

AN ERGONOMIC INTERFACE TO COMPUTE 3D MODELS USING PHOTOGRAMMETRY

Isabelle CLÉRY¹, Marc PIERROT-DESSEILLIGNY^{1,2}

¹Laboratoire MATIS, Institut Géographique National, Université Paris-Est
73, avenue de Paris, 94165 Saint Mandé, France
isabelle.clery@ign.fr

²Laboratoire Informatique Paris Descartes, équipe SIP
45 rue des Saints Pères, 75270 Paris Cedex 06, France
marc.pierrot-deseilligny@ign.fr

Keywords : photogrammetry, GUI, dense 3D modeling, professional applications

Abstract:

Many commercial softwares offer to create 3D models of objects or scenes from photographs only. However professionals, including archaeologists and architects, have numerous constraints that severely restrict the choice of such software, such as limited financial resources, little knowledge in photogrammetry and computing, professional imperatives that limit the time of acquisition and processing (particularly manual treatments), and finally specific expectations in terms of quality, precision and completeness of the modeling. The MATIS laboratory has developed a set of free, open source tools, which can perform all stages of photogrammetric projects processing and that provide dense and precise 3D models thanks to the rigorous photogrammetric approach. A graphical interface, which will be presented in this paper, is being developed to provide a unified and simplified access to these tools. Through the interface, the user can manipulate the images of the same site and control every step of processing that includes : the automatic calculation of tie points between images, the estimation of camera poses and the 3D models calculation. The results can be converted into depth maps or dense 3D points clouds (in standard ply format), and can be displayed directly by the GUI. Particular attention is paid to the quality of documentation and to the product portability on the main types of operating systems (Linux, Mac and Windows). This interface, although not finalized, already enables to perform the full treatment of certain types of surveys and is available on-line.

1. INTRODUCTION

Many solutions for 3D reconstruction from images have appeared encouraged by a strong demand from many areas such as archeology, geology, architecture, artwork expertise..., which are generally remote from photogrammetry and computer science. Indeed, 3D scanning from images is a complementary technique of laser, but cheaper, more convenient and discrete (thus better suited to sites with difficult access), of adjustable precision and especially that renders the scene colors. Among these programs we can name : Microsoft's Bundler, CMVS (PMVS2, [3]) usually used with Bundler, Photosynth based upon Bundler ([1]), AutoDesk's ImageModeler, Eos Systems' PhotoModeler and Arc3D ([2]). But the choice of such software is strongly constrained by the available resources and the applications targeted. In particular, free and ergonomic supply targeting non-photogrameters users for scientific applications is very small. In this context, the MATIS laboratory has made available a set of open-source tools for performing the entire processing of photogrammetric projects and an ergonomic GUI, that is currently under development, is presented here. After introducing the project context and the goal of the interface, the main features of the interface will be detailed, followed by some results obtained in different areas and finally the projected evolutions will be exposed.

2. CONTEXT

2.1 The laboratory tools

MATIS laboratory already has a set of tools that perform each step of the photogrammetric processing ([7]). The main tools used successively during the treatment, are presented below.

2.1.1 *Pastis*

The Pastis software (it means “Program using Autopano Sift for the Tie-points in the images”) uses the SIFT algorithm (Scale-invariant feature transform, [5]) to automatically find homologous points on different pairs of partially overlapping images. For a pair of images, SIFT first looks for points of interest on each image (they are areas of high gradient : edges of objects, tasks ...). Then it matches each found point of the first image with a point of the second image using to the likeness of their neighborhood. These tie-points have the advantage to be invariant under change of contrast (except saturation), scaling, image translation and low camera rotation, which lets the photographer the freedom to change brightness when shooting and to take pictures in different directions, but they are sensitive to blur and scene modifications (like changes in shadows, moving trees ...). Pastis runs the SIFT algorithm on several image pairs in parallel to accelerate the calculation and can handle very weighted images by dividing them up among blocks.

2.1.1 *Apero*

The *Apero* software (acronym for “Relatively Operational Experimental Photogrammetric Aerotriangulation”, [6]) computes the parameters of the poses of the camera (or aerotriangulation, including perspective summit and orientation calculation, and re-estimation of calibrations) from a set of tie points and approximate objective calibrations.

Apero uses a photogrammetric-only approach which produces a very precise aerotriangulation. More specifically, on the one hand, it uses exact equations to calculate the actual physical parameters (perspective summit and orientation, ...), contrary to the approach of computer vision which considers theoretical mathematical quantities (fundamental matrix) taking into account more parameters than necessary, that amounts to re-estimate the focal length for each image and then can cause a drift of aerial triangulation and thus imprecision of the final result. On the other hand, the photogrammetric approach allows to include in computing the re-estimation of complex objectives calibration (including fish-eye).

The calculation of the relative orientations (poses are relative to one “master” pose) and calibrations reestimation consists of two parts:

- from a “master image” (its pose is located at the origin of the reference and its orientation is the identity), an initial solution is calculated recursively one image by one ; the orientation of one image is given by the best solution obtained by both calculating the essential matrix followed by the RANSAC algorithm on the one hand, and the method of space resection followed by RANSAC on the other hand. Regularly, *Apero* makes a bundle adjustment on the already oriented images to avoid drifts of the model;
- the calculation of all parameters is processed by iterative least squares minimization of the sum on all tie-points of image errors of the previous solution. An image error for a point is function of images orientations and calibrations and is estimated as the difference between the tie-point coordinates in an image and the reprojection in this image of this tie-point after being thrown to the ground by beam intersection.

It is possible to include in this calculation:

- absolute georeferencing using GPS, both INS and GPS, or a global orientation (computed from a horizontal plane, an abscissa axis and a scale),

- initial relative orientation, which allows processing by blocks of images,
- varied types of calibrations (in particular, equidistant or equisolid fisheye models)

2.1.1 MicMac

MICMAC (acronym for “Multi Image Matches for Auto Correlation Methods”, [4]) is a software mapping images developed by Marc Pierrot-Desseilligny. MICMAC calculates among others depth maps from a set of tif images, their poses estimation and a depth interval (which corresponds to extreme values of the scene depth). Depth maps are images in which each pixel represents the depth of the corresponding 3D point seen from the point of view of an image called the “master image” (which has nothing to do with the one used for poses estimation) ; they can be directly converted to 3D points clouds. The principle of calculation is as follows :

- The calculation is done in image geometry (this amounts to generalize the epipolar geometry to multi-images) ; among the different geometries proposed by MicMac, it has the advantage that the reference image is strictly superimposable to the 3D model calculated, in particular the “master image” is a natural solution to allow the user to add information. The explored scene space is discretized to a one pixel-equivalent step in planimetry.
- MICMAC calculates for each point in this space a measure of likelihood (which is the sum of correlation coefficients for each pair of images).
- Finally, the surface is extracted by an energetic approach to minimize on the whole considered space a sum made up of a likelihood measure and a smoothing term in order to homogenize the surface.

This calculation is performed iteratively on sub-sampled images at increasing resolutions ; the result obtained in the previous resolution is used to predict the current step solution by defining a search area. This multi-scale/multi-resolution aspect leads to both a better robustness and a shorter execution time.

In addition, the correlation coefficient is invariant under change of contrast, so the result does not depend on the scene lighting.

MICMAC provides a map of correlation coefficients corresponding to the chosen depth for each pixel of the depth map to quantify its accuracy. MICMAC is very flexible and many other results can be provided: MNT calculation given in a reference map (it was originally developed for this purpose), calculation of parallax...

2.2. The interface project

MATIS laboratory wishes to make these tools available to the largest public in order to value them and at the same time answer the public service mission of IGN. The problem is that these tools are called with command line and are designed to treat a wide generality of sites, so they require a large number of parameters that have to be manually inserted into files in a predefined format. This complexity makes these tools hardly accessible to non specialists in photogrammetry and computer sciences. That is why all these tools have been gathered together in a GUI to allow any non-expert user to easily perform the entire process ([8]). Like the lab tools, this interface is free and open source in order to reach all categories of users interested in the potential of tools but discouraged by the complexity of their use.

3D models produced by these tools are very dense, need a shooting that follows a list of constraints and requires different parameters to ensure the desired accuracy. That's why the targeted public is mainly composed of professionals (photogrameters, archeologists, architects, geologists...), but the general public, who wish to occasionally get a very dense 3D model, is not excluded. This professional public has the specificity of having constraints associated with their job (economic, material, limited or even no Internet access while fieldworking), has a very specific expertise in its field, but generally limited knowledge in photogrammetry and has clear expectations for result rendering, related to a scientific inquiry or the imperatives of cultural heritage management. In particular, these users can accept to follow some simple

rules of photogrammetry during the shooting in order to significantly improve the quality of the model. The tools offered by the MATIS have therefore the advantage of being free, requiring only low-cost data (images from a camera), and producing dense, accurate models with a scalable resolution (the model resolution directly depend on the images resolution).

The GUI is designed to be ergonomic, easy to install, adapted to the generality of the data and multi-platform. To simplify data input, acquisitions are divided into shooting type (convergent, parallel, divergent...) that allows to set up default parameters ; remaining parameters are entered graphically (in form, by drawing or with 3D selection) thanks to Qt library, a cross-platform GUI library written in C++. The computation is processed step by step, each step corresponding to a previously presented tool which is called by the GUI. The grouping of these tools eliminates redundancy in configuration (images used, filenames reused for input-output). A window, which is designed with the OpenGL library for graphics rendering, enables to view the results in 3D (to control estimated poses or to display point clouds). Finally a full documentation is accessible both on-line and aside.

Although the interface is still under development, it already enables to perform the entire processing of a convergent shooting on Linux 32 bits and 64 bits and on Mac.

2. OVERVIEW OF THE GUI

At the beginnig, the user loads all his images ; acceptable formats are tif(f), all raw and jpg. The GUI main window is represented in figure 1.

2.2. Tie-points search (Pastis)

A window enables the user to enter Pastis parameters :

- the decomposition of images names to automate the process (a default one is provided)
- initial calibration parameters for each objective used (at least the focal length and the sensor size are required),
- the shooting type (convergent, parallel...)
- the images pairs where to do the search ; they are selected in a list
- the size to subsample images in order to accelerate the processing.

The interface links each image to its calibration, converts them to tif format, and then starts Pastis.

2.3. Poses estimation (Apero)

After tie-points computation, the user enter Apero parameters :

- the images to orient ; it enables to restart computation on a smaller set if all-image computation fails
- the Apero “master” image
- the global absolute orientation ; it can be either only relative, or given by an horizontal plan (defined by a mask), an abscissa axis (defined by a segment) and a scale (defined by a segment drawn in stereoscopy and its measure), or given by the absolute orientation of an image and a scale (this enables to link different projects) ; all these parameters are drawn in the interface (figure 2)
- if occurred an images set to reestimate a calibration by auto-calibration before poses estimation, in order to refine the initial calibrations supplied by the user
- whether computing by block of images : short focal length images are oriented first, then the global triangulation is done from the previous result ; this avoid drift in poses estimation
- whether computing a different calibration per image (when focus has been changed during shooting)
- whether filtering tie-points to accelerate computing and 3D rendering (if many were found)

The poses estimated and tie-points (converted to 3D points) can be displayed in a 3D viewing window (figure 3).

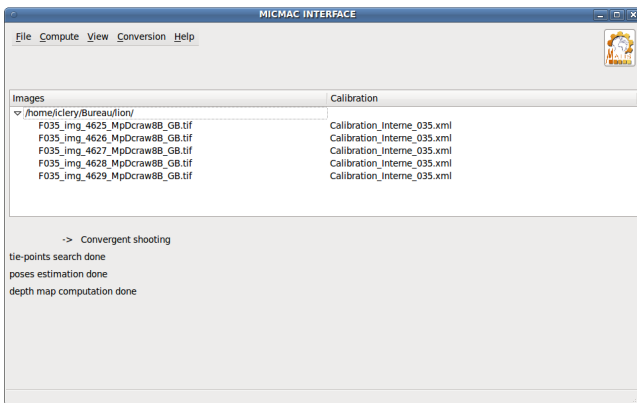


Figure 1: GUI main window

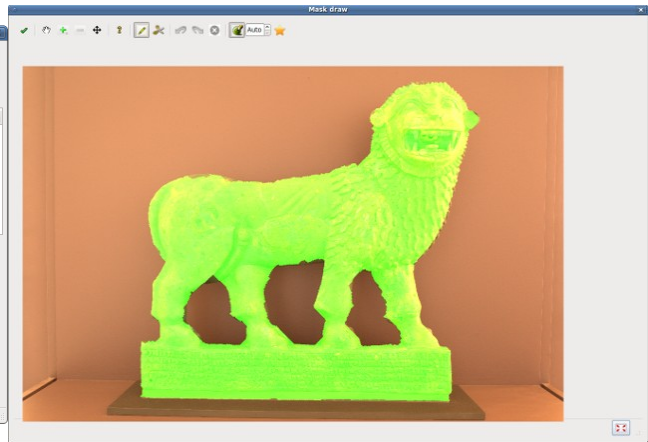


Figure 2: GUI drawing window
It is used with to draw masks and segments for Apero and MicMac steps

2.4. Depth maps and 3D models computation (MicMac)

Usually the scene has a quite complex relief and has to be represented by several point clouds made from depth map of different points of view, as a laser scanner would do. For each depth map to compute, the user enter the depth map parameters :

- the MicMac “master image”
- a mask of the scene to model in 3D, which can be drawn in the GUI (figure 2)
- a set of images used for the correlation step to retrieve the depth information ; all required images can be chosen in the 3D view of the GUI (figure 3)

The GUI calls MicMac iteratively to compute all depth maps at the same time.

It is also possible to compute a depth map in Euclidian geometry, which better suits to parallel shooting like building facades. When the “master image” and the set of images used for the correlation are selected, an assemblage table is calculated (it is a georeferenced mosaic of the previous images) so that the mask can be drawn on this assemblage table.

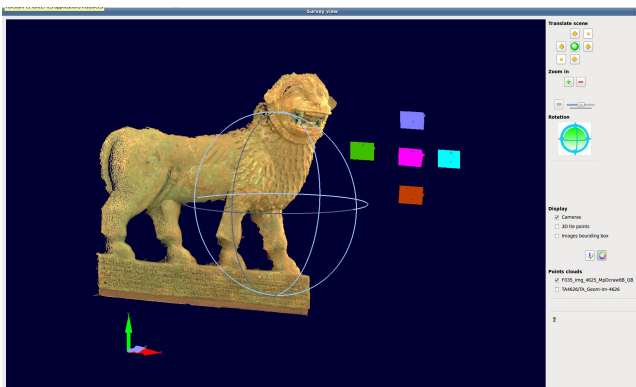


Figure 3: GUI 3D view window
It displays camera poses, 3D tie-points and MICMAC points clouds



Figure 4: Shaded representation of a depth map
Lighting can be configured and hypsometric colors can be added

2.5. Products obtained

The user can either convert the obtained depth maps to 3D points clouds in the normalized format ply (readable by MeshLab for example) and that are textured with the master image, or to a shadowed representation in tif format (figure 4), or he can display it as points clouds directly in the GUI (figure 3).

The shadowed representation provides a simple and intuitive visualization of depth information and usefully replaces a points cloud in absence of 3D display software, for non-dynamic presentations or publications (for which a 2D screen-shot of a 3D cloud is not interesting) or to transmit results in a light format.

3. SCENE MODELLING

3.1. Acquisition constraints

3D model computing requires image acquisition to verify photogrammetry usual constraints : images overlapping to ensure the presence of tie-points and so poses computing, stereoscopy (duplicate information from at least two different points of view) for “master images” of depth maps and an on the whole fixed scene (and no diachronism) to avoid too many mismatches resulting from Pastis. Lighting is not a problem as the used algorithms (SIFT, correlation coefficient) are invariant by lighting changes, but its consideration improves aesthetically the model texture rendering.

Some scenes give anyway poor results regardless of previous considerations. This is the case of badly textured objects like monochrome plaster, or of objects with specular reflection, like metallic objects or bronze. This is because MicMac produces dense models : each pixel of the “master image” gives one 3D point using a local texture comparator (the correlation coefficient), contrary to tools that process only highly contrasted pixels and interpolate the 3D model between them.; MicMac has then the advantage of being more accurate on well textured materials such as stones, earth, plants...).

3.2. Some results

The GUI has been tested with various scenes of different kind, texture and from different application fields. Some of example are presented above.

The first example is a stone elephant statue (China). Four depth maps (one from each side of the elephant) have been computed to model this statue. The 3D model presented above, like every following 3D model, is made of textured 3D points only (there is no triangulation).



Figure 5: Modelisation of the elephant
on the left : a picture of the right side of the elephant
in the middle : a shadowed representation of the corresponding depth map
on the right : the 3D model of the elephant

The second example is *La Cathedrale*, a stone sculpture of Rodin exposed in the Rodin museum (Paris, France). The shooting was taken as experiment of serialization of artwork modelling using light equipment (without laser scanner) and a simplified, standardize procedure to answer to the French Ministry of Culture will of monuments and cultural patrimony 3D digitisation. Because this artwork has a very complex form, the shooting is made of eight different depth maps.



Figure 6: Modelisation of *La Cathedrale*
on the left : a picture

in the middle : a shadowed representation of the corresponding depth map

on the right : the 3D model of *La Cathedrale*

The third example is a chair from *Le Petit Trianon* in Versailles, France. To help correlation on the monochrome white wood, a synthetic textured image has been projected on the chair with a video projector during the shooting. Then a picture was taken from the exactly same point of view than the “master image” without projecting synthetic texture in order to color the 3D model after depth map computation.



Figure 7: Modelisation of a chair
on the left : a picture

in the middle : a shadowed representation of the corresponding depth map

on the right : the 3D model of the chair

4. PESPECTIVES

Some users have already tested the interface, not many because the interface is still in development and not finished. Testers cover the fields of photogrammetry, archeology, architecture and geology. That allowed to improve the GUI in terms of user-friendliness, hardware compatibility and functionalities.

Several extensions are planned to better meet the needs of the public:

- Merge the set of 3D point clouds obtained from the different depth maps of a single project to harmonize the clouds, have a more fluid display of the 3D model and filter the noisy points,
- Calculate an unrolled depth map from parameters drawn in the 3D view (useful for conical or cylindrical objects like columns)
- And many others depending on demand ...

CONCLUSION

It was presented here a multi-platform graphical user interface for a set of free tools developed by the laboratory of IGN MATIS, that allows to process the entire photogrammetric chain from raw pictures to a textured 3D model and that is accessible to non-photogrammeters users. This interface brings together tools that search tie-points in image pairs, estimate the camera poses, compute depth maps and convert them into 3D points clouds. The density and accuracy of the obtained 3D models are suitable for professional applications such as archeology and architecture. The interface offers a simplified and essentially graphical setup, 3D visualization of the results at various stages of calculation and is provided with full documentation. The interface is still in development but is already downloadable in its beta version ([8]) and we are looking for testers to improve its functionalities depending on the desired applications.

REFERENCES

- [1] Snavely, N., Seitz, S., : Szeliski, R.: *Modeling the World from Internet Photo Collections*, International Journal of Computer Vision, 2007.
- [2] Vergauwen, M., Van Gool, L.: *Web-Based 3D Reconstruction Service*, Machine Vision Applications, 17, pp. 411-426, 2006.
- [3] Furukawa, Y., Ponce, J.: *Dense, and Robust Multi-View Stereopsis*, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, July 2007.
- [4] Pierrot-Desseilligny, M.: *MicMac, un logiciel pour la mise en correspondance automatique d'images dans le contexte géographique*, Bulletin d'Information Scientifique et Technique de l'IGN n° 77 (B1).
- [5] Vedaldi, A.: *An open implementation of the SIFT detector and descriptor*, UCLA CSD Tech. Report 070012, 2006
- [6] Pierrot-Desseilligny, M., Cléry, I.: *Apero, an open source bundle adjustment software for automatic calibration and orientation of set of images*, Proceedings of the ISPRS Commission V Symposium, Image Engineering and Vision Metrology, Trento, Italy, 2-4 March 2011
- [7] Pierrot-Desseilligny, M.: *MicMac, Aperro, Pastis and other beverage. The doc !*, <http://www.micmac.ign.fr/svn/micmac/trunk/Documentation/DocMicMac/DocMicMac.pdf>, 2011/03/21
- [8] Cléry, I. : *InterfaceMicmac, Documentation utilisateur*, http://www.micmac.ign.fr/svn/micmac_data/trunk/DocInterface/doc.pdf, 2011/04/15