APERO, AN OPEN SOURCE BUNDLE ADJUSTMENT SOFTWARE FOR AUTOMATIC CALIBRATION AND ORIENTATION OF SET OF IMAGES.

M. Pierrot Deseilligny\textsuperscript{a,b}, I. Clery\textsuperscript{a}
\textsuperscript{a} Institut Géographique National, Direction technique, laboratoire Matis, 94165 Saint Mandé
\textsuperscript{b} Laboratoire Informatique Paris Descartes, équipe SIP
marc.pierrot-deseilligny@ign.fr
Commission V, WG V/4

KEY WORDS: Photogrammetry, bundle adjustment, 3D modelization, open-source, cultural heritage.

ABSTRACT:
IGN has developed a set of photogrammetric tools, APERO and MICMAC, for computing 3D models from set of images. This software, developed initially for its internal needs are now delivered as open source code. This paper focuses on the presentation of APERO the orientation software. Compared to some other free software initiatives, it is probably more complex but also more complete, its targeted user is rather professionals (architects, archaeologist, geomorphologist) than people. APERO uses both computer vision approach for estimation of initial solution and photogrammetry for a rigorous compensation of the total error; it has a large library of parametric model of distortion allowing a precise modelization of all the kind of pinhole camera we know, including several model of fish-eye; there is also several tools for geo-referencing the result. The results are illustrated on various application, including the data-set of 3D-Arch workshop.

1. INTRODUCTION

1.1 Context
The tremendous development of cheap high quality digital camera and computational power of personnel computers has led, in the past decade, to a very active community of research in photogrammetry and computer vision. According, to the current literature, the “old cartographic dream” of modeling the word in 3D at scale one using only photo seems now almost affordable, at least from the algorithmic point of view...
In fact, despite the impressive results obtained by different teams, and the success of tools like (photosynth 2010) for touristic application, the use of 3D photomodelization by professional users, like architects, archaeologist, geomorphologists, as a cheap and light tool for measurement is still quite undeveloped compared to the potential of these technologies.
In this context, Institut Géographique National, (IGN, the French mapping agency), has decided in 2007 to delivers as open-source several photogrammetric software internally developed at the MATIS laboratory. The two main tools currently distributed are APERO, a program for computing orientation of images, and MICMAC a program for computing depth maps from oriented images. This paper presents the main characteristics of APERO and illustrates the results with different 3D models computed by MICMAC with the oriented image.

1.2 State of the art, algorithmic part.
Until the 2000 year, the photogrammetric and computer vision community have worked almost independently on the problem of 3D reconstruction out of stereo image acquisition (Forstner 2009). The photogrammetrists were working on well calibrated analogical (eventually scanned) images for metrological applications, using bundle adjustment in Euclidean space, in process involving a lot of human interaction (Kasser 2002).

During the same time, the computer vision community were working on low resolution images for real times application (without interaction) focussing on the relative position of objects, using mainly projective space (Faugeras 1994).
Since the begun of 2000 years, we have assisted to the emergence of high quality, relatively cheap, digital camera. In 2010, for 2000 Euros, one can affords a reflex camera with quite good photogrammetric performance and image quality (in term of resolution and SNR ratio) comparable to the first operational aerial digital camera that were used in years 2003 (Souchon 2010). This, naturally, has led to the idea that, taking the best of the both of photogrammetry and computer vision, it should be possible to have tools, fully automatic and precise, allowing 3D modelization of complex environment from image acquired by amateur camera.
Starting from the sate of art of the 2000 years, several drastic algorithmic improvements were required before achieving the objective of automatic modelization from image. Briefly speaking, three main challenges had to be, and in a great part have been, resolved:
• First of all, given two image recovering partially, the automatic detection of reliable and sufficiently dense tie points between both images had to be solved; this point has probably been the most important source of progress and, at the beginning of the last decade, tools like SIFT (Lowe 2004) or MSER (MATAS 2002) have emerged who have dramatically change our possibility;
• Second step, the automatic orientation of large block of images using only of tie points had to be solved; although for this step (almost) all the algorithmic component where known from long time, they had to be merged in reliable software; solution like bundler (snavely 2008) are now affordable;
• Last point, the dense automatic matching of oriented images had to be greatly improved; the possibility of using multi correlation techniques, because taking digital image...
is almost free of charge, added to a very dynamic community using sophisticated optimisation tools, as (Labatut 2007) or (Furukawa 2009), has led to very impressive results.

1.3 State of the art, operational solutions.

Once the basics algorithmic and software kernel exist, which was the case since around 2005, there is still some issue before usable tools are offered to people. These tools allowing professional practitioners to modelize and measure the word in 3D can take many different aspect; among the many initiatives existing now days, here is a brief review of the possible solutions:

- Web-service, this solution has the advantage of the easiness for the end-user (no program to install) and allow a cost synergy by sharing the CPU, several service exists (Arc3d 2010) or have existed (3dsee 2009, apparently closed), today they are generally free of charge; these web services are certainly a good window to promote these techniques; however these solution has several drawbacks for an operational use : it requires people to share their image which may be unacceptable, it requires to have an internet access (difficult during a mission in the desert…), the code used is generally not free (even not accessible as binary) and so cannot be improved nor finely adapted to the user needs;
- Classical software: we mean tools distributed by software editors that are generally nor open source nor free of charge; the drawback of these solutions it to pay the licenses of these software and, maybe more importantly, to have to install on your computer software on which you have no control; of course, for people who can accept these disadvantages, these solutions have the advantages of simplicity…
- Open source-free of charge software: to our opinion there is a real need for many practitioners, from the scientific community, in cultural heritage and environmental surveying, for such software; (OSP 2010) is a web site dedicated to presents these solution; the most known solution is based on bundler (snably 2008) and PMVS (Furukawa 2009);

1.4 Why yet another open source solution?

Today, there exists already some operational “open-source free of charge” software solution to compute 3D model out from a set of images; for example, one of the most well-known, bundler-PMVS, is easy to use and can produce appealing 3D clouds out of an unordered collection of images. These solutions generally have the advantages of the computer vision approach: they can work even when the images were acquired by people having very few ideas of photogrammetric constraints and they deliver data in cloud points formats easy to use in current software. However, to our knowledge, for most of these open source solution, the price to pay for the flexibility is an incomplete photogrammetric rigour in the formulation of equations that may lead to unacceptable precision for some application. More precisely:

- The orientation kernel is funded on the evaluation of fundamental matrix, leading to evaluate more values by camera than the 6 unknown necessary to physically parameterize a rigid movement in Euclidean space; this over parametrization, which for example allows systematically each image to have its own focal length, is classically the source of imprecision and derives;
- The calibration model proposed by these software, are relatively simple, most often limited to 1 to 3 coefficients of radial distortion; with light and low cost compact camera as can be embedded on small UAV, precise aero-triangulation of large set of images requires the possibility of more sophisticated distortion : at least decentric distortion and sometime arbitrary polynomial distortion; furthermore, with the open source solution we know, it is impossible to use images acquired by fish-eye, which is really unfortunate as fish-eye are extremely convenient for the quick photogrammetric modelization of global interior scenes; furthermore, for practical reason (zooming, focussing, change of lenses) it is generally impossible to calibrate the camera in labs and we need software that allow optionally the computation of, eventually high order, distortion directly on the data acquired on the field by self-calibration methods working during the bundle adjustment steps;
- For the matching steps, these softwares generally work on a subset of well contrasted points; this choice is quite pertinent for some 3D modelization of interior scenes, but it may be inadequate for other application where a dense matching is required.

For all the reason mentioned above, IGN arrived to the idea that the photogrammetric software that were developed internally in its labs, for aerial photogrammetry purpose, could be of interest in some application of the scientific community and it was decided in 2007, to make an open source deposit of MIMAC (MICMAC 2010), a software dedicated to aerial image matching. Since 2007, MicMac has evolved taking into account more complex scenes, it has then been completed by APERO, a photogrammetric bundle adjustment software for automatic orientation of images.

2. ALGORITHMIC ASPECT

2.1 General “philosophy” and protocol

Making 3D modelization out of unordered collection of images is a very topic popular now days and very impressive results are obtained by several teams working on crowd-sourcing out of images collected on the web. However, this is definitely not our goal.

The software we develop are rather meant to scientific using photogrammetry as a measuring tool for their current researches; these scientific can be, for example, architects, archaeologist, geomorphologists…

We think that such users are ready to pay the price of learning some simple rules and protocols in image acquisition if the respect of these constraints leads, at the end, in an easier and more precise modelization. As photogrammetrist have known for a long time, for a good quality 3D modelization, one day of acquisition on the ground require five days of work in the lab; so it may be valuable to take a bit more time taking the photo if its helps a lot of time in the lab…

Briefly speaking, the protocol to respect for terrestrial modelization is the following :

- For each desired cloud points take a “master” image, and 4 closed associated images (with low ratio base to height);
- Between each master image take a sufficient number of intermediary images to assure the connection during the orientation step; as a rule of thumb, if your are turning an object, it is perfectly sufficient to take an image each 15 degree;

For aerial modelization, the protocol is just the respect of the usual rules of “classical” aerial photogrammetry.
2.2 General pipeline of 3D process software.

The general pipeline of our 3D modelization process is quite classical an divided in 3 main steps:

- First steps compute tie points from all pair of images (eventually use heuristic or external information to know which image may have a part in common);
- Second steps compute relative orientations from the tie points and, if appropriate, convert the relative orientation in absolute orientation from auxiliary data;
- Third steps compute the dense matching from the oriented images using the MicMac software.

For the first step, for our research work we use a completely external solution: the sift++ implementation of sift algorithm (Vedaldi 2010) which has just been modified to work with large images. Sift generally gives us sufficient result when the images are taken respecting the protocol described bellow. However, the idea is that if required, for performance of legacy purposes, the user may use MicMac and Apero and replace sift by his own detector.

For the second and third step, APERO and MICMAC are completely home made software developed in C++. The main characteristics of MICMAC are described in (Pierrot Desseilligny 2006).

2.3 Main modules of APERO.

The mains modules of APERO are:

- A modules specialized in the computation of initial solution, it uses algorithm like essential matrix, space resection and a specialized scheduler to find an “optimal” tree (each image, except the “first” one having a set of “fathers” on which its orientation is built);
- A bundler adjustment module, classically based on a linearization of projection equation and Gauss-Newton or Levenberg-Markard iteration;
- Modules for absolute or scene-based geo-referencing;
- Several module for importing an exporting data : tie point, ground point, internal and external calibration, GPS if existing …

2.4 Computation of initial solution.

When doing a completely relative orientation (when there is only tie points), the user choice a first image which will arbitrarily fix the orientation and origin of coordinates. Then APERO enter a main loop, as long as there exist non oriented images:

- Select the next image, according to an a priori estimator of the stability that will result from the orientation computation;
- Compute a first orientation using algorithm leading to direct solutions; APERO test essential matrix with ransac and if there is sufficient multiple points try also the space resection algorithm with ransac; select a posteriori the best solution;
- Regularly make bundle adjustment on the already oriented image to avoid divergence.

As an estimator of the stability, we analyse for each image the cloud of tie points with already oriented images. We want to select an image where the tie points are numerous and homogeneously spread. The estimator we use is the smallest value of the inertial matrix of this cloud.

2.5 Bundle adjustment part.

The bundle adjustment part is quite classical (Triggs 2000). To be precise, for each tie point:

- we compute an estimation of the ground point by bundle intersection of all images where it is seen, using the current values of external and internal calibrations;
- we add a term to the minimization which is the sum of the reprojection in the images of the ground point, this term depends of three kind of unknown : the ground point, the external orientation, the internal orientation;
- this term is linearized as we have an estimator of each unknown and is added to a global quadratic form that will be minimized;

To avoid to have a system that growth in proportion of the number of unknown ground points, we use a “trick” current in aerial photogrammetry :

- if one has to solve the system (with block matrix notation):
  \[
  \begin{bmatrix}
  \lambda & B & 0 & X \\
  'B & C & D & Y \\
  0 & 'D & E & Z \\
  \end{bmatrix}
  \begin{bmatrix}
  U \\
  V \\
  W \\
  \end{bmatrix}
  =
  \begin{bmatrix}
  X \\
  Y \\
  Z \\
  \end{bmatrix}
  \]

in our case X is the unknown 3-vector corresponding to the ground point, Y represent the unknowns (internal and external) linked with X, and Z all the other unknown;

- then it is equivalent to solve the system where X has disappeared :

\[
\begin{bmatrix}
  C^{-1}B\lambda^{-1}B & D \\
  'D & E \\
  \end{bmatrix}
  \begin{bmatrix}
  Y \\
  Z \\
  \end{bmatrix}
  =
  \begin{bmatrix}
  V^{-1}B\lambda^{-1}U \\
  \end{bmatrix}
  
\]

So we have a system with a number of unknown roughly equal to 6 time the number of images (plus some internal calibration unknown). It can be very sparse, for example in aerial case, and is solved by cholesky method after suboptimal ordering of the unknowns using an “approximate minimum degree algorithm”.

When it’s adequate, many other observations can be integrated to the minimization. The most current :

- link to GPS observation of image center;
- links to known ground points;

2.6 Control of variable and energy function.

At each step of the minimization process, the user can select:

- Which unknown are free to evolve and which are temporarily frozen, generally it is a good idea to freeze the internal calibrations at the beginning of the optimization;
- A choice of weighting function corresponding to different (so called) robust norms;

2.7 Internal calibration model.

APERO propose a large set of internal calibration models:

- Distortion free camera;
- Radial distortion with polynomial model;
- Radial distortion with centric and affine parameters;
- Ebner’s and Brown’s model;
- Polynomial models, from degree 3 to 7;
- Fish-eye model made by combination of a theoretical model and a additional polynomial distortion, the theoretical model can be equidistant or equisolid;

The user can also select a combination of N previous models, in this case the N-1 first models are used as a pre-correction of measures and the last one is optimized.

2.8 Geo-referencing of solution.

For many professional applications, an arbitrary relative orientation is not what is needed; APERO offers different possibility to geo-reference the results:
• when embedded INS-GPS exists, they can be the starting point of compensation with attachments to these observations;
• when only embedded GPS exist, it can be used to transfer absolute orientation once relative orientation is computed;
• when at least three ground point are known and measured in two images, they can be used to transfer absolute orientation once relative orientation is computed;
• when no absolute orientation data are known, APERO offers the possibility to build an orientation having some physical significance by specifying an horizontal plane.
Of course, all these observations can also be used during the compensation process.

3. OPERATIONNAL ASPECT

3.1 License.

The sift software we use internally for our own research as a tie point generator for APERO is submitted to the sift patent, please refer to this license to see if you can use it. Else, one can easily replace SIFT by another tie points generator.
The APERO and MICMAC software are open-source software, they are released under the Cecill-B licence (Cecill 2005). This licence is essentially an adaptation to the French law of the L-GPL licence; we have chosen it because it is quite permissive, so if one wants to use these software it is most probable that there wont be any legacy issues.

3.2 XML Parameters.

The APERO software is very (to much ?) parametrical. However it is not an interactive software, all the parameters of the orientation process are stored in an XML file. These files, that contain currently one hundreds of tags, are structured in several section:
• A section describing the observation (tie and ground points, GPS …) and who refers to exterior files;
• A section describing the unknown and the way to compute their initial values;
• A section to describe the different steps of bundle adjustment, with weighting parameters that can be reset at each step;
• Several section for exporting results and controlling the optimization kernel.
Obviously, this kind of parameterisation is not very appealing for practitioners… In fact this is a photometric kernel and we regard this interface as a portability layer for developers who will write more user friendly interface. Two such interfaces are mentioned bellow.

3.3 Formal Code generator .

One important aspect of APERO implementation is that we have developed a formal code generator; starting from a formula describing a new parametric calibration, it generates automatically an optimized C++ files that compute the partial derivative of the projection formulas relatively to all the parameter.
This kind of approach is very useful for managing complex calibration like fish-eye where the derivative can take hundred of line.

3.4 Interface

As said before, the basic XML-interface is not user-friendly and would be difficult to use for practitioner. A project is currently being developed at IGN for encapsulating all the complexity of the XML file and allows to run interactively APERO and MICMAC. This interface can be downloaded at :
http://www.micmac.ign.fr/svn/micmac_data/trunk/DocInterface

The lab Map Gamsau has also developed an interface to APERO-MICMAC, as a plug-in Maya, it is specialized for architectural application. The documentation can be found here :
A site dedicated to this interface will open in march 2011 :
http://www.map.archi.fr/tapenade
The XML-low level interface will be documented at end of 2011.

4. EXPERIMENTS

In all this experiment we have used Sift as tie-point generator, APERO as image orienter and MICMAC as image matcher ; for 3D-arch data set where few dense matching are made, and most example only show the 3D resulting from tie points.

4.1 Interior scenes

As we said before, one useful characteristic of APERO is that can manage natively the fish-eye cameras, this is convenient for interior scene modelization because :
• If one do not need a high resolution for all the scene, one can cover the totality of the scene with very few images,
• It lead to photogrammetric block that are very solid,
The figure 2 shows the results of an experimentation for modelization of the underside of a bridge, the bridge image where acquired from a bark, 150 images were used, mixing fish-eye for the 4.2 UAV modelization of exterior scenes

Photogrammetry with UAV is interesting as it is often an intermediary case between the terrestrial case and the classical aerial surveying. Figure 3 et 4 present two UAV mission that were oriented with APERO:

- Figure 3, data has been acquired with a survey copter UAV (Survey copter 2010), able to embed 5kg of payload, the camera was a Nikon D300; we had a GPS and INS who were used as initial solution to the bundle adjustment; the model was created with 100 images;
- Figure 4, data have been acquired with a drone lehman (Lehman UAV 2010) a very light and cheap UAV, the camera is compact Olympus tough-8000 (182 g), we had no GPS nor INS, we have compute the relative orientation of the 200 images and used a small part of the ground with horizontal plane to give to the block an orientation physically coherent.

In both case the DTM was produced with MicMac using a ground geometry on the oriented images.

4.3 Modelization of objects

The pipeline SIFT-APERO-MICMAC has been used for modelization of small object by our lab, or our colleagues of Map-Gamsau (Godet 2010) over hundreds of examples.

4.4 Data-set of 3D-Arch

We have also tested our software on different data set from the 3D-Arch special session on “Markerless automated orientation of image sequences”:

- The piazza navona furnished by FBK;
- The karnak temple furnished by IGN;
- The “castle” furnished by EPL;
• Several data set furnished by Verone university; this data set contained an example on which we failed to compute an orientation;

**FBK Data set**:
This data set was made with all the image having the same calibration.
We have run APERO on this data set with the following strategy (which our default strategy on large dataset):

• Select a subset of image with some depth variation, and make a first self calibration of the camera;
• Make a global orientation of the whole set of images, using the previous internal calibration; freeze the calibration as long as all image are not oriented, free it at the end of adjustment,

Figure 7 show some results, they seems coherent although we had no time to make a numeric evaluation. The focal we computed is significantly different from the value given in the dataset.

**IGN Data set**:
This data set was made turning around a column of the hypostyle temple of Karnak. The camera was an average quality brige. Two focal length were used:

• At 5 mm we made global view of the colum, with sufficient overlapping
• At 15 mm, we made 24 sets of 4 view, optimized for dense matching;

The orientation was made in two step: first orientation of 5 mm and then orientation of the whole set, starting of the previous initial values for the 5mm.

Figure 8 represent the result of orientation. As this data set was acquired with our protocol, we were also able to do dense matching using MicMac. Figure 9 presents the result.
EPFL Data set:
We have used the castle set, as it contains few images, we made the calibration in one step. The results seems coherent, although there are very few points found by SIFT.

Verone University Data set:
We have used the Pozzovegiani and Dante set, the results seems coherent, although there are few sift points with these small image. We have try also the piazzaerbe dataset but were not able to compute a complete orientation; our SIFT detectors could not find sufficient number of point and APERO complained that the were several unconnected block of images.

5. CONCLUSION
This paper has given a brief description of APERO, a totally free and open source software for photogrammetric calibration of sets of images. APERO can compute initial value of orientation, from only tie points, using classical algorithm as essential matrix and can make fine compensation by bundle adjustment using rigorous photogrammetry. Although there are many quality open source software in computer vision, we think that there is a need, at least in the scientific community, for software offering both the possibility of computer vision and photogrammetry.
One of the weaknesses of APERO is currently it lacks of interface and documentation, however several interface are being developed and a first level of documentation will be delivered before end of 2011.

References:
Faugeras O., 1994, Three dimensional computer vision, MIT Press.


(CECILL 2005) : http://www.cecill.info/licences/Licence_CeCILL-B_V1-fr.html


3dsee 2009, http://3dsee.net/


Godet A., Pierrot Desesilligny M., de Luca L., Une approche pour la documentation graphique 3D d‘édifices patrimoniaux à partir de (simples) photographies », In Colloque Photogrammétrie au Service des Archéologues et des Architectes, SFPT-CIPA, Villeneuve lez Avignon, Septembre 2010.


Photosynth 2010, http://photosynth.net/


Tapenade 2011 : http://www.map.archi.fr/tapenade