

PRÉ-LOCALISATION DES DONNÉES POUR LA MODÉLISATION 3D DE TUNNELS : DÉVELOPPEMENTS ET ÉVALUATIONS

Christophe Heinkelé¹, Philippe Foucher¹, Emmanuel Moisan^{1,2}, Pierre Charbonnier¹

1: Cerema, Equipe-projet ENDSUM, 11 rue Jean Mentelin BP 9, 67035 Strasbourg

2: Laboratoire Icube UMR7357, INSA Strasbourg, 24 Boulevard de la Victoire, 67084 Strasbourg

Résumé

Le présent article décrit l'implémentation d'une méthode de pré-localisation d'images au sein de grands volumes de données dans des tunnels navigables et routiers. Elle repose sur une technique d'odométrie visuelle simplifiée, rapide et facile à mettre en œuvre et dont la précision est évaluée par comparaison avec des techniques plus conventionnelles. Nous expliquons comment cette méthode permet de former des sous-ensembles d'images se rapportant à une même coupe transversale de l'ouvrage et présentant un fort taux de recouvrement, appelés « tuiles ». Ces sous-ensembles permettent de faciliter la navigation dans les séquences ainsi que les traitements postérieurs, comme par exemple la reconstruction 3D par photogrammétrie.

Mots clés : Odométrie visuelle simplifiée, structuration des données, reconstruction 3D dans des ouvrages oblongues, évaluation quantitative des méthodes

Abstract

This paper focuses on the implementation of a method for pre-locating large volumes of images acquired in navigable and road tunnels. It is based on a simplified visual odometry technique that is fast and easy to implement and whose accuracy is assessed by comparison with more conventional techniques. We explain how this method allows forming subsets of images relating to the same cross-section of the structure and having a strong overlap, called "tiles". These subsets facilitate sequence navigation and subsequent processing, such as 3D photogrammetric reconstruction.

Keywords : Simplified visual odometry, data structuration, 3D-Reconstruction in oblong infrastructures, assessment

1. Introduction

Dans cet article, nous nous intéressons à la mise en œuvre et à l'évaluation de méthodes de positionnement d'images acquises dans des tunnels fluviaux ou routiers. Les approches proposées permettent la localisation des images collectées en tunnel et nous en présentons deux applications directes : la visualisation des données et la modélisation 3D.

Ces travaux s'inscrivent dans un projet plus global, qui vise à mettre au point des méthodes rapides et automatisées d'aide à l'inspection de tunnels routiers ou navigables, en s'appuyant au maximum sur des informations collectées dans des images. Le premier besoin pour le gestionnaire d'un tunnel est de disposer d'un outil de visualisation de l'ensemble de la voûte au moyen d'images localisées, voire sous la forme d'un modèle 3D. Cela lui permet, par exemple, d'effectuer des relevés d'objets d'intérêt afin de mieux suivre l'état de son infrastructure.

Nous disposons d'un prototype d'acquisition dynamique de séquences d'images, qui nous permet de collecter un corpus d'images de l'ensemble de la voûte de chaque tunnel. Pour des raisons pratiques, ce système nécessite plusieurs passages dans l'ouvrage afin de réaliser l'acquisition complète du tunnel. Nous proposons de regrouper les images disponibles en sous-ensembles spatialement cohérents, que nous appelle-

rons « tuiles ». Chaque tuile représente une sélection des images se rapportant à une même coupe transversale du tunnel, d'un piédroit à l'autre, et présentant un maximum de recouvrement entre-elles. Cet ordonnancement facilite la navigation dans les séquences et l'exploitation des images. L'enjeu de cet article est de proposer une méthodologie adaptée pour constituer ces tuiles dans un temps raisonnable.

Dans le cas des tunnels, le positionnement des données ne peut être obtenu à partir de signaux GPS. Nous ne disposons pas de centrales inertielle ou d'un autre système de repérage dans l'espace. Nous nous orientons donc vers des méthodes de positionnement basées sur l'image, *i.e.* l'odométrie visuelle. Dans les tunnels routiers, les odomètres mécaniques peuvent être facilement intégrés au système d'acquisition pour localiser les données. Pour les tunnels navigables, ces derniers ne peuvent pas être mis en œuvre, l'acquisition se faisant à partir de capteurs montés sur des bateaux. Nous proposons ici une méthode de localisation des données principalement basée sur l'image. Nous nous plaçons donc dans le contexte des techniques de *Structure from Motion* (SfM) (Szeliski, 2010). Plus précisément, nous proposons d'introduire une étape en amont, dans laquelle les images de plusieurs séquences ou issues de caméras différentes sont ordonnées et localisées le long de l'ouvrage.

La première contribution de cet article porte sur l'évaluation d'une méthode particulière d'odométrie visuelle, dite « odométrie visuelle simplifiée » (Albert et al., 2013; Charbonnier et al., 2014) qui permet d'assigner une donnée de localisation à chaque image d'une caméra de référence sur le porteur. Une part importante de ce travail concerne l'évaluation systématique de la méthode de pré-localisation sur une série d'expérimentations *in-situ*. La deuxième contribution est la proposition d'une méthode de ré-indexation de l'ensemble des images à partir du positionnement de cette caméra de référence pour former des tuiles, ainsi que son évaluation.

L'article est organisé de la façon suivante : La section 2 positionne le travail vis-à-vis de l'état de l'art dans le domaine. Dans la section 3, nous décrivons la méthodologie générale de notre démarche en détaillant notamment les méthodes de pré-positionnement. Dans la section 4, les résultats sont présentés et commentés. Quelques cas d'applications utilisant ces méthodes de pré-localisation seront présentés dans la section 5. Enfin, dans la section 6, nous concluons par quelques améliorations de la méthode.

2. État de l'art

Pour structurer et sélectionner les sous-ensembles d'images, deux options sont envisageables. La première consiste à utiliser l'horodatage des images et à identifier les groupes d'images acquis dans un intervalle de temps défini. Cependant, cette approche n'est pas pertinente, puisque des images d'une même portion de voûte de tunnel ne sont pas obligatoirement issues d'une même séquence d'acquisition et que la vitesse du système imageur varie d'une séquence à une autre. La seconde option, que nous avons retenue, consiste à regrouper les images selon leur localisation dans l'infrastructure, obtenue par odométrie visuelle.

L'orientation et l'alignement des images pour la reconstruction photogrammétrique d'un modèle 3D sont réalisés en mettant en correspondance des caractéristiques extraites dans chaque image. Lorsque le nombre d'images utilisées pour reconstruire le modèle est élevé, une recherche exhaustive de tous les correspondants possibles entre toutes les images entraîne rapidement des coûts prohibitifs en termes de temps de calcul et de ressource mémoire. Il est donc nécessaire de limiter les appariements aux quelques images qui ont effectivement un recouvrement entre elles. En fonction de l'information disponible sur les séries d'images obtenues, deux situations se présentent (Havlena, 2012) : aucune information d'ordonnancement n'est disponible pour les images (*unordered datasets*) ou les images peuvent être disposées dans un certain ordre grâce à une information spatiale ou séquentielle (*ordered datasets*).

Dans la première situation, la plus défavorable, il n'existe aucune information spatiale ou séquentielle permettant de situer les images les unes par rapport aux autres. Dans ce cas, les méthodes de structuration des

caractéristiques permettent de limiter les temps de calcul. Ainsi, dans (Brown et Lowe, 2003), pour reconstruire automatiquement des panoramas, chaque caractéristique est mise en correspondance avec ses k plus proches voisines grâce à un algorithme de type *kd-tree* (Beis et Lowe, 1997). Dans (Brown et Lowe, 2007), l'algorithme de reconstruction de panoramas est amélioré grâce à l'introduction d'une méthode de compensation du gain photométrique et d'étapes pour redresser les panoramas. Proche de notre application, ce type de techniques est utilisé dans (Bauer et al., 2015) pour modéliser la surface de la voûte d'un tunnel pendant sa phase de construction. Cependant, dans (Wang et al., 2017), les auteurs indiquent que le coût de calcul des méthodes de recherche de plus proches voisins à partir de *kd-tree* ou *kd-forest* peut devenir très important lorsque le nombre d'images ou de caractéristiques augmente significativement. Des stratégies sont possibles pour diminuer ces coûts : la réduction de la taille des images ou la limitation du nombre de caractéristiques, mais les performances peuvent s'en ressentir.

Dans la mesure du possible, l'introduction d'un *a priori* pour ordonner les images entre elles permet de limiter la recherche d'appariements aux quelques images présentant un recouvrement. Cet *a priori* peut être, par exemple, la position GPS des caméras pendant les acquisitions (Wu, 2013) ou l'ordre des images dans les séquences vidéo (Akbarzadeh et al., 2006; Cornelis et al., 2006). Dans (Cannelle, 2013), l'auteur se limite aux appariements d'images distantes de moins de 30 m entre elles, distance obtenue par des techniques de localisation par GPS. Dans nos expérimentations, les images sont acquises au cours de différents passages. L'information d'horodatage des images dans la séquence ne peut être exploitée seule pour définir l'ordonnancement. Les techniques d'odométrie visuelle, décrites dans (Scaramuzza et Fraundorfer, 2011; Paudel et al., 2014) consistent à relier des informations de capteurs 2D ou 3D vers des positionnements 2D ou 3D. L'odométrie visuelle peut être considérée comme un cas particulier de SfM (*Structure from Motion*).

Dans (Aqel et al., 2016), reprenant (Scaramuzza et Fraundorfer, 2011), les auteurs distinguent trois familles de méthodes : l'approche par détection de caractéristiques (*feature-based approach*), l'approche directe (*appearance-based approach*) et les méthodes hybrides. La première famille d'approche repose sur une extraction de caractéristiques dans les images, suivie d'une recherche d'appariements entre les images en fonction des caractéristiques détectées. Ce type d'approche est sensible à la texture des surfaces analysées. Dans des situations où l'image est faiblement texturée, par exemple les parois en béton, la détection des points d'appariement est parfois compliquée et peut engendrer des performances plus faibles pour l'odométrie. La seconde famille, dite directe, ne se heurte pas à ces difficultés. Elle se fonde sur les variations d'intensité des pixels entre deux images pour estimer le mouvement. Parmi les différentes approches directes, on retiendra celles utilisant la corrél-

lation croisée pour déterminer le décalage entre images. Comparée à la première famille de méthodes, l'approche directe présente le défaut d'être moins rapide (Scaramuzza et Fraundorfer, 2011). Cependant, il est possible de s'appuyer sur les propriétés de la transformée de Fourier pour accélérer les calculs de la corrélation croisée (Goshtasby et al., 1984). Dans le même ordre d'idée, le décalage observé entre deux images se traduit dans l'espace de Fourier par une différence de phase (Froosh et al., 2002). Cette technique de corrélation de phase est utilisée dans (Birem et al., 2018) pour réaliser une odométrie visuelle en (X,Y,Z). Cependant les auteurs constatent que cette méthode souffre d'un manque de robustesse lorsque la profondeur est grande. Dans notre application, cette situation n'est rencontrée que dans des cas rares (présence d'une galerie annexe). Enfin, les méthodes hybrides couplent les deux approches pour gagner en robustesse, comme dans (Scaramuzza et Siegwart, 2008), où les caractéristiques au sol servent à déterminer les translations et une méthode directe sur l'ensemble de la scène permet de déterminer les rotations du véhicule.

Dans la suite de cette contribution, nous proposons de mettre en œuvre la méthode d'odométrie visuelle simplifiée proposée dans (Albert et al., 2013; Charbonnier et al., 2014). Il s'agit d'une méthode d'odométrie visuelle 2D-2D, qui repose sur un calcul direct des translations entre deux images successives. L'orientation des capteurs ne peut être obtenue par cette approche. Nous évaluerons les performances de l'odométrie visuelle simplifiée en utilisant l'approche directe d'une part, et par détection de caractéristiques d'autre part.

3. Chaîne de traitements

La pré-localisation des images est faite ici de manière longitudinale au tunnel, dans un cadre mono-dimensionnel. Elle se fonde sur la notion d'abscisse curviligne, qui correspond à la distance de parcours mesurée, depuis une origine arbitraire, le long d'une courbe donnée. Pour une courbe finie de longueur L , elle est comprise entre 0 et L . L'abscisse curviligne accessible par odométrie (visuelle ou mécanique) est liée à la trajectoire du porteur au moment de l'acquisition. On supposera que celle-ci est maintenue aussi parallèle que possible au tracé du tunnel. Par ailleurs, nous considérerons par la suite les tunnels rectilignes ou à fort rayon de courbure, ce qui correspond à la quasi-totalité des tunnels fluviaux et à de nombreux tunnels routiers. Ainsi, toutes les trajectoires sont supposées rectilignes et parallèles à l'axe du tunnel et il suffit de recalculer les positions mesurées par rapport aux entrées (abscisse 0) et sortie (abscisse L) de ce dernier. Notons que, pour gommer les imperfections liées au non-respect de ces hypothèses simplificatrices, il est usuel de recalculer les positions par rapport à des repères implantés dans le tunnel, le plus souvent tous les 10 m.

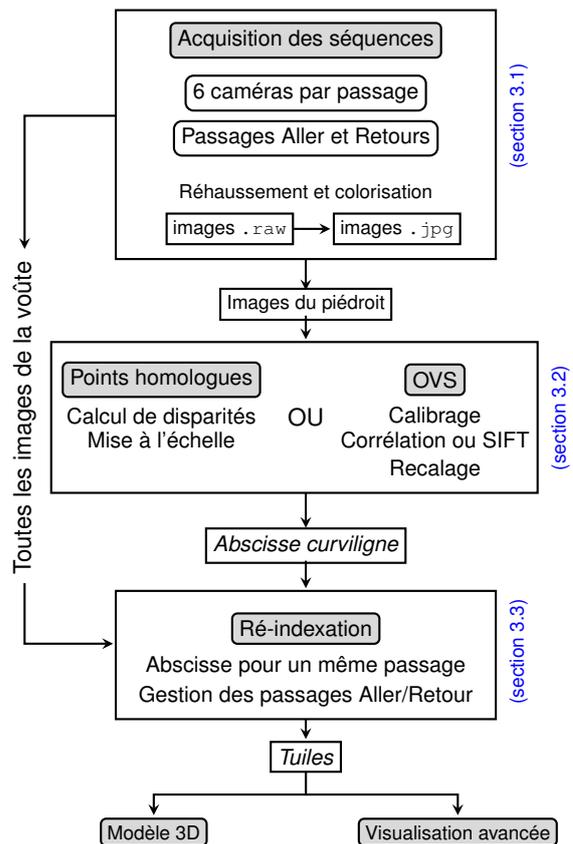


FIGURE 1: Schéma de la méthode globale expliquant l'enchaînement des différentes étapes. Les cases grisées indiquent l'intervention d'un opérateur.

La démarche générale de la méthode est représentée sur la figure 1, qui en illustre les trois étapes. La première consiste à acquérir les images au cours de campagnes expérimentales, puis à les rehausser pour faciliter les calculs postérieurs (voir sous-section 3.1). La seconde étape vise à pré-positionner en abscisse curviligne les images issues d'une caméra de référence. Ce pré-positionnement est d'abord calculé sur les images du piédroit (sous-section 3.2). Enfin la troisième étape consiste à reporter cette information à toutes les autres images en les ré-indexant par rapport à la caméra déjà pré-positionnée (sous-section 3.3) pour les organiser en tuiles.

3.1. Acquisitions des séquences

3.1.1. Système imageur

Le système imageur a été développé pour une application dans un tunnel fluvial, mais il a également été testé dans des tunnels routiers, la différence résidant dans le porteur. Les images à notre disposition ont ainsi été collectées sous forme de séquences dans des tunnels canaux ou routiers.

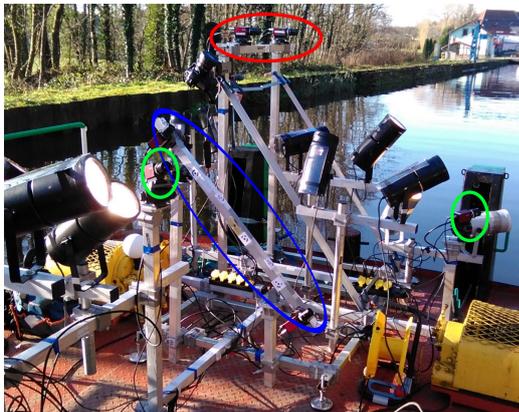
Le système imageur est modulaire et composé de 6 caméras montées sur un véhicule adapté au type de tunnel étudié (barge pour les tunnels canaux, véhicule léger pour les tunnels routiers). Dix projecteurs orientables¹

1. Lampes à décharge avec ballast électronique de 150W, de type « Elpro Circle 150 ».

sont disposés sur le porteur pour éclairer de manière diffuse la voûte en prenant soin de limiter les ombres portées (voir figure 2). Les caméras industrielles utilisées (Pike F210C d'Allied Vision Tech) sont équipées d'objectifs à focale fixe et ont une résolution de 1920×1080 pixels. Elles peuvent être orientées selon plusieurs directions de manière à collecter des images de l'ensemble de la voûte en plusieurs passages. L'acquisition est déclenchée de manière synchrone pour toutes les caméras selon une fréquence temporelle préalablement fixée. Une séquence comprend l'ensemble des prises de vues horodatées d'une caméra au cours d'un passage. Un passage contient donc six séquences d'images. Pour couvrir l'ensemble de la voûte d'un tunnel, il est nécessaire d'effectuer plusieurs passages dans chaque sens de traversée, comme cela est illustré sur la figure 3. Lors des acquisitions, une trajectoire aussi parallèle au piédroit et une vitesse aussi constante que possible sont maintenues.



(a)



(b)

FIGURE 2: (a) Vue générale du système d'acquisition des données sur le porteur dans le tunnel. (b) Ensemble des caméras du dispositif avec une paire stéréovision orientée vers le piédroit (ellipse rouge), une base stéréovision orientée vers la voûte (ellipse bleue) et 2 caméras en visée oblique (cercles verts).

Dans tous les tests et pour tous les passages, une paire de caméras en stéréovision est spécifiquement orientée vers le piédroit de la structure pour positionner le dispositif par odométrie visuelle (voir figure 3, acquisitions *b* et *c* ou *n* et *o*). Une deuxième paire de caméras

est orientée du rein de voûte jusqu'à la clef selon différents angles, de manière à couvrir l'ensemble de la surface : à chaque passage, cette seconde paire est orientée vers la voûte selon un nouvel angle (vues *e* : *f*, *h* : *g*, *i* : *j* et *k* : *l* de la figure 3). Les deux dernières caméras sont en visée oblique de manière à faciliter une éventuelle étape de reconstruction photogrammétrique (vues *a*, *d*, *m* et *p*).

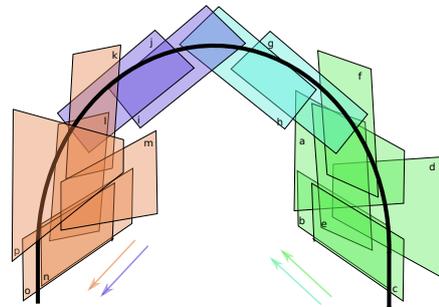


FIGURE 3: Schématisation des passages du système d'imagerie pour couvrir l'intégralité de la voûte. Les deux passages dans un sens (vert et bleu) permettent d'obtenir un côté du tunnel tandis que les passages dans le sens inverse (violet et orange) permettent d'obtenir l'autre côté.

La figure 3 montre que, en un point de l'axe du tunnel, 4 passages permettent d'obtenir une mosaïque complète couvrant la voûte d'un piédroit à l'autre.

3.1.2. Acquisition des images

Grâce à ce dispositif, nous avons pu recueillir un très grand nombre d'images pour chaque tunnel, dont un exemple est montré sur la figure 4. A titre indicatif, pour couvrir la voûte complète d'un tunnel canal d'une longueur de 500 mètres, plus de 50.000 images sont collectées, ces dernières présentant un fort recouvrement entre elles. Pour un intervalle donné en abscisse curviligne dans le tunnel, il y a donc un ensemble d'images (tuiles), issues de différents passages dont les recouvrements successifs permettent de couvrir la totalité de la voûte d'un piédroit à l'autre. Étant donnée la fréquence des acquisitions, il y a toujours un recouvrement entre deux images successives. De plus, la configuration des caméras permet le recouvrement entre deux passages.

3.1.3. Rehaussement et colorisation des images

Du fait des conditions expérimentales, les rendus et la luminosité des images sont très variables dans les tunnels. On note en particulier des problèmes de vignettage ou des hétérogénéités d'éclairage dans les images, ainsi que des changements de luminosité entre images de la séquence. Dans ces deux situations, un rehaussement de la luminosité des images a été réalisé par application d'une correction gamma (γ) et d'un filtre « flat-field » avant colorisation, mis en œuvre dans (Charbonnier et al., 2014). Pour cela, nous réalisons tout d'abord la moyenne globale d'une sélection d'images, en niveaux de gris. Cette moyenne globale de l'image



FIGURE 4: Exemples d'images de piédroit acquises en tunnel routier.

est ensuite appliquée à chaque image de manière à homogénéiser le rendu visuel des images tout au long de la séquence. Quelques précautions doivent être prises cependant dans la sélection des images. La présence d'éléments perturbateurs (câbles ou éléments sombres par exemple) peut fausser le filtre. De même, l'application d'un flou sur l'image moyenne obtenue permet d'éviter de rehausser certains contours, comme ceux orientés selon les maçonneries. Des exemples d'images avant et après traitement sont visibles sur la figure 5 pour deux tunnels. On y constate l'augmentation de la luminosité des images avec des détails beaucoup plus visibles. Comme attendu, nous remarquons également un rehaussement du niveau de bruit, en particulier sur les zones initialement très sombres. Par ailleurs, la luminosité est plus uniforme dans les images ce qui devrait faciliter les traitements pour la reconstruction, par exemple.

Cette phase de pré-traitement n'est pas entièrement automatisée à l'heure actuelle. La détermination des images permettant de constituer le filtre du « flat-field » nécessite l'intervention d'un opérateur, pour sélectionner manuellement les images ne contenant pas d'objets perturbateurs et en évitant les images des extrémités où la lumière naturelle est prépondérante. Il en va de même pour la correction γ qui demande quelques tests avant d'en fixer les paramètres.

Enfin, les images obtenues sont acquises à partir de caméras équipées de matrices Bayer qui sélectionnent pour chaque pixel un des trois canaux colorimétriques. Une phase de colorisation est donc nécessaire de manière à obtenir l'image en couleurs.

3.2. Calcul de l'abscisse curviligne

Au sein d'une même séquence, les images suivent un ordonnancement logique. En accumulant les inter-distances entre vues successives, il est possible de déterminer l'abscisse curviligne de chacune le long de la trajectoire de prise de vue. Ces inter-distances ne sont pas directement délivrées par le système d'acquisition, mais nous pouvons les calculer en utilisant l'information de recouvrement entre images successives. Les abscisses curvilignes ainsi déterminées sont alors recalées en utilisant par exemple l'information de longueur du tunnel.

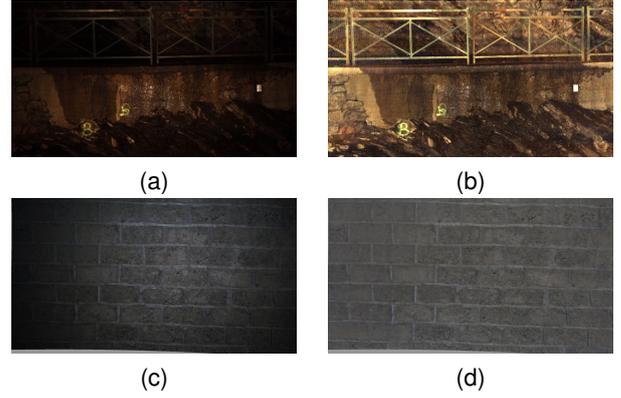


FIGURE 5: Exemples de 2 tunnels avant (gauche) et après (droite) corrections gamma et « flat-field ».

Deux méthodes ont été considérées pour déterminer les distances entre vues successives. Toutes deux utilisent les images acquises par le doublet stéréo visant le piédroit, la caméra de gauche étant prise comme caméra de référence. La première méthode repose uniquement sur un ensemble de points homologues communs à deux images successives de la séquence de la caméra gauche pour estimer la disparité liée au mouvement. La deuxième méthode, appelée Odométrie Visuelle Simplifiée (OVS) (Albert et al., 2013; Charbonnier et al., 2014) exploite la géométrie épipolaire du doublet pour estimer directement la distance en mètres entre acquisitions successives.

3.2.1. Pré-localisation par points homologues

La première méthode de localisation consiste à calculer des points d'intérêt entre deux images successives de la séquence, puis à les mettre en correspondance en associant les descripteurs extraits. Nous cherchons à estimer une translation selon un seul axe entre 2 poses, c'est pourquoi il n'est pas nécessaire de chercher un grand nombre de points homologues. Une fois une centaine de points appariés, la disparité globale entre images, d_M peut être estimée comme la moyenne des inter-distances entre points homologues. En pratique, il est plus robuste de considérer la médiane des inter-distances. Pour ce qui est des outils, nous utilisons le logiciel MicMac (Rupnik et al., 2017) et les points homologues sont extraits à l'aide de Sift++. A faible résolution, cette extraction est très rapide d'une image à l'autre, étant donné que l'on peut utiliser le caractère séquentiel des images.

N_d correspond au nombre d'images de la séquence acquise entre l'entrée et la sortie du tunnel. En appliquant ce procédé aux paires successives d'images de la séquence issue de la caméra de référence, on obtient un vecteur de disparités inter-images $\{d_M^i\}_{i=1\dots N_d-1}$. L'abscisse curviligne S_i associée à l'image N_i dans la séquence est obtenue par une simple mise à l'échelle :

$$S_1 = 0, \quad S_i = \frac{1}{K} \sum_{j=1}^{i-1} d_M^j \quad 2 \leq i \leq N_d \quad (1)$$

où $K = S_{N_d}/L$ est une constante qui permet d'assurer que l'abscisse curviligne de la dernière image correspond bien à la longueur physique L du tunnel.

La méthode par points homologues a l'avantage d'être très directe et ne nécessite pas de calibrage préalable de la caméra de référence. Pour exprimer le résultat dans le système métrique, une simple mise à l'échelle globale est effectuée par rapport à la longueur du tunnel.

3.2.2. Odométrie Visuelle Simplifiée (OVS)

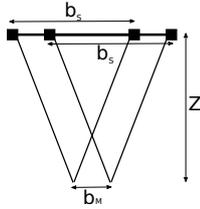


FIGURE 6: Schéma du principe de l'OVS

La seconde méthode permet de calculer directement dans le système métrique l'avancement entre deux images. Elle s'appuie sur un calibrage préalable des paires en stéréovision. Ce calibrage est effectué de manière classique à partir d'une mire plane (Szeliski, 2010) et permet d'extraire les paramètres intrinsèques (Zhang, 1999) et extrinsèques (Tsai, 1987) de la paire de caméras. À partir de ces paramètres, les images sont corrigées de la distorsion optique et une rectification épipolaire est appliquée. Les focales des deux caméras rectifiées sont alors identiques. Le but de l'odométrie visuelle simplifiée est de fournir une abscisse curviligne la plus précise possible à chaque image de la séquence. C'est une méthode directe qui fait l'hypothèse d'une surface plane. La contrainte majeure au bon fonctionnement de la technique réside dans le fait que le mouvement du porteur doit être parallèle à la base stéréovision, hypothèse dans laquelle nous nous plaçons.

Le principe de cette méthode consiste à déterminer l'abscisse d'une caméra par rapport à son abscisse précédente en calculant les mesures de profondeur par stéréovision en utilisant les deux caméras de piedroit de deux manières différentes : d'une part en tant que base stéréoscopique à l'instant d'une acquisition, d'autre part en n'en utilisant qu'une seule, à deux instants différents. En considérant des images rectifiées, la profondeur Z_S est donnée par :

$$Z_S = \beta \frac{b_S}{d_S} \quad (2)$$

où β est la focale exprimée en pixels, b_S est la distance entre les deux centres optiques des caméras (Szeliski, 2010), ou base stéréoscopique, et d_S est la disparité entre images (en pixel). β et b_S ont été estimés grâce au calibrage. La même formule s'applique pour le couple stéréoscopique constitué par deux acquisitions successives d'une des deux caméras (voir figure 6). En appliquant l'équation de la stéréovision rectifiée à cette paire,

la profondeur Z_M est définie par :

$$Z_M = \beta \frac{b_M}{d_M} \quad (3)$$

b_M est la grandeur que l'on désire déterminer, car il s'agit de la distance parcourue par le système d'imagerie entre deux acquisitions. Dans la double hypothèse où le mouvement entre deux images successives est parallèle à l'ouvrage et où la surface est plane, on peut considérer que $Z_S = Z_M$. Des équations 2 et 3, on obtient donc :

$$b_M = d_M \frac{b_S}{d_S} \quad (4)$$

où les disparités d_* sont en pixels et les distances b_* en mètres.

Calculer à la fois d_S et d_M est essentiel pour la mise en œuvre de l'OVS. Deux méthodes ont été testées pour calculer les disparités : par corrélation de phase ou par mise en correspondance de caractéristiques. La méthode de corrélation de phase utilise les propriétés de la transformée de Fourier pour calculer efficacement le déplacement entre deux images en maximisant la corrélation entre elles (on montre facilement que la transformée de Fourier du produit de corrélation des deux images présente un pic dont la position correspond aux coordonnées du vecteur déplacement). La corrélation pouvant s'implanter par un produit dans le domaine de Fourier, la méthode est très rapide. Le lecteur est invité à se référer à (Szeliski, 2010, pp. 341-344) pour plus de détails sur la méthode et à (Forsyth et al., 2002) pour une implantation de précision sub-pixellique. La robustesse de cette approche demande cependant à être analysée pour les images fortement bruitées. Le taux de recouvrement entre deux images est également un facteur qui peut avoir une influence dans le calcul des disparités par corrélation. Les performances de l'OVS par corrélation seront donc évaluées en fonction de ce facteur (voir sous-section 4.2.1). Afin de mieux respecter l'hypothèse de planéité (par exemple dans des cas difficiles comme le tunnel de Revin, voir Fig.5-b) et de diminuer encore les temps de calcul, on limite le calcul de corrélation à une zone de travail préalablement positionnée dans l'image. Par ailleurs, la présence ponctuelle d'une cavité ou d'un élément saillant peut engendrer des valeurs b_M erronées. Un lissage des résultats du mouvement sur la séquence peut éventuellement être envisagé. Dans le cas où l'OVS par corrélation de phase est défaillante, nous proposons de calculer les disparités par mise en correspondance de points SIFT (Lowe, 2004).

À partir de la séquence des inter-distances entre prises de vues de la caméra de référence, on obtient l'abscisse curviligne S'_i à la position i par accumulation, ce qui peut s'écrire :

$$S'_1 = 0, \quad S'_i = \sum_{j=1}^{i-1} \frac{d_M^j}{K_j} \quad 2 \leq i \leq N_d \quad (5)$$

où, pour faciliter l'analogie avec l'équation 1, on a posé $K_j = d_S^j/b_S$. On note que, contrairement à la méthode par points homologues, où le coefficient d'échelle était

constant, K_j varie en fonction notamment de la disparité d_S^j , donc de la distance au piédroit. Il dépend également de la base stéréoscopique b_S , dont l'estimation par calibrage est forcément imparfaite. La mesure obtenue par OVS est directement exprimée dans le système métrique. Toutefois, pour éviter l'accumulation d'erreurs le long de la séquence, une règle de proportionnalité est appliquée à toutes les abscisses curvilignes calculées par l'équation 5, pour mettre en cohérence la longueur de la trajectoire avec celle du tunnel.

3.2.3. Sélection des entrées/sorties du tunnel

Que ce soit dans l'approche par points homologues (voir 3.2.1) ou pour la méthode de l'OVS (3.2.2), il est nécessaire de déterminer une image de début et une image de fin, pour lesquelles l'abscisse est fixée : abscisse 0.0 pour le début et abscisse L (correspondant à la longueur du tunnel) pour la fin. Ce recalage global permet de faire correspondre les abscisses aux extrémités avec la longueur physique du tunnel : les deux méthodes fournissent ainsi des résultats dans le système métrique et comparables entre eux.

Notons qu'une erreur au cours de cette sélection d'images entrée/sortie entache le positionnement de l'intégralité des images de la séquence. En pratique, la longueur du tunnel est difficile à mesurer avec précision en l'absence de points de référence. Il y a alors deux possibilités : implanter des points de coordonnées connues à chaque entrée du tunnel ou s'appuyer sur la longueur donnée dans la documentation. Dans le premier cas, la longueur est connue précisément mais un travail de relevé topographique est éventuellement nécessaire pour déterminer les coordonnées des points de référence. Dans le second cas, aucun relevé n'est nécessaire mais une différence peut exister entre la longueur réelle et les distances calculées entre les images aux extrémités.

3.3. Ré-indexation

Jusqu'à ce stade, seules les images de la caméra gauche du doublet stéréoscopique possèdent une abscisse. L'objectif de cette étape est d'attribuer une abscisse à chaque image du tunnel, et plus précisément, au centre de chaque image. Nous expliquons d'une part comment on attribue une abscisse aux images prises lors du même passage et d'autre part, comment on établit la correspondance entre un passage aller et un passage retour.

3.3.1. Abscisse des images d'un même passage

Dans notre configuration des caméras, lors d'une acquisition d'un bloc d'images à un instant donné, les centres des prises de vue prises à cet instant se trouvent pour la plupart être proches des centres des images de piédroit. Pour ces images (par exemple c,e et f sur la figure 19), il suffit de leur assigner l'abscisse de l'image de la caméra de référence.

Par contre, les caméras obliques, donnant par exemple les vues a et d sur la figure 3, sont celles dont les centres de prise de vue sont éloignés de la caméra de référence.

C'est ce qui est représenté sur la figure 7 par des acquisitions successives de la caméra b et de la caméra d.

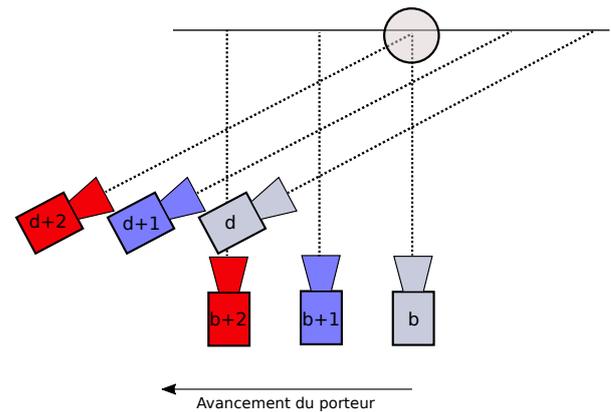


FIGURE 7: Les centres des prises de vue b et d ne coïncident pas : b correspond à d+2, donc b et d+2 ont la même abscisse.

En effet, les caméras obliques visent une zone située en amont ou en aval du porteur avec un recouvrement très faible avec l'image de la caméra de référence. Une visualisation de quelques images obliques permet d'optimiser la distance entre la zone visée par les caméras obliques (d, d+1 et d+2 sur la figure 7) et la zone visée par la caméra de référence (b, b+1, b+2 sur la figure 7). Cette opération est réalisée pour un passage aller et un passage retour car l'éloignement du système d'acquisition par rapport à la scène visée n'est pas obligatoirement le même. Pour une séquence de même type, aller ou retour, on considère que cette ré-indexation est constante.

3.3.2. Décalage entre les passages aller et retour

Les images des séquences aller ou retour présentent une rotation de 180° au niveau de la clé de voûte, ainsi qu'un léger décalage. Ainsi pour une même abscisse il n'y a pas de correspondance satisfaisante entre les images de la clé de voûte obtenues dans chaque sens. Ce décalage, systématique et constant, est dû à l'erreur de sélection des images d'entrée et de sortie du tunnel, comme évoqué dans le premier paragraphe de la sous-section 3.2.3. Il est donc nécessaire de recalibrer les abscisses afin de faire correspondre face à face les piédroits du tunnel. Cette opération s'effectue à une abscisse donnée pour une séquence, en sélectionnant visuellement une image de la clé de voûte prise dans l'autre séquence ayant le plus de recouvrement commun. La correction en abscisse consiste à estimer la différence entre l'abscisse *a priori* et celle déterminée visuellement, puis de propager cette correction à toutes les images de la séquence.

3.3.3. Constitution d'une tuile

Dans certains cas, de nombreux passages ont été effectués et le nombre potentiel d'images à une abscisse donnée peut être élevé. Cependant, nous avons limité à 16 le nombre d'images constituant une tuile, comme sur la figure 3. Parmi toutes les vues enregistrées par une même caméra au cours de différents passages dans un

sens de parcours, on sélectionne celle la plus proche de l'abscisse souhaitée. La distance entre deux tuiles successives est variable selon les expérimentations mais un recouvrement minimum de 50% en longitudinal est requis. Il faut garder à l'esprit qu'un recouvrement trop important augmente de manière inutile les temps de calcul.



FIGURE 8: Vue générale d'une entrée du tunnel de Revin. Le canal est en contrebas d'une piste cyclable protégée par un garde-corps.

4. Évaluation des méthodes

Cette section porte sur l'évaluation des méthodes de pré-positionnement réalisée à partir d'une série de campagnes expérimentales.

4.1. Sites expérimentaux

Pour analyser et comparer les performances des méthodes de localisation proposées, trois campagnes d'acquisition ont eu lieu dans le tunnel canal de Revin (08 Ardennes), la galerie de service du tunnel routier du Somport (64 Pyrénées-Atlantiques) et le tunnel routier de Rive-de-Gier (42 Loire). Les caractéristiques de ces infrastructures et les paramètres d'acquisition utilisés pour chaque expérimentation sont rassemblés dans le tableau 1. On remarque sur ce tableau la diversité des sites d'acquisition en termes de moyen d'accès, de revêtement et de géométrie. Le recouvrement entre deux images successives est de l'ordre de 95% pour le tunnel de Revin et d'environ 90% dans le tunnel de Rive-de-Gier. Pour les séquences acquises dans le tunnel du Somport, la fréquence d'acquisition est, pour des raisons techniques, de 2 Hz , conduisant à un recouvrement entre images voisines d'environ 65%. Dans ce tunnel, seules 5 portions de longueurs comprises entre 40 et 500 m ont été collectées.

Dans le tunnel de Revin, l'acquisition des données de la voûte au complet est particulièrement difficile puisque le canal est à une hauteur d'environ 1.8 m en contrebas d'une piste cyclable située sur le bajoyer (voir Fig. 8). La présence du garde-corps rend difficile la prise d'images du piedroit longeant cette piste. Ce tunnel est équipé d'un réseau de plaquettes catadioptriques géoréférencées par des méthodes classiques de topographie. Les

catadioptriques, détectables de manière automatique dans les images sont implantés tous les 10 m (à l'exception d'une inter-distance de 2 m entre les deux première plaquettes) sur le bajoyer ainsi qu'aux extrémités du tunnel sur le piedroit opposé. Nous considérons ce réseau de points, obtenu par des méthodes tachéométriques classiques, comme relevé de référence.

4.2. Mesure par OVS

4.2.1. Évaluation du calcul de la disparité

Nous voulons dans un premier temps évaluer la méthode de calcul des disparités par corrélation de phase. Pour cela, nous calculons, à chaque acquisition, les inter-distances entre deux images successives dans le tunnel de Revin.

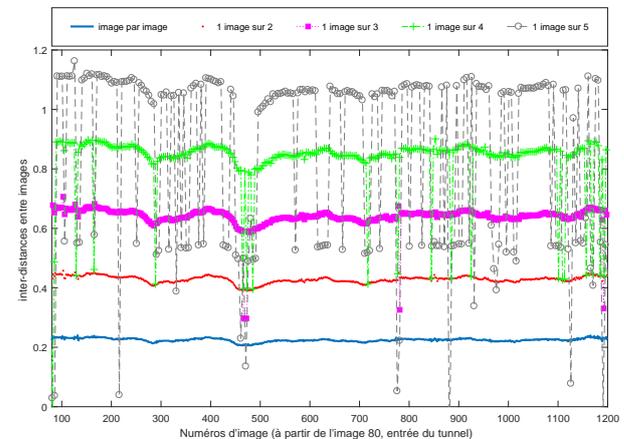


FIGURE 9: Impact du recouvrement entre images sur le calcul des disparités (et donc des inter-distances) par corrélation de phase.

Sur la figure 9, la courbe bleue montre les inter-distances calculées par la méthode de corrélation de phase de l'OVS. On observe que celles-ci sont relativement stables aux alentours de 21 cm entre 2 images et sont en cohérence avec les consignes d'acquisition fixées (5 Hz) à la vitesse d'avancement du bateau (proche de 1 m.s^{-1}). Rappelons que la luminosité des images a été rehaussée (section 3.1.3) engendrant par endroit un rapport signal sur bruit faible. Malgré cette qualité moyenne de l'image on note que le calcul des disparités par corrélation de phase semble cohérent et n'est pas affecté par l'amplification du niveau de bruit.

On souhaite également analyser l'influence du recouvrement entre deux images sur le calcul des inter-distances. Pour cela, ne considérons dans un premier temps qu'une acquisition sur deux. L'évolution des inter-distances est montrée sur la courbe rouge (figure 9), l'allure générale de la courbe d'inter-distance est similaire à celle de la courbe bleue. La valeur d'inter-distance entre deux images est logiquement doublée par rapport au cas initial lorsque toutes les images sont considérées. Nous pouvons continuer à réduire le recouvrement entre deux images, en ne prenant alors en compte qu'une

TABLE 1: Campagnes expérimentales en tunnels routiers et fluviaux.

Expérimentation	Longueur	type	Fréquence	Commentaires
Tunnel de Revin	224 m	canal	5 Hz	Catadioptrés
Tunnel de Rive-de-Gier	80 m	routier	1 acq./50 cm	
Tunnel Somport	3 km	routier	2 Hz	Quelques portions acquises dans une galerie de service

image sur trois (courbe mauve), une image sur 4 (courbe verte) et, enfin, une image sur 5 (courbe grise). On observe que l'inter-distance entre images augmente au fur à mesure que le recouvrement diminue. Cependant un comportement instable apparaît avec une chute de certaines valeurs d'inter-distance. Sur la courbe mauve, ces problèmes sont épisodiques (pour les numéros d'images 470, 471, 781 et 1196) mais ils deviennent plus fréquents sur la courbe verte, et *a fortiori* sur la courbe grise. Ainsi, lorsque le recouvrement entre images devient trop faible, le calcul de disparité par corrélation de phase n'est plus exploitable.

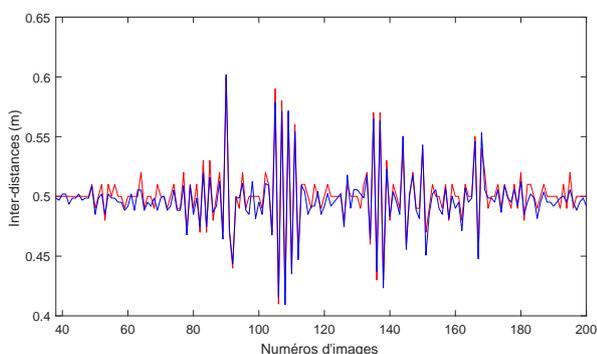


FIGURE 10: Comparaison des inter-distances entre deux acquisitions calculées par OVS (en bleu) et données par l'odomètre mécanique (en rouge) pour un passage dans le tunnel de Rive-de-Gier.

4.2.2. Comparaison entre corrélation de phase et SIFT

La comparaison porte sur les données enregistrées dans le tunnel du Somport. L'évolution de la distance entre deux images, calculée dans un premier temps par OVS (corrélation de phase) est représentée sur la figure 11-a. Les mesures d'inter-distances sont extrêmement fluctuantes et, en conséquence, totalement inexploitable dans notre application. Comme nous l'avons montré dans la première expérimentation (voir sous-section 4.2.1), le faible recouvrement entre 2 images (environ 65% pour cette expérimentation) ne permet pas d'employer la méthode de corrélation de phase pour mesurer le déplacement entre deux images. Dans ces situations, il est possible de calculer les disparités à partir d'extraction des SIFT. Cette méthode permet de déterminer les inter-distances y compris lorsque le recouvrement est faible. C'est ce qui a été entrepris dans les séquences du tunnel du Somport (Fig.11-b). Sur cette

courbe, on observe une inter-distance, dont la valeur médiane sur toute les données, est de $1,21\text{ m}$ conforme aux modalités de déclenchement (2 Hz) et à la vitesse du véhicule (environ 8 km.h^{-1}). Des pics, de faible intensité, tel celui représenté dans le rectangle vert de la figure 11-b, sont visibles à intervalle régulier, avec une fréquence accrue à la fin des acquisitions. D'autres pics de plus grande intensité (rectangle rouge) sont également visibles. Dans le premier cas, les pics correspondent à l'acquisition de cavités (Fig. 11-c) présents en nombre dans le tunnel, et de manière plus régulière dans la dernière partie du tunnel. Les pics de forte amplitude sont, eux, dus aux quelques galeries transversales (Fig. 11-d) dans le tunnel. Cependant ces erreurs sont largement identifiables dans la série de mesure d'inter-distances et peuvent être atténuées en lissant les résultats sur quelques mètres.

4.2.3. Comparaison avec un odomètre

Les mesures obtenues au moyen de l'OVS (par corrélation de phase) ont été comparées à celles d'un odomètre classique dans le tunnel routier de Rive-de-Gier. Les résultats sont donnés sur la figure 10. Qualitativement, les allures générales des deux courbes sont très proches avec une inter-distance de 50 cm environ. La différence entre les courbes est de l'ordre du cm , correspondant à l'ordre de grandeur des précisions de l'odomètre mécanique obtenues sur banc de mesure. Cette évaluation montre qu'il est possible d'utiliser l'OVS pour se positionner en abscisse curviligne en tunnel routier.

4.2.4. Comparaison avec des points de repère

Nous souhaitons également évaluer l'OVS à partir de mesures effectuées sur un bateau où les mouvements, notamment latéraux, sont plus amples. Pour cela, une expérimentation est menée dans le tunnel de Revin, équipé d'un réseau de points de référence relevés sur des catadioptrés par des techniques classiques de topométrie. Dans cette expérimentation, l'objectif est d'analyser les résultats de l'odométrie visuelle simplifiée par rapport à cet ensemble de référence.

Pour prendre en compte la distance entre deux catadioptrés sur les images, on sélectionne l'image où la plaquette est la plus proche du milieu de l'image et on ajoute le décalage δ^j entre la plaquette j et le milieu de l'image (Fig.12). Nous avons choisi de toujours considérer le bord inférieur droit de la plaquette comme référence pour calculer ce décalage.

Ainsi, pour $j \in [2, \dots, N_c]$, la distance cumulée $d_{j-1,j}$ entre le catadioptré $j-1$ et le catadioptré j est calculée

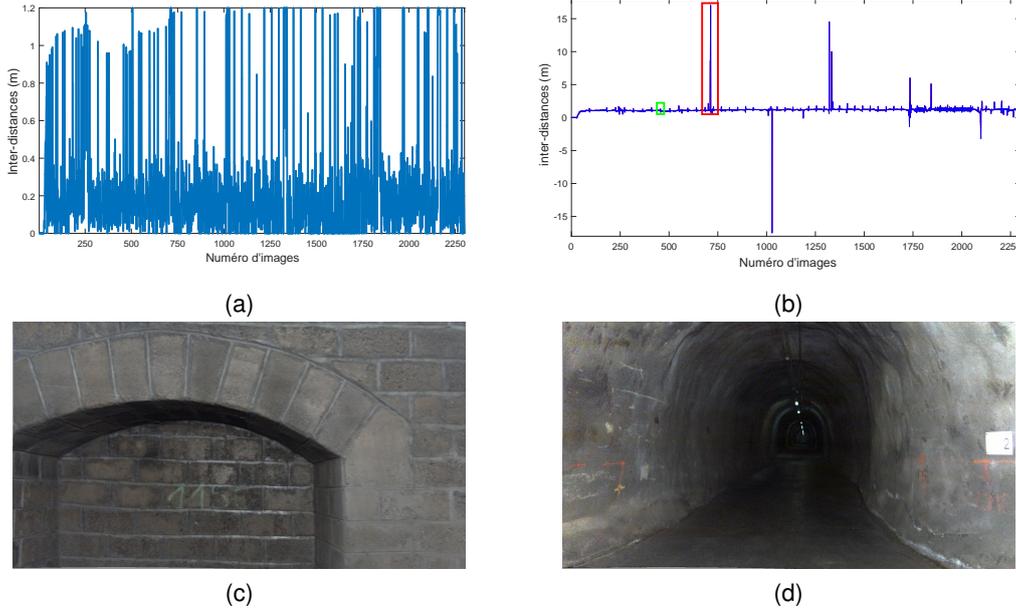


FIGURE 11: Calcul des Inter-distances par OVS dans le tunnel du Somport : (a) par corrélation de phase, (b) par extraction de caractéristiques SIFT. Les sauts à fréquence régulière, tel celui représenté dans le rectangle vert, sont dus à la présence de cavités (sous-figure (c)). On remarque que la fréquence d'apparition de ces cavités est plus élevée en fin d'acquisition. Les sauts ponctuels de grande valeur tel celui représenté dans le rectangle rouge sont liés à la présence de galeries transversales (sous-figure (d)).

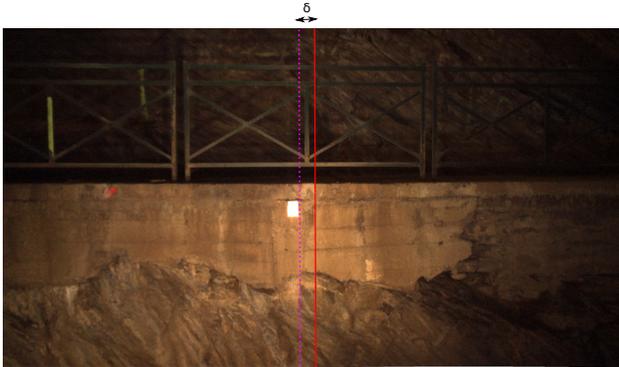


FIGURE 12: Catadioptré implanté sur le bajoyer et représentation du décalage δ entre le milieu de l'image (ligne rouge) et le bord inférieur droit du catadioptré (ligne pointillée en violet).

selon :

$$d_{j-1,j} = \sum_{k=1}^{N_j} b_M(k) + \delta^{j-1} - \delta^j \quad (6)$$

où N_j représente le nombre d'images entre le catadioptré $j-1$ et j . Ce nombre est variable mais proche de 50, c'est-à-dire qu'environ 50 images séparent deux catadioptrés (à raison d'une image tous les 20 cm environ). N_c représente le nombre de catadioptrés dans le tunnel ($N_c = 24$). $b_M(k)$ correspond à la distance inter-images, telle que définie (par b_M) dans l'équation 4.

Le résultat est montré sur la figure 13 pour un passage dans le tunnel. Dans une première analyse, on remarque que les deux courbes ont des allures simi-

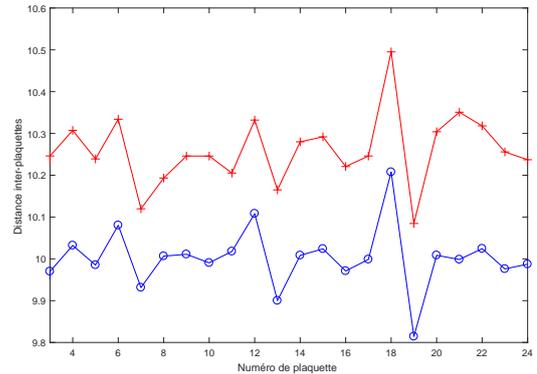


FIGURE 13: Comparaison des inter-distances $d_{(j-1,j)}$ calculées par OVS (en rouge) aux distances entre catadioptrés calculées à partir de leurs géo-référencements (en bleu) pour un passage dans le tunnel.

lares. Les variations d'inter-distance entre deux plaquettes suivent ainsi une trajectoire quasiment parallèle. L'autre élément marquant ressortant de l'analyse de ces courbes est qu'on observe un biais entre les deux approches. Ainsi, l'inter-distance mesurée par OVS est systématiquement sur-estimée par rapport aux mesures de référence.

La figure 14 montre le rapport entre les deux courbes. La moyenne de ces rapports est de 1,0298 avec un écart-type égal à 0,0023. Ce décalage peut donc être considéré comme constant. Sur cette expérimentation, nous pensons que le calibrage est à l'origine de cette

erreur systématique. Celle-ci n'est cependant pas bloquante pour le bon fonctionnement de la méthode car, constante, elle est très facile à compenser.

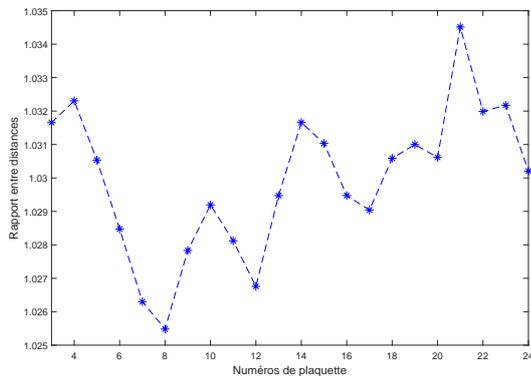


FIGURE 14: Rapport entre les 2 courbes de la figure 13. Les deux premières plaquettes, distantes de 2 m entre elles, ne sont pas considérées.

Cette compensation peut se faire à partir de la longueur totale du tunnel, qui est toujours connue, ou à partir de repères implantés dans le tunnel.

4.3. Comparaison des méthodes de positionnement

Dans un premier temps, nous comparons les inter-distances entre deux images calculées par la méthode par points homologues et l'odométrie visuelle simplifiée dans le tunnel de Revin (fig. 15).

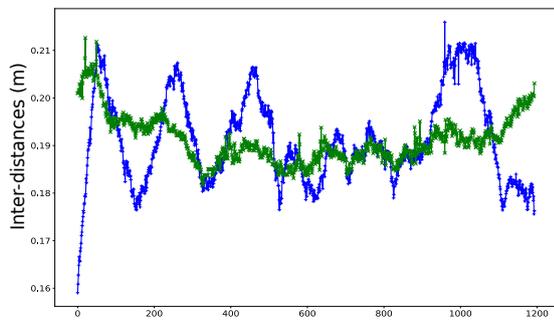


FIGURE 15: Comparaison des déplacements obtenus avec la méthode par points homologues (en bleu) et avec l'Odométrie Visuelle Simplifiée (en vert).

On a déjà vu dans la section 4.2.1 que, avec un fort recouvrement, les inter-distances mesurées par la méthode OVS étaient stables autour de 21 cm, et étaient cohérentes par rapport aux conditions expérimentales. On observe des variations locales beaucoup plus importantes avec la méthode par points homologues que pour l'OVS. Si la méthode par points homologues possède un avantage de simplicité car elle repose sur des images simplement colorisées qui n'ont pas été corrigées par un modèle de caméra, l'absence de métrique pose un problème important de dérive dans les calculs successifs de disparité. En effet, l'estimation de la translation du plan

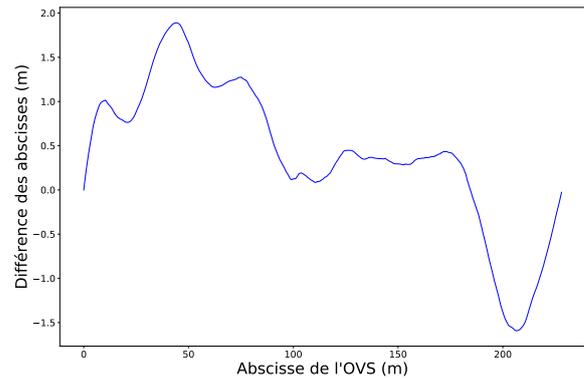


FIGURE 16: Différences des abscisses obtenues entre la méthode par points homologues et l'Odométrie Visuelle Simplifiée.

se trouve faussée en présence d'un objet saillant texturé par exemple et au contraire, si la surface est faiblement texturée, la recherche de points homologues peut s'avérer plus coûteuse car il faut augmenter le ratio de Lowe (Lowe, 2004). Les erreurs locales commises par cette méthode se propagent à l'ensemble des images lors de la mise à l'échelle finale sur toute la longueur du tunnel.

Nous savons d'après les résultats obtenus dans la section 4.2 que le pré-positionnement par l'OVS est décimétrique. La conséquence des erreurs locales des deux méthodes est comparée sur la figure 16, où le cumul de toutes les distances intermédiaires donne l'abscisse curviligne pour l'image considérée (la première image d'entrée possédant l'abscisse 0). Même si la première image et la dernière possèdent la même abscisse, au sein du tunnel les abscisses des images peuvent varier jusqu'à 3.5 m, ce qui est loin des objectifs décimétriques d'une méthode de pré-positionnement.

4.4. Positionnement des images de la tuile

Les performances de l'OVS étant connues, on peut désormais évaluer le positionnement de l'ensemble des images qui constituent une tuile. Une première visualisation des 16 images d'une même tuile, sélectionnées automatiquement, est proposée sur la figure 17 pour le tunnel de Revin. En dépit de la configuration complexe du tunnel, il existe toujours un recouvrement entre deux images voisines. Ainsi, à partir de toutes ces images, on peut parcourir la voûte d'un côté à l'autre du canal. Par ailleurs, on peut voir que la qualité des images rehaussées peut s'avérer très moyenne, ce qui illustre que la méthode du « flat-field » ne peut pas corriger entièrement les défauts d'éclairage. Ces zones fortement bruitées, correspondant à l'eau ou à la partie la plus lointaine des images dans les caméras obliques ne sont pas exploitées dans les calculs de reconstruction photogrammétrique.

Une analyse quantitative des résultats est également réalisée en complément de l'analyse visuelle. L'objectif est notamment d'évaluer le recalage des images de la clé de voûte, acquises dans les deux sens de passage. Pour cela nous nous appuyons sur le nombre de

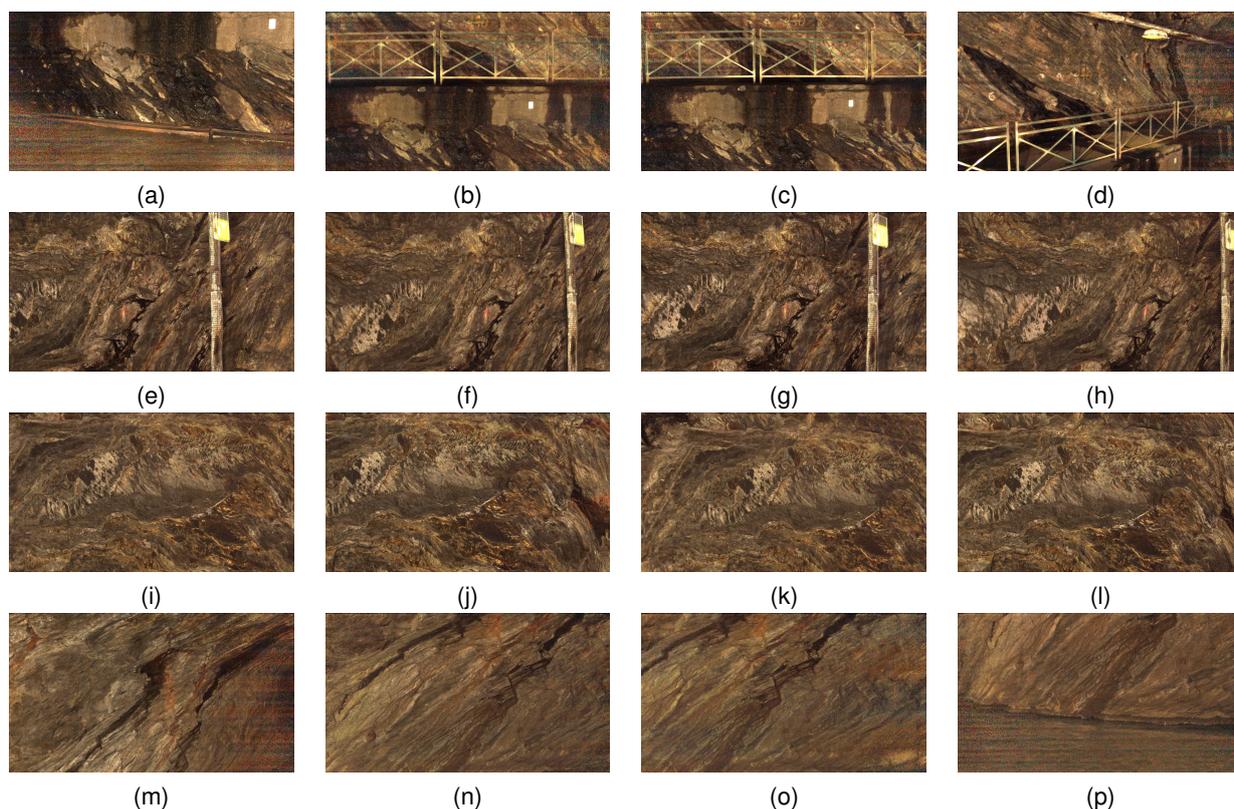


FIGURE 17: Ensemble des 16 images sélectionnées automatiquement pour une même tuile. Les lettres font référence au schéma de la tuile représentée sur la figure 3.

points homologues entre deux images d'abscisse curviligne proche. C'est le sens des représentations de la figure 18 où les points correspondants sont reliés par une ligne verte. La figure montre que les images à l'abscisse 100, quel que soit le passage, se correspondent de par leur taux de recouvrement important. Les faisceaux de points homologues en nombre important sont constitués d'un ensemble dense de droites parallèles, pour un même sens de passage : ce sont des images quasiment identiques légèrement translattées. Pour les faisceaux qui se coupent, il s'agit d'images de clef de voûte de sens de passage différents, d'où d'une part un plus faible nombre de points homologues et d'autre part l'existence d'une rotation entre les images. Les résultats restent similaires aux différentes abscisses testées : plusieurs centaines pour des sens opposés et plusieurs milliers pour un même sens de parcours. Ces correspondances montrent que le calcul d'orientation des images peut être mené avec succès.

4.5. Aspects algorithmiques

L'algorithme de pré-positionnement des images est codé en langage Python. Sa rapidité d'exécution dépend de la technique employée. Ainsi, l'odométrie visuelle simplifiée basée sur la corrélation de phase nécessite 20 minutes de calcul environ pour une séquence de 2300 images. Sur ce même nombre d'images, l'utilisation de caractéristiques SIFT nécessite 1 heure de calcul. La complexité informatique est surtout portée par les ac-

cès mémoire lorsque l'on charge à chaque fois l'image considérée. A ces temps de calcul, il convient d'ajouter les traitements manuels subsistants pour réaliser la totalité de la chaîne de traitement. Ces temps de réalisation dépendent de l'opérateur.

5. Applications

La méthode de pré-localisation décrite et évaluée dans les sections précédentes permet de positionner les images en abscisse curviligne dans ce type d'infrastructures. Nous avons appliqué directement cette approche de pré-localisation pour naviguer virtuellement dans le tunnel avec des outils de niveau de complexité différente : une visualisation et exploitation des données localisées (sous-section 5.1) et la modélisation 3D du tunnel par approche photogrammétrique (sous-section 5.2).

5.1. Visualisation

On peut concevoir un logiciel de visualisation de l'ensemble des images prises dans l'ouvrage, en se basant sur l'abscisse curviligne. A une position donnée, il est ainsi possible de représenter le déplié de l'ensemble de la voûte ce qui répond à un premier besoin exprimé par des exploitants et spécialistes de l'inspection des tunnels. Un tel déplié est représenté sur la figure 19.

Un autre objectif qui nous intéresse est de pouvoir détecter automatiquement des objets d'intérêt

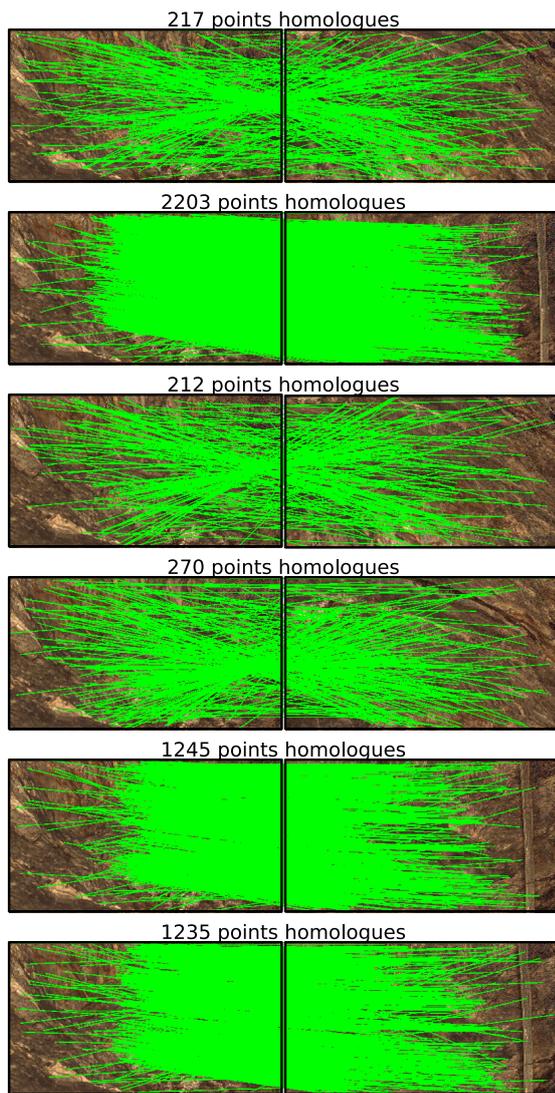


FIGURE 18: Correspondances sur les images de clefs de voûte entre un passage de référence à une abscisse donnée (colonne de gauche) et 6 autres passages à l'abscisse la plus proche (colonne de droite). Les lignes vertes relient les points homologues sur une paire d'images. Les paires des lignes 1,3 et 4 correspondent à un sens de passage opposé, tandis que les paires des lignes 2, 5 et 6 correspondent à un même sens de passage.

(désordres, équipements) dans les images (Foucher et al., 2016; Decor et al., 2019). Les méthodes d'apprentissage profond sont actuellement explorées pour répondre à cette problématique. Celles-ci nécessitent au préalable un travail de recensement et de labellisation conséquent des objets dans les images. Lorsque les éléments à saisir sont visibles dans plusieurs images, la localisation des images se reportant à la même surface permet de faciliter le relevé d'un même objet présent sur ces dernières et d'améliorer ainsi la base d'apprentissage.



FIGURE 19: Exemple d'un déplié à une abscisse donnée.

5.2. Modélisation 3D

Une modélisation 3D de la voûte complète du tunnel par approche photogrammétrique a également été réalisée. La donnée préalable de l'abscisse curviligne pour les images expérimentales permet d'optimiser très fortement les différents maillons de la chaîne des traitements photogrammétriques. Le pré-positionnement donne la possibilité de lancer uniquement des calculs de points homologues les plus pertinents entre les images par exemple : les calculs d'appariement inutiles sont évités, ce qui accélère cette première étape en photogrammétrie. De plus, étant donné la forte diminution de points homologues non pertinents dans le calcul de faisceaux, la phase d'orientation se trouve également facilitée.

Il est aussi possible d'envisager de regrouper les images en ensembles, appelés tronçons, composés de plusieurs tuiles juxtaposées. Ces tronçons sont traités indépendamment et agglomérés par la suite entre eux. Cette méthode permet de paralléliser les calculs photogrammétriques sur l'ensemble du tunnel. De cette manière, nous avons élaboré des modèles 3D de différents ouvrages en utilisant les logiciels MicMac (Rupnik et al., 2017) ou Photoscan\metashape². Pour ce faire, le regroupement des images propres à chaque tronçon a été réalisé de manière automatique à partir de leur abscisse curviligne, donnée par la méthode proposée. Les images sélectionnées sont celles qui présentent le meilleur recouvrement possible entre elles en évitant les redondances dues aux multiples passages dans le tunnel. L'objectif étant de reconstruire la surface de la paroi du tunnel, il est inutile dans cette étape d'accumuler des images similaires issues de passages différents, car cela alourdit le calcul de toutes les corrélations locales entre les images, et n'améliore pas nécessairement le résultat final. La figure 20 illustre la reconstruction d'un tronçon élémentaire du tunnel du Somport. La largeur correspond à la largeur d'une image sur le piedroit du tunnel (environ 2 m). En alignant les différents tronçons, on obtient un modèle plus complet du tunnel, comme le montre la figure 21 pour le tunnel de Revin. Dans cette figure, le tunnel est modélisé sur une longueur de 40 m environ. Avec le logiciel Photoscan\metashape, l'alignement des tronçons est fait à partir des positions et orientations des caméras (fonction *align*) par rapport au premier tronçon. Il n'y a donc pas d'étape globale d'ajustement de faisceaux.

Notons que les deux modèles présentés ont été densifiés à partir des fonctionnalités dédiées propres à chaque logiciel.

2. <https://www.agisoft.com/>

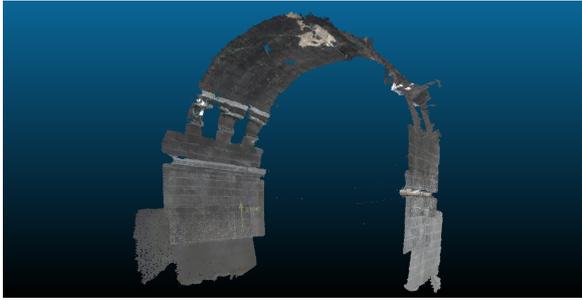


FIGURE 20: Reconstruction d'un tronçon en 3D d'une cavité dans le tunnel du Somport avec le logiciel MicMac. Les lacunes observées dans le modèle sont dues à la présence de cavités sur la voûte qui ne sont pas visibles dans l'image acquise.

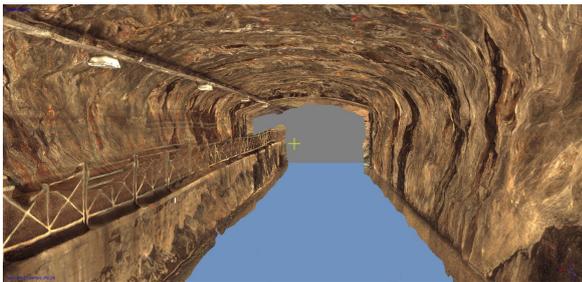


FIGURE 21: Reconstruction du tunnel de Revin avec le logiciel Photoscan.

6. Discussion et conclusion

Dans cet article, nous avons évalué une méthode de localisation en abscisse curviligne de données à partir d'une technique d'odométrie visuelle simplifiée. L'approche que nous avons mise en œuvre permet de connaître la distance d'une image par rapport à la précédente, distance calculée à partir de techniques de corrélation de phase ou, lorsque le recouvrement entre deux images est insuffisant, à partir d'une recherche de points en correspondance. Nous avons mis en évidence empiriquement l'efficacité de cette approche en comparant les inter-distances mesurées par OVS et avec une odométrie mécanique, d'une part, et par rapport à un canevas de points topographiques de référence, d'autre part. Il ressort de cette étude que la méthode de pré-positionnement proposée peut être utilisée pour positionner un objet dans une infrastructure oblongue où le GPS est inaccessible. La mesure de distance n'étant pas absolue, le positionnement est bien évidemment sujet aux dérives dues à l'accumulation d'erreurs. Cette limite n'est cependant pas propre à cette méthode et l'ajout de points de contrôle disposés sur la paroi du tunnel relevés avec un tachéomètre permet de faire diminuer cette dérive.

Nous avons signalé qu'un certain nombre d'opérations manuelles sont nécessaires avant d'appliquer l'algorithme sur la séquence. Nous pouvons classer ces interventions en deux catégories. La première relève d'une sélection d'images dans les séquences, à savoir celles correspondant aux extrémités du tunnel et les images

nécessaires à l'établissement du filtre pour la technique du « flat-field ». La deuxième catégorie demande un peu plus de technicité, à savoir le calibrage de la paire stéréovision, le recalage des séquences sur la clef de voûte et la zone d'intérêt (*i.e.* la troncature de l'image) pour que l'OVS fournisse des résultats satisfaisants. Pour chacune de ces interventions, ce n'est pas tant la tâche de paramétrage en elle-même qui peut s'avérer fastidieuse mais la vérification sur un petit jeu de données de la validité de ces paramètres. La discussion porte ainsi sur la possibilité et l'intérêt d'automatiser ces étapes manuelles. Pour le repérage des entrées et sorties, les extrémités de tunnel ne montrent pas toujours généralement une séparation franche au niveau du piédroit (voir figure 8) et la décision doit être apportée par l'utilisateur. La sélection des images pour le filtre doit aussi s'effectuer en veillant à ce qu'il n'y ait pas d'objets récurrents (chemin de câbles, tuyaux...) dans l'ensemble choisi, auquel cas le filtre se trouve faussé. Ces étapes demandent de revenir sur les images, mais leur automatisation demanderait un travail important et il nous semble difficile d'être exhaustif étant donnée la variabilité de configurations existantes de tunnels.

Pour les tâches de la deuxième catégorie, le calibrage est réalisé une seule fois sur des images prises sur une mire indépendamment des images du tunnel et peut être vu comme une donnée d'entrée de la méthode. Le choix de la zone pour paramétrer la corrélation est très rapide et mettre en place une méthode itérative qui évalue la surface la plus efficace pour la méthode ne semble donc pas très opportun. Le recalage sur la clef de voûte pourrait être lui automatisé par appariement de points homologues à une abscisse donnée entre toutes les images candidates. C'est là le point d'amélioration le plus pertinent, à notre sens, de la méthode en elle-même.

Pour conclure, il est intéressant de noter que notre système d'acquisition est figé lors d'un passage dans un tunnel et que les relations géométriques entre ces positions des caméras les unes par rapport aux autres n'ont pas été exploitées dans la présente méthode. Après une phase de calibrage, ces relations permettraient de localiser automatiquement, à la précision de l'OVS près, toutes les caméras. Cette approche méthodologique est plus évoluée que notre méthode de pré-positionnement et constitue une perspective très intéressante au travail présenté dans cet article.

Références

- Akbarzadeh, A., Frahm, J.-M., Mordohai, P., Clipp, B., Engels, C., Gallup, D., Merrell, P., Phelps, M., Sinha, S., Talton, B., et al., 2006. Towards urban 3D reconstruction from video. Dans : 3D Data Processing, Visualization, and Transmission, Third International Symposium on. IEEE, pp. 1–8.
- Albert, J., Charbonnier, P., Foucher, P., Muzet, V., Prybyla, D., Perrin, T., Grussenmeyer, P., Guillemin, S., Koehl, M., septembre 2013. Devising a visual inspection system for canal tunnels : Preliminary studies. Dans : International Archives of the Photogrammetry, Remote Sensing and spatial infor-

- mation sciences, XXIV International CIPA Symposium. Strasbourg, France, pp. 13–18.
- Aqel, M. O., Marhaban, M. H., Saripan, M. I., Ismail, N. B., octobre 2016. Review of visual odometry : types, approaches, challenges, and applications. SpringerPlus 5 (1), 1897.
- Bauer, A., Gutjahr, K., Paar, G., Kontrus, H., Glatzl, R., 2015. Tunnel surface 3D reconstruction from unoriented image sequences. arXiv preprint arXiv :1505.06237.
- Beis, J. S., Lowe, D. G., juin 1997. Shape indexing using approximate nearest-neighbour search in high-dimensional spaces. Dans : Conference on Computer Vision and Pattern Recognition. IEEE, San Juan, Puerto Rico, pp. 1000–1006.
- Birem, M., Kleihorst, R., El-Ghouthi, N., 2018. Visual odometry based on the fourier transform using a monocular ground-facing camera. Journal of Real-Time Image Processing 14 (3), 637–647.
- Brown, M., Lowe, D., 2003. Recognising panoramas. Dans : ICCV'03 Proceedings of the Ninth IEEE International Conference on Computer Vision. Vol. 2. pp. 1218–1225.
- Brown, M., Lowe, D. G., 2007. Automatic panoramic image stitching using invariant features. International journal of computer vision 74 (1), 59–73.
- Cannelle, B., 2013. Estimation de pose de grands blocs d'images panoramiques issues de systèmes de numérisation mobile. Thèse de doctorat, Université Paris-Est.
- Charbonnier, P., Foucher, P., Chavant, P., Muzet, V., Prybyla, D., Perrin, T., Albert, J., Grussenmeyer, P., Guillemin, S., Koehl, M., 2014. An image-based inspection system for canal-tunnel heritage. International Journal of Heritage in the Digital Era 3 (1), 197–214.
- Cornelis, N., Cornelis, K., Van Gool, L., 2006. Fast compact city modeling for navigation pre-visualization. Dans : Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on. Vol. 2. IEEE, pp. 1339–1344.
- Decor, G., Bah, M.-D., Foucher, P., Charbonnier, P., Heitz, F., mai 2019. Détection d'anomalies dans des images de tunnels par forêts d'arbres aléatoires ou par apprentissage profond. Dans : actes de la conférence ORASIS 2019. Saint-Dié-des-Vosges, France, format électronique.
- Foroosh, H., Zerubia, J., Berthod, M., 02 2002. Extension of phase correlation to subpixel registration. IEEE transactions on image processing : a publication of the IEEE Signal Processing Society 11, 188–200.
- Foucher, P., Bah, M., Charbonnier, P., Boulogne, C., Larive, C., juin 2016. classification automatique de défauts sur des images de tunnels par forêts d'arbres aléatoires. Dans : Congrès national sur la Reconnaissance de Formes et l'Intelligence Artificielle (RFIA). Clermont-Ferrand.
- Goshtasby, A., Gage, S., Bartholic, J., mai 1984. Two-stage cross correlation approach to template matching. IEEE Transactions on Pattern Analysis and Machine Intelligence 6 (3), 374–378.
- Havlena, M., juin 2012. Incremental structure from motion for large ordered and unordered sets of images. Ph.d thesis, Czech Technical University in Prague, Prague, République Tchèque.
- Lowe, D. G., 2004. Distinctive image features from scale-invariant keypoints. International journal of computer vision 60 (2), 91–110.
- Paudel, A. P., Habed, A., Vasseur, P., Paudel, A. P., Habed, A., Vasseur, P., D-d, I. K., Kweon, I. S., 2014. 2D-3D camera fusion for visual odometry in outdoor environments.
- Rupnik, E., Daakir, M., Pierrot Deseilligny, M., juin 2017. MicMac – a free, open-source solution for photogrammetry. Open Geospatial Data, Software and Standards 2 (1), 14.
- Scaramuzza, D., Fraundorfer, F., 2011. Visual odometry [tutorial]. IEEE Robot. Automat. Mag. 18 (4), 80–92.
- Scaramuzza, D., Siegwart, R., octobre 2008. Appearance-guided monocular omnidirectional visual odometry for outdoor ground vehicles. Trans. Rob. 24 (5), 1015–1026.
- Szeliski, R., 2010. Computer Vision : Algorithms and Applications, 1er Edition. Springer-Verlag, Berlin, Heidelberg.
- Tsai, R., 1987. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf tv cameras and lenses. IEEE Journal on Robotics and Automation 3 (4), 323–344.
- Wang, X., Zhan, Z., Heipke, C., 2017. An efficient method to detect mutual overlap of a large set of unordered images for structure-from-motion. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences 4, 191–198.
- Wu, C., 2013. Towards linear-time incremental structure from motion. Dans : 2013 IEEE International Conference on 3D Vision (3DV). IEEE, pp. 127–134.
- Zhang, Z., septembre 1999. Flexible camera calibration by viewing a plane from unknown orientations. Dans : Proceedings of the Seventh IEEE International Conference on Computer Vision (ICCV). Vol. 1. IEEE, Kerkyra, Grèce, pp. 666–673.