



Développement d'une solution de numérisation 3D temps réel via smartphone

■ Raphaël HAENEL

À l'heure où les smartphones et tablettes se démocratisent en tant qu'outil de travail, le monde des travaux publics se modernise. Simplifier le quotidien des chefs de chantier, est l'objectif que s'est fixé l'entreprise SYSLOR. Dans son objectif d'amélioration continue sur le thème de la création de jumeaux numériques, la start-up Messine a souhaité cristalliser ses recherches autour du développement d'une solution algorithmique privilégiant la rapidité et l'autocontrôle. Sur la base de photos RGB-D géoréférencées à l'aide de l'antenne GNSS conçue par SYSLOR, cette étude propose une approche calculatoire innovante permettant de concevoir un nuage de points très rapidement.

■ MOTS-CLÉS

Numérisation 3D, carte de profondeurs, GNSS, smartphone.

- une optimisation du processus de calcul pour atteindre une productivité en temps réel ;
- une accessibilité de l'acquisition pour tout un chacun.

État de l'art

Avec le développement des technologies, la reconstruction 3D connaît un essor fulgurant depuis ces dernières années. En effet, que ce soit par le biais d'une stéréo binoculaire ou plus généralement multivues, la reconstruction 3D est devenue un point de recherche prédominant à travers différents domaines d'utilisation. Dans notre cas, la reconstruction d'une scène en temps réel est souvent associée au domaine de la vision par ordinateur et de l'apprentissage machine qui représente deux secteurs évoluant très rapidement, notamment avec un apport très régulier de nouveaux algorithmes de plus en plus performants.

Pour être en mesure de pouvoir correctement synthétiser les différentes approches de reconstruction, il a été nécessaire, comme l'ont fait [Seitz et al., 2006], de définir des critères d'évaluation et de comparaison des méthodes. Tout d'abord, il est possible d'envisager une différenciation des approches en matière de représentation de la scène, puisque la géométrie d'un objet peut aisément être décrite selon :

- une grille d'occupation voxel (équivalent 3D du pixel), utilisée pour sa simplicité et sa faculté à se rapprocher de n'importe quelle surface ;
- un maillage polygonal (généralement triangulaire), très populaire pour des questions de simplicité de stockage et de praticité de visualisation des données ;

Contexte de l'étude

D'après les chiffres fournis par la FNEDRE (Fédération nationale des entreprises de détection de réseaux enterrés) pas moins de 100 000 accidents sont recensés chaque année en France, lors de travaux d'excavation, à cause d'une localisation imprécise des réseaux ou d'une mauvaise lecture de plans. C'est en ce sens que SYSLOR a développé ses solutions pour offrir aux acteurs du chantier une visualisation efficiente des données (en réalité augmentée notamment) et ainsi minimiser les risques d'accident, notamment en comprenant mieux la disposition des réseaux existants. Outre la visualisation, l'une des applications SYSLOR offre également la possibilité aux chefs de chantier d'effectuer eux-mêmes l'acquisition des tranchées, par voie photographique ou vidéo. L'objectif est de favoriser l'autonomie du chef de chantier pour des travaux simples et de réaliser plus rapidement le remblaiement des tranchées sur site. La problématique rencontrée est que le récolement des réseaux enterrés repose sur une procédure photogramétrique usuelle, qui bien qu'étant tout à fait légitime et fonctionnelle, se trouve être chronophage. Le temps d'envoi des données de même que le temps de calcul sont significatifs, et

constituent un réel frein pour ces utilisateurs. Ainsi, SYSLOR envisage de pallier cette problématique en développant une solution innovante de numérisation 3D des réseaux enterrés sous des contraintes temporelles qui viendront qualifier le système de "temps réel". Cette démarche de développement offrirait la possibilité de réduire drastiquement les risques relatifs à une mauvaise acquisition pouvant rendre le récolement de la tranchée complexe, voire irréalisable. En effet, la finalité affichée est de pouvoir proposer à un chef de chantier une solution lui permettant d'avoir un suivi permanent de la digitalisation pour qu'il soit en mesure de juger immédiatement de l'exhaustivité des données et de la qualité du modèle généré. Il pourrait en particulier directement procéder au remblaiement de la tranchée. Cela induit implicitement que le processus d'acquisition doit être relativement accessible pour un utilisateur non aguerri.

En résumé, les développements mis en œuvre au cours de ce projet ont dû être mis en adéquation avec un cahier des charges défini par les contraintes suivantes :

- une fiabilité calculatoire du processus soutenue efficacement par les performances offertes par les smartphones actuels ;





- une carte de profondeurs qui permet de se soustraire à un éventuel rééchantillonnage de la géométrie 3D.

De manière plus générale, il est également possible de subdiviser les approches de reconstruction en quatre classes selon la manière dont elles sont entreprises :

1^{re} classe : extraction et mise en correspondance de points caractéristiques qui serviront à ajuster une surface aux entités générées ;

2^e classe : calcul d'une fonction de coût associée à un volume 3D, puis extraction d'une surface à partir de celui-ci ;

3^e classe : évolution récurrente d'une surface pour minimiser une fonction de coût ;

4^e classe : calcul d'un ensemble de cartes de profondeurs.

Nous avons tenté, dans le cadre de notre étude, de synthétiser de manière la plus exhaustive possible les solutions existantes dans la littérature scientifique en couvrant tous les horizons possibles. Le cahier des charges défini dans le cadre de la problématique de notre sujet nous a permis de hiérarchiser les différentes méthodes recensées. Cette structuration d'intérêt a pour mission d'extraire les approches qui semblent être les plus intéressantes au regard de notre thématique. Ainsi, nous saurons vers où orienter nos futurs développements. Le récapitulatif de la mise en coïncidence

entre les méthodes étudiées et lesdits critères est présenté dans le *tableau 1*. Parmi l'ensemble des algorithmes que l'on retrouve dans la littérature scientifique, certains ont d'ores et déjà été implémentés sur smartphones et tablettes. Ce constat est rassurant et confirme qu'un traitement 3D est potentiellement embarquable sur ce type d'appareil. Néanmoins, la principale inconnue qui subsiste réside dans l'adaptabilité de ces applications dans notre contexte de numérisation des réseaux enterrés. En effet, jusqu'ici nous ne pouvons que conjecturer que ces solutions peuvent correctement s'accorder à ces objets linéaires particuliers. Nous avons décelé au cours de cette étude bibliographique l'intérêt majeur que portent les cartes de profondeurs pour la compréhension générale d'une scène. Ainsi, au vu de la synthèse générale menée, la méthode du "planesweep" et l'API d'ARCore semblent être tout à fait pertinentes pour la génération de cartes de profondeurs. De même, la méthode "marching cubes" paraît convenir pour une fusion adéquate des données.

Production des données de profondeur

Une carte de profondeurs, ou plus communément en anglais "depth map", est une image où chacun des pixels ne va pas contenir une information de



Figure 1. Illustration d'une carte de profondeurs.

couleurs, mais une donnée relative à la distance des surfaces des objets d'une scène à partir d'un point de vue particulier (figure 1). Elle caractérise une pseudo-reconstruction de l'espace qui ne représente pas directement une surface, mais plutôt un échantillon de valeurs discrètes qui permet de comprendre la géométrie de l'environnement capturée selon une certaine position de caméra.

La production de cartes de profondeurs peut provenir de l'utilisation de capteurs spécifiques LiDAR ou temps de vol (ToF). Mais, en l'absence de tels capteurs, ces données peuvent également provenir de l'étude d'une similitude colorimétrique, que l'on désigne aussi par critère de photo-consistance, entre les pixels d'une paire stéréoscopique. Différentes approches existent selon la méthode utilisée.

L'implémentation de la méthode du "planesweep" n'ayant pas abouti à l'obtention de résultats suffisamment pertinents pour une utilisation en temps réel, nous nous sommes très vite orientés vers l'API disponible sous ARCore.

API Depth d'ARCore

Le développement des technologies mobiles, notamment en matière de capteurs d'odométrie ou encore de SLAM (*Simultaneous Localization and Mapping*) a permis l'émergence et la démocratisation des applications de réalité augmentée. Les services Google, souhaitant toujours offrir une meilleure expérience à ses utilisateurs, ont développé une API intégrant les mesures d'une carte de profondeurs issue d'une caméra monoculaire [Valentin et al., 2019] à une cadence de 30 Hz. Ces données particulières permettent notamment de gérer les effets d'occlusion entre le monde réel et les éléments

	Temps réel	Optimisation des calculs	Adaptabilité au contexte des réseaux enterrés	Qualité de la reconstitution
3D-R2N2	*	*	*	*
Depth API ARCore	***	***	**	**
DORN	*	*	*	*
GEOMetrics	*	*	*	***
KinectFusion	***	***	**	*
Marching cubes	**	**	*	***
Mesh R-CNN	*	*	**	***
MobileFusion	*	*	**	***
MonoFusion	*	*	**	***
PatchMatch	***	**	*	*
Pix2Mesh	*	*	**	***
Planesweep	***	***	**	***
MVSNet	*	*	**	***
Raycasting	***	***	*	**
RoutedFusion	**	*	**	**
SurfelMeshing	*	**	**	***
Tétraédisation de Delaunay	*	**	*	*
TGV	**	**	*	**
Visual hull constraints	***	***	*	**

* : Mauvaise cohérence, ** : concordance acceptable, *** : adéquation particulière aux critères

Tableau 1. Mise en correspondance entre les méthodes présentées et les critères liés à notre étude.

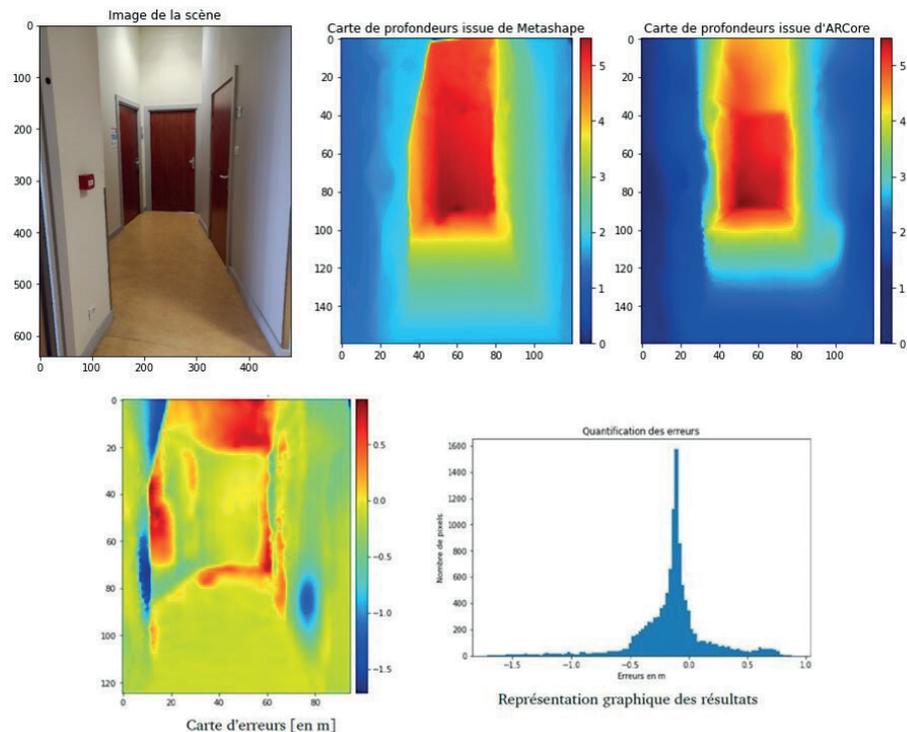


Figure 2. Analyse des résultats établis à partir d'ARCore.

virtuels. Cependant, il conviendra de noter que l'utilisation de cette application est soumise à une spécification d'ordre technique : le smartphone doit supporter ARCore.

Afin de comprendre comment sont générées les cartes de profondeurs de Google, il est nécessaire de s'intéresser au principe de fonctionnement global d'ARCore. Cette application repose sur trois concepts fondamentaux : le suivi de mouvement, la compréhension de l'environnement ainsi que l'estimation de la lumière.

À partir d'une approche similaire au SLAM conventionnel, des points caractéristiques sont détectés tout au long du déplacement afin de situer le smartphone dans l'espace. Grâce à une méthodologie analogue à RANSAC, ces points serviront ensuite à déterminer des plans moyens qui vont assurer la compréhension de l'environnement alentour. Ces différentes informations vont ensuite permettre de construire une carte de profondeurs à partir de l'optimisation hybride de deux algorithmes PatchMatch [Bleyer et al., 2011] et HashMatch [Fanello et al., 2017]. L'idée est d'alterner entre la génération d'une valeur de profondeur isolée (associée à une probabilité d'apparition), et sa propagation aux pixels avoisinants sous

l'hypothèse qu'une image est constituée de régions de profondeur constante. Comme pour la première méthode, nous avons analysé l'exactitude des résultats obtenus à partir de cartes de profondeurs extraites depuis le logiciel commercial Agisoft Metashape. Se basant uniquement sur des informations colorimétriques pour déterminer les valeurs de profondeur, une scène avec des surfaces homogènes et relativement peu texturées va induire des résultats moins précis. Quand bien même, notre étude a permis

de mettre en lumière des résultats très intéressants (figure 2) avec une bonne concordance des cartes. L'erreur constatée est de l'ordre d'une dizaine de centimètres pour un point de vue particulier.

Finalement, au terme de cette étude, nous sommes arrivés à la conclusion que l'API fourni par ARCore répond plus favorablement à nos attentes tant au niveau de la qualité des résultats que la cadence de production des données. C'est donc cette méthode qui a été privilégiée dans le cadre de nos développements.

Présentation de la preuve de concept développée

Grâce à l'intégralité des recherches menées, nous avons été à même de concevoir une chaîne de traitement optimale qui essaye d'intégrer simultanément les souhaits de l'entreprise, mais également les contraintes matérielles liées à l'utilisation d'un smartphone. Concrètement, la preuve de concept se décompose en six étapes majeures scindées en deux parties : la création des données 3D sous la forme d'un nuage de points (figure 3) et l'assemblage en un unique modèle (figure 6). En effet, chaque point de vue correspondant à une image, un unique nuage de points lui sera associé. Comme l'acquisition comptera plusieurs images, il conviendra alors d'assembler l'ensemble des nuages

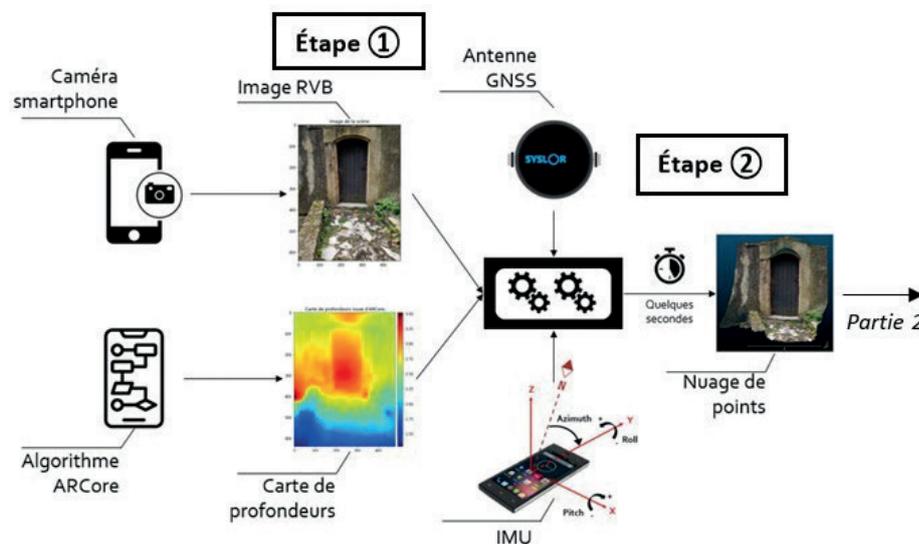


Figure 3. Organigramme de la première partie : création des données 3D.

entre eux tout en gérant les problématiques liées à la fusion.

■ Création des données 3D

Étape 1 : Cette première phase consiste à rassembler l'ensemble des données nécessaires au futur calcul. La première source de données provient de l'acquisition des données RGB-D, autrement dit, l'association entre l'image couleur extraite depuis la caméra d'un smartphone et la carte de profondeurs associée à la vue calculée à partir de l'API d'ARCore.

Pour géoréférencer directement le nuage de points, nous avons utilisé l'antenne GNSS SYSLOR, qui allie ergonomie et efficacité. Elle permet de géoréférencer les images avec une précision de l'ordre de 2 à 5 cm. Néanmoins, ce géoréférencement caractérise véritablement l'antenne GNSS et non pas directement la caméra du smartphone. Ce faisant, une transformation a été mise en œuvre afin de passer efficacement des coordonnées du centre de l'antenne aux coordonnées au centre de l'optique via des constantes appareils (figure 4).



Figure 4. Antenne GNSS SYSLOR.

À cela, nous venons ajouter des mesures angulaires produites par les capteurs inertiels présents dans un smartphone. Grâce à la combinaison d'un accéléromètre trois axes, d'un gyroscope ainsi que d'un magnétomètre, nous pouvons décrire l'orientation du smartphone dans l'espace au moyen d'une représentation eulérienne des angles, à savoir tangage, roulis et azimut.

Étape 2 : Une fois toutes les données nécessaires à notre disposition, il est possible de constituer un nuage de points géoréférencé à l'échelle. Il faut

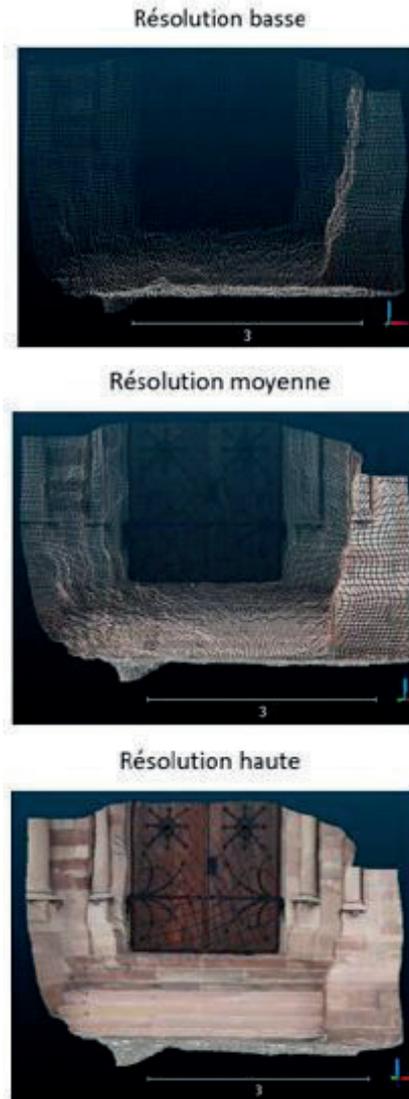


Figure 5. Impact de la résolution des images RGB-D sur le nuage créé.

savoir que les caractéristiques de ces nuages sont tributaires de la résolution des images RGB-D, qui varie généralement entre 160x120 et 640x480 pixels. Une résolution élevée va induire une

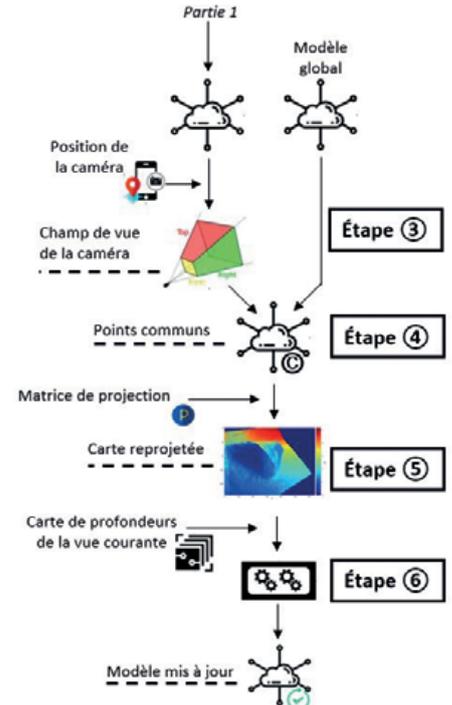


Figure 6. Organigramme de la seconde partie : assemblage en un unique modèle.

plus grande densité du nuage puisqu'à chaque pixel, sera attribué une projection 3D. Cela va induire une bien meilleure représentation visuelle des données avec un détail plus fin des éléments constitutifs d'une scène telle que représentée en figure 5.

Cependant, le temps de production se verra légèrement rallongé. Un compromis devra ainsi être trouvé afin de combiner un temps de traitement optimal et une représentation visuelle efficace. Sur l'aspect qualitatif, les modèles créés avec cette méthode sont intéressants puisque nous obtenons un écart moyen de 7 cm avec un écart-type de 6 cm par rapport à un modèle généré

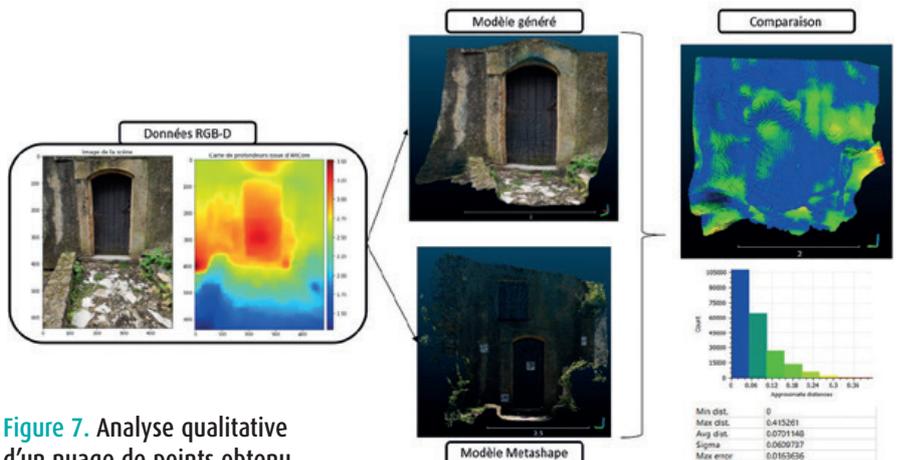


Figure 7. Analyse qualitative d'un nuage de points obtenu.

depuis Metashape (figure 7). Pour une première esquisse de résultat, ce dernier est assez prometteur d'autant plus que la génération de ce modèle s'effectue quasiment en temps réel.

■ Assemblage des données en un unique modèle

Étape 3 et 4 : Le recouvrement des images induit nécessairement une détermination multiple de certains points. Ainsi, l'objectif de ces étapes est d'extraire ces points particuliers afin d'être mis à jour par la suite. Pour cela, nous commençons par calculer le champ de vision associé à la caméra (*viewing frustum* en anglais). C'est une entité géométrique qui va contenir en son sein tous les éléments 3D observables depuis la caméra. Ensuite, l'étape 4 va tout simplement consister à réaliser une segmentation topologique des différents points à partir de la géométrie construite durant l'étape précédente.

Étape 5 : Avant de passer au processus de mise à jour en tant que tel, il est nécessaire de mettre en place l'environnement de comparaison. Un environnement 2D a été privilégié, car le traitement est moins complexe. Ainsi, les points extraits depuis l'étape 4 sont reprojétés sur le plan image de la vue physique en cours formant ainsi une carte reprojétée aussi appelée pseudo-carte de profondeurs.

Étape 6 : La mise à jour du modèle à construire dépend de la comparaison pixel à pixel entre la pseudo-carte de profondeurs et la carte associée à la vue courante. Il sera en particulier possible de venir compléter le modèle en cours de formation lorsqu'une zone inédite de la scène vient d'être couverte. Puis, dans les zones de redondance, des modifications seront entreprises selon différentes possibilités : mesure qui obstrue le modèle, conflit de visibilité ou encore raffinement du modèle lorsque les profondeurs associées aux deux cartes coïncident.

■ Exposition de résultats

En l'absence de chantier avec des réseaux enterrés visibles, nous avons cherché à concevoir une preuve de concept qui se veut généralisable à plusieurs environnements. La portée

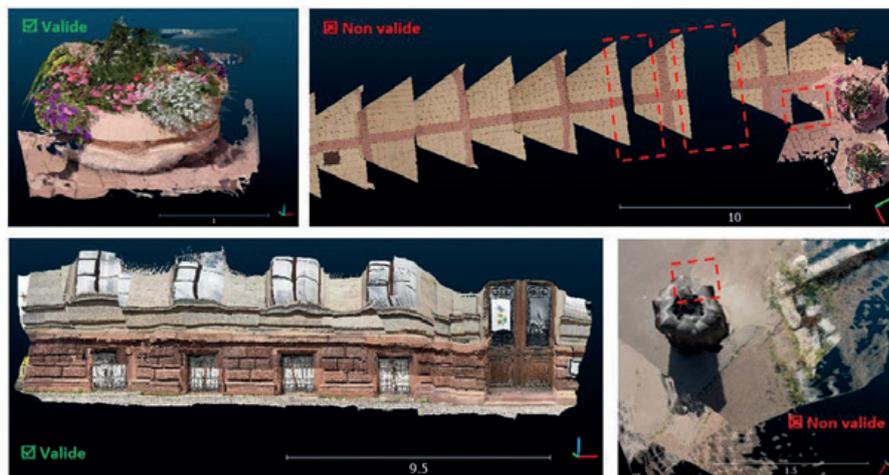


Figure 8. Modèles reconstruits témoignant de la qualité de l'acquisition.

d'acquisition devra néanmoins se limiter à 3-4 m. En effet, l'erreur d'estimation des données de profondeur croît quadratiquement avec la distance, ainsi au-delà de ce seuil, des déformations non négligeables sont observables dans le modèle 3D.

Quelques exemples de modèles reconstruits sont fournis en figure 8. Il apparaît clairement que l'exhaustivité des données dépend implicitement de la qualité de l'acquisition qui est bien illustrée sur l'ensemble des modèles permettant ainsi assez aisément de valider ou non le récolement effectué. Ce constat illustre les objectifs remplis au cours de cette étude. Néanmoins, pour l'heure, la productivité de tels modèles ne satisfait pas encore la notion de temps réel. De plus, des défauts persistent dans notre preuve de concept. En effet, on peut observer sur chaque nuage de points, des déformations importantes ou encore un assemblage approximatif. Ces deux phénomènes ont deux origines principales. D'une part, l'imprécision relative à la détermination des cartes de profondeurs causée soit par un suivi de mouvement pas assez appuyé ou une surface non propice à l'obtention de résultats viables. D'autre part, une dérive des capteurs inertiels du smartphone. Bien que ces derniers fournissent des données angulaires assez intéressantes, elles ne sont pas suffisamment robustes. Il conviendra donc, pour finaliser cette chaîne de traitement, d'améliorer ces points de même qu'optimiser certains traitements encore trop chronophages pour un applicatif en temps réel.

Conclusion

Concevoir une preuve de concept n'est pas chose aisée, encore moins lorsque l'on dispose de contraintes matérielles assez importantes. Ce projet de fin d'études a été l'occasion de se confronter à une problématique d'envergure qui traduit les ambitions d'innovation de start-up comme SYSLOR. Grâce à toutes les recherches et les essais menés, nous avons abouti à une première chaîne de traitement qui fonctionne. Globalement, les objectifs initiaux de ce projet ont été remplis puisque nous sommes capables de fournir un modèle 3D satisfaisant qui illustre de manière assez pertinente l'exhaustivité des données et atteste donc de la conformité de l'acquisition. Bien que le temps réel n'ait pas encore été concrètement atteint, la démarche pour y parvenir a été bien esquissée.

Perfectible, cette preuve de concept ne constitue que les prémices d'une future application clé en main qui verra le jour avec l'apport de matériels plus performants. Par matériel, nous entendons en particulier un capteur de profondeur qui, combiné aux mesures GNSS et aux mesures inertielles du smartphone, viendra annuler la dérive angulaire grâce à un recalage via l'environnement mesuré, mais également une meilleure description de la scène avec des données de profondeur plus précises.

Remerciements

Je souhaite remercier l'entreprise SYSLOR et plus particulièrement

▶ l'ensemble de ses collaborateurs, de m'avoir accordé leur confiance pour mener à bien ce projet d'envergure. Malgré la problématique de taille, nous avons réussi, grâce à un travail collaboratif, à esquisser de nouveaux outils de récolement qui, je l'espère, pourront être utilisés dans le futur. ●

Contact

Raphaël HAENEL
raphael.haenel@sylor.net

Bibliographie

Bleyer, M., Rhemann, C., and Rother, C. (2011). 'PatchMatch Stereo - Stereo Matching with Slanted Support Windows'. In Proceedings of the British Machine Vision Conference 2011, pages 14.1-14.11, Dundee. British Machine Vision Association. doi : 10.5244/C.25.14

Collins, R. (1996). 'A space-sweep approach to true multi-image matching'. In Proceedings CVPR IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pages 358-363, San Francisco, CA, USA. IEEE. doi : 10.1109/CVPR.1996.517097

Fanello, S. R., Valentin, J., Kowdle, A., Rhemann, C., Tankovich, V., Ciliberto, C., Davidson, P., and Izadi, S. (2017). 'Low Compute and Fully Parallel Computer Vision with HashMatch'. In 2017 IEEE International Conference on Computer Vision (ICCV), pages 3894-3903, Venice. IEEE. doi : 10.1109/ICCV.2017.418

Seitz, S., Curless, B., Diebel, J., Scharstein, D., and Szeliski, R. (2006). 'A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms'. In 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 1 (CVPR'06), pages 519-528, New York, NY, USA. IEEE. doi : 10.1109/CVPR.2006.19

Semler, Q., Mangin, L., Moussaoui, A., and Semin, E. (2019). 'Development of a low-cost centimetric GNSS Positioning solution for android applications'. International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences, XLII-2/W17 :309-314.

Valentin, J., Kowdle, A., Barron, J. T., Wadhwa, N., Dzitsiuk, M., Schoenberg, M., Verma, V., Cszasz, A., Turner, E., Dryanovski, I., Afonso, J., Pascoal, J., Tsotsos, K., Leung, M., Schmidt, M., Guleryuz, O., Khamis, S., Tankovitch, V., Fanello, S., Izadi, S., and Rhemann, C. (2019). 'Depth from motion for smartphone AR'. ACM Transactions on Graphics, 37(6) :1-19. doi: 10.1145/3272127.3275041

ABSTRACT

Key words: 3D scanning, depth map, GNSS, smartphone, on-the-fly reconstruction

When smartphones and tablets are becoming more and more popular as working tools, the world of public works is modernizing. Simplify the daily life of the site managers, is the objective that SYSLOR company has set. In its objective of continuous improvement on the theme of the creation of digital twins, the start-up Metzine wished to crystallise these researches around the development of an algorithmic solution privileging speed and self-control. On the basis of RGB-D photos georeferenced with the GNSS antenna designed by SYSLOR, this study proposes an innovative computational approach allowing to design a point cloud very quickly.

RIEGL VQ-840-G

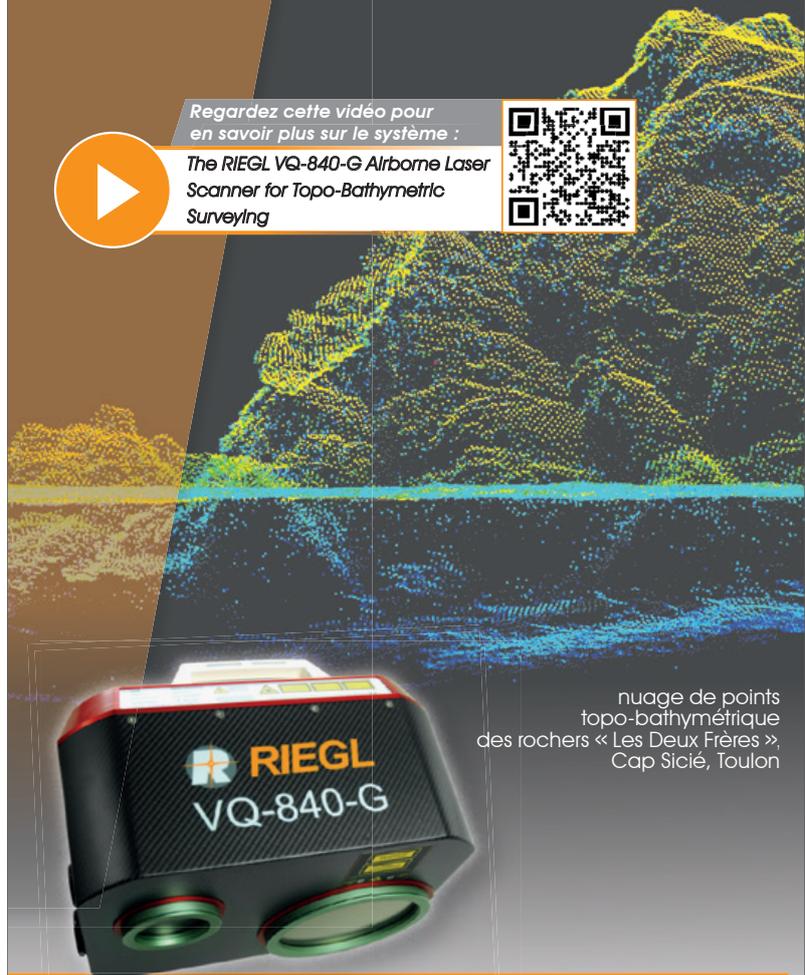
SCANNER LASER TOPO-BATHYMÉTRIQUE

DESTINÉ À L'EXPLOITATION SUR UN DRONE OU SUR TOUT TYPE D'AÉRONEF DÉPLOYÉ À BASSE ALTITUDE

Regardez cette vidéo pour en savoir plus sur le système :



The RIEGL VQ-840-G Airborne Laser Scanner for Topo-Bathymetric Surveying



nuage de points topo-bathymétrique des rochers « Les Deux Frères », Cap Sicié, Toulon

POUR DES LEVÉS DE HAUTE RÉOLUTION DES ZONES LITTORALES OU EN EAUX PEU PROFONDES

- parfaitement adapté à l'intégration sur un drone ou un hélicoptère
- utilisation d'un laser vert permettant de pénétrer dans l'eau jusqu'à plus de 2 profondeurs Secchi
- fréquence de mesure 50 kHz - 200 kHz
- centrale inertielle (INS) et appareil photo numérique intégrés

Visitez-nous à Nice



ISPRS Congress 2022 | Juin 6-11, 2022

Participez à notre présentation consacrée à la bathymétrie dans la session Technology Track : High-Resolution LIDAR Bathymetry from Different Platforms

Découvrez l'ensemble de la gamme de capteurs et systèmes RIEGL LIDAR www.riegl.com

newsroom.riegl.international

