

Rapport du projet recherche

Jean-françois Villeforceix

Gauthier Duponchel

15 février 2016

Table des matières

Présentation	3
Préambule	3
Contexte	4
1.1 Nom du projet	4
1.2 Laboratoire	4
1.3 Commanditaires	4
1.4 Durée	4
1.5 Données	4
Problématique et objectif	5
État de l’art	7
Corrélation image entière	7
3.1 Principe	7
3.2 Résultat	8
Détection de points d’intérêt puis appariement des points	9
4.1 Détection de points d’intérêt	9
4.1.1 Qu’est-ce qu’un point d’intérêt ?	9
4.1.2 Méthodes de détection ?	9
4.1.3 Présentation de détecteurs de points d’intérêt	9
4.1.4 Qualités des principaux détecteurs de points d’intérêt	10
4.2 Appariement de points d’intérêt	10
4.2.1 Objectif	10
4.2.2 Appariement par corrélation	11
4.2.3 Appariement via descripteurs de points	12
Tests et résultats	14
Détermination du vecteur translation	14
5.1 Protocole expérimental	14
5.2 Résultats globaux	15
5.2.1 Comparaison des méthodes	15
5.2.2 Résultats des filtrages	17
Conclusion	20

Présentation

Préambule

Dans le cadre de la formation de deuxième année du cycle ingénieur de l'ENSG, une découverte du domaine de la recherche est mise en place. Cette découverte prend forme au travers d'un projet recherche.

L'objectif de ce rapport est :

- d'exposer la problématique étudiée lors de ce
- de présenter les différentes techniques mises en places
- de présenter les résultats des tests que nous avons effectués

La seconde partie du rapport fait figure d'état de l'art et présente deux méthodes différentes d'appariement d'images.

- Première approche : Corrélation image entière
- Seconde approche : Détection de points d'intérêt puis appariement des points

Pour chacune des approches, on exposera le principe général avant de mettre en évidence comment il permet d'estimer le vecteur translation existant entre les images.

Contexte

1.1 Nom du projet

Étude de méthode d'appariement entre images générées à partir de donnée LIDAR

1.2 Laboratoire

IGN - MATIS (Méthodes d'analyse et de traitement d'images pour la stéréo-restitution)

1.3 Commanditaires

Arnaud Le-Bris (arnaud.le-bris@ign.fr)

Jean-Pierre Papelard (jean-pierre.papelard@ign.fr)

Bruno Vallet (bruno.vallet@ign.fr)

1.4 Durée

14 mercredis de septembre 2015 à février 2016

1.5 Données

Acquisition LIDAR terrestre (véhicule mobile « Stéréopolis »)

Ortho-LIDAR de résolution 4 cm.

Problématique et objectif

Le véhicule de numérisation mobile « Stéréopolis » acquiert images et nuages de points laser au niveau de la rue. Le traitement des données brutes utilise le résultat d'une trajectographie post-traitée pour produire le géo-référencement d'images de points dans le référentiel géographique officiel.

L'acquisition d'un chantier complet génère de nombreux multi-passages aux différents nœuds du réseau routier. A ces endroits, la précision des trajectographies ne permet pas une superposition parfaite des données produites.



Illustration des multi-passages

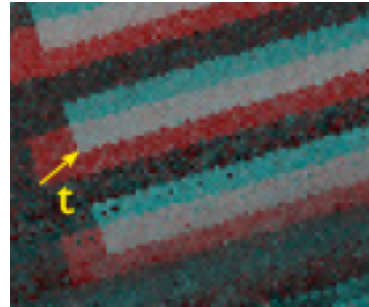


Illustration des décalages

Pour corriger ce problème, un processus de recalage absolu et relatif permet actuellement d'identifier manuellement des détails homologues dans les dalles d'ortho-images de la chaussée générées à partir des données LIDAR.

L'objectif de notre projet est ainsi de tester les différentes techniques photogrammétriques d'appariement d'images afin d'identifier celle qui pourra être implémentée dans un processus automatisé de recalage du « véhicule » aux lieux de multi-passages.

Comme nous l'avons précédemment évoqué, les images que nous traitons ont été géo-référencées via la trajectographie. Ainsi, on part de l'hypothèse que la transformation que l'on cherche à déterminer est un vecteur de translation qui permette de recaler précisément "les trajectoires" aux lieux de multi-passages.

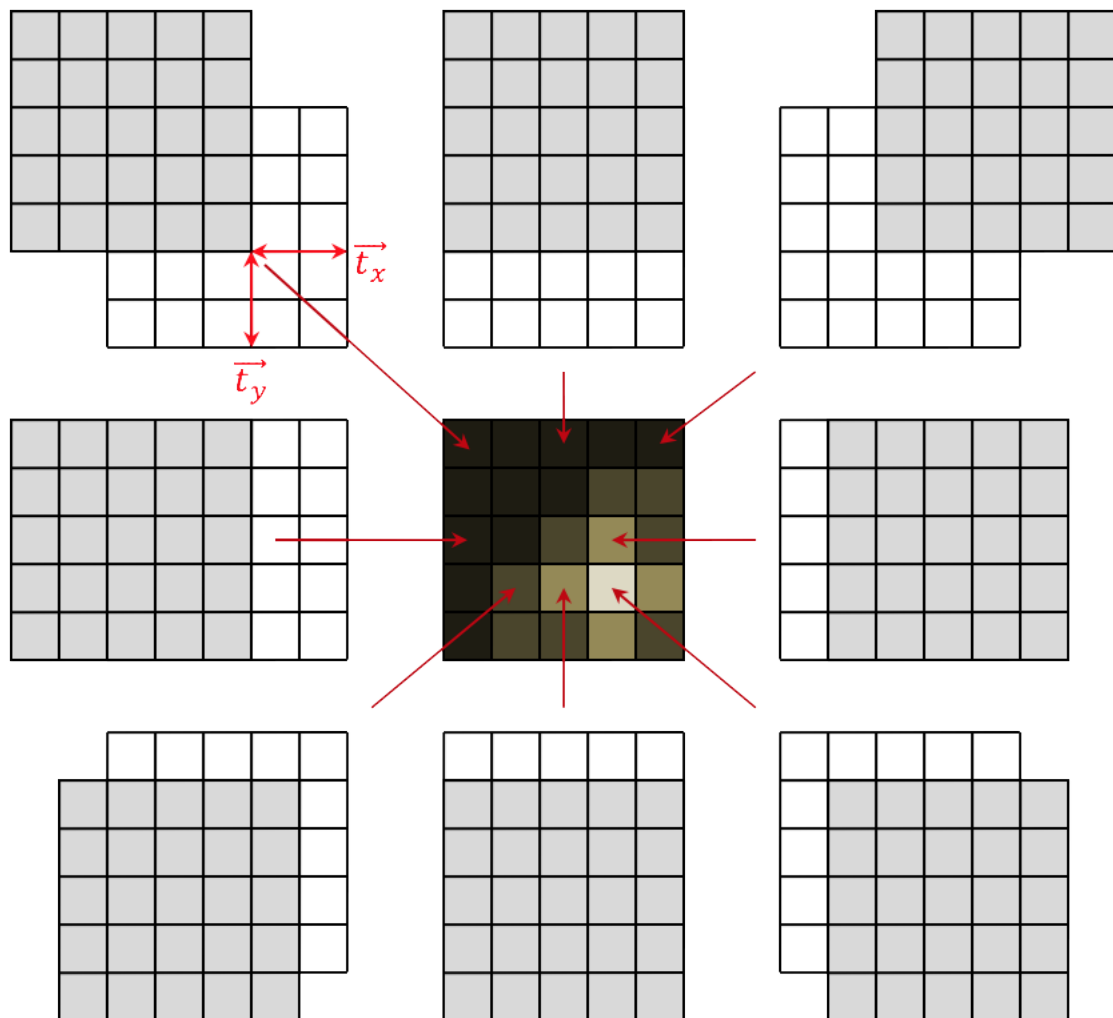
La première partie de ce rapport consistera en la présentation des différents processus photogrammétriques utilisés. La seconde sera elle dédiée à la présentation des tests effectués ainsi que leurs résultats.

État de l'art

Corrélation image entière

3.1 Principe

Cette technique est basée sur le calcul d'un score de corrélation global entre deux images. On commence par calculer ce score en prenant les deux images « brutes » on obtient un premier score de corrélation qu'on affecte au pixel central de l'image des scores que l'on génère. Ensuite, on translate pour différents $t_x; t_y$ la seconde image, on calcule les scores de corrélation que l'on affecte aux pixels correspondants de l'image résultante des scores de corrélation.



Principe de la corrélation image entière

3.2 Résultat

On génère ainsi une image des scores de corrélation. Dans cette image, plus le pixel est clair, plus le score de corrélation est grand. C'est donc le pixel le plus clair de cette image des scores qui correspond à la meilleure translation.

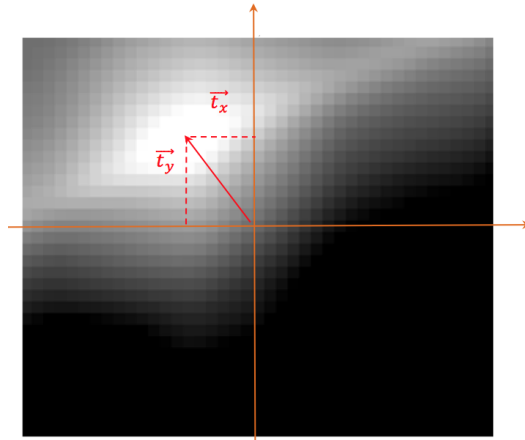


Image des scores de corrélation

Détection de points d'intérêt puis appariement des points

4.1 Détection de points d'intérêt

4.1.1 Qu'est-ce qu'un point d'intérêt ?

Les points d'intérêt sont des points « remarquables » qui correspondent à des doubles discontinuités de la fonction d'intensité de l'image (celles-ci pouvant être dues à des variations de la radiométrie (texture) des objets photographiés ou à des discontinuités de profondeur). Les points d'intérêt vont donc correspondre à des points isolés, des coins, des intersections, des « taches » ou « blobs ». Il y aura donc de bonnes chances de détecter ces mêmes points dans plusieurs images d'une même scène.

4.1.2 Méthodes de détection ?

De nombreuses méthodes pour détecter des points d'intérêt ont été proposées. Elles peuvent se classer en 3 familles :

1. Approches contours : il s'agit de détecter les contours dans une image puis d'extraire les points d'intérêt le long des contours en considérant les points de courbures maximales ainsi que les intersections de contours.
2. Approches intensité : l'idée est de s'intéresser directement à la fonction d'intensité des images pour en extraire les points de discontinuité.
3. Approches à base de modèles : les points d'intérêts sont identifiés dans l'image par mise en correspondance de la fonction d'intensité de l'image avec un modèle théorique de cette fonction au voisinage des points d'intérêts recherchés. Précis mais problème de répétabilité (approche moins générique) ?

Les approches « intensité » sont celles utilisées généralement. Les raisons en sont leur indépendance vis à vis d'une première étape de détection de contours (stabilité) ainsi que vis à vis du type de points d'intérêts (méthodes plus générales).

On se limitera donc dans cette étude aux détecteurs basés sur l'approche intensité.

4.1.3 Présentation de détecteurs de points d'intérêt

L'idée dans cette partie est de répertorier les différents algorithmes de détection de points d'intérêt et d'en présenter les principales caractéristiques et qualités. Pour une lecture rapide et globale, le résultat est présenté sous forme d'un tableau. Pour chaque méthode on identifie ainsi la nature des points détectés ainsi que les invariances du détecteur.

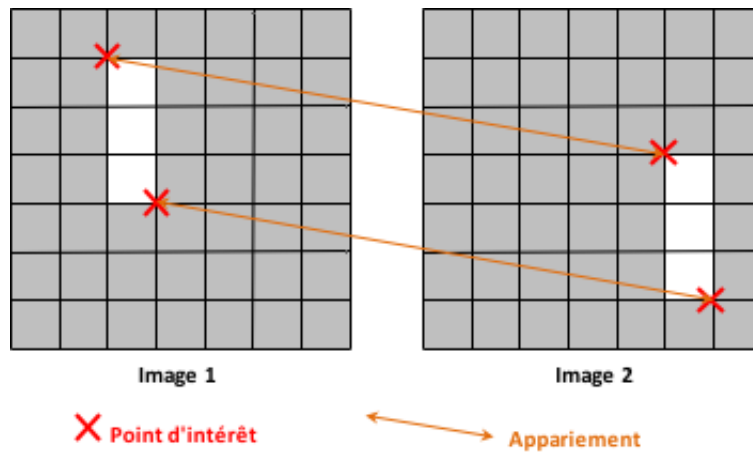
4.1.4 Qualités des principaux détecteurs de points d'intérêt

NOM	Points détectés			Invariant à			
	Coin	Point isolé	Tache	Rotation	Echelle	Légère déformation Affine	Déformation Affine
MORAVEC 1980	✓	✓					
HARRIS 1988	✓	✓		✓			
FAST 2005	✓	✓		✓			
HARRIS-LAPLACE 2004	✓	✓	✓	✓	✓	✓	
HESSIEN-LAPLACE 2004	✓	✓	✓	✓	✓	✓	
SIFT	✓	✓	✓	✓	✓	✓	
HARRIS AFFINE 2004	✓	✓	✓	✓	✓		✓
HESSIEN AFFINE 2004	✓	✓	✓	✓	✓		✓
ASIFT 2009	✓	✓	✓	✓	✓		✓

4.2 Appariement de points d'intérêt

4.2.1 Objectif

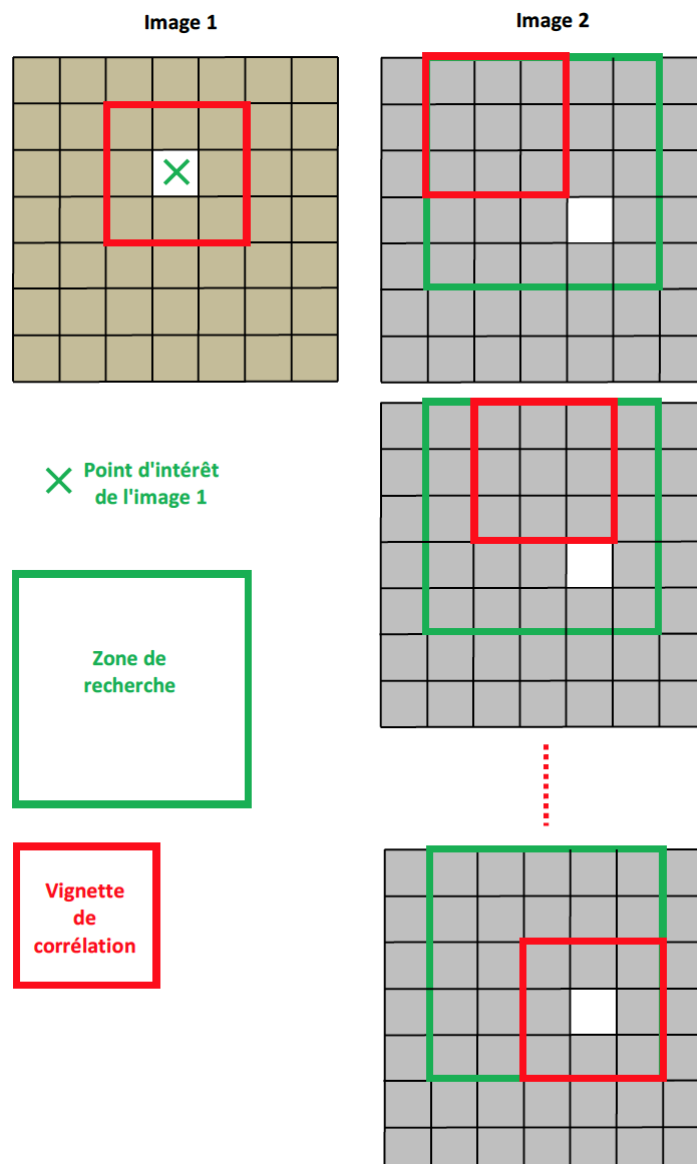
Après avoir détecté des points d'intérêts dans chacune des deux images, il faut dans un second temps les appairer. Ainsi dans l'exemple qui suit, on doit mettre en relation le coin supérieur gauche d'un passage piéton de l'image 1 avec son homologue dans l'image 2 et procéder de même pour le coin inférieur droit. Pour réaliser ces appariements deux approches s'offrent à nous : appariement par corrélation ou utilisation de descripteurs de points d'intérêt.



4.2.2 Appariement par corrélation

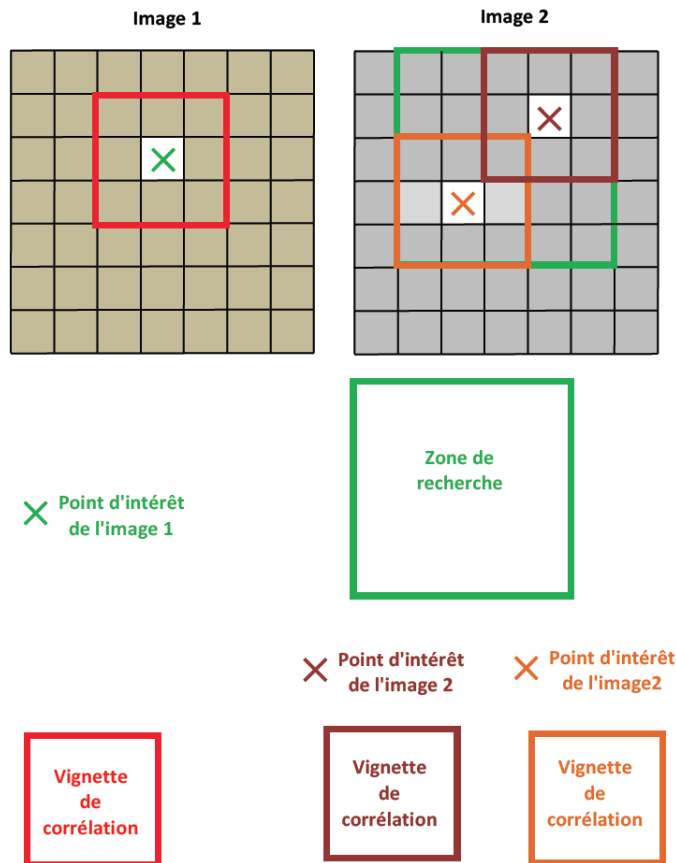
- Stratégie point::libre :

Ayant repéré un point d'intérêt dans l'image 1, on va chercher dans l'image 2 le point qui lui ressemble le plus au regard du score de corrélation. Pour cela on définit une zone de recherche dans l'image 2 (en vert sur la figure) dans la quelle on va déplacer la vignette de corrélation (en rouge sur la figure). Le pixel retenu sera celui pour lequel le score de corrélation entre la vignette de l'images 1 et les vignettes de l'image 2 sera maximal.



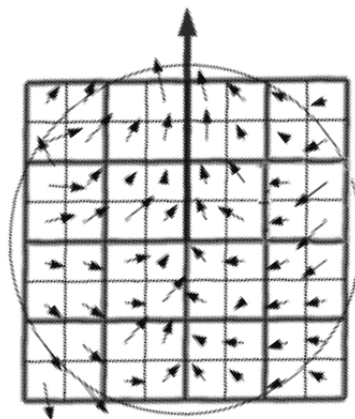
- Stratégie point::point :

Ayant repéré un point d'intérêt dans l'image 1 on définit une zone de recherche dans l'image 2 et on calcule les scores de corrélation entre la vignette placée sur le point d'intérêt de l'image 1 et les vignettes placées sur les points d'intérêt de l'image 2 qui se trouvent dans la zone de recherche. Le pixel retenu sera celui pour lequel le score de corrélation entre la vignette de l'image 1 et celles de l'image 2 sera maximal (ici le point bordeaux de l'image 2 obtient un score supérieur et est donc apparié).



4.2.3 Appariement via descripteurs de points

Lors de la détection de points d'intérêt, on localise les points d'intérêt et on calcule simultanément un descripteur qui caractérise le point et son voisinage. Ainsi le détecteur de points SIFT calcule pour chaque point d'intérêt repéré un descripteur basé sur le gradient de son voisinage. Ces descripteurs peuvent être calculés à différentes échelles et orientés selon la direction principale du gradient. Ainsi on obtient un processus invariant aux changements d'échelle et à la rotation.

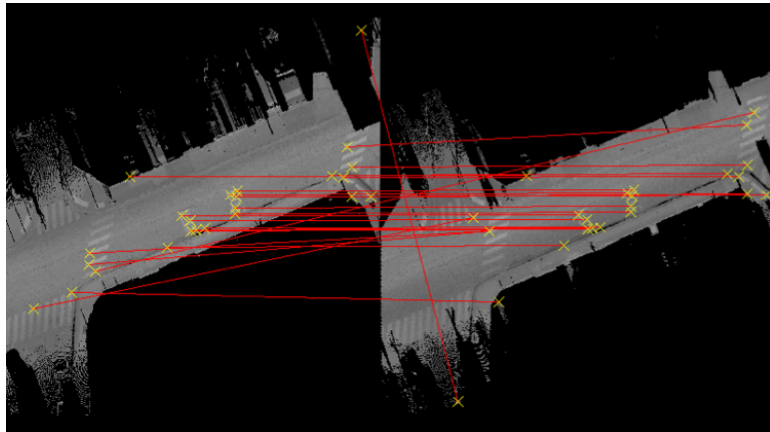


Descripteur de voisinage fondé sur le gradient

Tests et résultats

Détermination du vecteur translation

Après avoir détecté un ensemble de points d'intérêt, on a réalisé l'appariement de ceux-ci. La visualisation des appariements met en évidence l'existence d'appariements erronés. Il s'agit désormais donc de filtrer les appariements obtenus afin de ne conserver que ceux qui permettront de déterminer la translation moyenne existant entre image 1 et image 2.



Visualisation des appariements entre image 1 et image 2

5.1 Protocole expérimental

Les tests ont été menés suivant trois vagues au fur et à mesure de l'avancement des méthodes de calcul de la translation (codées en C++) :

1. des tests avec un petit jeu de données et cinq paires d'images
2. des tests avec le filtrage par écart-type sur neuf paires d'images censées rendre compte de tous les cas possibles (présence de masque, faible recouvrement inter-LIDAR, peu de détails, ...)
3. les mêmes tests que précédemment avec cette fois-ci les filtrages par bande-passante et avec prédicteur.

Le protocole générique pour construire les fichiers d'images suivait l'arborescence suivante :

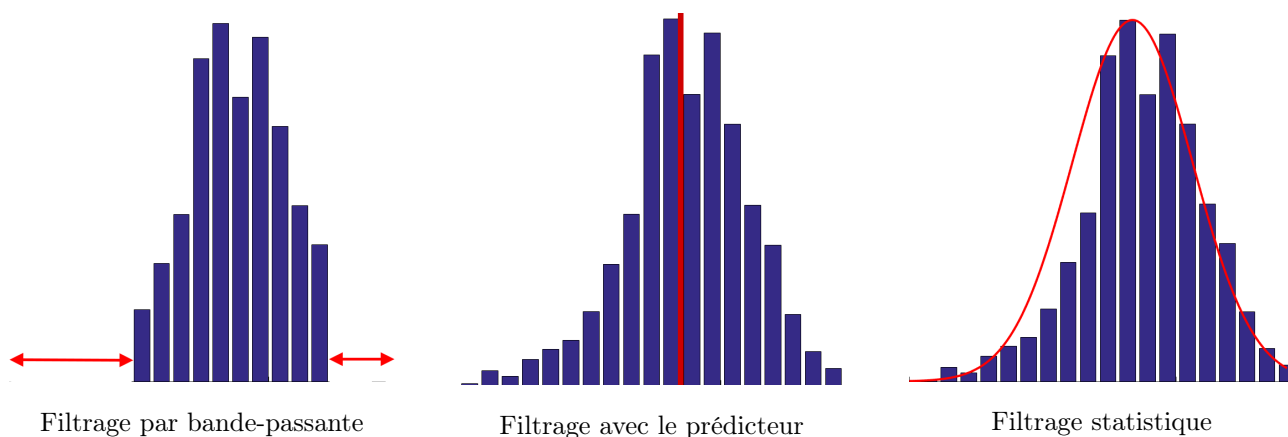
```
image1.image2/  
├── image1.tif  
└── image2.tif
```

A partir de cet arborescence, les tests créaient des fichiers de résultats selon une manière standardisée afin de faciliter la lecture ultérieure de ces fichiers :

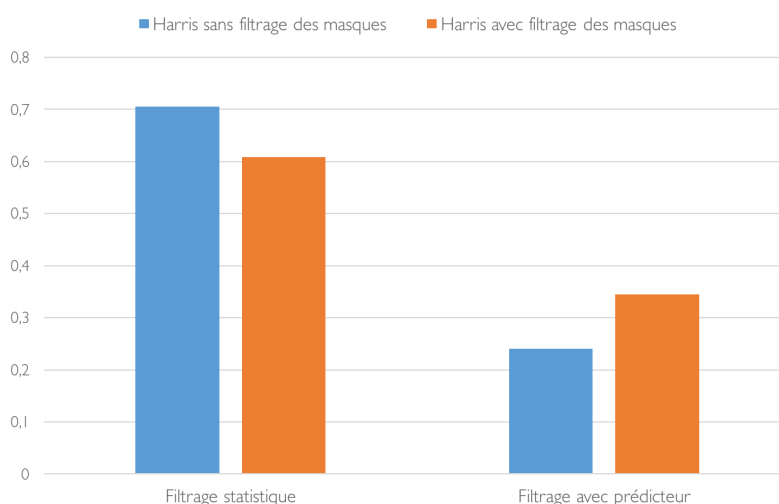
```
image1.image2/  
├── image1.image2.nom_méthode1.rest_filtrage  
└── image1.image2.nom_méthode2.analyses_filtrage
```

Les extensions `.rest_` correspondant aux fichiers contenant les points restants après filtrage, et les `.analyses_` résumant les résultats du test. Les extensions `_filtrage` correspondent à l'abréviation choisie pour la méthode de filtrage :

- `_desc` pour le filtrage avec prédicteur
- `_basic` pour le filtrage par bande-passante
- rien pour le filtrage par écart-type



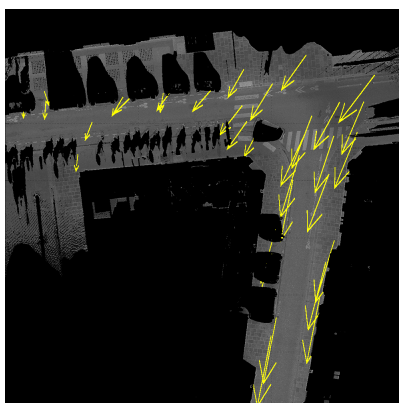
La méthode à choisir n'est pas seulement l'algorithme de détection ou de corrélation de points, c'est aussi la façon de traiter les résultats bruts et par essence même erronés. Après tous les calculs, il apparaît que le filtrage avec prédicteur est plus performant que le filtrage statistique (le filtrage par bande-passante ayant été quant à lui destiné aux tests préliminaires) :



Comparaison des filtrages dit « statistique » et « avec prédicteur »

Cependant il faut noter que le filtrage avec prédicteur a une limite à prendre en compte, à savoir que lorsque le prédicteur (i.e. le résultat de corrélation image entière) est faux, tout le traitement qui s'ensuit est faux.

L'autre point noir est le fait que parfois la translation entre deux images n'est pas unique comme le montre l'image suivante :

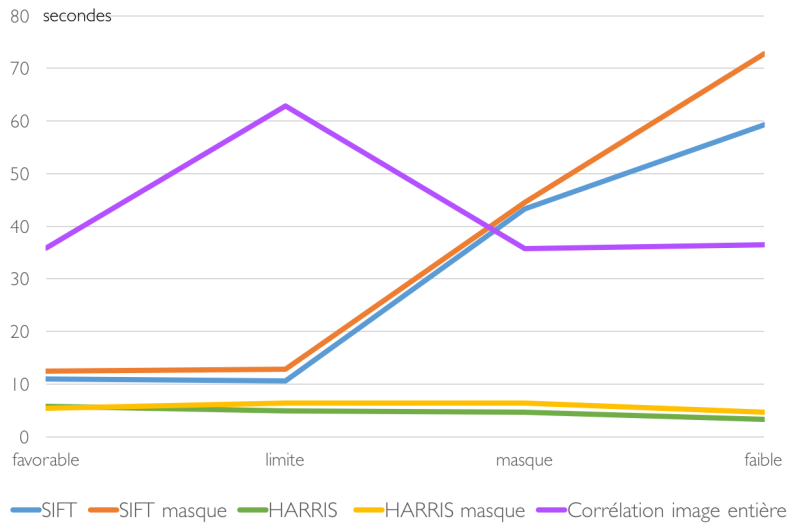


Cas d'une translation multiple entre deux images

C'est ainsi que nous privilégierons le filtrage fondé sur l'écart-type plus robuste dans la suite.

Critère de rapidité

Le dernier argument de poids dans le choix de la méthode à implémenter est la vitesse d'exécution d'ensemble du programme. Il faut contrôler les délais car sur un gros échantillon constitué de dizaines voire centaines d'images, les traitements sont décuplés et le temps joue alors un rôle important. Pour calculer ces délais et les extrapoler, nous avons utilisé un *timer* (TIMETHIS.exe) et obtenu les courbes suivantes :



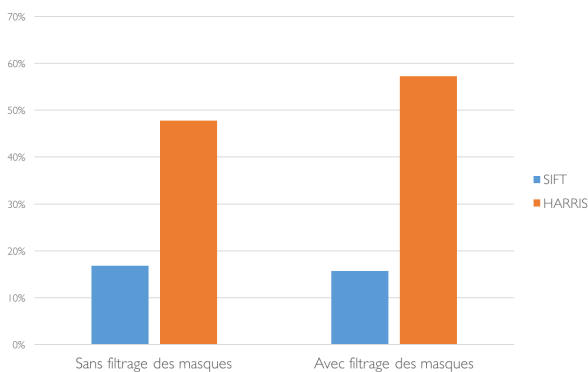
Mesures du temps d'exécution des différentes méthodes

On constate que Harris est le plus rapide, et que SIFT voit sa vitesse se dégrader avec la difficulté des cas rencontrés. On remarque également le temps élevé que met corrélation image entière à fournir la translation, ce qui nous conforte dans notre choix d'abandonner le filtrage avec prédicteur (n.b. le prédicteur est le résultat de corrélation image entière). Noter que le temps de calcul de la translation avec les exécutables que nous avons produits, est négligeable devant les temps évoqués ci-dessus.

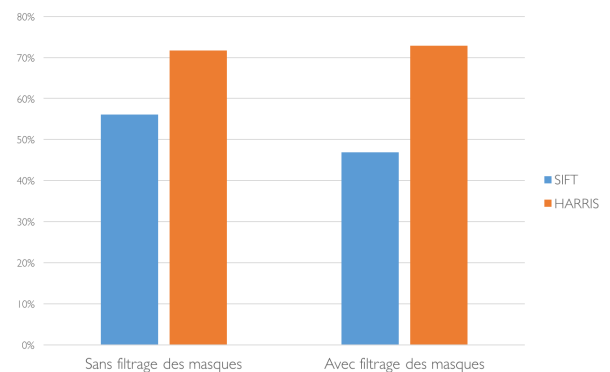
5.2.2 Résultats des filtrages

Les résultats exposés dans cette section concernent les filtrages. Le premier aspect discriminant des filtrages est le nombre de points éliminés :

Pourcentage de points supprimés



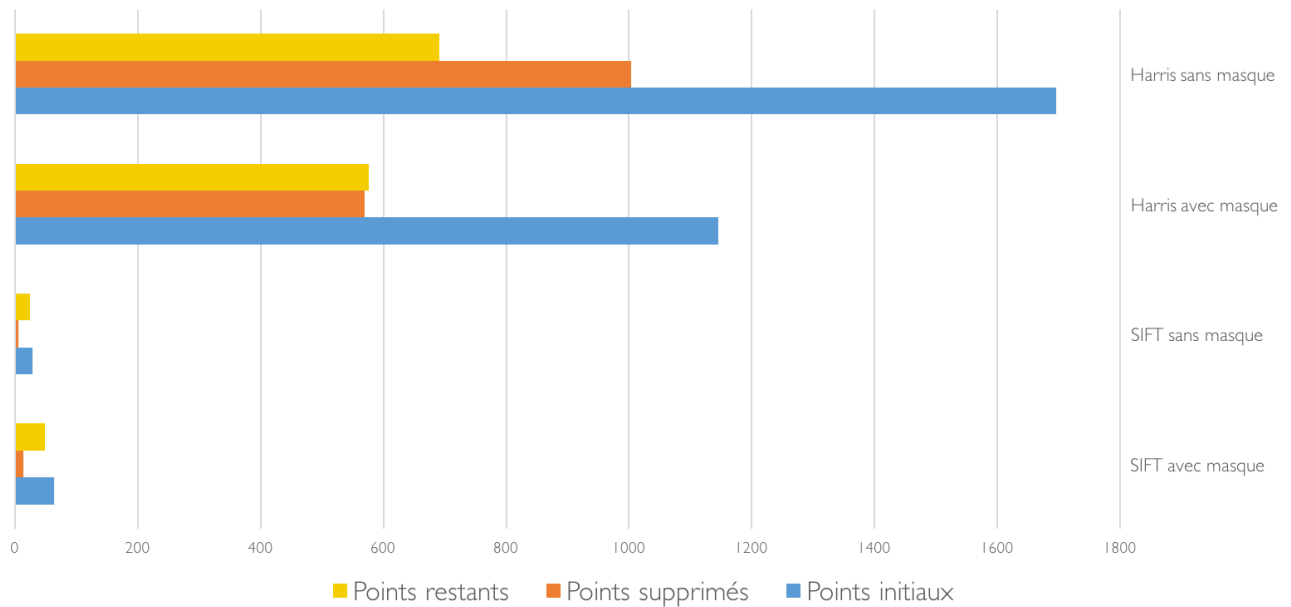
Filtrage par l'écart-type



Filtrage avec le prédicteur

On constate que le filtrage des masques n'a pas d'influence notable mais en revanche les algorithmes de détection de points sont nettement plus déterminants : dans le premier filtrage dit « par écart-type » environ 16 % des points de SIFT sont éliminés alors que près de 60 % des points de Harris sont considérés comme des erreurs. Ces écarts sont moins nets dans le cas du filtrage qui prend le résultat de corrélation image entière comme prédicteur. Néanmoins on observe qu'en proportions, il y a plus de points SIFT qui sont gardés que de points de Harris.

Pour ce qui concerne les effectifs à proprement parler, Harris donne en moyenne 1500 points d'intérêt, tandis que SIFT trouve 50 points. Sachant que Harris fournit plus de points en nombre que SIFT, même si beaucoup sont éliminés, ceux qui restent permettent de déterminer une translation précise et unique, contrairement à SIFT qui est limité par son faible échantillon de base.



Représentation des effectifs en fonction des différentes méthodes de filtrage

Conclusion

La constitution d'un état de l'art solide nous a permis d'avoir une idée de ce qui correspondait au mieux à notre problématique et l'objectif, qui était au début orienté sur les comparaisons des différentes méthodes existantes, a été redéfini au cours du projet lorsque nous avons pris le parti de créer des programmes qui cherchent non plus à apparier les images, mais aussi à esquisser le recalage des trajectographies du véhicule « Stéréopolis ». Nous avons pu ainsi déterminer la meilleure méthode pour répondre au problème d'automatisation du recalage en proposant :

Harris avec filtrage des masques et filtrage des erreurs par des méthodes statistiques

Ce projet a été pour nous l'occasion de découvrir la vie au sein d'un laboratoire de recherche sur le plan humain et professionnel. Nous avons connu toutes les étapes du cheminement du chercheur : d'abord la découverte de ses données et de son cahier des charges ; puis la prise en main des fonctionnalités existantes grâce à l'état de l'art ; ensuite les phases d'expérimentations qui donnent des résultats, ceux-ci conduisant à une reformulation partielle des fonctionnalités à produire (cycle qui se répète plusieurs fois avant d'arriver au programme satisfaisant) ; enfin la mise en forme et l'analyse qui conduisent au présent document.

Les perspectives de développement et d'intégration au sein du processus de recalage des trajectographies sont envisagées mais il reste de nombreux contrôles à établir afin de pouvoir passer en opérationnel.