Protocols and assisted tools for effective image-based modeling of architectural elements

Nicolas Nony[1], Livio De Luca[1], Aymeric Godet[1,2], Marc Pierrot-Deseilligny[2], Fabio Remondino[3], Alexandre Van Dongen[1], Mauro Vincitore[1]

[1] UMR CNRS/MCC 3495 MAP-Gamsau ENSA Marseille, France
[2] MATIS, IGN Paris, France
[3] 3D Optical Metrology (3DOM), Bruno Kessler Foundation (FBK), Trento, Italy

Abstract. Domains related to documentation, conservation, restoration and enhancement of heritage objects use more and more digital 3D surveying. Image-based methods, due to their low-cost and accessibility, constitute an interesting alternative to traditional laser scanning surveys. However, these tools, often based on computer vision methods, are either difficult to use or not sufficiently accurate for scientific purposes. Within this context, this paper presents a photogrammetric pipeline in order to prove every heritage agent with a suitable and efficient solution for digitizing architectural and archaeological elements, taking into account the entire process, from the acquisition to the 3D result.

Keywords: photogrammetry, acquisition protocols, case studies, 3D point cloud

1 Introduction

Nowadays it's clear that the reality-based 3D surveying and modeling of heritage sites and objects pose a number of problems related to their execution, in particular if the survey is not carried out by experts or by the person who needs the final 3D results. In many cases, managing a survey requires considering technical factors peculiar to specialists – equipment settings and configuration, lighting, onsite constraints, etc. – and others more related to the object under investigation and to the project's aims – accuracy, needs, final use, etc. Therefore the communication and collaboration between the two communities requires a common language that they both need to have. Yet, despite their common interest, their collaboration sometimes faces the gap that may exist between the user's needs and the technical difficulties peculiar of surveying methods: requirements in the two domains are not necessarily formalized in the same terms. Therefore, it's often reasonable and common that the person carrying out a 3D survey is the one who is going to use the 3D result as he/she is aware of the real necessities of the surveying operation. This is also demonstrated by the recent scientific publications where groups of archaeologists or geographers are dealing with surveying instruments and techniques in order to produce 3D models for their needs. Thus it appears that a good solution (hardware and software) is to allow non-expert people to take care of the entire surveying process as well as the realization of the different final supports (mainly 3D models, but also orthoimages, planes, cuts etc.). Surveying instruments like terrestrial laser scanners and digital cameras are getting very common and every day easier to use. Moreover, there are many freeware or low-cost tools available on the internet that can theoretically solve the 3D reconstruction problem from images. They are based on photogrammetric and computer vision methods and allow any user to automatically create a 3D model from photos with very few mouse clicks (e.g. Bundler-PMVS [1], 123D-Catch [2], etc.) However, since they are mostly designed for the general public and a minimal user interaction, they are mainly used to create nice-looking and low accuracy 3D results. Since the processes are fully automated and the acquisition protocols not precisely documented, users have very few control over the process, no statistical output of the process and the reliability is also very low. The result is mostly visual and is not workable in a scientific context, which requires high accuracy, precise metrics and dense 3D geometry.

MATIS Laboratory (IGN, France) developed a set of photogrammetric tools (PASTIS-APERO-MICMAC) which allows the creation of several 3D and metric results (depth maps, dense and color 3D point clouds and orthoimages) with also quality check parameters [3-6]. These results, because of their shape and quality, in terms of accuracy and resolution, are able to answer a large amount of questions asked by the heritage specialists: for instance, depth maps can be used for the analysis of engravings or eroded elements, precise orthoimages for studying building deterioration, metric point clouds to monitor and evaluate excavation of an archaeological area, etc. The developed packages are also open source and very flexible, which makes them massively usable by the whole scientific community, but also customizable for the non-experts. However, this accessibility is limited due to their complexity and therefore they are more suited to expert users with good technical and informatics knowledge. For these reasons, considering the potentialities of the developed tools and for the interest of the entire scientific and non-expert community, it seemed necessary to have a thorough reflection on the tools and propose something more useful for all the communities.

Therefore the TAPENADE project [4][7] (Tools and Acquisition Protocols for Enhancing Artifacts Documentation), developed as a collaboration between IGN (France), CNRS Marseille (France) and FBK Trento (Italy), aimed at realizing a tool chain available for every user at no cost and freely. This set of tools allow a user to realize the entire process, from the field acquisition to the creation of a 3D documentation suited to his/her needs. In practical terms, these tools are:

- specific acquisition protocols for each type of object (which equipment to use, how to take photos, etc.)
- a user-friendly interface implementing the different software for processing the images acquired during the field campaign.

2 Workflow

The whole processing is performed with the chain of Linux-based software PASTIS-APERO-MICMAC. It is a set of photogrammetric tools for automatic orientation and dense multi-image matching of digital images. The process consists of automated tie point extraction, bundle adjustment for camera parameters derivation, dense image matching for surface reconstruction and orthoimages generation.

2.1 Description of the pipeline

The entire image-based pipeline is divided into three main steps:

- tie-points extraction with PASTIS (Programme utilisant Autopano Sift pour les Tie Points dans les ImageS), based on the SIFT-match algorithm [8].
- internal and external camera parameter recovery with APERO (Aérotriangulation Photogrammétrique Expérimentale Relativement Opérationnelle) using a robust bundle adjustment [7];
- depth maps, point-clouds and orthoimages generation with MICMAC (Multi-Image Correspondance, Méthodes Automatiques de Corrélation) [5].

2.2 Technical matters

With respect to other existing open-source or web-based or low-cost software, the presented workflow consists of powerful, flexible and reliable packages derived from the photogrammetric community, therefore keeping always an eye to the accuracy of the final results and not only at the nice-looking achieved 3D data. The packages run from command lines with xml files containing all the input parameters (although some GUI are already existing or under development). The input file permits a very large amount of configurations and processing methods (e.g. fish-eye lens in the bundle adjustment, both convergent and nadir image during the matching, etc.) and the output data settings (such as point cloud density, DSM format, etc.) can be entirely defined and controlled by the user.

The other important matter is the various image formats taken into account, from low resolution JPG to un-compressed formats such as RAW or TIF. At last, there is no restriction about the pictures' resolution and size, or about the lens and cameras' characteristics as well as the dimension of the image block (from our experience, an UAV block of 1000+ images was processed).

3 Protocols and best practices

In order to ensure to the users a mostly automated process still providing optimum results, a series of specific acquisition protocols has been established. The protocols were created from a large amount of realized case studies, based on practical field works on different scenarios (indoor, outdoor) and using different digital cameras.

In particular, the experiments reported afterwards, focused on architectural objects which present different scales and problems: engravings, façades, entire buildings, stretches of streets, interiors, poor texture, illumination changes, etc. Different types of cameras (reflex or compact) and lenses (long and short focal lens, fisheyes) were employed as the final result quality depends on the quality of the pictures and therefore on the used equipment.

These experiences have been clearly summed up as case studies (Fig. 1) which describe and analyze the adopted methodology as well as the results. These case studies, available in the TAPENADE website, are more than a tutorial as they deeply describe the surveying problems and the developed solutions to achieve dense and accurate point clouds and orthoimages. Thanks to these experiences, some best practices are given for different scene/goal combinations.

Fig. 1. Several examples of case studies and protocols available on the TAPENADE website: 3D surveying of interiors, architectural elements, façades, etc.

3.1 Defining the 3D surveying and modeling needs

In every field work, the data acquisition must be preceded by a precise definition of the survey goals, i. e.

- which object are we surveying? Building, landscape, fragment, etc.?
- what will be the final output? Orthoimage, point cloud, depth map, etc.?
- which will be the final use of the 3D data? Accuracy, visualization, internet application, etc.?

These three elements are essential since they determine the modalities of data acquisition and processing.

3.1 Analyzing the scene

A knowledge of the scene (building, artifact, etc.) under investigation is mandatory to fully consider the conditions onsite. Indeed, it happens that local constraints do not allow a full achievement of the objectives. Therefore, a brief analysis of the scene and its environment in order to highlight the different constraints is necessary to understand the intrinsic constraints (complexity, morphology, material, dimensions, etc.) and external constraints (environment, occlusions, light, inaccessibility, etc.).

For example, illumination conditions and luminosity affect the speed and acquisition parameters of the camera and therefore whether a tripod is necessary or not (which slows down the image acquisition step.) The presence of obstacles implies a change in the acquisition location or the use different focal length. Also, homogeneous and transparent materials, reflections and windows cannot be correctly processed.

3.1 Image acquisition

The employed digital camera must be preferably calibrated in advanced following the basic photogrammetric rules in order to compute precise and reliable interior parameters [9]. Although APERO can perform self-calibration, it is always better to accurately calibrate the camera using a 3D object / scene (e.g. lab testfield or building's corner) with an ad-hoc camera network composed of a dozen of convergent images at different distances from the object, with orthogonal roll angles and covering the entire image format.

The shooting configuration can be convergent, parallel or divergent. It is also important to keep a reasonable base-to-depth (B/D) ratio: too small baselines guarantee more success in the automatic tie points extraction procedure but strongly decrease the accuracy of the final reconstruction [10]. 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. According to the principles described above and in case of complex monuments, possible images acquisition could be:

- acquire 3 to 5 images with a 80-90% overlap. These groups of images are called "correlation images" and are used to generate the dense point clouds. The overlap between each group has to be about 20%, in order to avoid holes in the final point cloud. These groups of images are connected between them thanks to a second series of images (called global orientation), covering the whole monument, with a 60-70% overlap (Fig. 2a).
- acquire a single sequence of images, covering the whole monument, with good overlap and B/D ratio. These images are used for both orientation and point cloud generation (Fig. 2b).
- same principle as above, adding e.g. some images with a longer focal length, covering the whole monument and with a 20% overlap. These images, oriented with the others, are projected on the depth map to generate orthoimages (Fig. 2c).

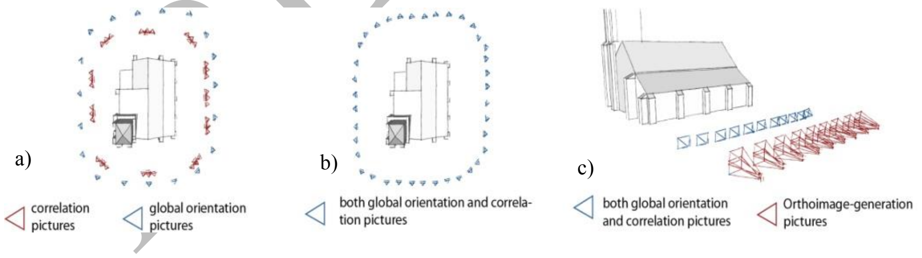

Fig. 2. Possible image networks for the surveying and modeling of a building.

4 A user-friendly interface

The acquisition step is only one part of the entire image-based 3D reconstruction pipeline and it requires some experience as described above. The successive image processing is also very crucial in particular if the user wants to have a control on the reconstruction phase.

As the developed programs run mainly from the Linux-shell and need some xml input files containing all the running parameters (tie points selection strategy, matching resolution, DSM output format, etc.), a graphical user interface (GUI) was created, in order to overcome these problems and provide for a more comfortable photogrammetric tool. The goal was to allow any user to easily and accurately manage a 3D reconstruction project by automating the configuration and execution procedures of the different processing step as much as possible, still keeping the control of the 3D reconstruction procedure.

4.1 General GUI description and data processing

An experimental interface was created as a Maya plug-in in MEL (Maya Embedded Language). Its functions follow the different stages of the data , allowing to handle all type of surveys, cameras, lens, etc., with a sufficient control on the various parameters. Each step of the processing has its own window, where the user can access to all the settings. The navigation through the images is made with lists and, for visibility's sake, the selected image is systematically displayed. Drawing tools for the matching masks have been created, together with a tool useful for scaling and GCPs import. From the command line the user can still check in real-time the state of progress of the processing.

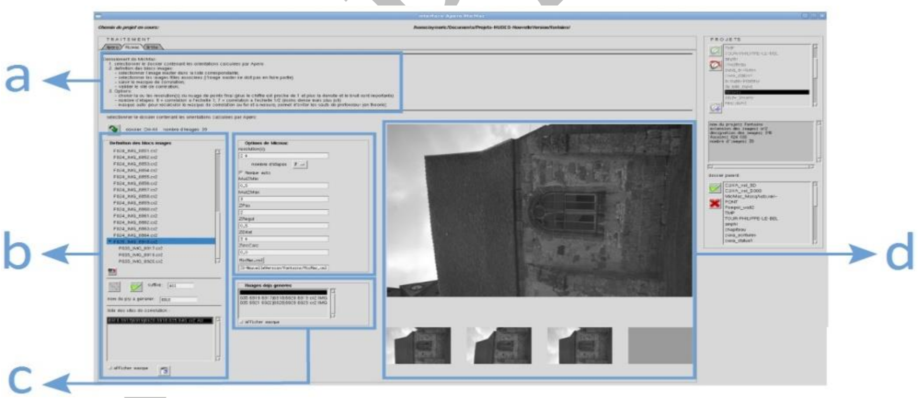

Fig. 3. GUI for the dense matching reconstruction step using MICMAC. (a): various explanations about point cloud generation; (b): the correlation tree, containing the list of all the available images; (c): the list of the previous point clouds generated within this project; (d): the current correlation images).

The first step consists in extracting tie points with TAPIOCA, a simplified version of PASTIS. This process is run with a Linux-shell, e.g.: `<Tapioca Line ".*NEF" 1500 5>` which means that each image is compared to the five previous and the 5 next images in order to find the homologues points.

The camera calibration can be done running TAPAS (e.g. <Tapas RadialExtended ".*NEF" OUT=Calib> which means the calibration is made with all the NEF images and the results is put in a folder named Calib) a simplified version of APERO [8]. TAPAS, starting from the already extracted tie points, takes into account several models of lens distortion, in particular Brown's and fisheye model. Finally, with APERO, a relative or absolute (with GCP's) orientation can be computed.

After the successful image orientation, MICMAC is used to derive dense point clouds or an orthoimage (Fig. 3). For the *point cloud generation* purpose, MICMAC considers the image block configuration (convergent or parallel images) and needs the user to define, in particular for convergent images, some reference images to start the multi-image matching procedure. The identification of these master images (together with matching masks to avoid possible blunders and non-desired matched areas) allows to better handle occlusions and complex scenarios, deriving color point clouds (Fig. 4a) or shaded surface models. For an *orthoimage generation*, the user can employ TARAMA, which allows to create a low resolution image of the whole scene. Then the matching mask is defined on a single image, containing the entire in which the orthoimage has to be computed (Fig.4.b).

a)

b)

Fig. 4. Two examples of the results achievable using the developed image-based pipeline and GUI: a 3D point cloud of the Fontain Church, France (processing time: ca 3 hours) and an orthoimage of a bas-relief on the Dendera temple, Egypt (processing time: ca 1 hour).

Finally, in order to be really accessible to the users, a new version of the GUI is under development in C++. Thus, the final GUI will be totally free and available on TAPENADE website.

5 Conclusions and future works

The article presented the development of a GUI for the photogrammetric open-source packages APERO-MICMAC. Different protocols and best practices have been presented (and more are available in the TAPENADE website) for the 3D reconstruction of facades, entire buildings, archeological fragments, etc. In order to complete this list, satisfy all the possible needs, solve the specific problems associated to the different fields and further develop the tools, we seek partnerships and feedbacks from various heritage specialists (restorators, preventive archeologists, museums curators, etc.). The purpose of the developed tools and GUI method is to create a kind of community with the aim of progressively enrich 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, the tools are not black-boxes, they allow the users to have a full control of the data processing, with quality check on the derived metric results. The drawback of other existing tools similar to APERO-MICMAC is the general low reliability of the procedure, the lack of accuracy and metrics and the final result, being them often a "one-bottom" tool. On the other hand, the authors presented the further developments of a photogrammetric 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.

References

1. M. Seitz, S., Snavely, N., Szeliski, R.: Photo Tourism: Exploring Photo Collections in 3D. In ACM Transactions on Graphics, SIGGRAPH (2006)
2. 123D-Catch website: http://www.123dapp.com/catch
3. Pierrot-Deseilligny, M., Clery, I.: Apero, an open source bundle adjustment software for automatic calibration and orientation of a set of images. Int. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 38(5/W16), Trento, Italy (2011)
4. Pierrot-Deseilligny, M., De Luca, L., Remondino, F.: Automated image-based procedures for accurate artifacts 3D modeling and orthoimage generation. Proc. CIPA (2011)
5. Pierrot-Deseilligny M., Paparoditis N.: A multiresolution and optimization-based image matching approach: An application to surface reconstruction from SPOT5-HRS stereo imagery. Int. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 36(1/W41) (2006)
6. Micmac 2012, Micmac Apero Documentation: http://www.micmac.ign.fr
7. TAPEnADe project (Tools and Acquisition Protocols for Enhancing Artifact Documentation) : http://www.tapenade.gamsau.archi.fr
8. Pierrot-Deseilligny, M., Clery, I.: Some possible protocols of acquisition for optimal use of the "APERO" open source software in automatic orientation and calibration. Proc. EuroCOW-2012 Conference, Spain (2012)
9. Remondino, F., Fraser, C.: Digital camera calibration methods: considerations and comparisons. Int. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 36(5), 266-272 (2006)
10. Voltolini, F., Remondino, F., Pontin, M., Gonzo, L.: Experiences and considerations in image-based-modeling of complex architectures. Int. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 36(5), pp. 309-314 (2006)