

UNE INTERFACE ERGONOMIQUE DE CALCUL DE MODÈLES 3D PAR PHOTOGRAMMÉTRIE

Isabelle Cléry, Marc Pierrot-Deseilligny

Université Paris-Est, IGN, Laboratoire MATIS 73, avenue de Paris, 94165 Saint Mandé
isabelle.clery@ign.fr ; marc.pierrot-deseilligny@ign.fr

Résumé

De nombreux logiciels du commerce proposent de créer des modélisations en 3D d'objets ou de scènes à partir uniquement de photographies. Cependant les professionnels, notamment les archéologues et les architectes, ont un certain nombre de contraintes qui restreignent fortement le choix du logiciel, telles que des ressources financières limitées, des connaissances peu approfondies en photogrammétrie et en informatique, des impératifs qui limitent le temps d'acquisition et de traitement, et enfin des attentes précises en terme de qualité, de précision et de complétude de la modélisation. Le laboratoire MATIS a développé un ensemble d'outils libres, open source, qui permettent d'effectuer l'ensemble des étapes du traitement d'un chantier photogrammétrique et qui fournissent des modèles 3D très denses et précis grâce à la rigueur de l'approche photogrammétrique ; ils représentent donc une solution intéressante pour ces professionnels du point de vue algorithmique. Cependant ils ont été développés pour des applications de recherche et sont donc peu ergonomiques pour le public non photogrammètre et non informaticien. C'est pourquoi une interface graphique est en cours de développement afin de fournir un accès simplifié et unifié à ces outils, notamment aux architectes et aux archéologues. Grâce à l'interface, l'utilisateur peut manipuler les images d'un même chantier et contrôler chaque étape du traitement qui comprend : le calcul automatique de points homologues entre les images, l'estimation des poses de la caméra et le calcul des modèles 3D. Les résultats peuvent être convertis en cartes de profondeur ombrées ou en nuages denses de points 3D (au format standard ply), et peuvent être affichés directement par l'interface. Pour offrir la possibilité de traiter des chantiers suffisamment variés, tout en masquant la complexité du paramétrage, la démarche retenue est de présenter à l'utilisateur un accès par grande famille de prises de vue, par exemple : chantier convergent, relevé de façades, chantier aérien sub-vertical, modélisation d'intérieur. . . Une attention particulière est apportée à la qualité de la documentation et à la portabilité du produit sur les principaux types d'ordinateur grand public (Linux, Mac et Windows). Cette interface, bien que non finalisée, permet déjà d'effectuer l'intégralité du traitement de chantiers de type convergent et est disponible en ligne. Des utilisateurs issus des différentes communautés de professionnels ciblées (archéologie, architecture, géologie, photogrammétrie architecturale) l'ont déjà testée et l'interface est régulièrement mise à jour afin de l'adapter aux remarques et demandes de ces testeurs.

Mots clés : Interface, chantier photogrammétrique, carte de profondeur, modèle 3D, estimation de poses.

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, 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 ; they represent an interesting solution for these professionals from an algorithmic point of view. Unfortunately they were developed for research applications and thus are not ergonomic for non-photogrammetry and non-computer-scientist public. Therefore a graphical interface 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 interface. To offer the possibility of treating varied enough sites, while hiding the complexity of settings, the chosen approach is to present the user with access by shooting family, for example : convergent shooting, survey of facades, sub-vertical aerial shooting, interior modeling . . . 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 surveys of convergent shooting and is available online. Users from the different targeted professional communities (archeology, architecture, geology, architectural photogrammetry) have already tested it and the interface is regularly updated to adapt it to these testers' comments and requests.

Keywords : GUI, photogrammetric treatment, depth map, 3D modelling, camera parameters estimation.

1. Introduction

De nombreuses solutions de reconstruction 3D à partir d'images uniquement se sont développées sous l'effet de la forte demande venant de nombreux domaines d'activité tels que l'archéologie, la géologie, l'architecture, l'expertise d'œuvres d'art., domaines généralement éloignés de la photogrammétrie et de l'informatique. En effet la numérisation 3D à partir d'images est une technique complémentaire du laser, mais moins coûteuse, plus pratique et discrète (donc mieux adaptée aux chantiers difficiles d'accès), de précision modulable, et surtout qui restitue les couleurs de la scène. Parmi ces logiciels, on peut citer :

- Bundler, logiciel libre, gratuit et multiplateforme de Microsoft, produit des modèles 3D peu denses pour des applications non professionnelles (tourisme). Photo Tourism est le projet à l'origine de PhotoSynth (accessible à l'adresse <http://photosynth.net>) et de Bundler (accessible à l'adresse <http://phototour.cs.washington.edu/bundler/>);
- CMVS (PMVS2, (Furukawa et Ponce, 2010)), logiciel libre, gratuit et multiplateforme issu de la recherche, produit des modèles 3D denses à partir d'images géoréférencées (avec Bundler par exemple), et s'adresse aux professionnels de la vision par ordinateur ;
- Photosynth (Snavely et al., 2008) s'appuie sur Bundler pour calculer les modèles 3D sur un serveur distant (ce qui assure la rapidité du calcul); il offre de plus un système de partage des fichiers ;
- ImageModeler, logiciel payant d'Autodesk, s'adresse aux professionnels utilisant déjà les autres logiciels d'Autodesk (géomètres...). Il est accessible à l'adresse <http://usa.autodesk.com/adsk/servlet/pc/index?siteID=123112&id=11390028>.
- PhotoModeler, logiciel payant d'Eos Systems (une version gratuite simplifiée est aussi disponible sous Windows), produit des modèles denses et texturés et s'adresse aux professionnels (archéologues, architectes...); un point négatif cependant est la forte implication de l'utilisateur dans la modélisation. PhotoModeler est accessible à l'adresse: <http://www.photomodeler.com>.
- Arc3D (Vergauwen et Van Gool, 2006), solution gratuite et multiplateforme, produit des modèles 3D via un serveur distant pour des applications non professionnelles (tourisme).

Cependant le choix d'un tel logiciel est soumis à de nombreuses contraintes liées au cadre d'utilisation,

aux ressources disponibles et aux applications, contraintes qui réduisent alors considérablement le choix du logiciel. En particulier, l'offre gratuite et ergonomique ciblant les utilisateurs non photogrammètres et en vue d'applications professionnelles, est très réduite. De plus il existe de nombreux algorithmes de reconstruction différents (méthodes volumiques ou surfaciques, utilisation de silhouettes, de l'ombrage, des textures ou de la photoconsistance ...) dont la qualité du résultat peut varier suivant la configuration de la scène (qualité de la texture sur l'ensemble de la scène, présence de concavités, variations de la luminosité, possibilité de saisie manuelle d'éléments géométriques de la scène...). Dans ce contexte, le laboratoire MATIS¹ de l'IGN a mis à disposition du public un ensemble d'outils open-source permettant d'effectuer l'intégralité du traitement de chantiers photogrammétriques, et dont une interface ergonomique, actuellement en cours de développement, est présentée ici. Après une présentation du contexte du projet et du principe retenu pour l'interface, les principales fonctionnalités de l'interface seront détaillées, suivies de quelques résultats obtenus dans différents domaines ; enfin les évolutions prévues seront détaillées.

2. Contexte

2.1. Les outils développés par le laboratoire MATIS

Le laboratoire MATIS de l'IGN dispose déjà d'un ensemble d'outils permettant chacun d'effectuer une étape du traitement de chantiers photogrammétriques. Les principaux outils, utilisés successivement lors du traitement, sont présentés ci-après.

2.1.1. Pastis

Pastis (Programme utilisant Autopano Sift pour les Tie-points dans les ImageS) utilise l'algorithme SIFT pour rechercher automatiquement des points homologues sur différents couples d'images se recouvrant partiellement. SIFT (Scale-Invariant Feature Transform) est un algorithme développé à l'extérieur du laboratoire MATIS (Vedaldi et Fulkerson, 2008). Pour un couple d'images donné, il recherche d'abord des points d'intérêts sur chaque image (ce sont des zones de fort gradient : coins d'objets, taches...). Puis il apparie chaque point d'intérêt d'une image avec ceux de l'autre image selon la ressemblance de leur voisinage. L'intérêt de cet algorithme est que les points homologues trouvés sont invariants par changement de contraste (sauf saturation), changement d'échelle, translation de l'image et par faible rotation de la caméra, ce qui laisse la liberté au photographe de modifier la luminosité lors de la prise de vue et de prendre les photos selon des directions différentes. En revanche ces points homologues sont sensibles au flou et aux mouvements (modification des ombres, déplacements d'objets en arrière-plan, arbres...).

¹Méthodes d'Analyses et de Traitement d'Images pour la Stéréo-restitution, <http://recherche.ign.fr/labos/matis/accueilMATIS.php>

La spécificité de Pastis est de lancer l'algorithme SIFT sur plusieurs couples d'images en parallèle pour accélérer le calcul.

2.1.2. *Apero*

Apero (Aérotriangulation Photogrammétrique Expérimentale Relativement Opérationnelle) calcule les paramètres des poses de la caméra (sommet de la perspective, orientation, réestimation de l'étalement) à partir d'un ensemble de points homologues et d'étalements approchés des objectifs utilisés. Apéro utilise une approche purement photogrammétrique, ce qui lui permet d'obtenir une aérotriangulation (estimation des paramètres des poses) très précise. En effet, il utilise des équations exactes afin de calculer des paramètres physiques réels (sommet de la pose, orientation...), contrairement à l'approche de la vision par ordinateur qui estime des grandeurs mathématiques théoriques (matrice fondamentale) comprenant plus de paramètres que nécessaire, ce qui revient à réestimer la focale de chaque pose et peut entraîner une dérive de l'aérotriangulation et donc une imprécision du résultat final. D'autre part, l'approche photogrammétrique permet d'inclure dans le calcul la réestimation d'étalements complexes des objectifs (notamment des fish-eye). Le calcul des orientations relatives et des étalements réestimés se décompose en deux parties:

1. Le calcul d'une solution initiale :

- L'utilisateur désigne une image "maîtresse" (pose située à l'origine du référentiel et d'orientation l'identité),
- Tant qu'il subsiste une image à orienter, la prochaine image à orienter est celle présentant le plus de points homologues avec les images déjà orientées et les mieux répartis dans l'image (Apero utilise la matrice d'inertie des points homologues),
- L'orientation de l'image est donnée par la meilleure solution obtenue par le calcul de la matrice essentielle puis l'algorithme RANSAC d'une part, et par la méthode de résection de l'espace puis RANSAC d'autre part,
- Régulièrement, on procède à un ajustement par faisceaux sur toutes les images déjà orientées pour éviter la dérive du modèle ;

2. Le calcul des paramètres par ajustement de faisceaux :

- Chaque point homologue est projeté au sol par intersection de faisceaux puis reprojecté dans chaque image,
- La somme des écarts images est calculée en fonction des coordonnées du point au sol, des poses et des étalements,

- Après linéarisation des équations précédentes au voisinage de la solution initiale et simplification, les poses et étalements sont calculés par moindres carrés itérés.

À chaque itération, l'utilisateur peut modifier les pondérations ou choisir quels paramètres réestimer. Il est possible d'intégrer dans ce calcul :

- Un géoréférencement absolu par des données GPS, INS et GPS, des points terrains ou un géoréférencement utilisateur (orientation donnée par un plan horizontal et un axe, associée à une échelle),
- Une orientation relative initiale, ce qui permet d'effectuer le traitement par blocs d'images,
- Des types d'étalements très variés (distorsion radiale, affine, décentrique, fish-eye de modèle équidistant ou équisolide...)

2.1.3. *MICMAC*

MICMAC (Multi Image Correspondances par Méthodes Automatiques de Corrélations (Pierrot-Deseilligny, 2007)) est un logiciel de mise en correspondance d'images développé par Marc Pierrot-Deseilligny. MICMAC calcule entre autres des cartes de profondeur à partir d'un ensemble d'images au format .tif, de l'estimation des poses de la caméra et d'un intervalle de recherche altimétrique (qui correspond aux valeurs extrêmes des profondeurs). Les cartes de profondeur sont des images dont chaque pixel représente la profondeur du point 3D correspondant, du point de vue d'une image dite de référence. Le principe de calcul est le suivant :

- Le calcul se fait en géométrie image (cela revient à peu près à généraliser la géométrie épipolaire au cas multi-images) ; parmi les différentes géométries proposées par MicMac, celle-ci a pour avantage que l'image de référence est rigoureusement superposable au nuage calculé; notamment cette image de référence est un support naturel pour permettre à l'utilisateur de saisir des informations. L'espace objet exploré est discrétisé à un pas équivalent de 1 pixel en planimétrie (c'est-à-dire pour chaque pixel de l'image).
- Puis MICMAC calcule pour chaque point de cet espace une mesure de vraisemblance (somme des coefficients de corrélation pour chaque paire d'images entre le pixel correspondant sur la première image et le pixel correspondant sur la seconde image).
- Enfin la surface est extraite par une approche énergétique visant à minimiser sur tout l'espace étudié une somme comprenant la mesure de vraisemblance et un terme de lissage afin d'homogénéiser la surface.

Ce calcul est effectué itérativement sur les images sous-échantillonnées à différentes résolutions croissantes, chaque étape utilisant le résultat obtenu à la résolution précédente pour prédire le résultat et ainsi définir une zone de recherche. Cet aspect multi-échelle/multi-résolution permet de gagner à la fois en robustesse et en temps d'exécution. Une particularité du coefficient de corrélation utilisé ici est d'être invariant par changement de contraste, si bien que le résultat ne dépend pas de la luminosité de la scène. MICMAC fournit aussi une carte des coefficients de corrélation pour chaque profondeur retenue afin de quantifier la précision de la carte en chaque point. MICMAC étant assez flexible, de nombreux autres résultats peuvent être fournis : calcul de MNT dans un repère cartographique donné (il a été développé initialement dans ce but), calcul de la parallaxe transversale pour une réestimation très précise des poses...

2.2. Le public visé

Le laboratoire MATIS souhaite mettre ces outils à la disposition du plus grand nombre dans le but de les valider et dans le même temps de répondre à la mission de service public de l'IGN. Les modèles 3D résultant de l'utilisation de ces outils sont très denses, requièrent une prise de vue qui suive une liste de contraintes et nécessitent divers paramètres qui assureront la précision souhaitée. C'est pourquoi le public visé est essentiellement constitué des professionnels, ayant des compétences ou non en photogrammétrie, mais sans toute fois exclure les utilisateurs issus du grand public qui souhaiteraient ponctuellement obtenir un modèle 3D très dense. Ce public a pour particularité d'agir dans un cadre professionnel avec les contraintes qui y sont liées (contraintes économiques, matérielles, accès Internet limité ou inexistant lorsque le traitement s'effectue sur le chantier), d'avoir des compétences très pointues mais généralement limitées en photogrammétrie et d'avoir des attentes précises en matière de rendu, liées à un questionnement scientifique ou à des impératifs de gestion du patrimoine. En particulier, ces utilisateurs peuvent accepter de respecter quelques règles simples de photogrammétrie lors de la prise de vue dans le but d'améliorer sensiblement la qualité du modèle. Les outils proposés par le MATIS présentent donc l'intérêt d'être gratuits, de ne nécessiter que des données peu coûteuses, et de produire des modèles précis, denses et de résolution modulable.

2.3. Le problème

Ces outils étant conçus pour traiter une grande généralité de chantiers, ils requièrent un grand nombre de paramètres que l'utilisateur doit insérer manuellement dans des fichiers selon un format prédéfini. Cette complexité rend ces outils difficilement accessibles pour un utilisateur non spécialiste en photogrammétrie et en informatique. C'est pourquoi l'ensemble de ces outils a été réuni dans une interface graphique afin de permettre à tout utilisateur non expert d'effectuer l'intégralité du

traitement d'un chantier photogrammétrique de manière simplifiée. De même que les outils auxquels elle donne accès, cette interface est libre et Open Source afin de toucher toutes les catégories d'utilisateurs intéressés par le potentiel des outils mais découragés par la complexité de leur utilisation.

2.4. Les buts et les contraintes de l'interface

À partir d'un ensemble d'images d'un même chantier, prises selon les méthodes traditionnelles de photogrammétrie (recouvrement suffisant entre les images successives pour assurer leur orientation relative, redondance suffisante de l'information pour permettre le calcul des cartes de profondeur par corrélation selon les différents points de vue souhaités), non traitées (c'est-à-dire issues directement de l'appareil photo et généralement au format .raw), l'interface doit permettre un accès simplifié à l'ensemble des outils précédents et une exploitation aisée des résultats obtenus, tout en intégrant en entrée les différents paramètres tels qu'un éventuel étalonnage, les différentes focales utilisées... Cette interface doit répondre aux contraintes suivantes:

- Ergonomie et accessibilité à un large public, notamment via une simplification du paramétrage ;
- Adaptation à la généralité des données (forme de la scène, agencement de la prise de vue, objectifs utilisés, format des images...) ;
- Portabilité de l'interface (sur les principaux systèmes d'exploitation: Linux, Windows et Mac) ;
- Facilité d'installation.

2.5. Principe retenu

La solution retenue pour simplifier la saisie des paramètres a été de les regrouper selon les différents types de chantier (chantier convergent, parallèle, divergent, et autres types de chantier) ; ainsi une partie des paramètres est remplie directement par l'interface. Les paramètres restants peuvent être soit saisis graphiquement à partir de fenêtres de dessin ou d'affichage en 3D, soit proposés par l'interface à partir d'un calcul statistique. A chaque étape du traitement, une fenêtre permet de saisir les paramètres utiles ; le programme correspondant est alors appelé par ligne de commande. Le regroupement des calculs en une interface unique évite la redondance du paramétrage (images utilisées, noms des fichiers réutilisés en entrée-sortie). En complément des calculs, une fenêtre de visualisation du chantier en 3D a été intégrée à l'interface afin de contrôler le résultat de l'estimation de pose et d'afficher les nuages de points 3D issus des cartes de profondeur. Cette fenêtre de visualisation a été conçue à l'aide de la bibliothèque de rendu graphique en 3D OpenGL². L'accessibilité de

²OpenGL est accessible à l'adresse : <http://www.opengl.org>

l'interface est assurée par une documentation complète disponible à la fois via la fonction d'aide en ligne et indépendamment en pdf. Afin d'assurer la portabilité de l'interface, celle-ci a été développée à l'aide de Qt4³, bibliothèque multiplateformes pour interfaces graphiques en C++. Par ailleurs, l'interface utilise la bibliothèque open-source elise, créée par le MATIS afin de développer les différents outils précédemment présentés, et regroupant diverses fonctions de calcul et de gestion des fichiers utilisés par les outils. Enfin un installateur a été créé afin de simplifier la mise en place de l'interface et de ses dépendances. L'interface est encore en cours de développement mais permet déjà d'effectuer l'intégralité du traitement d'un chantier de type convergent sous Linux.

3. Aperçu de l'utilisation

3.1. Démarrage

À l'ouverture de l'interface (Figure 1), le programme recherche les outils nécessaires à son fonctionnement. L'utilisateur importe les images du chantier. Les formats acceptés sont tif(f) et tous les formats raw et jpg.

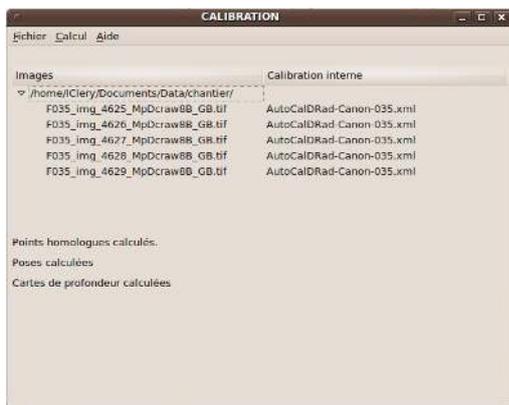


Figure 1 : Fenêtre principale de l'interface.

Chaque étape du traitement effectué via l'interface correspond à l'utilisation d'un des outils précédemment décrits.

3.2. Recherche de points homologues (Pastis)

L'utilisateur saisit d'abord les paramètres de la caméra :

- Le nom commun à toutes les images afin d'automatiser le processus (un nom est retrouvé et proposé par défaut),
- L'étalonnage (ensemble des paramètres internes de la caméra et de chaque objectif utilisé : focale, distorsion, point principal d'autocollimation, centre de distorsion). Si aucun étalonnage n'est disponible, seules la longueur focale et la taille du pixel (ou le format de l'appareil photo) sont obligatoires (Figure 2).

L'utilisateur choisit ensuite le type de chantier dont il s'agit dans une liste (Figure 3). Enfin l'utilisateur sélectionne les couples d'images pour lesquels seront recherchés des points homologues (une fonction permet de sélectionner les couples d'images voisines pour un voisinage donné) (Figure 4). La recherche de points homologues s'effectuant sur des images réduites en vue d'accélérer le calcul, l'utilisateur doit entrer une taille pour l'image sous-échantillonnée. L'interface associe chaque image à son étalonnage, convertit les images au format tif, puis recherche et apparie les points homologues pour les couples sélectionnés via le programme Pastis. Les coordonnées des points homologues sur les deux images sont regroupées par couple d'images dans des fichiers texte. L'avancement du traitement est indiqué dans la fenêtre principale à chaque étape du calcul (Figure 1). Par ailleurs une sauvegarde est effectuée à la fin de chaque étape afin de sauvegarder les paramètres et de permettre quand c'est possible la reprise du calcul au même point, même si l'étape n'est pas achevée.

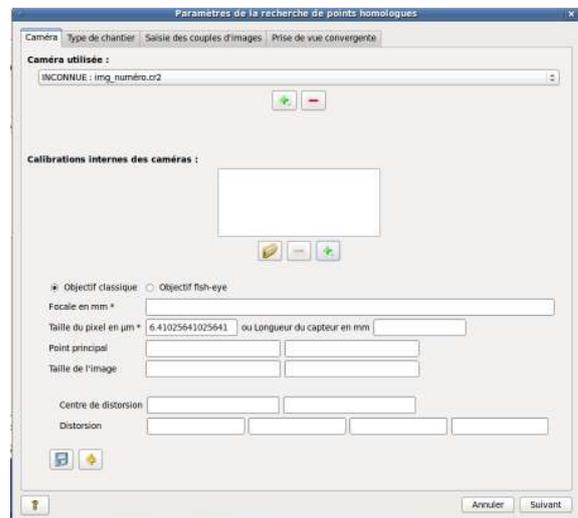


Figure 2 : Saisie des paramètres de la caméra.



Figure 3 : Saisie du type de chantier.

3.3. Estimation des poses de la caméra (Apero)

L'utilisateur sélectionne d'abord l'ensemble des images qu'il veut orienter. Ceci permet de relancer

³Qt est accessible à l'adresse : <http://qt.nokia.com>

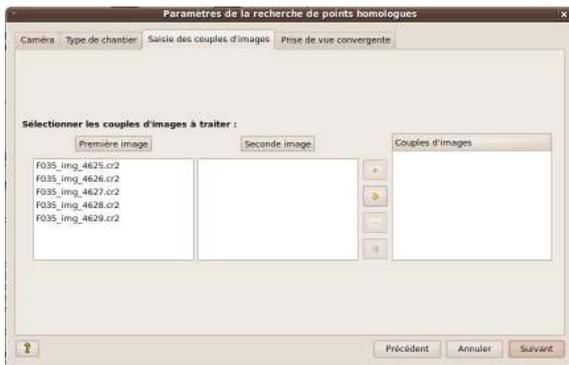


Figure 4 : Saisie des couples d'images.

le calcul sur un groupe réduit d'images dans le cas où trop peu de points homologues auraient été trouvés. L'orientation des caméras étant relative, l'utilisateur choisit l'image "maîtresse", c'est-à-dire la pose à l'origine du repère, relativement à laquelle les autres poses seront calculées (Figure 5). Apero définit l'échelle du chantier à partir de la base entre la pose correspondant à l'image "maîtresse" et une autre pose qu'il choisit lui-même. Il est aussi possible de fixer une orientation et une échelle globales au repère relatif précédent. Pour cela l'utilisateur saisit dans les images un masque du plan horizontal et l'axe des abscisses afin de fournir la rotation, ainsi qu'un segment pour lequel il fixe la longueur afin de fournir l'échelle du chantier. La possibilité de calculer une aérotriangulation absolue à partir de points terrains et de données INS et GPS existe dans Apero et sera prochainement implémentée dans l'interface. L'interface offre la possibilité d'effectuer un auto-étalonnage sur un sous-groupe d'images avant l'aérotriangulation proprement dite afin d'affiner les valeurs initiales des étalonnages fournis par l'utilisateur. Dans le cas d'images prises avec des objectifs de longueurs focales différentes, il peut être utile de réaliser l'aérotriangulation par blocs d'images de même longueur focale, afin d'éviter les divergences. L'utilisateur peut sélectionner via l'interface les images de courte focale qui seront orientées en premier, puis l'aérotriangulation globale sera effectuée à partir du résultat précédent. Il est aussi possible de réestimer un étalonnage propre à chaque image, ce qui est utile notamment si la mise au point a été modifiée au cours de la prise de vue. Enfin un filtrage des points homologues respectant leur dispersion est proposé afin d'optimiser le calcul et l'affichage. Le résultat de l'estimation des poses est visible dans une fenêtre de visualisation 3D où sont affichés les positions de la caméra, l'emprise des images et les points homologues trouvés en 3D. Différents outils permettent la navigation dans la vue, l'affichage des noms des images ou la coloration des points (couleurs hypsométriques, textures issues des images) (Figure 7).

Les paramètres (ré)estimés (sommets et orientations des poses, paramètres de l'étalonnage) sont regroupés



Figure 5 : Saisie de l'image "maîtresse".

par pose dans des fichiers xml.

3.4. Calcul des cartes de profondeur et modèles 3D (MicMac)

Les scènes au relief accentué ou de forme complexe présentent généralement des parties cachées (c'est le cas notamment des objets dont on fait le tour : statues..., et des objets présentant des concavités: bas-reliefs baroques...) ou plus simplement des zones obliques trop déformées par l'image pour assurer une résolution suffisante de la prise de vue (en particulier les bords des parties en relief). Il est alors préférable de réaliser des cartes de profondeur à partir de plusieurs points de vue différents afin de reconstruire les parties cachées ou trop déformées (le point de vue d'une carte de profondeur correspond au point de vue qu'aurait un scanner laser de même angle que le champ de l'objectif). L'utilisateur liste ici les différentes cartes qu'il veut calculer et saisit les paramètres pour chacune (Figure 6). Le programme MICMAC sera alors lancé autant de fois qu'il y a de cartes à créer avec les paramètres correspondants. Il sélectionne "l'image de référence", c'est-à-dire l'image dans la géométrie de laquelle la carte sera calculée. Il est possible aussi de la sélectionner à partir de la vue 3D (Figure 7). L'image de référence n'est pas obligatoirement l'image "maîtresse" utilisée dans la partie orientation.

3.4.1. Masque de l'image de référence

Le calcul nécessite par ailleurs un masque de la partie intéressante de la scène afin de ne pas cartographier l'arrière-plan (où la corrélation est généralement difficile et crée des artefacts). L'interface permet soit d'importer une image *.tif noir et blanc, soit de créer un masque grâce à une fenêtre de dessin où l'utilisateur peut tracer des polygones à ajouter ou à découper du masque. Une

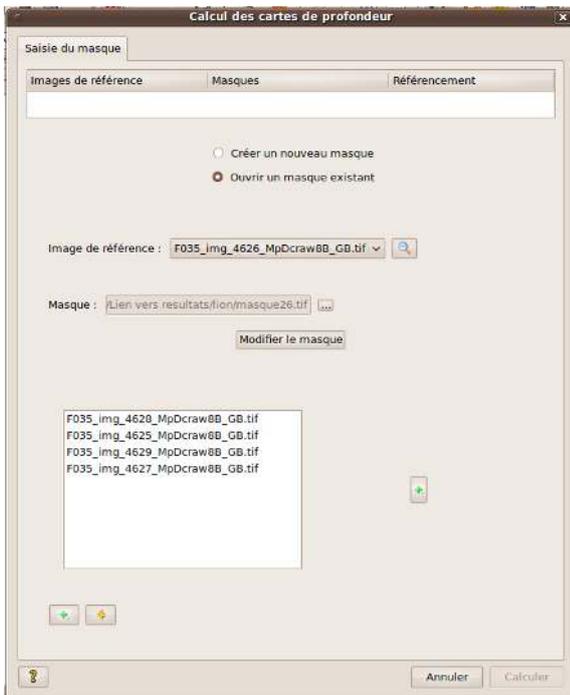


Figure 6 : Saisie des paramètres d'une carte de profondeur (image de référence, masque et images utilisées pour la corrélation).

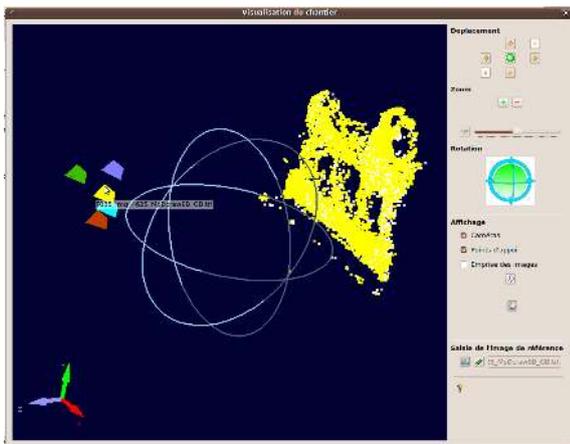


Figure 7 : Fenêtre de visualisation du chantier en 3D après estimation des poses de la caméra. Illustration de la saisie de l'image de référence pour une carte de profondeur dans la fenêtre de visualisation 3D du chantier.

fonction permet de suivre les contours de l'objet grâce à un calcul de plus court chemin pondéré par le gradient de l'image (Figure 8). Un autre outil permet de calculer un masque automatique (plein ou intégrant des "trous") de l'objet par triangulation des points homologues trouvés précédemment (Figure 9).

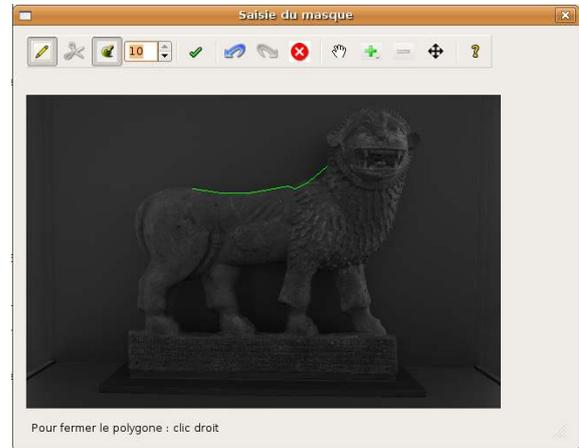


Figure 8 : Dessin d'un polygone pour la saisie du masque de l'image de référence à l'aide de l'outil de suivi des contours.

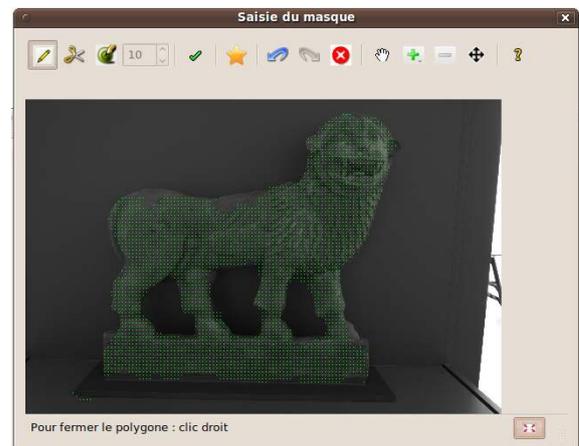


Figure 9 : Exemple de masque automatique avec gestion des "trous" calculé par la fonction de la fenêtre de dessin.

3.4.2. Images pour la corrélation

La carte de profondeur est créée par corrélation entre l'image de référence et d'autres images qui permettent de retrouver l'information sur la troisième dimension. L'utilisateur doit sélectionner ce groupe d'images soit à partir de la liste des images, soit dans la vue 3D, soit à l'aide d'un calcul statistique sélectionnant les images ayant le meilleur recouvrement avec l'image de référence (Figure 10).

3.4.3. Produits obtenus

Le résultat de ce calcul est une série de cartes de profondeur et de nuages 3D à différentes résolu-

tions. Les formats de ces données n'étant pas lisibles directement (les cartes de profondeur sont en 16 bits et les nuages au format hexadécimal), il est possible de les convertir via l'interface. Les nuages de points 3D obtenus sont au format *.ply et texturés avec l'image de référence. Cependant, afin de ne pas imposer l'installation de logiciels supplémentaires, ces nuages sont affichables directement dans la fenêtre de visualisation 3D de l'interface (Figure 11). Les cartes de profondeur converties au format tif représentent en niveaux de gris la scène sous un éclairage donné, réglable par l'utilisateur et éventuellement associé à une coloration hypsométrique. Ces cartes de profondeur ont l'avantage d'apporter une visualisation simple et intuitive de l'information sur la troisième dimension, et peuvent remplacer avantageusement un nuage de points :

- En l'absence de logiciel d'affichage 3D,
- Pour les présentations non dynamiques ou les publications (pour lesquelles l'imprimer-écran 2D d'un nuage 3D présente peu d'intérêt),
- Ou encore, pour transmettre les résultats grâce à leur volume réduit.



Figure 10 : Carte de profondeur avec ombrage.

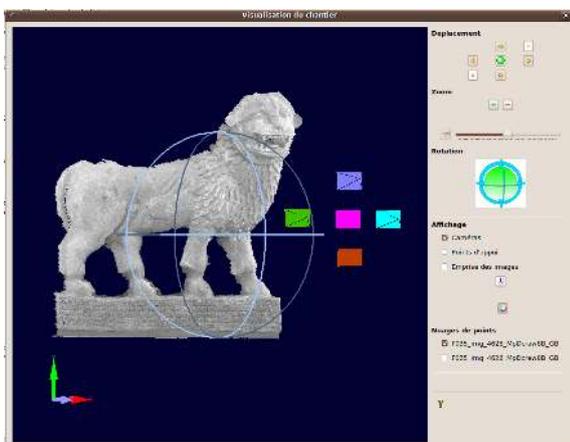


Figure 11 : Modèle 3D affiché dans la fenêtre de visualisation 3D de l'interface.

3.5. Contraintes d'acquisition

3.5.1. Contraintes géométriques

Le bon déroulement du calcul et la qualité du modèle 3D dépendent de l'acquisition. C'est pourquoi la prise de vue doit obéir aux deux règles suivantes:

- Règle pour l'aérotriangulation: les images voisines doivent se recouvrir partiellement et présenter un angle de prise de vue faible, ce qui devrait permettre de trouver des points homologues d'une image à la suivante (partie Pastis) et ainsi de les orienter relativement (partie Apero) ;
- Règle pour la modélisation 3D: chaque point de vue qui sera modélisé en 3D (partie MicMac) doit être constitué d'une image de référence et d'un ensemble d'images peu espacées (relativement à la distance à la scène) et telles que chaque pixel de l'image de référence soit vu dans au moins une autre de ces images (ce qui permet d'avoir la troisième dimension pour chaque pixel).

En pratique, dans le cas où l'utilisateur veut modéliser intégralement un objet par photogrammétrie terrestre (figure 12), il en fera le tour (prise de vue dite "convergente") et créera une carte de profondeur pour chaque face (ou selon les concavités de l'objet), de la même manière qu'avec un scanner laser. Chaque carte sera constituée d'une image de référence (point de vue du nuage de points) et de quatre images prises en croix encadrant cette image de référence (cette disposition permet d'éviter les parties cachées, notamment les côtés des aspérités) (règle 2). Puis il reliera les points de vue par des images intermédiaires (règle 1) prises tous les 10 à 20 degrés selon la courbure de l'objet. Bien sûr, ceci n'est pas l'unique méthode possible répondant aux critères précités.

À ces contraintes dues au traitement s'ajoutent les contraintes de l'utilisateur. Tout d'abord, la résolution du modèle 3D est la même que la résolution des images utilisées dans la partie MicMac. Ensuite, pour une meilleure précision, il faut éviter de modifier la mise au point d'une image à l'autre (pour un même objectif utilisé), en veillant à conserver une bonne netteté, ce qui permet de calculer un étalonnage unique pour toutes ces images et donc de réduire le nombre de paramètres (bien sûr, l'interface prend quand même en compte les changements de mise au point s'il y a lieu). Enfin, pour une précision optimale, il peut être utile soit de fournir un étalonnage précis de la caméra, soit de prendre quelques photos supplémentaires d'une zone très texturée de la scène afin de calculer via l'interface un étalonnage précis par auto-calibration.

3.5.2. Contraintes radiométriques

Les algorithmes n'étant pas sensibles aux changements de contraste, la qualité de la lumière sera réglée

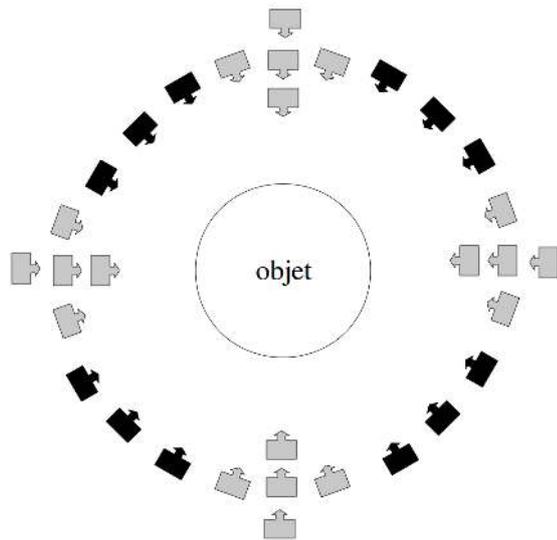


Figure 12 : Exemple de méthode de prise de vue pour modéliser un objet complet (vue dans le plan horizontal). Les poses en gris de la caméra correspondent aux images utilisées pour les cartes de profondeur, celles en noir correspondent aux images intermédiaires utilisées pour l'aérotriangulation. Le dessus et le dessous de l'objet s'acquièrent de la même manière.

uniquement par les contraintes de l'utilisateur. En particulier, les images de référence utilisées pour les modèles 3D servent à texturer ces modèles et devront être égalisées radiométriquement avant texturation des modèles. Il faudra par ailleurs s'assurer qu'il ne s'agit pas d'un type de scène défavorable pour les algorithmes utilisés, comme expliqué dans la partie Résultats.

4. Exemples de résultats

L'interface a été validée par des chantiers provenant de différents domaines d'application. Quelques exemples sont donnés ci-dessous.

4.1. Exemple de sculpture

La statue ci-dessous est une statue d'un éléphant photographiée en Chine.



Figure 13 : Photographie de l'éléphant.

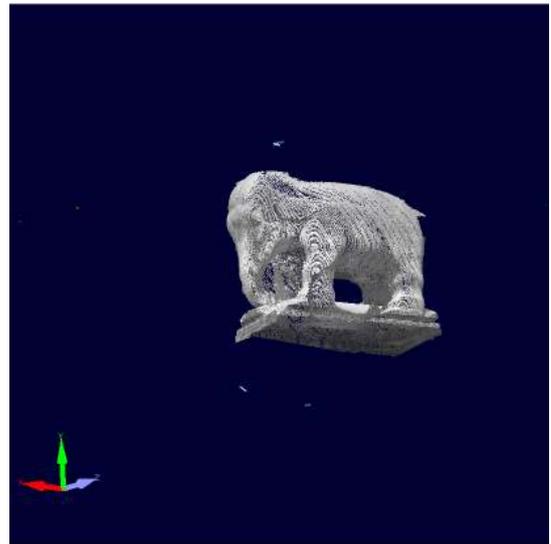


Figure 14 : Modèle 3D.

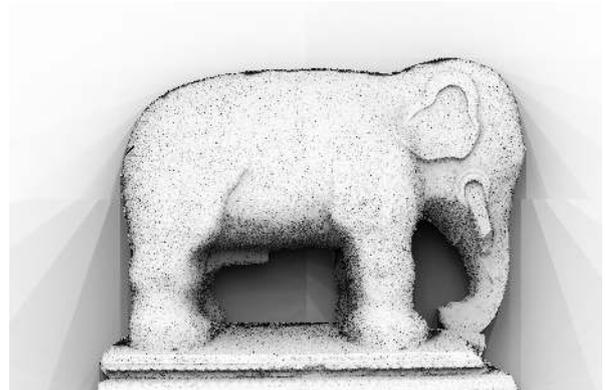


Figure 15 : Carte de profondeur ombrée du profil droit.



Figure 16 : Carte de profondeur ombrée de la face avant.

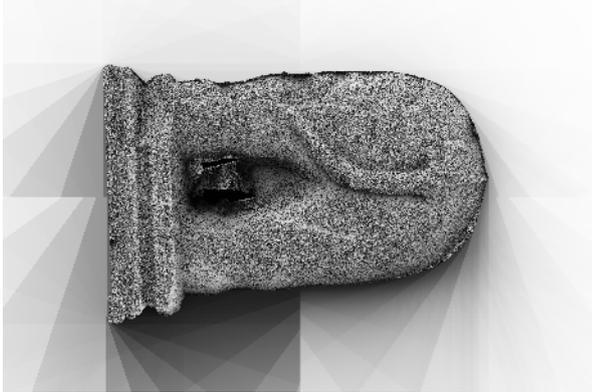


Figure 17 : Carte de profondeur ombrée de la face arrière.



Figure 18 : Carte de profondeur ombrée du profil gauche.

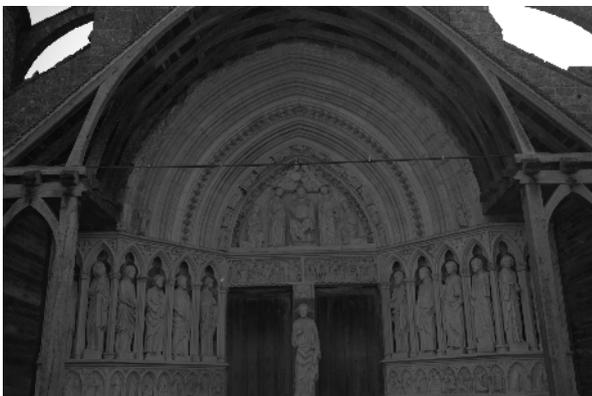


Figure 19 : Photographie du portail.

4.2. Exemple de façade

Le portail photographié ci-dessous est celui de l'église de Rampillon (Seine-et-Marne).



Figure 20 : Carte de profondeur ombrée du portail.

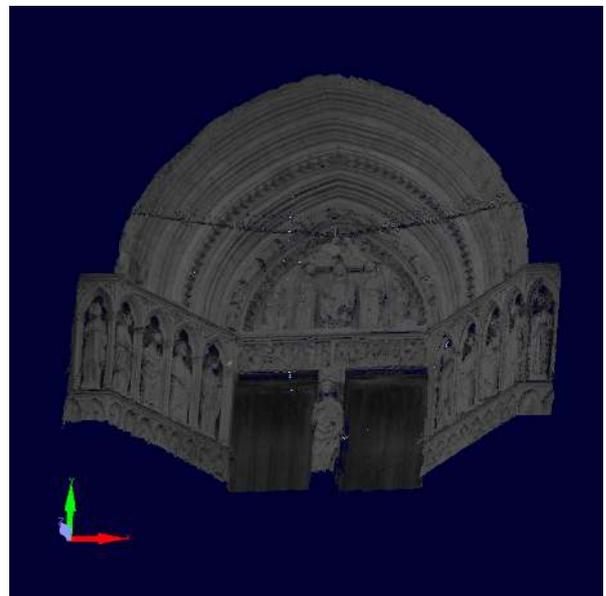


Figure 21 : Modèle 3D du portail.

4.3. Temps de calcul

Le temps de calcul dépend essentiellement:

- Du nombre et de la taille d'images : pour la conversion des images (nombre d'images), la recherche des points homologues (nombre de couples et taille des images sous-échantillonnées), la corrélation (nombre d'images utilisées pour une carte et taille des images), mais aussi pour la conversion des cartes de profondeur, l'affichage dans la fenêtre 3D.
- Du nombre de processeurs de l'ordinateur : certains calculs (conversion des images, recherche des points homologues, calcul des cartes de profondeur) utilisent l'ensemble des processeurs pour traiter les données en parallèle.

Beaucoup d'autres éléments peu quantifiables entrent en jeu :

- La texture des images : pour des images très texturées, beaucoup de points homologues seront trouvés ce qui allongera le temps de calcul de l'estimation de poses (mais améliorera la précision des cartes de profondeur).
- La rapidité de convergence du calcul de la carte de profondeur, qui dépend de l'aspect lisse de la scène (solution initiale proche de la solution finale) et de la texture de la scène (coefficients de corrélation bien différenciés).

Le temps de calcul ne peut donc pas être fourni précisément pour chaque chantier en fonction de données quantifiables. Cependant pour en donner un aperçu, les temps obtenus pour chaque étape du traitement sont indiqués dans le tableau 1 à titre indicatif pour deux exemples typiques de chantiers convergents.

4.3.1. Exemple avec l'éléphant

Il s'agit d'un ensemble de 68 images de taille 5634*3753 pixels, prises avec une focale de 35 mm. L'ordinateur utilisé a 4 processeurs. La recherche des points homologues s'effectue sur un voisinage des 7 images les plus proches (486 couples), et sur des images sous-échantillonnées de taille 1500*999. Les images étant bien texturées, de nombreux points homologues sont trouvés (jusqu'à 4000, en moyenne 500 par couple). Les cartes de profondeur sont calculées avec les 4 images les plus proches. Cet exemple correspond au cas le plus courant de modélisation d'un objet complet.

4.3.2. Exemple avec le lion

Il s'agit d'un ensemble de 5 images prises avec le même appareil photo. Les paramètres utilisés et l'ordinateur sont les mêmes que dans l'exemple précédent. Cet exemple correspond au chantier minimal utilisable pour des plans ou des surfaces d'objets de taille réduite (les gravures par exemple, dont la représentation fine du relief facilite la lecture ou l'étude du matériau).

4.4. Résultats attendus selon le type de scène et la configuration de la prise de vue

4.4.1. Types de scène défavorables

Du fait de l'algorithme utilisé et de la généralité des chantiers traités, certaines scènes peuvent donner de mauvais résultats quelle que soit la méthode de prise de vue. Ce sont d'ailleurs les cas d'échec classiques des méthodes à base de photogrammétrie: absence de texture et diachronisme. La méthode de mise en correspondance de MicMac est une approche "dense" où chaque pixel est apparié, contrairement à des outils comme PMVS qui n'apparient que les pixels de fort contraste et interpolent le modèle entre les pixels appariés ; si l'approche dense a l'avantage de donner des

résultats plus précis sur les matériaux texturés (pierres, terres, végétaux), elle a par contre l'inconvénient d'être plus bruitée sur les matériaux peu texturés (plâtre monochrome) ou spéculaires (objets métalliques ou en bronze). D'autre part, une modification importante de la scène d'une image à l'autre (y compris de l'arrière-plan et des parties non comprises dans le masque) peuvent altérer le calcul des poses de la caméra (étape Apéro), puisque celui-ci s'effectue à partir des points d'intérêt trouvés dans les images et appariés entre eux (étape Pastis).

4.4.2. Précision des résultats selon la configuration de la prise de vue

Il n'y a pas de formule prédéfinie pour estimer la précision attendue des résultats. En effet, elle dépend :

- De la résolution des images ;
- De leur texture : une image bien texturée aura un coefficient de corrélation plus grand et il y aura moins de risque d'erreur ;
- Et de la base utilisée (écart entre les sommets des poses successives de la caméra) : une base plus grande par rapport à la distance entre la caméra et la scène apporte une plus grande précision, dans la limite de déformations suffisamment faibles d'une image à l'autre pour assurer la corrélation.

La meilleure précision attendue (cas d'une corrélation parfaite en chaque pixel) peut-être estimée par la formule : précision = 0,5 * résolution * distance / base.

5. Évolutions de l'interface

5.1. Tests sur des utilisateurs

Quelques utilisateurs ont testé l'interface, peu nombreux du fait que l'interface est récente et son développement non terminé. Les testeurs couvrent les domaines de la photogrammétrie architecturale, de l'archéologie, de l'architecture et de la géologie. Il en est ressorti un sentiment de bonne ergonomie globale, de simplicité due à l'unification des logiciels en un seul, d'automatisation très poussée (recherche automatique de points homologues, pas de sélection des facettes des objets à modéliser...) et une documentation très claire et utile. La documentation utilisateur de MICMAC est accessible à cette adresse: <http://www.micmac.ign.fr/svn/micmac/trunk/Documentation/DocMicMac/DocMicMac.pdf>. La documentation utilisateur de l'interface est accessible à cette adresse: http://www.micmac.ign.fr/svn/micmac_data/trunk/DocInterface/doc.pdf.

Cependant il est dommage que l'utilisateur doive valider chaque étape avant de poursuivre, la sélection du masque est parfois laborieuse, les paramètres de la caméra ne sont pas intuitifs ou bien les utilisateurs ne savent pas toujours comment les obtenir, les messages

| | Éléphant | Lion |
|--|--|--------|
| Étape 1 (conversion des images et recherche des points d'intérêts) | 12 min | <1 min |
| Étape 2 (calcul des poses et conversion des points d'intérêts en 3D) | 10 min | <1 min |
| Étape 3 (calcul des cartes de profondeur) | 52 min (12 min faces et 15 min côtés) | 8 min |
| Total | 1 h 30 | 10 min |

Tableau 1 : Temps de calcul pour les exemples.

d'erreur ne sont pas toujours compréhensibles, le protocole d'acquisition n'est pas toujours connu ou compris, et il est dommage que les nuages de points ne soient pas convertis en données plus maniables (orthophotos, nuages triangulés, coupes...) Quelques améliorations ont été apportées à la demande des utilisateurs, telles que la mise à l'échelle du modèle 3D, l'adaptation à certains systèmes (Linux 64 et 32 bits, virtual box) et tailles d'écrans, l'optimisation de certains calculs et la documentation (au sujet de l'obtention des paramètres, des erreurs fréquentes ou de l'acquisition des données). Enfin certaines demandes ne seront pas suivies du développement correspondant, telles que la sélection de l'ensemble des paramètres au départ afin de tout calculer en une seule étape, pour des raisons de complexité de programmation et d'adaptation du calcul aux données (par exemple le calcul d'une carte de profondeur dépend de quelles images ont été correctement orientées).

5.2. Extension prévues

Différentes extensions sont prévues afin de répondre au mieux aux besoins des différents utilisateurs :

- Adapter l'interface aux autres types de chantiers, notamment aux prises de vue parallèles du type façades de bâtiment pour lesquelles il est préférable de calculer les cartes de profondeur dans une géométrie du terrain et non dans la géométrie d'une pose ;
- Fusionner les nuages de points 3D issus des différentes cartes de profondeur calculées pour un chantier, afin d'harmoniser les nuages, de fluidifier l'affichage du modèle 3D et filtrer les points bruités ;
- Calculer une carte de profondeur à partir d'un déroulé saisi dans la vue 3D ;
- Et bien d'autres en fonction de la demande...

6. Conclusion

Il a été présenté ici une interface graphique ergonomique et multiplateformes pour un ensemble d'outils libres développés par le laboratoire MATIS de l'IGN, permettant d'effectuer l'ensemble de la chaîne de traitement de chantiers photogrammétriques et accessible à tout utilisateur, photogramètre ou non. Cette interface actuellement en cours de développement permet de rechercher des points homologues entre des couples d'images, d'estimer les poses de la caméra et de calculer des cartes de profondeur et des modèles 3D à partir d'un ensemble d'images issues d'un même chantier. La densité et la précision des modèles 3D obtenus conviennent aux applications professionnelles telles que l'archéologie et l'architecture. L'interface propose une saisie simplifiée et essentiellement graphique du paramétrage, une visualisation 3D des résultats à différentes étapes du calcul et est accompagnée d'une documentation complète. L'interface étant destinée à un large public, elle est actuellement téléchargeable dans sa version bêta et recherche des testeurs afin d'améliorer ses fonctionnalités en fonction des applications souhaitées.

Références

- Furukawa, Y., Ponce, J., 2010. Accurate, dense, and robust multi-view stereopsis. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32 (8), 1362–1376, PMVS est accessible à l'adresse : <http://grail.cs.washington.edu/software/pmvs/>.
- Pierrot-Deseilligny, M., 2007. MicMac, un logiciel pour la mise en correspondance automatique d'images dans le contexte géographique. *Bulletin d'Information Scientifique et Technique de l'IGN* 77, www.micmac.ign.fr.
- Snavely, N., Seitz, S., Szeliski, R., 2008. Modeling the world from internet photo collections. *International Journal of Computer Vision* 89 (2), 189–210.
- Vedaldi, A., Fulkerson, B., 2008. VLFeat: An open and portable library of computer vision algorithms. <http://www.vlfeat.org/>.
- Vergauwen, M., Van Gool, L., 2006. Web-Based 3D reconstruction service. *Machine Vision Applications* 17, 411–426, Arc3D est accessible à l'adresse : <http://homes.esat.kuleuven.be/~visit3D/webservice/v2/>.