AUTOMATED IMAGE-BASED PROCEDURES FOR ACCURATE ARTIFACTS 3D MODELING AND ORTHOIMAGE GENERATION

Marc PIERROT-DESEILLIGNY\textsuperscript{a}, Livio DE LUCA\textsuperscript{b}, Fabio REMONDINO\textsuperscript{c}

\textsuperscript{a} Institut Géographique National (IGN), Paris, France  
Email: marc.pierrot-deseilligny@ign.fr

\textsuperscript{b} Laboratoire MAP-Gamsau (CNRS/MCC), Marseille, France  
Email: livio.deluca@gamsau.archi.fr

\textsuperscript{c} 3D Optical Metrology Unit, Bruno Kessler Foundation (FBK), Trento, Italy  
Email: remondino@fbk.eu

Keywords: Photogrammetry, 3D Modeling, Bundle Adjustment, Orthoimage, Open-source

Abstract: The accurate 3D documentation of architectures and heritages is getting very common and required in different application contexts. The potentialities of the image-based approach are nowadays very well-known but there is a lack of reliable, precise and flexible solutions, possibly open-source, which could be used for metric and accurate documentation or digital conservation and not only for simple visualization or web-based applications. The article presents a set of photogrammetric tools developed in order to derive accurate 3D point clouds and orthoimages for the digitization of archaeological and architectural objects. The aim is also to distribute free solutions (software, methodologies, guidelines, best practices, etc.) based on 3D surveying and modeling experiences, useful in different application contexts (architecture, excavations, museum collections, heritage documentation, etc.) and according to several representations needs (2D technical documentation, 3D reconstruction, web visualization, etc.).

1. INTRODUCTION

The creation of 3D models of heritage and archaeological objects and sites in their current state requires a powerful methodology able to capture and digitally model the fine geometric and appearance details of such sites. Digital recording, documentation and preservation are demanded as our heritages (natural, cultural or mixed) suffer from ongoing attritions and wars, natural disasters, climate changes and human negligence. In particular the built environment and natural heritage have received a lot of attention and benefits from the recent advances of range sensors and imaging devices \cite{1}-\cite{3}. Nowadays 3D data are a critical component to permanently record the form of important objects and sites so that, in digital form at least, they might be passed down to future generations. This has generated in the last decade a large number of 3D recording and modeling projects, mainly led by research groups, which have realized very good quality and complete digital models \cite{4}-\cite{8}. Indeed remote sensing technologies and methodologies for Cultural Heritage 3D documentation and modeling \cite{10} allow the generation of very realistic 3D results (in terms of geometric and radiometric accuracy) that can be used for many purposes like historical documentation, digital preservation and conservation, cross-comparisons, monitoring of shape and colors, simulation of aging and deterioration, virtual reality/computer graphics applications, 3D repositories and catalogues, web-based visualization systems, computer-aided restoration, multimedia museum exhibitions and so on. But despite all these potential applications and the constant pressure of international heritage organizations, a systematic and targeted use of 3D surveying and modeling in the Cultural Heritage field is still not yet employed as a default approach. And when a 3D model is generated, it is often subsampled or reduced to low-resolution model for online visualization or to a 2D drawing due to a lack of software or knowledge in handling properly 3D data by non-expert. Although digitally recorded and modeled, our heritages require also more international collaborations and information sharing, to make them accessible in all the possible forms and to all the possible users and clients, e.g. via web \cite{11}. Nowadays the digital documentation and 3D modeling of Cultural Heritage should always consist of \cite{2}:

- Recording and processing of a large amount of 3D (possibly 4D) multi-source, multi-resolution, and multi-content information;
- Management and conservation of the achieved 3D (4D) models for further applications;
- Visualization and presentation of the results to distribute the information to other users allowing data retrieval through the Internet or advanced online databases;
- Digital inventories and sharing for education, research, conservation, entertainment, walkthrough, or tourism purposes.
The article deals with the first of the aforementioned items. An automated image-based 3D modeling pipeline for the accurate and detailed digitization of heritage artifacts is presented. The developed methodology, composed of open-source photogrammetric tools, is described with examples and accuracy analyses.

1.1 Why another open-source, automated, image-based 3D reconstruction methodology?

Today different image-based open-source approaches are available to automatically retrieve dense or sparse point cloud from a set of unoriented images (e.g. Bundler-PMVS, Microsoft Photosynth, Autodesk Photofly, ARC3D, etc.). They are primarily based on computer vision methods and allow the generation of 3D information even if the images are acquired by non-expert people with no ideas of photogrammetric network and 3D reconstruction. Thus the drawback is the general low reliability of the procedure and the lack of accuracy and metrics in the final results, being useful primarily for visualization, image-based rendering or LBS applications. On the other hand the authors are developing a photogrammetric web-based open-source pipeline, based on solid principles and guidelines, in order to derive precise and reliable 3D reconstructions useful for metric purposes in different application context and according to several representation needs.

2. SURVEYING AND 3D MODELING

Nowadays there are a great number of sensors and data available for digital recording and mapping of visual Cultural Heritage [10]. Reality-based 3D surveying and modeling is meant as the digital recording and 3D reconstruction of visual and existing scenes using active sensors and range data [6], passive sensors and image data [14], classical surveying (e.g. total stations or GNSS), 2D maps [15] or an integration of the aforementioned methods. The choice or integration depends on the required accuracy, object dimensions, location constraints, instrument’s portability and usability, surface characteristics, working team experience, project budget, and final goal of the survey and so on. Optical range sensors like pulsed (Time-Of-Flight), phase-shift and triangulation-based (light sheet or pattern projection) instruments have received in the last years a great attention, also from non-experts, for 3D surveying and modeling purposes. Range sensors directly record the 3D geometry of surfaces, producing quantitative 3D digital representations (point clouds or range maps) in a given field of view with a defined measurement uncertainty. Range sensors are getting quite common in the surveying community and heritage field, despite their high costs, weight and the usual lack of good texture. There is often a misuse of such sensors simply because they deliver immediately 3D point clouds neglecting the huge amount of work to be done in post-processing in order to produce a geometrically detailed and textured 3D polygonal model. On the other hand, passive optical sensors (like digital cameras) provide for image data which require a mathematical formulation to transform the 2D image features into 3D information. At least two images are generally required and 3D data can be derived using perspective or projective geometry formulations [14][17]. Image-based modeling techniques, mainly photogrammetry and computer vision, are generally preferred in case of lost objects, simple monuments or architectures with regular geometric shapes, small objects with free-form shape, point-based deformation analyses, low budget terrestrial projects, good experience of the working team and time or location constraints for the data acquisition.

2.2 Standards and best practice for 3D modeling issues

Best practices and guidelines are fundamental for executing a project according to the specifications of the customer or for commissioning a surveying and 3D modeling project. The right understanding of technique performances, advantages and disadvantages ensure the achievement of satisfactory results. Many users are approaching the new surveying and 3D modeling methodologies while other not really familiar with them require clear statements and information about an optical 3D measurement system before investing. Thus technical standards, like those available for the traditional surveying or CMM field, must be created and adopted, in particular by all vendors of 3D recording instruments. Indeed most of the specifications of commercial sensors contain parameters internally defined by the companies. Apart from standards, comparative data and best practices are also needed, to show not only advantages but also limitations of systems and software. As clearly stated in [18], best practices help to increase the chances of success of a given project.

3. AUTOMATED IMAGE-BASED 3D RECONSTRUCTION

The developed photogrammetric methodology for scene recording and 3D reconstruction is presented in detail in the next sections. The pipeline consists of automated tie point extraction, bundle adjustment for camera parameters derivation, dense image matching for surface reconstruction and orthoimages generation. The single steps of the 3D reconstruction pipeline have been investigated in different researches with impressive results in terms of automated markerless image orientation [19]-[21] and dense image matching [22]-[25].
3.1 Camera calibration protocol and image acquisition

The pipeline is primarily focused on terrestrial applications, therefore on the acquisition and processing of terrestrial convergent images of architectural scenes and heritage artifacts. The analysis of the site context relates to the lighting conditions and the presence of obstructions (obstacles, vegetation, moving objects, urban traffic, etc.). The former influences the shooting strategy and the exposure values, the latter is essential in order to perform multiple acquisitions and to select the correct camera focal length. The choice of focal lengths has a direct influence on the number of acquisitions and the resolution of the final point cloud, therefore it is always better to know in advance the final geometric resolution needed for the final 3D product. The employed digital camera must be preferably calibrated in advance following the basic photogrammetric rules in order to compute precise and reliable interior parameters [26]. Although the developed algorithms and methodology can perform self-calibration (i.e. on-the-field camera calibration), it is always better to accurately calibrate the camera using a 3D object / scene (e.g. lab testfield or building’s corner) following the basic photogrammetric rules: a dozen of convergent images at different distances from the object, with orthogonal roll angles and covering the entire image format. Each objective and focal length employed in the field must be calibrated. The images acquired for the 3D reconstruction should have an overlap around 80% in order to ensure the automatic detection of tie points for the image orientation. The shooting configuration can be convergent, parallel or divergent. Convergent images ensure to acquire possible hidden details. If a detail is seen in at two images, then it can be reconstructed in 3D. It is also important to keep a reasonable base-to-depth (B/D) ratio: too small baselines guarantee more success in the automatic tie point’s extraction procedure but strongly decrease the accuracy of the final reconstruction. The number of images necessary for the entire survey depends essentially on the dimensions, shape and morphology of the studied scene and the employed focal length (for interiors fish-eye lenses are appropriate). Figure 1 illustrates the possible acquisition schemas according to three different contexts: external, internal, façade.

3.2 Image triangulation

For the orientation of a set of terrestrial images, the method relies on the open-source APERO software [27]. As APERO is targeted for a wide range of images and applications, it requires some input parameters to give to the user a fine control on all the initialization and minimization steps of the orientation procedure. APERO is constituted of different modules for tie point extraction, initial solution computation, bundle adjustment for relative and absolute orientation. If available, external information like GNSS/INS observations of the camera perspective centers, GCPs coordinates, known distances and planes can be imported and included in the adjustment. APERO can also be used for camera self-calibration, employing the classical Brown’s parameters or a fish-eye lens camera model. Indeed, although strongly suggested to use previously calibrated cameras, non-expert users may not have accurate interior parameters which can therefore be determined on-the-field. The typical output of APERO is an XML file for each image with the recovered camera poses.

3.3 Surface measurement with automated multi-image matching

Once the camera poses are estimated, a dense point cloud is extracted using the open-source MicMac software [28]. MicMac was initially developed to match aerial images and then adapted to convergent terrestrial images. The matching has a multi-scale, multi-resolution, pyramidal approach (Figure 2) and derives a dense point cloud using an energy minimization function. The pyramidal approach speeds up the processing time and assures that the matched points extracted in each level are similar. The user selects a subset of “master” images for the correlation procedure. Then for each hypothetic 3D points, a patch in the master image is identified, projected in all the neighborhood images and a global similarity is derived. Finally an energy minimization approach, similar to [22] is applied to enforce surface regularities and avoid undesirable jumps.
3.4 Point cloud generation

Starting from the derived camera poses and multi-stereo correlation results, depth maps are converted into metric 3D point clouds (Figure 3). This conversion is based on a projection in object space of each pixel of the master image according to the image orientation parameters and the associated depth values. For each 3D point a RGB attribute from the master image is assigned.

3.4 Orthoimage generation

Due to the high density of the produced point clouds, the orthoimage generation is simply based on an orthographic projection of the results. The final image resolution is calculated according to the 3D point cloud density (near to the initial image footprint). Several point clouds (related to several master images) are seamless assembled in order to produce a complete orthoimage of surveyed scene (Figure 4).
3.5 Informatics implementation and GUI

APERO and Mic-Mac can be used as stand-alone programs in a Linux OS shell. The algorithms are also available with an end-user GUI with dedicated context interfaces:

- a general interface for the APERO-MicMac chain, developed at the IGN;
- a specific interface for the entire 3D reconstruction, integrated into NUBES Forma (Maya plug-in) [11], developed at the CNRS MAP-Gamsau Laboratory (Figure 5a). Starting from the automatic processing results, this application allows to:
  - collect 3D coordinates and distances;
  - generate dense 3D point clouds on demand (globally or locally);
  - extract relevant profiles by monoplotting (rectified image / point cloud);
  - reconstruct 3D architectural elements using interactive modeling procedures;
  - extract from the oriented images and project onto the 3D data high-quality textures.
- a web-viewer for image-based 3D navigation and point clouds visualization, developed at the CNRS MAP-Gamsau Laboratory (Figure 5b) for the visualization of the image-based 3D reconstructions produced with APERO/MicMac procedures [29]. The viewer allows to jump between the different image points of views, back-projecting the point clouds onto the images. It consist on a simple PHP-based web site (that user can publish on his/her own server) containing a folder for the 2D content (images), a folder for the 3D content (point clouds, polygons, curves) and a table with the camera parameters.

![GUIs integrated into NUBES Forma (Maya plug-in) developed at CNRS MAP-Gamsau Laboratory for image-based 3D reconstruction (a) and a web-viewer for image-based 3D navigation and point clouds visualization (b).](image)

Figure 5: GUIs integrated into NUBES Forma (Maya plug-in) developed at CNRS MAP-Gamsau Laboratory for image-based 3D reconstruction (a) and a web-viewer for image-based 3D navigation and point clouds visualization (b).

4. THE TAPENADE PROJECT

In order to give access to the realized procedure and software to a large number of users requiring metric 3D results (architects, archaeologist, conservators, etc.), the TAPEnAdE project (Tools and Acquisition Protocols for Enhancing the Artifact Documentation) [30] was started. This project aims to develop and distribute free solutions (software, methodologies, guidelines, best practices, etc.) based on the developments mentioned in the previous sections and useful in different application contexts (architecture, excavations, museum collections, heritage documentation, etc.) and according to several representation needs (2D technical documentation, 3D reconstruction, web visualization, etc.). The project would like to define acquisition and processing protocols following the large set of executed projects and the long-lasting experience of the authors in 3D modeling applications and architectural documentation. Examples, protocols and processing tools are available in the project web site.

5. CASE STUDIES

Figure 7 and Figure 8 present some examples related to different contexts (architectural exterior, building interiors, architectural element, archaeological excavation, museum object) and the relative 3D point clouds or orthoimages derived with the presented methodology. Detailed information on several other case studies is available on the TAPEnAdE web site [30].

5.1 Accuracy and performance evaluation

The results achieved with the 3D reconstruction pipeline described before were compared with some ground-truth data to check the metric accuracy of the derived 3D data. Figure 6 shows a geometric comparison and accuracy evaluation of the proposed methodology. A set of 8 images depicts a Maya relief (roughly 3×2 m) acquired with a calibrated Kodak
DSC Pro SRL/n (4500×3000 px) mounting a 35 mm lens. The ground-truth data were acquired with a Leica ScanStation 2 with a sampling step of 5 mm. The generated image-based point cloud was compared with the range-based one delivering a standard deviation of the differences between the two datasets of ca 5 mm.

Figure 6: Examples of the geometric comparison with ground-truth data. Original scene (left), derived 3D point cloud (center) and deviation map for a Maya bas-relief (std = ca 5 mm).

6. CONCLUSIONS

The article presented an open-source set of tools for accurate and detailed image-based 3D reconstruction and web-based visualization of the metric results. The image processing for 3D reconstruction is fully automated although some interaction is possible for geo-referencing, scaling and to check the quality of the results. The methodology is very flexible and powerful thanks to the photogrammetric algorithms. Different type of scenes can be reconstructed for different application contexts (architecture, excavations, museum collections, heritage site, etc.) and several representations can be delivered (2D technical documentation, 3D reconstruction, web visualization, etc.). The purpose of the developed method is to create a community with the aim of progressively enriches the performance and the relevance of the developed solutions by a collaborative process based on user feedbacks. If compared to other similar projects and products, TAPEnADE aims to deliver open-source tools which can be used not only online (web-based) and with highly reliable and precise performances and results.

7. ACKNOWLEDGMENTS

The results showed in this article are based on multiple contributions coming from students, engineers and young researchers. Authors want acknowledge in particular Isabelle Clery (IGN) and Aymeric Godet (IGN/MAP) for their contributions on informatics implementation; Nicolas Nony (MAP), Alexandre Van Dongen (MAP), Mauro Vincitore (MAP), Nicolas Martin-Beaumont (IGN/MAP) and Francesco Nex (FBK Trento) for their essential contributions on the protocols and case studies.

Figure 8: Examples of 3D reconstruction and orthoimages generation of complex architectural structures.
<table>
<thead>
<tr>
<th>Information</th>
<th>Object / Scene</th>
<th>3D reconstruction</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Architectural exterior</strong></td>
<td>![Image]</td>
<td>![Image]</td>
</tr>
<tr>
<td>- Dimensions: ca 15x20x20 m</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Number of images: ca 160</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Processing time: ca 3 hours</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Representation: 3D model/ortho</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Architectural interior</strong></td>
<td>![Image]</td>
<td>![Image]</td>
</tr>
<tr>
<td>- Dimensions: ca 20x30x15 m</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Number of images: ca 200</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Processing time: ca 4 hours</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Representation: 3D model</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Architectural elements</strong></td>
<td>![Image]</td>
<td>![Image]</td>
</tr>
<tr>
<td>- Dimensions: ca 30x150 cm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Number of images: ca 150</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Processing time: ca 3 hours</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Representation: 3D model</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Museum object</strong></td>
<td>![Image]</td>
<td>![Image]</td>
</tr>
<tr>
<td>- Dimensions: ca 15x5x20 cm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Number of images: ca 50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Processing time: ca 1.5 hours</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Representation: 3D model</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Building façade</strong></td>
<td>![Image]</td>
<td>![Image]</td>
</tr>
<tr>
<td>- Dimensions: ca 60x18 m</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Number of images: ca 40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Processing time: ca 2 hours</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Representation: 3D model/ortho</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Archaeological excavation</strong></td>
<td>![Image]</td>
<td>![Image]</td>
</tr>
<tr>
<td>- Dimensions: ca 15x2 m</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Number of images: ca 180</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Processing time: ca 4 hours</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Representation: 3D model/ortho</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 7: Examples of 3D metric reconstructions achieved with the presented open-source pipeline.
8. REFERENCES


