Detailed 3D modelling of castles
El-Hakim, Sabry; Gonzo, I.; Voltolini, F.; Girardi, S.; Rizz, A.; Remondino, F.; Whiting, E.

International Journal of Architectural Computing, 5, 2, 2007-06

National Research Council Canada
Institute for Information Technology

Conseil national de recherches Canada
Institut de technologie de l'information

NRC · CNRC

Detailed 3D Modelling of Castles *

El-Hakim, S., Gonzo, I., Voltolini, F., Girardi, S., Rizz, A., Remondino, F., Whiting, E.
June 2007

* published in the International Journal of Architectural Computing. Volume 5, Number 2. June 2007. pp. 199-220. NRC 49290.

Canada

Detailed 3D Modelling of Castles

Sabry El-Hakim, Lorenzo Gonzo, Francesca Voltolini, Stefano Girardi, Alessandro Rizzi, Fabio Remondino and Emily Whiting

Detailed 3D Modelling of Castles

Sabry El-Hakim, Lorenzo Gonzo, Francesca Voltolini, Stefano Girardi, Alessandro Rizzi, Fabio Remondino and Emily Whiting

Digitally documenting complex heritage sites such as castles is a desirable yet difficult task with no established framework. Although 3D digitizing and modelling with laser scanners, Photogrammetry, and computer aided architectural design (CAAD) are maturing, each alone is inadequate to model an entire castle in details. We present a sequential approach that combines multiple techniques, each where best suited, to capture and model the fine geometric detail of castles. We provide new contributions in several areas: an effective workflow for castle 3D modelling, increasing the level of automation and the seamless integration of models created independently from different data sets. We tested the approach on various castles in Northern Italy and the results demonstrated that it is effective, accurate, and creates highly detailed models suitable for interactive visualization. It is also equally applicable to other types of large complex architectures.

1. Introduction

Digital 3D documentation of heritage buildings like castles is important for conservation, education, and virtual visits. Castles are unique and fascinating structures with varying architectural styles that evolved over the centuries. As they were built for defensive purposes, they made the most of readily available natural protective settings like hilltops and ridges, and were fortified with rings of high walls, strong gates, moats, and towers, as in Figure 1. This impedes data acquisition and modelling, particularly with cumbersome equipments like laser scanners

► Figure 1: Examples of castle structures and locations.

1.1. The 3D-Arch project

A joint project between ITC in Italy and NRC in Canada, named 3D-Arch, for modelling castles in Trentino province in northern Italy was launched in 2005. The goal is to capture complete geometric details on exterior and interior of castle buildings and model them with high resolution 3D triangular meshes for accurate documentation and photo-realistic visualization. Courtyards and walls should also be modelled. The cost and data collection time must fit the limited budget and granted access time. The preliminary study of the project [1] revealed several problems. Due to typical castle locations, it is hard to find adequate number of places from which to capture images or scans. Also, due to the complexity and variety of buildings, the assumptions made on standard architecture to automate 3D modelling, such as parallelism, perpendicularity, or symmetry, are often inappropriate. To solve these problems, multiple techniques such as Photogrammetry, surveying, laser scanning, and floor plans, must be used [2]. This requires developing techniques to seamlessly combine different models together and remove overlaps and fill gaps between them to create one model suitable for documentation and visualization.

1.2. Capturing fine geometric details

Models with fine geometric details are essential for realistic rendering under varying angles and illumination. The resolution of the rendered model should ideally match what is perceptible by the human eyes on a real visit. There are many ways to create virtual representations, like panoramas and walkthrough animations, but a textured 3D geometric model is the most desirable since it allows unrestricted interactive visualisation and manipulation and offers the user the freedom to choose any viewpoint at various lighting conditions. But even when rich texture maps are used, without fine geometric details the model will exhibit smooth, flat-looking, surfaces and polygonised silhouettes that are easily detected by the human eye. It is therefore important that the model captures fine 3D details to guarantee realistic experience even at close up. Available technologies for this purpose include laser scanning and Image-based modelling (IBM) such as dense stereo matching and shape from shading (SFS), texture, or shadows. Laser scanners promise to provide detailed and accurate representation of any shape [3]. Combined with colour information, either from the scanner itself or from a digital camera, realistic models can be created. Laser scanners remain costly, bulky, hard to use, influenced by surface properties, and impractical on many castle grounds. Due to object size, shape, and occlusions, it is usually necessary to use multiple scans from different locations to cover every surface. Aligning and integrating the different scans requires special tools and expertise. Also, scanners produce huge number of points, even on perfectly flat surfaces, yet the points really needed for reconstruction, like corners and edges, are often missed. IBM techniques can produce accurate and realistic-looking models but remain highly interactive since fully automated methods are still unproven in real applications. The needed interactivity will inevitably limit the amount of details a model can have.

1.3. The approach

Laser scanning will be used on selected parts, like inside rooms with geometrically interesting walls and ceilings, to directly capture geometric details otherwise difficult to capture from images. All other models will be produced with IBM since it entails low cost hardware and uncomplicated data capture. Dense stereo matching can be used to add fine geometric details however, occlusions, lack of texture, and light variations between images are persistent problems especially with widely separated views. In our approach, these problems are reduced by using a basic geometric model to narrow the search for matches and setting up logical rules for selecting surface regions where the matching can be reliably applied. In areas where stereo matching is ineffective, shape changes may be resolved from intensity or shade variations. Changes in captured light intensity due to surface reflection properties and position and orientation relative to the light source and the camera are direct functions of surface slope — the larger the angle the darker the surface. However, SFS is not accurate since it is based on crude assumptions to resolve its under-determined formulation. Thus we only implement SFS to bring out supplementary details such as stones *relative* to a basic model. Our IBM approach works as follows:

- Create a basic model of a structure and segment it into separate surface patches.

- Apply either stereo matching or SFS on each segment depending on its texture properties.
- Remaining segments, where both techniques are unsuccessful, revert to the basic model.

The approach has the following advantages:

- It uses a small number of images, which is convenient for castles.
- Interactivity is limited to measuring limited seed points to create and segment the basic model.
- Fine details are captured fully automatically from stereo matching and SFS, using the basic model to provide constraints and increase reliability.

Although Bi-directional Reflectance Distribution Function (BRDF) and inverse global illumination are essential to render the model under varying lighting conditions, they are beyond the scope of this paper. Also techniques to visualise huge 3D models in real time will not be discussed. The rest of the paper is organised as follows. The next section discusses relevant previous work. Castles used for the examples shown throughout the paper are then described, followed by description of the proposed approach in section 4. In section 5, we focus on the approach to capture fine details from images. Accuracy evaluation is outlined in section 6 followed by concluding remarks.

2. Previous work

There is a growing body of work on digital 3D modelling of large architectures like castles. We give a short overview of related work, organized by the implemented technique.

2.1. CAAD techniques using floor plans

This is the traditional and remains the most common approach. In the case of plans created directly in digital form from measurements, file formats such as AutoCAD or DXF allow the information to be arranged in separate layers, each containing different type of elements. Significant research has been devoted to convert existing line drawings to CAAD representation [4]. Creating 3D models may be done interactively on simple floor plans: closed poly lines are used to define solid portions of the model, differentiating between walls and interior space. These areas are then extruded to the desired wall height. However, this method can be time consuming and error prone on complex buildings thus, semi-automatic techniques were developed [5]. One technique modelled heritage buildings using an automatic building generation method based on a library of predefined building blocks [6]. For colour details, many use synthetic textures which yield computer-generated look, but some use textures from images which offer more realistic appearance [7].

2.2. Photogrammetry and surveying

Some Photogrammetry and image-based techniques are available. The *Façade* system [8] starts with an interactive step to construct the overall shape using pre-defined blocks. A second automated step adds details using stereo matching. The medieval fortress Kufstein, Austria, was modelled with images taken from a helicopter and on the ground plus surveying and CAAD software [9]. Similar approaches were used to model castles in Germany [10] and Sicily [11]. Tourist slides and surveying were used to model the Holy Sepulchre church in Jerusalem using Photogrammetry combined with existing floor plans [12]. The methods are still labour intensive and the projects were reported to take several years to complete.

2.3. Automated and semi-automated image-based techniques

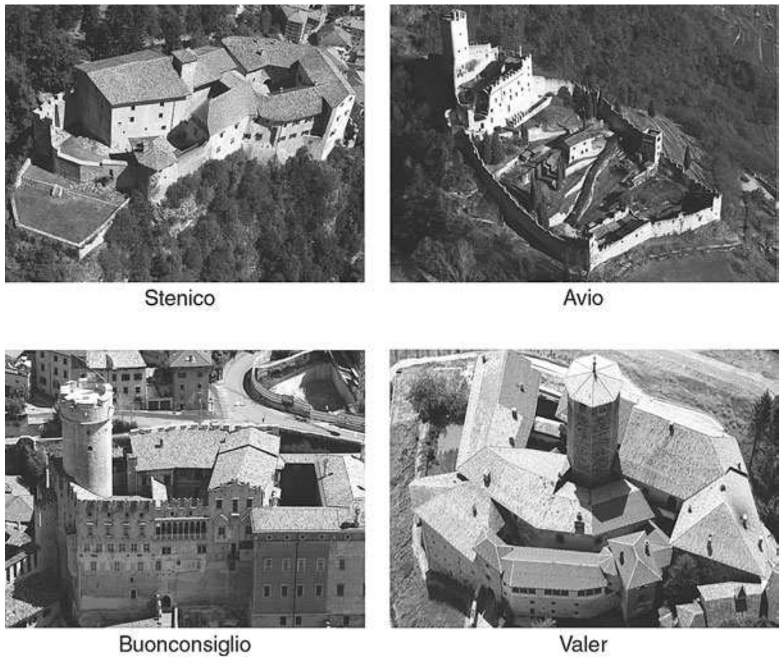
These techniques focus on the automatic recovery of camera calibration and orientation parameters plus 3D coordinates of feature points. However, acquiring points suitable for modelling and segmenting the point clouds into topologically meaningful groups remain interactive [13]. Fine details are usually added from dense stereo matching following the determination of camera parameters. Closely spaced images, like videos, are required for successful matching, which may be difficult for castles. Even if accessibility is not an issue, covering a castle with closely spaced image sequences is impractical. Also, accuracy becomes an issue on long image sequences due to error propagation unless points are tracked in many images. Wide base-line stereo matching has been attempted [14] but no practical results on complex structures or performance evaluation are yet available. To our knowledge, no castle model was completed based purely on fully automated techniques, although small sections have been demonstrated [15]. Techniques attempting full automation of the modelling process [16] rely on constraints of surface shapes and assumed relationships between surfaces and many require vanishing points from sets of parallel lines. Since in most castles these assumptions do not apply, such techniques are ineffective. A practical solution is to incorporate limited interaction on regular shapes such as arches, columns, doors and windows [17]. With a small number of interactively measured seed points, the remaining points defining the model of those elements can be added automatically.

2.4. Laser scanning and combination of multiple techniques

Laser scanning is increasingly being used for large site digitization [18]. Due to the drawbacks mentioned in section 1, results vary widely from one published work to another and in many sites it is combined with other techniques. For example, Photogrammetry was utilized for outlines and main shapes and laser scanning for fine geometric details [19, 20].

3. The castles of Trentino

Each of the castles reported in this paper presented different challenges, which gave us a great opportunity to advance our approach into a general system that can be applied to most other castles. Figure 2 shows aerial views of the *Buonconsiglio*, *Avio*, *Stenico*, and *Valer* Castles.



► Figure 2: Selected castles of Trentino, Northern Italy.

Each castle has assorted elements, as depicted in Figure 3:

- Buildings and towers of mixed styles from different periods organized around courtyards.
- Inner and outer tall thick walls and arched gates.
- Renaissance loggias, some with rich frescos such as the Romanino loggia at Buonconsiglio castle.
- Chapels and rooms with significant frescos, such as the "Cycle of the months" medieval frescos at Buonconsiglio castle.
- Partially destroyed buildings and ruins, such as those found in the Avio castle.

Aerial and ground images were taken with the 14 mega-pixel *Kodak*® DCS Pro SLR-camera at different focal settings depending on available space

► Figure 3: Images of various castle parts.

around each building. This took an average of 4 hours in each site. Some parts, such as chapels and frescoed rooms, were scanned with *Leica*® HDS3000 3D laser scanner at 1 cm resolution and 4–6 mm accuracy. Digital images were used for texturing the scanned surfaces. One day was usually spent in surveying of 100–120 points evenly distributed on outside walls with *Leica*® total station. We also obtained existing floor plans in AutoCAD DWG format. We will describe our approach in details using results from the four castles.

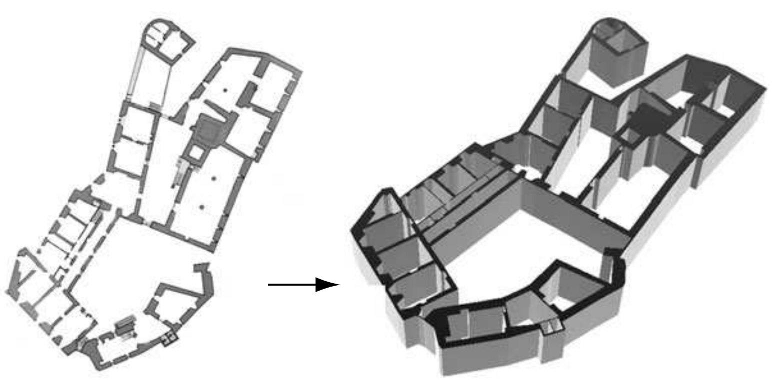
4. The squential approach for castle documentation

Our approach is summarized as follows:

1. Use floor plans, which exist for most castles, to create a non-detailed less accurate model with approximate heights.
2. Pre-calibrate high resolution digital camera at various settings.
3. Survey evenly distributed points on the outer walls to define the reference coordinate system, scale, and connect independently-created models.
4. Take images from a helicopter and create an overall model of building exteriors, walls, courtyards, and main grounds. This model will substitute most of the model created in step 1.
5. Take ground images to model parts not visible from aerial images and add fine geometric details.
6. Use medium-range (1m–50 m) laser scanner to capture details of selected interiors.
7. Assemble and integrate all models.

4.1. Approximate model from floor plans

Floor plans are used to add sections missed by imaging or scanning and for model assembly (section 4.5). To convert the floor plans into a 3D model, semantic information, such as room identities and connecting openings, is necessary followed by walls extrusion to heights taken from the sensor-based models, Figure 4. Window and door insertions are carried out where needed. We use *AutoCAD*® software for this process.

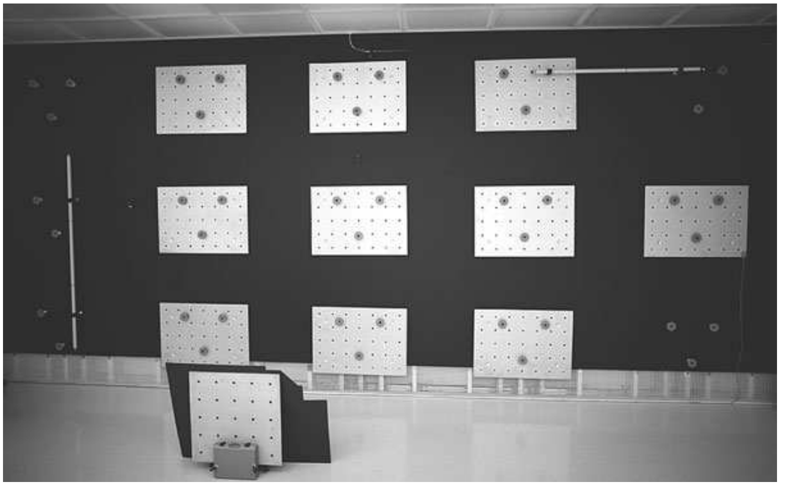


◀ Figure 4: Floor plans.

4.2. Camera calibration

Precise camera calibration including all lens distortion parameters is crucial for 3D capture. The camera can be self-calibrated from the project images or calibrated separately at a different location or a laboratory [21]. Self-calibration is necessary if camera settings are unknown and vary between images, but to achieve accurate and reliable calibration that includes all distortion parameters, certain geometric configurations of images are needed. Since this is often not possible at castle sites, it is sensible to take the images at known fixed settings then calibrate in the lab at those settings using surveyed points like those in Figure 5.

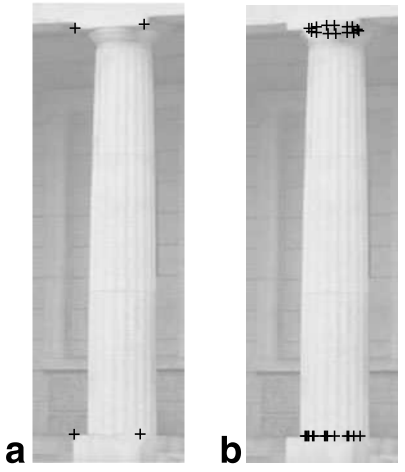
► Figure 5: Calibration targets.



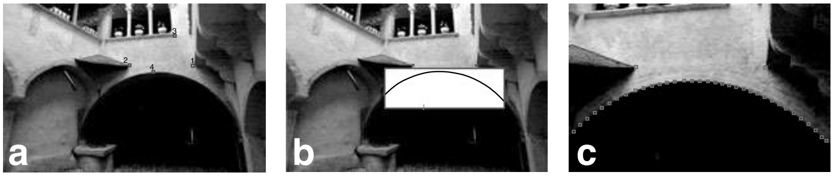
4.3. Basic models from aerial and ground images

Aerial images are used to create overall model of buildings and grounds while the façades detail and the inside of rooms require images taken from ground level. We create basic models of surface elements such as planer walls, columns, windows, and arches using an approach initially developed in [17]. For example, a column is automatically constructed from four seed points, two on the corner of the top crown and two on the corners of its base, Figure 6. From these points, the radius of the column and direction of its axis can be computed. The ratio between the upper and the lower circle is usually set to 0.85. 3D points on top and bottom circles of the column can then be automatically added. For reconstructing arches, first a plane is fitted to seed points on the wall, Figure 7a. A morphological edge detector [22] is applied to the region, Figure 7b, and points at constant interval along the arch are automatically sampled. Using image coordinates of these points in one image, the known image parameters, and the plane parameters, the 3D coordinates are computed and projected on the image, Figure 7c. For windows and doors we need four outside corner points and one point on the inside surface. By fitting a plane to the corner points and a plane parallel to it at the surface point, the complete window or door is created.

Table 1 displays the number of seed points required to be measured interactively for various shapes. This semi-automatic procedure requires an average of 1–2 days of work by one person for each model. Figure 8 shows snap shots of castle models created by this technique.



◀ Figure 6: Seed points (a) and automatically added column points.



◀ Figure 7: Arch modelling.

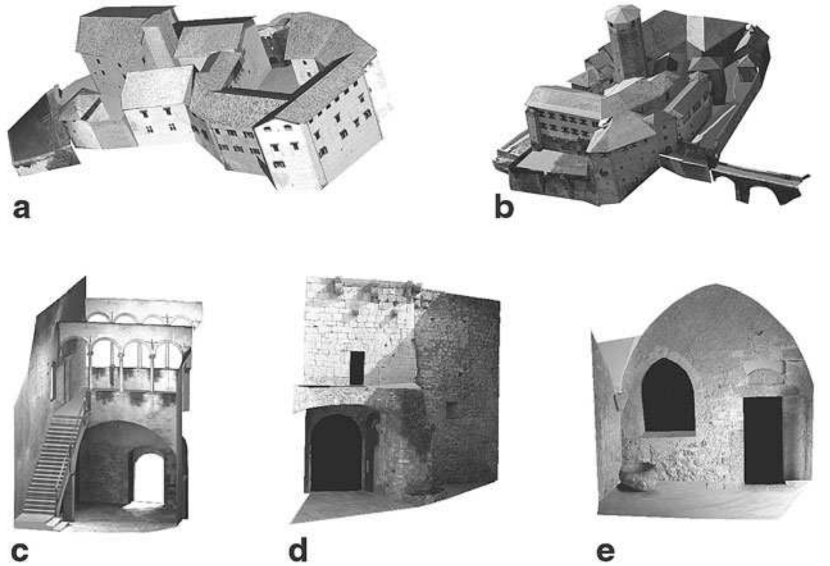
Element	Seed Points
Plane	3: none-linear
Column or cylinder	4: 2 on base, 2 on crown
Window or door	5: 1 in each corner, 1 inside
Block attached to wall/plane	1 per corner
Arch	3: on front surface/plane

◀ Table 1: Seed points required for different building elements.

4.4. Detailed models from laser scanning

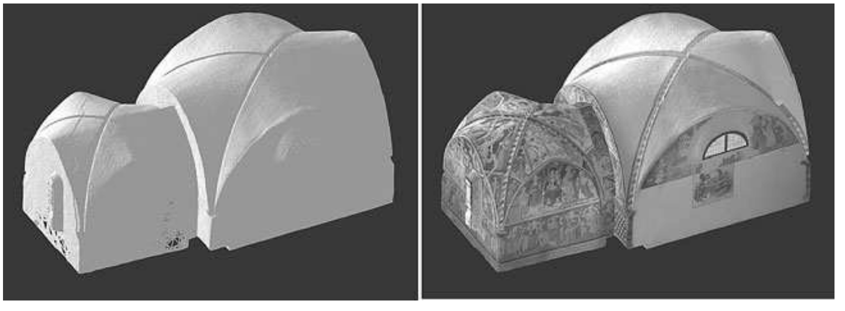
Indoor rooms with interesting geometric and texture details are good candidates for laser scanning since the scanner can be better set up in an indoor environment. One of the most important issues is to determine the spatial resolution to be used. This depends on the desired level of details, the type of scanner, and the room size. The number of scans is affected by the spatial resolution, surface shape complexity and occlusions. We use PolyWorks© software for multi scan registration and for creating the triangular mesh and our own software for texture mapping. Examples of such models for some castle rooms are shown in Figures 9–11.

► Figure 8: Examples of basic models from aerial and ground images.

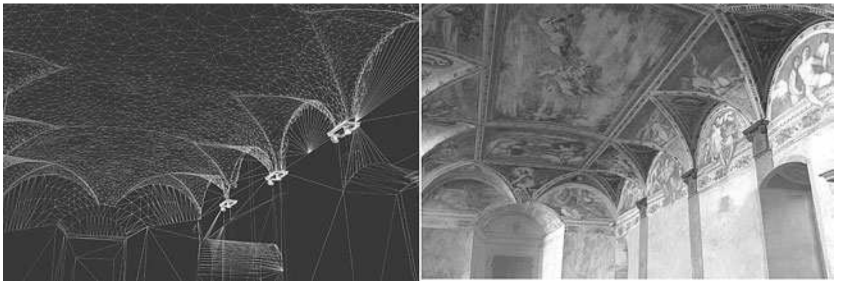


► Figure 9: Wire-frame, shaded, and textured model of Valer castle chapel altar.

► Figure 10: Wire-frame and textured model of Valer castle chapel.

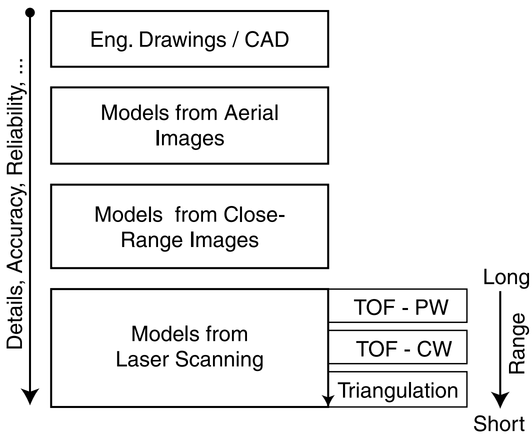


► Figure 11: Wire-frame and textured model of Romanino room, Buonconsiglio castle.



4.5. Model assembly and integration

The procedure is hierarchical, by data source, where details, accuracy and reliability increase as we advance from one data level to the next, Figure 12. In overlap regions, data in one level overrides data in previous levels. Engineering drawings provide the least detailed or accurate level since they may not reflect changes that happened since their completion. Accuracy, details and reliability are better with models from aerial images and even better from ground image. A higher level of details can be achieved by Time-Of-Flight (TOF) – Pulsed Wave (PW) laser scanners, and accuracy improvement by the TOF- Continuous Wave (CW) scanners. Triangulation-based laser scanners, if used at close range, provide the most detailed, accurate and reliable data.



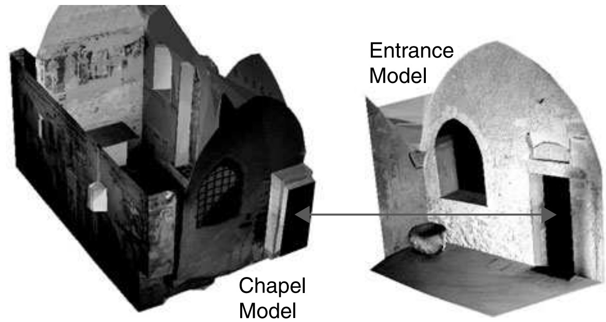
◀ Figure 12: Data level hierarchy.

To combine models from different data sets into one model appropriate for reconstruction and visualization:

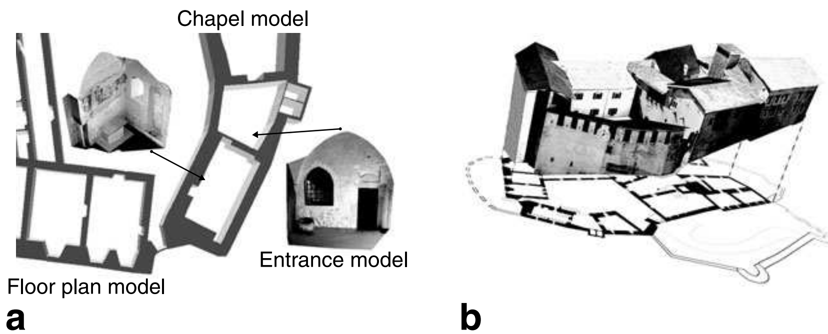
- They must have the correct scale and orientation.
- Joint primitives; surfaces, edges, and vertices, from adjacent models must match perfectly.
- No gaps, redundant surfaces or intersecting edges are acceptable.
- Textures on adjacent models should integrate seamlessly [23].

Some of these issues are not dealt with in commercial software, thus special techniques and software tools were developed. The models are registered manually in modelling software such as *3ds max*® using common points or portals, Figure 13. The process will be unnecessary if the individual models were directly created in the same coordinate system using surveyed points. When models have no visible common points with other models or no surveyed points, floor plans are used for positioning those models as in Figure 14a. Any remaining gap between models is filled from the floor plans, Figure 14b.

▶ Figure 13: Model registration with common portal.



▶ Figure 14: Model assembly with floor plans.



Once all models are registered, they are integrated according to the hierarchical order shown in Figure 12. The mesh is adjusted to remove redundant surfaces from the lower level models, and in some cases Boolean operations are required to create holes for openings. Lower level models are re-triangulated to account for point changes and to create holes where detailed models are inserted.

5. Adding the fine details from images

From section 4, we have a basic model but without fine geometric details. Using this model as a guide and knowing camera calibration and orientation parameters, we developed an automatic procedure to model fine details with high-resolution meshes to achieve accurate documentation and photo-realistic visualisation. Several techniques are used, each where best suited:

1. The basic model created above is segmented into separate surfaces patches.
2. For patches with regular shape, fit an implicit function (e.g. plane, cylinder or quadric) using seed points and apply a relative stereo matching technique.
3. More complex or irregular segments with unknown approximate function use an absolute multi-image matching technique.
4. For segments unsuited for stereo matching, apply depth from shading (DFS) technique.

We now describe each technique in some detail.

5.1. Dense stereo matching

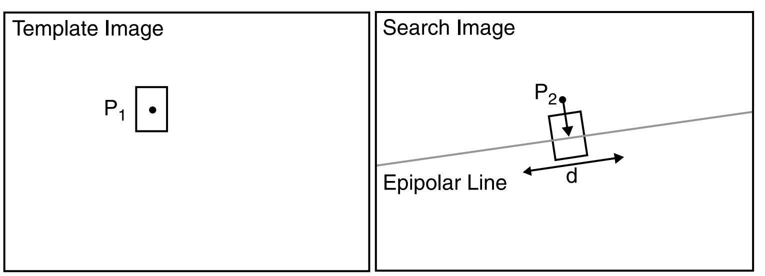
An overview of dense stereo matching techniques can be found in [24]. Occlusions, lack of texture, and light variations between images are persistent problems for stereo matching especially with widely separated views. Template-based stereo matching works best when sufficient texture variations or localised features are present on the surface. Therefore, the first rule we use to select the areas where stereo matching will apply is intensity-level analysis of the template window. This includes mean, standard deviation, and second derivative of the grey-levels of the pixels in the window. If those are higher than preset thresholds, the stereo matching will proceed otherwise we consider the region to be too uniform for stereo matching and switch to DFS, which works best on smoothly shaded surfaces.

Relative stereo matching

In our approach, the problems are reduced by: using the basic model to narrow the search for matching; and set up logical rules for selecting image regions where the matching can be reliably applied. For each segment with known fitted function:

1. A high-resolution mesh of triangulated 3D points, one per pixel, is placed automatically on each segment based on its fitted function.
2. Using preset rules, the coordinates of the points in a segment are adjusted using stereo matching. We only compute the difference between the coordinates from the basic model and the final coordinates from the stereo matching.

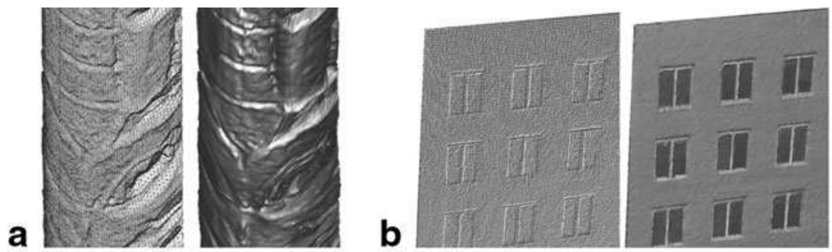
The stereo matching is based on minimizing the normalised squares of the difference between the template and the search window. The search is done along the epipolar line and we also limit the search to a disparity range computed from the basic model. For example in Figure 15, point P_1 in the template image has a corresponding point P_2 in the search image that is



◀ Figure 15: Stereo matching with search constraints.

computed directly from the basic model. Based on maximum depth variation (roughly preset), we can easily compute the region on the epipolar line (distance d) where we limit the search. The window in the search image is re-sampled to take into account the difference in orientation between the two images and surface orientation of the basic model. This accounts for the geometric variations between these two images and gives accurate and reliable results. We now apply another rule to accept or reject the matched points. If the best-matched window differs from the template by more than a threshold (pre-set experimentally), the matching is considered unreliable or invalid (e.g. the region is occluded in the right image) and the system reverts to the basic model point, which is point P_2. Figure 16 shows two examples, using (a) cylinder and (b) plane as basic models.

► Figure 16: Relative stereo matching examples an (a) carved cylinder and (b) planer façade.



Absolute multi-image matching

The stereo matching approach presented above, although fast (seconds or few minutes per segment) and effective, requires an approximate surface shape. However, for irregular, complex or damaged surfaces, the approximate shape is unknown. Therefore, an extended, albeit slower, more global approach that does not require knowledge of an approximate surface has been developed. Entire objects or full images can be used rather than segments. It is based on non-linear least-squares estimation that solves for several parameters including the matched pixel location and photometric images differences [25, 26]. It uses more than two images to increase its accuracy and reliability by simultaneously matching the point in all the images it appears in. It is a coarse-to-fine hierarchical solution with automatic quality control.

The approach illustrated in Figure 17 performs three mutually connected steps:

► Figure 17: The absolute multi-image matching procedure.

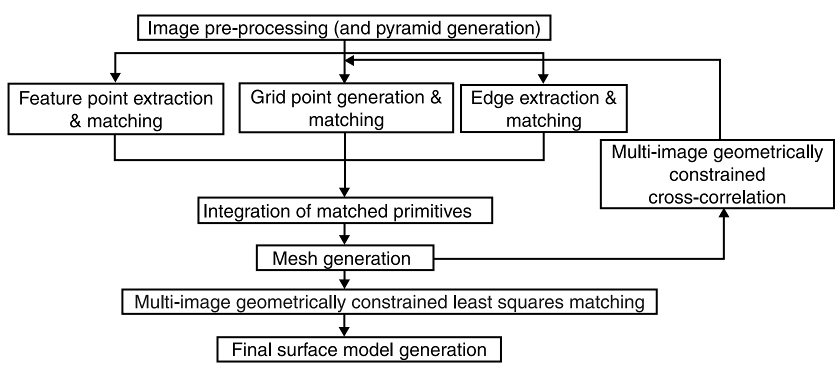
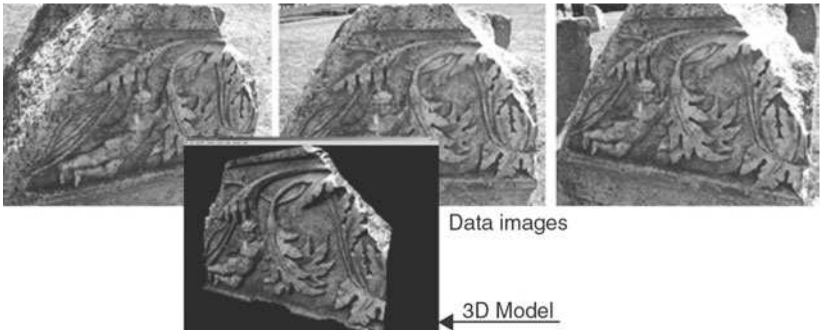


Image pre-processing: the set of available images is processed with an adaptive smoothing filter in order to reduce the effects of the radiometric problems such as strong bright and dark regions and optimizes the images for subsequent feature extraction and image matching. Generating image pyramids, which are several versions of the image with progressive spatial resolutions, follows.

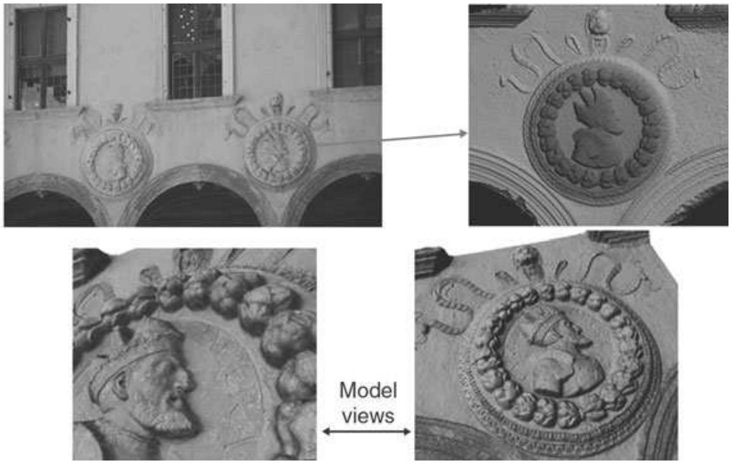
Multiple Primitive Multi-Image (MPM) matching: this part utilizes a coarse-to-fine hierarchical matching strategy for accurate and robust surface reconstruction. Starting from the low-density features in the lowest resolution level of the image pyramid, the MPM matching is performed with two or more images, incorporating multiple matching primitives (feature points, edges, and grid points/pixels). Feature points are suitable to generate dense and accurate surface models but they suffer from problems caused by noise, occlusions, and discontinuities. Edges generate coarser but more stable models as they have higher semantic information and they are more tolerant to image noise. The MPM performs three operations in each pyramid level: features and edges extraction and matching, integration of matching primitives and initial mesh generation. Within the pyramid levels, the matching is performed with an extension of the standard cross-correlation technique while only in the last (original) level a multi-photo geometrically constrained LSM [27] is performed. The multi-image matching is guided from object space and allows reconstruction of 3D objects from all available images simultaneously. The high redundancy from multiple images allows also automatic detection and removal of mismatches through consistency checking within a small neighbourhood. Moreover, at each pyramid level, a triangular mesh is reconstructed from the matched features using the constrained Delauney triangulation method. The mesh is used in the subsequent pyramid level for derivation of approximations and adaptive computation or self-tuning of the matching parameters. They are automatically determined by analyzing the results of the higher-level image pyramid matching and using them at the current pyramid level. These parameters include the size of the correlation window, the search distance and the threshold values. The adaptive determination of the matching parameters results in higher success rate and less mismatches.

Refined matching: a modified Multi-Photo Geometrically Constrained Matching (*MPGC*) and the Least Squares B-Spline Snakes (*LSB-Snakes*) methods are used to achieve potentially sub-pixel accuracy matches and identify some inaccurate and possibly false matches. This is applied only at the original image resolution level. The surface derived from the previous MPM step provides well enough approximations for the two matching methods and increases the convergence rate. Figure 18 and 19 show example results.

► Figure 18: Unstructured shape modelled with absolute multi-image matching.



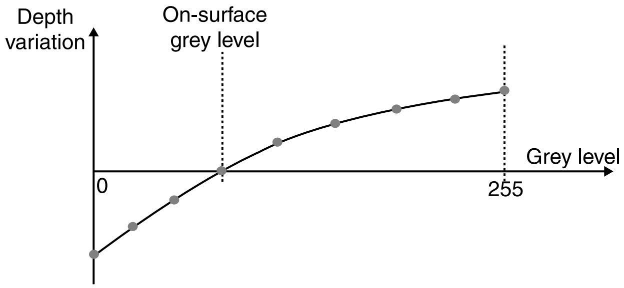
► Figure 19: Detailed model from Buonconsiglio castle.



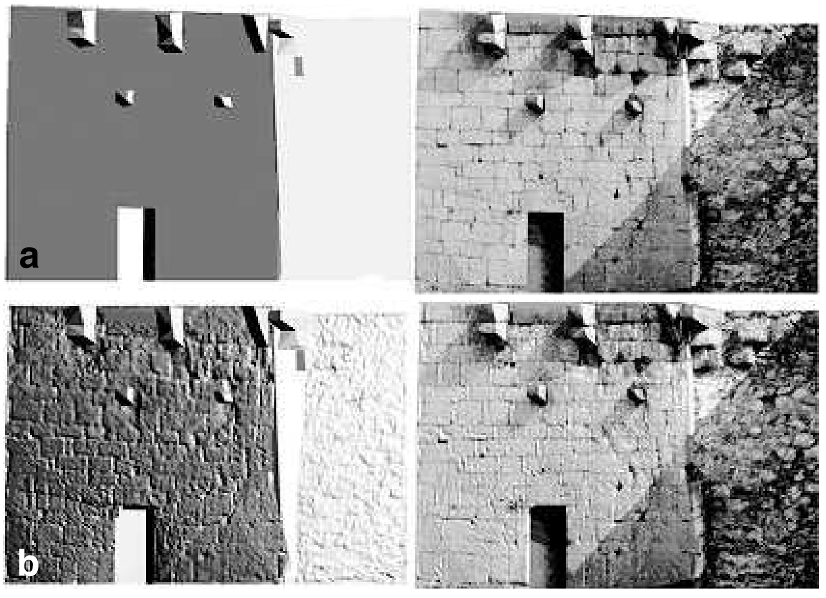
5.2. Depth from shading

DFS is applied where grey-level variations are not adequate for stereo matching and sections appearing only in a single image. Standard shape from shading techniques [28] lacked success in determining absolute 3D shape of objects in actual applications due to its ill-posed formulation and unrealistic assumptions, such as the camera looks orthogonally at a Lambertian surface and there is only one single light source located at infinity. Our approach computes the depth directly, rather than surface normal. It is applied to a work image: a grey-level version of the original with some pre-processing such as noise removal filtering and editing of unwanted shades and elements. Using known depth and grey level at 8–10 points determined interactively we form a curve describing the relation between grey-levels and depth variation from the basic model, Figure 20.

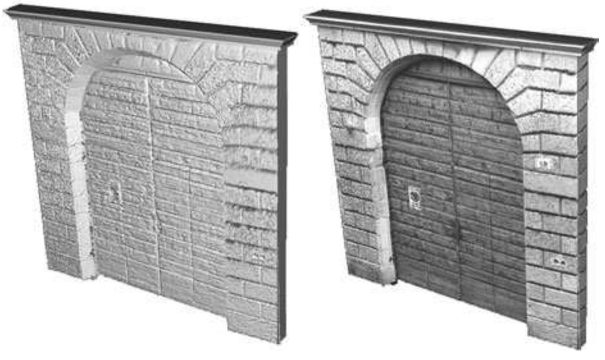
The curve intersects the grey-level axis at the average intensity value of points actually falling on the basic model. By adjusting the curve the results can be instantly reviewed. We adjust the coordinates of the grid points on the surface of the basic model segment according to shading using this curve. We now have a triangulated grid of points whose coordinates are altered from the initial basic model to account for the fine details. Since DFS will usually produce less accurate results, it will only be applied to parts where exact dimensions are not necessary, such as stones, bricks, and wood, Figure 21–22.



◀ Figure 20: Grey-level versus depth variation relative to basic model.



◀ Figure 21: Basic model (a), appearance improvement with DFS (b).

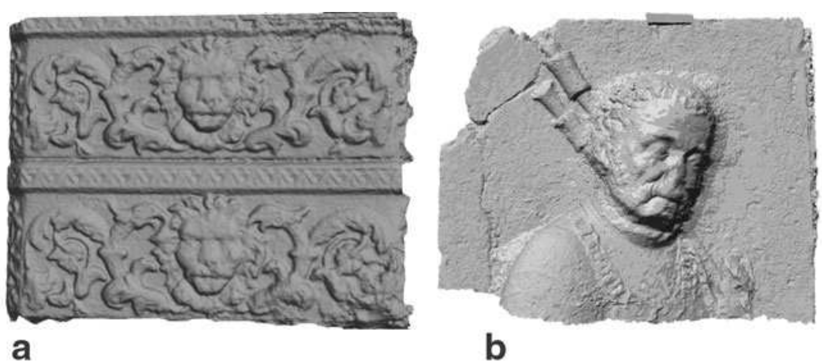


◀ Figure 22: Example of DFS.

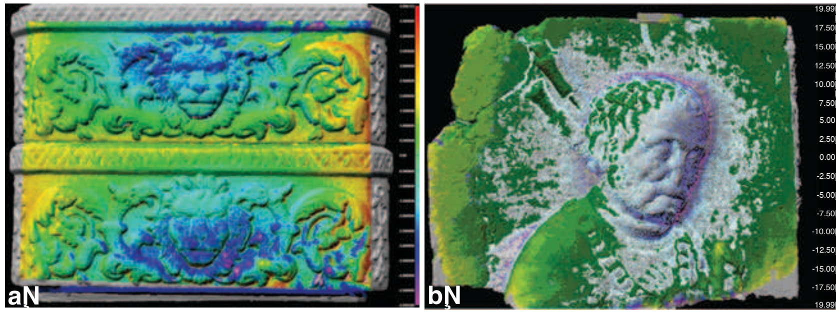
6. Accuracy evaluation

Surveyed points on the buildings façades are used as ground truth to evaluate the accuracy of the models, both individual models and integrated models. Here we give the results from the Stenico castle as a typical example. The castle buildings extend over about 100 meters by 64 meters area and have up to 35 meters height. The aerial images were taken from a range of 65 meters to 120 meters. The image average field of view was 160 meters wide. One pixel in the image corresponds to about 8 cm when using 9 mm focal length. The accuracy is assessed by the variance-covariance matrix of the computed 3D points, and validated by surveyed checkpoints that had 3 mm accuracy. The achieved accuracy averaged: **17** *mm* (X), **15** *mm* (Y), and **16** *mm* (Z). This is one part in 10,000 and represents 0.2 pixels. For the ground-level models, we achieved accuracies of: **1** *mm* to **2** *mm*, which give average relative accuracy of one part in 6,000 and represents 0.3 pixels. The superior relative accuracy reached from aerial images, even though they were taken from much longer ranges, is attributed to the better geometric configuration compared to ground images. Ground images had less than ideal locations due to tight spaces around most buildings and inside rooms. Results from other castles were similar, except for Valer castle which had less accuracy on ground models (**2** *mm* to **3** *mm*) due to its tighter spaces.

► Figure 23: (a) Lab test model and (b) on-site model from dens stereo matching.



► Figure 24: Colour-coded difference between stereo matching and accurate laser scanner.



We extensively tested our matching approach on several test objects, one of which is in a lab environment shown in Figure 23a, and several on castle wall carvings as the one shown in Figure 23b. For ground truth, the objects were scanned with a triangulation-based laser scanner with 0.5 mm accuracy. The same objects were then modelled with both matching techniques described in 5.2 and 5.3. To compare these models with ground truth data, we used PolyWorks© Inspector software. Sample results are illustrated in Figure 24a and 24b. The average difference between the scanned model and the image-based model was **0.5** *mm* to **0.7** *mm*, for all data sets and both matching techniques.

7. Conclusions

Through the experience of 3D modelling of several different castles in Northern Italy, a clear workflow has been established. Using images taken from a low-flying helicopter and from ground, combined with existing floor plans and limited surveying and laser scanning, we can completely model a complex castle with realistic details. Semi-automated modelling and model assembly techniques were developed. The time required to collect the data was only 2–3 days per castle and the data processing and modelling averaged 1–2 month, which is much less than reported similar work. To capture fine geometric details from images, we presented a sequential multi-stage segment-based procedure starting with basic model of surface elements then adding details with relative stereo matching, absolute multi-image matching, and depth from shading, each where best suited. The results have shown that our semi-automated approach creates highly detailed models with accuracy in the millimetre range and at low cost.

Acknowledgements

This project is funded by the Autonomous Province of Trento. Dr. Franco Marzatico, the director of Castelli Trentini, Count von Spaur, owner of Castel Valer, and Fondo Italiano per l'Ambiente owner of Castel Avio kindly allowed us access and provided the floor plans.

References

1. El-Hakim, S. F., Gonzo, L., Girardi, S., Picard, M., Whiting, E., Photo-realistic 3D reconstruction of castles with multiple-sources image-based techniques, in: *Proc. ISPRS XXth Congress,* 2004, Istanbul, 120–125.
2. El-Hakim, S. F., 3-D modelling of complex environments, in: El-Hakim, S. and Gruen, A., eds., *Proc. Videometrics and Optical Methods for 3D Shape Measurements,* SPIE Vol. 4309, 2001, San Jose, CA, 162–173.
3. Blais, F., Review of 20 years of range sensor development, *Electronic Imaging,* 2004, 13(1), 232–240.
4. Kernighan, B. W., Van Wyk, C. J., Extracting geometric information from architectural drawings, in: *Proc. Workshop on Applied Computational Geometry (WACG),* 1996, 82–87.

5. Lewis, R., Sequin, C., Generation of 3D building models from 2D architectural plans, *Computer-Aided Design,* 1998, 30(10), 765–779.
6. Dikaiakou, M., Efthymiou, A., Chrysanthou, Y., Modelling the walled city of Nicosia, in: Arnold, D., Chalmers, A., Niccolucci, F., eds., *4th International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage (VAST'2003),* Eurographics, UK, 2003, 57–66.
7. Foni, A. E., Papagiannakis, G., Magnenat-Thalmann, N., Virtual Hagia Sophia: restitution, virtualization and virtual life simulation, in: *UNESCO World Heritage Congress,* 2002.
8. Debevec, P., Taylor C. J., Malik, J., Modelling and rendering architecture from photographs: A hybrid geometry and image-based approach, in: *Proc. of SIGGRAPH,* 1996, ACM, 11–20.
9. Hanke, K., Oberschneider, M., The medieval fortress Kufstein, Austria – an example for the restitution and visualization of cultural heritage, in *Proc. ISPRS Commission V Symposium,* 2002, Corfu, Greece, 530–533.
10. Kersten, T., Pardo, C. A., Lindstaedt, M., 3D acquisition modelling and visualization of north German castles by digital architectural Photogrammetry, in: *Proc. ISPRS XXth Congress,* Istanbul, 126–131.
11. Bacigalupo, C., Cessari, L., Survey techniques and virtual reality for recovery plan of a fortified Mediterranean town, in: *Int. Workshop Vision Techniques for Digital Architectural and Archaeological Archives,* Ancona, 2003, 40–44.
12. Georgopoulos, A., Modatsos, M., Non-metric bird's eye view, in: *Proc. ISPRS Commission V Symposium,* 2002, Corfu, 359–362.
13. Gibson, S., Hubbold, R. J., Cook, J., Howard, T. L. J., Interactive reconstruction of virtual environments from video sequences, *Computers & Graphics,* 2003, 27, 293–301.
14. Strecha, C., Tuytelaars, T. and Van Gool, L., Dense Matching of Multiple Wide-baseline Views. In: *IEEE Proceedings of ICCV,* 2003, Vol. 2, 1194–1201.
15. Pollefeys, M., Van Gool, L., Vergauwen, M., Verbiest, F., Cornelis, K., Tops, J. and Koch, R., Visual modelling with a hand-held camera, *Int. Journal of Computer Vision,* 2004, 59(3), 207–232.
16. Werner, T., Zisserman, A., New technique for automated architectural reconstruction from photographs, in: *Proc. 7th-European Conf. on Computer Vision,* 2002, vol. 2, 541–555.
17. El-Hakim, S. F., Semi-automatic 3d reconstruction of occluded and unmarked surfaces from widely separated views, in: *Proc. ISPRS Commission V Symposium,* 2002, Corfu, 143–148.
18. Allen, P. K., Troccoli, A., Smith, B., Murray, S., Stamos, I., Leordeanu, M., New methods for digital modelling of historic sites. *IEEE Computer Graphics and Applications,* 23(6), 2003, 32–41.
19. El-Hakim, S. F., Beraldin., Picard, M. and Godin, G., Detailed 3D reconstruction of large-scale heritage sites with integrated techniques, *IEEE Computer Graphics & Applications,* 2004, 23(3), 21–29.
20. Borg, C. E., Cannataci, J. A., Thealasermetry: a hybrid approach to documentation of sites and artifacts, in: *Proc. CIPA-ISPRS Workshop on Scanning for Cultural Heritage,* 2002, Corfu, 93–104.
21. Remondino, F., Fraser, C., Digital camera calibration methods: considerations and comparisons, in: *Proc. ISPRS Commission V Symposium,* 2006, Dresden, 266–272.
22. Lee, J., Haralick, R., Shapiro, L., Morphologic edge detection. *IEEE J. Robotics and Automation,* 1987, 3(2), 142–156.

23. El-Hakim, S.F., Gonzo, L., Picard, M., Girardi, S., Simoni, A., Paquet, E., Viktor, H., Brenner, C., Visualization of highly textured surfaces, in: Arnold, D., Chalmers, A., Niccolucci, F., eds., *4th International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage (VAST),* Eurographics, UK, 2003, 231–240.
24. Scharstein, D. and Szeliski, R., A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms. *Int. Journal of Computer Vision,* 2002, 47(1–3), 7–42.
25. Gruen, A., Adaptive least squares correlation: a powerful image matching technique. *South African Journal of Photogrammetry, Remote Sensing and Cartography,* 1985, 14(3), 175–187.
26. Remondino, F., Zhang, L., Surface Reconstruction Algorithms for Detailed Close-Range Object Modelling, in: *Proc. ISPRS Commission III Symposium,* 2006, IAPRS, Vol. XXXVI, Part 3, 117–123.
27. Gruen, A., Baltsavias, E., Adaptive least squares correlations with geometrical constraints, in: *Proc. of SPIE,* Vol. 595, 1986, 72–82.
28. Zhang, R., Tsai, P.-S., Cryer, J. E., Shah., M, Shape from shading: A survey. *IEEE Trans. PAMI,* 1999, 21(8), 690–706.

Sabry El-Hakim
National Research Council
Visual Information Technology Group
1200 Montreal Rd., Ottawa, Ontario, Canada

sabry.el-hakim@nrc-cnrc.gc.ca

Lorenzo Gonzo, Francesca Voltolini, Stefano Girardi, Alessandro Rizzi
Centre for Scientific and Technological Research
ITC-irst
Trento, Italy

(lgonzo, fvoltolini, girardi, rizziale)@itc.it

Fabio Remondino
Swiss Federal Institute of technology
Institute of Geodesy and Photogrammetry
ETH-Zurich, Switzerland

fabio@geod.baug.ethz.ch

Emily Whiting
MIT
Department of Architecture
Cambridge, MA, USA

ewhiting@mit.edu