MULTI-STEP AND MULTI-PHOTO MATCHING FOR ACCURATE 3D RECONSTRUCTION

M. Previtali, L. Barazzetti, M. Scaioni

Politecnico di Milano, Department of Building Environment Science and Technology
Via M. d'Oggiono, 18/a, 23900 Lecco, Italy
email: mattia.previtali@mail.polimi.it, (luigi.barazzetti, marco.scaioni)@polimi.it,
web: http://www.icet-rilevamento.lecco.polimi.it/

Commission III – WG/1

KEY WORDS: Accuracy, Feature Extraction, Matching, Orientation, Robust Estimation, Surface Measurement

ABSTRACT:

The paper presents an automated procedure for surface reconstruction from digital images. This method was developed for close-range photogrammetric applications, with a particular attention to terrestrial free-form objects that can be modelled with point clouds extracted from images. Therefore, the paper is not directly concerned with architectural elements, where objects feature breaklines and discontinuities that are preferably modelled with manual measurements. The implemented algorithm (MGCM+) integrates two image matching techniques developed in Photogrammetry and Computer Vision in order to obtain metric results in an automated way. Different strategies were exploited to successfully combine both strategies, along with several new improvements. Starting from a set of images and their orientation parameters a preliminary seed model is extracted by using a patch-based algorithm (PMVS). Then, a multi-photo refinement via LSM (MGCM) improves the precision of results and provides a statistical evaluation through a variance-covariance matrix.

1. INTRODUCTION

Nowadays there is an intense research activity aimed at developing new strategies for object reconstruction from images. Several approaches were developed in both Photogrammetry and Computer Vision (CV), keeping in mind different requisites such as accuracy, completeness, automation, reliability, and so on. Furthermore, the typology of the analyzed objects is diverse as well. For instance, in Furukawa et al. (2010), Frahm et al. (2010) and Strecha et al. (2010) some methods for "large scale city modelling" are illustrated, showing impressive results. Several thousands of images are automatically processed in order to obtain 3D models of vast areas containing complex buildings.

Photogrammetrists could state that these kinds of reconstructions are not good for mapping purposes, as results are often incomplete and are not accompanied by statistical analyses. Typical problems can be (i) modelling of architectural objects, where breaklines should be matched in order to obtain sharp edges, (ii) use of uncalibrated cameras and images downloaded from the Internet, without any consideration about network design, (iii) lack of a geodetic network for stability control in the case of large blocks, with GCPs used as pseudo-observations in bundle adjustment.

Accuracy during image orientation becomes a point of primary importance in Photogrammetry. Indeed, a photogrammetric bundle adjustment is supposed to ensure the metric quality of the final result. This is partially in contrast with a CV bundle adjustment, as best summed up in Snavely et al. (2008), where the functioning of Bundler is described: "most SfM methods operate by minimizing reprojection error and do not provide guarantees on metric accuracy".

This different point of view is motivated by the use of the final 3D model. In a few words, the approaches are diverse because purposes are diverse.

However, in close-range photogrammetry a growing number of CV methods is receiving great attention. For the image

orientation phase, it is now available on the market the new PhotoModeler 2011 (EOS, Canada – www.photomodeler.com), that is the first (photogrammetric) commercial package capable of orienting target-less images in a fully automated way. The mathematical model used during bundle adjustment is a typical photogrammetric approach, but the operator for image matching (SIFT in this case) and the strategies for outlier rejection (based on the *fundamental* or *essential matrices*) come from CV. In addition, there are also other solutions for automatic orientation in close-range, where different techniques (e.g. *Least Squares Matching* - LSM) are integrated to improve precision and reliability (Barazzetti et al., 2010; Pierrot Deseilligny and Clery, 2011; Roncella et al., 2011).

This means that the combined use of techniques developed in both disciplines allows one to obtain accurate results in a fully automated way. This is now a reality for image orientation only, while with the work presented in this paper we would like to extend the concept also towards 3D modelling (for some specific categories of objects).

In the field of close-range photogrammetry some (multi-image) commercial software for surface reconstruction are available today. Most of them are derived from Aerial Photogrammetry (e.g. CLORAMA - Remondino et al., 2008; LPS eATE - www.erdas.com). In their original implementations, these are able to extract a digital surface model (DSM), that is a 2.5D representation of the ground. On the other hand such 2.5D models are adequate for airborne mapping, but they feature evident limits in close-range surveys, because they cannot handle scenes at 360°. In addition, problems arise for DSM cells having multiple depth values. On the other hand, the accuracy obtainable with these methods is noteworthy, especially thanks to the use of sub-pixel *area-based matching* (ABM) procedures.

As previously mentioned, we would like to present a methodology for surface measurement in the case of 3D objects. The aim is to obtain models useful for photogrammetric surveys. Architectural objects with sharp breaklines (e.g.

facades) are not considered here because their detailed modelling can best be accomplished with interactive measurements, through the identification of basic geometric shapes. We focus instead on 'free-form' objects, i.e. objects that can be modelled with meshes generated from point clouds.

The proposed matching procedure is divided into two steps. First of all a seed model is created with a *patch-based image matching* technique and then ABM operators densify and refine the point cloud. Multiple images are simultaneously used to detect and remove outliers with the analysis of the light ray intersections in 3D space. As the core for global processing is the combined use of *Multi-photo Geometrically Constrained Matching* (MGCM) and other methods (that essentially allow 3D processing), we called the procedure MGCM+.

2. 3D RECONSTRUCTION PIPELINE WITH MGCM+

In this section a complete pipeline for the geometrical reconstruction of large-scale objects and scenes, using high resolution images in a multi-view stereo framework, is described. This method incorporates the high-end matching algorithms developed in Photogrammetry and CV.

As can be seen in the flowchart in Figure 1, a block of suitable images (in terms of network geometry and image resolution) is needed. All images must be captured by using calibrated cameras in order to improve the precision of the final 3D measurements. It is out of the scope of this paper to illustrate the geometric characteristics that the image block should have (overlap, external or internal constraints, relative angles between images and the like). To obtain accurate and reliable surface measurements, each portion of the whole object must be covered by at least 3-4 images to exploit the potential of multi-photo matching. In addition, the length of any baseline has to be selected according to a compromise between the precision in the depth direction (large baselines are better) and the limitation of image deformations required by ABM (short baselines and small view angles).

Image orientation is the second prerequisite of MGCM+. It can be performed by manual or automatic procedures, but the latter have the advantage to generate a denser point cloud of tie points. This can be used to initialise MGCM+ with an approximate surface. On the contrary, if this initial model is not sufficiently dense, an alternative solution is applied to derive the seed model. This is mainly based on the algorithm proposed by Furukawa and Ponce (2010), as illustrated in subsection 2.1. The advantage of this method is its independence from the reference frame adopted and the capability of working without any initial rough model of the object.

The MGCM algorithm combines (i) LSM based on intensity observations with (ii) collinearity conditions used as geometrical constraints for the determination of all object point coordinates. The introduction of the collinearity constraints and the opportunity to simultaneously match multiple scenes increase the matching reliability.

MGCM is a matching technique presented by Grün (1985), with new improvements in Grün and Baltsavias (1988) and Baltsavias (1991). One could say that this technique is old. However, it is still the most precise method for image coordinate measurement and it is therefore appropriate for metric purposes. Here the theoretical background of MGCM is not explained in detail, but the main aspects of this technique are outlined to highlight its advantages with respect to other ABM methods. In addition, some limitations of the original formulation are pointed out.

Let us consider a block of images depicting an object. One of them is selected as 'master image' according to a specific criterion (see subsect. 2.2). A set of points (x_{Pk}, y_{Pk}) found by means of an *interest operator* (or nodes of a regular grid) are used to extract a set of 'templates' i.e. square patches with a side of a few pixels. Each of them is reprojected on the other images of the block by exploiting a rough DSM of the object. A squared window ('slave') is extracted around each reprojected point for each generic image i, obtaining a total number of n possible candidates.
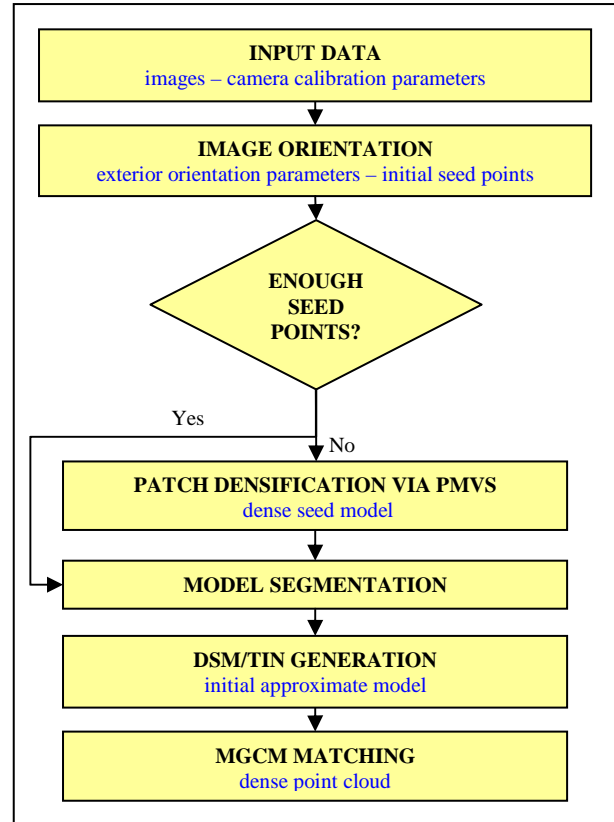

Figure 1. Workflow of MGCM+.

The geometric deformation between the 'template' and each 'slave' is modelled using an affine transformation, which locally approximates quite well perspective deformations. Then the 'template' is compared with all corresponding 'slaves'. The relationship describing the intensity values of each pixel in the 'template' is given by the discrete function $f(x,y)$, and the n 'slaves' are represented by functions $g_1(x,y)$, ..., $g_n(x,y)$. An intensity observation equation for each pixel of the 'template' and the corresponding pixel on the 'slave' i is written as follows:

$$f(x,y) - e_i(x,y) = g_i^0(x,y) + g_{xi}\,da_{10i} + \\ + g_{xi}\,x_o\,da_{11i} + g_{xi}\,y_o\,da_{12i} + g_{yi}\,da_{20i} + \\ + g_{yi}\,x_o\,da_{21i} + g_{yi}\,y_o\,da_{22i} \tag{1}$$

where the unknown quantities are corrections to the parameters of the affine transformation da_{jki}. The coefficient $g_i^0(x,y)$ is the observed value in the approximate position of the 'slave', while g_{xi} and g_{yi} are the partial derivatives of the function $g(x,y)$. Numerically, the derivatives correspond to row and column

gradients. The function $e_i(x,y)$ gives the residual error with respect to the affine model.

In addition, it is also possible to take into account some radiometric transformations (in many cases using a linear formulation) between 'template' and 'slave'. However, in our case we usually disregard this radiometric compensation to limit the parameters and we prefer to operate with a preliminary radiometric equalization at local level.

The MGCM combines the intensity observation equations (1) with the collinearity condition. In fact, for a pinhole (central perspective) image k the constraint between the generic object point $(X_p=[X_p\ Y_p\ Z_p]^T)$ and its corresponding 2D point (x_{pk},y_{pk}) on the image k is given by the well-known collinearity equations:

$$x_{pk} = -c_k \frac{r_{1k}^T\left(X_p - X_{ok}\right)}{r_{3k}^T\left(X_p - X_{ok}\right)} \triangleq -F_k^x \qquad (2)$$

$$y_{pk} = -c_k \frac{r_{2k}^T\left(X_p - X_{ok}\right)}{r_{3k}^T\left(X_p - X_{ok}\right)} \triangleq -F_k^y$$

where c_k is the principal distance, X_{0k} is the vector expressing the perspective centre coordinates, $\mathbf{R}_k=[r_{1k}\ r_{2k}\ r_{3k}]^T$ is the rotation matrix. Image coordinates (x_{pk},y_{pk}) are computed with respect to the principal point.

If both interior and exterior orientation (EO) parameters of each station are known, eq.s 2 can be rewritten as follows:

$$\Delta x_{pk} + F_k^x + x_{pk}^0 = 0 \qquad (3)$$

$$\Delta y_{pk} + F_k^y + y_{pk}^0 = 0$$

The unknown parameters in eq.s 3 are shifts $(\Delta x_k,\ \Delta y_k)$ and object point coordinates (X_p). After their linearization, Eq.s 3 become:

$$\Delta x_{pk} + \frac{\partial F_k^x}{\partial X}dX + \frac{\partial F_k^x}{\partial Y}dY + \frac{\partial F_k^x}{\partial Z}dZ + F_k^{x(0)} + x_{pk}^0 = 0 \qquad (4)$$

$$\Delta y_{pk} + \frac{\partial F_k^y}{\partial X}dX + \frac{\partial F_k^y}{\partial Y}dY + \frac{\partial F_k^y}{\partial Z}dZ + F_k^{y(0)} + y_{pk}^0 = 0$$

Shifts allow one to link both sets of eq.s 1 and 4, because $\Delta x_{pk}=da_{10}$ and $\Delta y_{pk}=da_{20}$ for the same set of images and point P. Therefore, the resulting joint system can be solved using conventional Least Squares solution schemes (see Baltsavias, 1991).

MGCM presents some important advantages with respect to other traditional automatic matching techniques used in Photogrammetry and CV. Compared to the normal LSM, where it is possible to match simultaneously only a couple of images, the MGCM obtains highly redundant results thanks to the collinearity constraint that permits combined multi-image matching. This reduces multiple solutions in case of repetitive textures and helps overcome possible occlusions thanks to the chance to view the object from multiple stations. In addition, 3D object point coordinates are directly computed together with their theoretical accuracies. Recently, an extensions of the standard cross-correlation technique have been developed, obtaining the so called *Geometrically Constrained Cross-Correlation* (GC[3]) (Zhang and Grün, 2006). This technique uses the collinearity condition in a way similar to MGCM. However, the perspective changes in close-range data can cause some

troubles to the correlation strategy, while the affine transformation between the template and each slave increases the potential of MGCM algorithm in case of convergent images. The application of our method requires a decomposition of the object into 2.5D regions. For each of them a DSM or a TIN (triangulated irregular network) structure are interpolated starting from a set of seed points (subsect. 2.1). In particular, the choice of the subset of images for the measurement of each portion of the object is a crucial task. Inside this problem, a key aspect is which image could better serve as 'template' (subsect. 2.2). All portions reconstructed in the local reference systems are finally joint together to derive a unique 3D model. It is important to mention that the recombination of the point clouds is rigorous as the rigid body transformations employed are known exactly. With this approach, even though the core matching strategy is still 2.5D, any 3D shape could be potentially reconstructed.

2.1 Seed model generation

As the measurement of the object surface with the MGCM algorithm needs an initial approximation, an intermediate step was added to obtain a preliminary seed model. In the case the EO parameters have been computed by using an automatic procedure, one could use all tie points matched with feature-based matching (FBM) operators. However, in some cases their number is not sufficient (e.g. with texture-less objects) or their distribution in the images can be really variable, leaving some empty areas.

The importance of a good seed model is remarkable not only for the geometric quality of the final product, but also in terms of CPU time as it can limit the search along the 3D light ray, reducing the number of trials during the translation of the correlation window. Lastly, tie-point coordinates are usually incorporated into a photogrammetric bundle adjustment. If the number of point correspondences used for image orientation becomes significant, there is a consequent increment of the computational cost. According to the authors' experience some (few) close-range photogrammetric packages can process several thousand of image points, but when there are more than 100,000 image coordinates, the computation of a rigorous bundle solution based on collinearity equations could become quite difficult, especially with standard PCs.

For this reason a limited number of tie points with a good distribution and geometric multiplicity is still the best compromise during the orientation phase. The generation of a seed model is carried out in a new matching phase, where EO parameters are kept fixed in order to exploit the geometric constraint due to collinearity.

As our method was developed for 3D objects, we exploit the *patch-based matching* (PMVS) approach proposed by Furukawa and Ponce (2010). Their procedure was incorporated into our matching pipeline in order to generate a low resolution initial model. This choice is motivated by the robustness of the method that combines multiple images during the dense matching step: if at least three images are processed simultaneously, blunders and spurious points can be removed by analysing the local data redundancy. In addition, the method is able to work with 3D objects and does not require any manual measurement.

With these considerations in mind, the use of an intermediate procedure in the reconstruction pipeline could be seen as a drawback and a lack of originality. On the other hand, it is quite difficult to find (or develop) an open source implementation much better than Furukawa and Ponce's code. In addition, as mentioned in the introduction, this work is partially based on

techniques developed by different authors. Nowadays, some of these have reached a significant level of maturity, while others still need improvements to become useful in practise. Our contribution tries to combine different procedures and this is not a trivial task. Moreover the processing pipeline should be able to exploit only the best properties of each technique. For instance, there are two fundamental drawbacks found with the implementation available:

- it is difficult to manage several high resolution images used at their original size; and
- the input data are expressed with the P-matrix camera model (Hartley and Zisserman, 2003), while a photogrammetric bundle adjustment provides EO parameters.

The solution for the second point can be found in Barazzetti (2010), where an explicit relationship between both orientation datasets is illustrated for the case of distortion-free images. The first problem can be overcome by using compressed images with an opportune modification of the projection matrices. This solution is viable because we are interested in the creation of a seed model. In several scientific work images are subsampled during dense matching without introducing a coarse-to-fine approach that considers the original data. This is an approximation that degrades the quality of the final result, and it assumes an increasing importance due to the technological improvement of digital cameras, with geometric resolutions superior to 14 Mpix even for low-cost compact sensors. Our solution to this problem is described in the next section, where images are always used at their original size without any loss of precision.

2.2 Approximate geometrical model handling

An important limit in the MGCM algorithm is the need of an approximate position of the object points, along with a preliminary location of the homologous point positions in the slave images. To overcome this problem a seed model of the object is derived by using the method described in the previous section. In the current implementation of MGCM+ the object surface is approximated by using a DSM oriented with respect to a reference plane. In the case of complex 3D objects which do not feature a 2.5D geometry (likewise the topographic surface in mapping projects), the whole patch-based point cloud is segmented in approximated 2.5D regions. As things stand now, the segmentation of the point cloud is performed in a manual way (this solution is fast and simple for many objects). Each model is processed in a separate way and the final point clouds are then connected together. This is also good for parallel computing. In any case, the use of a more flexible 3D data structure like a TIN will be added soon, because it is the best solution for complex objects.

A set of points is then defined in the 2D regular grid of each DSM. The cell size can be set by the user according to the resolution of the images. The approximate elevation of each point with respect to the reference plane is selected by using interpolation techniques. In a second stage, this elevation will be estimated by the MGCM L.S. solution. A back-projection of each grid point on all the available views is then carried out to select the 'master' and 'slaves' images and it also provides the set of initial positions of homologous points.

However, although the DSM used can be a rough approximation of the real surface, the homologous points defined by the back-projection principle can be very far from their true positions. For this reason, additional intermediate

points are set up on the projective ray connecting the DSM cell and the image point in the 'master' image. The number of additional points along the projective ray and their inter-distances are both parameters estimable on the basis of the approximated surface model quality. For each point defined along the projective ray the 'slave' image patches are derived using the collinearity principle. According to this approach, a set of approximated candidate positions for the L.S. solution of the system is found. At this stage, both sets of eq.s of kind (1) and (4) are set up in order to compute the corrections for image and object point positions. The partial variance factors σ_{0i}^2 for each individual patch are also estimated. The process is iterated until the corrections are negligible. The last problem concerns the choice, among all candidate solutions computed along the projective ray, of the correct matching. In particular, it is considered as correct match the one minimizing the mean variance factor.

2.3 Selection of images and LSM approximate parameters

As remarked above, one of the weak points in the original MGCM formulation is the selection of 'master' and 'slave' images. Generally the problem is solved in this way: given the set of images to be processed, an image (usually the central one) is manually picked up as 'master'. Consequently, all the other images will serve as 'slaves'. The manual choice of a fixed 'master', obviously, is not the best criterion. This is mainly due to a couple of reasons: (i) if an object, approximated with a 2.5D model, is not entirely visible in a single image, multiple processing with different 'master' images is needed; (ii) in terrestrial surveys there are some lateral views of an object, capturing areas occluded in the central image. They give an important contribution to the final reconstruction of the object. This contribution would be completely neglected by using a 'fixed master' approach.

Nevertheless, the alternate use of all photos as 'template' is not a good solution because of the huge CPU time needed to complete global processing. Finally, in close-range applications the perspective deformations between different images can be so large that the affine model between 'master' and 'slave' images could become inadequate.

For all these reasons an optimization in the image selection phase is needed. To start with, a selection based on the information derived from the approximated model is accomplished. For a specified point in the DSM, all images in which the point is visible are considered with a simple back-projection. The selection of the 'master' is then carried out inside this set. The surface normal direction in correspondence to the considered object point is computed, then this direction is compared to all photo normals where the point is visible. The image whose normal is closer to the surface normal direction is chosen as 'master'. With this strategy we can easily handle also 2.5D objects that are not completely visible in a single image, without requiring the intermediate interaction of the user.

An optimization is also mandatory for what concerns the selection of 'slaves' images. In fact, in many cases the perspective deformations can become a problem for LSM. For this reason we limit the number of possible 'slaves' only to those where LSM can provide good results. Also in this case the choice is operated using the approximated surface model. For each point in the original model we consider the shape of the DSM cell containing the same points in different images. If a large geometrical deformation occurs, the shape of the DSM cell in the images presents significant changes. Therefore we back-project the DSM cell containing the object points in all images and we compare the cell changes between the defined

'master' image and the other 'slave' candidates. In particular we consider two parameters: 'cell area' and 'cell shape'. If the area of the back-projected DSM cell in a 'slave' is less than half of the cell area in the 'master', the variation between images, owing to both perspective changes or scale variations, is considered too large and the point is not processed. However, in some cases even if the area does not vary too much from two images a significant perspective variation could occur. In this case the "shape" of the cell changes in a significant way. To recognize this situation we consider the inclination of the back-projected DSM cell on the images. If angular variations between the 'template' and a candidate 'slave' are superior to 40% the image is rejected.

Finally, it is important to find a set of approximate parameters for the affine transformation between 'template' and 'slaves'. After selecting the 'slave' images as described above, the DSM cell is known in different images. This information can be used to compute initial values for rotation, affinity and scale parameters for the LSM, simply using an affine transformation between the back-projected DSM cell in the 'master' and 'slave' images. As shown in Balsavias, (1991) the significance of the shaping parameters in the affine transformation can be evaluated with their correlations. In fact, high correlations among the parameters of the affine model and the others might indicate their non-determinability. In our case we are particularly interested in evaluating the significance of shears and scales as approximate values. The correlations between similar shaping parameters (scales - shears) and the correlations between shape parameters in the same direction have an essential importance. At this stage two approaches can be used to evaluate these correlations: a deterministic approach and a statistical one.

In the first case all parameters can be considered as highly correlated if their correlation coefficients exceed a fixed threshold. This means that one of the two correlated parameters can be assumed as not significant and the one with the larger variance should be excluded.

In many cases the use of a fixed threshold for the definition of high correlations can be a real challenge and could lead to a poor solution. Here, a statistical approach becomes more suitable. In particular, it is possible to assume that the parameters have a multivariate normal density distribution. Under this hypothesis, and after fixing a significance level for the test, the correlation of the shaping parameters can be verified using *Hotelling's test* (Baltsavias, 1991). If the test fails, there are correlations, otherwise all parameters can be considered as statistically uncorrelated. A further investigation should be carried out to determine which coefficients are effectively correlated. This check can be done with a test that imposes the null hypothesis $\rho = 0$.

As can be noticed, in the statistical approach no threshold needs to be set at the beginning. It is necessary to fix only the significance level for the test. Therefore, this is the default procedure in our method.

3. EXPERIMENTS

Shown in figures 2a-b-c are three close-range objects modelled from multiple convergent images (12 Mpix) by using the proposed method. The first example (a) is a 3D object that was divided into three portions to fit the 2.5D requisite. The approximate DSM (its first region is shown with a colour-map representation) was the starting point to extract 1.3 million points roughly. No blunders were found at the end of matching phase and a final mesh was created after the combination of all point clouds. As previously mentioned, the alignment of partial reconstructions does not introduce new errors, as each rigid-body transformation is known.

In this case a rigorous accuracy analysis was impossible, as a reference dataset was not available. However, this method allows a statistical evaluation with the covariance matrix, offering the standard deviations of all 3D points.

The second dataset (b) can be clearly modelled without any partitioning. The point cloud obtained from 5 convergent images is made up of 0.5 million points. Blunders were correctly removed by MGCM+. For this dataset we carried out a comparison between another point cloud generated using Leica Photogrammetry Suite (LPS) - eATE. This software was developed for aerial mapping purposes and can be considered a well-assessed tool for object reconstruction, based on *semi-global matching* (Hirschmüller, 2008). Both meshes were aligned using the ICP registration algorithm implemented in Geomagic Studio. The range of the error bar is ±13 mm, while the standard deviation of discrepancies is ±0.7 mm. The object is 0.9 m wide.

The example in Figure 2c comprehends 7 images capturing a bas-relief 1 m wide. Also in this case the object can be easily modelled using a single partitioning, offering the opportunity for a new comparison with LPS-eATE. This gave a discrepancy of about 0.4 mm, while the error bar ranges from [-13; +13] mm.

Some other datasets for multi-view stereo evaluation are available on the Internet (provided by Strecha). They are quite challenging because of a repetitive texture and several breaklines. These are typical objects for which usually a manual reconstruction gives better results. The façade of the building in Figure 2d was modelled using 25 images, that were oriented in order to obtain photogrammetric EO parameters (although the camera-matrices are available). The output of the multi-image matching phase with an incorporated block sub-division are 2.7 million points. For the second dataset (Figure 2e), 11 images were employed to obtain 1.2 million points that were interpolated to create a mesh. A scale factor was then fixed to remove this ambiguity, and the model was aligned with the laser mesh with ICP obtaining a discrepancy (in terms of standard deviation) of about ±12 mm. In any case, during this comparison we included all areas that were not visible in the images, where there are evident gaps in the photogrammetric model. Here, the distances between the model (that are not errors but only empty areas) are superior to 10 cm, and caused a global worsening. The analysis was repeated only for small portions of the model in order to avoid this problem, estimating a std.dev. of about ±5.6 mm. The obtained value is comparable to the ground sample distance (GSD) of the images and the sampling step during MGCM+.

Figure 2f shows a 360° reconstruction from a set of 32 images around a statue, confirming the suitability of the method for 3D objects. The small objects (g) is instead made of marble. Although it is well-known that this material is prone to produce noisy results, the visual reconstruction seems good.

Figure 2h shows a geotechnical case, where a rock face was surveyed at different epochs in order to monitor its stability and discover potential risks (e.g. rockfalls). To accomplish this task the metric content of the model is essential. In addition, as data must be compared to obtain a multi-temporal analysis, images must be registered to the same reference system with some GCPs incorporated into the bundle adjustment. The comparison with a laser model, after removing disturbing elements such as vegetation, revealed a discrepancy of about ±5 mm, i.e. the nominal precision of the laser scanner used.
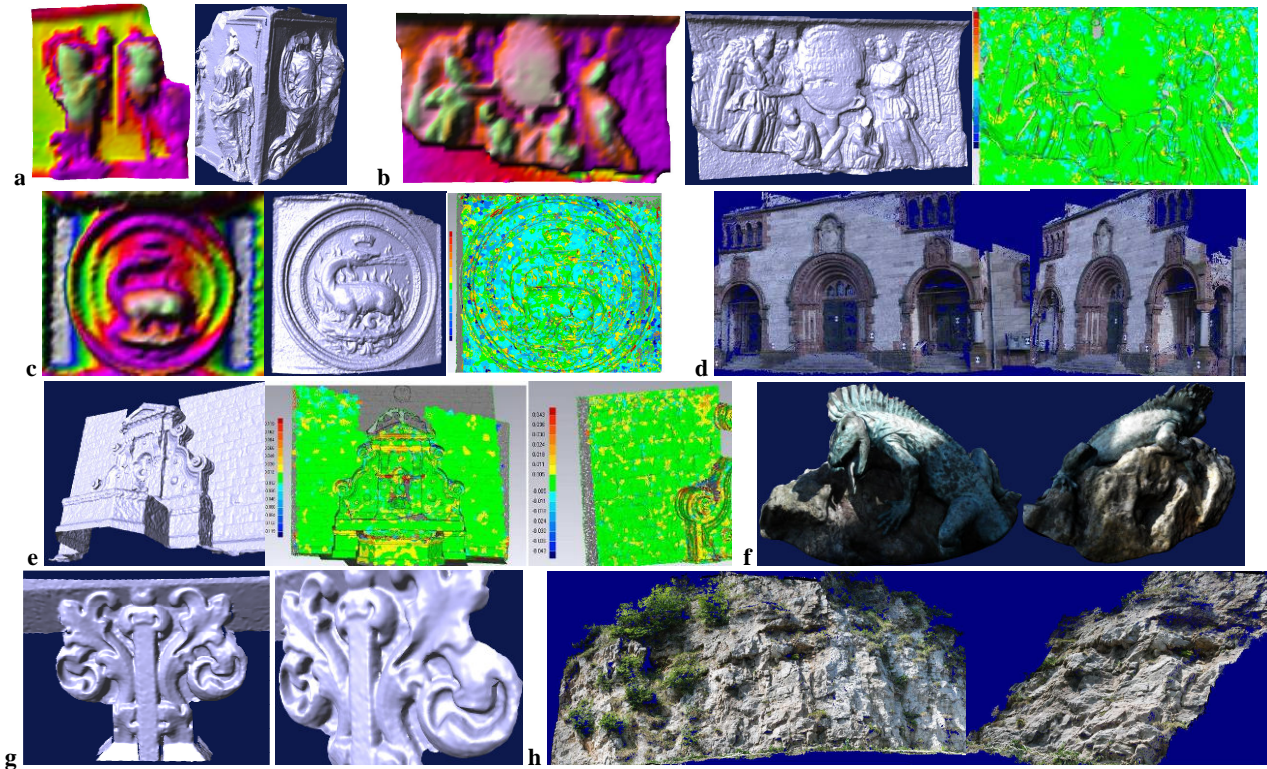
Figure 2. Some results obtained with MPGC+ for the reconstruction of 2.5D and 3D objects.

4. CONCLUSIONS

The paper presented an automated pipeline for multi-view reconstruction of close-range objects. The final aim was to setup a software able to model free-form objects from images featuring good characteristics in terms of resolution, overlap and network geometry. The reconstruction process is automated, starting from image orientation phase up to the generation of a dense point cloud. Partitioning of the object is the only manual task, although an automated solution is under development. Here, we are not interested in the final step, i.e. mesh generation, as several commercial and open source solutions are available to accomplish this task.

An important aspect of this work is the joint use of CV and Photogrammetry techniques. In particular, we think that MGCM is a powerful matching method, as it is very robust, precise and invariant with respect to affine deformations (after setting good initial values). On the contrary, other CV methods can automate the typical photogrammetric workflow. In a few words, the advantages of both disciplines are combined, while shortcomings are reduced.

There are some limits in our approach, like the use of a 2.5D DSM as initial model. A solution based on a 3D TIN as seed model is under investigation. Another limit is the manual segmentation of complex objects. In our future work we will try to eliminate this step or, at least, to introduce an automatic segmentation procedure.

REFERENCES

Baltsavias, E.P, 1991. Multiphoto Geometrically Constrained Matching. Ph. D. thesis, Inst. of Geodesy and Photogrammetry, ETH Zurich, Switzerland, Mitteilungen No. 49, 221 pp.

Barazzetti, L., 2010. A trip to Rome: physical replicas of historical objects created in a fully automated way from photos. Proc. of HCITOCH. *Lecture Notes in Computer Science*, 6529: 63-72 (2011).

Barazzetti L., Remondino, F. and Scaioni M., 2010. Extraction of accurate tie points for automated pose estimation of close-range blocks. *ISPRS Technical Commission III Symposium on Photogrammetric Computer Vision and Image Analysis* , 6 pp.

Frahm, J. et al. 2010. Building Rome on a cloudless day. In: Proc. of ECCV 2010, 14 pp.

Furukawa, Y. and Ponce, J., 2010. Accurate, dense, and robust multi-view stereopsis. *IEEE Trans. PAMI*, 32(8): 1362-1376.

Furukawa, Y., Curless, B., Seitz, S.M. and Szeliski, R., 2010. Towards Internet-scale Multi-view Stereo. Proc. of IEEE Conf. CVPR'10, 8 pp.

Hartley, R., and A. Zisserman, 2003. *Multiple View Geometry in Computer Vision – 2nd Ed.* Cambridge University Press, UK.

Hirschmüller, H., 2008. Stereo processing by Semi-Global Matching and Mutual Information. *IEEE Trans. PAMI*, 30(2): 328–341.

Grün, A., 1985. Adaptive least square correlation – a powerful image matching technique. *S. Afr. J. of. Photogrammetry, Remote Sensing and Cartography*, 14(3), pp 175-187.

Grün, A. and Baltsavias, E.P. 1988. Geometrically Constrained Multiphoto Matching. *PE&RS*, 54(5), pp. 663-671.

Pierrot Deseilligny, M. and Clery, I., 2011. APERO, an open source bundle adjustment software for automatic calibration and orientation of set of images. *IAPRS&SIS*, 38(5/W16), on CD-ROM.

Remondino, F., El-Hakim, S.F., Grün, A. and Zhang, L., 2008. Turning images into 3-D models - development and performance analysis of image matching for detailed surface reconstruction of heritage objects. *IEEE Signal Processing Magazine*, 25(4), pp. 55-65.

Roncella, R., Re, C. and Forlani, G., 2011. Performance evaluation of a structure and motion strategy in architecture and cultural heritage. *IAPRS&SIS*, 38(5/W16), on CD-ROM.

Snavely, N., Seitz, S.M. and Szeliski, R., 2008. Modeling the world from internet photo collections. *IJCV*, 80(2), pp. 189-210.

Strecha, C.,Pylvanainen, T. and Fua, P., 2010. Dynamic and Scalable Large Scale Image Reconstruction. Proc. of CVPR'10, 8 pp.

Zhang, L., and Grün, A., 2006. Multi-image matching for DSM generation from IKONOS imagery. ISPRS Journal, 60, pp. 195-211.