points identification, but keeping always an eye on the accuracy of the final results, thus a tradeoff between very short baselines (i.e. fully automation possible but theoretically worst precisions) and wide-baseline (i.e. fully automation not feasible but higher precision) is inevitable.

The consequences of the increasing success of image orientation algorithms are slowly appearing and, as far as photogrammetry is concerned, may lead to profound changes. Although the goal of complete automation in the photogrammetric pipeline is still far away (and probably not fully required), especially as far as map production is concerned, it removes one of the most critical stages in the 3D restitution procedure. For non-expert users, automated image orientation is a very welcome step, probably much more important than automation in camera calibration. Thus with well-designed interactive tools to support the plotting and mapping phase, completing a photogrammetric survey, at least for not-too-complex objects, is now a task open to a larger group of potential users.

Optimization of precision and reliability of photogrammetric networks has been deeply investigated in aerial photogrammetry as well as in industrial close-range photogrammetry [11]. This topic should also be part of the processing pipeline of automatic orientation. Network design, including ground control point density and distribution, must also be discussed in the new framework: best practice rules [12-13] in "standard" cases might be reviewed or adapted to newly acquisition procedures based on convergent and largely overlapping images. The large variety of architectural shapes to be surveyed and modeled makes it difficult to find simple rules to ensure the block quality as well as to run realistic simulations prior to the actual survey, although complex surveys can be broken down in smaller elementary ones. It is important, however, to provide guidelines for image acquisition when using automatic procedures for image orientation, because this has consequences for both processing time as well as, more fundamentally, for the block geometry characteristics and therefore for the quality of the restitution. Finally, it is also important to look at the influence that a larger-than-necessary image overlap (as it is normally the case) might have on e.g. DSM generation using multi-image techniques.

The paper reviews the latest advances in automated orientation of targetless images, with particular focus on the methodology developed by the authors. Section 2 presents the key features of the implemented tie point extraction strategies with some details about software implementation. In section 3 some examples of automatic image orientation of blocks with different geometric configurations are reported to highlight the variety of cases where the developed methodology can be successfully applied. In section 4 a discussion on the open issues is presented while section 5 concludes the article.

## 2. PROCEDURES FOR AUTOMATIC IMAGE ORIENTATION

Many approaches have been developed in the past years in the photogrammetric and CV communities for the automated orientation of large and complex image sequences. The former discipline focuses primarily on precision and reliability of the results [14-16]. The latter is more concentrated in the orientation of large blocks (thousands of images) taken with uncalibrated cameras without any metric purpose except 3D visualization and image browsing [17-19].

The main problems preventing the presence of reliable and precise commercial approaches are the presence of convergent images, unpredictable baselines and scale variations, lighting changes, repetitive patterns, homogeneous textured areas, etc. These effects, along with complex object configurations, give strong perspective deformations and make the automated tie point identification a challenging task, especially compared to aerial photogrammetry applications. Today Automatic Aerial Triangulation (AAT) has reached a significant level of development and several software packages are available on the market [20]. This is a fundamental difference with respect to close-range applications, where there is only one commercial package (PhotoModeler 2011, released at the end of October 2010) capable of performing this task without placing markers in the scene.

In recent years some research solution were developed in the field of close-range photogrammetry [21-24]. Although all algorithmic implementations are based on different methods, they follow a general scheme (Fig. 1). The identification of the image correspondences is carried out by using particular combinations of image pairs or triplets, that are extracted from the original block according to its network geometry. These are then processed with *area-based matching* (ABM) and *feature-based matching* (FBM) algorithms, the latter invariant to variations in scale and rotations (around the optical axis of the camera) and thus more robust to affine deformations. During the FBM phase several procedures developed in CV are employed to obtain a linear formulation of the relative orientation problem [25]: the *fundamental matrix* is used for image pairs, while for image triplets the *trifocal tensor* remains the best choice. As automated approaches can produce several mismatches, robust estimators [8] are usually used to find and remove any possible outlier. The following step is the exhaustive analysis of all image combinations in order to increase the multiplicity of the generic point and to find the image coordinates for the whole block. Exterior orientation parameters can then be estimated with a *photogrammetric bundle adjustment*. An external datum is fixed with additional metric data, such as ground control points (GCPs) or GPS/GNSS information. Alternatively, an inner constraint is used to remove the rank deficiency of the bundle solution (*free-network adjustment*). In the next sections, these phases are illustrated and discussed more in detail.

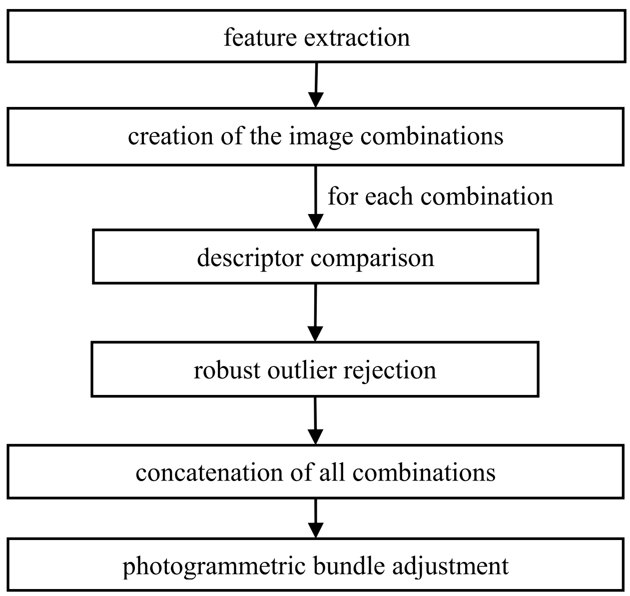



Fig. 1. Generic workflow for automated image orientation in close-range photogrammetry.

### 2.1 Image acquisition

A very common image configuration in close-range photogrammetry is the *sequence*, i.e. a single strip whose images feature a large overlap. An *ordered image sequence* (Fig. 2a-b) has a proper network geometry which simplifies the identification of the image correspondences which leads to a reduction of the CPU time. A large overlap (i.e. short baselines) between consecutive images allows a reduction of the perspective differences and thus better matching results, limiting the number of outliers. If a single strip is not sufficient to cover the entire object, multiple sequences can be combined to obtain a *regular block* (Fig. 2c). In some cases it is quite difficult to acquire regular blocks (e.g. UAV systems without autonomous navigation devices or images acquired by non-experts) and the resulting network geometry could feature a unordered distribution of the images which form a *sparse block* (Fig. 2d). These last situations give an increment of the computational cost because there is no preliminary information about the image combinations that share tie points: it is thus necessary to process all image pairs. In [26] a method based on the generation of a visibility map is proposed. It provides a sort of connectivity graph between the images, which can be created with some navigation information and an approximate model of the object. Alternatively, a quick processing of compressed images can help in determining the initial pairs.

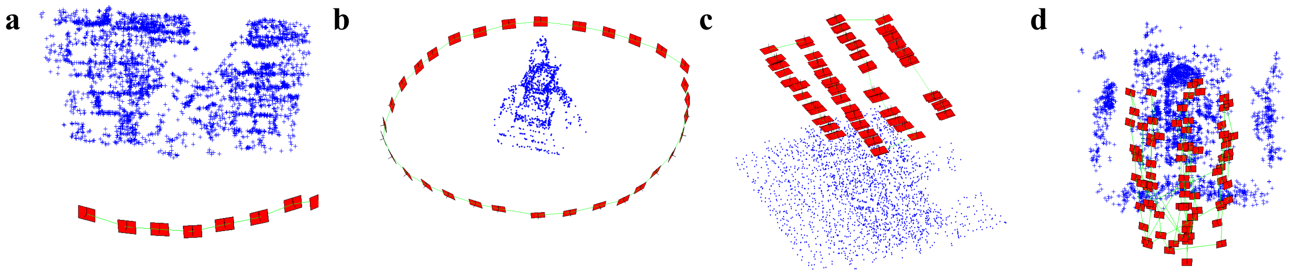



Fig. 2. Examples of network geometries: a) open sequences, b) closed sequences, c) regular block, d) sparse block.

### 2.2 Tie point extraction with FBM

Most FBM techniques used for image orientation are based on the use of detector/descriptor operators. The detector is capable of finding interest points in the images, while the descriptor associates a vector of information to each single detected point. One of the most valuable property for an automated operator is its repeatability, which means the capability of extracting the same feature under different viewing conditions. The use of these kinds of operators provide several features for each image. Some points will be visible for an image only, whereas others will be matched in two or more images. During the matching phase homologous points will be labelled in order to obtain a structured system similar to those used in aerial photogrammetry.

Nowadays there exist two feature-based operators generally employed for automated image orientation: SIFT [4] and SURF [6]. The *Scale Invariant Feature Transform* (SIFT) algorithm provides highly distinctive features invariant to image scaling and rotations. The *Speeded Up Robust Features* (SURF) algorithm was designed to be a fast distinctive point detector and descriptor. From the practice, SURF gives comparable results to SIFT but it has a shorter computational time. Both SIFT and SURF associate a descriptor to each extracted image feature. The descriptor is a $n$-dimensional vector (in general $n$ is equal to 64 or 128) that describes the feature. Corresponding points can be found by simply comparing the descriptors, using rigorous (slow) procedures (e.g. *quadratic matching*) or fast but approximate methods (e.g. a *kd-tree search*).

### 2.3 Outlier rejection

FBM operators like SIFT and SURF can provide a good number of corresponding points. However, some outliers can be found and should be removed from the dataset. Most implementations uses robust estimation strategies based on the epipolar constraint encapsulated into the fundamental matrix **F** or an essential matrix **E** (if interior orientation parameters are known). Both matrices provide a linear relationship between the image points of a stereo pair, without requiring any information about the object.

The **F**-matrix [3] is generally estimated using 7 corresponding points [27] while the **E**-matrix needs only 5 image correspondences [28]. Other (more photogrammetric) methods are instead based on the exhaustive analysis of the relative orientation solution, that provides multiple results among which the real one must be selected [29].

The study of the epipolar geometry of an image pair is not always sufficient to discard all wrong points: mismatches lying on the epipolar lines cannot be detected with this approach, in particular in presence of repeated patterns. The use of the *trifocal tensor* **T** gives better results, especially in case of sequences. The trifocal tensor allows the analysis of an image triplet using the intersection of three light rays and can be estimated from 7 image correspondences. This condition is not equivalent to the estimation of two (or three) fundamental matrices as there is a projective ambiguity. The image configurations which defeat the trifocal tensor are rare, therefore the procedure is more robust. Alternative approaches for the analysis of a generic triplet exploit the model coordinates estimated for two image pairs. These 3D points should be equivalent except for an overall ambiguity, that is a 7-parameter transformation. It is also possible to estimate a local bundle solution for the triplet, where outliers are removed with *data snooping* techniques.

The computations of **F**, **E** or **T** need to be combined with robust methods as they can detect any possible outlier in the dataset. The percentage of mismatches can be relevant and *high breakdown point estimators* are usually employed. These methods use the minimal number of data to estimate the model, then the remaining points are compared to verify their consistence. This process is repeated several times according to the probability of finding a dataset that does not contain outliers. There exist several approaches to complete this task. For instance, RANSAC estimates the number of trials and then extracts a minimum number of correspondences to calculate the model. It can deal with a significant number of outliers but needs a preliminary threshold *T* according to the data precision. The *Least Median of Squares* (LMedS) [8] has a lower breakdown point (50%), but does not require any setting of the threshold *T*. MAPSAC [30] solves the problem related to the threshold *T* of RANSAC. As a high *T* leads to weak solutions, MAPSAC associates to the inliers a score that depends on how well they fit the data, while in the standard RANSAC formulation they have a constant score. It has been demonstrated that MAPSAC leads to a significant improvement of the solution [24].

### 2.4 Tie point concatenation and bundle adjustment

Once corresponding points for each image pair (or triplet) are available, they are organized into tracks and the comparison of the numerical values (pixel) of all image points give the set of image correspondences for the entire block or sequence.

After the matching of all image pair combinations, the number of tie points can be reduced according to their multiplicity (i.e. the number of images in which the same point is visible). A regular grid is projected onto each image and for each cell only the point with the highest multiplicity is stored [14, 26]. This is useful for large blocks composed of images with a good texture.

The detected image correspondences are then used for the image orientation step within a rigorous photogrammetric bundle adjustment. Although the robust estimators described in the previous section can remove a significant number of wrong correspondences, *data snooping* techniques are often necessary to refine the results and eliminate, within the bundle iterations, the remaining outliers. With the bundle adjustment the user can derive the exterior orientation parameters and 3D object coordinates of the measured tie points and then proceed to further processing for DSM generation, orthophoto production, feature extraction, etc.

## 3. EXAMPLES

In this section some examples are shown to witness the impact of new orientation techniques in different applications. More specifically, photogrammetry is receiving large attention in both geological and geotechnical fields, especially for surface reconstruction issues, while great interest is also returning from the architectural community. Range-based methods (in particular terrestrial laser scanning) are able to provide dense and accurate point clouds and can be considered the current state of the art for such applications. However, their high costs and difficult portability make image-based methods an efficient and economical alternative [31-33] (Example 1), in particular when orthophotos are needed.

UAVs are becoming a very popular image recording system as they offer great surveying advantages with respect to traditional aerial flights. UAV blocks have a hybrid network geometry between aerial and terrestrial blocks. Therefore, most software packages for automatic aerial triangulation cannot be employed (Example 2).

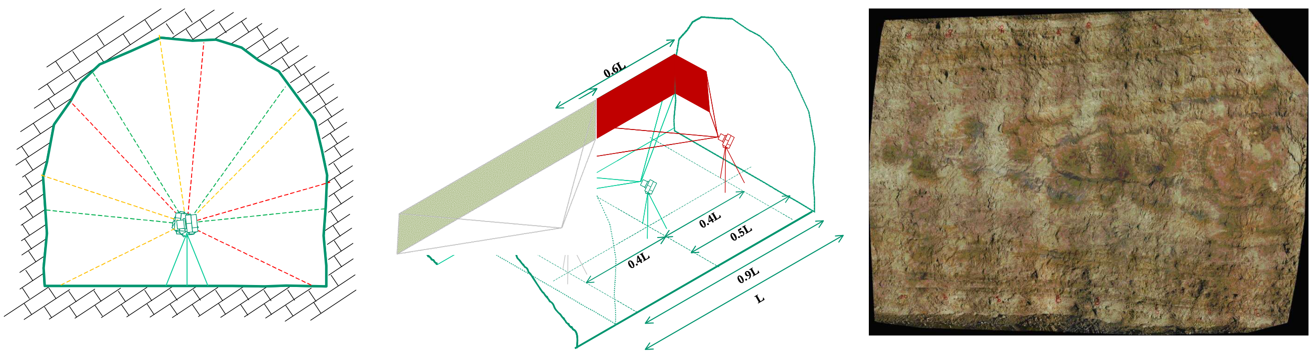
Finally, two examples (3-4) from mobile mapping sequences and complex architectural areas are reported to show the potentials of the developed automated orientation techniques even in case of complex and long image sequences.

### 3.1 Example 1 - Underground quarry

In the last few years a growing demand of high-resolution DSMs for stability analysis of rock faces in rock mechanics applications has occurred. Successful applications of photogrammetry to this aim has been already reported in [33-34]. Recently the method has also been applied to underground quarries. Since the excavation usually proceeds in a continuous cycle, survey time should be minimized. In the case study presented here the photogrammetric survey covered a stretch of tunnel of ca. 130 m. The gallery cross-section has abutments about 3 m high and a semicircular vault with a radius of about 8 m. The resolution required for the DSM was 5 mm, while the accuracy needed to achieve enough information on the small rock discontinuities was of a few millimeters in each model. A Nikon D100 camera (6 Mpx) with a focal length of 18 mm was used for the survey providing an object resolution of 2 mm/pix on the abutment and 3 mm/pix on the vault.

To limit the acquisition time, a device mounted on a tripod was built to quickly rotate the camera to preset zenithal angles, so that the appropriate overlap between the longitudinal strips can be obtained (Fig. 3). The tripod is positioned in the longitudinal plane of symmetry of the tunnel, with the head mount oriented in such a way that the camera rotates in a plane vertical and normal to the tunnel axis. For each tripod station, 6 images are taken, each covering approximately a 12×8 m patch of the object. At the end of the procedure 6 longitudinal strips (with 80% longitudinal overlap and about 20% side-lap) are produced. A total of ca. 110 images per strip (i.e. 660 images in total) were captured.

The fact that the acquisition procedure approximates a panoramic camera and therefore an approximate homography exists between strips [3] has been exploited to connect the 6 strips in a single block. In this way all the points extracted in the overlap area of two longitudinal strips are transferred to the adjacent one to tie the strips together. After the registration of all the images with a bundle adjustment and the successive DSM generation, a detailed orthophoto of the internal surface of the tunnel (first 40 m) was created (Fig. 3c).




a b c

Fig. 3. Underground quarry: a), b): sketch of image acquisition scheme; c) orthophoto of the internal surface of the tunnel (first 40 m).

### 3.2 Example 2 - UAV block

An important and emerging surveying technology is represented by UAV photogrammetry Today it is not really clear how to classify UAV projects: sometimes they are meant as aerial projects, but in some cases images are more similar to those of close-range applications, as they feature wide baselines and convergent viewing angle. The growing interest for these flying platforms is motivated by several new applications. UAV surveys allow acquisition of high resolution images thanks to the possibility of flying at very low altitudes. Manned aircrafts or helicopters rarely reach these altitudes. The possibility of mounting digital cameras on UAVs can be useful for capturing images of objects or scenes that are difficult to be surveyed with standard aerial images. Typical examples are building façades, dams, archaeological excavation areas, vertical rock faces [35], etc.

An example of UAV-based project is shown in Fig. 4. 70 images (4288×2848 pix) were acquired, in collaboration with ETH Zurich, using an UAV-system of Survey Copter over the archaeological Maya area in Copan (Honduras). The block features a structure similar to an aerial survey but has a larger scale. In addition, several problems were found because of the uniform texture of the ground and the occlusions created by the dense vegetation. The automatic orientation technique provided 57,000 image points (roughly) and a final RMSE (reprojection error after the computation of the bundle adjustment) of about 0.62 px [26]. The oriented images were then used for orthophoto production and for the 3D modeling of some archaeological structures of the area.

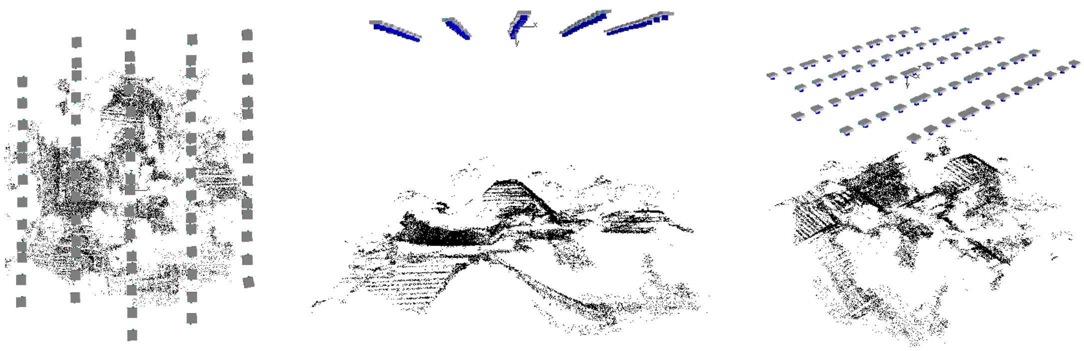


Fig. 4. Automatic orientation results for the UAV block over Copan (Honduras): 3D points and camera poses.

### 3.3 Example 3 - Mobile mapping sequence

During a survey mission, a mobile mapping vehicle acquires (at least) two synchronous image sequences, e.g. acquired with the front cameras. The orientation of such sequences with automatic techniques (disregarding the navigation data) will face unfavorable conditions: (i) the successive image centers will be approximately aligned with the optical axis of the cameras; (ii) large image scale variations as well as large perspective differences might frequently arise between consecutive images; (iii) a large part of the image frame (i.e. the road surface) has characteristics (texture) that can negatively affect the automatic recognition of homologous features.

To study the performance of automated orientation algorithms in such critical cases, a country road image sequence was analyzed (see Fig. 5a) [36]. 91 image pairs covering about 300 m were oriented (Fig. 5b) with a particular implementation that simultaneously considers both sequences of left and right cameras and enforces the constant relative orientation of synchronous pairs. To assess the quality of the block orientation, after the matching/orientation procedure the discrepancies between the exterior orientation parameters evaluated by the navigation system and those obtained from the photogrammetric adjustment were computed. To geo-reference the block the first and last image pairs were fixed to the INS/GNSS positions. Table 1 shows the RMSE and the maximum differences between the exterior orientation parameters and those from the inertial navigation system, with values in accordance with the typical project requirements.

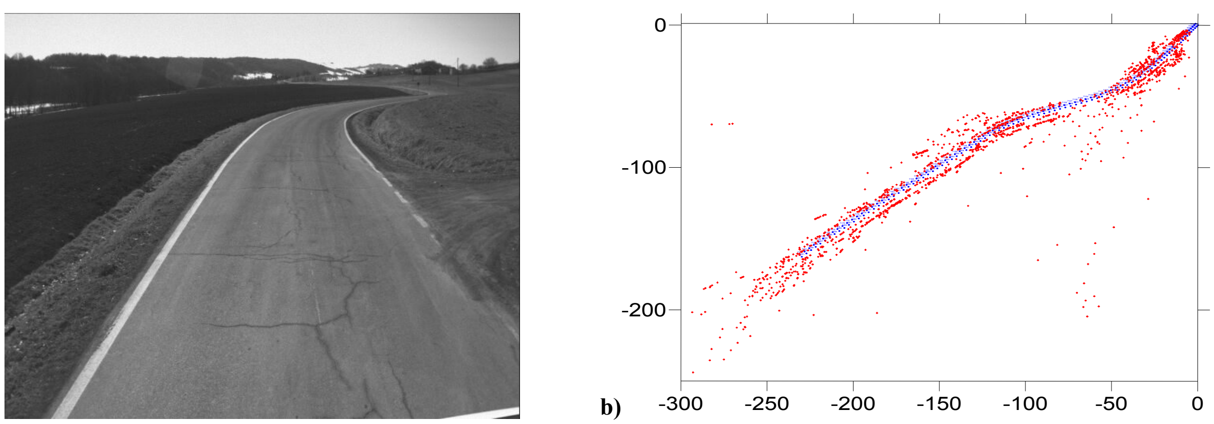



Fig. 5. The first image of the sequence (a); tie points (red) and camera stations (blue) - vehicle drives from NE to SO (b).

| | ΔX (m) | ΔY (m) | ΔZ (m) | Δω (deg) | Δϕ (deg) | Δκ (deg) |
|---|---|---|---|---|---|---|
| **RMSE** | 0.39 | 0.30 | 0.08 | 0.52 | 0.19 | 2.30 |
| **MAX** | 0.70 | 0.51 | 0.20 | 1.36 | 0.51 | 4.73 |

Table 1. RMSE and maximum differences between the exterior orientation parameters from images and navigation data.

### 3.4 Example 4 - Complex image sequences

One of the most challenging situations for an automated orientation procedure is the extraction of reliable image correspondences from scenes featuring low-texture objects and repetitive elements. From a practical point of view, typical examples are architectural objects and building facades, where a restrictive threshold during the comparison of the descriptors with the ratio test removes many good image correspondences. This means that the threshold should be modified according to the texture of the images.

A typical example is shown in Fig. 6 [14]. A sequence of 52 images was acquired with an amateur camera in Piazza del Campidoglio (Rome). The images form a closed sequence and represent a very complex case with several repetitive elements and moving objects (pedestrians). Images were automatically matched delivering the necessary correspondences for the successive bundle adjustment which was performed with Australis. The software removed several wrong data during the iterations of the bundle and the final RMSE was 1.19 px, with approximately 11,700 3D points. Beyond statistical results, it is important to verify the shape of the recovered camera poses. Although metric data for a more advanced accuracy analysis were not available, the visual comparison with an aerial view of the square demonstrates a good consistence.

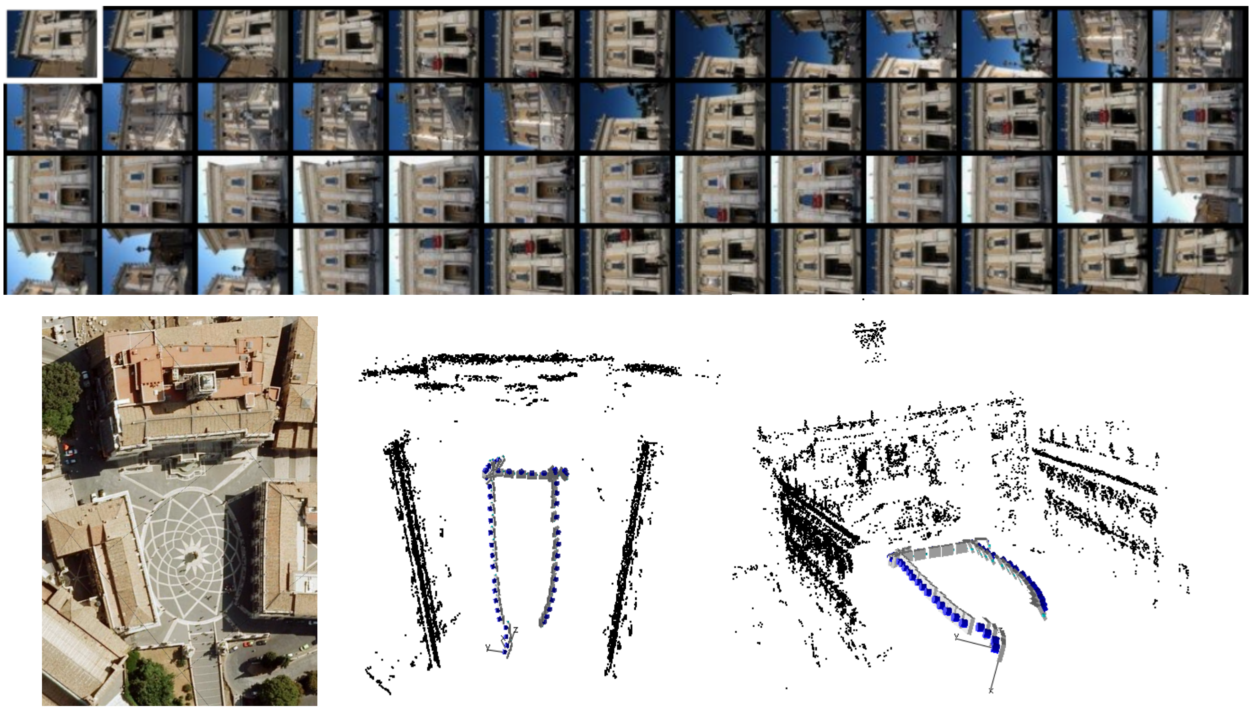


Fig. 6. Automatic orientation results: images and recovered cameras with 3D points.

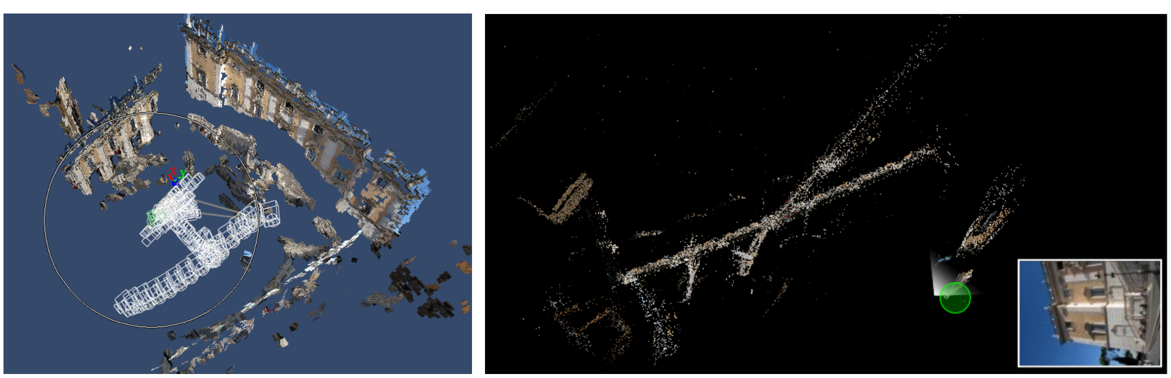


Fig. 7. Wrong and incomplete orientation results achieved with Photofly (left) and Photsynth (right).

Using the same images, the automated camera poses estimation was performed with two well-known web-based CV tools: Autodesk Photofly and Microsoft Photosynth. Both software can determine the orientation parameters in an automated way and can be employed for 3D reconstruction purposes. The processing of the entire dataset was quite fast but the achieved results were not satisfactory, as shown in Fig. 7. Both tools failed to orient the sequence, most probably due to the repeated patterns in the scene and the employed bundle adjustment solution.

# 4. DISCUSSION

As demonstrated throughout the examples reported in the previous section, nowadays automatic orientation of sequences or blocks in close-range photogrammetry is an effective solution for a wide category of end-users and applications. On the other hand, it is worth considering that the most developments in this field do not come from the photogrammetric community. Many R&D efforts have been concentrated to achieve this aim by companies and academies interested to applications like location-based services, augmented reality, visualization or image browsing, which are mainly CV topics. The economic and intellectual potentiality of stakeholders related to these services is much larger than the one involved with traditional applications in close-range photogrammetry, like cultural heritage documentation, image metrology and so on. But disregarding in which community innovations have arisen, many problems are shared between CV and photogrammetry. Consequently, the technical developments in both disciplines are strictly connected, although they feature two largely different approaches which could be identified by the following keywords: photogrammetry is more focused on precision, reliability, completeness; CV focuses more on automated procedures that can be executed "in batch" by a computer without any user interaction

In the following subsections a few issues about problems and perspectives introduced by automated orientation of image sequences in close-range applications will be analyzed in more detail.

## 4.1 Gathering new users

Since some years the use of coded targets has introduced a first simplification in the photogrammetric pipeline [37-38] while recently commercial software packages started to incorporate *targetless automatic orientation* procedures (e.g. PhotoModeler 2010, Eos Inc., Canada). These innovations, coupled with *dense image matching* reconstructions [39] and *automatic texture mapping* [40], allow the creation of photorealistic image-based 3D models comparable to those achievable using 3D laser scanners and related processing software [41]. Furthermore these new improvements are expected to introduce innovative opportunities for photogrammetric applications. First of all, an enlargement of the potential users is expected. The chance to skip the manual orientation of images, which is replaced by a software procedure, could attract more practitioners with a low photogrammetric background. Tasks like design of data acquisition and DSM generation still require the knowledge of the theoretical background and a good surveying experience, but some new operations, that were strictly achievable by expert operators, could be performed by non-experts.

The same occurred in recent years with the introduction of terrestrial laser scanners (TLS) as surveying instruments. This gave a great interest to surveyors as such new instruments were able to acquire dense 3D point clouds in a straightforward and fast manner. In a similar way the increasing of reliable and precise automated procedures in close-range photogrammetry is supposed to attract many users, due to the fact that this technique is faster at data acquisition stage, cheap and able to deliver good 3D results. Nevertheless, the enlargement of the user should not lead to an abuse of the photogrammetric technique: a proper knowledge of the potential and limits of the method is fundamental.

## 4.2 Manual or automatic orientation?

Which criteria have to be considered to choose between a manual or automatic procedure for image orientation? What kind of image sequence allows to derive automatically the orientation parameters? The answer to these questions must entail different aspects. The choice must be made before data acquisition and keeping in mind the goal of the surveying project. Indeed, the data acquisition procedure is influenced by the employed image orientation method, as well as by the local geometry of the environment to be surveyed. Thus images acquired under very large baselines would allow, theoretically, a good precision of the object coordinates but would require interactive measurements of the tie points.

Nevertheless the final precision and accuracy of the estimated orientation parameters is approximately the same with both approaches, as reported in [24], also in the case of targetless procedures. The accuracy of a photogrammetric project mainly depends on the external constraints, whose measurement is generally carried out manually, except when targets are used.

### 4.3 Aspects influencing the orientation quality and result

#### *4.3.1 Datum setup and block stability*

In aerial photogrammetry the definition of the geodetic datum relies on the external constraints (GCPs and GNSS/INS observations). These contribute to the block stability as well. In addition, the geodetic datum is strictly required for mapping purposes. On the other hand in close-range applications, the setup of an arbitrary local datum is sufficient for many applications. The option of free-network adjustment is often used and a linear distance between two points in the photogrammetric model can be adopted to fix the scale ambiguity [42]. In both situations, the scale and datum definition are still interactive phases.

In a photogrammetric block, the stability of the solution depends on two categories of points:

- tie points (TPs), which guarantee the precision of the orientation solution; they should be numerous, well distributed on the object and in the image space and they should feature a high multiplicity;
- ground control points (GCPs), which have the role to fix the datum and to control the block deformation.

But understanding the interaction between the block geometry and the GCP structure is a more complex task in close-range than in aerial photogrammetry. This is also due to the similar precision of surveying and photogrammetric measurements. An accurate preliminary design of the block geometry would be useful to this aim, although this requires the knowledge of the approximate shape of the object [43].

#### *4.3.2 Precision and reliability of image coordinates*

The intrinsic precision of image coordinates mainly depends on the adopted measurement techniques. FBM algorithms used for automated orientation purposes can lead to sub-pixel precision. Residuals in the range 0.5-0.8 px after the bundle adjustment are reported in [24]. The application of area-based matching (ABM) algorithms to refine the FBM results at a later stage is supposed to improve the precision [44]. *Least Squares Matching* [45] (LSM) is usually the preferred technique, even though it is computationally expensive. A further consideration is deserved about multi-photo extensions of LSM (like MPGC [46]). In fact, once image orientation has been achieved, this technique plays a key-role during the object reconstruction stage, because it exploits some geometric constraints to reduce the risk of blunders. In this case, constraints based on orientation parameters that still have to be refined could make more difficult the convergence of LS iterations based on the radiometric content of the images. A pairwise LSM between homologues points extracted with FBM algorithms, followed by a consistency check based on homologous rays, would be ideal. Finally, a validation of tie points is still possible during the bundle adjustment iterations, where data snooping technique is included.

The precision of the final object coordinates depends on the tie points measurement precision, but also on their multiplicity and spatial distribution [47]. Generally, the repeatability of FBM operators, i.e. the capability of tracking a point along many images, is lower than the one of a human operator. This drawback becomes more relevant with convergent images.

A contribute to overcome this problem is given by the acquisition of an image block where the basic image combinations (pairs and triplets) which are established during automatic FBM can be connected, so that the same point can be tracked along more images. Another approach could be the integrating of corner detectors (like FAST [48]) in the FBM strategy [24].

## 5. CONCLUSIONS

The article reviewed the state-of-the-art in automated targetless orientation of large and complex image blocks, with examples from close-range photogrammetric projects. Automation is nowadays feasible in the different steps of the photogrammetric pipeline (camera calibration, image orientation, DSM generation, texture mapping or orthophoto generation). In particular automated image orientation is practicable in most of the projects and under many network configurations, not only for visualization purposes as in the past but also for metric applications. These achievements are opening new possibilities for close-range applications. Automation is also helping photogrammetry to re-acquire importance with respect to TLS systems, employed mostly by non-expert users. But the higher the level of automation, the higher the probability to encounter a misuse of the technique, due to a lack of proper surveying expertise. In any case, automated orientation techniques do not solve for the problem of geo-referencing and datum definition, which are still highly interactive and moreover are still depending on the quality of the acquired image data. This is a very important factor as everything is locked to this and the results are strictly related to the image source, i.e. to the acquisition's procedure and the employed camera as well as the user capability and knowledge. Although some improvements are still feasible, in particular in the FBM domain, R&D in the automated orientation phase should now move to the creation of some protocols and standardizations of the acquisition procedure (like for TLS [49]) instead of continuing to demonstrate algorithmic capabilities in processing thousands of images and delivering sparse point clouds useless for metric and 3D reconstruction applications. Particular attention should be focus on the block geometry, because this should encompass two requirements that can lead very often to controversial solutions: (i) the need of adequate ratios between image baselines and camera-object distances in order to obtain enough precision, especially in depth direction; (ii) the need of large overlaps between the images for image matching techniques (for either orientation or surface reconstruction). The definition of some standard rules for the most common configurations of terrestrial blocks would help the (non-expert) users carry out block design. Another important issue is related to the use of external constraints. In fact, if definition of an external datum is not strictly required in all applications (like it is in geological applications, for instance), a free-network adjustment can provide an optimal solution, in particular for small and stable blocks. On the other hand, the use of GCPs might increase the block stability in case of large and complex blocks. In addition, in many applications GCPs would simplify the definition of the block geometry, especially when the survey is operated by non expert users. A typical example of this is the photogrammetric survey of the internal courtyard of a building with four facades facing on it. The use of enough GCPs would enable the adoption of 4 independent photogrammetric projects, without the need of more images to link the different sides of the blocks. On the other hand, this solution is more time consuming at the acquisition time and requires the use of a theodolite for GCP measurements, which might be an additional cost. But if the users wants to avoid or to limit the use of GCPs, he should have enough experience to setup a proper block geometry. The definition of standard protocols would however help also in this case.

**Related references**


[1] Beardsley, P., Torr, P., Zisserman, A., "3D model acquisition from extended image sequences". Proceedings of ECCV'96, Lecture Notes in Computer Sciences, Vol. 1065, pp. 683-695 (1996).

[2] Pollefeys, M., Koch, R., Van Gool, L., "Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters," International Journal of Computer Vision, 32(1), 7-25 (1999).

[3] Hartley, R.I., Zisserman A., [Multiple View Geometry in Computer Vision],Cambridge University Press, UK, (2000).

[4] Lowe, D., "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, 60(2), 91-110 (2004).

[5] Mikolajczyk, K., Schmid, C., "Scale and Affine Invariant Interest Point Detectors," International Journal of Computer Vision, 60(1), 63-86 (2004).

[6] Bay, H., Ess, A., Tuytelaars, T., Van Gool, L., "SURF: Speeded up robust features," Computer Vision and Image Understanding, 110(3), 346-359 (2008).

[7] Fischler, M.A., Bolles, R.C., "Random Sample Consensus: a paradigm for model fitting with applications to image analysis and automated cartography", Commun. ACM 21, 381-395 (1981).

[8] Rousseeuw, P.J., Leroy, A.M., [Robust Regression and Outlier Detection], John Wiley, New York (USA), 352 pp. (1987).

[9] Torr, P., Murray, D., "The development and comparison of robust methods for estimating the fundamental matrix," International Journal of Computer Vision, 24(3), 271-300 (1997).

[10] Sturm, P., Ramalingam, S., Tardif, J.-P., Gasparini, S., Barreto, J., "Camera Models and Fundamental Concepts Used in Geometric Computer Vision", Foundations and Trends in Computer Graphics and Vision, 6(1-2), 1-183 (2010).

[11] Mason, S.O., "Expert system-based design of close-range photogrammetric networks," ISPRS Journal of Photogrammetry and Remote Sensing, 50(5), 13-24 (1995).
[12] Kraus, K., [Photogrammetry], Vol. 2. Dümmler, Bonn, Germany, 466 pp. (1997).
[13] Waldhaeusl, P., Ogleby, C., "3x3-rules for simple photogrammetric documentation of architecture," Int. Archives of Photogrammetry and Remote Sensing, Melbourne, Australia, Volume XXX, Part 5, 426-429, (1994).
[14] Barazzetti, L., Remondino, F. and Scaioni M., "Accurate and automated orientation of complex image sequences," Int. Archives of Photogrammetry Remote Sensing and Spatial Information Sciences, 38-5/W16, on CD-ROM (2011).
[15] Pierrot-Deseilligny, M., Clery, I., "APERO, an open source bundle adjustment software for automatic calibration and orientation of set of images," Int. Archives of Photogrammetry Remote Sensing and Spatial Information Sciences, 38-5/W16, on CD-ROM (2011).
[16] Roncella, R., Re, C., Forlani, G., "Performance evaluation of a structure and motion strategy in architecture and cultural heritage," Int. Archives of Photogrammetry Remote Sensing and Spatial Information Sciences, 38-5/W16, on CD-ROM (2011).
[17] Snavely, N., Seitz, S.M., Szeliski, R., "Modelling the world from internet photo collections," International Journal of Computer Vision, 80(2), 189-210 (2008).
[18] Frahm, J., Georgel, P., Gallup, D., Johnson, T., Raguram, R., Wu, C., Jen, Y., Dunn, E., Clipp, B., Lazebnik, S., Pollefeys, M., "Building Rome on a Cloudless Day," Proc. of ECCV (2010).
[19] Gherardi, R., Toldo, R., Garro, V., Fusiello A., "Automatic camera orientation and structure recovery with Samantha," Int. Archives of Photogrammetry Remote Sensing and Spatial Information Sciences, 38-5/W16, on CD-ROM (2011).
[20] Büyüksalih, G. and Li, Z.,. "Practical experiences with automatic aerial triangulation using different software packages", Photogrammetric Record, 18(102), 131-155 (2005).
[21] Läbe, T., Förstner, W., "Automatic relative orientation of images," Proc. of the 5th Turkish-German Joint Geodetic Days, 6 pp., on CD-ROM (2006).
[22] Remondino, F. and Ressl, C., "Overview and experience in automated markerless image orientation," Int. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 36(3), 248-254 (2006).
[23] Roncella, R., "Sviluppo e applicazioni di tecniche di automazione in fotogrammetria dei vicini", Ricerche di Geomatica, Mestre, pp. 119-128 (2007).
[24] Barazzetti, L., Remondino, F., Scaioni M., "Orientation and 3D modelling from markerless terrestrial images: combining accuracy with automation," The Photogrammetric Record, 25(132), 356-381 (2010).
[25] Förstner, W., "New orientation procedures," Int. Archives of Photogrammetry Remote Sensing and Spatial Information Sciences, 33(3), 297-304 (2000).
[26] Barazzetti, L., Remondino F., Scaioni M., Brumana R., "Fully Automatic UAV Image-Based Sensor Orientation," Int. Archives of Photogrammetry Remote Sensing and Spatial Information Sciences, 38(1), 6 pp., on CDROM (2010).
[27] Armangué, X. and Salvi, J., "Overall view regarding fundamental matrix estimation," Image Vision Computing, 21, 205-220 (2003).
[28] Nister, D., "An efficient solution to the five-point relative pose problem," IEEE Trans PAMI., 26(6), 756-770 (2004).
[29] Pozzoli, A., Mussio, L., Scaioni, M., "A solution for the general case of three-image orientation," Int. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 55, Part B3, 992-997 (2004).
[30] Torr, P., "Bayesian model estimation and selection for epipolar geometry and generic manifold fitting," International Journal of Computer Vision, 50(1), 35-61 (2002).
[31] Roncella, R., Forlani, G., "Modelli digitali di pareti in alta quota: metodologie e casi di studio", Proc. SIFET, Vol. 1, p. 77-86 (2009).
[32] Ferrero, A.M., Forlani, G., Roncella, R., Voyat, I.H., "Advanced geostructural survey methods applied to rock mass characterization," Rock Mechanics and Rock Engineering, 1, 1-35 (2008).
[33] Zucca, F., Remondino, F., Zizioli, D., Meisina, C., "Slopes survey and analysis using photogrammetrically derived Digital Surface Models," Proc. of RSPSoc 2009 Annual Conf., pp. 494-500 (2009).
[34] Roncella, R., Forlani, G., Remondino, F., "Photogrammetry for geological applications: automatic retrieval of discontinuity in rock slopes," Proc. of Videometrics VIII, SPIE IS&T Electronic Imaging, 17-27 (2005).
[35] Eisenbeiss, H., "The autonomous mini helicopter: a powerful platform for mobile mapping," Int. Archives of Photogrammetry Remote Sensing and Spatial Information Sciences, 37(B1), 977-983 (2008).

[36] Forlani, G., Roncella, R., Remondino, F., "Structure and motion reconstruction of short mobile mapping image sequences," Proc. of the 7th Conf. on Optical 3D Measurement Techniques, Vol I, 265-274 (2005).
[37] Ganci, G., Hanley, H., "Automation in videogrammetry", Int. Archives of Photogrammetry Remote Sensing, 32(5), 53-58 (1998).
[38] Cronk, S., Fraser, C., Hanley, H., "Automatic metric calibration of colour digital cameras," Photogrammetric Record, 21(116), 355-372 (2006).
[39] Vu., H. H., Keriven, R., Labatut, P., Pons, J.-P., "Towards high-resolution large-scale multi-view stereo," Proc. IEEE Conf. CVPR'09, 1430-1437 (2009).
[40] Abdelhafiz, A., [Integrating Digital Photogrammetry and Terrestrial Laser Scanning], PhD thesis, Institute for Geodesy and Photogrammetry, Technical University Braunschweig, Germany. ISBN 3-926146-18-4. Online on Deutsche Geodätische Kommission (DGK), München 2009, ISBN 978-3-7696-5043-3 (2009).
[41] Cignoni, P., Scopigno, R., "Sampled 3D models for CH applications: an enabling medium or a technological exercise?," ACM Journal. on Computers and Cultural heritage, 1(1) (2008).
[42] Papo, H.B., Perelmuter, A., "Free Net Analysis in Close-Range Photogrammetry," Photogrammetric Engineering and Remote Sensing, 48(4), 571-576 (1982).
[43] Fraser, C.S., "Network Design Considerations for Non-Topographic Photogrammetry," Photogrammetric Engineering and Remote Sensing, 50(8), 1115-1126 (1984).
[44] Remondino, F., "Detectors and descriptors for photogrammetric applications," Int. Archives of Photogrammetry Remote Sensing and Spatial Information Sciences, 36(3): 49-54 (2006).
[45] Grün, A., "Adaptative least squares correlations: a powerful matching techniques," South African Journal of Photogrammetry Remote Sensing and Cartography, 14(3), 175-187 (1985).
[46] Grün, A., Baltsavias, E.P., "Geometrically constrained multiphoto matching," Photogrammetric Engineering and Remote Sensing, 54(5): 633-641 (1988).
[47] Fraser, C., Network design. In [Close-Range Photogrammetry and Machine Vision], K.B. Atkinson (Ed.), Whittles Publishing, Caithness, Scozia (UK), 256-282 (1996).
[48] Rosten, E., Drummond, T., "Machine learning for high-speed corner detection," Proc. EECV2006, 430-443 (2006).
[49] Beraldin, J. A., Picard, M., Bandiera, A., Valzano, V., Negro, F., "Best Practices for the 3D Documentation of the Grotta dei Cervi of Porto Badisco, Italy", Three-Dimensional Imaging, Interaction, and Measurement, SPIE Vol. 7864, pp. 78640J-78640J-15 (2011).