HAL
open science

Traitement automatique de time-lapses stéréoscopiques de milieux naturels : application à la sureveillance de glaciers alpins

Héla Hadhri

To cite this version:

Héla Hadhri. Traitement automatique de time-lapses stéréoscopiques de milieux naturels : application à la sureveillance de glaciers alpins. Traitement des images [eess.IV]. Université Savoie Mont Blanc, 2020. Français. NNT : 2020CHAMA012 . tel-03534542

HAL Id: tel-03534542

https://theses.hal.science/tel-03534542

Submitted on 19 Jan 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Thèse présentée pour obtenir le grade de docteur

Université de Savoie Mont-Blanc

Laboratoire LISTIC
École doctorale SISEO

Discipline : STIC - traitement de l'information

Traitement automatique de time-lapses stéréoscopiques de milieux naturels

application à la surveillance de glaciers alpins

Par : **Héla HADHRI**

Sous la direction de Emmanuel Trouvé, Professeur des Universités
Sous la co-direction de Abdourahmane Atto, Maître de conférences
Sous la supervision de Flavien Vernier, Maître de conférences

Membres extérieurs:

Président du jury : Marie-Odile Berger, Dir. de recherche, INRIA / LORIA, Nancy.

Examinateur : Antoine Rabatel, Chargé de recherche, IGE, Grenoble.

Rapporteur : Cédric Demonceaux, Prof. des Universités, Le2i, Le Creusot.

Rapporteur : Jean-Philippe Malet, Dir. de recherche, EOST / IPGS, Starsbourg.

Date de soutenance : 25/06/2020

Résumé

Au cours de la dernière décennie, la photogrammétrie multi-vues et les techniques de suivi d'objets basées sur des données d'imagerie proximale ont permis de générer, avec un coût relativement faible, divers types de données décrivant directement ou indirectement la forme et la dynamique tridimensionnelles de milieux intérieurs, urbains et naturels. En milieux naturels, des évolutions technologiques et méthodologiques restent nécessaires afin de prouver le potentiel de tels outils pour des applications géo-morphologiques précises, métriques et automatiques. En effet, la plupart des techniques existantes ne sont pas particulièrement conçues et pensées afin de traiter des données de hautes définitions spatiale et temporelle, entachées d'aléas photométriques et atmosphériques importants et caractérisées par une visibilité variable et un déplacement dense et semi-rigide.

Nous essayons dans cette thèse de résoudre ces problématiques en proposant une chaîne de traitements automatique de séries temporelles d'images (time-lapses) stéréoscopiques pour l'observation de mouvements gravitaires tels que le déplacement de glaciers alpins. Nous étudions tout d'abord les différentes stratégies de traitement possibles selon une taxonomie à plusieurs niveaux des données à disposition : monoscopiques et stéréoscopiques, monodates et bidates, temporelles et multi-temporelles. Nous proposons ensuite une suite complémentaire de traitements automatiques robustes qui repose sur un choix justifié de représentations et de modèles géométriques et temporels régissant le lien entre ces différents types de données. En particulier, nous évaluons l'incertitude via une approche géométrique et maximisons la cohérence des estimations de vélocité métrique issues des briques fondamentales de la chaîne de traitements via l'optimisation d'une fonctionnelle tenant compte de descripteurs spatio-temporels, d'une contrainte de fermeture temporelle des champs de déplacements intra time-lapse et enfin du sens et de la stationnarité de l'écoulement de glaciers alpins.

Ces travaux s'inscrivent dans le contexte du réchauffement climatique et son impact sur la cryosphère. La surveillance proximale continue de glaciers alpins via un système d'observation dédié ouvre en effet de nouvelles perspectives quant à une quantification d'indicateurs métriques corrélés à l'évolution des températures et des précipitations. Nous appliquons les méthodes proposées sur des données acquises sur le site du glacier d'Argentière dans le massif du Mont-Blanc. Les séries temporelles de suivi métrique géoréférencé de la vélocité d'une partie de ce glacier sont ainsi générées avec une estimation de leurs incertitudes sur l'ensemble des données acquises en 2018.

Les résultats montrent qu'il est possible de développer des applications robustes de surveillance proximale géo-photogrammétrique en milieux climatiques extrêmes et non-contrôlés sous réserve de prise en compte des caractéristiques de la mesurande et des connaissances a-priori sur la scène observée.

Abstract

Over the past decade, multiview photogrammetry and object tracking techniques based on proximal imaging data have made it possible to generate, at a relatively low cost, various types of data describing directly or indirectly the three-dimensional shape and dynamics of indoor, urban and natural environments. In natural environments, technological and methodological developments are still needed to prove the potential of such tools for precise, metric and automatic geomorphological applications. Most of the existing techniques are not particularly designed and thought out in order to process data of high spatial and temporal definition, marred by important photometric and atmospheric artifacts and characterized by variable visibility and a dense and semi-rigid displacement.

In this thesis, we try to solve these problems by proposing an automatic processing chain of stereoscopic time-lapse images for the observation of landslides such as the flow of alpine glaciers. We first study the different possible strategies for processing stereoscopic time-lapses of natural environments according to a multi-level taxonomy of the available data : monoscopic and stereoscopic, monodate and bidate, temporal and multi-temporal. We then propose a complementary suite of robust automatic processings based on a justified choice of geometric and temporal representations and models governing the link between these different types of data. In particular, we evaluate data uncertainty via a geometric approach and maximize the coherence of the metric velocity estimates derived from the fundamental bricks of the processing chain via the optimization of a functional taking into account spatio-temporal descriptors, a temporal closure constraint of intra-timelapse displacement fields and finally the direction and stationarity of the flow of alpine glaciers.

This work is motivated by the current context of global warming and its impact on the cryosphere. The continuous proximal monitoring of alpine glaciers via a dedicated observation system opens up new perspectives for the quantification of metric indicators correlated to the evolution of temperature and precipitation. We apply the proposed methods to data acquired on the Argentière glacier site in the Mont-Blanc Massif. Tracking time series of the geo-referenced metric velocity of part of this glacier are thus generated with an estimate of their uncertainties on the data set acquired in 2018.

The results show that it is possible to develop robust geo-photogrammetric proximal monitoring applications in extreme and uncontrolled climatic environments, provided that the characteristics of the measurand and a-priori knowledge of the observed scene are taken into account.

Remerciements

Je tiens tout d'abord à remercier cordialement l'agence nationale de la recherche pour avoir financé le projet Pheonix [1]- Parcimonie, observations non-stationnaires de grandes dimensions, modélisation des séries chronologiques d'images et télédétection - dans lequel s'inscrit cette thèse.

Je remercie Pr. Emmanuel Trouvé de s'être impliqué dans ma thèse malgré son emploi du temps et pour ses conseils qui ont beaucoup alimenté ma réflexion, Dr. Flavien Vernier pour son aide précieuse et sa présence et Dr. Atto pour sa bienveillance et pour les débats que nous avons pu avoir. Merci à vous trois pour votre soutien et de m'avoir donné l'occasion de vivre cette expérience très enrichissante.

jugéJe tiens à remercier Pr. Cédric Démonceaux et Pr. Jean Philippe Malet d'avoir rapporté ce travail de recherche. Les remarques et critiques que vous m'avez prodigué m'ont permis de perfectionner ma réflexion sur différents aspects du sujet traité. Je remercie Pr. Marie Odile berger et Pr. Antointe Rabatel d'avoir jugé mon travail. J'ai particulièrement apprécié le débat pluridisciplinaire qui a pu avoir lieu le jour de la soutenance.

Je remercie Dr. Xavier Bodin et Emmanuel Malet de l'équipe du laboratoire EDYTEM avec qui j'ai eu l'immense plaisir de sortir en mission terrain. Les moments passés en altitude resteront à jamais gravés dans ma mémoire.

Je remercie également toute l'équipe pédagogique du département QLIO de l'IUT d'Annecy de m'avoir permis une très agréable expérience en enseignement.

À Ozzy, Etienne, Lina, Mat, Batal, Nadia, Sirine, Narjess et Emna ; merci d'avoir été ma famille pendant les moments heureux et moins heureux des différentes phases de la thèse.

Enfin et surtout, merci à ma famille pour avoir toujours cru en moi : tous les éloges du monde ne sauraient traduire l'amour et la gratitude que je vous porte.

1. https://anr.fr/Projet-ANR-15-CE23-0012

En hommage à mon oncle Jalel fervent scientifique et avant tout dont la bonté et l'humour illuminaient la vie des Hadhri toutes générations confondues.
Je te dédie les efforts et l'énergie que j'ai fourni dans l'accomplissement de cette thèse de doctorat.
Puisses-tu reposer dans nos coeurs et illuminer nos âmes.

Table des matières

Introduction générale et contexte **3**

I Étude de l'existant pour le traitement de time-lapses stéréoscopiques de milieux naturels **8**

1 Concepts et modèles mathématiques 9
1.1 Introduction 9
1.2 Reconstruction par inférence de mesures 10
1.2.1 Inversion de modèles paramétriques 10
1.2.2 Régularisation 13
1.2.3 Pondération 14
1.2.4 Prises en comptes d'observations aberrantes 15
1.3 Transformations géométriques et invariants 15
1.4 Évaluation des incertitudes 17
1.4.1 Définition d'incertitude 17
1.4.2 Propagation analytique des incertitudes 17
1.4.3 Propagation numérique des incertitudes 18
1.4.4 Dilution géométrique de la précision (GDOP) 19
1.5 Conclusion 20
2 Méthodes monodates et bidates monoscopiques 21
2.1 Introduction 21
2.2 Caractérisation des données images de milieux naturels 22
2.3 Sélection et prè-traitements 23
2.3.1 Méthodes basées sur un modèle physique 23
2.3.2 Méthodes basées texture 25
2.3.3 Méthodes basées apprentissage 29
2.4 Estimation du déplacement des pixels 29
2.4.1 Approches denses 29
2.4.2 Approches éparses 35
2.5 Alignement paramétrique 39
2.6 Conclusion 41
3 Méthodes monodates et bidates stéréoscopiques 43
3.1 Introduction 43
3.2 Caractérisation des données stéréoscopiques en milieux naturels 44

3.3 Instrumentation et modélisation de la caméra perspective 44
3.3.1 Modèle optique d'un système d'observation 44
3.3.2 Instrumentation et système d'observation 47
3.4 Reconstruction non-métrique 48
3.4.1 Homographies et projections perspectives 48
3.4.2 Modélisation homogène 50
3.4.3 Vision stéréoscopique 50
3.4.4 Contraintes stéréoscopiques 52
3.4.5 Triangulation 56
3.5 Reconstruction de champs de vélocités *East, North, Up* 57
3.5.1 Reconstruction métrique 57
3.5.2 Géo-référencement 58
3.5.3 Projection de vélocité pixellique sur un modèle 3D 58
3.6 Optimisation et robustesse 59
3.6.1 Lors de l'estimation des contraintes stéréoscopiques 59
3.6.2 Lors de la triangulation 60
3.6.3 Lors de l'utilisation de coordonnées homogènes 61
3.7 Estimation de l'incertitude dans les mesures 3D 61
3.8 Conclusion 63
4 Méthodes de traitements temporels et multi-temporels 65
4.1 Introduction 65
4.2 Caractérisation des données temporelles issues de time-lapses 66
4.3 Méthodes classiques pour le traitement de séries temporelles 67
4.4 Méthodes de traitement de séries multi-temporelles 68
4.4.1 Symétrie bidate 69
4.4.2 Symétrie multitemporelle 69
4.5 Conclusion 71

II Chaîne de traitements proposée 75
5 Site expérimental et stratégie 77
5.1 Introduction 77
5.2 Description du site et motivation 77
5.2.1 Site surveillé 77
5.2.2 Motivation 79
5.3 Instrumentation du site et données acquises 84
5.4 Stratégie globale 89
5.5 Conclusion 91
6 Traitements monodates et bidates monoscopiques 93

6.1 Introduction 93
6.2 Pré-traitements 93
6.2.1 Sélection automatique des données à traiter 94
6.3 Estimation de la vélocité des pixels 97
6.3.1 Stratégie de recalage 97
6.3.2 Calcul de vélocité pixellique 99
6.4 Conclusion 106
6.5 Perspectives 107
7 Traitements monodates et bidates stéréoscopiques 109
7.1 Introduction 109
7.2 Mise en correspondance et calcul de la pose des caméras 109
7.3 Modèle numérique de terrain et stratégie adoptée 115
7.3.1 Reconstruction 3D 115
7.3.2 Stratégie adoptée pour la modélisation du terrain 116
7.4 Mesures bidates stéréoscopiques 118
7.5 Propagation d'incertitude 123
7.5.1 Modèle proposé 123
7.5.2 Résultats d'incertitude tridimensionnelle 126
7.6 Conclusion 127
7.7 Perspectives 131
8 Proposition de traitements en séries temporelles 133
8.1 Introduction 133
8.2 Données temporelles 134
8.2.1 Données de modélisation temporelle de l'écoulement glaciaire 134
8.2.2 Mesures d'écoulement glaciaire non régularisées 135
8.3 Traitements temporels 140
8.3.1 Méthode proposée 140
8.3.2 Paramétrisation, résolution et résultats 161
8.4 Conclusion 167
8.5 Perspectives 171

Conclusion générale et perspectives **173**

Liste des notations

$\mathrm{A}_{\mathbf{n}}^{\mathbf{2}}\mathbf{t}$	Ensembles (arrangements) de données temporelles bidates sur n dates
2×(2D+t)	Deuxième type de données bidates stéréoscopiques
$\epsilon_{\tau_{\text{sim}}}$	Erreur spatio-temporelle quadratique moyenne des mesures de vélocité (τ =LF) et de position (τ =CM) sur les données simulées.
γ	La pondération donnée au terme de pénalisation de la variation totale
$\hat{\mathbf{F}}_\tau$	Une des composantes régularisées d'une série chronologique de suivi sur les dates T
$\hat{x}$	Estimation de la variable aléatoire x
$\hat{x}_0$	Estimation initiale de la variable aléatoire x
$\hat{x}_{LS}$	Estimation de la variable aléatoire x obtenue par minimisation de l'erreur quadratique
$\mathbb{X}_{MNT}$	Nuage de points retenu en tant que Modèle Numérique de Terrain minimisant la dispersion de reproductibilité
$\mathbf{F}_0$	Vecteur contenant l'ensemble des données $\mathrm{A}_{\mathbf{n}}^{\mathbf{2}}\mathbf{t}$
OF	Champs de vélocités calculés avec la méthode du flot optique
SF	*Scene-flow* / Flot de scène, i.e, vecteur de déplacement ou de velocité tridimentionnelles
$\mathbf{x}'_c$	Projection, sur le plan image de la caméra c, d'un point de contrôle détecté sur le plan image de la caméra c'
$\mathbf{x}_c$	Point de contrôle détécté sur la caméra c
$\mathcal{E}(k_x, k_y)$	Énergie associée à une image dans le domaine fréquentiel
$\mathcal{E}(x, y)$	Énergie associée à une image dans le domaine spatial
$\mathscr{N}$	Nombre d'images dans les time-lapses acquis : 6 images par jour.
μ_Z^2	Carte d'incertitude de triangulation en profondeur
μ_{XY}	Carte d'incertitude de triangulation en position
Φ_τ	Application régissant les différentes contraintes de cohérence temporelle entre les différents arrangements d'images d'un time-lapse : Φ_{LF} ou Φ_{CM}

Π	Plan épipolaire
$\propto$	Proportionnel à
Σ	Matrice de variance / covariance
σ	Écart-type
ϱ	Fonction de coût
A	Application de mesures entachées d'un bruit additif
A_∞	Lumière ambiante d'une scène observée
$A_{d-\theta}$	Matrice de co-occurence
B	Baseline spatiale ou distance géodésique entre deux caméras d'un système d'observation stéréoscopique
CC	Corrélation-croisée centrée normalisée
F_τ	La composantes τ d'une série chronologique de suivi sur un time-lapse
F_c	La distance focale d'une caméra c en pixels
f_c	La distance focale métrique d'une caméra c
$J(\mathbf{x})$	Radiance d'une scène observée
J_x	Matrice jacobienne d'un opérateur par rapport à x
K_c	Matrice intrinsèque d'une caméra c
l_c	Ligne épipolaire associée à une caméra c
N	Nombre d'images dans les time-lapses journalier acquis : 1 image retenue par jour.
n	Nombre d'images dans les time-lapses journalier acquis ayant été retenues après classification.
P_c	Matrice projection d'une caméra c
R_T	Matrice de rotation/translation entre deux caméra
S_c	Centre optique d'une caméra c
$t(\mathbf{x})$	Facteur de transmission de la lumière sur une scène observée
T_G	Indice de visibilité des crevasses du glacier d'Argentière.
T_S	Taux d'enneigement sur une zone locale de l'image
W	Matrice de pondération

W_0	Matrice de pondération des observations $\mathbf{F}_0$
w_{s_Δ}	Fenêtre rectangulaire de taille s et translatée de $(\frac{\Delta}{2}, \frac{\Delta}{2})$
X	Point 3D obtenu par triangulation stéréoscopique
y	Données observées
(2×2D)+t	Premier type de données bidates stéréoscopiques
2×2D	Données mono-date stéréoscopiques
2D+t	Données bidates monoscopiques
#n 2D	Ensemble de n images dans un time-lapse monoscopique
CM	Approche *Common-Master* pour la construction d'une contrainte de fermeture temporelle du type SBAS (Small BAseline Subset)
DMDF	Méthode de calcul de la velocité bidate proposée : DeepMatching et DeepFlow
E	La matrice essentielle d'un système d'observation stéréoscopique
GDOP	Méthodes de dilution géométrique de la précision (*Geometric Dilution Of Precision*)
LF	Approche *Leap-Frog* pour la construction d'une contrainte de fermeture temporelle du type SBAS (Small BAseline Subset)

Introduction générale et contexte de la thèse

Contexte et motivation

La caractérisation de la topographie et de la dynamique d'objets géophysiques a fait historiquement appel à des observations qualitatives et quantitatives effectuées sur le terrain qui ont progressivement évolué vers des techniques de géo-positionnement et d'imagerie satellitaires ou encore d'outils de reconstruction tels que les scanners laser terrestres à précision submillimétrique.

Ces dernières décennies, les avancées en Intelligence Artificielle et les besoins de caractérisation et de modélisation de structures tridimensionnelles d'une façon continue et automatique, par exemple en navigation robotique ou encore en réalité augmentée, ont conduit à l'émergence de méthodes basées sur l'imitation de la perception visuelle humaine du mouvement parallaxe et de la texture à partir de données d'imagerie satellitaires et optiques.

Il existe effectivement de nos jours des modèles géométriques qui permettent d'extraire la topologie surfacique d'objets visualisés à travers un appareil photo en mouvement et une multitude de méthodes statistiques et paramétriques qui permettent une description locale ou globale d'images numériques qui peuvent être déployées de façon continue afin d'améliorer les techniques d'observation de sites naturels.

Plus précisément, les modèles de géométrie épipolaire et de description robuste et invariante de la structure spatiale des pixels ainsi que les avancées théoriques et technologiques quant à la prise en compte du bruit de mesure et aux techniques d'optimisation convexe et non-convexe permettent théoriquement une modélisation 3D dynamique à partir d'une succession d'images acquises via un système d'observation stéréoscopique optique.

Dans un contexte d'industrialisation progressive de la planète, la surveillance proximale continue de glaciers alpins, objets géophysiques réagissant relativement rapidement aux changements climatiques locaux, via un système d'observation dédié ouvre de nouvelles perspectives. Une quantification d'indicateurs métriques précis corrélés à la hausse de la température de la cryosphère et de l'atmosphère, principalement induite par les émissions de gaz à effet de serre d'origine anthropique, pourrait effectivement servir à une prise de conscience de la gravité de la situation et à la prévention des risques directs et indirects.

Problématique

La problématique réside principalement dans la réalisation que la plupart des applications et outils existants en vision par ordinateur et en géo-photogrammétrie ne soient pas particulièrement pensés et conçus afin de traiter automatiquement une masse de données acquises dans un environnement non-contrôlé tout en fournissant en sortie des résultats métriques, cohérents et précis.

Afin de répondre à cette problématique, nous proposons dans cette thèse une étude des différentes stratégies possibles et une chaîne de traitements détaillée dédiée à des données "time-lapses stéréoscopiques" de hautes résolution et définition, spatiale et temporelle, entachées d'aléas photométriques et atmosphériques importants, caractérisées par une visibilité variable et un déplacement dense et non-rigide, tel le déplacement glaciaire, et enfin une topographie difficilement paramétrique.

Structure de la thèse

Ce manuscrit est divisé en deux parties. La première partie constitue une étude théorique selon une taxonomie à plusieurs niveaux des outils qui permettent d'implémenter une application de suivi optique stéréoscopique où nous nous plaçons d'emblée dans un contexte de surveillance proximale de milieux naturels. Le niveau le plus grossier de cette taxonomie, dépeint sur le tableau 1, est induit par les différents types de données présents dans un time-lapse stéréoscopique et définit les différents chapitres de chaque partie.

La partie I est ainsi structurée en 3 chapitres détaillant respectivement les méthodes de traitement de données bidimensionnelles (chapitre 2), tridimensionnelles (chapitre 3) et temporelles (chapitre 4) précédés par un préambule (chapitre 1) dans lequel nous exposons des concepts mathématiques qui reviendront de façon récurrente dans le reste du manuscrit. La seconde partie du manuscrit se veut le miroir de la première avec 3 chapitres, 6, 7 et 8, détaillant respectivement les méthodes proposées pour traiter des données réelles d'observation de glaciers alpins avec, dans chaque chapitre, une caractérisation détaillée des données dont il est question, une justification des choix de traitement et les résultats obtenus. Dans le chapitre de préambule de la partie II (chapitre 5), nous expliciterons davantage le contexte dans lequel nous nous plaçons, les détails pratiques sur le système d'observation stéréoscopique déployé, son installation, la prise de mesures et une vue d'ensemble sur la stratégie adoptée.

Cette stratégie repose sur la construction d'une chaîne de traitements dont les briques sont complémentaires et suffisantes pour traiter un ensemble de données time-lapses stéréoscopiques d'une manière quasi automatique, de la sélection et du pré-traitement des données pertinentes jusqu'à la régularisation et au débruitage qui permettent de rendre exploitables les mesures. Comme nous le verrons dans la deuxième partie, une brique de

Données			approche :
Bidimensionnelles	#N 2D	Ensemble de N images	Traitements monodates monoscopiques
	2D+t	Deux images de la même scène prises à des temps différents	Traitements monoscopiques bidates
Tridimensionnelles	2×2D	Un couple stéréoscopique d'images i.e. deux images de la même scène prises en même temps avec une perspective différente	Traitements stéréoscopiques (reconstruction 3D)
	(2×2D)+$\mathbf{t}$	Deux couples stéréoscopiques d'images prises à des temps différents	Traitements bidates stéréoscopiques - approche 1
	2×(2D+$\mathbf{t}$)	Deux images prises à des temps différents et leurs homologues stéréoscopiques respectifs	Traitements bidates stéréoscopiques - approche 2
Temporelles	$\mathrm{A}_{\mathbf{n}}^{\mathbf{2}}\mathbf{t}$	Ensembles (arrangements) de données temporelles bidates mesurables sur un time-lapse contenant des acquisitions à traiter sur n dates	Traitements redondants et symétriques sur une série temporelle d'images
	$\mathbf{t}$	Séries temporelles, i.e. chronologiques, de mesures monodates ou bidates	Traitements en séries temporelles

TABLE 1 – Les différents types de données dans un time-lapse stéréoscopique.

sélection manuelle de points d'ancrage sur les images et de classification très grossière de l'ensemble des données à traiter selon leur aspect saisonnier sera à effectuer une fois afin de rendre le reste des traitements automatiques.

Enfin, un chapitre de conclusion viendra clore ce manuscrit en revenant sur les hypothèses retenues tout au long de la thèse et en énonçant les perspectives pour la continuation du développement du domaine de la vision par ordinateur pour l'observation continue de milieux naturels complexes.

Contribution

Nous nous situons dans la continuité des travaux existants dans le développement de chaînes de traitements optiques, stéréoscopiques et continues appliquées à l'observation des glaciers [1][2][3][4]. Afin de créer de la robustesse, nous avons combiné des ressources conceptuelles et théoriques d'une manière inédite en posant les objectifs suivants :

— Formuler un protocole d'acquisition et de pré-traitements des données tenant compte de l'instabilité des paramètres intrinsèques et extrinsèques du système d'observation dans le temps.

— Sélectionner automatiquement l'ensemble des données à traiter parmi les données acquises (données monodates) en fonction de la visibilité et de la texture du glacier.

— Déployer une procédure de suivi appropriée sur les données bidates monoscopiques. Le but ici est de disposer d'un outil capable de mesurer la vélocité sous-pixellique entres deux images de très haute résolution dont l'aspect (texture, couleur, dégradation, etc) est très variable, en intégrant des heuristiques nous permettant de disqualifier les mesures incohérentes dans le reste des traitements.

— Estimer la topologie métrique de la scène observée afin de remonter à des déplacements métriques et géo-référencés (East, North, Up). Plusieurs stratégies de reconstruction 3D (données stéréoscopiques) et de calcul de la vélocité métrique (couples bidates stéréoscopiques) sont alors étudiées selon les données utilisées conjointement.

— Estimer l'incertitude des mesures effectuées sur le plan image en fonction de la profondeur d'observation et de la distance entre les deux caméras.

— Générer des séries temporelles de vélocité métrique régularisées. Ceci comprend le développement d'une méthode de maximisation de la cohérence des estimations de vélocité glaciaire bidate s'appuyant sur l'optimisation d'une fonctionnelle tenant compte de descripteurs spatio-temporels, d'une contrainte de fermeture temporelle des champs de déplacement intra time-lapse et enfin du sens et de la stationnarité de l'écoulement glaciaire projeté sur le plan image.

La thèse proposée diffère des travaux existants essentiellement par son degré d'automatisation, réalisé grâce à la construction de descripteurs et d'heuristiques décrivant la qualité

des données et des mesures, dans l'utilisation de méthodes de suivi pixelliques robustes, dans la méthode de calcul de cartes d'incertitudes et enfin dans la prise en compte de mesures de vélocité sur plusieurs échelles temporelles.

Première partie :

Étude de l'existant pour le traitement de time-lapses stéréoscopiques de milieux naturels

Dans cette première partie, une étude des outils mathématiques fondamentaux ainsi que de l'état de l'art pour la conception d'une chaîne de traitements de time-lapses stéréoscopiques pour l'observation de milieux naturels est réalisée. Nous introduisons les travaux étudiés d'une manière incrémentale selon le stade de traitement : Tout d'abord les concepts et modèles qui permettent de décrire et d'analyser de telles données d'une manière générale (chapitre 1). Nous étudions ensuite les outils mathématiques et algorithmes spécifiques pour les traitements monodate et bidates monoscopiques (chapitre 2), monodates et bidates stéréoscopiques (chapitre 3) et enfin les traitements en séries temporelles et multi-temporelles (chapitre 4).

Chapitre 1

Concepts et modèles mathématiques pour le traitement de time-lapses stéréoscopiques

Sommaire du chapitre

1.1 Introduction **9**

1.2 Reconstruction par inférence de mesures **10**

1.2.1 Inversion de modèles paramétriques 10

1.2.1.1 Méthodes statistiques et probabilistes 11

1.2.1.2 La méthode des moindres carrés 11

1.2.1.3 Méthodes d'optimisation itératives 12

1.2.2 Régularisation 13

1.2.3 Pondération 14

1.2.4 Prises en comptes d'observations aberrantes 15

1.3 Transformations géométriques et invariants **15**

1.4 Évaluation des incertitudes **17**

1.4.1 Définition d'incertitude 17

1.4.2 Propagation analytique des incertitudes 17

1.4.3 Propagation numérique des incertitudes 18

1.4.4 Dilution géométrique de la précision (GDOP) 19

1.5 Conclusion **20**

1.1 Introduction

Dans ce chapitre, nous allons introduire des notions et des modèles mathématiques qui reviendront de façon récurrente dans l'étude des outils permettant le traitement de time-lapses stéréoscopiques de milieux naturels. Nous aborderons tout d'abord la notion de reconstruction d'un modèle à partir d'observations numériques bruitées (section 1.2), notion très présente dans cette thèse, notamment pour le co-recalage des données time-lapses, la reconstruction 3D stéréoscopique ou encore le traitement de séries temporelles.

Nous verrons ensuite quelques modèles utilisés en vision par ordinateur (section 1.3) et enfin les méthodes qui vont nous permettre de quantifier la précision des résultats issus de la chaîne de traitements proposée (section 1.4).

1.2 Reconstruction par inférence de mesures

La notion de reconstruction occupe une place importante en traitement de l'information et plus particulièrement pour le traitement de time-lapses stéréoscopiques. En effet, les grandeurs x recherchées dans un signal peuvent ne pas être accessibles directement et on cherchera à les inférer à partir d'un ensemble d'observations y accessibles d'un modèle liant les deux grandeurs via des contraintes.

Un ensemble cohérent de telles contraintes est généralement lié à un modèle physique ou à une suite de mesures où chaque contrainte résulte d'un capteur. Cette définition rejoint la notion d'inversion de modèles paramétriques appelée les "problèmes inverses". Typiquement, la reconstruction 3D à partir d'observations images et de contraintes de projection [5, Chapitre 6], la déconvolution de signaux [6], la reconstruction de cartes de déplacements en télédétection [7][8], les techniques tomographiques ou encore certaines techniques géodésiques [9] impliquent des problèmes inverses.

A noter que lorsque la reconstruction est faite pour retrouver certains paramètres du modèle, on parlera plutôt de régression ou d'apprentissage mais que les approches sont similaires.

1.2.1 Inversion de modèles paramétriques

Soit A une application de mesures entachées d'un bruit additif modélisée ϵ comme suit :

$$\begin{aligned} A : \mathbb{M} &\longrightarrow \mathbb{D} \\ x &\longmapsto \tilde{y} = y + \epsilon. \end{aligned} \tag{1.1}$$

L'opérateur A à inverser peut être linéaire ou non-linéaire et bien ou mal conditionnée. Cette notion de conditionnement associé à un problème est une heuristique de la difficulté numérique de la résolution de celui-ci. D'un point de vue probabiliste, cela correspond à l'amplification de la variance, ou incertitude, par l'opérateur A.

Les 3 grandes familles les plus courantes d'outils qui permettent l'étude et/ou la résolution d'un problème inverse sont :

— les méthodes statistiques et probabilistes,
— la méthode des moindres carrés,
— les méthodes d'optimisation itératives.

1.2.1.1 Méthodes statistiques et probabilistes

Étant donné la nature stochastique du bruit de mesure , les problèmes inverses peuvent être vus comme des problèmes d'inférence statistique où on cherchera à retrouver une estimation $\hat{x}$ dont la densité de probabilité, ou uniquement le mode, prédit le mieux la densité de probabilité des observations y. On maximisera alors l'inférence Bayésienne $P(x;y) = P(x|y) \propto P(y|x)P(x)$ entre les variables aléatoires multi-variées x et y ou plus simplement la fonction de vraisemblance $P(x;y) = P(y|x)$ si $P(x)$ est supposée uniforme :

$$\hat{x} = \underset{x}{argmax} P(x;y). \tag{1.2}$$

Les contraintes du modèle encodé dans A et des connaissances a-priori sur $\hat{x}$ sont respectivement utilisées afin de paramétriser $P(y|x)$ et $P(x)$. L'inférence statistique $P(x;y)$ produit une probabilité qui s'interprète comme le degré d'adéquation d'une entrée x, hypothétique, aux observations $\tilde{y}$ qu'elle engendre et sa maximisation peut donc être assimilée à la minimisation d'une fonction de coût qui traduit l'erreur d'adéquation entre $A(x) = \tilde{y}$ et y.

Dans ce sens, la deuxième et la troisième famille de méthodes permettent de minimiser une fonction de coût ϱ définie directement comme suit :

$$\hat{x} = \underset{x}{argmin} \quad \varrho(A(x) - y), \tag{1.3}$$

où ϱ est une fonction de coût convenablement choisie. Plus particulièrement, la maximisation du log de la fonction de vraisemblance $P(x;y) = P(y|x)$ et la minimisation d'une fonction de coût définie par $\varrho(.) = ||.||_2^2$ (somme des erreurs au carré) sont théoriquement équivalents sous une hypothèse de bruit ϵ Gaussien. Cette optimisation est connue sous le nom de "moindres carrés".

1.2.1.2 La méthode des moindres carrés

La résolution des problèmes inverses par minimisation du carré du coût ϱ défini comme la norme $L = 2$ est l'approche la plus standard pour les problèmes sur-déterminés. Dans le cas d'une notation matricielle du problème (opérateur A linéaire ou notation homogène), l'équation 1.3 devient :

$$\hat{x}_{LS} = \underset{x}{argmin} \quad ||Ax - y||_2^2 = ||\epsilon||_2^2, \tag{1.4}$$

ce qui permet de retrouver directement $\hat{x}_{LS}$ via la pseudo-inverse de A définie par $A^\dagger = (A^T A)^{-1} A^T$ dans le cas où l'opérateur linéaire est sur-détérminé et $A^T(AA^T)^{-1}$ dans le cas sous-détérminé. Avec cette configuration, une estimation du degré de condtionnement

est retrouvée via le coefficient k défini par :

$$k = \frac{\sigma_{\max}}{\sigma_{\min}}, \tag{1.5}$$

où $\sigma_{\max}$ et $\sigma_{\min}$ sont les valeurs singulières maximales et minimales de la matrice décrivant l'opération linéaire induite par A. Dans la pratique, ce calcul peut s'avérer lourd, surtout si la matrice A contient des lignes ou colonnes pleines qui induise une matrice $A^T A$ non-éparse, et la solution $\hat{x}_{LS}$ peut être retrouvée d'une façon plus optimale via la décomposition orthogonale de A tel la décomposition en valeurs singulières (SVD) ou la décomposition QR. De plus, une SVD tronquée permet la prise en compte uniquement des valeurs singulières supérieures à un seuil pour améliorer le conditionnement du problème (composantes principales).

Autrement, si le problème ne possède pas de représentation linéaire, la résolution consiste à approcher le modèle par approximations linéaires locales et à affiner $\hat{x}_{LS}$ à partir d'une estimation initiale $\hat{x}_0$ et par itérations successives [10]. Par exemple, la méthode dite de Newton permet d'exprimer l'équation 1.4 localement comme suit :

$$\Delta\hat{x}^i_{LS} = \underset{x}{argmin} \quad J^T J \Delta x + J^T \Delta\epsilon_0, \tag{1.6}$$

où $\Delta\hat{x}^i_{LS}$ est la solution linéaire à l'itération i qui permet d'incrémenter $\hat{x}_{LS}$ et J la matrice Jacobienne de ϵ par rapport x. Concernant les problèmes sous-déterminés, une infinité de solutions $\hat{x}_{LS}$ avec un coût résiduel $||\epsilon||_2^2$ nul existent. Il convient alors de contraindre l'espace parcouru ou de définir une fonction de coût plus adéquate afin de restreindre le champ des solutions possibles. Ces deux approches sont possible via la famille des méthodes dites d'optimisation avec ou sans contraintes.

1.2.1.3 Méthodes d'optimisation itératives

Comme pour les méthodes directes, pour des observations y données, $\hat{x}$ est la solution où la fonction de coût $\varrho(\epsilon(x))$ atteint un minimum. La nécessité de procéder par itérations numériques (on parle ici principalement de l'utilisation de solver) vient principalement de l'éventuelle non-convexité de l'espace de recherche associé à un opérateur non-linéaire. Utiliser les méthodes d'optimisation itératives facilite en outre :

— l'injection une estimation initiale x_0 de $\hat{x}$ choisie judicieusement (proche du minimum global),

— l'utilisation d'une fonction de coût adaptée aux hypothèses sur le bruit de mesure, tel que $\varrho(\epsilon) = ||\epsilon||_1$ si le bruit est supposée suivre une distribution de Laplace (présence de valeurs aberrantes par exemple), etc.

— le rajout des contraintes linéaires ou non-linéaires, d'égalités ou d'inégalités sur $\hat{x}$ afin de restreindre le champ des solutions,

— le rajout d'un terme de régularisation en plus du terme d'adéquation aux données,
— l'optimisation de la résolution, en terme de mémoire de travail, en présence de modèles éparses.

Ces propriétés sont également intéressantes dans les cas linéaire et convexe et plus particulièrement en présence d'une quantité importante d'observations, où une SVD peut user voire dépasser la mémoire de travail, lorsque le problème est mal conditionné ou encore lorsque la matrice A est éparse [11]. Il existe une multitude de méthodes numériques de résolution itérative de problèmes inverses qui tournent autour des méthodes suivantes :

— la méthode des gradients conjugués pour les problèmes linéaires,
— les méthodes de descente de gradient et ses variantes où l'espace de recherche est parcouru en suivant les pentes les plus raides de ϱ jusqu'à un minimum,
— la méthode de Newton et ses variantes (par exemple Gauss-Newton ou quasi-Newton) basée sur des approximations locales de l'opérateur à inverser qui permettent le calcul d'une direction de descente,
— la méthode de Levenberg-Marquadt [12][13] de résolution de problèmes de moindres carrés non-linéaires largement utilisée en vision 3D (chapitre 3) et qui se situe entre les méthodes de Newton et celles de descente de gradient,
— les algorithmes du type Simplexe, approche géométrique dont l'avantage est la non-nécessité de calcul de dérivées,
— les méthodes de parcours aléatoire tel que la méthode RANdom SAmple Consensus,
— les algorithmes, pas très optimaux, du type "Brute Force" où la solution est trouvée par recherche exhaustive,
— etc.

Dans ce type d'approches, les dérivées sont soit calculées à partir d'une forme analytique soit par des algorithmes numériques de différences finies. Quoique beaucoup plus simple, les méthodes des différences finies peuvent potentiellement bloquer la progression près d'un minimum en raison d'une précision insuffisante. Quand la fonction de coût n'est pas quadratique ou une valeur absolue, elle peut par exemple être une "Huber loss", une "Cauchy loss", ou n'importe quelle fonction de coût adaptée à la nature de ϵ.

1.2.2 Régularisation

Régulariser des données revient à les rendre régulières et conformes à une connaissance a-priori de ceux-ci. Plus particulièrement, pour les techniques d'apprentissage et les problèmes inverses, la régularisation est le processus qui consiste à ajouter des informations sur la solution afin de résoudre un problème mal posé ou d'éviter un sur-ajustement à des observations y bruitées. Par exemple, la méthode de Levenberg-Marquadt [12][13] inclut un tel terme de régularisation où la minimisation des moindres carrés de l'erreur

quadratique prend la forme augmentée suivante :

$$\hat{x} = \underset{x}{argmin} \quad ||A(x) - y||_2^2 + ||\Gamma x||_L^2. \tag{1.7}$$

Cette formulation est appelée régularisation de Tikhonov avec Γ un multiple de la matrice identité tel que : $\Gamma = \lambda I$. $\lambda > 0$ est le facteur d'amortissement qui permet d'équilibrer entre un premier terme d'attache aux données et un deuxième terme de lissage et ainsi d'éviter des solutions ayant une norme L trop importante résultat d'un mauvais conditionnement. Dans l'algorithme Levenberg-Marquadt, la norme utilisée est la norme $L = 2$ et λ est ajusté à chaque nouvelle itération.

Comme pour la maximisation d'une l'inférience Bayésienne introduite plus-haut, cette formulation régularisée prend en compte une hypothèse de départ sur la distribution de $\hat{x}$ dont la plage de variation dépend du choix de λ. Dans ce sens, une paramètrisation $\Gamma = \lambda W \neq \lambda I$ peut être vue comme la pondération de la distribution de $\hat{x}$.

Il est cependant parfois assumé qu'une restriction de la norme pondérée de x n'est pas suffisante pour amortir les oscillations indésirables dans la solution [14] et un terme de pénalisation sur la norme des *variations* de $\hat{x}$ est rajouté en plus ou à la place. Ce terme de restriction est appelé la variation totale qui est définie par :

$$v_{tot}(x) = ||x_{i+1} - x_i||_L = ||\nabla x||_L, \tag{1.8}$$

où ∇ est l'opérateur différentiel vectoriel. Restreindre la variation totale de $\hat{x}$ permet en effet de favoriser une solution où les variations d'un point à l'autre sont douces (par exemple avec une tendance à pente douce) par rapport à une solution sujette à des variations brusques.

1.2.3 Pondération

Bien que $\hat{x}$ soit calculée avec une prise en compte de la nature du bruit et éventuellement certaines contraintes sur les valeurs ou la distribution de valeurs de la solution désirée, celle-ci omet la nature hétéroscédastique des observations $y_{\in \mathbb{M}}$. Pour modéliser un bruit de mesure non-uniforme, l'équation 1.3 devient :

$$\hat{x} = \underset{x}{argmin} \quad \varrho(WA(x) - Wy). \tag{1.9}$$

Avec la formulation de l'équation 1.9, un coût défini comme les moindres carrés devient une distance de Mahalanobis $||.||_W^2$ qui a pour effet de standardiser les observations si la matrice W est diagonale et de les décorréler si W contient en plus des termes non-diagonaux. Ainsi, W peut être la matrice de précision ou encore avoir des éléments diagonaux constitués de scores normalisés attribués à chaque observation.

1.2.4 Prises en comptes d'observations aberrantes

L'algorithme RANSAC ou RANdom SAmple Consensus est une amélioration possible de tout problème d'inversion ou d'apprentissage. Cette méthode est préconisée en présence de valeurs aberrantes dans les observations. Afin de retrouver des paramètres pertinents avec la plupart de ces observations, un sous-ensemble de y est d'abord sélectionné afin d'estimer un premier modèle et le reste des observations sont classifiées en tant que inliers (pertinents) ou outliers (non-pertinents) par rapport à ce modèle selon une fonction de coût.

Le modèle est retenu si il induit un nombre suffisant d'inliers et est ré-estimé à partir d'un nouvel ensemble d'observations sinon. Ce schéma est répété sur plusieurs itérations et le modèle retenu sera celui dont l'erreur d'adéquation aux données pertinentes et le nombre d'outliers sont faibles. Un inconvénient d'une estimation via RANSAC est qu'elle suppose de fixer un ensemble de seuils et conditions d'arrêt spécifiques au problème traité et aux données.

1.3 Transformations géométriques et invariants

La notion d'invariance est très présente en traitement d'image 2D et 3D. À l'origine, la notion d'invariant mathématique est relative à un changement de repère en géométrie. En traitement d'image cela signifie que les attributs d'une image 2D ou d'un modèle 3D conservent leur identité après avoir subi une transformation géométrique. Cette identité est généralement liée à des descripteurs locaux ou globaux. Par exemple, un algorithme de traitements 2D est dit invariant à un changement d'échelle s'il permet de décrirer l'image de la même manière avant et après une mise à l'échelle, par exemple via sous-échantillonnage ou par modification de la distance d'acquisition.

Les transformations géométriques perceptibles sur les données 2D et 3D peuvent être non-rigides, telles qu'un écoulement de fluide ou une déformation élastique, ou rigides, tel que le déplacement d'un objet inextensible parallèlement à la caméra. Une modélisation paramétrique rigide par morceaux est aussi possible. Une taxonomie des modèles paramétriques est donnée dans le tableau 1.1 selon le nombre de degrés de liberté (DdL) et de l'invariance engendrée.

Transformation	Invariance	DdL en 2D	DdL en 3D	Rigidité
Translation	distances, angles	2	3	✓
Rotation	distances, angles	1	3	✓
Affine	parallélisme	6	12	✗
Homographie ou application projective	birapports	8	15	✗

Table 1.1 – Modèles paramétriques en vision dans les espaces 2D et 3D.

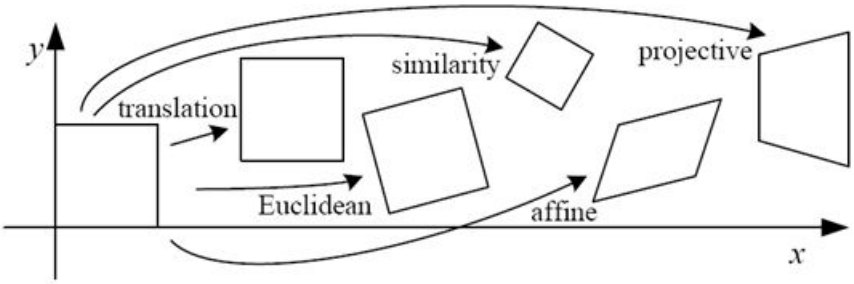


FIGURE 1.1 – Modèles paramétriques en vision dans l'espace 2D.

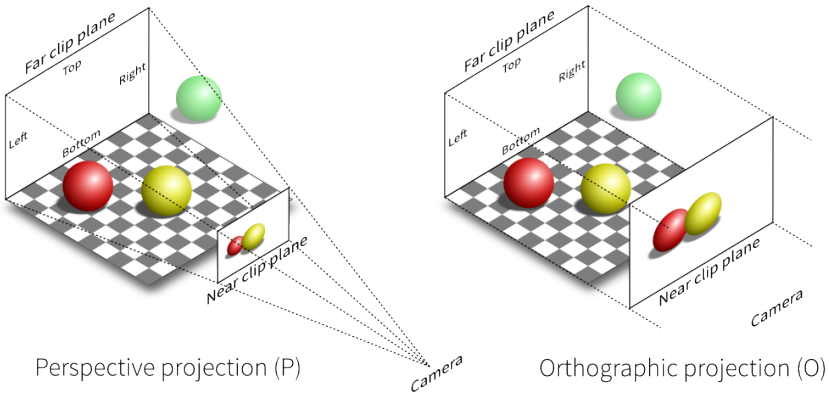


FIGURE 1.2 – Modèles de projection 2D/3D. Gauche : Projection perspective. Droite : Projection orthographique.

Une translation couplée avec une rotation modélise alors une transformation Euclidienne à 3 DdL (une isométrie) et une transformation affine inclut en plus un paramètre de cisaillement homogène qui rend la déformation non-rigide (voir figure 1.1). Une déformation géométrique peut aussi modéliser la projection d'un espace 3D dans un espace 2D (acquisition d'une photo) ou son inverse (reconstruction 3D). Les différentes projections possibles entre les espaces $\mathbb{R}^3$ et $\mathbb{R}^2$ sont données dans le tableau 1.2 et illustrée par la figure 1.2).

Transformation	Invariance	DdL	Type de reconstruction
Orthographique	ratios de distances, parallélisme, perpendicularisme	5	métrique ou Euclidienne
Affine	parallélisme	8	affine
Perspective	incidence	11	projective

TABLE 1.2 – Modèles paramétriques de projection de l'espace 3D dans l'espace 2D.

Certaines des transformations dans les tableaux 1.1 et 1.2 sont non-linéaires. Nous verrons cependant dans le reste de cette première partie qu'il est possible de modéliser les transformations 2D par un système matriciel linéaire de taille 3×3, les transformations 3D par des systèmes de taille 4×4 et les transformations $2D/3D$ par une application linéaire de taille 3×4 dites matrices de projection. Cette linéarisation, dite par notation

homogène, permet en outre de considérer plusieurs de ces transformations géométriques avec un modèle unique issu du produit matriciel des applications impliquées.

1.4 Évaluation des incertitudes

1.4.1 Définition d'incertitude

La définition formelle du terme "incertitude de mesure" par le VIM (Vocabulaire International des termes fondamentaux et généraux de métrologie) [15] est la suivante : "Paramètre, associé au résultat d'un mesurage, qui caractérise la dispersion des valeurs qui pourraient raisonnablement être attribuées au mesurande". Dans ce sens, les erreurs-type et les incertitudes peuvent être estimées par :

— comparaison à des valeurs "vérité terrain" connues sur un ensemble de données test, qui reste la méthode la plus directe pour estimer l'erreur probable,
— à partir de la distribution statistique de séries de mesurages : écart-type expérimental sous hypothèses de répétabilité (même expérience) et de reproductibilité (changement des conditions),
— avec un raisonnement sur les intervalles de tolérance et les sources d'erreurs à partir, par exemple, de jugements d'experts ou d'informations de bibliographie,
— par propagation analytique ou numérique de l'incertitude en entré,
— avec un raisonnement sur la dilution de la précision à travers le processus de mesurage.

Tandis qu'une mesure directe de l'erreur sur certaines observations permet d'identifier l'erreur globale probable sur le reste des données, un raisonnement sur la loi de probabilité de la mesurande permet une identification précise du biais, ou erreur systématique, et de la dispersion statistique, ou erreur aléatoire. Cette erreur aléatoire est au moins égale à la précision. Dans ce qui suit nous allons détailler les approches que nous estimons potentiellement utiles pour évaluer l'évolution de l'incertitude à travers la chaîne de traitement qu'on proposera.

1.4.2 Propagation analytique des incertitudes

Considérons le problème direct qui prend comme entrée x, comme sortie y et comme modèle A. Une estimation analytique de la propagation de la dispersion de x vers y est possible via une analyse de sensibilité de A qui, à l'instar du coefficient de conditionnement k, quantifie l'influence d'une entrée sur la sortie d'un opérateur. Celui-ci est obtenu via un développement de Taylor au premier ordre et est défini comme suit :

$$\Sigma_y = J_x \Sigma_x {J^T}_x, \tag{1.10}$$

où J_x est la matrice Jacobienne qui contient les coefficients de sensibilité de l'opérateur A par rapport à x. Lorsque les contraintes engendrées par les lignes de A sont découplées, le cas idéal, J_x est diagonales et il n'y a généralement pas une grande propagation de l'incertitude via la mesure. Pour les opérateurs non-linéaires, l'équation 1.10 reste une approximation locale au premier ordre. L'incertitude peut ensuite être estimée par les écart-types contenus dans Σ_y qui contiendront, ou non, un terme de corrélation. Ces méthodes de type probabiliste ont notamment été promue par le le Guide pour l'Expression de l'Incertitude de mesure (Guide to the expression of Uncertainty in Measurement (GUM)), devenu document de référence en théorie des incertitudes et nous référerons donc dans le reste du manuscrit à ce type de méthode par méthodes GUM. Dans le GUM, il est aussi préconisé d'estimer une incertitude expérimentale (dite de type A) qui peut comprendre une incertitude de répétabilité estimée par l'écart-type observé en effectuant plusieurs fois le mesurage sur les mêmes données. Dans le cas Gaussien, cet écart-type explique 68.2% des réalisations présentes dans les données mesurées y et il convient de multiplier le résultat par un coefficient c pour obtenir une incertitude dite élargie. Par exemple, pour $c = 2$, l'incertitude est alors exprimée avec un taux de confiance à $95,4\%$. Les limites des méthodes de propagation analytique de la variance sont :

— l'hypothèse d'un processus Gaussien qui peut être insuffisante pour certains problèmes,
— le choix du facteur c est guidé par le théorème central limite qui, encore une fois, peut être insuffisant si la mesurande ne suit pas une loi normale,
— estimation issues d'un approximation au premier ordre et donc biaisée si l'opérateur est fortement non-linéaire.

1.4.3 Propagation numérique des incertitudes

L'évaluation de l'incertitude liée à un ou à une série d'opérateurs de mesure peut aussi être estimée par simulation numérique du processus de mesurage. Il s'agit des algorithmes du type Monte-Carlo. Ces méthodes numériques sont basées sur des simulations logicielles qui permettent d'étudier la propagation des distributions de probabilités et nécessite d'injecter un grand nombre d'échantillons de valeurs pour chacune des grandeurs d'entrée. Selon la loi des grands nombres, moyenner les sorties des simulations permet d'identifier une estimation de l'espérance mathématique et de l'écart-type qui définissent la distribution empirique de y.

Cette méthode est assez coûteuse et certaines études proposent des méthodes pour réduire le nombre d'échantillons ou pour compléter la simulation avec une analyse de sensibilités. L'avantage de cette méthode est qu'elle est applicable sur l'ensemble des briques de la chaîne de traitements d'un coup mais celle-ci présentent également quelques limites :

— le choix du générateur de nombres aléatoires et du nombre d'échantillons peut faire varier les résultats,
— une analyse de sensibilités a-posteriori,
— certaines lois ne sont pas aisément simulables ou déductibles,
— l'approche est non-analytique,
— enfin et surtout, un coût informatique très élevé surtout pour des chaînes complexes nécessitant plusieurs propagations comme dans notre cas.

1.4.4 Dilution géométrique de la précision (GDOP)

Une famille de méthodes similaires est basée sur un calcul géométrique d'air ou de volume à l'intérieur duquel y peut se trouver [16] (méthodes GDOP pour *Geometric Dilution Of Precision*). Ce principe est notamment utilisé par le système de géolocalisation américain Global Positioning System (GPS) afin d'estimer l'incertitude sur une mesure de position issue de la trilatération à partir de 2 ou plusieurs satellites. Cette estimation permet en outre d'optimiser le nombre d'observations à trianguler. Plus particulièrement, dans le cas de triangulation de 2 observations x_1 et x_2, comme dans le cas de mesures stéréoscopiques, les mesures ponctuelles accompagnées de leurs incertitudes sont considérées comme des surfaces et les rayons de triangulation comme des cônes tridimensionnels ou aplatis. L'aire d'intersection de ces cônes est alors une estimation de l'incertitude de la mesure triangulée (figure 1.3).

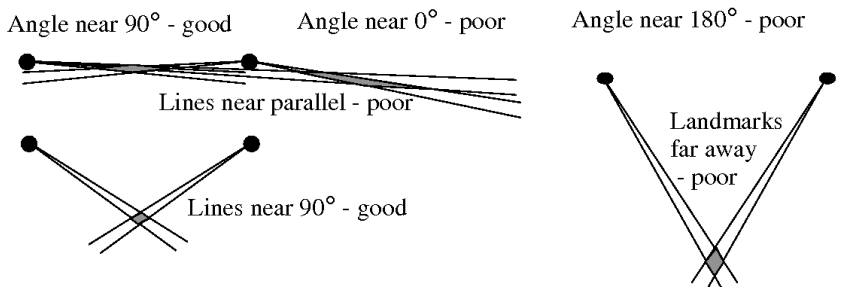


FIGURE 1.3 – L'aire du quadrilatère en gris est une estimation de l'incertitude de la masure triangulée. Figure tirée de [17]. Plus la surface grise est grande plus le point triangulé est imprécis (ou incertain).

Les sources d'erreurs et les techniques d'estimation et/ou de minimisation de l'incertitude pour les traitements de time-lapses stéréoscopiques seront détaillées dans les chapitres 2 et 3 puis dans le chapitre 7 où nous proposons un algorithme basé sur une approche géométrique pour l'estimation de l'incertitude 3D.

1.5 Conclusion

Dans ce premier chapitre, nous avons introduit différentes notions mathématiques qui permettent de poser une chaîne de traitements de time-lapses stéréoscopiques et d'en quantifier l'incertitude. Par la suite, l'information utile sera reconstruite, et sa dispersion propagée, d'une manière incrémentale à partir d'observations brutes ou pré-traitées, et ce, par l'ajustement de modèles théoriques et de contraintes sur les observations. Cette notion de reconstruction à partir d'observations numériques sera présente d'une façon plus spécifique dans le reste des chapitres de l'état de l'art pour les traitements monoscopiques, stéréoscopiques puis temporels et les modèles paramétriques 2D, 3D et 2D/3D présentés serviront à poser certains problèmes à inverser.

Chapitre 2

Méthodes monodates et bidates monoscopiques

Sommaire du chapitre

2.1 Introduction **21**
2.2 Caractérisation des données images de milieux naturels **22**
2.3 Sélection et prè-traitements **23**
2.3.1 Méthodes basées sur un modèle physique 23
2.3.2 Méthodes basées texture 25
2.3.2.1 Méthodes statistiques 25
2.3.2.2 Méthodes fréquentielles 27
2.3.3 Méthodes basées apprentissage 29
2.4 Estimation du déplacement des pixels **29**
2.4.1 Approches denses 29
2.4.1.1 Approches basées sur des mesures de similarité 30
2.4.1.2 Approches variationelles 33
2.4.2 Approches éparses 35
2.4.2.1 Détection et description de points d'intérêts 35
2.4.2.2 La description de points d'intérêt 38
2.4.2.3 Mise en correspondance 39
2.5 Alignement paramétrique **39**
2.6 Conclusion **41**

2.1 Introduction

Ce chapitre est dédié à l'étude des méthodes existantes pour le traitement d'ensemble et de couples d'images acquises avec un même appareil-photo sur deux dates différentes. Pour ce faire, nous commencerons tout d'abord par caractériser les données #N 2D (ensemble d'images non ordonnées) et 2D+t (deux images d'une même scène prises avec la même perspective sur deux dates différentes) (voir tableau 1) qu'on peut observer en milieux naturels et nous introduirons ensuite les méthodes adaptées pour chacune des étapes principales d'une chaîne de traitements monodates et bidates monoscopiques, à savoir :

— la sélection d'un ensemble de n images a propager (données #n 2D) à partir d'un ensemble de N images acquises (données #N 2D) et leur pré-traitement,
— l'estimation du déplacement des pixels entre deux images (données 2D+t),
— et/ou l'estimation d'un modèle paramétriques de déformation entre deux images (données 2D+t).

Nous nous plaçons ainsi d'emblée dans un contexte de données acquises en milieux naturels afin de sélectionner les outils adaptés à déployer de façon générique pour générer de données de suivi géophysique.

2.2 Caractérisation des données images de milieux naturels

Tout d'abord, les images de milieux naturels, étant des données de détection proximale de modélisation à grande échelle spatiale, seront généralement de haute définition (en nombre de pixels) et résolution (en taille de pixel au sol) afin de permettre des calculs géodésiques précis.

Données monodates monoscopiques #N 2D et #n 2D Les données #N 2D désignent l'ensemble des N images acquises d'un time-lapses considérées sans ordre particulier. Les défis majeurs du d'images #N 2D de milieux naturels par rapport à des images en milieux intérieurs sont notamment les aléas atmosphériques et météorologiques, on cite :

— la diffusion des ondes lumineuses dans les aérosols et les nuages présents entre le système d'observation et la scène qui réduit la visibilité, altère l'image et peut occulter la scène. Exemple : brume, brouillard, pluie, pollution, etc.
— l'emplacement du soleil qui engendre une différence dans l'aspect des images. Exemple : les ombres portées causant une différence des les gradients, l'interaction des ondes lumineuses avec le sol selon des angles d'incidence différents causant une différence dans l'intensité des pixels,
— la présence de neige qui entraîne un manque de texture,
— le changement de saison et d'autres épisodes climatiques plus ou moins courts qui entraînenst aussi une décorrélation photométrique.

Nous verrons dans le préambule de la deuxième partie de ce manuscrit les aléas spécifiques à l'observation de glaciers alpins. Afin de propager le minimum d'erreur au reste des briques de traitements, seulement n images, avec $n < N$, serviront au calcul des champs de déplacement pixellique. Celles-ci constituent alors les données # n 2D et devront disposer d'une bonne visibilité au niveau de l'objet d'intérêt.

Données bidates monoscopiques 2D+t Ces données désignent un ensemble # 2 2D de deux images, parmi les données sélectionnées, prises à des temps différents avec la même perspective (appareil-photo fixe) où la baseline temporelle (temps écoulé entre deux acqui-

sitions) ainsi que l'ordre dans les acquisitions sont importants. Ces données permettent le calcul du déplacement ou velocité de la scène en pixels. Un système d'observation monoculaire en milieu naturel étant souvent soumis à des changements de températures brutaux, son instrumentation interne (système de lentilles) et externe (contraction et dilatation thermiques du support de l'installation) sera probablement instable et causera un changement de perspective, souvent imperceptible à l'oeil nu, entre les images du time-lapse. Cette différence géométrique doit être compensée par un co-recalage global des zones fixes des images du time-lapse entre elles afin de ne pas biaiser les mesures de suivi.

2.3 Sélection et prè-traitements

On s'intéresse dans cette section aux outils et modèles permettant de caractériser la visibilité sur une image. Nous répartissons alors les méthodes envisagées en deux catégories :

— les méthodes basées sur un modèle physique de la visibilité,

— les méthodes basées sur la description de la texture.

Il existe en effet, comme nous allons le voir, des modèles physiques permettant de quantifier la diffusion des ondes lumineuses capturées par un appareil-photo qui peuvent être inversés afin de reconstruire une image non dégradée. D'autre part, caractériser la composition spatiale ou spectrale d'un ensemble d'images via un descripteur peut servir à détecter les images statistiquement aberrantes sur un time-lapse et, le cas échéant, les écarter ou les corriger.

2.3.1 Méthodes basées sur un modèle physique

Narashiman & al. [18][19][20] publient un des premiers travaux dans la littérature en vision par ordinateur qui s'intéresse de plus près au problème de la vision optique en milieu extérieur et à la difficulté d'appliquer les méthodes classiques directement sur des images affectées par les conditions météorologiques. Dans ces travaux, une liste des phénomènes qui viendraient réduire la visibilité en milieux naturels est fournie et un modèle physique de la diffusion et de l'atténuation des ondes lumineuses dans la brume est estimé. Ce modèle physique montre que dans des conditions d'acquisition brumeuses le contraste ainsi que l'intensité de la couleur de l'image décroisent exponentiellement avec la distance entre le système d'observation et l'objet observé (Figure 2.1).

Ce résultat est maintenant généralisé aux milieux similaires tels que les milieux aquatiques [21] ou en présence de pluie [22] et on parlera de "*dehazing*" ou de dé-brumage pour décrire n'importe quelle méthode qui vise à compenser des phénomènes atmosphériques qui causent une diffusion homogène et isotrope des ondes avant qu'elles n'atteignent la

FIGURE 2.1 – Exemple d'images acquises dans des conditions brumeuses. L'altération de l'image par la diffusion des ondes lumineuses dans la brume est fonction de la profondeur. La brume altère la couleur de l'image avec une teinte uniforme (ici grise-claire).

caméra. Une image est alors modélisée par une somme pondérée comme suit :

$$I(\mathbf{x}) = J(\mathbf{x})t(\mathbf{x}) + A_\infty(1 - t(\mathbf{x})), \tag{2.1}$$

où $J(\mathbf{x})$ est le paramètre à reconstruire et modélise la radiance (ou luminance énergétique) de la scène, A_∞ est la lumière ambiante et $t(\mathbf{x})$ est le facteur de transmission défini en fonction de la profondeur Z comme suit :

$$t(\mathbf{x}) = \exp^{-\beta Z(\mathbf{x})}. \tag{2.2}$$

Dans l'équation 2.1, la quantité $J(\mathbf{x})t(\mathbf{x})$ modélise donc le fait que la lumière qui émane de l'objet est atténuée par une diffusion dans l'atmosphère agrégée le long de la ligne de visée (loi de Beer-Lambert) et $A_\infty(1 - t(\mathbf{x}))$ modélise l'altération de la teinte de l'image $I(\mathbf{x})$ par la lumière ambiante A_∞ créée par la brume. Depuis, de nombreuses approches basées sur ce modèle physique ont été proposées [23][24][25][26][27][28]. Ces approches se différencient principalement les unes des autres par les hypothèses et les contraintes qui servent à estimer le paramètre de transmission $t(\mathbf{x})$, la lumière ambiante A_∞ et les techniques de reconstruction de $I(\mathbf{x})$ à partir de l'équation 2.1. Une reconstruction robuste comprend l'utilisation de filtres adaptés pour lisser le passage d'un niveau de profondeur à un autre en post-traitement. Ceci se fait par des techniques très similaires au soft-matting [29] ou segmentation avant\arrière-plan. De façon générale, la lumière ambiante A_∞ est calculée à partir de l'intensité des pixels $\mathbf{x}$ les plus corrélés avec la dégradation radiométrique observée, et donc les plus lointains. He & al. [25] proposent la technique qui s'est avérée la plus pertinente notamment en introduisant la notion de "dark-channel" pour calculer A_∞ et $t(\mathbf{x})$. La notion de "dark-channel" repose sur l'hypothèse que les pixels les plus corrélés avec la dégradation observée sont les pixels où aucun canal RVB n'est sombre. La rationalisation derrière cette hypothèse est qu'un objet de couleur aura forcément un canal sombre et qu'au contraire les objets à l'horizon auront des intensités lumineuses relativement similaires sur tous les canaux car fortement corrélées avec $A_\infty(\mathbf{x})$

qui est une teinte uniforme. Nous renvoyons le lecteur vers les travaux de Fattal & al. [28] pour plus de détails sur l'estimation de l'altération causée par des conditions brumeuses et la reconstruction par "dehazing".

Bien que les méthodes basées sur le modèle physique de la diffusion des ondes aient été jugées efficaces pour pré-traiter des données de mileux naturels, nous estimons qu'elles souffrent de certaines limitations pour des traitements # N 2D de façon générique, à savoir :

— l'hypothèse "dark-channel" pour le calcul de $A(\mathbf{x})$ n'est pas valide en présence d'une importante quantité de pixels blancs (Figure 2.2a),
— le modèle du facteur de transmission $t(\mathbf{x})$ n'est pas valide à grande échelle spatiale car la température est susceptible de varier en fonction de l'altitude causant de la brume à certains niveaux d'altitude et pas d'autres (Figure 2.2b),
— l'équation 2.1 ne gère pas les scénarios où la scène est partiellemet ou complètement occultée.

(a) (b) (c)

FIGURE 2.2 – Exemples de conditions d'acquisitions ou le modèle physique de l'équation 2.1 n'est pas valide. 2.2a : Hypothèse "dark-channel" non valide. 2.2b : Modèle physique de la formation de la brume non valide. 2.2c : Modèle physique de la formation de la brume non valide - Altération non-homogène.

Il faut donc adapter ces méthodes aux données avant de les envisager.

2.3.2 Méthodes basées texture

Un coefficient décrivant la visibilité d'une image peut être quantifié directement via l'information texturale contenue dans celle-ci. Le raisonnement ici est que le flou causé par la brume, ainsi que l'occultation, altèrent la composition spectrale et radiométrique (par exemple le contraste) de la scène observée.

2.3.2.1 Méthodes statistiques

La méthode la plus simple pour caractériser une texture donnée est d'effectuer des mesures statistiques calculées à partir de sa distribution de niveaux de gris, d'intensités lumi-

neuses RGB ou d'une autre représentation de l'image telle que, par exemple, par encodage dans l'espace colorimétrique TSV (Teinte, Saturation, Valeur). Ces mesures statistiques basées "histogramme" peuvent fournir des estimations de différents ordres caractérisant la texture. De surcroît, ces approches peuvent servir à pré-traiter les images par diverses techniques de transformation de la distribution de l'intensité des pixels, telle l'égalisation d'histogramme [30] ou la correction Gamma [30][31]. De telles égalisations permettent un rehaussement du contraste et des couleurs. Pour la classification, ces mesures statistiques s'avèrent cependant souvent limitées en pratique car insensible à la présence ou non de corrélation spatiale et donc pas assez discriminantes pour une classification. Pour pallier à cette limite, des méthodes statistiques prenant en compte la relation entre différents pixels peuvent être utilisées.
Parme une ces méthodes nous pouvons trouver les matrices de co-occurence proposées par Haralick et al [32]. Ces matrices, qu'on notera $A_{d-\theta}$ sont de dimension $l \times l$ où l est le nombre de niveaux d'intensité lumineuse pris en compte. Chaque cellule (i, j) de la matrice $A_{d-\theta}$ représente alors le nombre de couples de niveaux d'intensité de valeurs (i, j) séparés par une distance d dans la direction θ présents dans l'image. Cette méthode permet ensuite de calculer une dizaine de descripteurs de texture d'ordre plus ou moins élevé uniquement à partir de la matrice $A_{d-\theta}$. Parmi ces descripteurs, on trouve le moment anglaire d'ordre 2, la variance, l'auto-corrélation, l'entropie de Shannon, etc. Pour reprendre les mêmes exemples, le moment anglaire d'ordre 2, calculé en sommant le carré des éléments de $A_{d-\theta}$, est capable de fournir un indice sur l'homogénéité d'une texture et la description de l'auto-corrélation donne un indice sur la dépendance linéaire des pixels les uns par rapport aux autres. Les formules et le reste des descripteurs proposés peuvent être retrouvés dans les travaux originaux [32]. Cette méthode est utilisée principalement en télédétection [33] [34] [35] en outre pour une segmentation inter-image. Cette mesure est cependant limitée car basée sur les intensités lumineuses et donc sensible aux variations radiométriques.

Un descripteur similaire moins sensible aux variations radiométriques est le "motif binaire local" ou "local binary pattern" rendu populaire par le travail de Ojala & al. [36]. Pour chaque pixel de l'image, la relation de supériorité d'un pixel avec ses voisins est encodée dans un descripteur binaire. Cet encodage se fait d'une manière ordonnée sur des voisinages circulaires de différentes tailles (Figure 2.3). Les voisinages considérés sont décrits par leur rayon R, ou distance entre chaque paire de pixels à comparer, et le nombre de pixel p, comme illustré par la figure 2.3. En ne considérant que les scénarios de la première ligne de la Figure 2.3 et en faisant varier la fenêtre d'analyse, ce descripteur est rendu invariant par mise à l'échelle et par rotation.

Les régularités de l'image peuvent aussi être capturées par les gradients spatiaux de celle-ci. A ce titre, un autre exemple important d'approche basée "histogramme" pour la caractérisation de la texture et le HOG [37] ou Histogramme de Gradients Orientés.

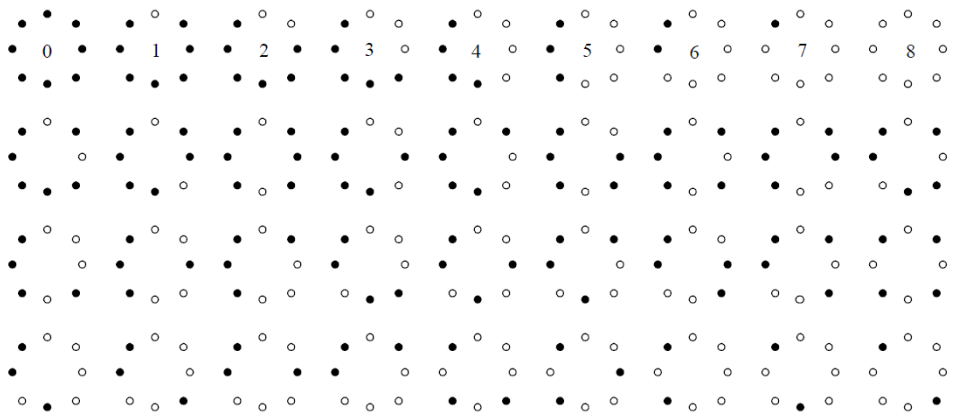


FIGURE 2.3 – Voisinages circulaires de rayons R avec $p = 8$ pour la construction du descripteur "motif binaire local". Blanc : 1. Noir : 0. La description du voisinge 4 est $g = 00001111_2$ et décrit un bord ou arête. La description du voisinage 8 est $g = 11111111_2$ et décrit une tache circulaire sombre. Seul les scénarios de la première ligne sont invariants par rotation. Figure tirée de [36].

Comme son nom l'indique, cet histogramme est construit à partir des gradients spatiaux de l'image. Des cellules rectangulaires ou circulaires sont construites et chaque pixel de la cellule vote pour une classe de l'histogramme, en fonction de l'orientation du gradient à ce point. Le choix est porté sur la classe la plus proche sur 9 classes possibles et le vote du pixel est pondéré par l'intensité du gradient en celui-ci. Les expériences faites par Dalal et Triggs [37] ont montré que la meilleure performance était obtenue pour des blocs rectangulaires contenant 3×3 cellules de 6×6 pixels chacune. La normalisation en amont des ces blocs permet d'éviter les disparités dues aux variations d'illumination.

2.3.2.2 Méthodes fréquentielles

La décomposition spectre uneale d'une image est une technique assez classique qui permet de fournir des éléments directement exploitables pour caractériser les différents schémas redondants (pseudo périodiques) de la texture présente dans une image. Cette décomposition est possible via :

— la décomposition en série de Fourier, où les niveaux d'intensité lumineusee une d'une image sont projetés (par isométrie) sur une base orthonormée Hilbertienn unee de polynômes trigonométriques [38],

— la décomposition en ondelettes, généralisation de la décomposition en séries de Fourier à court terme où les différentes valeurs de pixels sont projetée unes sur une base plus générale de l'espace de Hilbert $L^2(\mathbb{R}^2)$ [39].

Ainsi, l'analyse de Fourier discrète permet de définir une notion de densité spectrale ou d'énergie associée à une image. Cette énergie est le module de la réponse complexe suivante :

$$\mathcal{E}(k_x, k_y) = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} I(n,m) \exp^{-j2\pi(\frac{nk_x}{N} + \frac{mk_y}{M})} . \tag{2.3}$$

Par définition, cette multiplication dans le domaine spectral est équivalente à une convolution circulaire dans le domaine spatial. A ce titre, et pour une image de taille $N \times M$, la formule 2.3 peut être vue comme le résultat d'une convolution avec un banc de filtres de tailles $N \times M \times 2$ composé des parties réelles et imaginaires de la base de Fourier restreinte à des fréquences (k_x, k_y) allant de $(0,0)$ à $(N-1, M-1)$. Cette énergie est globale et afin d'incorporer une analyse spatiale dans la décomposition spectrale, une analyse de Fourier à court terme, ou fenêtrée, permet d'estimer une composition spectrale glissante sur les différentes zones de l'image comme suit :

$$\mathcal{E}(x,y) = I(k_x, k_y) \otimes w_{s_\Delta} \exp^{-j2\pi(\frac{nk_x}{N}+\frac{mk_y}{M})} . \tag{2.4}$$

Parmi les fenêtres w_{s_Δ} habituelles, on trouve la fenêtre rectangulaire qui va permettre de restreindre l'analyse à une zone spatiale de taille s autour de $(n \pm \frac{\Delta}{2}, m \pm \frac{\Delta}{2})$, une Gaussienne, les fenêtres de Hamming, Hann, Blackman, etc. qui permettent une restriction et une modulation spatiales de la convolution. En particulier, le filtre de Gabor constitue une analyse de Fourier où la fonction w_{s_Δ} est une Gaussienne. Tandis qu'un filtre de convolution issu de la base de Fourier est paramétré par sa fréquence et son orientation, la construction d'un banc de filtres de Gabor prendra en plus en compte les écart-types σ_x et σ_y de la Gaussienne w_s ainsi que l'échelle s qui définie la taille de la fenêtre analysée. La figure 2.4 montre deux exemples de filtres de Gabor et leurs paramètres. Les filtres

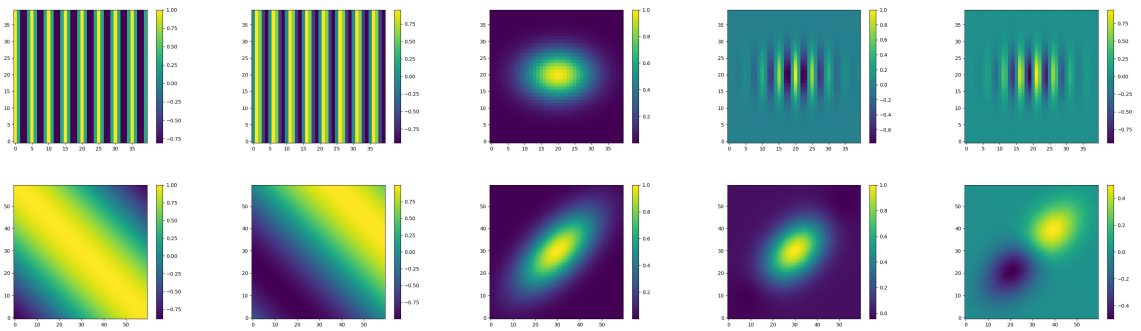


FIGURE 2.4 – Un filtre de Gabor d'échelle s est le produit d'une Gaussienne définie par les écart-types σ_x et σ_y et d'une sinusoïde complexe définie par sa frèquence f et sa phase θ. Exemple du haut : $s = 40$, $\sigma_x = 7$, $\sigma_y = 5$, $k_x = \frac{1}{5}$, $k_y = 0$ et $\theta = 0$. Exemple du bas : $s = 60$, $\sigma_x = 15$, $\sigma_y = 7$, $k_x = k_y = \frac{1}{100}$ et $\theta = \frac{\pi}{4}$.

ainsi définis sont d'autant plus sélectifs en fréquence que leur support spatial est large, et inversement. Ce compromis rejoint le principe général d'incertitude d'Heisenberg-Gabor. Cette représentation jointe fréquentielle et spatiale est associée à la famille des atomes temps-fréquence (espace-fréquence pour les images) et il existe une multitude de familles de base orthonormées Hilbertiennes appelée les ondelettes. Les ondelettes, d'une façon générale, s'avèrent mieux adaptées que les approches statistiques pour les signaux large bande (telles que des images de milieux naturels) et ont eu beaucoup de succès en trai-

tement d'image grâce à leur détection efficace de l'échelle et de la rotation d'une texture qui permet de créer un descripteur robuste. Au delà de leur robustesse, ces méthodes fréquentielles permettent des traitements plus rapides, surtout pour des données à large échelle, car effectués dans l'espace fréquentielplus parcimonieux. Les techniques fréquentielles permettent aussi de faire un rehaussement de l'image à travers la transformation de la distribution fréquentielle des pixels [40].

Une étude intéressante de la relation entre les différentes classes "ensoleillé", "nuageux", "neigeux", "pluvieux" et "brumeux" obtenues avec différentes méthodes basées modèle et texture pour la classification d'images de milieux naturels est faite dans les travaux de Wei-Ta Chu & al. [22].

2.3.3 Méthodes basées apprentissage

D'autres méthodes basées apprentissage profond par réseaux de neurones convolutifs [41][22] de bout-en-bout sont aussi proposées afin de quantifier la brume (et donc la visibilité) à partir d'une base d'images. A ce titre, on pourra trouver une comparaison de techniques de détection de brume faite dans le cadre du projet NTIRE2018 (pour "New Trends in Image Restoration and Enhancement") [42], où la plupart des méthodes proposées sont basées sur un réseau convolutif profond couplé avec le modèle de la formation de la brume présenté plus tôt. Par ailleurs, l'apprentissage de la texture par réseaux de convolution est similaire dans le fond et la forme à l'approche classique de construction d'un banc de filtres de Fourier, de Gabor ou d'ondelettes [43]. De ce fait, une extension naturelle des approches basées apprentissage profond par réseau de convolution est d'inclure une composante spectrale dans l'analyse via des outils plus classiques. Très récemment, une grande quantité d'architectures profondes incluant un opérateur fréquentiel est proposée, par exemple [44][45][46], etc. D'une manière générale, coupler réseaux profonds et filtres fréquentielles revient à apprendre les coefficients des filtres qui caractérisent le mieux une texture.

2.4 Estimation du déplacement des pixels

Dans cette section nous présenterons les méthodes de traitement des données 2D+t envisagées au cours de cette thèse. Nous les avons réparti selon la densité spatiale des cartes de déplacement qui en résulterait.

2.4.1 Approches denses

Nous divisons ces approches en 2 catégories :

— les méthodes basées sur une mesure de similarité,

— les méthodes variationnelles.

2.4.1.1 Approches basées sur des mesures de similarité

Une estimation dense du déplacement des pixels entre une image I_0 de taille $N \times M$ et une image I_1 est possible via la comparaison de celles-ci selon une mesure de similarité. Parmi ces mesures de similarité, on trouve l'information mutuelle [47] et différentes mesures impliquant des L-distances, telles une erreur Euclidienne ou une différence absolue. La méthode représentative de cette famille d'algorithmes est la corrélation-croisée centrée normalisée, ou CC, définie comme suit :

$$CC(I_0, I_1)(\Delta x, \Delta y) = \frac{\sum\limits_{x,y}[I_0 - \bar{I}_0].[I_{1_\Delta} - \bar{I}_{1_\Delta}]}{\sqrt{\sum\limits_{x,y}[I_0 - \bar{I}_0]^2}\sqrt{\sum\limits_{x,y}[I_{1_\Delta} - \bar{I}_{1_\Delta}]^2}}, \tag{2.5}$$

où $I - \bar{I}$ est l'image I dont les valeurs d'intensité lumineuse ont été centrées par soustraction de la valeur moyenne, I_{1_Δ} dénote l'image I_1 translaté spatialement par $(\Delta x, \Delta y)$, "." est la multiplication point par point (ou produit matriciel de Hadamard) et enfin $\sum_{x,y}$ dénote la sommation des valeurs du champs obtenu. Dans l'équation 2.5, le numérateur est une mesure de la dépendance linéaire (ou encore de la distance angulaire) entre les images I_0 et une version translatée de I_1 et la normalisation par le dénominateur permet de créer de l'invariance à l'intensité des pixels. Après normalisation, $\underset{\Delta x, \Delta y}{argmax}(|CC(I_0, I_1)|) \in [-1, 1]$, fournit la translation globale de I_1 qui corrèle le plus avec I_0, et, comme sur la Figure 2.5, il convient d'utiliser une version locale de la CC afin d'estimer la translation de différents patchs de I_{1_Δ} sur toute l'image I_0. La CC locale permet

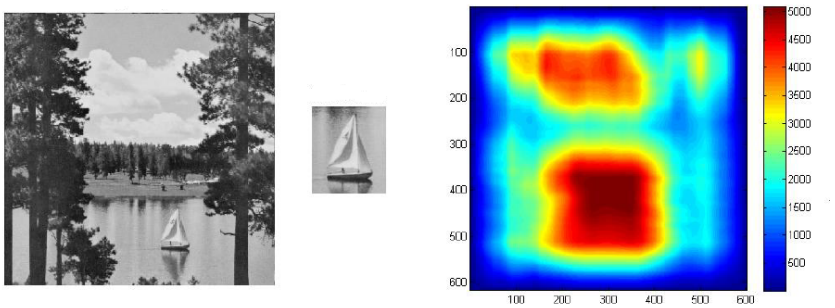


FIGURE 2.5 – La corrélation-croisée non normalisée permet de trouver la translation (u, v) (emplacement) du patch local I_1 (milieu) par rapport à l'image I_0 (à gauche) qui est le maximum de la carte de corrélation-croisée (à droite).

de ce fait, et dans une certaine mesure, d'estimer un mouvement non-rigide approximé par des translations par morceaux voire un champ dense lorsque déplacement est calculé en chaque pixel. Afin d'optimiser ce calcul il est communément suggéré :

— de limiter la fenêtre de recherche si une translation maximale peut être supposée,

— d'éviter les calculs redondants de corrélation croisée en déduisant les valeurs de corrélation en un point (u, v) des valeurs calculée en $(u, v-1)$ ou $(u-1, v)$ [48] ,
— d'effectuer les calculs dans le domaine de Fourier (de par l'équivalence de l'opérateur corrélation avec une convolution).

Ces techniques de corrélation dense peuvent être implémentées de telle sorte que le calcul de similarité soit invariant au changement d'échelle et à la rotation [49][50].

En particulier, Revaud & al. [50] ont récemment proposé une approche basée sur une mesure de corrélation similaire à l'équation 2.5. Cette méthode, nommée DeepMatching (DM), est robuste au changement d'illumination, d'échelle et à la rotation de l'image I_0 par rapport à I_1. Elle repose sur la construction d'une pyramide de cartes de corrélations pour chaque patch, dit patch atomiques, de 4×4 pixels. Contrairement aux approches multi-échelles proposées auparavant [51][49] où I_0 et I_1 sont directement transformées afin de considérer la corrélation sur plusieurs échelles et sous diverses transformations géométriques, avec l'approche DeepMatching seule la base de la pyramide, ici notée CC_0, est calculée directement à partir des images. Les niveaux supérieurs de la pyramide sont ensuite agrégés récursivement par une suite d'opérations sur CC_0, telle que :

$$\begin{aligned}
\text{Max-pooling } 3 \times 3 : \quad & CC(p) \leftarrow \max_{m \in \{-1,0,1\}^2} CC(p+m), \\
\text{Sous-échantillonnage par } 2 : \quad & CC(p) \leftarrow CC(2p), \\
\text{Shift par t} : \quad & CC(p) \leftarrow CC(p-t), \\
\text{Transformation en puissance } \lambda : \quad & CC(p) \leftarrow CC^{\lambda}(p).
\end{aligned} \tag{2.6}$$

Comme on peut le voir sur la figure 2.6, pour une carte de corrélation CC_0 de patchs atomiques (4×4), le niveau supérieur CC_1 (8×8) est d'abord obtenu par max-pooling et sous-échantillonage afin d'avoir la taille adéquate tout en gardant les valeurs de corrélation les plus élevées. Un shift de 1 pixel est effectué pour des raisons pratiques afin que les coordonnées de la carte de corrélation soient cohérentes entre elles (pyramide centrée comme celle de la figure 2.6) et la transformation en puissance $CC \leftarrow CC^{\lambda}$ permet d'introduire une non-linéarité qui met l'accent sur les corrélations les plus importantes. Ce processus est ré-itéré jusqu'au niveau le plus grossier de la pyramide. Finalement, afin d'obtenir une unique carte de corrélation dense est de remonter à une carte de déplacements atomiques Δ_{DM} et qui prend en compte les calculs à toutes les échelles, un "backtracking" à partir de ce sommet de la pyramide est effectué en défaisant ces opérations sur les points ayant le plus corrélé (Figure 2.7).

Pour un point donné, la somme des corrélations sur tout le chemin de "backtracking" constitue son score de corrélation. Il est important de noter que ces étapes dites "dense" ne fournissent pas forcément un déplacement en tout point car le résultat est généralement filtré afin de ne garder que les estimations pertinentes.

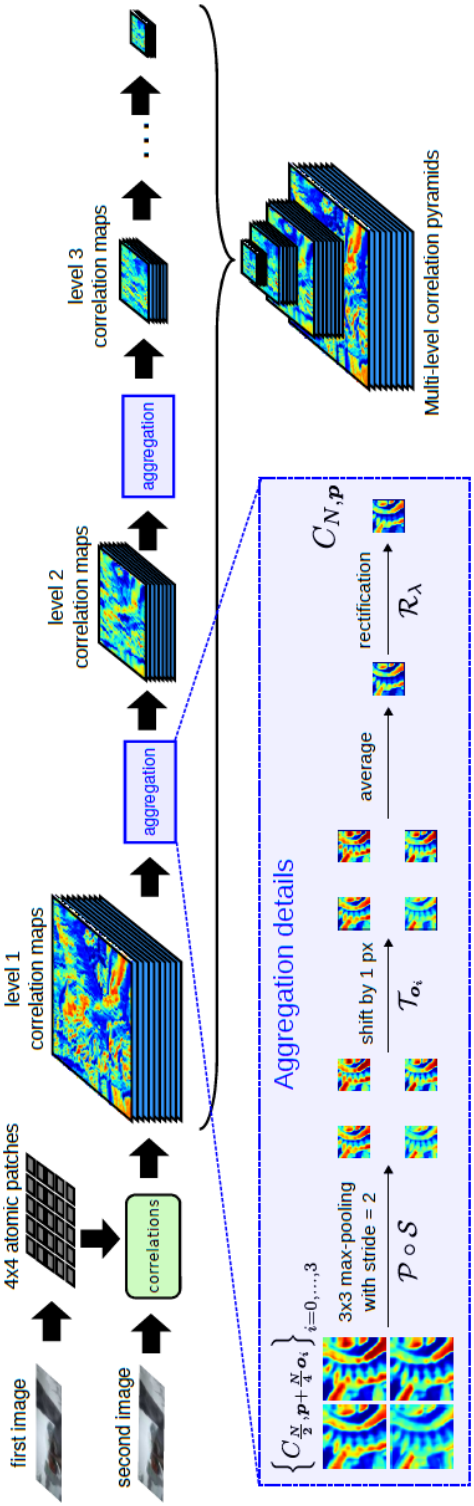


FIGURE 2.6 – Construction d'une pyramide de corrélation avec l'approche DeepMatching [50].

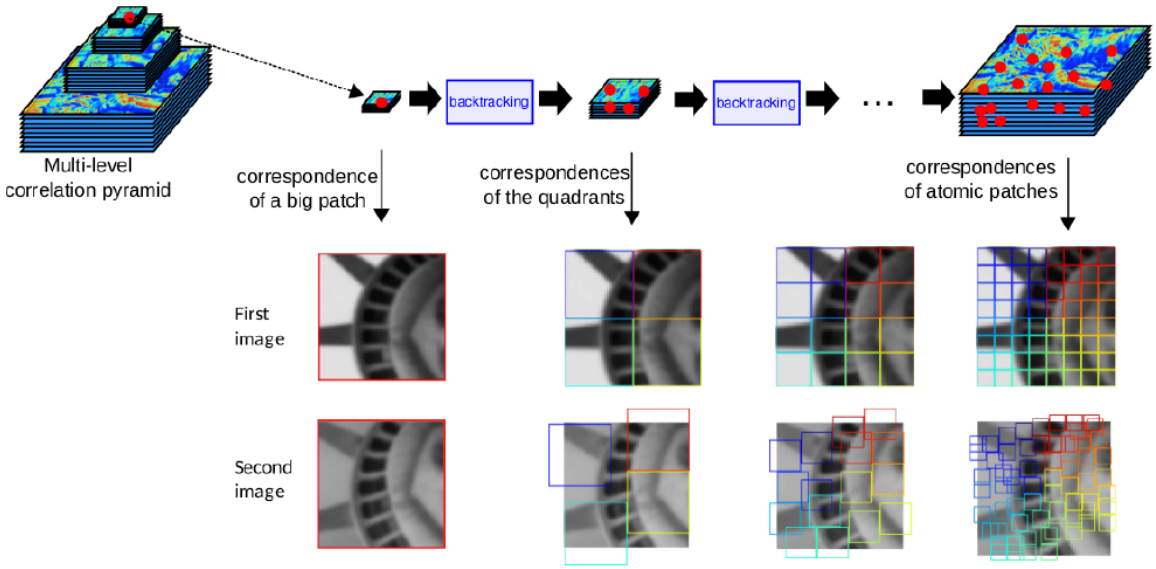


FIGURE 2.7 – "Backtracking" à partir du sommet de la pyramide de corrélation selon l'approche DeepMatching [50].

2.4.1.2 Approches variationelles

La deuxième famille de méthodes denses est la famille des méthodes variationnelles, aussi appelées méthodes différentielles ou flot optique. Le flot optique, concept introduit par Horn & al. [52], est défini comme le champ sde vélocités $\mathbf{OF} = (u, v) = (\frac{dx}{dt}, \frac{dy}{dt})$ mesuré à partir de la variation de la luminance entre 2 images de la même scène. L'hypothèse de départ de l'algorithme du flot optique est alors la conservation de l'intensité lumineuse définie comme suit :

$$I(x, y, t) = I(x + u, y + v, t + \Delta t). \tag{2.7}$$

Après un développement de Taylor au premier ordre de la quantité $I(x+u, y+v, t+\Delta t)$, l'équation 2.7 devient :

$$\begin{aligned} \begin{bmatrix} I_x & I_y \end{bmatrix} \begin{bmatrix} \dfrac{dx}{dt} \\ \dfrac{dy}{dt} \end{bmatrix} + I_t \approx 0 \\ \Leftrightarrow (\nabla I)^T.\mathbf{OF} \approx -I_t, \end{aligned} \tag{2.8}$$

avec $I_c = \frac{\partial I}{\partial c}$. La particularité d'une telle équation est que celle-ci est une équation d'Euler-Lagrange bi-dimentionnelle dont l'étude analytique est déjà maîtrisée. Le problème d'op-

timisation est ainsi défini comme suit :

$$\begin{aligned} &\underset{u,v}{argmin}\varrho(|I(x,y,t) - I(x+u,y+v,t+1)|^2) \\ &\underset{u,v}{argmin}\varrho(|(\nabla I)^T.\mathbf{OF} + I_t|^2). \end{aligned} \tag{2.9}$$

Cette linéarisation n'est valide que sous l'hypothèse d'un déplacement image de faible amplitude et linéaire par rapport aux coordonnées image. De plus, l'hypothèse de conservation de l'intensité lumineuse est un modèle incomplet, car il ne prend en compte que l'information radiométrique et que le problème défini est sous-déterminé. Il convient alors de procéder itérativement et par injection de contraintes supplémentaires (Chapitre 1, section 1.2.1.3). Dans l'approche classique [52], la fonction coût est la norme $L = 2$. Des améliorations génériques possibles sont :

- l'injection d'une contrainte de lissage spatial qui suppose que les pixels voisins varient de la même manière [52],
- L'injection d'une contrainte de conservation des gradients pour relaxer l'hypothèse de départ,
- procéder selon une approche multi-échelle afin de prendre en compte les déplacements importants,
- l'inclusion de termes de second ordre [53],
- l'inclusion des contraintes locales de mise en correspondance [50],
- l'inclusion un terme de symétrie, ou cohérence temporelle, dans la résolution [54],
- la prise en compte de contraintes fréquentielles spatiales et temporelles [55],
- etc.

En particulier, l'approche proposée par Weinzaepfel & al. [50], nommée DeepFlow (DF), repose en plus de l'approche classique [56], sur un terme de conservation des gradients, un terme de lissage spatial (minimisation des gradients) et sur une étape préalable de corrélation par DeepMatching afin d'initialiser l'optimisation non-linéaire. Le problème d'optimisation obtenu est le suivant :

$$\begin{aligned} \underset{u,v}{argmin}(\varrho_0(|I(x,y,t) - I(x+u,y+v,t+1)|^2)+ \\ \alpha.\varrho_1(|\nabla I(x,y,t) - \nabla I(x+u,y+v,t+1)|^2)+ \\ \beta.(\varrho_2(|\nabla u|^2 + |\nabla v|^2)) + \gamma.(\varrho_3(\Delta_{DM}(x,y))). \end{aligned} \tag{2.10}$$

où α, β et γ sont les termes de pondération des différentes contraintes et ∇ est l'opérateur gradient. Contrairement à l'approche initiale où la minimisation est faite dans le sens des moindres carrés, la fonction coût utilisée par l'algorithme Deepflow est $\varrho_i(s^2) = \sqrt{s^2 + \epsilon^2}$, avec $\epsilon = 0.001$. De par sa ressemblance avec la norme L_1, cette fonction est convexe et robuste aux valeurs aberrantes et le terme de régularisation ϵ règle le problème de non-dérivabilité autour de 0 [57].

D'une manière générale, tous les algorithmes de flot optique découlent du modèle initiale avec un terme d'attache aux données, où une constance de certaines propriétés de l'image est assumée, et un terme de lissage qui décrit la granulosité spatiale des cartes de vélocité résultantes. Plus de détails sur les approches basées sur le flot optique, les différents termes de données et de régularisation, la fonction coût associée et leurs résolutions sont disponibles dans les travaux [58][59][57] et [50]. En outre, Trianshu & al [60] montrent que les méthodes variationnelles sont plus à même de décrire le mouvement d'un fluide que les méthodes basées sur de la corrélation. En effet, les contraintes utilisées pour le calcul du flot optique sont tirées de l'équation de conservation de la quantité de mouvement, principe fondamental de la dynamique des fluides.

2.4.2 Approches éparses

Quand les images sont de taille importante et que la déformation radiométrique et géométrique entre les deux images à comparer est importante, les approches denses peuvent être coûteuses et générer beaucoup de valeurs aberrantes. Pour plus de parcimonie et de pertinence il convient d'utiliser des méthodes éparses qui vont fournir une description mathématique des images en seulement quelques points, dits points d'intérêt, suivi d'une mise en correspondance dans l'espace des descripteurs. Si besoin, les mesures éparses peuvent être "densifiées" par exemple par interpolation ou par estimation d'un modèle paramétrique en aval.

2.4.2.1 Détection et description de points d'intérêts

Tout d'abord, les points d'intérêt à décrire sont, par définition, des points saillants dans l'image riches en information spatiale, historiquement associés à la notion de coins dans l'image. A l'instar des bâches ou coins réflecteurs souvent utilisés en télédétection de par leur visibilité dans le spectre optique ou pour leur forte réponse aux ondes électromagnétique en imagerie radar, un point d'intérêt adéquat pour le traitement de time-lapses de milieux naturels possède les propriétés suivantes :

— reproductible, c.à.d détecté sur toutes les images du time-lapse,
— sa détection est robuste aux changements photométriques,
— sa détection est robuste aux changements d'échelle afin de capturer un déplacement vers la caméra,
— sa description est faite par rapport à des caractéristiques singulières sur l'image, c.à.d qu'on évitera de considérer des points de contrôle corrélés avec un motif pseudo-périodique sur l'image,
— sa détection est bien réparties sur la scène observée, ce qui nécessite un consensus avec le point précédent.

On évitera alors les méthodes de caractérisation de la texture précédemment cités pour

définir des points d'intérêt. Il est à noter qu'il est courant de trouver des points d'intérêt provenant de primitives non ponctuelles et on parlera indifféremment de points d'intérêt, de "blob", "feature" ou de région d'intérêt. En détection proximale, ces régions peuvent en outre être introduites manuellement dans l'image afin d'ancrer certaines coordonnées spatiales, comme par exemple par la mise en place de bâches marquées, ou encore via des objets géophysiques singuliers et visibles sur les images, a détecter automatiquement ou manuellement.

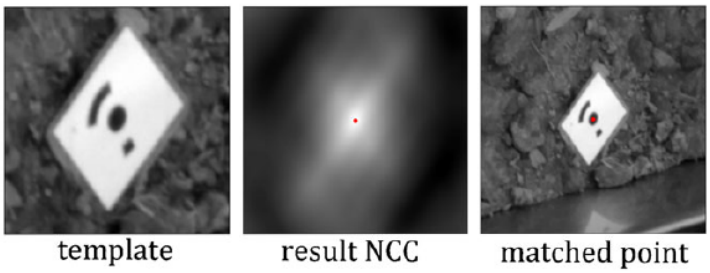


FIGURE 2.8 – Exemple de point d'intérêt singulier placé sur une scène observée par détection optique proximale. Figure tirée des travaux de Eltner et al. [61].

Concernant les méthodes automatiques, la robustesse désirée peut être créée directement lors de la détection ou lors de la construction des descripteurs. Parmi les pratiques d'usage lors de la détection et de la description de points d'intérêt robustes, on peut citer :

— l'utilisation des gradients spatiaux de l'image,
— la construction d'espaces d'échelles par filtrage passe-bas et sous-échantillonage,
— la détection et la description de régions co-variantes,
— l'utilisation de méthodes statistiques basées sur des histogrammes locaux.

Gradients spatiaux Les méthodes les plus classiques ainsi que la plupart des méthodes modernes reposent sur les gradients de l'image, ou les contours. Les gradients sont mathématiquement intéressants, sont relativement faciles à calculer, car constituent des zones de discontinuité de l'intensité des pixels et sont invariants aux altérations radiométriques, car ne prennent en compte que la différence entre des pixels voisins. D'une manière générale, les dérivées sont approximées par convolution avec des filtres binaires, par exemple de Canny ou Sobel, qui permettent de calculer les variations spatiales horizontales et verticales sur un voisinage de 3×3 pixels. Comme pour le cas de la texture, les orientations locales des gradients peuvent aussi être estimées sur différentes échelles et par rapport à différentes orientations par convolution avec des filtres du type Gabor adaptés [62] ou par la construction d'un espace d'échelle. Une fois les gradients estimés, les points d'intérêt peuvent en être déduits en prenant le centroïde des contours, les extrémités, les maximas de l'amplitude, l'intersection de deux contours [63], les minima des rayons de courbure [64][64], etc.

La structure locale des gradients peut aussi être capturée via la mesure suivante [65] :

$$AC_w(I)(\Delta x, \Delta y) = \sum_{x,y} [I_w - I_{w_\Delta}]^2. \tag{2.11}$$

La mesure dans l'équation 2.11 contient un terme d'auto-corrélation de $[I - I_{w_\Delta}] = [I - I(x \pm \Delta x, y \pm \Delta x)]$ sur un voisinage w centré autour de (x, y). Cette mesure permet en effet de capturer les régularités de l'image et, comme pour la CC, cette estimation s'effectue localement sur une fenêtre glissante. Une forte mesure AC sur un voisinage w traduit donc un fort gradient dans la direction induite par $(\Delta x, \Delta y)$ et il convient donc d'effectuer le calcul sur plusieurs directions afin de discriminer les points d'intérêt du type coin des contours.

Les travaux de Harris & al. [65] ont notamment introduit la notion de matrice d'auto-corrélation, ou matrice des moments d'ordre 2, nommée A, avec le développement sur l'équation 2.11 suivant :

$$I_w(x + \Delta x, y + \Delta x) \approx I_w + \begin{bmatrix} I_{w_x} & I_{w_y} \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta x \end{bmatrix},$$

d'où

$$\begin{aligned} AC(I) &= \left[\sum_{x,y} \begin{bmatrix} I_{w_x} & I_{w_y} \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \right]^2, \\ &= \begin{bmatrix} \Delta x & \Delta y \end{bmatrix} \sum_{x,y} \begin{bmatrix} I_{w_x}^2 & I_x I_{w_y}, \\ I_{w_x} I_{w_y} & I_{w_y}^2 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}, \\ &= \begin{bmatrix} \Delta x & \Delta y \end{bmatrix} A \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}. \end{aligned} \tag{2.12}$$

Pour un voisinage et une direction induite par $(\Delta x, \Delta y)$ donnés, l'amplitude du gradient peut ensuite être déduite directement à partir de la matrice A et plus particulièrement via l'amplitude de ses valeurs propres. Les travaux qui ont suivi ceux de Harris & al. se différencient par la manière de capturer l'information sur les valeurs singulières et de paramétrer le seuil [66][67][68][69]. En somme, cette détection est invariante par rotation, du fait de l'invariance des valeurs singulières de A, mais pas par changement d'échelle. Pour cette raison, il convient de les utiliser dans un schéma multi-échelles [66][68][67].

Espaces d'échelle Les espaces d'échelles sont une représentation formelle du signal image, introduite en 1991 [70], qui permet de créer de l'invariance à l'échelle pendant la détection de points d'intérêt. Cette méthode consiste à représenter l'image I sur différentes octaves et échelles, où les différentes échelles sont obtenues par floutage Gaussien progressif et

les octaves par sous-échantillonage. Cette méthode a amorcé de multiples travaux qui ont entre autres abouti à la notion de différence de Gaussiennes utilisée par la méthode SIFT [71] et ses variantes [72][73] et qui montrent que la différence entre deux versions de l'image sur deux échelles différentes constitue un filtre passe-bande idéal pour détecter les points et les contours saillants à une échelle donnée. Ce filtre est aussi une approximation de l'opérateur Laplacien tel que proposé par Lindeberg [74].

La détection de régions co-variantes Les opérateurs SIFT et les coins de Harris couplés avec une représentation en espaces d'échelles constituent des détecteurs robustes, invariants aux altérations radiométriques, à la rotation ainsi qu'à un changement d'échelle uniforme. Les travaux de Linderberg & al. [75] ont permis la généralisation de ce résultat aux transformations affines (rotation + échelle + *cisaillement*). La méthode proposée est basée sur la notion de régions co-variantes ellipsoïdales obtenues par adaptation affine itérative de régions circulaires. Afin de compléter ces travaux, Mikolajczyk et Schmidt [68] proposent une nouvelle formulation des détecteurs de coins de Harris multi-échelles et affinement adaptés. Ainsi, dans la nouvelle formulation, les noyaux Gaussiens uniformes utilisés pour la construction de l'espace d'échelle sont considérés comme des distributions multivariées. Les vecteurs propres et les valeurs propres de la matrice de covariance Σ_{xy} sur la région inspectée définissent alors la rotation, le cisaillement et l'échelle de l'ellipsoïde [75]. Chaque région détectée est enfin normalisée comme sur la Figure 2.9.

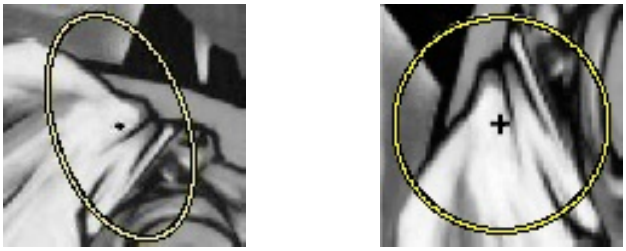


FIGURE 2.9 – Point d'intérêt détecté dans sa région co-variante (gauche) et après normalisation (droite) [76].

Cette invariance locale aux transformations affines est très intéressante car une transformation spatiale affine par morceau peut, dans une certaine mesure, modéliser un mouvement non-rigide [77][78] ou une différence de perspective entre deux images. Plus de détails sur les détecteurs robustes sont disponibles dans les travaux de Mikolajczyk, Schmidt & al. [79][68][80][76].

2.4.2.2 La description de points d'intérêt

Afin d'apparier l'ensemble des point d'intérêt détectés sur deux images, un descripteur doit être associé à chaque point. Ce descripteur peut provenir directement de l'intensité

des pixels ou celle des gradients, des mesures qui ont permis sa détection ou être calculé dans une étape en aval. Parmi les méthodes les plus robustes on trouve l'histogramme de gradients orientés [37], ou HOG, qui vient compléter une détection invariante. Le descripteur HOG a été utilisé avec la méthode SIFT [71] et, comme introduit dans la section 2.3.2.1 de ce chapitre, permet de caractériser la distribution locale des gradients. Si le voisinage est affinement normalisé, les directions de gradients encodées seront absolues et permettront un description invariante par transformation affine.

2.4.2.3 Mise en correspondance

L'homologation de points, ou régions, d'intérêt dépend de la stratégie de mise en correspondance. La mise en correspondance peut par exemple se faire via une des mesures de similarité définies plus tôt. La plupart des méthodes robustes sont néanmoins basée sur une classification plus-proche-voisin via la construction d'un arbre KD (arbre décisionnel k-dimensionnel) dans l'espace des descripteurs. Des méthodes de détection du proche voisin approximatif sont néanmoins plus adaptée aux calculs en grandes dimensions. Celles-ci ne reposent pas sur une recherche exhaustive points ou régions homologues mais sur l'optimisation de la recherche au sein de l'arbre-KD [81] ou encore sur l'hypothèse qu'un appariement aléatoire peut être raffiné efficacement afin de trouver un plus-proche-voisin approximatif cohérent [82]. D'autre part, en présence d'une importante quantité de bruit, la mise en correspondance génère beaucoup de faux positifs et une étape de désambiguïsation, ou de validation, est nécessaire. Beaucoup de méthodes existent, par exemple par RANSAC [83], partitionnement (clustering) [84], optimisation [85] ou encore via des contraintes géométriques ou temporelles. Une méthode simple est de poser un seuil sur la mesure de similarité suffisante entre les descripteurs à apparier. Une variante proposée par Lowe [71] repose aussi sur un seuil, mais sur le deuxième plus proche voisin. Dans cette méthode, un appariement est considéré valide si $\frac{D_1}{D_2} <$ seuil, D_1 étant la distance du point d'intérêt à apparier à son plus proche voisin et D_2 sa distance au deuxième plus proche voisin.

2.5 Alignement paramétrique

L'appariement de points entre d'une I_i avec une image maîtresse I_m peut aboutir à un alignement paramétrique des deux images. L'alignement paramétrique consiste à trouver un modèle, parmi les modèles décrits dans la partie 1.3, qui une fois inversé, aligne I_m et I_i. Cet alignement peut être local, par exemple de la taille de la fenêtre de corrélation ou entre les voisinages de points d'intérêt homologues, ou global, auquel cas on parlera de recalage [86]. Typiquement, un modèle global sert à compenser le mouvement de la caméra entre les différentes images d'un time-lapse. Dans ce cas, le co-recalage ne devrait

prendre en compte que les zones fixes des images à recaler (en masquant le reste) et le modèle paramétrique adéquat est une homographie qui approxime un faible changement de perspective (en réalité, il modélise un mouvement du plan image avec une caméra fixe). Quelque soient l'approche envisagée, globale ou locale, et le modèle paramétrique A, le problème est formulé comme suit :

$$\begin{aligned} I_m(\mathbf{x}) &\approx AI_i(\mathbf{x}) \\ &\Leftrightarrow \\ I_m(\mathbf{x}) &\approx I_i(W(\mathbf{x};\mathbf{p}_A)), \end{aligned} \tag{2.13}$$

où $W(\mathbf{x};\mathbf{p}_A)$ sont les nouvelles coordonnées $\mathbf{x}' = \begin{pmatrix} x' \\ y' \end{pmatrix}$ sous-pixelliques obtenus après transformation avec les paramètres $\mathbf{p}_A$ de A et $I_i(W(\mathbf{x};\mathbf{p}_A))$ est l'image I_i après interpolation des intensité de pixels sur les nouvelles coordonnées. L'obtension de la nouvelle image transformée $I_i' = I_i(W(\mathbf{x};\mathbf{p}_A))$ peut se faire d'une façon directe :

$$\begin{aligned} \mathbf{x}' &= W(\mathbf{x};\mathbf{p}_A), \\ I_i'(\mathbf{x}') &\leftarrow I_i(\mathbf{x}), \end{aligned} \tag{2.14}$$

ou inverse :

$$\begin{aligned} \mathbf{x} &= W(\mathbf{x}';\mathbf{p}_A)^{-1}, \\ I_i'(\mathbf{x}') &\leftarrow I_i(\mathbf{x}). \end{aligned} \tag{2.15}$$

Une transformation inverse est plus optimale dans le sens où $W((x,y);\mathbf{p}_A)$ génére empiriquement plus une couverture non régulière dans l'image re-échantillonnée I_i' (dû à des coordonnées en dehors de l'image). Pour un co-recalage global d'un ensemble d'images ou d'un time-lapse, le choix de l'image maîtresse I_m peut être arbitraire, par exemple la première image classifiée visible pendant la phase de pré-traitements, ou basé sur les différentes heuristiques calculées sur les images I_i.

L'approche classique pour le recalage bidate, appelé LK pour ses inventeurs Bruce D. Lucas et Takeo Kanade [87], consiste à minimiser itérativement la différence entre l'image $I_i'(\mathbf{x}')$ transformée et l'image maîtresse $I_m(\mathbf{x})$ dans le sens des moindres carrés et pour les paramètres p_A de A, tel que :

$$\underset{\Delta\mathbf{p}_A}{argmin} \sum_{\mathbf{x}} [I_m(\mathbf{x}) - I_i(W(\mathbf{x};\mathbf{p}_A + \Delta\mathbf{p}_A))]^2. \tag{2.16}$$

La fonctionnelle à minimiser (équation 2.16) est non-convexe et non-linéaire (même si $W(\mathbf{x};\mathbf{p}_A)$ est linéaire) car $I(\mathbf{x})$ est généralement non-linéaire en $\mathbf{x}$ et comme toute

optimisation non-linéaire, il est important de disposer d'une bonne estimation initiale de A.

Plus concrètement, à chaque itération un incrément $\mathbf{p}_A \leftarrow \mathbf{p}_A + \Delta p$ est calculé. Le calcul de l'incrément Δp se fait dans le sens de la descente de gradients [87] tel que :

$$\Delta p = H^{-1} \sum_{\mathbf{x}} \left[\nabla I \frac{\partial A}{\partial \mathbf{p}_A} \right] [I_m(\mathbf{x}) - I_i(W(\mathbf{x}; \mathbf{p}_A + \Delta \mathbf{p}_A))], \tag{2.17}$$

avec :

$$H = \sum_{\mathbf{x}} \left[\nabla I \frac{\partial A}{\partial \mathbf{p}_A} \right]^T \left[\nabla I \frac{\partial A}{\partial \mathbf{p}_A} \right], \tag{2.18}$$

où H est une approximation de la matrice Hessienne du modèle paramétrique A. Bien que cet algorithme soit très efficace, sa limite réside dans le fait que la matrice Hessienne H soit calculée à chaque itération. L'algorithme LK a ensuite été amélioré dans sa version "compositionnelle" où les auteurs ont montré l'équivalence de l'incrémentation $W(\mathbf{x}; \mathbf{p}_A + \Delta \mathbf{p}_A)$ avec $W(W(\mathbf{x}; \Delta \mathbf{p}_A); \mathbf{p}_A))$, ce qui permet de remédier à la limitation citée [88]. Cette approche, ainsi que ses variantes, est largement l'approche la plus utilisée dans la pratique. En particulier, une variante proposée par Bartoli [89] est spécialement robuste aux variations photométriques par incorporation d'un terme α dans le calcul de $I'_i(\mathbf{x}')$ tel que :

$$\mathbf{x}' = W((\mathbf{x}; \mathbf{p}_A), \alpha), \tag{2.19}$$

où α capture la différence des intensités de pixels entre I_i et I_m et est incrémentée à chaque itération. En d'autres termes, avec cette formulation, le recalage n'est pas seulement géométrique mais en plus radiométrique. Une autre variante proposée par Evangelidis & al. [90] propose d'utiliser un terme de corrélation-croisée normalisée au lieu d'une moindres carrés pour quantifier la distance entre I_i et I_m à chaque itération. Le problème devient :

$$\underset{\Delta \mathbf{p}_A}{argmax} \;\; CC(I_m, I_i(W(W(\mathbf{x}; \Delta \mathbf{p}_A); \mathbf{p}_A))), \tag{2.20}$$

et est résolut d'une façon similaire. Comme introduit dans la section 2.4.1.1 de ce chapitre, cette métrique est aussi robuste aux variations radiométriques.

2.6 Conclusion

Dans ce deuxième chapitre de l'état de l'art, nous avons vu un panel des méthodes permettant de traiter des time-lapses monoscopiques de milieux naturels. Ces méthodes incluent les techniques de classification d'un ensemble d'images selon la pertinence de l'information qu'elles contiennent et celles de calcul du déplacement pixellique sur un couple 2D+t.

L'information pertinente sur une image de milieux naturels peut être quantifiée par la détection des conditions météorologiques sur la scène observée via un modèle physique ou encore par des méthodes d'analyse harmonique qui permettent de quantifier la texture spatiale observée. Une classification permet ensuite de rejeter les images ayant un indice de visibilité aberrant afin qu'elles n'induisent pas de biais sur le reste des mesures de la chaîne de traitement.

Quant au calcul du déplacement, plusieurs types d'approches existent dans l'état de l'art. Celles-ci peuvent être basées sur l'intensité des pixels, tel que les méthodes de corrélation croisée, sur les gradients spatiaux, tel que les méthodes éparses de détection de points clés saillants, ou encore sur la minimisation d'une fonction coût englobant des contraintes photométriques, différentielles et de régularisation telles que les méthodes variationnelles de calcul du flot optique. Notamment, une méthode nommée DeepFlow [50] est classée parmi les méthodes ayants obtenues l'erreur la moins importante dans certaines études comparatives [91] et est dite robuste face aux textures répétitives.

Chapitre 3

Méthodes monodates et bidates stéréoscopiques

Sommaire du chapitre

3.1 Introduction **43**

3.2 Caractérisation des données stéréoscopiques en milieux naturels **44**

3.3 Instrumentation et modélisation de la caméra perspective **44**

3.3.1 Modèle optique d'un système d'observation 44

3.3.2 Instrumentation et système d'observation 47

3.4 Reconstruction non-métrique **48**

3.4.1 Homographies et projections perspectives 48

3.4.2 Modélisation homogène 50

3.4.3 Vision stéréoscopique 50

3.4.4 Contraintes stéréoscopiques 52

3.4.5 Triangulation 56

3.5 Reconstruction de champs de vélocités *East, North, Up* **57**

3.5.1 Reconstruction métrique 57

3.5.2 Géo-référencement 58

3.5.3 Projection de vélocité pixellique sur un modèle 3D 58

3.6 Optimisation et robustesse **59**

3.6.1 Lors de l'estimation des contraintes stéréoscopiques 59

3.6.2 Lors de la triangulation 60

3.6.3 Lors de l'utilisation de coordonnées homogènes 61

3.7 Estimation de l'incertitude dans les mesures 3D **61**

3.8 Conclusion **63**

3.1 Introduction

Ce chapitre est dédié à l'étude des méthodes existantes pour le traitement d'un couple stéréoscopique (2×2D) d'images ainsi que d'un quadruple stéréo-temporel tel que présenté sur la figure 3.1 ((2×2D)+t et 2×(2D+t)).

Pour ce faire, nous nous intéresserons dans un premier temps au modèle de caméra adopté et aux outils mathématiques couramment utilisés pour la reconstruction et la cartographie d'une scène observée par un système d'observation optique. Nous introduirons alors les fondamentaux de la géométrie projective, les modèles qui en découlent et comment les inverser à l'aide de primitives dans les images. Nous aborderons ensuite les méthodes qui permettent de calculer des déplacements métriques et géo-référencés à partir d'un quadruple stéréo-temporel et les algorithmes robustes qui permettent d'obtenir des résultats avec un minimum d'erreur dans la plage d'incertitude.

Nous finirons par les méthodes qui permettent de propager l'erreur pixellique des mesures 2D et 2D+t vers les mesures 3D et 3D+t.

3.2 Caractérisation des données stéréoscopiques en milieux naturels

Données monodates stéréoscopiques 2×2D Une paire stéréoscopique d'images correspond à deux images de la même scène prises sous une perspective différente et préférablement au même moment. La particularité de telles données pour la cartographie de milieux naturels est le choix limité pour le positionnement du système d'observation dû aux contraintes terrain (notamment en milieu alpin) et la difficulté de calibrer avec un objet étalon conçu en laboratoire telle une mire de calibration à dimensions précisément connues. Ceci est notamment dû à la distance à laquelle la scène est généralement observée.

Données bidates stéréoscopiques (2×2D)+t et 2×(2D+t) Ces données servent au calcul du déplacement (ou vélocité) tridimensionnel et métrique. Comme pour les déplacements pixelliques en milieux naturels, ceux-ci seront généralement denses et non-rigides. Les données (2×2D)+t font référence à deux représentations 3D de la scène observée à deux dates différentes tandis que les données 2×(2D+t) font référence à deux champs de déplacements ou de vélocité pixelliques (droite & gauche) observées avec une différence de perspective (figure 3.1).

3.3 Instrumentation et modélisation de la caméra perspective

3.3.1 Modèle optique d'un système d'observation

Une caméra, ou appareil photo, est un capteur de lumière composé d'un système de lentilles qui font converger les rayons vers un seul point avant de les capturer sur un plan image. La source de lumière peut être observée directement ou émettre des ondes qui

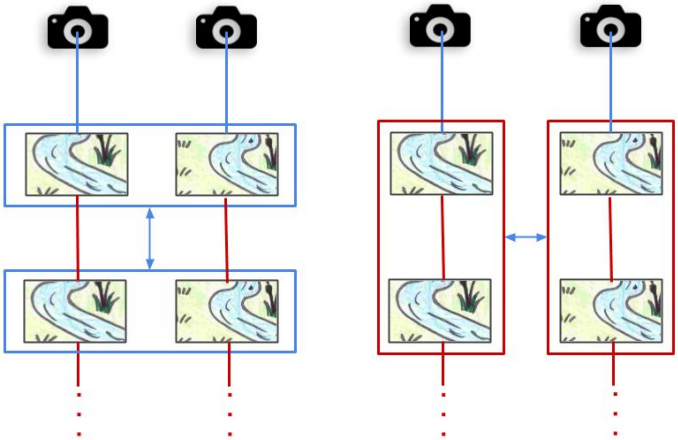


FIGURE 3.1 – Différents types de données pour le calcul du déplacement métrique sur un modèle tridimensionnel. Gauche : (2×2D)+t. Droite : 2×(2D+t).

vont subir une réflexion sur les objets observés et une diffusion dans l'atmosphère. Elle se compose d'un objectif qui focalise la lumière de la scène et d'un boîtier qui comprend le mécanisme de capture et stockage de l'image sur un support numérique (capteur CCD ou CMOS) ou physique (film photographique). Cette composition remonte au modèle dit "sténopé" ou "caméra obscura" utilisé dès l'antiquité pour observer les éclipses solaires dans une chambre noire (figure 3.2).

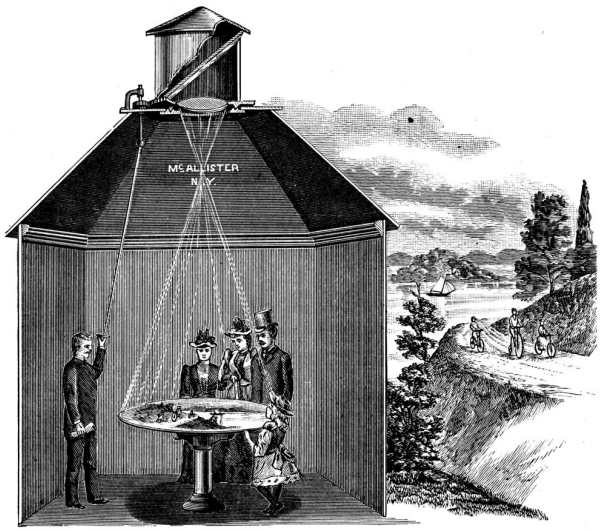


FIGURE 3.2 – Modèle sténopé ou camera obscura. Catalogue, William Y. McAllister, New York, c. 1890.

L'ouverture du modèle sténopé (figure 3.2) équivaut au point ou zone de convergence de l'objectif d'une caméra moderne et est le centre optique ou centre de projection, qui va inverser la projection observée. Le modèle optique permettant d'étudier le phénomène d'acquisition d'une image (et ainsi de l'inverser afin de reconstruire la scène observée) place

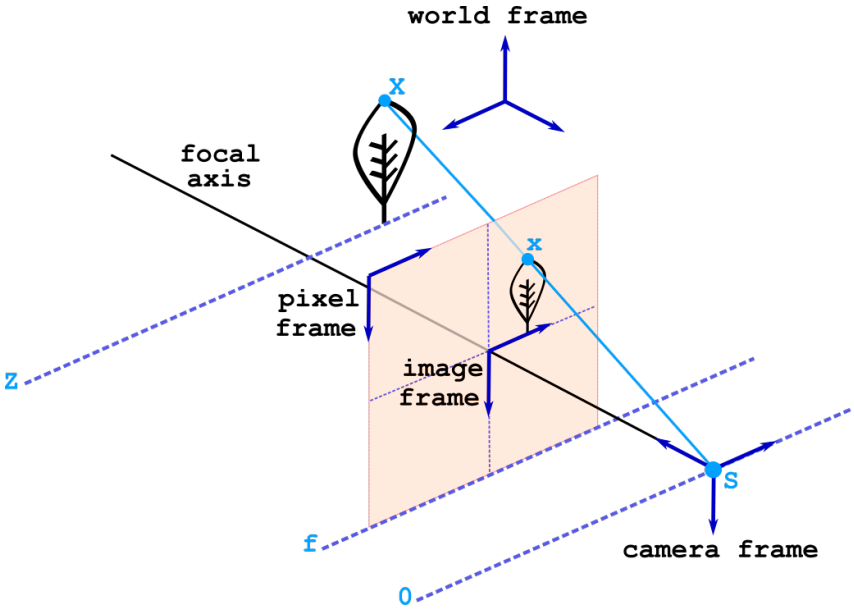


FIGURE 3.3 – Modélisation des paramètres de la caméra. S est le centre de projection de la caméra situé sur le plan de profondeur 0. Le plan image est situé sur le plan de profondeur égale à la focale f et peut être décrit par le repère des coordonnées en pixel $(\mathbf{x}, \mathbf{y})$ ou le repère image $(\mathbf{u}, \mathbf{v})$. Il contient le point image $\mathbf{x} = (x, y)^T$. Le point imagé $\mathbf{X} = (X, Y, Z)^T$ est situé sur le plan de profondeur Z.

le plan image entre le centre de projection et l'objet imagé pour plus de simplicité. Le modèle adopté en vision par ordinateur (figure 3.3) prend alors en comptes les paramètres suivants :

— la distance focale f, qui est la distance définie par la projection orthogonale du centre de projection sur le plan image,

— le point central $\mathbf{x_0} = (x_0, y_0)$, qui est l'intersection de l'axe focal et du plan image supposés orthogonaux.

Les équations qui en découlent sont (Thalès) :

$$\begin{aligned} X &= \tfrac{w_c}{w_i} u = \tfrac{Z}{F_p} u, \\ Y &= \tfrac{w_c}{w_i} v = \tfrac{Z}{F_p} v, \end{aligned} \tag{3.1}$$

où $\frac{Z}{F_p}$ est alors la résolution au sol en mètres par pixel qui permet une mise à l'échelle, F_p est la distance focale convertie en pixels de la même manière que dans l'équation 3.1 :

$$F_p = f \frac{w_i}{w_c}, \tag{3.2}$$

w_i et w_c sont respectivement la largeur de l'image en pixels et la largeur du capteur CCD photosensible et u et v sont les coordonnées pixels converties en coordonnées image par translation telle que :

$$\begin{aligned} u &= x - x_0, \\ v &= y - y_0. \end{aligned} \tag{3.3}$$

Utiliser les hauteurs au lieu des largeurs dans les équations 3.1 et 3.2 donne une conversion quasi équivalente puisque dans le modèle sténopé $f = f_x = f_y$.

3.3.2 Instrumentation et système d'observation

Les paramètres du modèle sténopé sont plus exactement les paramètres intrinsèques de la caméra et sont propres à l'objectif utilisé. Les varier peut changer radicalement les images observées, c'est pourquoi il est crucial d'opter pour un objectif capable de réaliser l'angle et profondeur de champ désirés (figure 3.4). Pour l'acquisition d'un time-lapse sur une portée temporelle importante telle que pour une observation continue de milieux naturels, il est important d'opter pour un objectif à focale fixe pour limiter la dérive des paramètres dans le temps. Ceci est d'autant plus vrai que les conditions climatiques du milieu en question sont extrêmes.

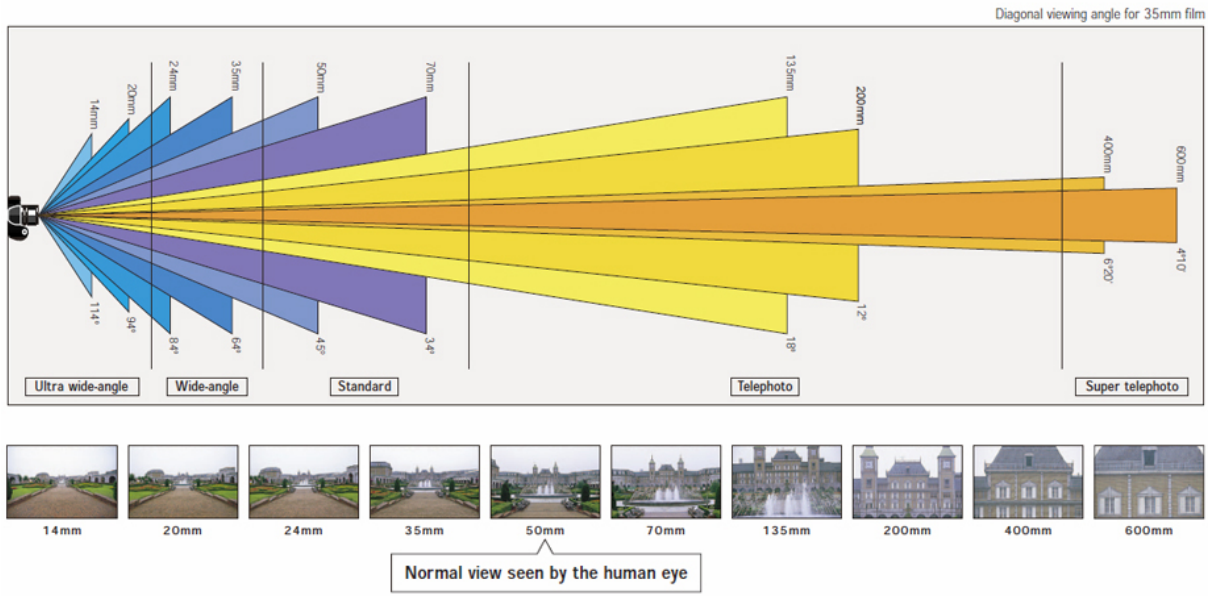


FIGURE 3.4 – Champs de vision et focale. Image tirée de [92].

D'autres grandeurs que la focale permettent d'influencer l'acquisition d'une image, à savoir :

— la vitesse d'obturation, qui détermine la durée d'exposition à la lumière,
— l'ouverture, qui dépend de la focale et du diamètre de la lentille et qui détermine la profondeur de champ ou zone de netteté (zone de points de convergence),
— la sensibilité ISO, ou sensibilité à la lumière.

Pour l'observation de milieux naturels, un grand-angle est préconisé afin de capturer la totalité de l'objet d'étude à une distance importante. De plus, avec un grand-angle la profondeur de champ est plus importante à ouverture égale par rapport à une plus longue focale. Cela permet d'observer la quasi-totalité de la scène avec netteté. D'autre part, un grand angle accentue les perspectives, mais un angle supérieur à $\sim 95°$ peut créer une

distorsion dite en barillet avec une petite focale et un compromis doit donc être fait.

3.4 Reconstruction non-métrique

Il y a une multitude de méthodes pour aboutir à une reconstruction 3D non métrique. On peut par exemple citer :

— la reconstruction multi-vues (Structure from motion) ou stéréo non calibrée [93][5],
— la reconstruction monoscopique à partir du changement de luminosité, de la texture ou des ombres [94][95][56] (shape from shading),
— la reconstruction à partir de caméras RGB-D [96] ou encore de capteurs laser,
— la reconstruction à partir de la perspective d'une image (intersection de droites parallèles, points de fuite) [97],
— la reconstruction à partir d'un modèle de la scène (Shape from Template) [98],
— etc.

Ces méthodes permettent en effet de capturer la structure de la scène observée, pour la plupart à une transformation projective 3D près. Dans ce qui suit, nous allons nous intéresser aux modèles utilisés en reconstruction 3D stéréoscopique.

3.4.1 Homographies et projections perspectives

Le modèle de l'équation 3.1 n'est pas linéaire. Utiliser les paradigmes de la géométrie projective permet une modélisation linéaire du modèle de caméra présenté. La forme linéaire des transformations géométriques $2D$, 3D et $2D/3D$ introduites dans la partie 1.3 émane aussi de cette notion. Il convient donc d'introduire brièvement certaines propriétés de la géométrie projective.

La géométrie projective provient initialement de la généralisation du théorème de Thalès aux faisceaux de droites (ensembles de droites concourantes) qui a mené à la notion du birapport tel celui illustré par la figure 3.5. Sur cette illustration, nous pouvons imaginer que les droites (A, D) et (A', D') sont des plans image. Ils sont alors liés par une homographie $2D$ ($H_{3\times3}$) définie comme une projection centrale modélisée par une application *linéaire* d'un l'espace projectif $\mathbb{P}^2$ sur lui-même. Cette représentation linéaire est une alternative à la représentation Euclidienne et modélise les notions intuitives de perspective et d'horizon grâce à l'utilisation de coordonnées homogènes.

Les coordonnées homogènes sont des coordonnées cartésiennes auxquelles une dimension, l'échelle projective, est rajoutée telle qu'un point A' de coordonnées cartésiennes (x, y) sera modélisé par un faisceau (x, y, w). L'échelle projective w est sa distance par rapport au centre de projection et tous les points sur le même faisceau seront considérés équivalents : $A' \sim A = \lambda A'$. Lors de l'acquisition d'une image, une "dé-homogénisation"

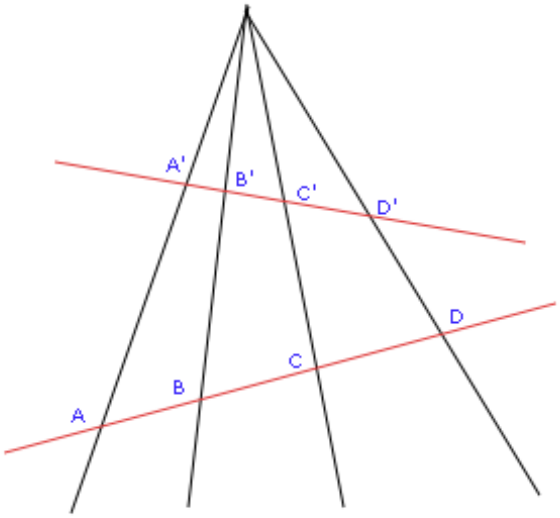


FIGURE 3.5 – Intersection d'un faisceau de 4 droites avec deux droites non-parallèles définissant deux bi-rapports égaux.

ρ^{-1} a lieu comme suit :

$$\rho^{-1}\begin{pmatrix} x \\ y \\ w \end{pmatrix} \sim \begin{pmatrix} \frac{x}{w} \\ \frac{y}{w} \end{pmatrix}. \tag{3.4}$$

Dans ce cas, A sera perçu comme étant plus petit que A' sur le plan image et les points à l'infini convergeront vers un point de fuite. La "dé-homogénisation" est explicitement utilisée par les logiciels de conception d'objets et de films 3D et implicitement par les peintres de la renaissance pour accentuer les perspectives lors de la projection centrale d'un objet 3D sur un plan $2D$.

L'avantage principal de l'utilisation de cette notation linéaire est que toute transformation géométrique est alors décomposable en une suite de transformations élémentaires à savoir, une translation, une rotation, un cisaillement et une homothétie (cette décomposition n'est pas unique).

D'une façon générale, une homographie $H_{n\times n}$ est toute application linéaire bijective d'un espace projectif $\mathbb{P}^n$ sur lui même. Par exemple, à l'instar d'une homographie $H_{3\times 3}$ qui permet de recaler deux images ayant une légère différence de perspective (figure 3.5), une transformation $H_{4\times 4}$ permet de recaler deux volumes 3D et possède les mêmes propriétés d'invariance. D'une façon encore plus générale, une projection perspective est l'application linéaire d'un espace projectif $\mathbb{P}^n$ vers un espace $\mathbb{P}^m$. L'acquisition d'une image constitue donc une projection perspective de $\mathbb{P}^3$ vers $\mathbb{P}^2$ et la reconstruction 3D son inverse.

L'algorithme de référence pour inverser une application linéaire d'un espace projectif $\mathbb{P}^n$ dans un espace projectif $\mathbb{P}^m$ (par régression) est appelé DLT (Direct Linear Transform) [99]. Avec l'algorithme DLT, le problème linéarisé est résolue par SVD en imposant

une contrainte de norme unitaire sur le vecteur contenant les paramètres de H. Cette contrainte permet de relaxer le problème en fixant arbitrairement l'échelle projective associée à celui-ci. Cet algorithme est notamment utilisé pour le recalage homographique monoscopique.

3.4.2 Modélisation homogène

La transformation linéaire qui découle des équations 3.1 peut donc être reformulée sous sa forme homogène comme suit :

$$\begin{pmatrix} u \\ v \\ F_p \end{pmatrix} \approx \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \tag{3.5}$$

et, en incluant certains paramètres supplémentaires, peut être exprimée matriciellement comme suit :

$$\lambda \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} F_p & s & x_0 \\ 0 & kF_p & y_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}. \tag{3.6}$$

Cette transformation peut être décomposée en une série de déformations $2D$ à savoir : une translation pour remonter à $(u, v)^T$ suivie d'une mise à l'échelle (équation 3.1) et d'un cisaillement :

$$\lambda \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & x_0 \\ 0 & 1 & y_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} F_p & 0 & 0 \\ 0 & kF_p & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & F_p/s & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}. \tag{3.7}$$

Contrairement au modèle sténopé de départ, la notation homogène permet d'inclure des facteurs de distorsion tels que le facteur de non-orthogonalité s (cisaillement) et le facteur k qui permet de décrire des pixels non rectangulaires et permet ainsi une mise à l'échelle anisotrope. Les facteurs de distorsions sont parfois modélisés avec un ordre supérieur et viennent généralement d'un défaut de la caméra ou du processus de numérisation, d'un post-traitement des images à reconstruire (comme un recadrage ou re-échantillonnage) ou de caméras spéciales telles que pour l'observation de l'espace. Dans ce qui suit, la matrice 3×3 contenant les paramètres intrinsèques d'une caméra c (équation 3.6) sera notée K_c.

3.4.3 Vision stéréoscopique

La stéréoscopie est l'ensemble des techniques mises en œuvre afin de reproduire une perception du relief à partir du mouvement parallaxe entre deux images planes (2×2D)

acquises avec deux caméras perspectives c et c'. En posant c comme la caméra "maîtresse" ou de référence, il convient d'exprimer les grandeurs $\mathbf{X} = (X, Y, Z)^T$ dans un même repère afin de fusionner les observations gauches et droites lors de la résolution du système induit par K_c. Ce repère peut être l'un des deux repères associés aux caméras, un repère choisi arbitrairement entre les deux caméras ou un repère géo-référencé dans le cas d'applications géodésiques.

Quel que le repère choisi, une première étape consiste à trouver la relation géométrique entre les deux repères caméras. Cette relation est une transformation tridimensionnelle Euclidienne, ou rigide, définie comme suit :

$$\lambda \begin{pmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{pmatrix} = \begin{pmatrix} & & & T_x \\ & R_{3\times3} & & T_y \\ & & & T_Z \\ 0 & 0 & 0 & 1 \end{pmatrix}_{c'\to c} \begin{pmatrix} X_{c'} \\ Y_{c'} \\ Z_{c'} \\ 1 \end{pmatrix}. \tag{3.8}$$

Les points de l'espace peuvent ensuite être exprimés par rapport au repère de la caméra maîtresse c tel que :

$$\lambda \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}_{c'} = K_{c'} \begin{pmatrix} & & T_x \\ & R_{3\times3} & T_y \\ & & T_Z \end{pmatrix}_{c'\to c} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}, \tag{3.9}$$

$$\lambda \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}_{c} = K_{c} \begin{pmatrix} & & 0 \\ & I_{3\times3} & 0 \\ & & 0 \end{pmatrix}_{c\to c} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}. \tag{3.10}$$

La transformation unitaire dans l'équation 3.10 est la projection perspective $H_{3\times4}$ dans sa forme la plus simple qui fait correspondre la caméra référence à elle-même (avant de convertir en coordonnées pixels) et celle de l'équation 3.9 inclut une translation vers la deuxième caméra et une rotation 3D qui peut, entre autres, être exprimée avec les angles d'Eulers ϕ, ψ et θ comme suit :

$$R_{3\times3} = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\psi & 0 & \sin\psi \\ 0 & 1 & 0 \\ -\sin\psi & 0 & \cos\psi \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}. \tag{3.11}$$

L'équation 3.11 décrit une série de 3 rotations à savoir : une précession autour de l'axe optique, une nutation et une rotation propre. La matrice des paramètres de rotation et de translation, appelés paramètres extrinsèques ou pose de la caméra, sera notée R_T et le

produit $K_c R_T$, la matrice de projection, sera notée P_c pour une caméra c donnée.

3.4.4 Contraintes stéréoscopiques

La matrice fondamentale, présentée pour la première fois par Quang-Tuan Luong [100], est la transformation projective qui va nous permettre d'établir la contrainte stéréoscopique qui lie 2 points homologues vus par les deux caméras. Elle peut être calculée à partir de traitements purement $2D$, sans aucune connaissance a priori des paramètres (intrinsèques et extrinsèques) de la caméra. Similairement aux traitements cités dans la section 2, le calcul de la matrice fondamentale commence par une mise en correspondance, dense ou éparse, des deux plans image. La matrice fondamentale est issue de la géométrie épipolaire qui est l'étude géométrique qui part des hypothèses de co-planarité des points homologues, $\mathbf{x}_{c'}$ et $\mathbf{X}$. Les équations de projection perspective des deux caméras c et c', ou gauche et droite, couplées avec les contraintes épipolaires donne lieu à l'équation suivante :

$$\mathbf{x}_{c'}^T F \mathbf{x}_c = 0. \tag{3.12}$$

En effet, considérons un point 3D $\mathbf{X}$ projeté en $\mathbf{x}_{c'}$ et $\mathbf{x}_c$; comme nous pouvons le voir sur la figure 3.6, ces 3 points ainsi que les centres des deux caméras S_c et $S_{c'}$ sont co-planaires et appartiennent au plan épipolaire noté Π (en vert dans la figure 3.6) . L'intersection de Π avec les deux plans images donnent les droites épipolaires $l_{c'}$ et l_c. L'épipole d'une image prise par la caméra c noté $\mathbf{e_c}$, est le point par lequel passe toutes les droites épipolaires sur cette image. Il correspond à la projection du centre de l'autre caméra dans l'image en question et peut très bien avoir des coordonnées pixel en dehors de l'image (négatives par exemple).

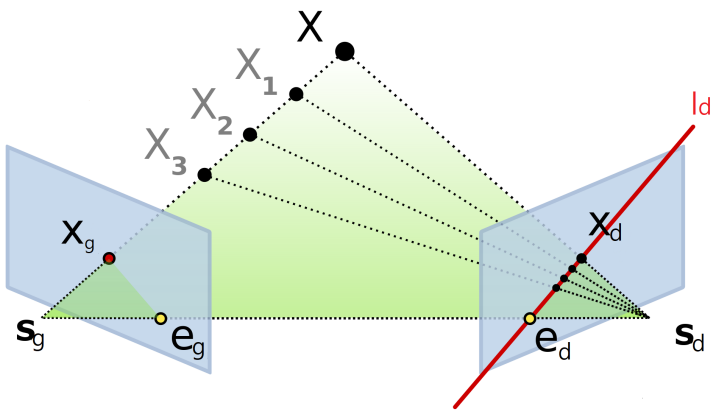


FIGURE 3.6 – Géométrie épipolaire avec, par exemple, $c = g$ et $c' = d$.

L'application qui va relier toutes les lignes épipolaires vues par la droite à leurs points

image homologues sur l'image gauche est donc encapsulée dans F comme suit :

$$l_c = F^T \mathbf{x}_{c'} \Leftrightarrow l_{c'} = F\mathbf{x}_c, \tag{3.13}$$

où l_c est la ligne épipolaire homologue à $\mathbf{x}_c$ et qui passe par $\mathbf{x}_{c'}$, et vice-versa avec une matrice fondamentale équivalente. Elle représente donc une application linéaire d'un espace projectif bidimensionnel vers un espace projectif unidimensionnel, et a une dimension 3×3 pour un rang de 2 (matrice singulière à déterminant nul). La matrice fondamentale peut être utilisée de différentes manières dans les chaînes de traitements tridimensionnels. De façon générale, elle permet une rectification, implicite ou explicite, des images sur un plan image commun. Elle peut par exemple très facilement être calculée à partir des matrices de projection P_c et $P_{c'}$ [5] obtenues via une procédure de calibration ou étalonnage préalable de la caméra. Cet étalonnage permet en effet de retrouver les paramètres intrinsèques et extrinsèques du système d'observation stéréoscopique grâce à une mire de calibration ayant au moins 6 triplets $\mathbf{x_c}$, $\mathbf{x_{c'}}$ et $\mathbf{X}$ connus avec précision. Il convient ensuite de calculer F et de retrouver les points homologues entre les deux vues lors d'une application ultérieure afin de les reconstruire. En photogrammétrie le problème est un peu différent puisqu'une calibration préalable est le plus souvent impossible.

Le problème de la photogrammétrie est donc plutôt de retrouver $P_{c'\to c} = K_{c'} R_{T_{c'\to c}}$ à partir de la mise en correspondance $2D$, avec c l'indice de la caméra prise comme référence. Deux cas de figure se présentent alors :

— les matrices K_c et $K_{c'}$ sont totalement inconnues,

— les matrices K_c et $K_{c'}$ sont connues ou approximées.

Paramètres intrinsèques totalement inconnus : Ce cas de figure peut arriver quand les images à reconstruire proviennent d'une source inconnue, par exemple retrouvées sur internet, ou que les paramètres intrinsèques sont trop instables pour envisager une approximation. Une estimation de $P_{c'\to c}$ peut être réalisée en résolvant un système linéaire composé de N points homologues liés par la relation de l'équation 3.12.

Pour ce faire, et pour des points homologues $\mathbf{x_c}$ et $\mathbf{x_{c'}}$ sous leur forme homogène unitaire, respectivement, $(x_c, y_c, 1)^T$ et $(x_{c'}, y_{c'}, 1)^T$, la matrice fondamentale F sera d'abord convertie en un vecteur 9×9 tel que :

$$F \to f_u = [F_{00}F_{01}F_{02}F_{10}F_{11}F_{12}F_{20}F_{21}F_{22}]^T, \tag{3.14}$$

et l'équation 3.12 sera reformulée comme suit :

$$[x_{c'}x_c, x_{c'}y_c, x_{c'}, y_{c'}x_c, y_{c'}y_c, y_{c'}, x_c, y_c, 1]f_u = 0. \tag{3.15}$$

L'équation 3.15 vient des calculs de produits vectoriels dans un espace préhilbertien

(*inner-product*) et est à résoudre, ou inverser, pour f_u. Pour qu'une solution existe, il convient de former un système d'au moins 7 points homologues non colinéaires puisque F possède 7 degrés de liberté (8 degrés de liberté pour une homographie et un degré en moins en forçant la contrainte de singularité). On obtient alors le système suivant :

$$\begin{pmatrix} x^1_{c'}x^1_c & x^1_{c'}y^1_c & x^1_{c'} & y^1_{c'}x^1_c & y^1_{c'}y^1_c & y^1_{c'} & x^1_c & y^1_c & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x^7_{c'}x^7_c & x^7_{c'}y^7_c & x^7_{c'} & y^7_{c'}x^7_c & y^7_{c'}y^7_c & y^7_{c'} & x^7_c & y^7_c & 1 \end{pmatrix} f_u = 0 \\ \Leftrightarrow \\ A_{f_{7\times 9}} f_u = 0, \tag{3.16}$$

qui correspond au système minimal à résoudre par une variante robuste et dédiée de l'algorithme DLT, présenté plus tôt. Nous verrons dans la section 3.6 la méthodologie qui permet d'obtenir une solution optimale à ce problème. La matrice de la caméra $P_{c'\to c}$ peut ensuite être estimée comme suit :

$$P_{c'\to c} = K_{c'} R_{T_{c'\to c}} = \begin{pmatrix} & e_{c'_x} \\ S_y F & e_{c'_y} \\ & 1 \end{pmatrix}, \tag{3.17}$$

avec S_y une matrice anti-symétrique quelconque et $e_{c'} = (e_{c'_x}, e_{c'_y}, 1)^T$ l'épipole contenu dans l'image c tel que :

$$F^T e_{c'} = 0 \quad \text{(noyau de } F^T\text{)}. \tag{3.18}$$

La matrice de projection de la deuxième caméra est la projection perspective définie dans l'équation 3.10. Un bon choix de la matrice anti-symétrique S_y, suggéré par les auteurs [100], est :

$$S = [e_{c'}]_\times = \begin{pmatrix} 0 & -1 & e_{c'_y} \\ 1 & 0 & -e_{c'_x} \\ -e_{c'_y} & e_{c'_x} & 0 \end{pmatrix} \tag{3.19}$$

où la notation dans 3.19 vient de la notation matricielle d'un produit vectoriel.

et cette méthode vient du théorème qui stipule qui si F est une matrice fondamentale correspondant à un système d'observation stéréoscopique ayant pour matrices de projections P_c, alors $P_{c'}^T F P_c$ est une matrice anti-symétrique. Nous renvoyons vers la source [100] ainsi qu'à [5] pour plus de détails et de preuves. Il est important de noter que F est une transformation issue de calculs purement projectifs et de ce fait, est invariante par transformation projective de l'espace 3D : c'est ce qui est connu sous le nom d'ambiguïté projective d'un système non-calibré (K_c inconnue). D'autre part, $P_{c'\to c}$ est aussi une transformation projective et est connu à un facteur d'échelle prés.

Paramètres intrinsèques connus ou approximés La matrice des paramètres intrinsèques de la caméra est connue avec précision si une calibration préalable est possible, autrement une approximation [101], plus ou moins robuste selon le système d'observation, peut être calculée comme suit :

— la longueur focale f provient de notre connaissance de l'objectif de caméra utilisé,
— le point central est considéré être au milieu de l'image,
— les paramètres de distorsions sont négligés.

Dans ce cas, il convient de passer par la matrice dite Essentielle E présentée pour la première fois par Longuet Higgins [102]. La relation entre E et F est comme suit :

$$E = K_{c'}^T F K_c. \tag{3.20}$$

Les deux matrices se rapportent à la géométrie épipolaire entre les deux vues et la différence réside dans le fait que la matrice fondamentale soit définie dans l'espace des coordonnées de l'image originale tandis que la matrice essentielle est en coordonnées normalisées, c'est à dire sur les coordonnées $\hat{\mathbf{x}}_c$ translatés et à l'échelle telle que :

$$\hat{\mathbf{x}}_c = K_c^{-1}\mathbf{x}_c. \tag{3.21}$$

Retrouver E peut se faire à partir de la matrice fondamentale (équation 3.20), ou en composant un système linéaire similaire à celui de l'équation 3.16 mais avec les nouvelles coordonnées normalisées. L'avantage d'utiliser E au lieu de F est que celle-ci possède 5 et non 7 degrés de liberté. Cette relaxation vient de la nouvelle contrainte sur E, en plus de celles héritées de F, qui stipule que les 2 valeurs singulières de E sont égales. Le cas minimal pour résoudre la version normalisée de l'équation 3.16 ne nécessite donc que 5 correspondances. La pose de la caméra c', $R_{T_{c'\to c}}$, peut ensuite être calculée avec l'algorithme suivant :

1. décomposition en valeurs singulières de E :

$$E = U\Sigma V^T \tag{3.22}$$

,

2. définir

$$W = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{et} \quad Z = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{3.23}$$

,

3. $[T]_\times = UW\Sigma U^T$ ou $[T]_\times = -UZU^T$,
4. $R_{3\times3} = UWV^T$ ou $R_{3\times3} = UW^TV^T$.

Ainsi :

$$R_{T_{c'\to c}} = \begin{pmatrix} & T_x \\ R_{3\times 3} & T_y \\ & T_Z \end{pmatrix}_{c'\to c} . \tag{3.24}$$

Cet algorithme, présenté plus en détail dans [5] et [102], donne lieu à N estimations de E, et donc de $R_{T_{c'\to c}}$ qui peuvent être classés en 4 types de configuration dont un seul, celui où les points de l'espace $\mathbf{X}$ se situent en face des deux caméras, est valide.

3.4.5 Triangulation

Une fois les matrices de projection déterminées, il convient de reconstruire tous les points homologues par triangulation pour obtenir un nuage de point. Dans le cas où les matrices de projections ont été obtenues via la matrice fondamentale F, la reconstruction est à une projection perspective près. En d'autres termes, il existe une transformation projective dans l'espace 3D telle que, pour tout $\mathbf{X}$:

$$\mathbf{X_m} = H_{4\times 4}\mathbf{X}, \tag{3.25}$$

où $\mathbf{X_m}$ est la vraie reconstruction Euclidienne ou métrique sur laquelle des mesures géodésiques sont possibles. Si la reconstruction a été obtenue via la matrice essentielle E , cette ambiguïté est levée, seule reste l'ambiguïté sur l'échelle et $H_{4\times 4}$ est alors une transformation rigide ou isométrie 3D. Dans les deux cas, aucun calcul géodésique n'est possible. Ces reconstructions restent utiles dans certaines applications où seule importe une appréciation qualitative du modèle 3D.

L'algorithme de triangulation le plus simple, de par sa linéarité, est formulé comme suit : pour chaque paire de points homologues $\mathbf{x_c} \approx \lambda(x_c, y_c, 1)^T$ et $\mathbf{x'_c} = \lambda'(x'_c, y'_c, 1)^T$, $(c, c') = (d, g)$, l'équation $\mathbf{x_c} = P_c\mathbf{X}$ est tout d'abord réécrite sous la forme d'un produit vectoriel $\mathbf{x_c}\times P_c\mathbf{X} = 0$ [5]. Cette écriture permet de se débarrasser du facteur de profondeur projective λ et donne lieu à 3 équations de la forme :

$$\begin{cases} x_c({p_c^3}^T\mathbf{X}) - ({p_c^1}^T\mathbf{X}) = 0 \\ y_c({p_c^3}^T\mathbf{X}) - ({p_c^2}^T\mathbf{X}) = 0 \\ x_c({p_c^2}^T\mathbf{X}) - y_c({p_c^1}^T\mathbf{X}) = 0 \end{cases} , \tag{3.26}$$

avec ${p_c^i}^T$ la i-ème colonne de P_c. Seulement 2 des équations dans le système d'équations 3.26 sont linéairement indépendantes et seront donc utilisées. Afin de combiner les informations issues des deux caméras c et c', et trianguler le point de l'espace $\mathbf{X}$, il convient d'empiler

les équations issues des deux caméras comme suit :

$$\begin{pmatrix} x_c p_c^{3^T} - p_c^{1^T} \\ y_c p_c^{3^T} - p_c^{2^T} \\ x_{c'} p_{c'}^{3^T} - p_{c'}^{1^T} \\ y_{c'} p_{c'}^{3^T} - p_{c'}^{2^T} \end{pmatrix} \mathbf{X} =\Leftrightarrow A\mathbf{X} = 0. \tag{3.27}$$

La résolution pour $\mathbf{X}$ est linéaire et peut se faire avec l'algorithme DLT ou en omettant l'échelle projective de $\mathbf{X}$ (en la fixant à 1). L'inconvénient de cette relaxation est que en assumant $\mathbf{X} = (X, Y, Z, 1)^T$, les points proches de l'horizon et ayant un facteur d'échelle proche de 0 ne seront plus représentés ce qui revient à perdre les avantages introduits par la représentation homogène des points de l'espace.

Quelle que soit la méthode adoptée, la résolution de l'équation 3.27 consiste en la méthode de triangulation directe minimale étant données 2 matrices de projection P_c et P'_c. Nous verrons par la suite (section 3.6) d'autres méthodes qui permettent une triangulation itérative par minimisation d'autres critères géométriques non-linéaires et qui permettent de raffiner la reconstruction projective et ses paramètres. Nous renvoyons vers les travaux d'origine [99] pour plus de preuves et de détails.

3.5 Reconstruction de champs de vélocités *East, North, Up*

3.5.1 Reconstruction métrique

Plusieurs approches permettent de retrouver une reconstruction métrique sur laquelle des mesures géodésiques sont possibles. Si aucune information de calibration n'est disponible, il convient d'avoir recours à une méthode stratifiée qui va tout d'abord retrouver une reconstruction dite affine puis Euclidienne à l'aide de primitives sur le nuage de points telle que des plans à l'infini et des coniques [5]. Autrement, des points de contrôle au sol permettent soit de contraindre la reconstruction durant l'étape de triangulation, soit d'appliquer une transformation projective de correction. De telles mesures sont souvent disponibles pour l'observation des milieux naturels. À défaut de pouvoir calibrer en amont, cette méthode permet donc de calibrer en aval. Une transformation projective 3D possède 15 degrés de liberté et peut donc être calculée à partir de 5 mesures de points au sol dont au moins 4 non co-planaires : chaque point 3D fourni 3 équations linéairement indépendantes pour la résolution de le l'équation 3.25. La reconstruction Euclienne est retrouvée soit en transformant le modèle projectif avec $H_{4\times4}$, soit en reconstruisant les points image $\mathbf{x_c}$ avec une nouvelle matrice de projection $PH_{4\times4}^{-1}$.

Toutefois, tandis qu'une simple mise à l'échelle peut être réalisée efficacement par minimisation d'une distance entre les deux nuages de points (données GPS et données triangulées), par exemple via l'algorithme itératif ICP (Iterative Closest Point) [103]) ou

via la solution analytique proposée par Horn [104], une telle transformation projective est plus communément estimée à partir des données 2D pour ensuite être inférée en 3D [105]. Hormis lors d'un recalage avec les données de calibration, ce type de méthodes de recalage dans l'espace 3D peut s'avérer utiles lors de la compensation du mouvement non désiré de la caméra entre deux reconstructions ou encore si l'estimation de données intrinsèques et/ou extrinsèques est modifiée après reconstruction. Il convient alors de détecter la projection des points d'ancrage sur au moins un des plans image.

3.5.2 Géo-référencement

La reconstruction métrique ainsi obtenue peut être exprimée dans un repère géo-référencé par extrapolation des points d'ancrage au sol incorporés. Il convient soit de choisir un repère de géo-référencement Euclidien, soit de construire une table de correspondance (look-up table) référençant tous les points $\mathbf{X}$. Il est préconisé d'avoir des mesures GPS bien réparties sur l'ensemble des zones fixes de la scène observée pour un géo-référencement optimal.

3.5.3 Projection de vélocité pixellique sur un modèle 3D

La technique de l'état de l'art qui permet une telle fusion de données bi et tridimensionnelles est connue sous le nom de "*scene-flow*", présentée pour la première fois par Sundar Vedula [106]. Le "*scene-flow*" **SF** est l'analogue tridimensionnel du flot optique. On peut donc dire que, théoriquement le flot optique est la projection du "*scene-flow*" sur le plan image d'une caméra c. Il existe une multitude d'approches qui permettent de calculer un champ de déplacements tridimensionnels selon les données disponibles et l'application.

Exploitation des données (2×2D)+t Exploiter les données (2×2D)+t revient à mesurer la différence directement entre deux nuages de points obtenus à des dates différentes afin de détecter le changement/déplacement survenu entre deux dates [107]. L'inconvénient avec une telle approche est que les méthodes de calcul du déplacement image $2D$, que nous considérons plus matures et moins sensibles que la reconstruction 3D, ne sont pas du tout exploitées. Les ambiguïtés projectives, d'une part, et celle dans la définition d'un plus-proche-voisin, d'un autre, risquent en effet d'introduire un biais important sur les mesures de déplacements métriques particulièrement dans le cas non calibré en présence de reliefs importants.

Exploitation des données 2×(2D+t) L'une des techniques les plus simples revient à mettre à l'échelle les champs de vélocités $2D$ à l'aide de la profondeur pour obtenir des valeurs en mètres [2] [108]. Dans ce cas, l'équation 3.1 est utilisée avec une carte de profondeur

issue d'un calcul de disparité stéréoscopique. L'avantage de cette méthode est qu'une construction de nuage de points n'est pas nécessaire. L'inconvénient est que la stéréoscopie n'est exploitée que pour la génération de la profondeur et qu'aucune triangulation tridimensionnelle n'est réellement réalisée. Cette technique est aussi connue sous le nom de "*range-flow*" ou de reconstruction $2.5D$. D'autre part, les travaux de Vedula [106] ont introduit une méthode variationnelle, extension de la méthode du calcul de flot optique, pour le calcul des champs de déplacements 3D. La différence entre ces travaux et la technique du flot optique est que le ou les termes de régularisation sont plus adaptées aux données 3D et contiennent les paramètres du système d'observation stéréoscopique tel que les matrices de projection P_c, ou directement la profondeur Z.

Ainsi, l'énergie à minimiser est analogue à celle du flot optique, mais conduit à un résultat stéréscopiquement cohérent. L'estimation des mesures tridimensionnelles et de la vélocité par la minimisation d'une fonctionnelle variationnelle peuvent être découplés [109] [110] ou joints [111] [112] [113]. Nous pouvons aussi trouver dans la littérature une fusion $2D/3D$ basée sur l'apprentissage de la profondeur avec des réseaux convolutifs [114] ou des filtres de Kalman [115] (appelée reconstruction $6D$). Ainsi, les calculs peuvent donc être réalisés sur les images (méthodes variationnelles 2×(2D+t)), directement sur les nuages de points [116] [117], les modèles de maillage qui en découlent [118] ou à l'aide d'images issues de caméras RGB-D (tel que la Kinect) [119] [120] où le "*scene-flow*" est directement résolu en utilisant les données de profondeur pour contraindre le mouvement dans le domaine de l'image. A noter que d'une façon générale l'utilité principale des algorithmes de calcul du "*scene-flow*" est de pouvoir construire des termes de pondération ou des contraintes en fonction de la profondeur qui permettent d'éviter les erreurs le long des discontinuités de la surface reconstruite.

Une taxinomie quasi complète des méthodes de calcul du déplacement tridimensionnel selon le type de données, les contraintes imposées et l'énergie minimisée est publiée dans les travaux de Zike Yan & al. [121].

3.6 Optimisation et robustesse

3.6.1 Lors de l'estimation des contraintes stéréoscopiques

La robustesse d'un algorithme de reconstruction 3D stéréoscopique commence par la robustesse de la détection et de la mise en correspondance des points $\mathbf{x_c}$ et $\mathbf{x_{c'}}$. On préconisera alors une mise en correspondance adaptée aux données (chapitre 2). D'autre part, les précautions suivantes sont aussi à considérer.

Solution unique Comme introduit dans la section 3.4, la résolution linéaire du système $A_{f_{7\times 9}}$ pour retrouver la matrice fondamentale nécessite au moins 7 points homologues $\mathbf{x_c}$

et $\mathbf{x_{c'}}$. Le problème $A_{f_{7\times 9}} f_u = 0$ (équation 3.16), appelé le "7-point algorithm", donne entre 1 à 3 solutions de F dans le cas idéal, plus en présence de bruit. Pour qu'une solution unique existe (toujours à un facteur d'échelle près), le rang de la matrice A_f doit être au moins égal à 8. Pour cela il convient de rajouter une paire de points homologues supplémentaires. Dans ce cas, une solution unique existe (dans le sens strict si les données ne sont pas bruitées et dans le sens des moindres carrés sinon). Cet algorithme, appelé le "8-point algorithm" est présenté et défendu plus en détail dans les travaux de Zisserman & al. [122].

Pertinence Pour retrouver une matrice F qui tient compte de la géométrie de toute la scène, il convient d'utiliser toutes les correspondances calculées conjointement avec le "8-point algorithm". Pour cela, l'algorithme RANSAC, présenté dans le chapitre 1, est utilisé afin d'obtenir une matrice fondamentale qui engendre le moins d'observations aberrantes.

Concernant l'homologue calibrée de F, la matrice essentielle E, un raisonnement similaire existe ; "5-point algorithm". Une étude sur le nombre de solutions obtenues et la résolution optimale est publiée dans les travaux de Nister [123].

3.6.2 Lors de la triangulation

L'équation linéaire 3.27 pour la triangulation n'a pas vraiment de sens géométrique et il convient d'utiliser une métrique géométrique en aval pour raffiner la solution obtenue. Il existe une multitude d'erreurs géométriques possibles [124] qui peuvent être convexes ou non-convexes. Par exemple, l'erreur de reprojection peut être définie comme la distance entre le point $\mathbf{x_c}$ qui a servi à retrouver P_c, et sa reprojection $\hat{\mathbf{x}}_\mathbf{c}$ à partir de $\mathbf{X}$ sur le plan image (équation 3.28).

$$\begin{cases} \hat{\mathbf{x}}_\mathbf{c} = \frac{P_c^1 \mathbf{X}}{P_c^3 \mathbf{X}} \\ \hat{\mathbf{y}}_\mathbf{c} = \frac{P_c^2 \mathbf{X}}{P_c^3 \mathbf{X}} \end{cases} \tag{3.28}$$

L'erreur de reprojection est alors définie comme suit :

$$||d(\mathbf{x}_c, \hat{\mathbf{x}}_\mathbf{c}) + d(\mathbf{x}_{c'}, \hat{\mathbf{x}}_{\mathbf{c'}})||_2^2. \tag{3.29}$$

On appelle ajustement des faisceaux, ou "*bundle-adjustment*", la minimisation de l'erreur de reprojection avec comme variables de décision à la fois le point de l'espace $\mathbf{X}$ mais aussi par rapport à P_c et $P_{c'}$. Cette fonctionnelle est fortement non-linéaire et la recherche de minima est généralement initialisée par la solution linéaire (équation 3.27) et retrouvée par la méthode itérative de Levenberg-Marquadt présentée brièvement dans le chapitre 1. L'algorithme d'ajustement des faisceaux est présenté en détail dans les travaux de Triggs & al. [125].

La minimisation d'une erreur géométrique peut aussi être accompagnée d'un terme de régularisation qui va introduire une contrainte sur la profondeur Z.

3.6.3 Lors de l'utilisation de coordonnées homogènes

Selon Hartley & al. [5] [122], quelle que soit l'application projective à résoudre, une normalisation des coordonnées homogènes $\mathbf{x_c}$ de manière qu'elles soient de l'ordre de grandeur de 1 ($||\mathbf{x_c}|| = \sqrt{2}$) et une translation vers le centroïde de l'image permettent une régression optimale. En effet, la normalisation améliore le conditionnement numérique des systèmes projectifs linéaires pour des raisons de poids des observations (sans normalisation, le poids d'une cordonnée est positivement corrélé par la coordonnée elle-même). En particulier, Chojnacki & al. [126] montrent que la normalisation rend l'erreur algébrique minimisée par les algorithmes linéarisés plus proches d'erreurs géométriques telles que l'erreur de reprojection.

3.7 Estimation de l'incertitude dans les mesures 3D

Les ambiguïtés projectives d'un système d'observation non-calibré peuvent être vues comme l'incertitude systématique liée à une reconstruction 3D. Nous avons vu que cette ambiguïté est levée par le passage à une reconstruction métrique géo-référencée.

D'autre part, la propagation analytique de la variance (équation 1.10) des mesures $2D$ permet une estimation de l'écart-type μ, ou incertitude, de l'ensemble des points 3D $\mathbf{X}$ obtenus par triangulation [127][128][129][130][131][132]. Sous une hypothèse d'indépendance des erreurs de mesures de la distance focale, de la baseline et des coordonnées 3D, le développement de Taylor au premier ordre des équations 3.1 donne :

$$\begin{aligned}
\mu_X^2 &= (\tfrac{X}{B})^2\mu_B^2 + (\tfrac{B}{D})^2\mu_u^2 + (\tfrac{X}{D})^2\mu_D^2, \\
\mu_Y^2 &= (\tfrac{Y}{B})^2\mu_B^2 + (\tfrac{B}{D})^2\mu_v^2 + (\tfrac{Y}{D})^2\mu_D^2, \\
\mu_Z^2 &= (\tfrac{Z}{B})^2\mu_B^2 + (\tfrac{Z}{f})^2\mu_f^2 + (\tfrac{Z^2}{Bf})^2\mu_D^2.
\end{aligned} \tag{3.30}$$

Ce développement montre que les erreurs de profondeur croissent en quadrature avec celle-ci et sont inversement proportionnelles à la baseline spatiale B.

Les méthodes basées sur un modèle de bruit Gaussien et une projection linéaire sont cependant théoriquement biaisées, car la distribution d'un point 3D obtenu par triangulation n'est pas Gaussienne. Une formulation simplifiée de la vraisemblance $P(\mathbf{X} \mid \mathbf{x}_c, \mathbf{x}_{c'})$ permet en effet d'avoir une idée sur la forme de la distribution d'un point $\mathbf{X}$ dans l'espace à partir d'observations $2D$ dé-corrélées, tel que :

$$P(\mathbf{X} \mid \mathbf{x}_c, \mathbf{x}_{c'}) \propto P(\mathbf{x}_c, \mathbf{x}_{c'} \mid \mathbf{X}) = P(\mathbf{x}_c \mid \mathbf{X})P(\mathbf{x}_{c'} \mid \mathbf{X}). \tag{3.31}$$

En supposant que les coordonnées pixel $\mathbf{x}_c$ soient corrompues par un bruit de mesure

gaussien et sont donc supposées suivre une distribution normale de la forme :

$$P(\mathbf{x}_c \mid \mathbf{X}) = \frac{1}{\sqrt{2\pi\mu_c^2}} e^{-\frac{(\mathbf{x}_c - E(\mathbf{x}_c))^2}{2\mu_c^2}}. \tag{3.32}$$

Le résultat est reporté sur la figure 3.7. Il existe d'autres méthodes analytiques incluant

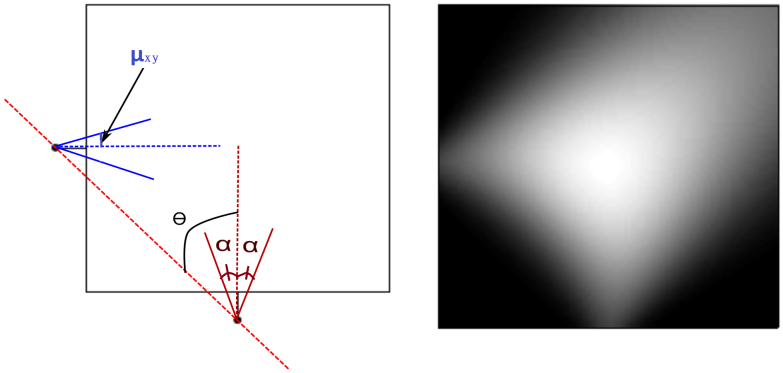


FIGURE 3.7 – Pdf d'un point 3D triangulé à partir de deux observations $2D$. Gauche : système d'observation. Droite : Fonction de densité de probabilité obtenue par inférence Bayésienne. Figure tirée et adaptée de [5]. μ représente l'incertitude ou écart-type.

l'utilisation de paradigmes de géométrie projective, tels que l'invariance des bi-rapports, pour dériver un estimation de la borne de Cramer-Rao [128]. Dans les travaux de Olague et & al [132], les auteurs dérivent un modèle empirique de l'incertitude 3D à partir du modèle sténopé et d'autres contraintes géométriques qu'ils minimisent afin d'obtenir la pose de caméra qui induit un minimum de variance.

Une évaluation numérique de l'incertitude est aussi possible en aval et permet la prise en compte d'une distribution empirique de $\mathbf{X}$. Une comparaison des incertitudes obtenues via une simulation Monte-Carlo avec ceux obtenus via une propagation de covariances [133] montre des résultats similaires. L'inconvénient de ce type de méthodes est que l'estimation de l'incertitude est effectuée pendant une phase de calibration en conditions de laboratoire qui permet de faire varier les paramètres d'acquisition. Pour des applications d'observation de milieux naturels, l'erreur type est classiquement quantifiée en utilisant les incertitudes des mesures dGPS qui ont servi au passage à une reconstruction métrique [134] et par quantification des déplacements en zones supposées fixes.

Pour finir, la dilution de la précision $2D$ via la triangulation peut aussi être estimée géométriquement à partir de mesures de l'angle α du cône de projection et de la profondeur Z (figures 3.7 et 3.8). Cette approche est utilisée par Olagues & al. [16] pour la quantification de l'incertitude sur une mesure de position triangulée à partir de capteurs de distances, par Gallup & al. [131] afin d'obtenir une reconstruction avec une résolution constante tout en changeant la baseline, par Alonzo [17] afin de quantifier l'incertitude sur la position d'un robot à vision stéréoscopique et enfin par Peng & al. [135] afin de limiter le nombre d'observations redondantes lors d'une reconstruction multi-vue. Demonceaux [136]

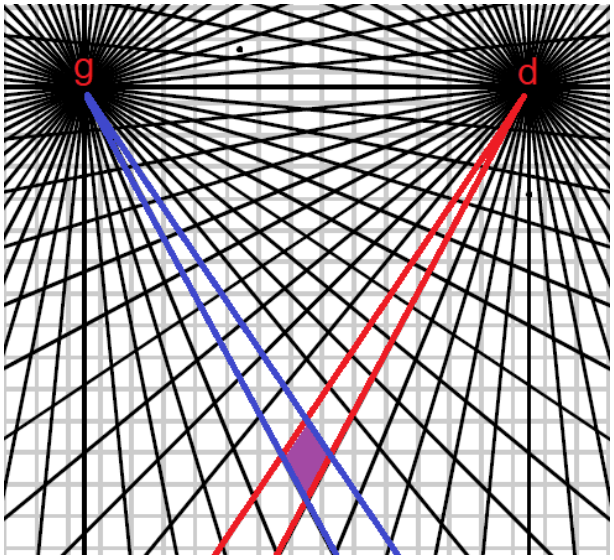


FIGURE 3.8 – Estimation géométrique de l'incertitude. Le quadrilatère violet indique la zone d'incertitude à partir d'une triangulation à partir des centres $(c, c') = (g, d)$.

a déployé une méthode similaire d'intersection géométrique de volumes projetés par deux systèmes d'observation, une caméra sphérique du type du type *Fish-eye* et un système de projection de laser circulaire, pour estimer la localisation 3D à six degrés de liberté d'un drône sur le plan image de la caméra. La modélisation via la sphère d'équivalence permet en effet d'utiliser invariablement les modèle de projection perspective avec une caméra sphérique et ainsi de procéder à une projection de cônes de révolution du système d'observation hybride et calibré caméra optique / laser afin de remonter à la trajectoire tridimensionnelle d'un objet observé.

Pour finir, l'erreur de reprojection résiduelle et une évaluation qualitative par des experts selon les connaissances a priori de la surface observée donnent aussi une indication sur la qualité d'une reconstruction 3D.

3.8 Conclusion

Dans ce chapitre, nous avons vu les fondamentaux des techniques de reconstruction stéréoscopique de surfaces et de champs de déplacements 3D. Ainsi, des primitives 2D sont estimées par des techniques similaires à celles citées pour les traitements monoscopiques (chapitre 2) et sont ensuite utilisées afin de modéliser la géométrie des caméras et ainsi reconstruire la profondeur en tous points et le nuage de points 3D qui en découle.

Cette reconstruction est idéalement raffinée via une phase de minimisation non-linéaire dont le but est de rendre l'ensemble des paramètres (intrinsèques et extrinsèques) du systèmes d'observation, les points 2D ainsi que les point 3D cohérents entre eux.

Pour remonter à une surface sur laquelle des mesures géodésiques telles que des distances et des déplacements métriques sont possibles, une phase de calibration en amont

ou en aval est nécessaire. Celle-ci permet en outre de géo-référencer le nuage de points obtenu afin, par exemple, de l'encoder en lignes de niveaux d'altitudes. Une dernière étape consiste à estimer l'incertitude liée aux résultats 3D et 3D+t ainsi obtenus. Pour cela, il convient de propager, géométriquement ou analytiquement, la variance des différentes mesures 2D et 2D+t qui ont servi à leur calcul.

Chapitre 4

Méthodes de traitements temporels et multi-temporels

Sommaire du chapitre

4.1 Introduction **65**

4.2 Caractérisation des données temporelles issues de time-lapses **66**

4.3 Méthodes classiques pour le traitement de séries temporelles **67**

4.4 Méthodes de traitement de séries multi-temporelles **68**

4.4.1 Symétrie bidate 69

4.4.2 Symétrie multitemporelle 69

4.5 Conclusion **71**

4.1 Introduction

Ce dernier chapitre de l'état de l'art est dédié à l'étude des méthodes envisagées pour le traitement de la dimension temporelle des mesures effectuées sur un time-lapse stéréoscopique de milieux naturels. Pour ce faire, nous allons tout d'abord classifier ces données selon leur répartition sur l'axe temporel et ensuite selon le formalisme utilisé afin de décrire l'évolution de la quantité suivie dans le time-lapse. Nous verrons ensuite les méthodes envisagées afin de les étudier et les traiter.

Les données multi-temporelles en question font référence à celles issues de traitements bidates 2D+t et/ou 3D+t introduits dans les chapitres précédents (chapitres 2 et 3). Les méthodes présentées dans ce chapitre peuvent alors constituer la dernière brique d'une chaîne de traitements de time-lapses à l'issue de laquelle des résultats finaux (sous forme de séries temporelles) et intermédiaires pourront être interprétés par des experts du milieu naturel observé.

4.2 Caractérisation des données temporelles issues de time-lapses

Données temporelles t Les mesures sur un time-lapse stéréoscopique peuvent donner lieu à des séries temporelles de champs de déplacements ou velocité pixelliques, métriques et géo-réferencées. Les particularités de ces données sont :

— un nombre important de séries temporelles (en tous pixels et en toute date) qui peuvent constituer des événements associés à un espace probabilisé et ainsi permettre une modélisation statistique de la donnée et du bruit,

— la possibilité d'une modélisation Lagrangienne ou Eulerienne du déplacement ou vélocité observéespour la construction des séries temporelles,

— comme évoqué précédemment, du fait que les séries temporelles de mesures bidates soient issues d'une longue suite de traitements effectués sur des données de milieux naturels, elles sont généralement entachées d'une importante quantité de bruit et nécessitent d'être régularisées avant d'être exploitables,

— une mesure de vélocité ou de déplacement étant bidate est théoriquement décomposable en une somme de vélocité ou déplacements à portée temporelle plus courte.

Données multi-temporelles $\mathrm{A}_{\mathbf{n}}^{\mathbf{2}}\mathbf{t}$ Nous appelons données multi-temporelles les mesures temporelles qui ne sont pas forcement séquentielles dans le temps. Cela comprend donc l'ensemble des mesures possibles entre toutes les combinaisons de paires dates. Ces mesures bidates possèdent la particularité de pouvoir modéliser des données temporellement symétriques. Prenons par exemple un vecteur vélocité **OF**, ou un encore flot de scène **SF**, calculé entre l'image à l'index i et l'image à l'index $i+2$ du time-lapse, noté $\mathbf{OF}_{i\,,i+2}$ (ou $\mathbf{SF}_{0,2}$), celui-ci est théoriquement symétrique par rapport au champs $\mathbf{OF}_{2,0}$ et au champs issu du calcul $\mathbf{OF}_{2,1} + \mathbf{OF}_{1,0}$. Ces champs possèdent tous trois une portée temporelle Δt égale à 2. On parlera dans la suite de fermeture temporelle une boucle temporelle dans laquelle la somme des déplacements est théoriquement nulle.

Il est ainsi possible de créer de la redondance dans les observations et intéressant d'exploiter cette contrainte de fermeture temporelle afin de discriminer la donnée consistante du bruit de mesure et des valeurs aberrantes. $\mathrm{A}_n^2 = n(n-1)$ arrangements de mesures bidates sont possibles sur des données du type time-lapse à n dates.

Formalismes Lagragien et Eulerien pour la construction des séries temporelles Comme pour le calcul du flot optique (équation 2.8), ces deux formalismes font référence à l'équation de conservation de la quantité de mouvement utilisée en mécanique des fluides afin de suivre l'évolution d'un fluide en écoulement dans le temps (vélocité, température, etc.).

Avec un formalisme Lagrangien, les séries temporelles sont construites suivant la trajectoire d'un pixel ou objet sur l'image (figure 4.1). Cette approche permet en outre de

prendre en compte les problématiques d'occlusion de l'objet observé, par exemple par inférence de la position de plusieurs objets suivis, et est particulièrement facile à mettre en oeuvre lorsque la méthode utilisée pour le calcul du déplacement bidate est éparse.

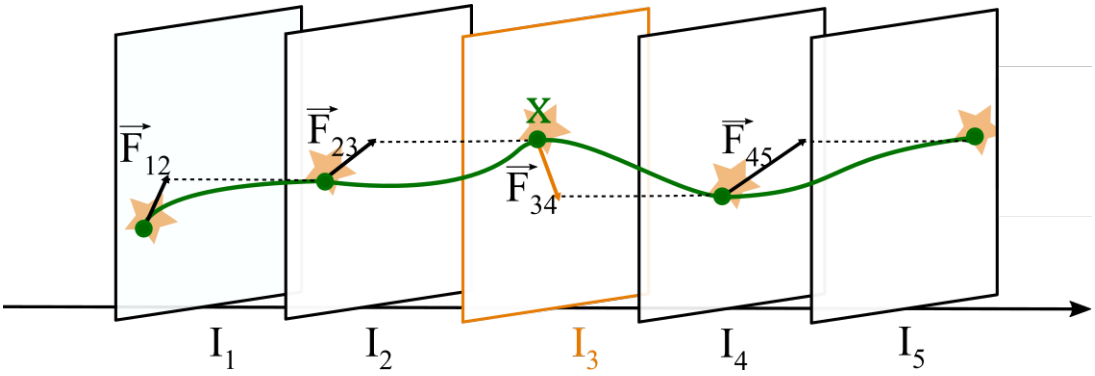


FIGURE 4.1 – Construction d'une série temporelle de suivi sur un time-lapse selon un formalisme Lagrangien.

Construire les séries temporelles selon un formalisme Eulérien revient à suivre l'évolution de la vélocité sur un pixel donnée (figure 4.2) sans prise en compte de l'advection (qui est le transport d'une quantité telle que la chaleur ou la velocité par le mouvement du milieu environnant) de la quantité de mouvement suivie. Cette approche est plus facile à mettre en œuvre lorsque les champs de vélocités sont denses et permet d'éviter les erreurs d'interpolation le long de la trajectoire.

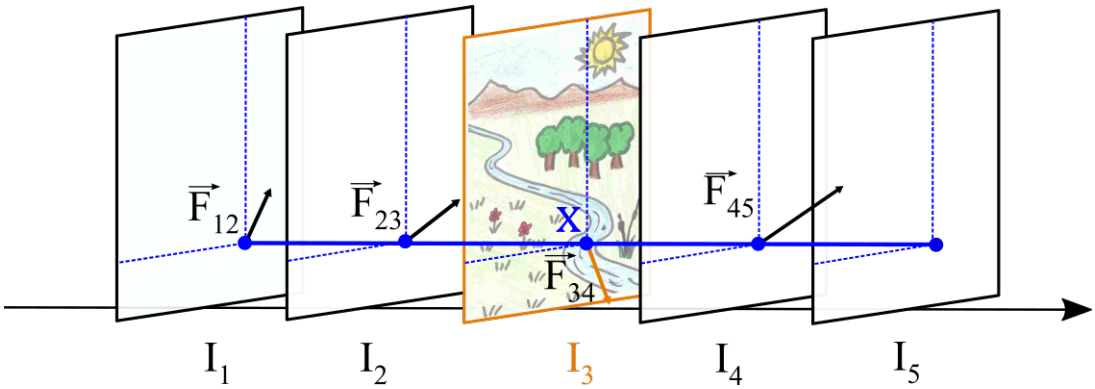


FIGURE 4.2 – Construction d'une série temporelle de suivi sur un time-lapse selon un formalisme Eulerien.

4.3 Méthodes classiques pour le traitement de séries temporelles

Le traitement de séries temporelles tourne autour des axes suivants :

— la détection de formes redondantes par auto-corrélation ou techniques symboliques [137],

— la régularisation par lissage [4], par exemple par filtrage-bas,
— la régression vers un modèle polynomiale par apprentissage,
— la décomposition spectrale avec des méthodes du type Gabor (ondelettes, etc),
— l'étude des composantes principales de l'auto-covariance par exemple par décomposition SVD,
— la modélisation ARMA et ses variantes où le signal temporel est décomposé en une tendance (espérance mobile), une ou des saisonnalités et une troisième composante non-stationnaire qui permet d'en caractériser la variance,
— le filtrage de Kalman, estimateur récursif basé sur la construction d'un historique d'observations, et ses variantes,
— l'utilisation de techniques telles que la collocation ou le krigeage spécialement adaptées aux données spatio-temporelles, techniques notamment utilisées dans des travaux similaires [3],
— etc.

Typiquement, en présence d'un grand nombre d'événements et sous une hypothèse Gaussienne, l'étude des moyennes, variances et co-variances permet, respectivement, une identification de l'évolution moyenne, de la dispersion et des dépendances linéaires dans l'espace et le temps. Une étude complémentaire des composantes principales de l'auto-covariance permet d'identifier le nombre de composantes informatives qui expliquent le plus de variances et réduire une série temporelle en ses composantes principales est aussi un moyen de la débruiter par la reconstruction d'un signal plus informatif [138][139][140].

Des techniques d'apprentissage ont également été utilisées pour apprendre le comportement attendu de séries temporelles afin de les régulariser [141][142][143][144] et il a été montré que les réseaux convolutifs entièrement connectés sont plus adaptés aux données temporelles car permettent de procéder d'une manière sémantique [145].

4.4 Méthodes de traitement de séries multi-temporelles

En présence de bruit et de valeurs aberrantes, des erreurs de symétrie, ou de fermeture temporelle, peuvent avoir lieu sur les données multi-temporelles issues de time-lapses. Ce type d'erreurs peut être minimisé afin de régulariser les séries temporelles séquentielles résultantes.

4.4.1 Symétrie bidate

Soit un champ de vélocités **F** qui peut être issu d'un calcul de flot optique (**OF**) ou de flot de scène (**SF**). Une contrainte de symétrie peut être décrite comme suit :

$$\begin{gathered}\phi : I_i, I_j \longmapsto \mathbf{F}, \\ \phi = |argmin\ \varrho(I_i, I_j)| \quad \Longleftrightarrow \quad \phi^{-1} = |argmin\ \varrho(I_j, I_i)|,\end{gathered} \tag{4.1}$$

où Φ^{-1} est l'application réciproques de Φ. En d'autres termes, la contrainte de symétrie temporelle stipule que si Φ est une mesure de champs de vélocité ou de déplacement issue de la minimisation de la différence d'une image I_i avec une image I_j, alors il existe une application réciproque Φ^{-1} qui fournit un champ équivalent (de même amplitude) par minimisation de la différence de I_j avec I_i.

Les travaux de Proesmans & al. [146] sont les premiers à construire une telle contrainte afin d'adresser les problèmes de discontinuités sur les champs flot optique. Une contrainte de symétrie est alors rajoutée à l'énergie variationnelle du flot optique. Trouvé & al [147] proposent des fonctionnelles variationnelles qui permettent une mise en correspondance qui fournit des résultats symétriques. Dans les travaux de Alvarez & al. [54], la symétrie est renforcée en aval du calcul variationnel et des problèmes d'interpolation de l'approche Lagrangienne pour la construction d'une telle contrainte sont relevés.

Dans une approche différente pour obtenir une symétrie dans les mesures de déplacements pixelliques non-rigides, les deux images impliquées dans la mesure bidate sont déformées et comparées dans un espace de référence abstrait choisi pour être entre les espaces natifs des images, dit espace intermédiaire ou "mid-space" [148][149][150]. De telles approches minimisent la fonction de coût par rapport à deux transformations qui amènent les deux images d'entrées dans cet espace intermédiaire et la composition des deux transformations fournit ensuite un déplacement symétrique. Le nombre de degrés de libertés est alors doublé et des contraintes supplémentaires sont nécessaires pour obtenir une transformation symétrique cohérente.

4.4.2 Symétrie multitemporelle

Une contrainte de fermeture, ou symétrie, temporelle peut aussi être construite non seulement entre deux paires d'images, mais le long d'une trajectoire temporelle. Dans ce sens, Volz et al [151] proposent d'adapter la formulation classique de la fonctionnelle variationnelle du flot optique par la construction d'un terme de conservation de l'intensité lumineuse dont la portée Δt est supérieure à 2 dates. Ainsi, pour une trajectoire comme celle de la figure 4.1, la contrainte de constance de données bidate (équation 2.7) en

fonction de $\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}$ et de $\mathbf{F}_{i,j}$, $i \in \{0,1,2,3,4\}$, $j = i+1$ est augmentée comme suit :

$$\begin{aligned} I_1(\mathbf{x} - \mathbf{F}_{23} - \mathbf{F}_{12}) &= I_2(\mathbf{x} - \mathbf{F}_{23}), \\ I_2(\mathbf{x} - \mathbf{F}_{23}) &= I_3(\mathbf{x}), \\ I_3(\mathbf{x}) &= I_4(\mathbf{x} + \mathbf{F}_{34}), \\ I_4(\mathbf{x} + \mathbf{F}_{34}) &= I_5(\mathbf{x} + \mathbf{F}_{34} + \mathbf{F}_{45}). \end{aligned} \tag{4.2}$$

Les équations 4.2 sont utilisées conjointement, avec une pondération inversement proportionnelle à la portée temporelle, afin de construire une fonction de coût dont la norme $L = 1$ est minimisée pour retrouver simultanément tous les champs flots optiques concernés.

— débruiter et régulariser les observations temporelles séquentielles,

— régulariser l'échantillonnage temporel, c.à.d la reconstruction de données sur les dates manquantes.

Soient A_n^2 mesures de déplacement bidates calculées à partir d'un time-lapse à n date, les contraintes de fermeture temporelle du type SBAS peuvent être divisées en deux catégories :

— les contraintes du type "*Leap-Frog*"(LF) ou de proche en proche :

$$\begin{gathered} \forall\, a \in\ [1,..,n]\,, \forall\, \Delta t \in [1,..,n-a]\ \forall\, \Delta t' \in\ [1,..,a-1] \\ -\mathbf{F}_{a,a+\Delta t} + \sum_{i=a}^{a+\Delta t-1} \mathbf{F}_{i,i+1} = \epsilon_{\mathrm{LF}}, \\ \mathbf{F}_{a,a-\Delta t'} + \sum_{i=a-1}^{1} -\mathbf{F}_{i,i+1} = \epsilon_{\mathrm{LF}}, \end{gathered} \tag{4.3}$$

— les contraintes du type "*Common-Master*" (CM) ou à image maîtresse commune :

$$\begin{gathered} \forall\, a \in\ [1,..,n]\,, \forall i \in [a,..,n]\,, \forall \Delta t \in [1..n-i] \\ \mathbf{F}_{i,i+\Delta t} + \mathbf{F}_{a,i} - \mathbf{F}_{a,i+\Delta t} = \epsilon_{\mathrm{CM}}. \end{gathered} \tag{4.4}$$

Il en résulte des applications linéaires ayant pour entrée l'ensemble des observations multi-temporelles et pour sortie une série temporelle de vélocité/déplacement $\mathbf{F}_{i,i+1}$ pour les contraintes du type LF, ou de position $\mathbf{F}_{a,j} = \int_a^j \mathbf{F}_{i,i+1}$ (cumul de déplacements) par rapport à la date a, pour les contraintes de type CM. Une bijection existe entre les deux types de sorties. Les problèmes d'inversion ainsi définis sont convexes et sont résolus par SVD sous une hypothèse de bruit de mesure Gaussien.

Les méthodes basées sur la technique SBAS proposées dans la littérature se différencient les unes des autres par les étapes qui servent à calculer les observations bidates à régulariser. Par rapport à l'approche classique où une inversion par SVD génère une série

temporelle $\mathbf{F}_{a,j} = \int_a^j \mathbf{F}_{i,i+1}$ régularisée, on trouve les différences suivantes :

— l'utilisation d'une stratégie *Leap-Frog* [152][153],
— l'utilisation d'une norme plus adaptée aux données telle qu'une norme $L = 1$ [153][154],
— l'addition d'une contrainte supplémentaire sur l'évolution de la série temporelle [140],
— l'addition d'un terme de régularisation du type Tikhonov (introduite en section 1.2.2) pour contraindre la norme de la solution [155],
— l'inversion simultanée des différentes composantes du déplacement tridimensionnel pour une reconstruction d'un flot de scène directement régularisé [155].

Plus de détails sur la méthode SBAS et ses variantes peuvent être retrouvés dans les travaux de Casu & al. [156] et de Doin et al. [157]. Dans le tableau 4.1 sont recensées certaines méthodes où la symétrie temporelle a été maximisée.

	plus que 2 dates	Formalisme	pendant/après le calcul de la velocité
Proesmans & al.[146]	✗	Lagrangien	pendant
Trouvé & al.[147]	✗	Lagrangien	pendant
Alvarez & al.[54]	✗	Lagrangien	après
Les méthodes mid-space [150]	✗	Lagrangien	pendant
Volz et al[151]	✓	Lagrangien	pendant
La méthode SBAS [156]	✓	Eulérien	après

TABLE 4.1 – Méthodes de maximisation de la symétrie temporelle.

4.5 Conclusion

Dans ce chapitre, nous avons vu certaines approches de l'état de l'art qui permettent de traiter des données temporelles telles qu'on peut en construire à partir de time-lapses stéréoscopiques de milieux naturels. Ces données seront alors issues d'une suite de mesures bruitées sur l'ensemble des données 2D+t puis 3D+t contenant de l'information pertinente. Leur régularisation est cruciale afin de fournir des données exploitables par les experts du milieu naturel observé. Outre les techniques en séries temporelles classiques, une multi-temporalité sus-définie permet d'exploiter une contrainte de symétrie sur les observations bidates à portée temporelle égale. Cela permet de régulariser les données séquentielles par minimisation de la fermeture temporelle impliquant 2 (aller-retour entre une date i et une date j) ou plusieurs (boucle ou réseau entre les deux dates i et j) mesures de déplacement ou de velocité. La méthode de minimisation de la fermeture temporelle sur un réseau d'observations bidates sera revue dans la deuxième partie de ce manuscrit où nous proposerons un algorithme spécifiquement adapté aux données time-lapses pour l'observation d'un mouvement dense et non-rigide tel que celui d'un glacier.

Conclusion de première partie

Dans cette première partie du manuscrit, nous avons introduit, d'une manière incrémentale, les outils fondamentaux pour l'étude des aspects monoscopique, stéréoscopique puis temporel qui nous a permis de concevoir une chaîne de traitements automatiques de time-lapses stéréoscopiques de milieux naturels.

Nous avons tout d'abord introduit dans le chapitre 1 certains modèles et notions mathématiques qui permettent d'aborder la problématique de reconstruction de données exploitables à partir d'une ou d'une série de mesures sur des données bruitées. Cela comprends les théories qui permettent l'ajustement d'observations numériques à un modèle théorique ainsi que les techniques d'estimation et de minimisation du bruit et des observations aberrantes.

Dans le chapitre 2, nous avons commencé par caractériser les données monoscopiques qu'on peut avoir sur un time-lapse et soulevé les différentes problématiques liées à leur traitement, à savoir : les perturbations météorologiques et atmosphériques qui réduisent la quantité de données pertinente, les variations radiométriques très présentes en milieux naturels et la difficulté liée à l'estimation d'un mouvement non-rigide tel qu'un glissement de terrain ou encore un mouvement fluidique. Plusieurs taximonies ont été proposées afin d'exposer les méthodes étudiées selon la tache à entreprendre. Les méthodes de prétraitements basées sur un modèle physique / sur la texture, les méthodes bidates denses / éparses, etc. Ces briques 2D et 2D+t consistent alors principalement à éliminer, par classification, les images où la visibilité est réduite, à éventuellement rehausser voire corriger les images à traiter pour enfin estimer le déplacement/velocité des pixels entre les images pertinentes avec compensation du mouvement de la caméra. A cette fin, des méthodes robustes aux variations radiométriques et à la différence d'échelle sur les zones mobiles sont de rigueur afin d'aboutir à des estimations cohérentes.

Dans le chapitre 3, nous avons introduit des notions de géométrie projective multivue qui permettent de reconstruire une surface tridimensionnelle et des déplacements métriques à partir de données bidates stéréoscopiques. Nous avons notamment abordé les aspects spécifiques aux données time-lapses stéréoscopiques de milieux naturels, à savoir : les différents types de données possibles entre deux paires stéréoscopiques, la calibration avec des mesures terrains, le choix d'un modèle numérique de terrain et l'estimation de séries temporelles de déplacement / vélocité métrique et géo-référencé.

Pour finir, nous avons consacré le chapitre 4 aux traitements en séries temporelles et multi-temporelles, dernière brique de la chaîne de traitements à concevoir. Ces données sont ainsi issues d'une série de mesures monodates et bidates, monoscopiques puis

stéréoscopiques acquises par le système d'observation installé en milieu naturel et ces traitements, du fait de la nature et de la quantité de données et celle de mesures, doivent permettre une réduction du bruit et des aberrations par exploitation d'une cohérence temporelle. Nous avons de ce fait commencé par exposer les différents types de données temporelles qu'on peut construire à partir des mesures bidates sur un time-lapse pour l'observation d'un mouvement géodésique dense ou quasi-dense et nous avons ensuite étudié différents modèles temporels et particulièrement ceux pouvant décrire des mesures de déplacement sur lesquelles une redondance théorique peut être exploitée afin de construire une erreur de cohérence temporelle en tout pixel à minimiser.

Deuxième partie :

Chaîne de traitements proposée

Dans cette deuxième partie, nous allons nous intéresser à la mise en oeuvre d'une chaîne de traitements de time-lapses stéréoscopiques. Pour ce faire, nous introduirons d'abord le site expérimental, nos motivations derrière sa surveillance et les détails liés à l'acquisition des données timel-lapses stéréoscopiques (chapitre 5). Nous présenterons ensuite nos propositions et contributions spécifiques pour les traitements monodates et bidates monoscopiques (chapitre 6), monodates et bidates stéréoscopiques (chapitre 7) et enfin temporels et multi-temporels (chapitre 8).

Chapitre 5

Site expérimental et stratégie de traitement adoptée

Sommaire du chapitre

5.1 Introduction **77**

5.2 Description du site et motivation **77**

5.2.1 Site surveillé 77

5.2.2 Motivation 79

5.3 Instrumentation du site et données acquises **84**

5.4 Stratégie globale **89**

5.5 Conclusion **91**

5.1 Introduction

Ce chapitre constitue un préambule à la deuxième partie du manuscrit. On y exposera des détails sur notre site expérimental, le site du glacier d'Argentière, la motivation derrière sa surveillance ainsi que des détails sur le système d'observation stéréoscopique instrumenté et les données time-lapses acquises. Nous présenterons également l'enchaînement de traitements proposés pour traiter les différents types de données en reprenant d'une façon plus détaillée le tableau 1. Ces différentes briques de traitement monoscopiques, stéréoscopiques et temporels feront respectivement l'objet des chapitres qui suivront.

5.2 Description du site et motivation

5.2.1 Site surveillé

La chaîne de traitements proposée dans cette thèse sera appliquée à la surveillance de glaciers alpins et en particulier le glacier d'Argentière (figure 5.1). Le glacier d'Argentière [1], situé aux coordonnées géographiques 45°55'N, 6°57'E, est un glacier alpin dit tempéré, par opposition aux glaciers froids dont la température est négative et qui collent au lit

1. http://glaciers-climat.fr/Glacier_d_Argentiere/Glacier_d_Argentiere.html

rocheux. Situé dans le massif du Mont-Blanc, il fait ∼10km de long pour une largeur de ∼550m et se situe entre 1600 m (zone de chute de séracs ou front du glacier) et 3200 m (zone d'accumulation) d'altitudes (voir le schéma de la figure 5.2).

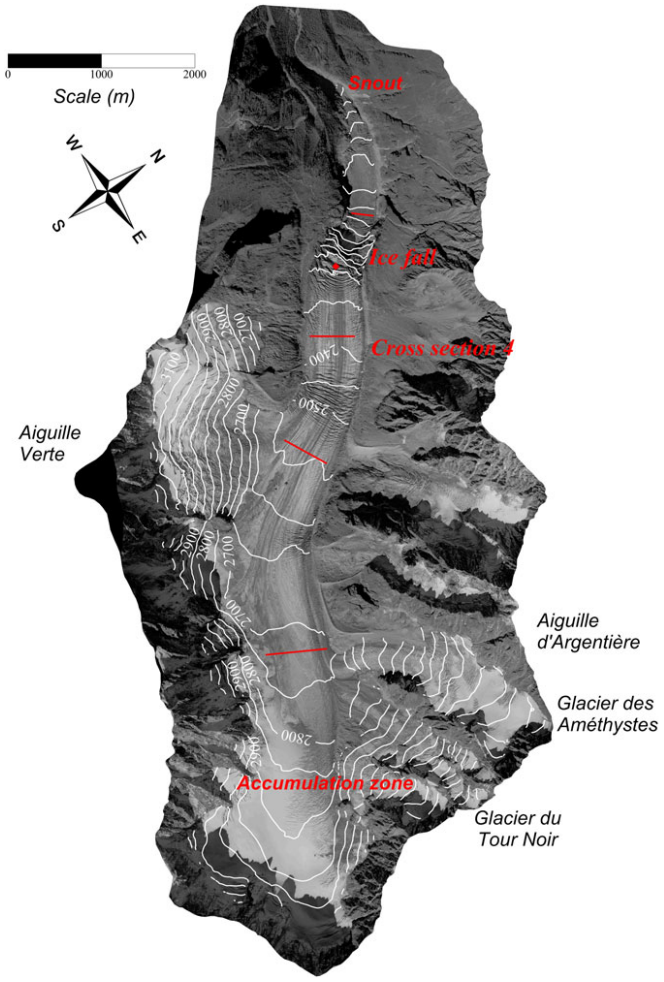


FIGURE 5.1 – Carte du glacier d'Argentière issue de photographies aériennes acquises en 2003 [158] et lignes de niveau d'altitude. Celui-ci s'écoule dans la direction Nord-Ouest.

La figure 5.2 illustre la géomorphologie d'un glacier alpin comme celui d'Argentière. La zone d'accumulation en amont alimente le glacier via de la neige accumulée et éventuellement l'écoulement d'autres glaciers. À plus basse altitude se situe la zone d'ablation soumise à des températures plus tempérées et généralement en écoulement constant. L'eau de fonte s'écoule à la fois en surface, en profondeur (entre le glacier et son lit rocheux) et parfois même au travers en forant des tunnels à travers le glacier. Celle-ci quitte le glacier en aval donnant naissance à des torrents, des rivières et des lacs. La rupture de la glace en mouvement, due à la forme du lit rocheux et aux différentes zones d'accélération ou de flexion, crée des crevasses sillonnant latéralement le glacier et déclenche des chutes de séracs au front du glacier. Il est aussi admis que la vélocité et le bilan de masse d'un glacier sont régis non seulement par la température, mais aussi par les précipitations à

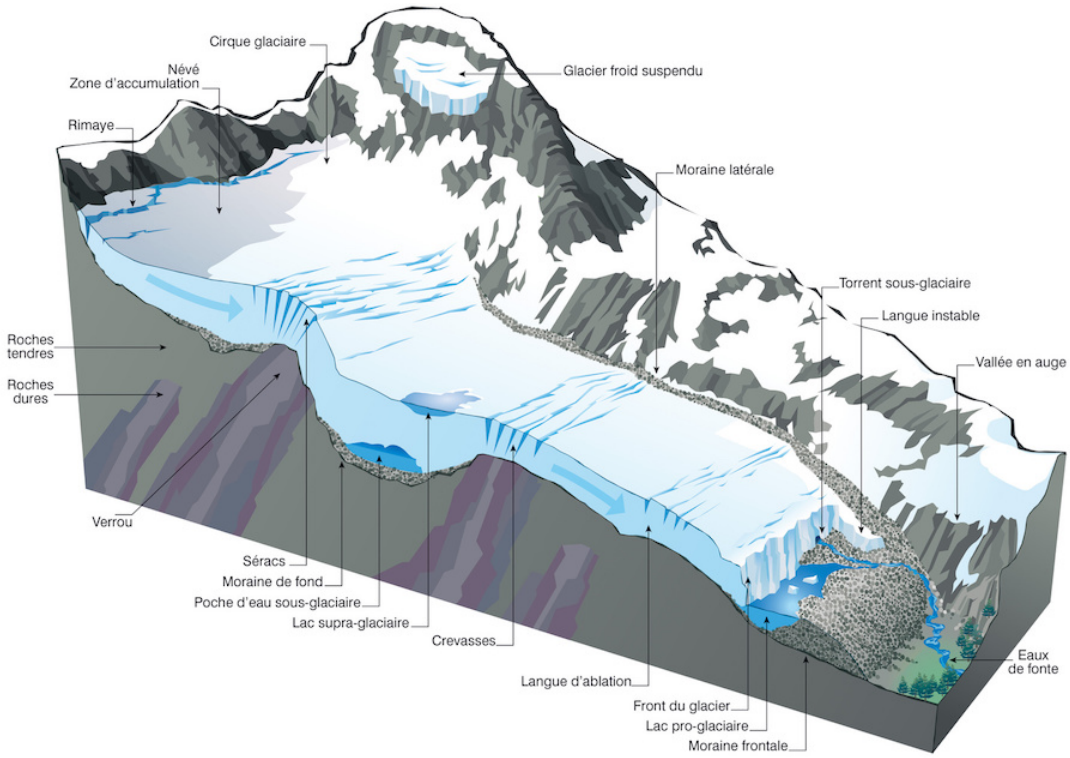


FIGURE 5.2 – Schéma d'une coupe longitudinale d'un glacier alpin. Figure publiée par le site `http://www.geoglaciaire.net`

cause de l'intensification des glissements basaux par lubrification hydraulique, ainsi que de son albédo. Tous ces facteurs font que l'écoulement glaciaire n'est généralement pas continu, mais se fait généralement par pics d'accélérations abrupts qui surgissent après un épisode climatique déclencheur [158].

L'écoulement du glacier d'Argentière a été étudié à plusieurs reprises par des techniques glaciologiques [159], par mesures GPS [160], télédétection radar [158][161][160][162], imagerie optique proximale [163] ou encore télédétection optique [1][164] (figure 5.3) afin d'effectuer une ou un ensemble de mesures métriques de vélocité bidates.

Dans le tableau 5.1 sont reportés des résultats de vélocités surfaciques journalières mesurées à environ 2400m d'altitude (cross-section 4 sur la figure 5.12), ainsi que les bilans de masses utilisés dans les travaux de Vincent et al [165] et des données d'archives météorologiques publiées par Météo France dans la station de Bourg-Saint-Maurice.

5.2.2 Motivation

La surveillance de glaciers est motivée par la prise de conscience progressive de la hausse des températures en tant qu'indicateur des effets locaux du réchauffement climatique d'origine principalement anthropique (figure 5.4).

En effet, des bilans de masse et bilans de masse surfacique de glaciers du monde entier sont majoritairement négatifs [167][165] et indiquent une accélération de la fonte des glaces

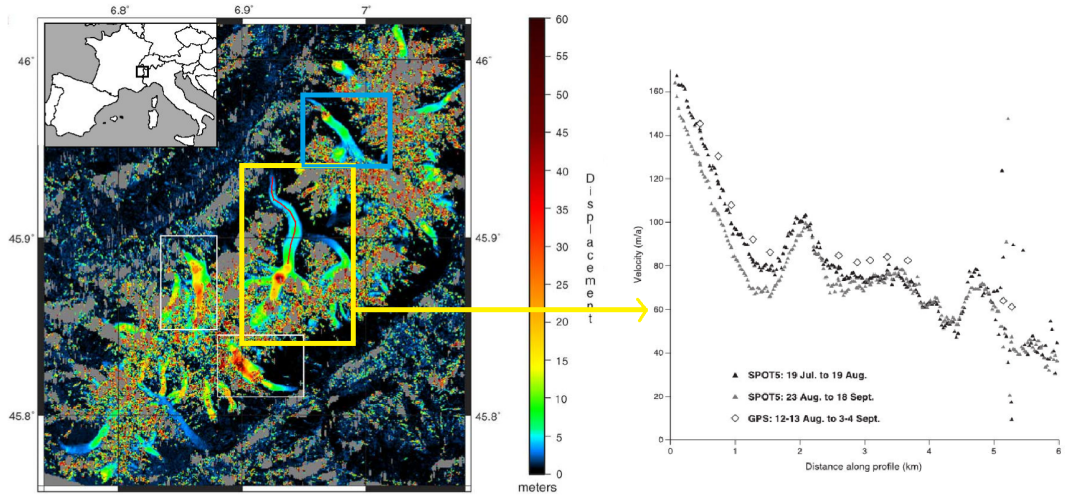


FIGURE 5.3 – Gauche : Champ de déplacement surfacique de glaciers du Mont-Blanc entre le 23 août et le 18 septembre 2003 d'après Berthier et al.[1]. Les glaciers encadrés correspondent d'Ouest à Est aux glaciers des Bossons, de Brenva, de la Mer de glace et d'Argentière (carré bleu). La ligne rouge représente le profil tracé le long du glacier de la Mer de glace (carré Jaune). Droite : Vitesses surfacique le long du profil longitudinal obtenu par corrélation d'images optiques (triangles noirs et gris) et par mesures GPS (losanges blancs). L'origine de l'axe des abscisses correspond à la partie du glacier proche des chutes de séracs du Géant dont l'écoulement est plus rapide.

	1996	2003	2007	2011	2013	2018
Bilan de masse moyenné entre les altitudes 2380 et 2600m en m.w.e.a^{-1} [165]	-3.0	-4.9	-2.6	-4.2	-3.7	-
Cumul de précipitations $mm.a^{-1}$	1225.5	713.7	1020.4	934.6	1014.7	1058.5
Températures moyennes ° C	9.7	11.2	10.4	11.4	9.9	11.9
Vélocité en surface (cm.j^{-1}) à 2400m d'altitude	24-41 [163], 30 [158][161]	23-50 [1][164] 27[158][161]	18 [158][161]	12-20(Avril), 15-25 (Août) [160]	23(G), 18(O), 15-30(SX) [166], 16 [158] ,	à calculer

TABLE 5.1 – Vélocités surfaciques journalières du glacier d'Argentière à partir de différentes sources. [163] : Mesures interférométriques (DInSAR) en mars. [158][161] : mesures GPS sur une année. [1][164] : mesures SPOT5 en août/sept. [160] : mesures GPS en avril puis en août. [166] : multiples sources ; G pour GPS en septembre, O pour imagerie Otique en septembre et SX pour mesures radar (TerraSAR-X) en octobre.

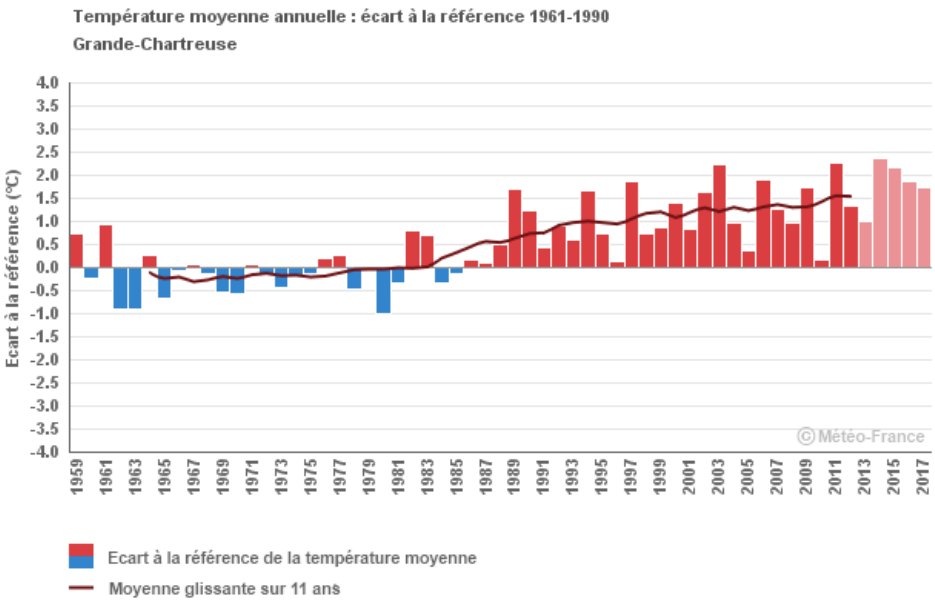


FIGURE 5.4 – Écarts de la température moyenne annuelle en France par rapport à une référence (moyenne entre l'année 1961 et 1990). Cette figure, tirée des archives de mesures de météo France, montre une évidente hausse des températures dans la région Auvergne-Rhône-Alpes.

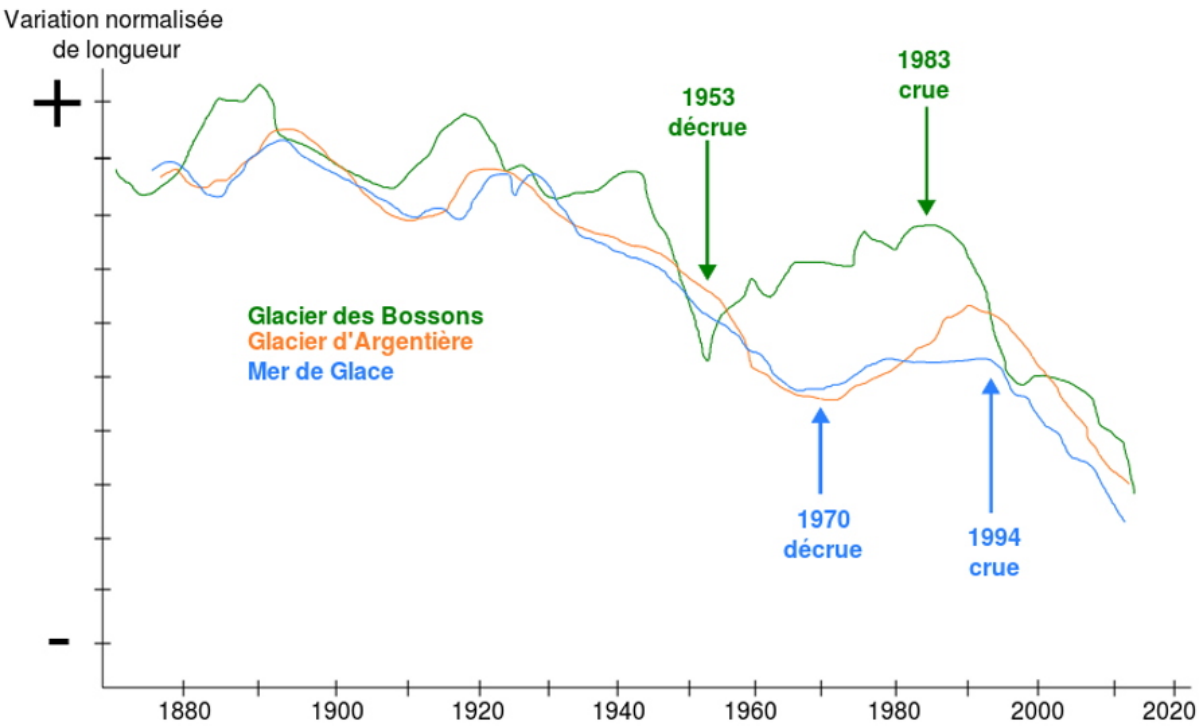


FIGURE 5.5 – Variation de longueur des glaciers des Bossons, d'Argentière et de la Mer de Glace depuis 1880. Figure publiée par Thibault Lorin.

en corrélation avec la hausse des températures ces dernières décennies. Celle-ci se traduit par une perte en masse glaciaire à la fois en épaisseur (ablation : perte d'altitude) et en longueur (recul du front).

Plus précisément, la surveillance de glaciers par des techniques d'imagerie time-lapse (optique ou radar) et de reconstruction stéréoscopique de la surface peut être déployée pour les raisons suivantes :

— Un time-lapse permet une observation continue (quotidienne ou multi-quotidienne) du glacier difficilement réalisable sur terrain (inaccessibilité, danger lié aux crevasses, etc.) tel que pour l'observation du cycle annuel des bandes de Forbes [168][169] (alternance de bandes arquées claires et sombres de périodicité annuelle), de la formation des crevasses, d'événements marquants tels que de fortes chutes de séracs ou encore pour une appréciation qualitative de l'écoulement et du recul du front glaciaire.
— Les champs pixelliques de vélocité, denses ou quasi denses, permettent, avant même de passer en unité métrique, une étude rhéologique du glacier telles que par identification des zones de flexion et d'accélération [170] ou encore de la loi de comportement, etc.
— La vélocité métrique est une variable pouvant servir au calcul de bilans de masse surfacique ou de l'épaisseur d'un glacier [164].
— Un modèle numérique de terrain issu des calculs stéréoscopiques permet une étude topographique et morphologique de la zone glaciaire, par exemple, par la construction de cartes de lignes de niveau.
— Des séries temporelles de vélocité métrique permettent une modélisation temporelle de la dynamique glaciaire et des études de corrélations par exemple avec des données d'archives météorologiques ou d'autres observations satellitaires.
— L'étude de la fonte de glaciers alpins en tant que ressources en eaux peut être réalisée pour des raisons économiques. Par exemple, le glacier d'Argentière sert à approvisionner le barrage-réservoir à vocation hydroélectrique d'Emossons via une galerie en écoulement libre de 8,55 km de long [171].
— Des prédictions sur la fonte des eaux peuvent être réalisées pour des raisons de sécurité afin de prendre les mesures adéquates lors d'événements dangereux tels que la destruction de zones environnantes par des chutes de séracs ou des inondations causées par les lacs glaciaires et les canaux pro-glaciaires [172].
— La fonte de glace peut amener à faire des découvertes paléontologiques d'empreintes et de spécimens bien préservés par la glace depuis le dernier ère glaciaire.

Nous pouvons par exemple voir sur la figure 5.6 une vue sur le front du glacier d'Argentière à partir du village du même nom situé directement en aval. On remarque un recul remarquable du front par rapport au siècle dernier.

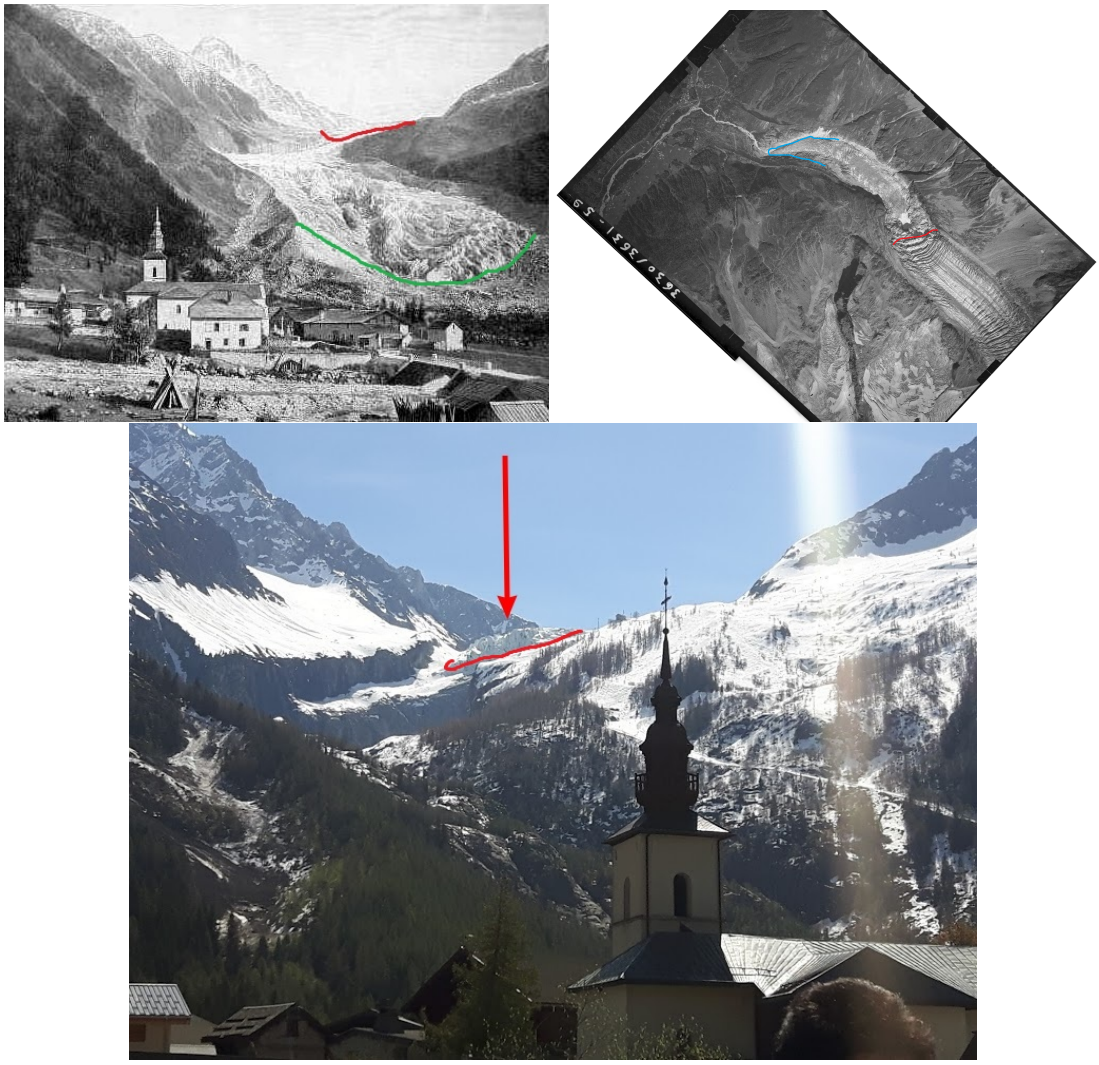


FIGURE 5.6 – Le village d'Argentière. Gauche : année 1860, photo prise par Aimé Civiale de la Société de géographie, disponible à la bibliothèque de l'ETH de Zurich. Droite : année 1949, photo tirée du géoportail de l'Institut national de l'information géographique et forestière (IGN). Bas : photo prise le 27 avril 2018. De nos jours, la partie visible du glacier s'arrête vers la flèche rouge (chute de séracs), mais il subsiste une partie dans une gorge étroite à plus basse altitude (non visible sur les time-lapses stéréoscopiques).

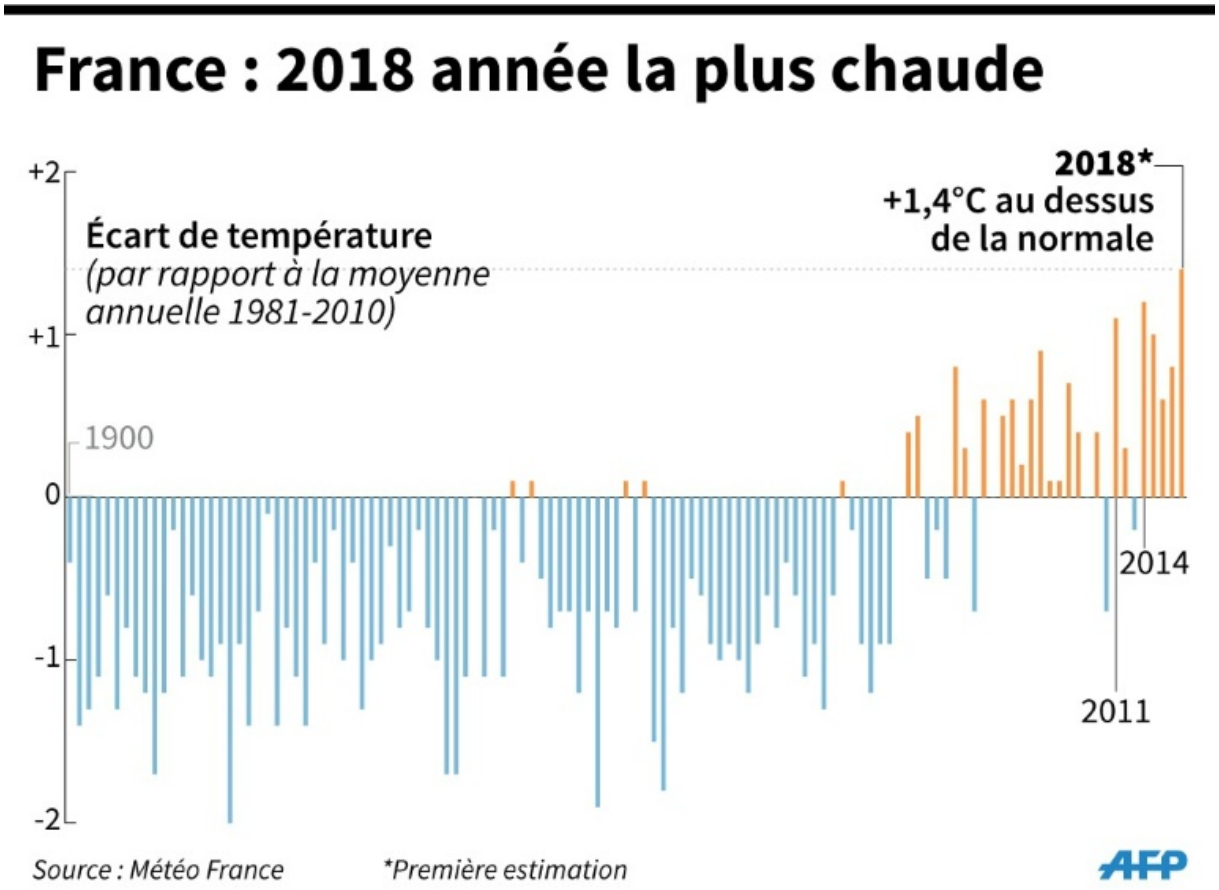


FIGURE 5.7 – 2018 "année la plus chaude en France métropolitaine ($+1,4°C$) depuis le début des mesures en 1900, devant 2014 ($+1,2°C$) et 2011 ($+1,1°C$)" selon les archives de MétéoFrance (source Agence France-Presse))

5.3 Instrumentation du site et données acquises

Les données à traiter se situent entre décembre 2017 et octobre 2018, période admise particulièrement chaude selon les archives de MétéoFrance (voir figure 5.7).

Le système d'observation stéréoscopique utilisé est composé de deux appareils photo Nikon D810 équipés d'un objectif à focale fixe qui, comme évoqué dans le chapitre 3, permet une moindre dérive des paramètres dans le temps. Ceux-ci possèdent tous deux les caractéristiques intrinsèques suivantes :

— un capteur CCD d'une surface de 35.9 × 24 mm pour des images de 7360 × 4912 pixels (36M pixels),
— une distance focale de 20 mm,
— une ouverture de f/2.8,
— une profondeur de champ infinie (focus sur le plan hyperfocal à 4.7m),
— un angle de champ de 83.7°.

Les deux appareils photo sont montés sur un support métallique (figure 5.8) et sont programmés pour prendre 6 photos par jours (à $7h$, $10h$, $13h$, $16h$, $19h$ et $22h$ avec une

dérive éventuelle de l'horloge).

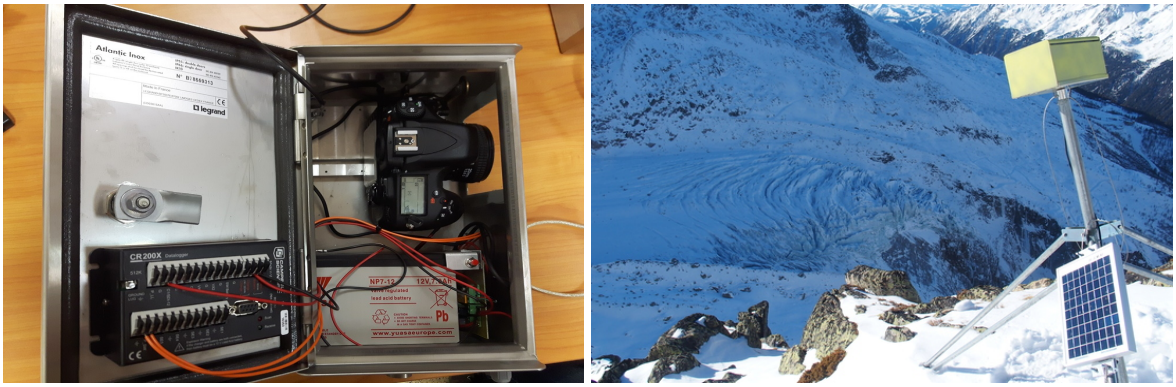


FIGURE 5.8 – Appareil-photo droit programmé, installé dans un boîtier, alimenté par un panneau photovoltaïque et monté sur une installation métallique avec une vue rasante sur le glacier d'Argentière.

La figure 5.9 représente un échantillon des données monoscopiques (appareil-photo gauche) ainsi acquises sur le site du glacier d'Argentière. Une observation qualitative de l'ensemble des images du time-lapse permet d'ores et déjà de noter :

— Une variabilité photométrique importante même entre des images prises le même jour ou à la même heure à un jour d'intervalle.

— Une couche de neige plus ou moins épaisse qui apparaît sur une ou une suite d'images. En effet, étant donné les changements brutaux de climat en haute altitude, celle-ci peut être due à un événement neigeux isolé ou annoncer le début d'un épisode neigeux plus long, par exemple, en hiver. Une couche de neige fine créera alors de nouveaux gradients sur l'image tandis qu'une couche épaisse couvrira les gradients originaux et entraînera une absence de texture.

— Une occlusion fréquente de l'objet observé par des nuages, de la brume, mais aussi par des artefacts lumineux, des gouttelettes d'eau ou encore du givre sur le boîtier de l'appareil photo. Cela doit donc être pris en compte lors de la sélection automatique de données à traiter et fait que les deux images d'un couple stéréoscopique (deux images prises en même temps par les appareils gauche et droit) n'auront pas forcement la même visibilité.

— Les ombres portées des montagnes environnantes et des objets géophysiques présents sur la scène créent des gradients spatiaux pouvant biaiser les traitements bidates (2D+t). En effet, ceux-ci vont changer de position selon l'emplacement du soleil et de ce fait selon le moment de la journée et de l'année.

— Malgré la rigidité du support, l'installation subit un mouvement connu sous le nom de "effet Tour-Eiffel" qui crée un changement de perspectives entre les différentes prises de vu.

Il est donc nécessaire de co-recaler les données entre elles, et judicieux de concevoir/employer des outils de traitements 2D et 2D+t qui reposent un minimum sur l'in-

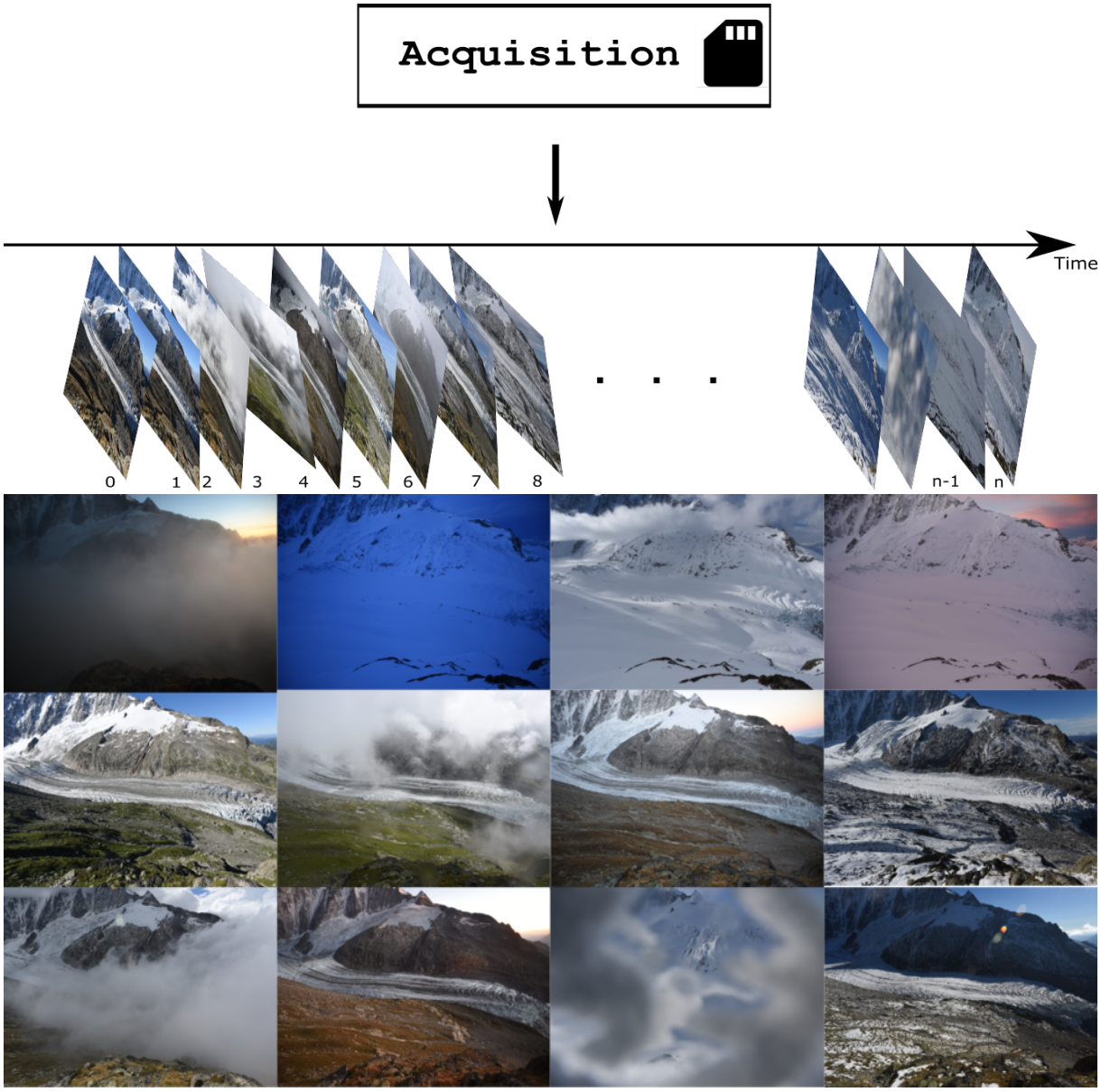


FIGURE 5.9 – Haut : ensemble d'images non recalées. Bas : exemples d'acquisitions (appareil-photo gauche) représentatifs des données traitées.

FIGURE 5.10 – Couple stéréoscopique avec une bonne visibilité.

tensité lumineuse des pixels et principalement sur les gradients statiques de la scène. Ces gradients peuvent par exemple être liés à la forme des montagnes se situant en arrière-plan (Aiguille Verte, Les Droites) ou encore aux crevasses du glacier qui possèdent l'avantage de rester marquées même en présence d'une importante quantité de neige.

D'autre part, en haute montagne la paramétrisation stéréoscopique du système d'observation, i.e. les paramètres extrinsèques, est contrainte par la morphologie du terrain. En effet, l'installation métallique de chaque appareil-photo doit se faire à l'abri des éboulements et couloirs d'avalanches, et dans une zone relativement facile d'accès lors des sorties terrain. Ainsi, les appareils photo ont été installés en hauteur avec une vue rasante sur le glacier à, respectivement, 2631 et 2684 mètres d'altitude (figures 5.12 et 5.10) tels que l'installation stéréoscopique possède les caractéristiques suivantes :

— une baseline de 154m,
— une distance minimale (profondeur) du glacier de 664m pour l'appareil photo droit et de 600m pour l'appareil photo gauche,
— une distance maximale du glacier de 3291 m pour l'appareil photo droit et de 3114 m pour l'appareil photo gauche,
— un sens d'écoulement du glacier quasi parallèle au plan image de l'appareil photo gauche afin que le déplacement métrique 3D soit plus fidèlement modélisé par le déplacement pixellique 2D.

Ces paramètres intrinsèques et extrinsèques engendrent une résolution spatiale moyenne de 10 cm sur le glacier et allant jusqu'à 50 cm sur les montagnes lointaines (figure 5.11).

Les données de calibration et de géoréférencement sont des mesures de GPS différentiel (dGPS) et ont été acquises lors de plusieurs sorties terrain en rives gauche et droite. Celles-ci sont illustrées sur le modèle numérique de terrain de la figure 5.12 et sont mesurées avec une incertitude moyenne de 0.01 m (mesure de répétabilité Gaussienne, i.e la variance sur plusieurs mesures dans les mêmes conditions).

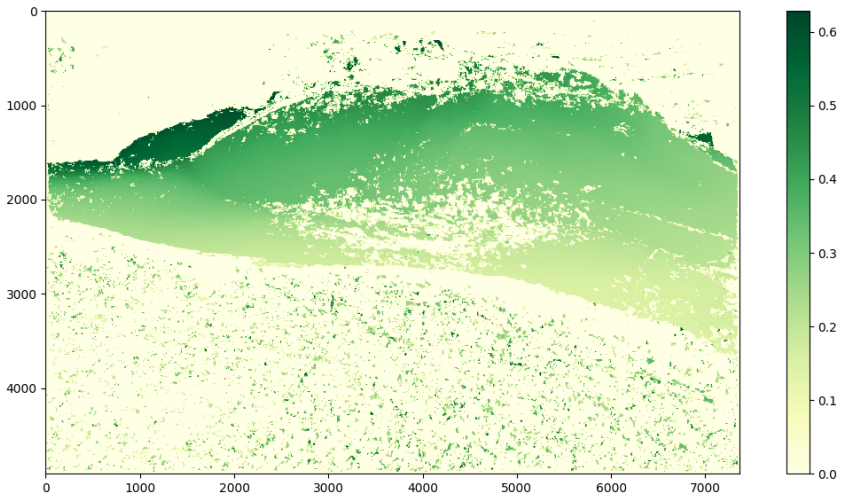


FIGURE 5.11 – Carte de résolution spatiale en mètres (point de vue de l'appareil-photo gauche).

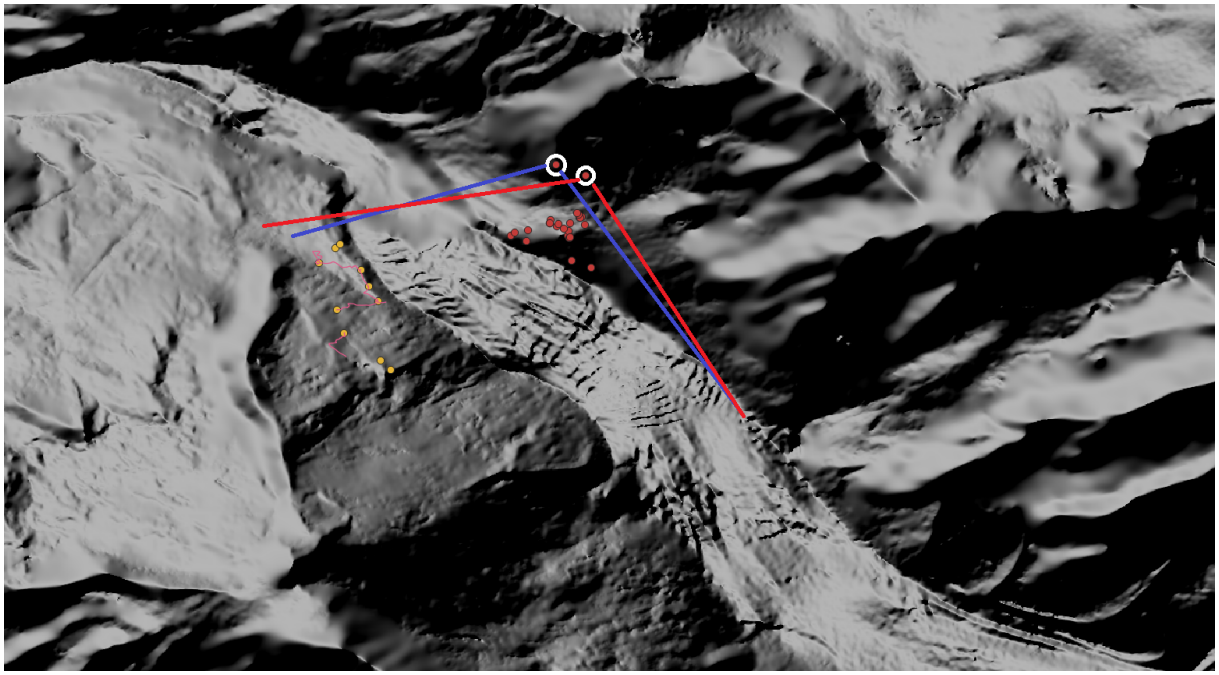


FIGURE 5.12 – Données in-situ (ronds) et système d'observation stéréoscopique vus sur un modèle numérique de terrain produit par l'Institut National de l'information géographique et forestière (IGN) avec une résolution spatiale de 5 m. On remarquera que le MNT n'est pas bien reconstruit (lisse) dans la zone de chute de séracs.

5.4 Stratégie globale

La stratégie globale proposée dans cette thèse est décrite dans le schéma de la figure 5.13 et le tableau 5.2. Après acquisition de l'ensemble des données image et ca-

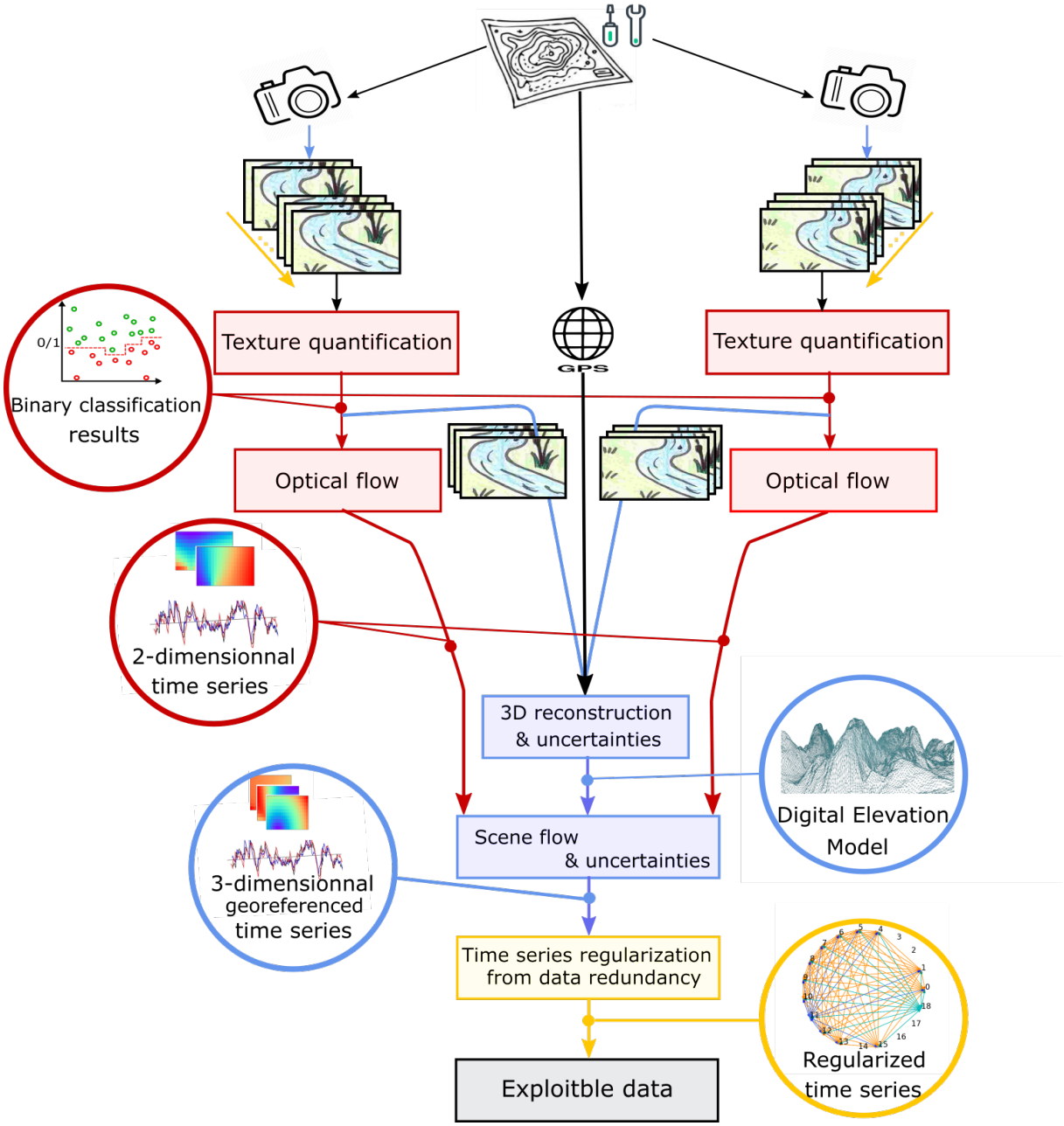


FIGURE 5.13 – Stratégie globale proposée pour le traitement de time-lapses stéréoscopiques pour l'observation des milieux naturels : mise en oeuvre pour l'observation du glacier d'Argentière. En rouge : traitements #N 2D et 2D+t. En bleu : les traitements 2×2D, 2×(2D+t) et (2×2D)+t. En jaune : les traitements en séries temporelles et multi-temporelles.

libration, les time-lapses gauche et droit sont d'abord traités indépendamment afin de quantifier la visibilité du glacier via sa texture spatiale sur l'ensemble des $\mathscr{N}$ images de chaque time-lapse et ainsi sélectionner une série temporelle de n images qui couvre le plus

Données	utilisées pour :
#$\mathscr{N}$ 2D	Quantification de la texture du glacier & classification binaire de la visibilité
2D+t	Mise en correspondance quasi-dense et & flot optique
2×2D	Reconstruction 3D de r couples d'images & calcul de l'incertitude de triangulation
(2×2D)+t	Choix d'une reconstruction 3D minimisant la dispersion
2×(2D+t)	Reconstruction 3D de champs flot optique (scene-flow) & propagation d'incertitudes
$\mathrm{A}_{\mathbf{n}}^{\mathbf{2}}\mathbf{t}$	Régularisation des séries temporelles 3D par maximisation de la cohérence temporelle entre les n images retenues
t	Séries temporelles (chronologiques) de suivi de mouvement géodésique

TABLE 5.2 – Les différents traitements proposés pour les différents types de données dans un time-lapse stéréoscopique. Tableau analogue au tableau 1. Ici $\mathscr{N}$ est le nombre de la totalité des images acquises par time-lapse, N est le nombre d'images dans une série temporelle monoscopique à pas de temps journalier et n est le nombre d'images retenues pour les mesures de suivi bidate.

de dates (N dates au total) par classification binaire : glacier visible / non-visible (Texture quantification). Le déplacement pixellique est ensuite estimé par mise en correspondance quasi-dense et flot optique variationnel (*optical flow* sur la figure 5.13).

Un sous-ensemble de r couples stéréoscopiques où les indices de visibilité préalablement calculés est le plus important est ensuite sélectionné afin de construire un modèle numérique de terrain possédant un minimum de variance. L'incertitude de triangulation est quantifiée à ce stade via une approche géométrique (voir section 3.7) que nous proposons et qui sera détaillée dans le chapitre 7. Le modèle projectif obtenu est géoreferencé avec les mesures dGPS qui servent en outre de contraintes de calibration lors du passage à une reconstruction métrique. Ce modèle numérique du glacier peut déjà servir afin d'inspecter qualitativement la zone surveillée par exemple par la construction de cartes de lignes de niveau d'altitude, latitude et longitude plus précises que celles obtenues par imagerie satellitaire (*3D reconstruction & uncertainties* sur la figure 5.13).

Il servira ensuite à calculer le flot de scène (scene-flow), ou déplacement 3D, par triangulation. L'approche adoptée permet en effet de remonter à un vecteur vélocité à 3 composantes à partir de 4 composantes modélisant les déplacements horizontaux et verticaux sur les plans images gauche et droite. Grâce au géoréférencement dGPS, ses composantes peuvent être relatives à un système de coordonnées géographiques par exemple le *East, North, Up* (ENU) beaucoup plus parlant pour les géologues qu'un repère arbitraire ou encore lié à une des deux caméras. L'incertitude en tout point du champ de vélocité 3D est estimée à ce stade par propagation de la variance des mesures pixelliques et les résul-

tats finaux sont exprimés avec une incertitude jointe de triangulation et de propagation (*Scene-flow & uncertainties* sur la figure 5.13).

Pour finir, la dernière brique de la chaîne de traitements automatiques proposée constitue notre contribution principale. À ce stade, des séries temporelles de vélocités métriques sont construites. Celles-ci souffrent d'une quantité importante de bruit et de valeurs aberrantes liés principalement à la qualité des données acquises et à l'accumulation d'erreurs de mesures. Ces séries chronologiques de vélocités métriques en tout point sont alors régularisées par exploitation de la redondance de mesures multitemporelles telles qu'on peut en construire à partir de time-lapses. En outre, cette régularisation, inspirée de la méthode SBAS introduite dans la section 4, est pondérée par les différentes heuristiques calculées tout au long de la chaîne de traitement, comprend un terme d'attache aux données, modélisant une contrainte de fermeture temporelle, et un terme de lissage (*Time series regularization from data rendundancy* sur la figure 5.13).

À l'issu de tous ces traitements, des séries temporelles géoréférencées ENU peuvent être exploitées afin de faire des corrélations, des prédictions ou une modélisation de la dynamique glacière. Cette étape ne sera pas décrite dans ce manuscrit, mais des détails supplémentaires sont disponibles dans le chapitre publié en collaboration avec le laboratoire Edytem (UMR5204) [173].

Tous les traitements effectués sur les time-lapses stéréoscopiques sont effectués sur des coordonnées image normalisées tel que décrit dans la section 3.6.3.

5.5 Conclusion

Dans ce chapitre nous avons explicité la motivation d'appliquer les méthodes étudiées et proposées pour le traitement automatique de séries temporelles de couples stéréoscopiques à l'observation du glacier d'Argentière situé dans le massif du Mont-Blanc. Le site a été instrumenté lors de plusieurs sorties terrain et les données acquises sont récupérées sur place puis traitées au sein du laboratoire LISTIC pour ensuite être transmises à des glaciologues et géomorphologues au besoin.

Des pré-traitements sur l'ensemble des $\mathscr{N}$ images acquises ont été effectués afin d'avoir une classification binaire sur les conditions météorologiques qui influent sur la visibilité (brume, neige, nuages, occlusion, etc.) sur les N dates considérées. Les traitements 2D+t servent à co-recaler les images du time-lapses entre elles et à estimer le déplacement des pixels. La visualisation et l'étude de champs de déplacements bidates permettent en outre de poser des hypothèses rhéologiques sur l'écoulement glaciaire. Les traitements 2×2D, (2×2D)+t et 2×(2D+t) permettent la construction d'un modèle numérique de terrain géoréférencé et de remonter vers des déplacements métriques. Enfin, les traitements en séries temporelles et multi-temporelles sur l'ensemble $\mathrm{A}^{\mathbf{2}}_{\mathbf{n}}\mathbf{t}$ d'arrangement de couples monoscopiques permettent de générer des résultats régularisés qui satisfont une contrainte

de fermeture temporelle. Cette contrainte, inspirée de techniques d'interférométrie différentielle radar (InSAR), a été revisitée afin de permettre l'exploitation d'observations entachées d'une grande quantité de bruit de mesure et de valeurs aberrantes et afin d'incorporer des connaissances a-priori sur la scène observée et les différentes mesures et leurs incertitudes.

Chapitre 6

Proposition de traitements monodates et bidates monoscopiques

Sommaire du chapitre

6.1 Introduction **93**

6.2 Pré-traitements **93**

6.2.1 Sélection automatique des données à traiter 94

6.2.1.1 Construction d'indices de visibilité 94

6.2.1.2 Stratégie de classification et sélection des images pertinentes 97

6.3 Estimation de la vélocité des pixels **97**

6.3.1 Stratégie de recalage 97

6.3.2 Calcul de vélocité pixellique 99

6.4 Conclusion **106**

6.5 Perspectives **107**

6.1 Introduction

Dans ce chapitre, nous exposerons les détails liés aux outils proposés pour traiter les données brutes monodates et bidates monoscopiques. Nous présenterons tout d'abord la technique implémentée afin de pré-traiter l'ensemble des $\mathscr{N}$ images dans un time-lapse et ensuite celles qui nous a permis de calculer le déplacement pixellique ou vélocité entre des paires d'images pré-traitées (données 2D+t), en l'occurrence, le co-recalage, la mise en correspondance quasi-dense et le flot optique variationnel. La problématique est alors de proposer des techniques robustes, adaptées au données à haute résolution spatiale et temporelle et qui ne nécessitent, dans la mesure du possible, aucune intervention manuelle.

6.2 Pré-traitements

Les prè-traitements constituent les traitements effectués sur les données $\#\mathscr{N}$ 2D et $\#N$ 2D, $\mathscr{N}$ étant l'ensemble des données acquises et N le nombre de dates considérées lors de la construction d'un time-lapse à pas de temps fixe, $\Delta t = 1$ jour dans notre cas.

Ceux-ci sont donc constitués de la phase de sélection de la donnée la moins bruitée à propager.

6.2.1 Sélection automatique des données à traiter

La stratégie adoptée afin de sélectionner les images impliquées dans les mesures de vélocité repose sur la construction de descripteurs qui permettent une discrimination automatique d'images dont la zone d'intérêt est trop corrompue. Comme évoqué plus tôt (chapitre 5), les causes majeures qui entraîneraient une visibilité réduite en milieu alpin sont une occlusion de l'objet géo-physique surveillé et un manque de texture lié à la présence d'une épaisse couche de neige. Pour ce faire, nous proposons deux heuristiques très simples :

— la détection des crevasses du glacier via un descripteur de texture T_G,

— la quantification des pixels blancs via un descripteur T_S.

Ces indices sont ensuite utilisés conjointement afin d'effectuer la classification binaire : "glacier visible" / "glacier non-visible".

6.2.1.1 Construction d'indices de visibilité

Les crevasses arquées observées au niveau de la zone frontale d'un glacier alpin tel que le glacier d'Argentière (figure 5.9 et 6.1) possèdent une périodicité quasi stationnaire qui, contrairement aux restes des attributs spatiaux du glacx er, restent marquées tout au long de l'année. Nous avons donc cherché à construire un détecteur de la fréquence, l'échelle et l'angle des parties supérieures et inférieures des franges.

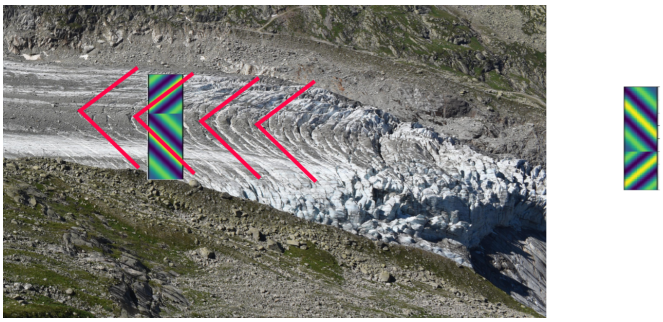


FIGURE 6.1 – Détection de fréquences spatiales présentes dans l'image par convolution avec un filtre passe-bande de type Gabor. Une convolution avec le noyau de Gabor G_{s_Δ} (représentation spatiale) équivaut à une corrélation avec le noyau superposé au glacier à droite. La convolution est effectuée dans le domaine spectral pour un accélérer de calcul.

L'indice de visibilité des crevasses T_G est la somme de la réponse obtenue en sortie d'une convolution spatiale à fenêtre glissante (équation 2.4) sur la région du glacier avec un filtre de type Gabor. Celui-ci est construit manuellement par empilage de deux noyaux de Gabor adaptés aux franges crevassées du glacier (voir figure 6.1). Pour chaque image

I de l'ensemble des données monoscopiques et pour un noyau de convolution G_{s_Δ} tel que celui de la figure 6.2, cet indice est exprimé comme suit :

$$T_G = \sum_{xy} |I(x,y) \otimes G_{s_\Delta}|. \tag{6.1}$$

On peut voir sur la figure 6.2 des exemples de résultats de cette convolution. Quand le glacier est visible, la réponse énergétique est importante (exemples du haut et du milieu) et quand celui-ci est occulté, celle-ci chute significativement.

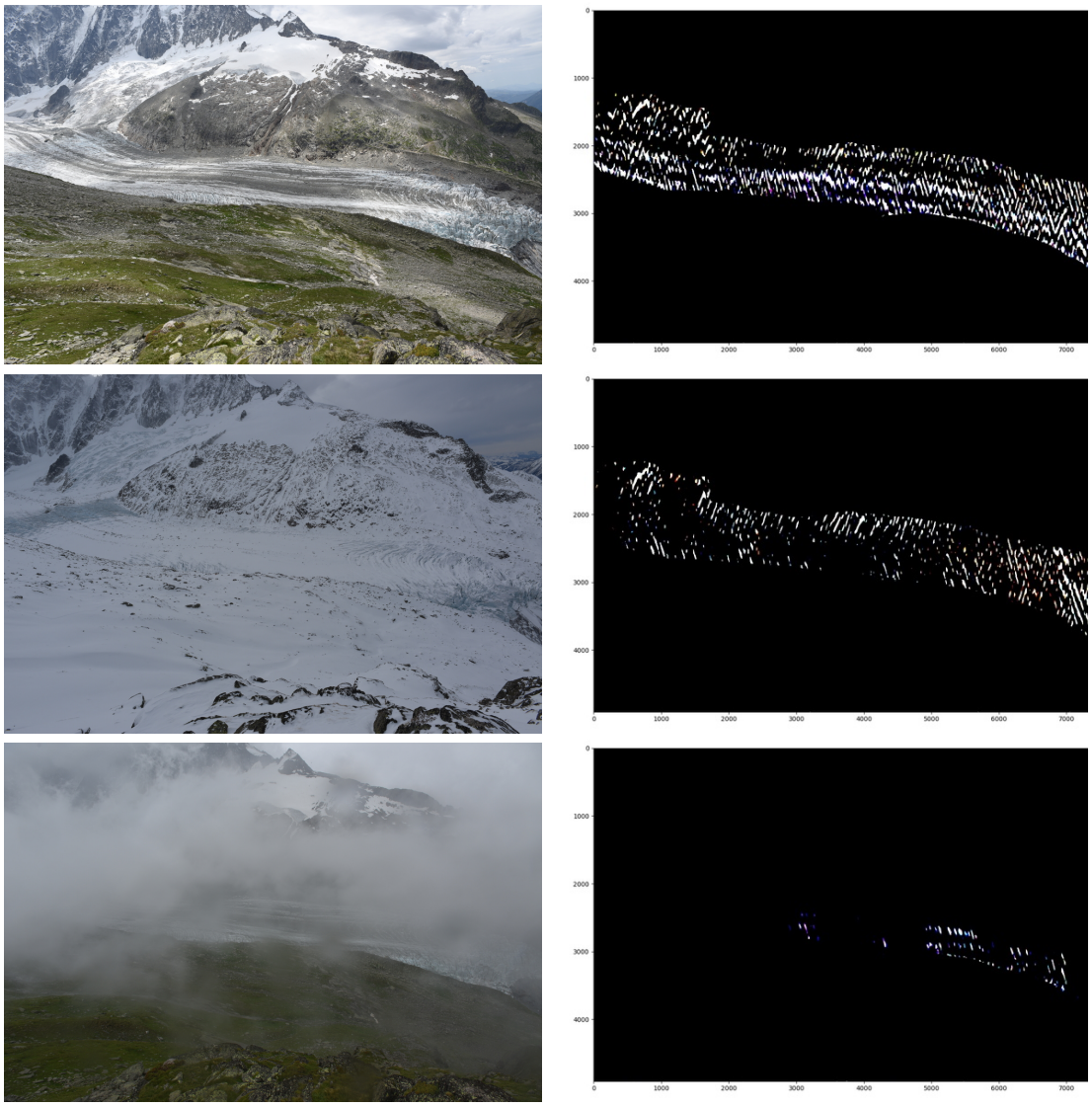


FIGURE 6.2 – Résultats de convolution spatiale avec un noyau de détection de texture sur la zone glaciaire. Les exemples du haut et du milieu ont été classifiés "crevasses visibles" et l'exemple du bas "crevasses non-visibles". Les zones fixes sont masquées.

Le deuxième indice de visibilité est construit afin de discriminer les images où les crevasses sont visibles, mais où le glacier est enfoui sous une épaisse couche de neige. Celles-ci pourraient en effet introduire un biais important dans les mesures dû à l'absence

et la variabilité des gradients. La méthode la moins coûteuse reste la quantification des pixels de couleur claire sur une zone de la scène qui n'est recouverte que lors d'épisodes neigeux assez importants pour engendrer de l'erreur. Un cadre de délimitation (bounding box) de taille 10 × 10 pixels située en rive gauche du glacier s'est expérimentalement montré suffisant afin de détecter ces images.

Les descripteurs de crevasses et de neige sont calculés sur les $\mathscr{N}$ images acquises qui, pour rappel, sont constituées de 6 images par jour. Le but de la première brique de la chaîne de traitements est de sélectionner au mieux la meilleure image par jour sinon aucune. Une série temporelle journalière d'images contenant l'image ayant le meilleur indice T_G par jour est d'abord considérée. La figure 6.3 montre les indices de visibilité T_G et T_S journaliers ainsi obtenus sur le time-lapse gauche de $N = 318$ jours. En haut de la

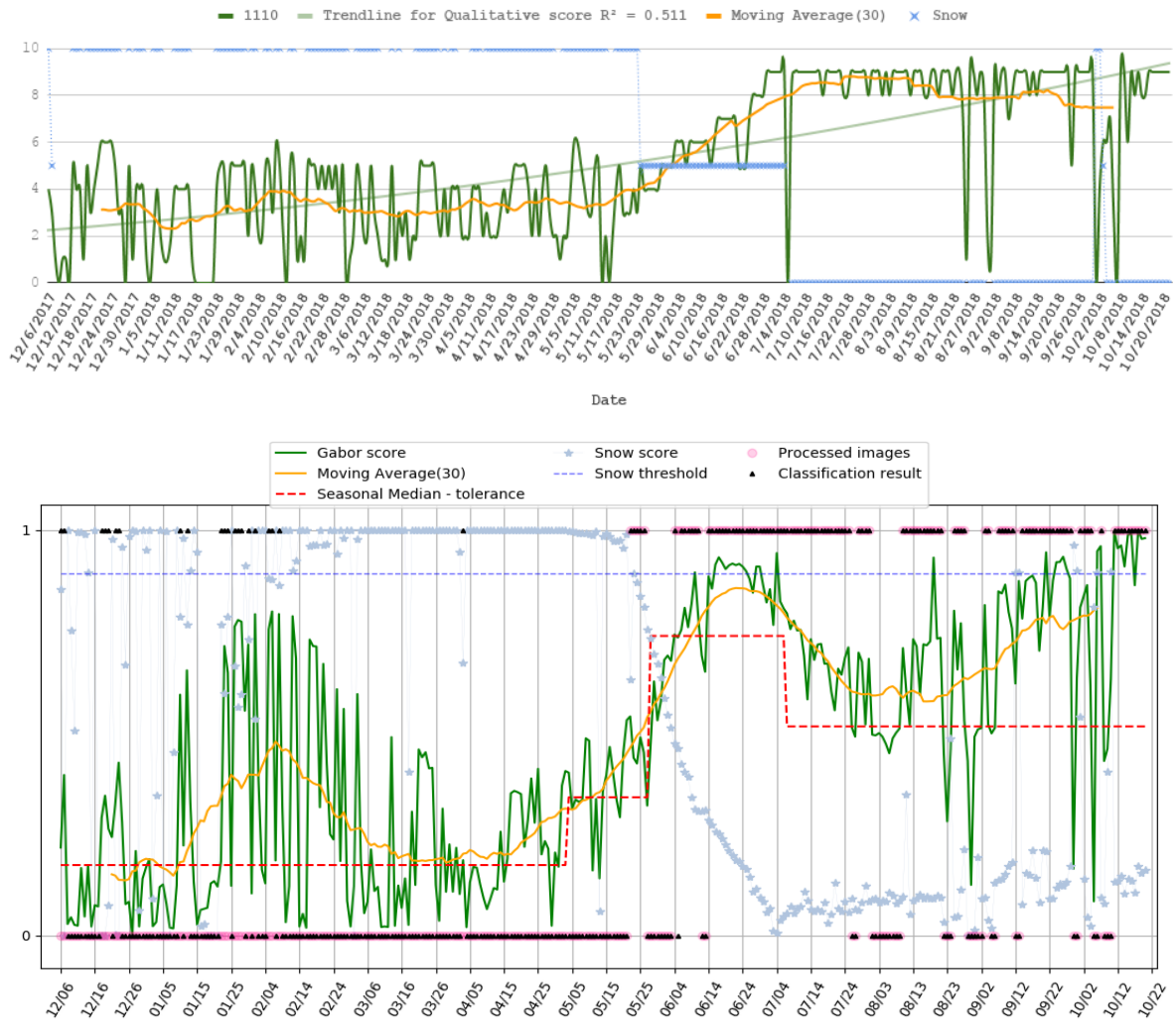


FIGURE 6.3 – Indices de visibilité calculés à partir d'un ensemble d'images #n 2D : meilleur indice sur chaque jour. En bleu : taux de neige T_S, en vert : indice de visibilité des crevasses T_G, en orange : moyenne mobile sur 30 dates de T_G, en pointillés bleus : seuil sur T_S, en pointillés rouges : seuil sur T_G. Les saisons ont été réparties comme suit : 06 décembre, 03 mai, 18 mai, 05 juillet, 22 octobre. Les résultats obtenus (triangles noirs pour les résultats de la classification et ronds rose pour les images réellement retenues) sont mentionnés avec un indice binaire.

figure on peut voir une vérité terrain construite manuellement par observation qualitative des images et qui va pouvoir valider notre méthode. Pour ce faire, à chaque image a été attribué un indice de visibilité des crevasses allant de 0 : "zéro-visibilité" à 10 : visibilité parfaite et un indice T_S de 10 si l'image n'est pas exploitable à cause de la neige, 5 si le glacier est enneigé, mais reste plus ou moins visible et de 0 en cas d'absence de problèmes liés à la neige. En bas de la figure 6.3 est illustrée la série temporelle d'indices journaliers de visibilité obtenus.

6.2.1.2 Stratégie de classification et sélection des images pertinentes

La classification consiste à appliquer deux seuils au-delà desquels les indices de visibilité seront considérés aberrants par rapport aux restes des indices. Ce seuil a été fixé empiriquement pour T_S et comme suit pour T_G :

$$\zeta_{G_s} = \text{median}(T_{G(t_s)}) - \text{tolérance}(s). \tag{6.2}$$

La tolérance mentionnée dans dans la définition de ζ_{G_s} (équation 6.2) est fixée empiriquement en fonction de la saison s puisque afin de considérer qu'un indice de visibilité T_G est aberrant, il faut que celui-ci soit comparé à des indices calculés sur des images similaires. Les images sont alors classées manuellement par aspect global : "saison des neiges" du 06 décembre au 03 mai, "saison pluvieuse" du 03 au 18 mai, "saison entre-deux" du 18 mai au 05 juillet et "aspect estival" du 05 juillet eu 22 octobre (date de récupération des données) ce qui nous donne 4 seuils différents (en rouge en bas de la figure 6.3). Les résultats de classification des données récupérées en 2018 sont reportés sur la figure 6.3. On voit alors que la plupart des images acquises pendant la saison "saison des neiges" ne seront pas propagées à cause d'un taux de neige excessif et beaucoup d'images pendant la "saison pluvieuse" ont été disqualifiées par manque de texture liée à la fréquence des franges glaciaire. Une observation qualitative de ce résultat montre des épisodes de pluie diluvienne entraînant une visibilité en effet très réduite.

6.3 Estimation de la vélocité des pixels

6.3.1 Stratégie de recalage

Afin qu'une coordonnée spatiale corresponde au même endroit physique le long du time-lapse, un co-recalage des images par rapport à une image maîtresse est nécessaire. Une homographie 3×3 modélise alors une légère différence de perspective monoscopique (voir section 1.3 du chapitre 1) sous l'hypothèse d'un mouvement panoramique. Celle-ci est estimée par corrélation sur des zones fixes (masquage du glacier). Les 8 paramètres de l'homographie entre chaque image I_i et l'image maîtresse I_m sont estimés par minimisation

du critère de corrélation-croisée (CC) proposé par Evangelidis & al. [90], introduit dans l'équation 2.20, avec une mesure corrélation croisée (CC).

La problématique est alors la suivante : compte tenu de la décorrélation des images entre elles avec le temps, i.e. qu'une image acquise en janvier sera très différente d'une image acquise en août, comment procéder pour qu'un time-lapse sur une année puisse être co-recalé automatiquement par rapport au même plan image. La décorrélation est mise en évidence sur la figure 6.4 où on voit une chute symétrique de la mesure de similarité en zones fixes (DM) avec la portée temporelle.

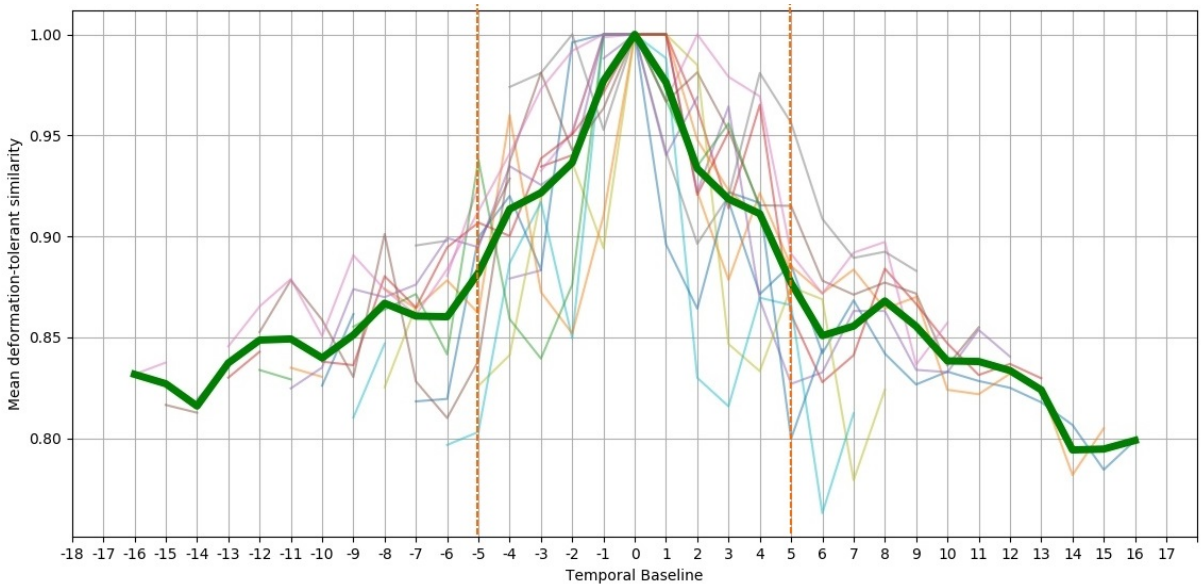


FIGURE 6.4 – Mise en évidence de la décorrélation des zones fixes des images avec la portée temporelle en jours (temporal baseline). L'indice de corrélation normalisée en ordonnée provient d'une mise en correspondance avec DeepMatching. En vert : moyenne sur toutes les images d'un jeu de données time-lapse journalier constitué 17 images. On remarque une décorrélation avec une tendance décroissante avec la portée temporelle.

Comme dépeint sur la figure 6.5, la stratégie proposée repose sur une image maîtresse glissante. Encore une fois, la classification manuelle des données #n 2D selon l'aspect global des images s'est montré utile afin de considérer des sous time-lapses ayant relativement le même aspect. L'image maîtresse initiale est alors l'image qui possède l'indice de visibilité des crevasses le plus haut sur la première saison (image 0 sur la figure 6.5).

Suite à cela, la dernière image recalée lors de la saison initiale est glissée dans la première "saison" et est considérée comme l'image maîtresse de celle-ci. Le processus est réitéré sur l'ensemble des données #n 2D jusqu'à co-recalage de l'ensemble des données. Avec cette stratégie, les différentes homographies sont estimées avec un minimum d'erreur de recalage. Contrairement à une stratégie de recalage où chaque image I_i est recalée par rapport à une image I_{i-1}, avec la définition d'une image maîtresse pour chaque saison, le co-recalage peut être effectué simultanément (task parallelism) afin d'optimiser les temps de calcul. On peut voir sur la figure 6.6 l'effet du recalage sur les mesures bidates. On

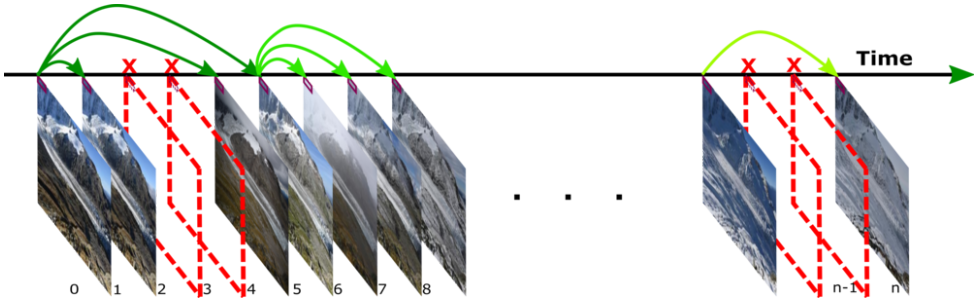


FIGURE 6.5 – Stratégie proposée pour recalage d'un time-lapse monoscopique. Celle-ci repose sur une image maîtresse par saison et permet d'éviter d'éventuels problèmes de décorrélation. L'image 0 servira pour comme image maîtresse pour le recalage des images 1, 2, 5 et 6 appartenant à la première saison. Limage 6 sera ensuite utilisée dans la saison suivante comme image maîtresse et ainsi de suite.

remarquera que l'effet du mouvement de la caméra induit un déplacement non rigide en zones fixes et qu'une erreur résiduelle minime subsiste.

6.3.2 Calcul de vélocité pixellique

La stratégie implémentée (qu'on nommera DMDF pour DeepMatching & DeepFlow) repose sur une estimation de flot optique par DeepFlow [1] (algorithme introduit en section 2.4.1.2) initialisée par une mise en correspondance quasi-dense par DeepMatching [2] (algorithme introduit en section 2.4.1.1). Une optimisation de l'usage de la mémoire est indispensable afin de traiter l'ensemble des images à haute résolution classifiées "glacier visible". Pour ce faire, nous avons opté pour :

— Une mise en correspondance quasi-dense à pleine résolution suivi d'un sous-échantillonage par floutage Gaussien 3×3 (seulement $30\% = 0.54\%\times0.54\%$ de l'image est gardé) avant la phase de calcul de flot optique et d'un sur-échantillonage des champs de vélocité obtenus. Les coordonnées homologuées par DeepMatching doivent alors être adaptées afin d'initialiser le calcul variationnel. Cette stratégie permet d'exploiter la résolution initiale afin de corréler les hautes fréquences spatiales présentes sur l'image, d'accélérer significativement le calcul variationnel tout en évitant les problèmes de saturation de la mémoire. Le sur-échantillonnage entraîne un champ de vélocité lisse en adéquation avec notre connaissance a priori sur le mouvement glaciaire.

— Une corrélation-croisée adaptative et locale lors de la phase de DeepMatching. Le rayon de recherche du patch atomique homologue est alors égal à $5\times\Delta t$, Δt étant la portée temporelle ou la différence en jours entre les deux dates considérées.

— Lors du calcul variationnel de l'équation 2.10, la contrainte de conservation de l'intensité lumineuse est moins pondérée que la contrainte de conservation des

1. https ://thoth.inrialpes.fr/src/deepflow/
2. http ://lear.inrialpes.fr/src/deepmatching/

gradients spatiaux : $\varrho_0 << \varrho_1$. Le terme de lissage ϱ_2 quant à lui est relativement faible du fait que la carte obtenue soit sur-échantillonnée en aval.

— Comme pour la phase de co-recalage, toutes les mesures bidates sont effectuées en parallèle (task parallelism) afin d'exploiter entièrement les multiples processeurs sur les machines de calcul.

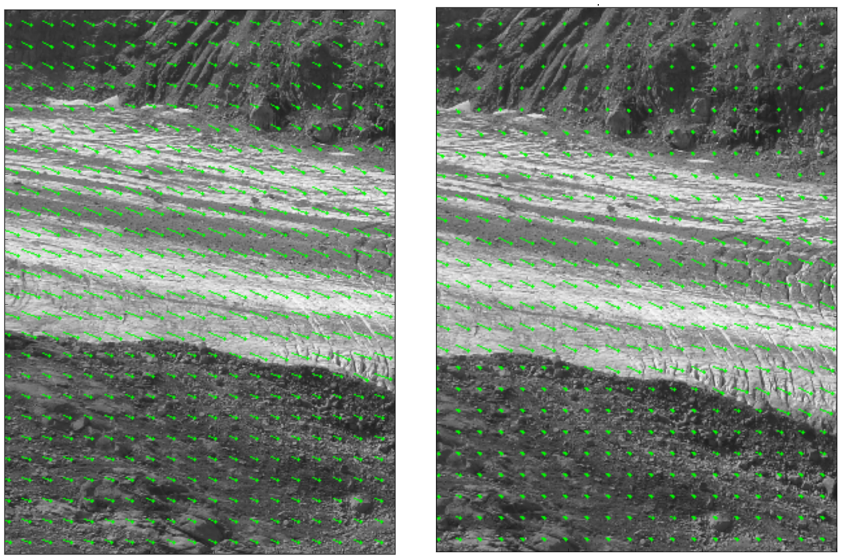
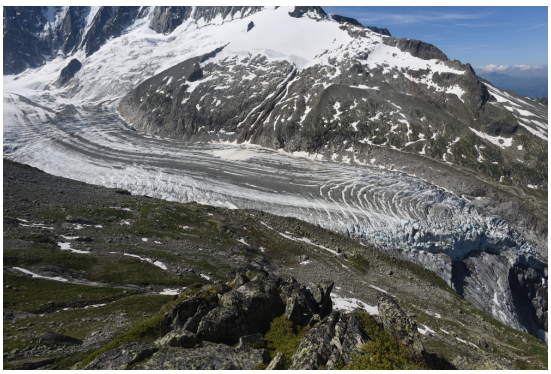


FIGURE 6.6 – Effet du recalage sur le déplacement pixellique. Gauche : sans recalage, Droite : après co-recalage des deux images concernées dans le calcul bidate avec l'image maîtresse.

Évaluation qualitative des mesures de vélocité Les figures 6.7 et 6.8 montrent un exemple de champs de vélocité pixellique calculés (figure 6.8) sur un quadruple stéréo-temporel (figure 6.7). Les mesures bidates obtenues entre le 30 août 2018 et le 02 septembre 2018 sur les données acquises par les deux appareils-photo sont qualitativement cohérentes, car le déplacement en zones fixes est quasi nul. On remarque que l'écoulement est structuré selon un champ de vitesse parabolique : la vélocité surfacique est quasi nulle aux bords du glacier et maximale à mi-hauteur (la loi de Poseuille dans sur un support ouvert). On peut également voir que la composante verticale (composante x) du vecteur vélocité est moins importante sur une partie du glacier située vers les coordonnées $(2000, 2000)$. Nous avançons que celle-ci est due à une zone relativement plate entre 2500 et 2700 mètres d'altitude (hypothèse confirmée par les mesures tridimensionnelles de la figure 5.1). Directement en aval de cette zone plate, on peut repérer une zone de flexion du glacier (voir également la figure 6.9). La vélocité pixellique du glacier suspendu des Rognons est également perçue. Son déplacement est moins important et majoritairement vertical de par le dénivelé important en cette zone.

La vélocité dans les zones du ciel et des montagnes lointaines est erronée à cause d'un déplacement occulté dans le premier cas et de la présence de brume dans le deuxième. Ceci met en évidence l'impact de la présence de brume qui altère non seulement l'intensité

(a) Photo gauche du 08 juillet 2018



(b) Photo droite du 08 juillet 2018

(c) Photo gauche du 07 août 2018

(d) Photo droite du 07 août 2018

FIGURE 6.7 – Quadruple stéréo-temporel datés du 08 juillet 2018 et du 07 août 2018.

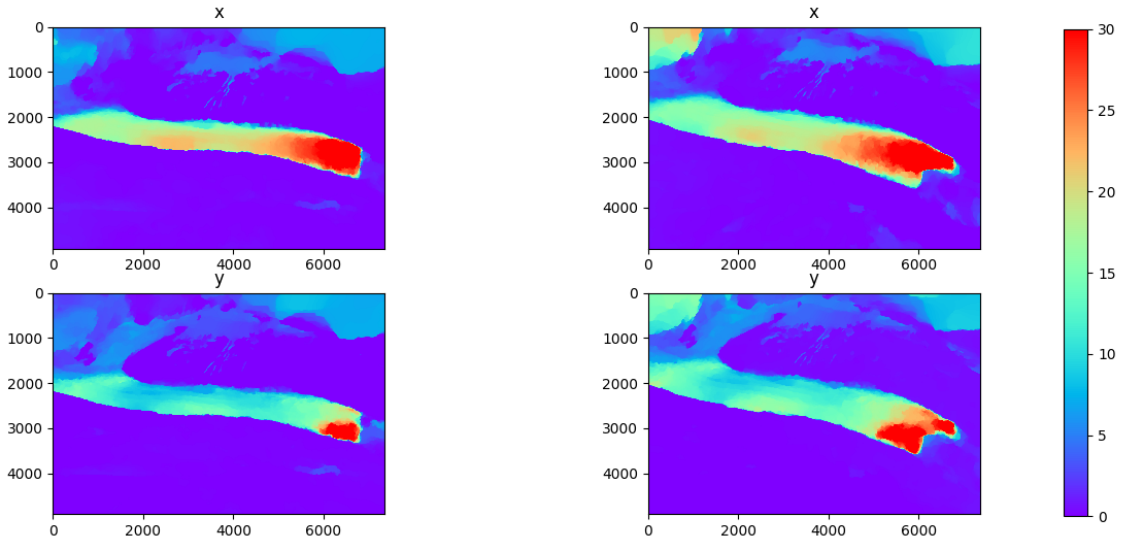


FIGURE 6.8 – Champs de vélocité pixellique dense calculés entre le 08 juillet 2018 et le 07 août 2018 : vélocité horizontale (composante x), vélocité verticale (composante y). Un épisode neigeux a eu lieu entre le deux dates.

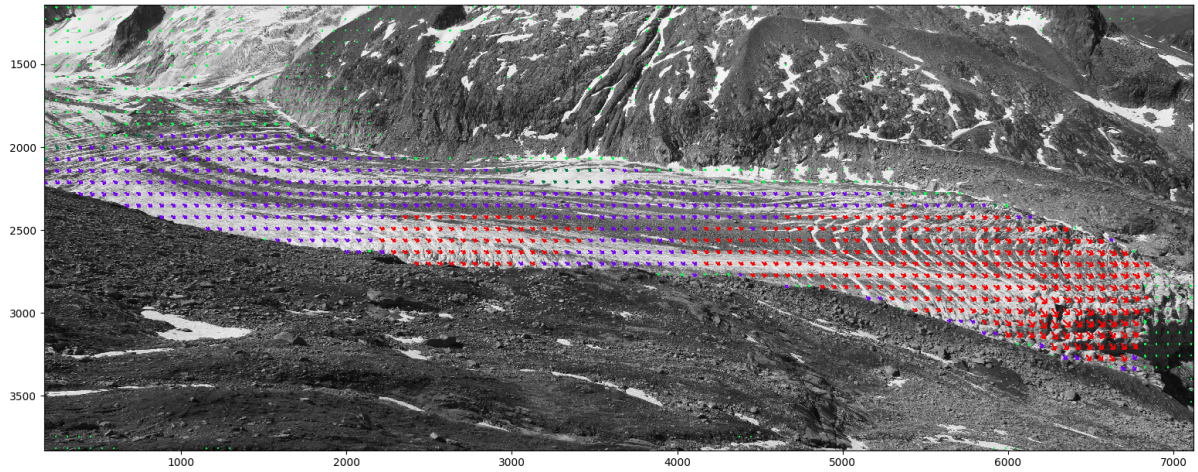


FIGURE 6.9 – Exemple de carte de vélocité pixellique permettant la mise en évidence des différentes zones d'accélération. La couleur des vecteurs est selon leurs amplitudes.

lumineuse mais aussi les gradients sur un une image. Un zoom sur le front du glacier (figure 6.10) montre la précision spatiale achevée par l'algorithme déployé. Ce champ a été mesuré sur une date où il y a une chute de séracs importante vers les coordonnées $(x, y) = (400, 300)$ qu'on peut repérer grâce à un déplacement quasi horizontal sur toute la zone en amont.

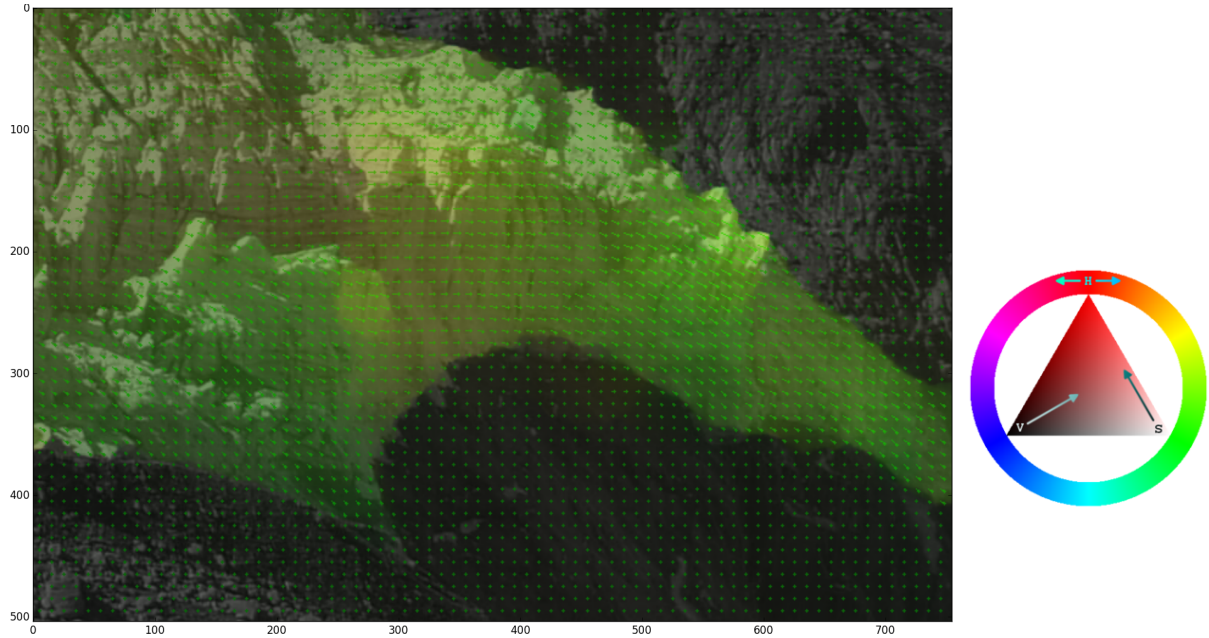


FIGURE 6.10 – Zoom sur la vélocité sur la zone de chutes de séracs avec encodage dans l'espace colorimétrique HSV (Hue = angle, Saturation = 0 %, Value = amplitude du vecteur vélocité) et flèches dont l'amplitude est doublée. Point de vue de l'appareil-photo gauche.

On voit sur l'exemple de la figure 6.11 qu'une mesure de corrélation-croisée, comme proposé dans les travaux similaires existants [1][48][4][2], souffre d'une granularité qui est moins conforme avec un déplacement glaciaire quasi laminaire. Après inspection qualita-

tive des déplacements résiduels en zones fixes, on note que cette granularité est majoritairement due aux ombres portées de rochers et autres reliefs et objets géologiques présents sur la scène. Qui plus est, la CC s'est montrée plus prône aux erreurs dues aux variabilités photométriques et climatiques, est théoriquement inadaptée car non invariable à l'échelle et ne peux donc pas capturer le déplacement vers la caméra et enfin est plus coûteuse en ressources informatiques.

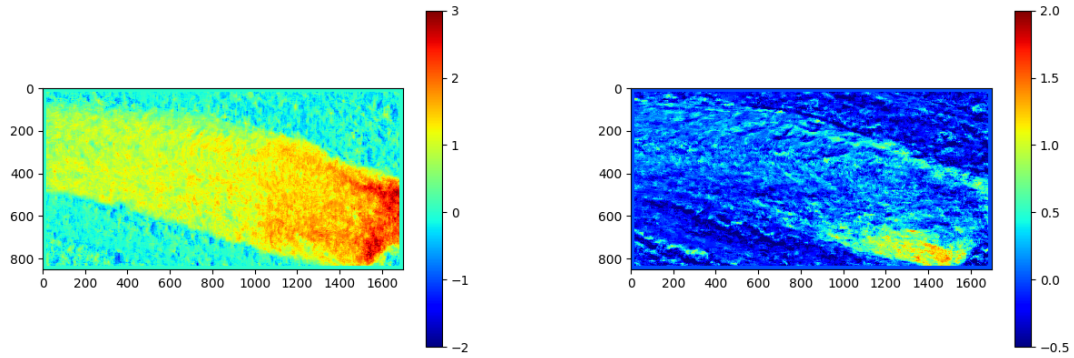


FIGURE 6.11 – Carte de déplacement typique obtenue par corrélation croisée centrée normalisée (CC) (résultat obtenu avec couple ayant de bons indices de visibilité). Gauche : déplacement horizontal. Droite : déplacement vertical.

D'autre part, tandis que les résultats obtenus par la méthode DMDF suggèrent un ratio signal sur bruit relativement faible sur les couples spatio-temporels disposant d'une bonne visibilité, une inspection de la série de résultats sur une grande portée temporelle montre, comme sur l'exemple de la figure 6.12, une quantité importante de bruit dans les champs de vélocité obtenus et ce particulièrement en considérant un pas de temps unique (par exemple journalier). Ceci est dû à la présence de données très hétérogènes, en termes d'apparence climatique et photométrique, malgré l'adaptation de la chaîne de traitement, sur un si longue portée temporelle (voir figure 6.13).

Évaluation quantitative des mesures de vélocité L'erreur probable sur la zone du glacier a été estimée par la construction d'un time-lapse où on dispose de la vraie série temporelle de cartes de vélocité. Pour ce faire, nous avons d'abord extrait une sous-imagette maîtresse sur la zone du glacier à pleine résolution (1100 × 1500 pixels) que nous avons ré-échantillonnée à l'aide de paramètres variables dans l'espace et le temps. Afin de modéliser au mieux un mouvement non-rigide tel qu'on peut l'observer sur le glacier, nous avons opté pour quatre modèles de transformation géométriques différents (parmi les modèles paramétriques de la section 1.3) :

— un modèle de transformation isométrique sous-pixellique : $M1$,
— un modèle de transformation affine (globale) légère : $M2$,
— un modèle de transformation affine (globale) importante : $M3$,

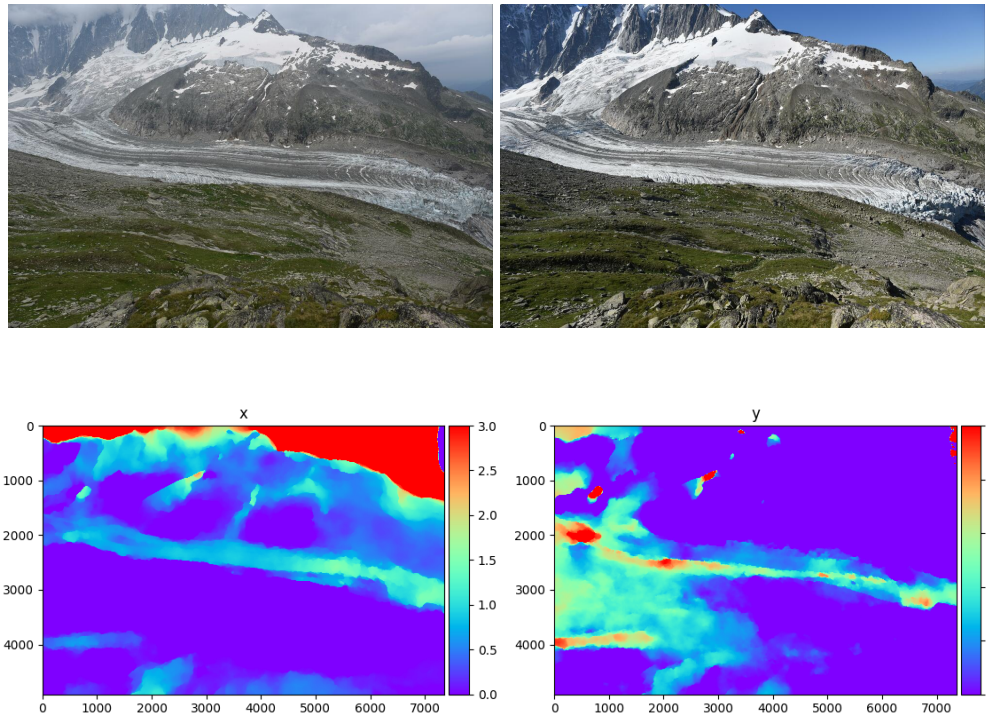


FIGURE 6.12 – Champs de vélocité pixellique dense calculés entre le 28 et le 29 juillet 2018. Gauche : vélocité horizontale, droite : vélocité verticale. Cet exemple illustre la sensibilité des calculs de vélocité pixellique à la présence de bruit atmosphérique.

— un modèle de transformation affine par morceaux : $M4$.

Quant aux zones statiques, celles-ci ont été simulées en prenant différentes zones statiques issues des données réelles.

Ainsi nous obtenons un ensemble de time-lapses de tests dont la difficulté de reconstruction de la vélocité est croissante. Le modèle $M4$, affine par morceau (à gauche sur la figure 6.14), modélise alors grossièrement le mouvement d'un fluide semi-rigide soumis à des compressions et des dilatations locales. Nous évaluons les séries temporelles de vélocité selon l'erreur quadratique suivante :

$$\epsilon_{sim} = \sum_{\Omega} \sum_{t=0}^{5} \frac{1}{N_{\Omega}} ||\tilde{\mathbf{F}}_{t,t+1} - \mathbf{F}_{t,t+1}||_2^2, \tag{6.3}$$

avec Ω le domaine spatiale constitué de N_{Ω} pixels de la zone glaciaire (patch central de taille 500×1000) des sous-imagettes et t leur domaine temporel.

L'évaluation de la méthode de calcul de la vélocité pixellique proposée a donné les résultats reportés sur le tableau 6.1. Nous voyons que l'erreur obtenue et de l'ordre du centième de pixel, sauf pour le jeu de données obtenu avec le modèle $M3$. Celui-ci dispose en effet de la déformation non-rigide la plus importante et nous remarquons qu'une corrélation croisée échoue particulièrement à retrouver un tel mouvement. La mesure de

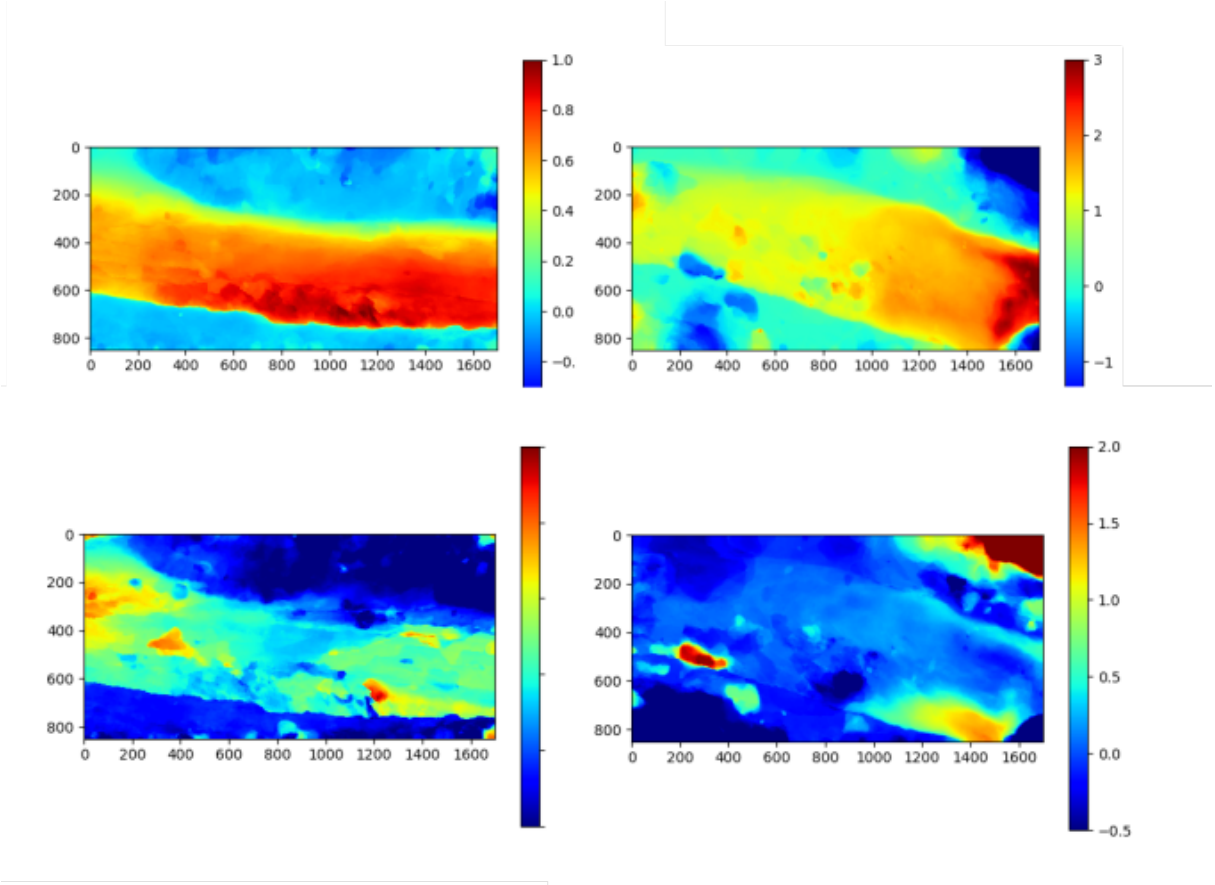


FIGURE 6.13 – Cartes de vélocité moyennées sur un ensemble de mesures bidates à pas de temps journalier (données classifiées "glacier visible"). Zooms sur le point de vue de l'appareil-photo gauche. Haut : déplacement horizontal. Bas : déplacement vertical. Droite : Museau du glacier. Gauche : zone en amont vers les coordonnées $(x, y) = (1500, 1500)$ sur l'image originale.

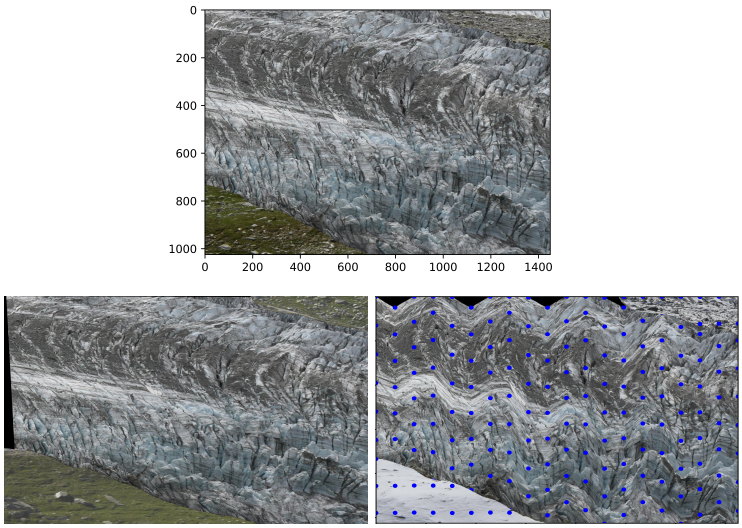


FIGURE 6.14 – En haut : sous-image maîtresse. Gauche : sous-image ré-échantillonnée avec un modèle affine. Droite : sous-image ré-échantillonnée avec un modèle affine par morceaux.

Jeu de données	CC	DMDF
$M1$	0.115	0.001
$M2$	0.633	0.084
$M3$	2.438	0.104
$M4$	1.710	0.043

TABLE 6.1 – Erreurs ϵ_{sim} en pixel obtenues sur les jeux de données de tests avec une corrélation croisée centrée et normalisée (CC) et la méthode proposée (DMDF).

CC présente également de nombreuses valeurs aberrantes sur le jeu de données 4.

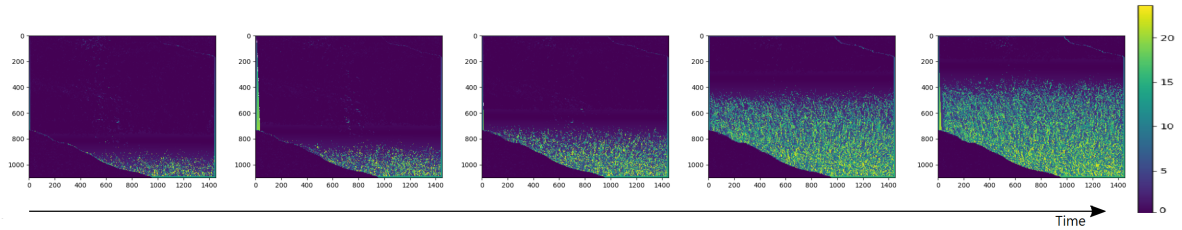


FIGURE 6.15 – Évolution de l'erreur entre les résultats obtenus par CC et la vérité terrain avec le modèle $M3$ à déformation affine croissante (modèle $M3$). La non-rigidité du jeu de donnée, ainsi que le bruit de mesure, augmentent avec les axes vertical (y) et temporel.

La figure 6.15 illustre le lien entre le bruit dans la mesure de corrélation-croisée CC et la non-rigidité engendrée par un cisaillement croissant dans le jeu de donnée $M3$. Cette erreur peut aller jusqu'à 20 pixel là où algorithme invariant va, dans une certaine mesure, capturer la transformation affine.

6.4 Conclusion

Dans ce chapitre, nous avons présenté les données monoscopiques et proposé une chaîne de traitements monodates puis bidates adaptés. Nos propositions incluent une détection de la texture spatiale liée aux franges du glacier ainsi que de la neige présente sur l'image afin de sélectionner automatiquement le maximum d'information pertinente sur une échelle intra et inter quotidienne. Cette sélection se fait par détection des images aberrantes parmi les images prises dans les mêmes conditions climatiques. Afin de compenser le mouvement indésirable de la caméra, un co-recalage par homographie 3×3 est effectué sur l'ensemble des images pertinentes selon une stratégie avec image maîtresse glissante. Cette stratégie permet de limiter la dérive de l'erreur de recalage dans le temps car celle-ci est généralement causée par un manque de corrélation entre deux images séparées par une portée temporelle importante (par exemple lors d'un changement de saison).

Le déplacement des pixels est estimé avec des algorithmes robustes de mise en correspondance puis de flot optique appelés respectivement DeepMatching et DeepFlow proposées par Weinzaepfel & al. [50]. Ces deux méthodes se sont montrées plus robustes qu'une

corrélation croisée centrée normalisée, utilisée dans les travaux similaires [1][2][3], dans un contexte où nous voulons traiter le plus de dates possibles. Les méthodes variationnelles sont en effet plus à même de capturer un mouvement fluidique (semi-rigide) que les mesures de similarité croisée et ce en raison de l'hypothèse de conservation de la quantité de mouvement.

Les mesures bidates restent néanmoins très sensibles aux conditions météorologiques et les valeurs aberrantes observées sont qualitativement corrélées avec la présence de nuages et de brume et, malgré une sous-pondération du terme de conservation de l'intensité lumineuse, avec la différence photométrique sur la zone glaciaire.

6.5 Perspectives

Concernant la phase de sélection d'images, il serait plus judicieux, en termes de temps de calcul, de considérer une corrélation du filtre de Gabor centrée longitudinalement sur le glacier et non sur l'ensemble de l'image. Les paramètres du filtre de Gabor détectant les fréquences désirées pourraient en outre être appris via un réseau de convolution profond sur l'ensemble des images acquises.

D'autre part, afin que la chaîne de traitements proposée soit plus automatique, le masque du glacier qui sert à quantifier la texture et au recalage pourrait être calculé par segmentation au lieu de considérer un masque manuel grossier. Cela pourrait par exemple se faire par le calcul d'un champ flot optique lisse sur un couple bidate monscopique non recalée et une segmentation par rapport à l'amplitude du déplacement. Dans ce sens, l'indice T pourrait être aussi retrouvé sur une zone détectée automatiquement tel que dans l'approche proposée dans la thèse de F. M. NEYER et al.[3] publiée en 2016. Il serait aussi intéressant de considérer une stratégie avec image maîtresse glissante qui soit automatique, par exemple via une description quantitative de la couleur dominante sur une région de l'image, par injection de données d'archive météorologique ou encore via des méthodes de classification en milieux naturels telles qu'on peut en trouver dans ces travaux [42].

Une perspective intéressante serait par exemple de développer une méthode de "*dehazing*" par morceaux et qui prend en compte l'importante différence d'altitude en milieux montagneux. Afin d'aborder les problématiques liées à l'hétérogénéité des données, la méthode DMDF pourrait être perfectionnée en considérant des paramètres et des poids adaptatifs, i.e. fixés par rapport aux images concernées dans le calcul. Concernant le calcul du flot optique, il serait intéressant de considérer plusieurs échelles temporelles dans les mesures bidates afin d'augmenter notre base d'observations. En effet, une portée temporelle plus grande entre deux images peut avoir l'avantage de posséder un plus grand ratio signal sur bruit. Cette optimisation fera l'objet du chapitre 8.

Chapitre 7

Proposition de traitements monodates et bidates stéréoscopiques

Sommaire du chapitre

7.1 Introduction **109**

7.2 Mise en correspondance et calcul de la pose des caméras **109**

7.3 Modèle numérique de terrain et stratégie adoptée **115**

7.3.1 Reconstruction 3D 115

7.3.2 Stratégie adoptée pour la modélisation du terrain 116

7.4 Mesures bidates stéréoscopiques **118**

7.5 Propagation d'incertitude **123**

7.5.1 Modèle proposé 123

7.5.2 Résultats d'incertitude tridimensionnelle 126

7.6 Conclusion **127**

7.7 Perspectives **131**

7.1 Introduction

Dans ce chapitre, nous allons exposer les outils déployés et ceux proposés afin de traiter les données monodates et bidates stéréoscopiques. Nous verrons, dans un premier temps, l'approche que nous déployons afin d'obtenir, à partir de time-lapses stéréoscopiques, un modèle numérique de terrain calibré, géo-référencé et dont la dispersion est minimale. Nous exposerons dans un second temps la méthode implémentée afin de remonter à des déplacements métriques par fusion de ce modèle avec les mesures de vélocité pixellique. Nous présenterons enfin dans la section 7.5 le modèle proposé afin d'estimer la dilution géométrique de la précision pixellique par triangulation stéréoscopique.

7.2 Mise en correspondance et calcul de la pose des caméras

La chaîne de traitements implémentée afin d'estimer la pose de la caméra c' par rapport à la caméra maîtresse c sur une paire monodate stéréoscopique se présente comme suit :

1. une détection, sur la paire d'images, des coins de Harris multi-échelles et affinement adaptés, tels que proposés dans les travaux de Mikolajczyk et Schmidt [68] (introduit en section 2.4.2.1),
2. une description des points d'intérêt basée sur des histogrammes de gradients orientés locaux (HOG) [37] (introduit en section 2.4.2.2),
3. une mise en correspondance par plus proche voisin approximatif [81] (introduit en section 2.4.2.3),
4. le calcul de la matrice essentielle E décrivant la géométrie épipolaire de notre système stéréoscopique à l'aide de l'algorithme algorithme des 5 correspondances [5] (introduit en section 3.4.4),
5. l'estimation de la pose de la caméra droite dans le repère de la caméra gauche à partir de la matrice essentielle E (introduit en section 3.4.4) qui consistera notre matrice de projection.

Ce choix a été fait empiriquement par observation qualitative des résultats obtenus avec plusieurs méthodes de l'état de l'art. Tout d'abord, plusieurs détecteurs et descripteurs de points d'intérêt ont été testés et ont donné des résultats de détection dont la précision est de l'ordre du centième de pixel. Les détecteurs de coins de Harris multi-échelles affinement adaptés, couplés avec une description via des histogrammes de gradients locaux, ont été retenus pour leur simplicité, leur bonne répartition sur la zone glaciaire et leur invariance par rapport à un changement de perspective localement linéaire (transformation spatiale affine par morceaux).

Lors de la mise en correspondance, une robustesse supplémentaire est introduite en posant un seuil de Lowe (section 2.4.2.3) et en filtrant les homologations dont l'erreur de symétrie stéréoscopique est supérieure à un second seuil. Cette erreur est définie, sans aucune contrainte géométrique, comme la distance Euclidienne d'un point $\mathbf{x}_g$ détecté sur l'image gauche et homologue à $\mathbf{x}_d$ et $\mathbf{x'}_g$, le point homologue à $\mathbf{x}_d$ en inversant le rôle des deux images dans le couple stéréoscopique.

D'autre part, le géo-référencement manuel de certaines coordonnées image sur les deux vues nous a permis d'ajouter une dizaine de correspondances sous-pixelliques. En effet, lors de la prise des mesures GPS in-situ, des bâches d'environ un demi mètre carré ont été posées et une prise de mesure dGPS s'est faite au centre de gravité de celles-ci. Afin de sonder encore plus le terrain, d'autres points faciles à repérer, tels des rochers de grande taille ou encore des reliefs visibles sur les images (par exemple l'intersection de ruisseaux ou de failles) ont également été géo-référencés avec précision. Des exemples de points de tels points sur illustrés sur la figure 7.1. Étant donnée qu'un recalage intra time-lapse a été effectué en aval des traitements tridimensionnels, cette brique manuelle ne sera à effectuer que sur un seul couple stéréoscopique et le reste des traitements restera alors automatique pour répondre à la problématique énoncée.

FIGURE 7.1 – Exemples d'images acquises lors d'une mission terrain pour la collecte de mesures GPS in-situ en rive droite (haut) et gauche (bas). Lors des sorties terrain et étant donné que le site étudié s'étend sur des kilomètres, les points à mesurer doivent tout d'abord être repérés conjointement sur les deux points de vue gauche et droit et sur un modèle numérique de terrain afin de naviguer efficacement. L'ensemble des données in-situ peuvent être visualisées sur le modèle numérique de terrain de la figure 5.12.

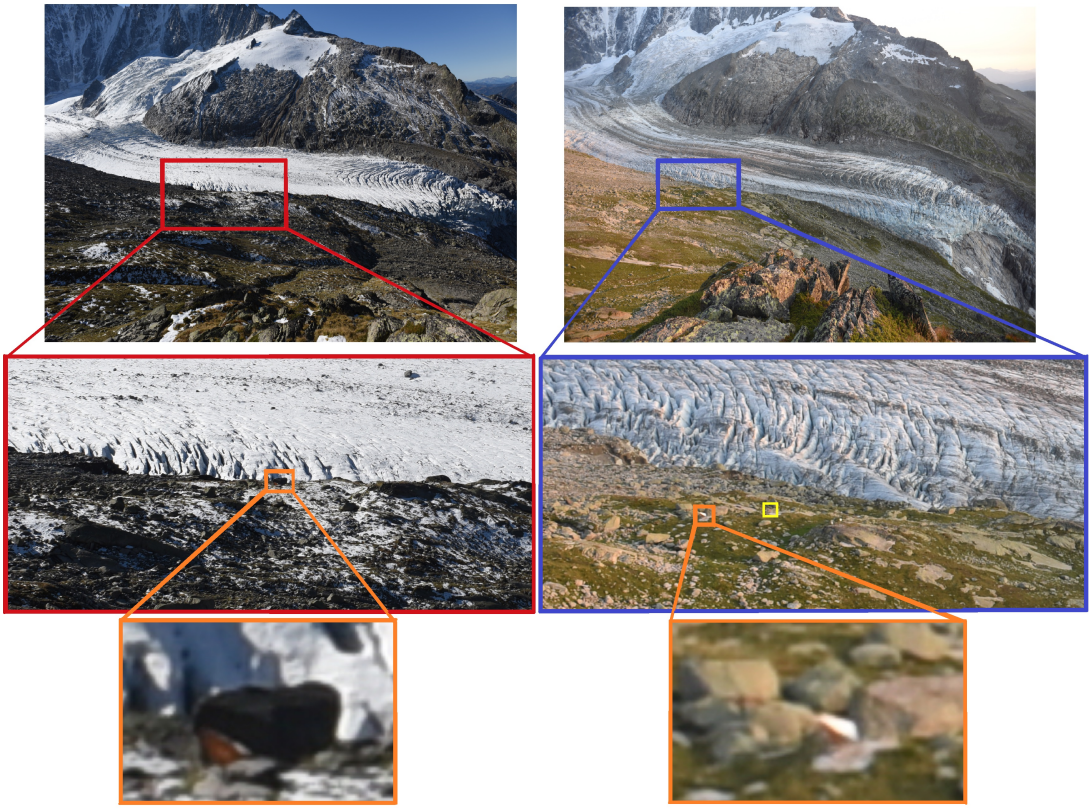
Des acquisitions témoins permettent ensuite de repérer ses objets saillants avec une précision sous-pixellique depuis les deux points de vue et de les recaler sur l'image maîtresse de chaque time-lapse. Comme on peut le voir sur la figure 7.2, étant donné l'échelle de la scène, la plupart de ces objets ne seront visibles que sur une dizaine de pixels et nous estimons qualitativement que la sélection manuelle de ces coordonnées est entachée Etant donnée qu'un recalage intra time-lapse a été effectué en amont d'une incertitude de l'ordre du tiers de pixel.

Lors de l'estimation de la matrice essentielle E, à 5 degrés de liberté, l'ensemble des correspondances détectées automatiquement (plus proches voisins approximatifs) et manuellement (pixels géolocalisés et recalés) ont été pris en compte par itérations RANSAC (algorithme introduit dans la section 1.2.4). La fonction coût est alors définie comme la distance Euclidienne maximale d'un point de contrôle à la ligne épipolaire qui lui est associée. Le critère de convergence des itérations RANSAC est alors défini de telle sorte que le modèle retrouvé est celui qui maximise les observations considérées "inliers" en zone glaciaire. En effet, en plus d'être la zone où nous voulons avoir le plus de cohérence, celle-ci est située à mi-chemin entre la zone lointaine (rive gauche) où la perspective entre les deux vues est plus similaire, mais où l'information tridimensionnelle est pauvre et les zones plus proches (rive droite) où la déformation perspective est difficilement modélisable avec une transformation affine par morceaux, principalement parce que les reliefs prononcés induisent une déformation parallaxe importante et beaucoup d'occlusions. Cette différence de perspective en rive proche est mise en évidence avec la visualisation d'un coin réflecteur situé sur le bord du glacier d'Argentière sur la figure 7.3). Au contraire, la zone

FIGURE 7.2 – Exemple de points repérés et géo-références sur une acquisition témoin. Les points marqués avec une flèche rouge désignent des objets géophysique saillants trouvés sur le site d'Argentière et les zones de couleur orange sont des bâches placées sur le site le jour de l'acquisition témoin.

du glacier a des profils d'altitude transversale et longitudinale assez lisses et une texture plus détectable par description affinement adaptée et est donc plus à même de générer les bonnes estimations stéréoscopiques. Les résultats affichés à cette étape de la chaîne de traitements figurent sur les figures 7.4 et 7.5.

(a) Photo d'un coin réflecteur prise sur terrain lors de la prise de mesure.



(b) Vue du coin réflecteur à partir des points de vue gauche et droite.

FIGURE 7.3 – 7.3a : Installation d'un coin réflecteur sur terrain. 7.3b : Vue du coin réflecteur dans 7.3a par les deux appareil-photos. Mise en évidence de l'échelle des points d'intérêt à détecter et de la différence de perspective locale particulièrement importante en rive droite du glacier (rive proche) . Ceci souligne la nécessité d'utiliser une détection ou une description de points de contrôle affinement adaptée pour la mise en correspondance stéréoscopique. On remarque que l'aspect de la scène en rive droite (visualisations zoomé au maximum du bas) change radicalement avec le point de vue à cause de la topographie (lit rocheux creusé du glacier, pente raide, reliefs importants, etc.) et des paramètres extrinsèques des caméras (entre autres, une différence d'environ 50m d'altitude).

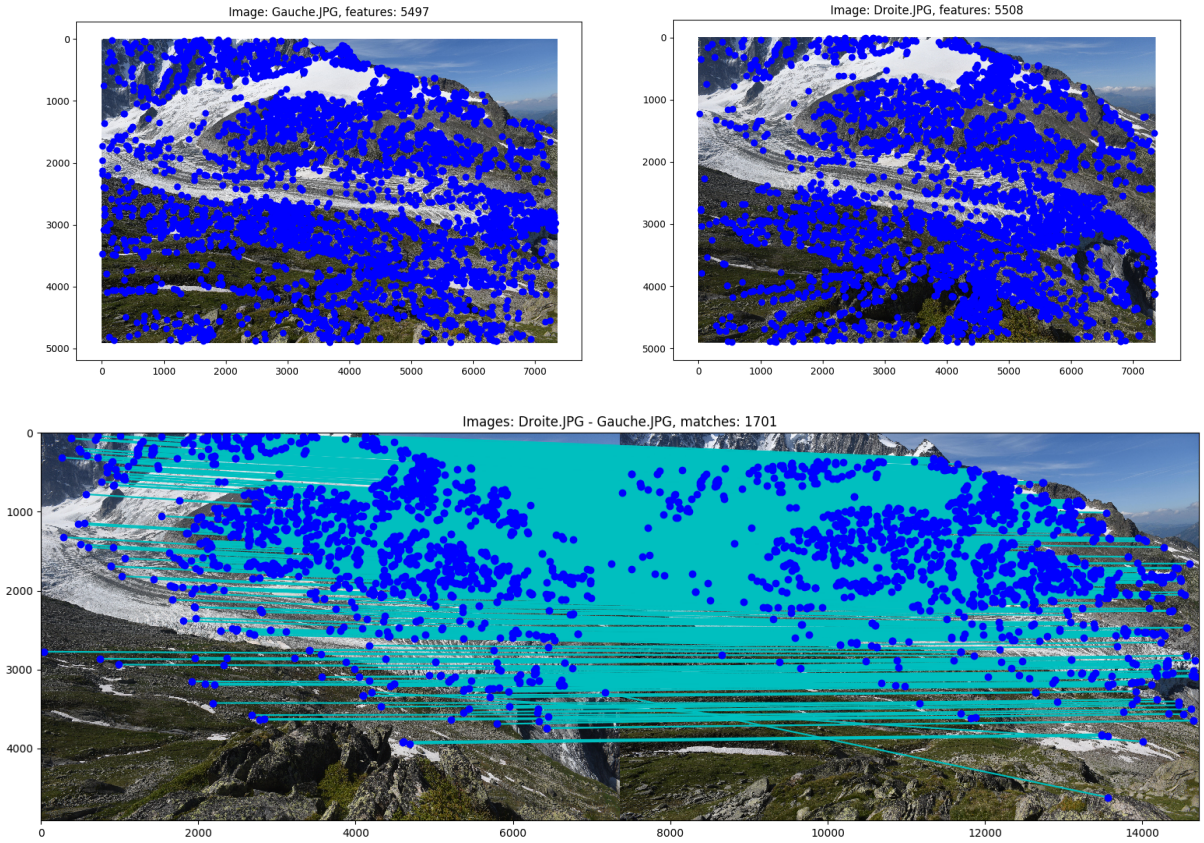


FIGURE 7.4 – Détection (haut) et mise en correspondance (bas) de coins de Harris multi-échelles et affinement adaptés sur un couple stéréoscopique avec une bonne visibilité.

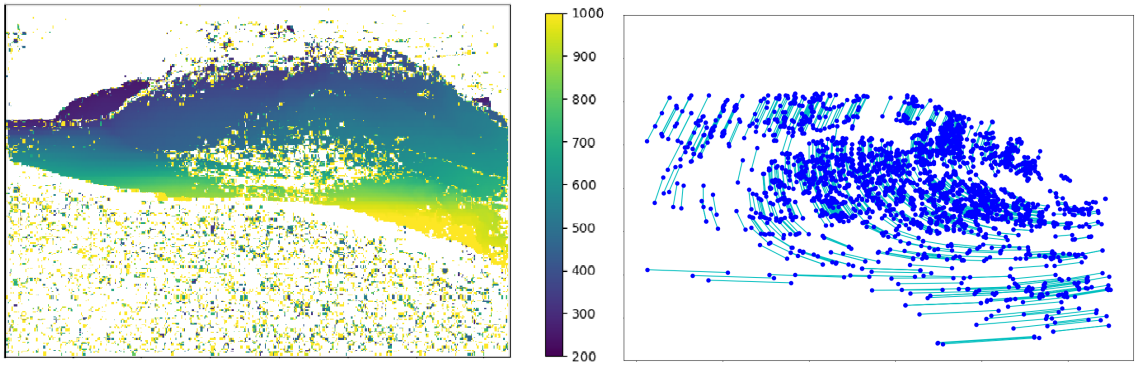


FIGURE 7.5 – Mise en évidence du mouvement perspective par visualisation de la carte de disparité (gauche) et des correspondances retenues après itérations RANSAC (droite). Malgré les mesures in-situ en rive droite du glacier (rive proche du système d'observation stéréoscopique), nous remarquons un manque de mise en correspondances dans cette zone proche après raffinement. L'étude tridimensionnelle de la scène observée se fera alors à partir de la profondeur minimale du glacier.

7.3 Modèle numérique de terrain et stratégie adoptée

7.3.1 Reconstruction 3D

Une fois les matrices de projection sur une paire stéréoscopique estimées, il convient d'effectuer une triangulation pour obtenir un modèle numérique de terrain $\mathbf{X}_i$ constitué de points tridimensionnels exprimés dans le repère de la caméra c de référence puis dans le repère des données de géo-referencement. Celle-ci est tout d'abord effectuée via l'algorithme linéaire DLT (Direct Linear Transform) décrit en section 3.4 de ce manuscrit qui suppose les hypothèses suivantes :

— les paramètres intrinsèques (focale et point central) n'ont pas subi de dérive depuis l'étalonnage par le constructeur,

— les pixels sont régulièrement espacés sur le plan image et les distorsions optiques liées au grand angle de vue sont négligeables.

Ces hypothèses sont a-priori assez réalistes avec une caméra moderne et un objectif à focale fixe et font qu'on reste dans le cas d'une reconstruction 3D calibrée (reconstruction calculée à partir de la matrice essentielle E et non Fondamentale F), mais peuvent néanmoins induire du biais dans la triangulation linéaire et aboutir à un nuage de points souffrant d'une légère distorsion projective et/ou en barillet. Il est donc nécessaire d'affiner davantage le résultat obtenu. Pour ce faire, une minimisation de l'erreur de reprojection (ajustement des faisceaux) par la méthode itérative de Levenberg-Marquadt est préconisée et peut, dans notre cas, être déployée des deux façons suivantes ; un ajustement des faisceaux avec, comme variables de décision, les coordonnées tridimensionnelles et les paramètres intrinsèques et extrinsèques et :

1. comme contraintes, les données de calibration et de géo-référencement,
2. suivi d'une mise à l'échelle du nuage de points avec les données de calibration et de géo-référencement.

Tandis qu'avec la première approche l'ambiguïté projective et sur l'échelle sont résolues jointement, avec la deuxième ceci se fait séquentiellement afin de retrouver un modèle 3D métrique. Dans les deux cas, une initialisation avec la solution obtenue par triangulation DLT est indispensable pour faire converger le *solver* étant donné que la fonctionnelle minimisée (équation 3.29) est hautement non convexe et en particulier en considérant un grand nombre de variables de décision.

Le choix est porté sur une résolution séquentielle des ambiguïtés tridimensionnelles. La deuxième approche a en effet fourni des résultats plus denses, plus cohérents avec notre connaissance à priori de la scène et moins dispersés dans le temps. La mise à l'échelle a été réalisée via la solution analytique proposée par Horn [104]. Cette disparité dans les résultats peut provenir de fait de la non-convexité de la fonction coût associée

à l'ajustement de faisceaux qui, sous l'effet de contraintes strictes, échoue à trouver un modèle cohérent en tout point.

7.3.2 Stratégie adoptée pour la modélisation du terrain

Malheureusement, même avec la méthode envisagée, les mesures de modèles numériques de terrain obtenus à différentes dates souffrent d'une pauvre reproductibilité en zones fixes et nous observons une dispersion des points triangulés. Un estimation de l'erreur de reproductibilité associée à un modèle 3D peut être définie comme suit :

$$\epsilon_{\text{repro}_i} = \frac{1}{T} \sum_{j=1}^{T} \sum_{\Omega_{ZF}} d_{\text{C2C}}(\mathbb{X}_i, \mathbb{X}_j). \tag{7.1}$$

L'équation 7.1 traduit la distance point par point (C2C pour "*cloud to cloud*") moyenne entre un nuage de points $\mathbb{X}_i$ et T autres nuages de points sur la zone spatiale Σ_{ZF} en zone fixe proche glacier. La figure 7.6 montre qu'une telle mesure est de l'ordre du mètre en rive proche du front glaciaire et croit sur les zones lointaines moins résolues et peu contraintes par les données GPS.

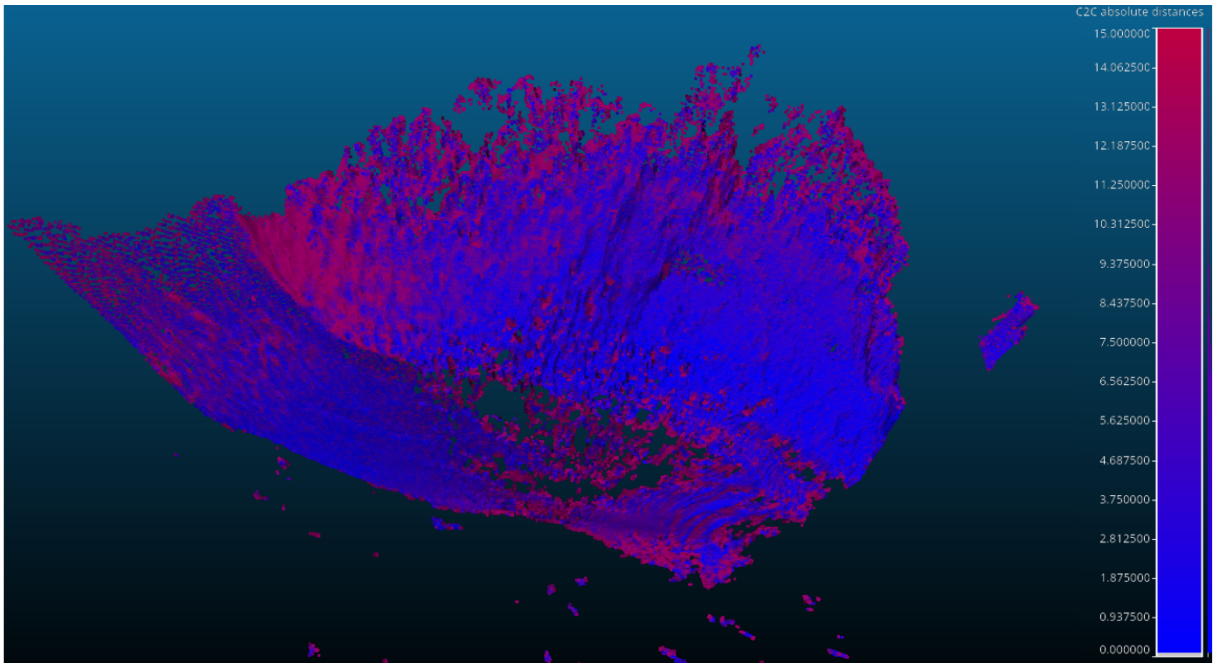


FIGURE 7.6 – Zoom sur une carte 3D de distance point par point moyenne sur 10 nuages de points en mètres (avec un mouvement des caméras compensé lors de la brique de traitements 2D). L'écart entre les zones fixes est une estimation de l'erreur de reproductibilité.

Nous notons en effet en ces zones lointaines un décollement des surfaces reconstruites entraînant une incertitude de reproduction sur un voisinage de 10 points tridimensionnels égale à 12.5mètres. Cette mesure est aussi particulièrement importante sur les zones

crevassées ce qui met en évidence la sensibilité de mesures effectuées directement entre nuages de points sans un passage par les données bidimensionnelles. Du fait de cette dispersion, calculer un modèle numérique de terrain à chaque date introduirait un biais non négligeable dans les mesures de velocité métrique. La stratégie (2×2D)+t pour les traitements bidates stéréoscopiques est alors rejetée.

D'autre part, considérer un modèle unique, moyen ou médian point par point, risquerait d'induire un MNT dispersé et spatialement incohérent. Nous adoptons alors une approche 2×(2D+t) afin de ne pas noyer le signal des mesures bidates métriques dans le bruit pendant la phase de triangulation 3D. Pour ce faire, au lieu d'estimer un modèle tridimensionnel à chaque date impliquée, une unique projection (définie par les deux matrices de projection P_g et P_d) sera choisie afin de modéliser la morphologie de la scène. En faisant les hypothèses d'un déplacement spatialement continu, d'une déformation géomorphologique négligeable, y compris sur la zone glaciaire, cette approche reste réaliste. De plus, celle-ci est conforme avec une modélisation Lagrangienne de l'écoulement glaciaire (voir section 4.2).

Le modèle $\mathbb{X}_{MNT}$ retenu est celui qui minimise la distance moyenne définie par l'équation 7.1 sur un ensemble de 10 paires stéréoscopiques disposant de la meilleure visibilité. Nous voyons sur la figure 7.7 que celui-ci est spatialement compact et bien répandu sur la scène observée et à partir de la profondeur du glacier sauf sur certaines zones où nous observons un manque de points triangulés entraînant un trou dans le modèle numérique de terrain. Ceci est principalement du aux occlusions causées par les reliefs montagneux. Celui-ci est obtenu avec une erreur de reprojection résiduelle moyenne de 3.10^{-4} pixel.

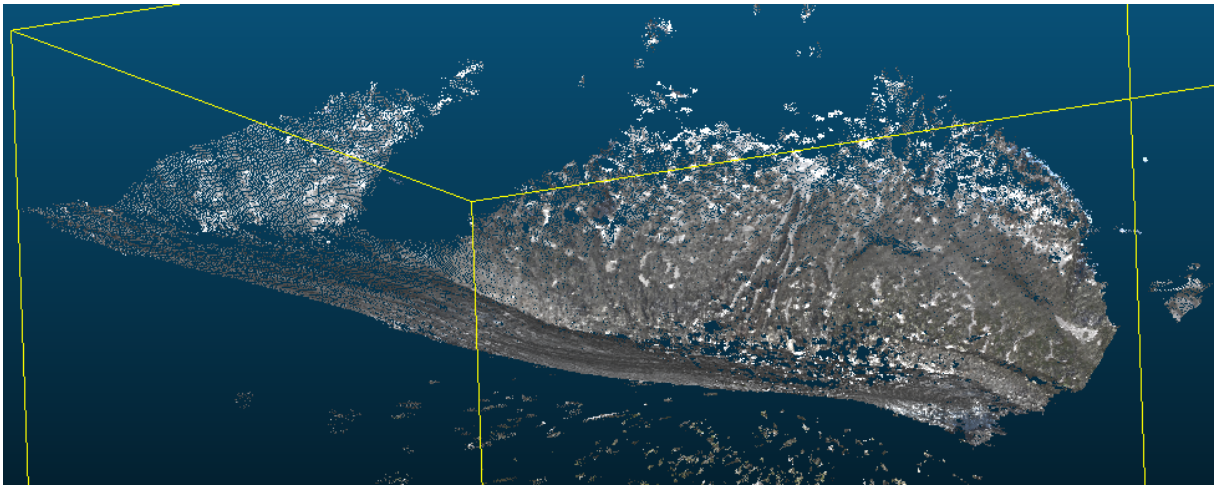


FIGURE 7.7 – Nuage de points 3D $\mathbb{X}_{MNT}$, obtenu par triangulation DLT, ajustement de faisceaux et recalage rigide avec les données de contrôle, minimisant l'erreur de reproductibilité.

L'évaluation qualitative de $\mathbb{X}_{MNT}$ peut être complétée par le traçage des lignes de niveaux telles que celles de la figure 7.8 et notamment les niveaux d'altitude qui permettent une comparaison visuelle à une vérité terrain tel qu'illustré sur les figures 7.9.

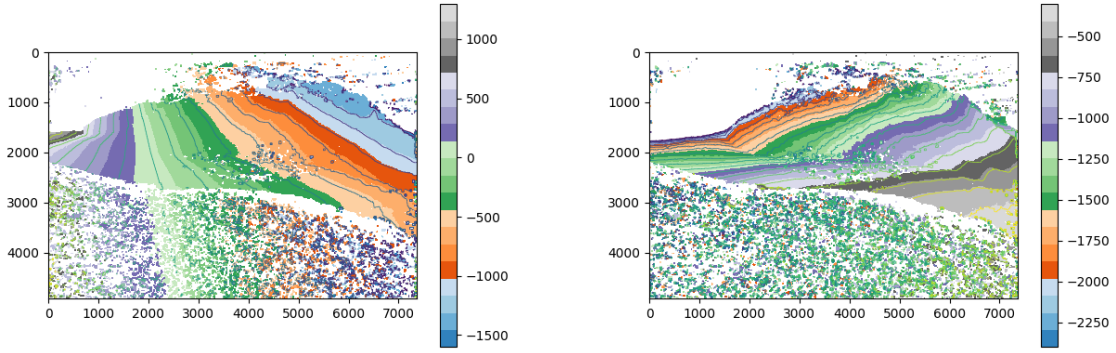


FIGURE 7.8 – Coordonnées géo-référencées (exprimées dans un repère arbitraire à partir des coordonnées DMS (degrés, minutes, secondes)) du modèle numérique de terrain. Gauche : longitude. Droite : latitude.

7.4 Mesures bidates stéréoscopiques

La méthodologie adoptée repose ainsi sur une approche 2×(2D+t) (voir section 3.5.3). Celle-ci est basée sur la fusion de mesures bidates monoscopiques avec les mesures monodates stéréoscopiques, toutes deux effectuées en amont (voir figure 3.1). Étant donné que le flot observé est continu (écoulement semi-rigide) et sous l'hypothèse d'une surface qui ne se déforme pas dans le temps (perte d'épaisseur imperceptible du glacier tempéré), le calcul du "*scene-flow*", ou flot de scène, est directement accessible par projection sur le modèle numérique de terrain. En effet, sous réserve que le mouvement des caméras ait été compensé et que les matrices de projection P_g et P_d soient connues (intrinsèques connues) et qu'elles aient été raffinées et corrigées avec les données de recalage GPS, il suffit, à partir d'une paire de flots optiques calculée à partir d'un quadruple stéréo-temporel, d'effectuer une triangulation linéaire (algorithme DLT de résolution par moindres carrées présenté en section 3.4.1) des paires sous-pixeliques $(\mathbf{x_g}, \mathbf{x_d})$ et $((\mathbf{x_g} + \mathbf{F}_{\mathbf{g}_{t,t+\Delta t}}), (\mathbf{x_d} + \mathbf{F}_{\mathbf{d}t,t+\Delta t}))$ homologués stéréscopiquement. Ces paires définissent en effet les coordonnées de départ et d'arrivée des deux vecteurs vélocité $\mathbf{F_g}$ et $\mathbf{F_d}$, projections du vecteur "*scene-flow*" sur les plans images gauche et droite. La différence entre les deux coordonnées 3D obtenues après "dé-homogénisation" (équation 3.4) définie alors le vecteur de déplacement métrique du point $\mathbf{X}$ entre les dates t et $t + \Delta t$ (voir illustration de la figure 7.10). Pour rappel, cela consiste en la résolution du problème linéaire suivant :

$$\begin{pmatrix} x_g p_g^{3^T} - p_g^{1^T} \\ y_g p_g^{3^T} - p_g^{2^T} \\ x_d p_d^{3^T} - p_d^{1^T} \\ y_d p_d^{3^T} - p_d^{2^T} \end{pmatrix} \mathbf{X} = 0, \tag{7.2}$$

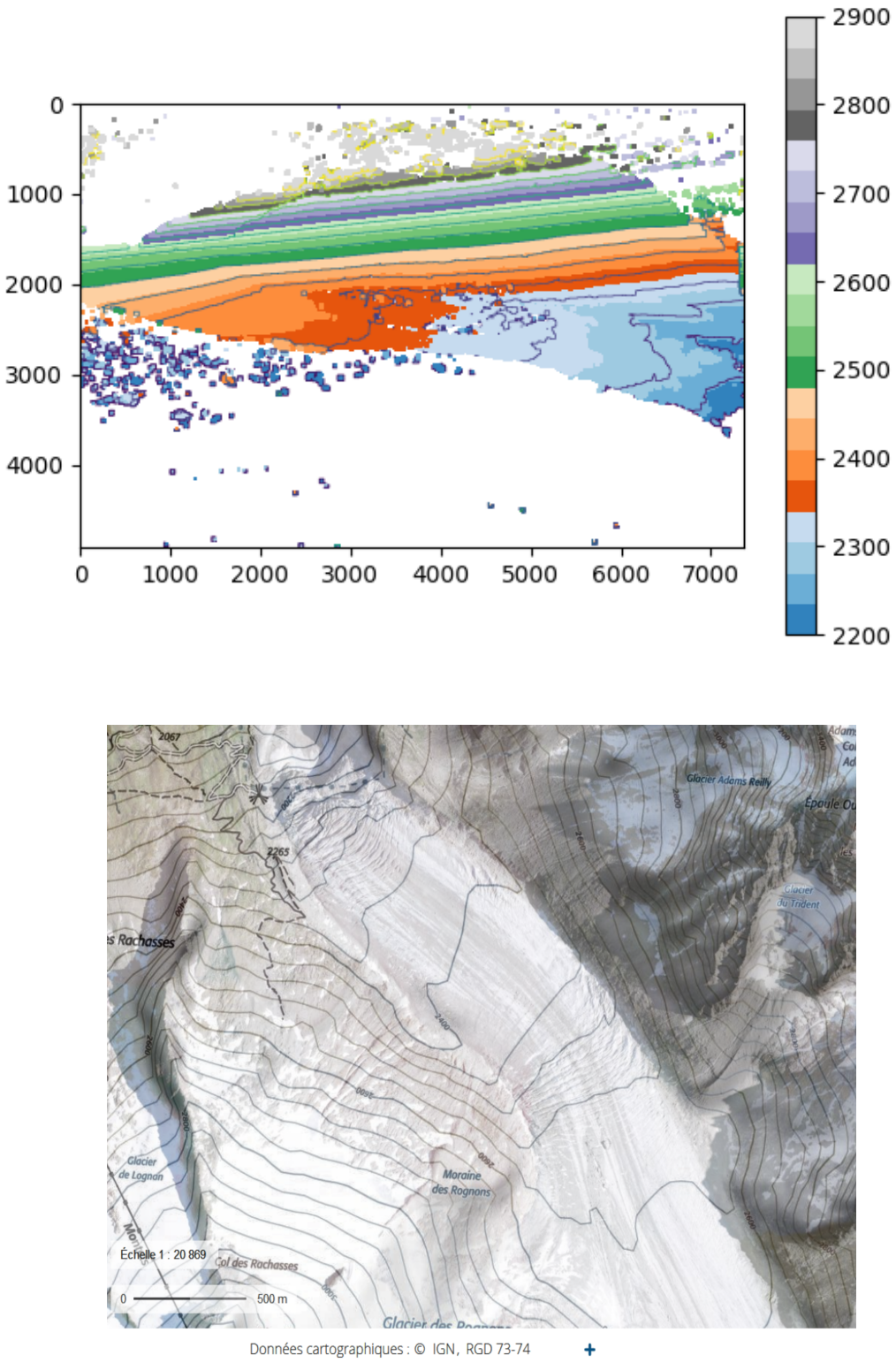


FIGURE 7.9 – Coordonnées geo-référencées du modèle numérique de terrain : lignes de niveau d'altitudes. Haut : modèle obtenu par triangulation à partir d'une paire stéréoscopique d'images. Bas : modèle publié par l'Institut national de l'information géographique et forestière (IGN) basé sur un MNT obtenu par imagerie optique aéroportée avec une résolution de 20cm/pixel.

pour la coordonnée tridimensionnelle $\mathbf{X}$, où $p_c^{i^T}$ est la i-ème colonne de la matrice de projection P_c.

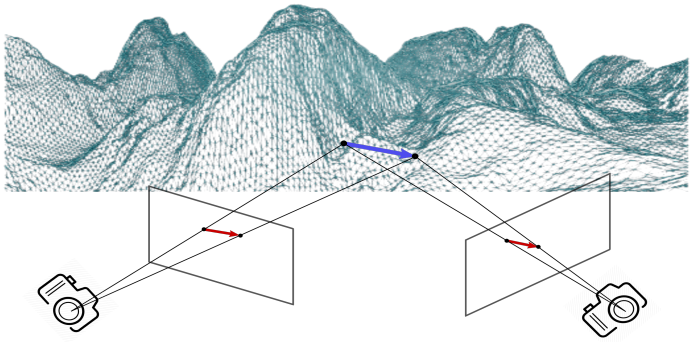


FIGURE 7.10 – Illustration d'un vecteur flot de scène "*scene-flow*" projeté sur deux plans images gauche et droite. En rouge : vecteurs vélocité pixellique ou flot optique. En bleu : vecteur vélocité métrique ou flot de scène.

Les cartes de déplacement East-North-Up associées aux deux paires de la figure 6.8 calculées entre le 08 juillet 2018 et le 07 août 2018 sont illustrées sur la figure 7.11. Les hypothèses rhéologiques faites sur l'écoulement glaciaire à partir de l'observation des

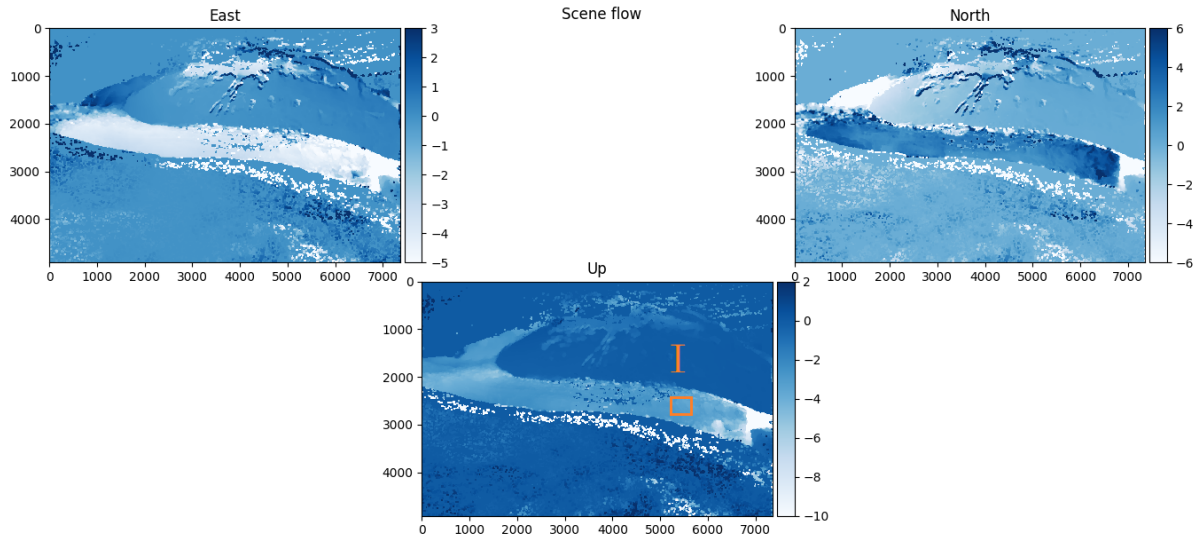


FIGURE 7.11 – Exemple d'une carte de vélocité métrique obtenue entre 08 juillet 2018 et le 07 août 2018 (Δt = 30 jours). La vélocité résiduelle en zones statiques, quoique pouvant décrire une éventuelle fonte de glace, nous donne une idée sur la qualité d'une carte de vélocité.

cartes de vélocité pixellique se sont vues confirmées après la légère correction de parallaxe induite par prise en compte de la profondeur. Cette mise à l'échelle nous a en outre permis de calculer un ensemble de cartes de vélocités entre les images retenues, et ce en distribuant les données d'une manière parallèle sur les unités de calcul.

Comme évoqué plus tôt, le meilleur moyen d'inspecter une carte de vélocité et d'émettre des hypothèses sur la nature du bruit de mesure reste la construction d'indicateurs décri-

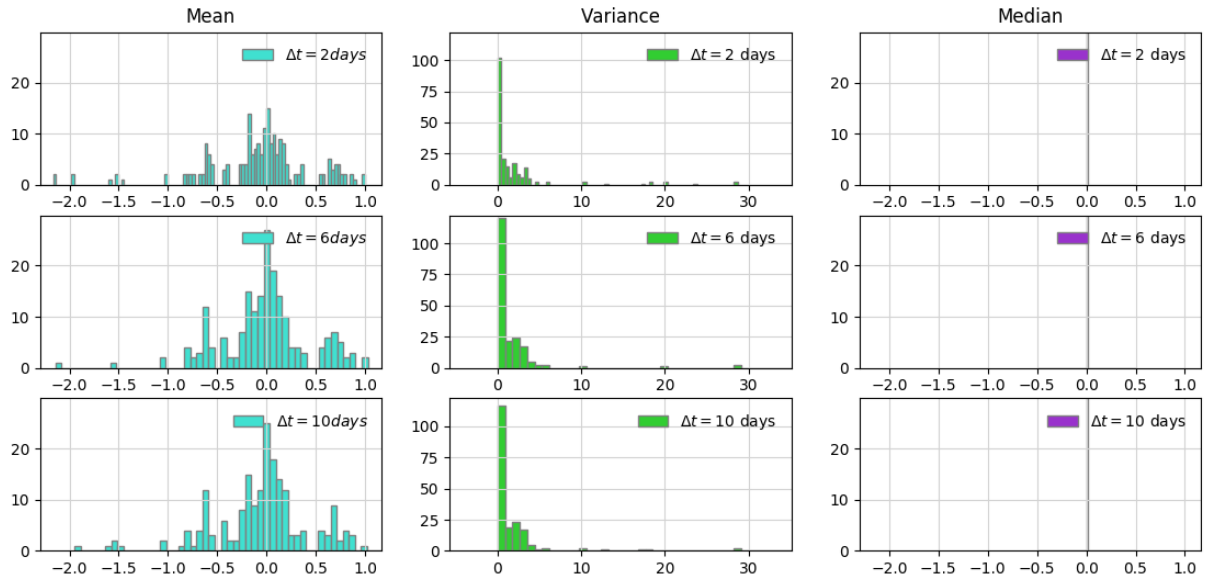


FIGURE 7.12 – Erreur-types en zones supposées statiques sur différentes cartes de vélocités métriques avec une portée temporelle différente : composante "East". mètre, mètre carrés et mètre.

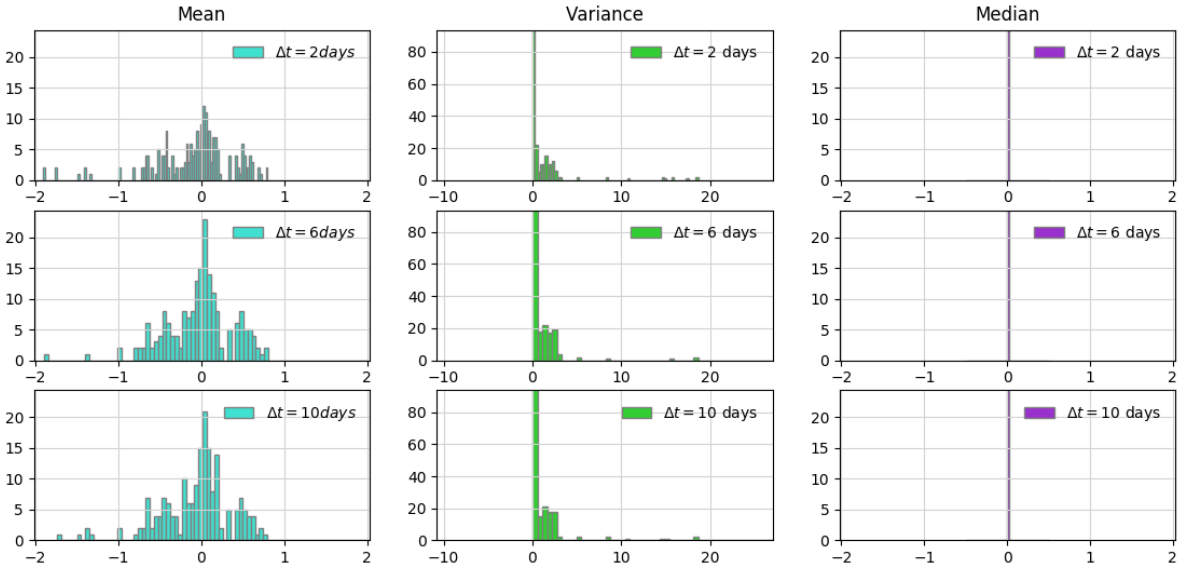


FIGURE 7.13 – Erreur-types en zones supposées statiques sur différentes cartes de vélocités métriques avec une portée temporelle différente : composante "North". mètre, mètre carrés et mètre

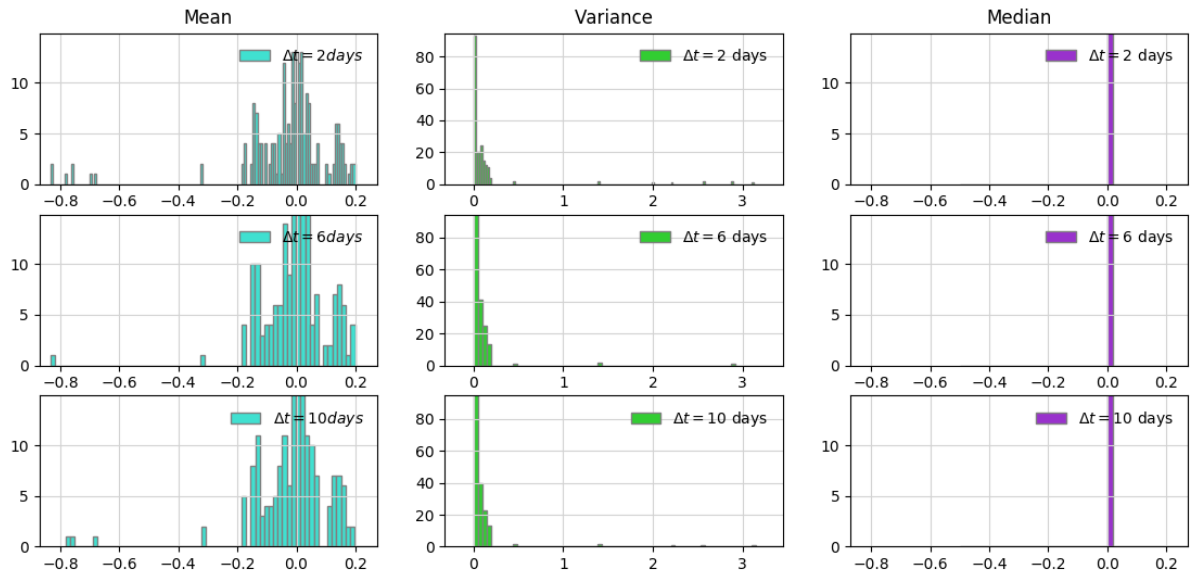


FIGURE 7.14 – Erreur-types en zones supposées statiques sur différentes cartes de vélocités métriques avec une portée temporelle différente : composante "Up". mètre, mètre carrés et mètre

vant la vélocité résiduelle en zones statiques. La distribution des mesures là où le signal est supposé être nul en plusieurs pixels nous informe en effet sur la présence de biais et sur la dépendance entre le bruit et la portée temporelle, i.e. si le bruit de mesure doit être considéré relatif à l'amplitude du déplacement ou absolu. Nous voyons sur les figures 7.12, 7.13 et 7.14 que la médiane en zone fixe est toujours nulle. Ceci nous rassure donc sur le fait que les différentes briques de traitements n'aient pas introduit de l'erreur systématique. La distribution des erreurs moyennes sur les zones fixes quant à elle est centrée en 0 et ne s'amplifie pas avec la portée temporelle. Nous nous retrouvons néanmoins dans le cas d'une distribution centrée en zéro, i.e. indicatrice d'un bruit aléatoire, à queue lourde (particulièrement à gauche), c.à.d. que des réalisations erronées se produisent plus fréquemment qu'une distribution normale ne le laisserait supposer. D'autre part, l'écart-type de la moyenne de la vélocité résiduelle en zones fixes nous donne une idée sur l'erreur-type expérimentale. Cette mesure est cependant dangereuse car pouvant indiquer un déplacement réel, telle une fonte de neige ou un nuage. La variance de cette mesure est alors indicatrice d'un bruit de mesure aléatoire puisque une déplacement réel sera généralement quasi-stationnaire. Après élargissement par un facteur 2 de l'erreur-type observée en zones fixes, celle-ci est égal à 0.7m pour les composantes "*East*" et et "*North*" et de 0.19m pour la composante "*Up*"(figures 7.12, 7.13 et 7.14).

7.5 Propagation d'incertitude

7.5.1 Modèle proposé

Nous allons maintenant présenter le modèle élaboré afin d'estimer la dilution géométrique de la précision spatiale des points images triangulés (voir section 3.7). Celui-ci est inspiré des techniques introduites dans la section 3.7 [16][131][17][135] illustrées par la figure 3.8 et tient compte de la plage d'incertitudes liée aux mesures 2D triangulées, de l'angle de lancé de rayon, et enfin la profondeur Z du point 3D dont nous voulons estimer l'incertitude de triangulation.

Soit $\mathbf{u}_c = (u, v)_c^T$ les coordonnées d'un point de contrôle exprimé dans le repère image d'une caméra c. La plage d'incertitudes $(\mu_u, \mu_v)_c$ associée à celui-ci sera amplifiée lors de la triangulation stéréoscopique. Cette amplification se fait selon un cône dont le sommet et le centre de projection $\mathbf{S}_c$ et dont l'intersection avec le plan image forme l'ellipse d'incertitude de demi-axes principaux de tailles μ_u et μ_v tel qu'illustré par la figure 7.15.

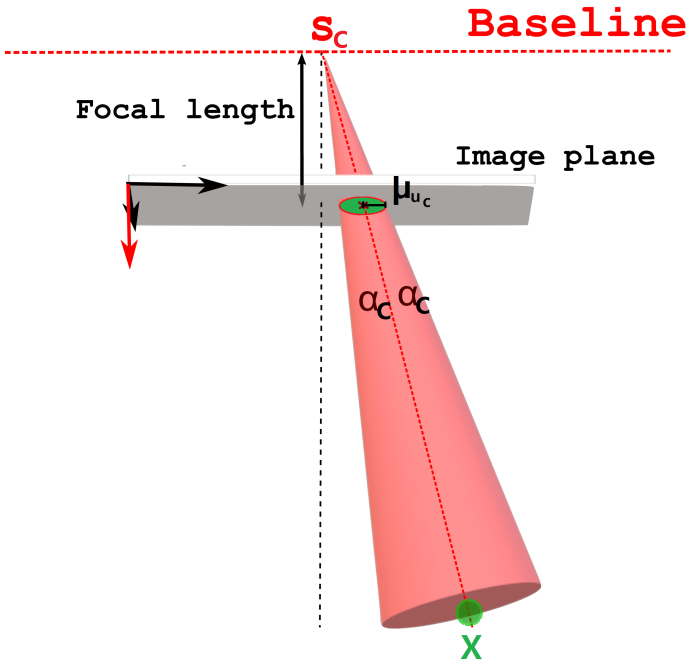


FIGURE 7.15 – L'ellipse d'incertitude 2D de demis axes principaux de tailles μ_u et μ_v, associée à un point de contrôles $\mathbf{u}_c = (u, v)_c^T$ exprimé dans le repère du plan image de la caméra c (ellipse verte) est induite par un cône d'angle α_c dont le sommet et le centre de projection $\mathbf{S_c}$ de la caméra c.

Le volume défini par l'intersection des deux cônes d'incertitude de centres S_c et $S_{c'}$ à la profondeur Z donne le volume d'incertitude d'un point X triangulé tel qu'illustré sur la figure 7.16. Celui-ci est invariant par rotation des rayons lancés autour de l'axe de la base spatiale et si nous ne considérons que le quadrilatère 2D $Q_1Q_2Q_3Q_4$, nous rejoignons la modélisation publiée par Hartley et Zisserman sur la plage de variation d'un

point tridimensionnel triangulé (figure 3.7). Dans notre modèle, l'erreur de reconstruction en profondeur Z est alors la longueur de la diagonale Q_2Q_4 et l'erreur en position est la longueur de la diagonale Q_1Q_3. Le modèle proposé repose en plus sur les hypothèses

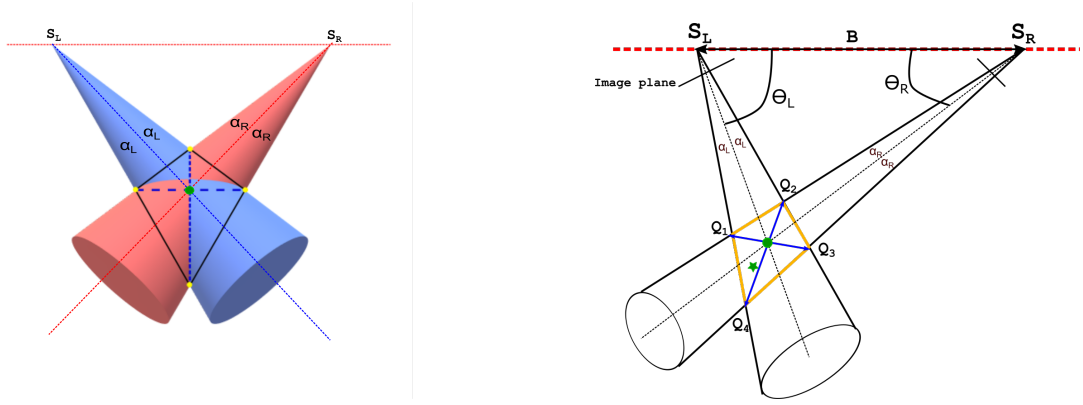


FIGURE 7.16 – Les deux cônes d'incertitude gauche et droite de sommet S_c et $S_{c'}$, $(c, c') = (L, R)$, centrés autour des rayons lancés pour la triangulation stéréoscopique, s'intersectent en un volume dont la coupe transversale est le quadrilatère $Q_1Q_2Q_3Q_4$.

suivantes :

$$(\mu_u, \mu_v)_g = (\mu_u, \mu_v)_d$$
$$\alpha < 0.1\text{rad}.$$

Les différents angles impliqués peuvent être calculés comme suit :

$$\begin{aligned} \alpha_c &= \arccos\big(\mathbf{U}_c^T\mathbf{U}_c^+ / ||\mathbf{U}_c||_2||\mathbf{U}_c^+||_2\big), \\ \theta_c &= \arccos\big(\mathbf{F}_c^T\mathbf{U}_c / ||\mathbf{F}||_2||\mathbf{U}_c||_2\big), \end{aligned} \tag{7.3}$$

où les différents vecteurs de l'équation 7.3 sont exprimés dans le repère de chaque caméra c tel qu'illustré par la figure 7.17 et sont définis comme suit :

$$\begin{cases} \mathbf{F}_c = (0, 0, f_c)^T & \mathbf{U}_c = (u_c, v_c, f_c)^T, \\ \mathbf{U}_c^+ = (u_c + \mu_u, v_c + \mu_v, f_c)^T & \mathbf{u}_c = (u_c, v_c) = (x_c - x_0, y_c - y_0), \end{cases} \tag{7.4}$$

avec f_c la distance focale de la caméra c.

Enfin, des mesures trigonométriques sur les différents triangles du modèle de la figure 7.16, couplées avec des approximations des petits angles (approximations de Gauss) nous ont permis de calculer les longueurs suivantes :

$$\begin{aligned} \overline{Q_2Q_4}^2 &= (2\alpha B)^2 \\ &\quad \big((G'(a+b))^2 + (D(a-b))^2 + 2G'(a+b)D(a-b)\cos(\theta_g + \theta_d) / (a^2 - b^2)^2\big), \\ \overline{Q_1Q_3}^2 &= \big(2\alpha B / a - b\big)^2 \big(G'^2 + D'^2 + 2G'D'S\big), \end{aligned} \tag{7.5}$$

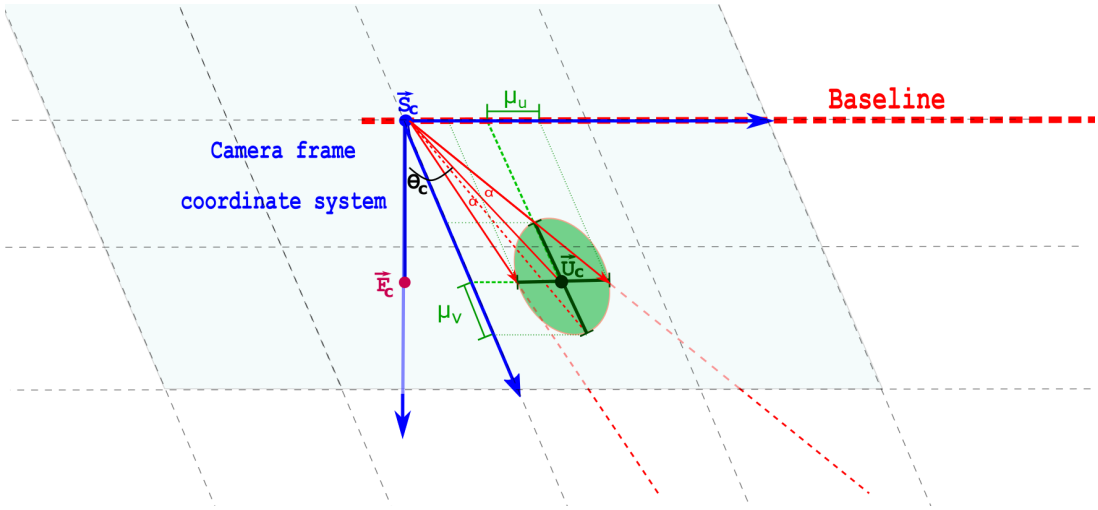


FIGURE 7.17 – Modèle géométrique et notations adoptées afin d'estimer les dimensions du quadrilatère d'incertitude de triangulation. α est le demi-angle du cône d'incertitude et θ_c est l'angle de lancé de rayon à partir d'une caméra $c \in (g, d)$.

avec :

$$\begin{cases} a = \sin^2(\theta_g + \theta_d) & b = 2\alpha \sin(\theta_g + \theta_d)\cos(\theta_g + \theta_d) \\ D = \sin(\theta_d) + \alpha\cos(\theta_d) & D' = \sin(\theta_d) - \alpha\cos(\theta_d) \\ G = \sin(\theta_g) + \alpha\cos(\theta_g) & G' = \sin(\theta_g) - \alpha\cos(\theta_g) \\ S = \cos(\theta_g + \theta_d) + 2\alpha\sin(\theta_g + \theta_d) & \end{cases}, \tag{7.6}$$

où $\alpha_g = \alpha_d = \alpha$ compte-tenu des hypothèses posées plus haut et B est la baseline spatiale (distance entre les deux caméras).

La figure 7.18 illustre l'évolution des incertitudes en profondeur telle que définie plus haut en fonction des angles de lancé de rayon θ_g et θ_d. Conformément à la figure 3.8, celle-ci tend vers l'infini quand les angles de visée $(\theta_c, \theta_{c'})$ ont une somme qui tend vers π (par exemple quand ils tendent vers $(0, \pi)$, $(\pi, 0)$ et $(\pi/2, \pi/2)$). Une divergence plus rapide est observée vers $(0, 0)$ et (π, π). Les expérimentations ont aussi montré que la taille des zones de divergence croit avec la croissance de l'incertitude 2D (μ_u, μ_v) et que la fonction d'incertitude en position se comporte de la même manière que celle de l'incertitude en profondeur avec des valeurs légèrement moins importantes et une divergence moins rapide.

Étant donné le nuage de points $\mathbb{X}_{MNT}$ obtenu via un système d'observation stéréoscopique calibré et retenu en tant que modèle numérique de terrain, l'algorithme proposé se présente alors comme suit : Ce calcul peut être optimisé par distribution parallèle des données sur les unités de calculs.

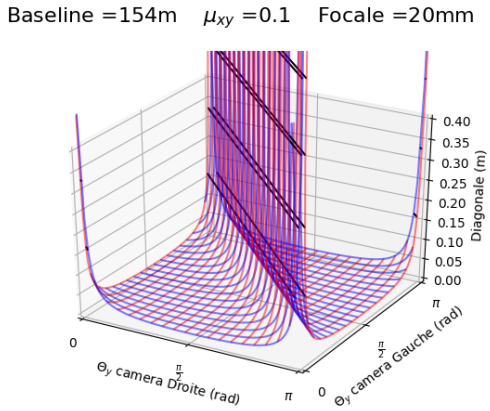


FIGURE 7.18 – Lignes de niveau de l'incertitude en profondeur en fonction des angles de visée pour $\mu_u = \mu_v = 0.1$ pixels.

Entrée : K_g, K_d, P_g, P_d, B, $\mu_u = \mu_v$, $\mathbb{X}_{MNT}$
Sortie : Carte d'incertitude en profondeur $\mu_Z = \overline{Q_2Q_4}/2$ en tout point, carte d'incertitude en position $\mu_{XY} = \overline{Q_1Q_3}/2$ en tout point
Tant que : $X \in \mathbb{X}_{MNT}$
(x_g, y_g) = Reprojection(P_g, X)
(x_d, y_d) = Reprojection(P_d, X)
Calculer α, θ_g, θ_d en fonction de (x_g, y_g), (x_g, y_g), K_g et K_d à l'aide l'équation 7.3
Calculer $\overline{Q_1Q_3}$ and $\overline{Q_2Q_4}$ en fonction de α, θ_g, θ_d à l'aide l'équation 7.5

7.5.2 Résultats d'incertitude tridimensionnelle

L'incertitude tridimensionnelle des mesures métriques peut être propagée depuis l'incertitude des vecteurs flot optique, auquel cas nous estimons la plage de variation des coordonnées d'arrivée du vecteur flot de scène, ou encore être associée à la dilution par triangulation de la précision des points d'intérêt homologués stéréoscopiquement. Dans ce deuxième cas, l'incertitude métrique correspond plutôt à la plage de variation des coordonnées tridimensionnels du nuage de points $\mathbb{X}_{MNT}$. Ces deux incertitudes sont illustrées sur la figure 7.19.

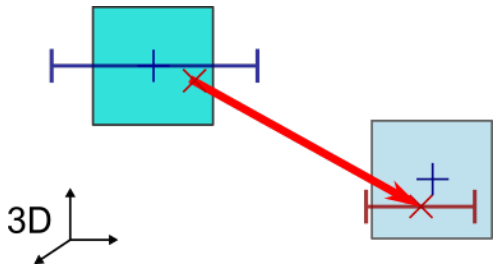


FIGURE 7.19 – Incertitudes tridimensionnelles. À gauche en bleu : incertitude liée à la coordonnée d'un point triangulé. À droite en rouge : incertitude liée à la mesure de vélocité métrique associée au pixel de départ.

Nous estimons que ces deux incertitudes bidimensionnelles se situent entre un dixième

et un tiers de pixel de par les résultats obtenus par simulation sur des données simulées (tableau 6.1) ainsi que la précision achevée lors de la sélection manuelle des points d'ancrage (coordonnées images des objets géo-référencés tels que les bâches, etc.). Les résultats avec ces deux bornes sont reportés sur les figures et 7.20 et 7.21. Sur ces figures, les cartes d'incertitude sont générées par la méthode de propagation analytique de la variance (méthode GUM) qui, pour rappel, donne l'analyse de sensibilités du modèle sténopé suivante :

$$\begin{aligned} \mu_X^2 &= (\tfrac{X}{B})^2\mu_B^2 + (\tfrac{B}{D})^2\mu_u^2 + (\tfrac{X}{D})^2\mu_D^2, \\ \mu_Y^2 &= (\tfrac{Y}{B})^2\mu_B^2 + (\tfrac{B}{D})^2\mu_v^2 + (\tfrac{Y}{D})^2\mu_D^2, \\ \mu_Z^2 &= (\tfrac{Z}{B})^2\mu_B^2 + (\tfrac{Z}{F})^2\mu_f^2 + (\tfrac{Z}{D})^2\mu_D^2, \end{aligned} \tag{7.7}$$

ainsi que via l'algorithme proposé plus haut (méthode GDOP). L'écart-type obtenu est ensuite doublé afin d'exprimer une incertitude élargie exprimée avec un niveau de confiance à 95%. Ces résultats montrent que les données métriques qui découlent de mesures sous-pixelliques ont une précision légèrement en dessous de la résolution de l'image (voir figure 5.11) en zone proches et qui croit sur les zones lointaines, et ce en toute composante "*East*", "*North*" et "*Up*" et avec les deux méthodes GUM et GDOP.

Tandis qu'avec la méthode GUM, l'incertitude croient linéairement avec la profondeur conformément au modèle linéaire utilisé, avec la méthode GDOP, nous notons une divergence sur les zones lointaines. Cette divergence théoriquement observée par traçage du modèle construit en fonction des angles de visée (figure 7.18) se situe au niveaux des sommets montagneux de haute altitude et une partie du glacier. Ces zones, visualisées grâce à un encodage couleur avec une échelle logarithmique, correspondent en effet à des angles de visée θ_g et θ_d, tels que définis sur le modèle de la figure 7.16, dont la somme tend vers π. Nous notons également que la croissance de l'incertitude métrique en fonction de l'incertitude pixellique est corrélée positivement avec la profondeur (c.-à-d. que la dilution de la précision est plus importante pour les zones lointaines à incrément de l'incertitude pixellique égal) et est plus marquée avec la méthode GDOP que la méthode GUM.

Nous reportons dans le tableau ci-après les estimations au niveau de la zone marquée par un I sur la figure 7.11 : Nous notons également que, pour la composante "*Up*", tandis que l'incertitude est moins importante sur les zones à incertitude finie, la zone de divergence est quant à elle plus importante.

7.6 Conclusion

Dans ce chapitre, nous avons présenté les de traitements tridimensionnels de la chaîne de traitements proposée. Ceux-ci reposent sur des briques fondamentales de reconstruction

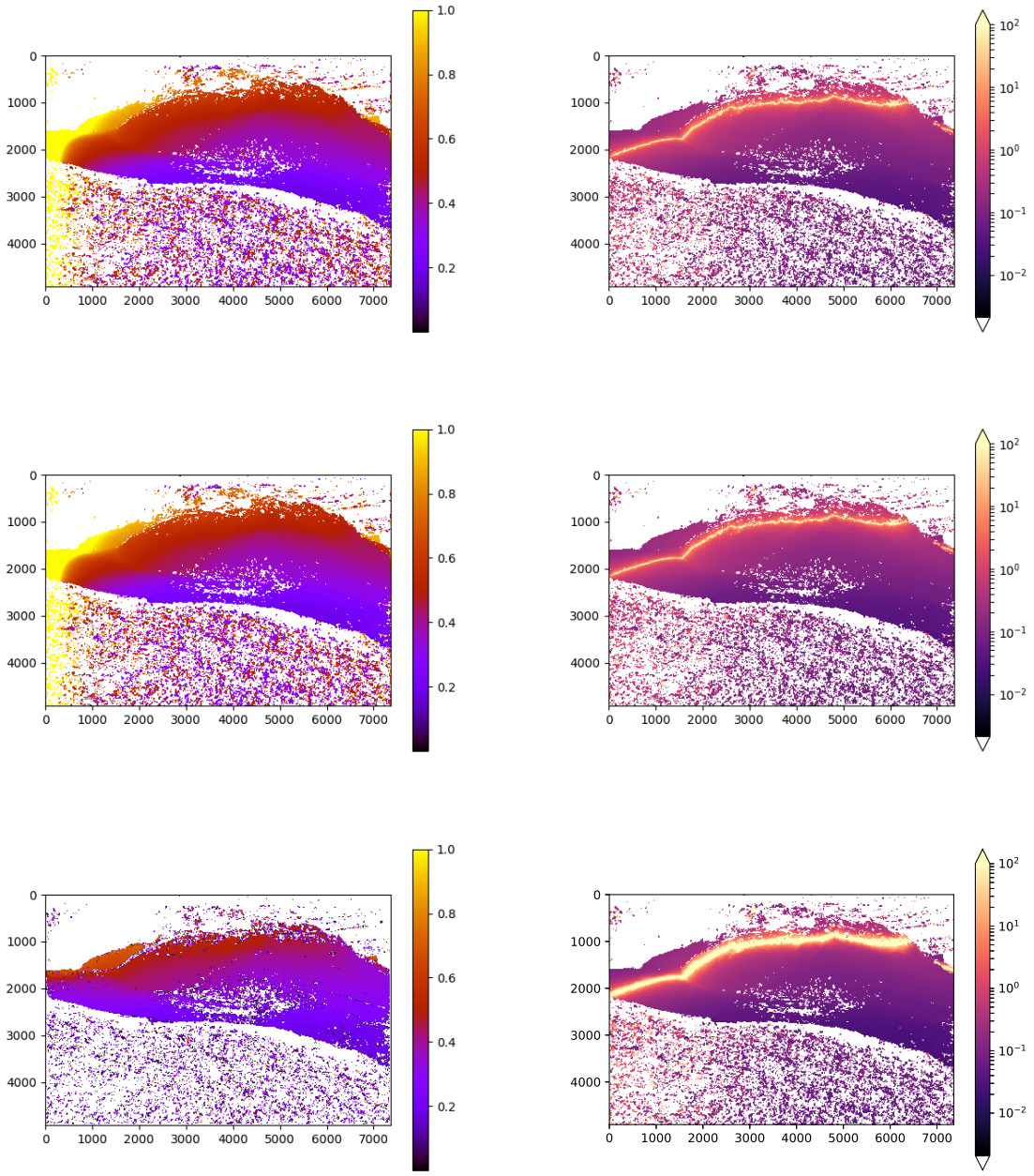


FIGURE 7.20 – Cartes d'incertitude sur la coordonnée d'arrivée du vecteur flot de scène projeté sur le modèle $\mathbb{X}_{MNT}$ obtenues avec $(\mu_u, \mu_v)_c = (0.1, 0.1)$ pixels pour la composante *East* (haut), *North* (milieu) et *Up* (bas). Gauche : propagation analytique de la variance (méthode GUM), échelle linéaire. Droite : propagation géométrique de la précision (méthode GDOP), échelle logarithmique. Résultats élargis (écart-type élargi expliquant 95% des réalisations) et en mètre.

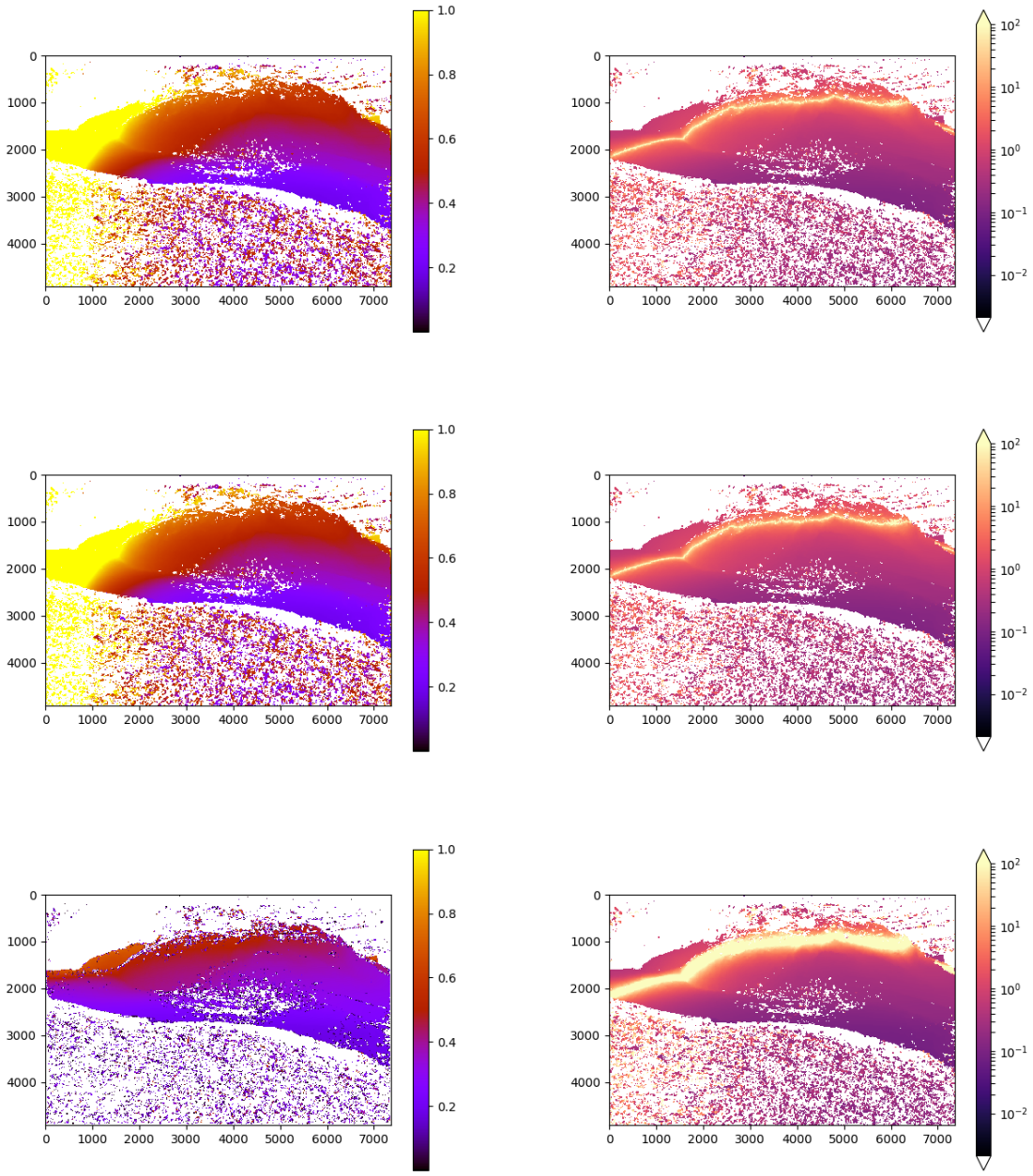


FIGURE 7.21 – Cartes d'incertitude sur la coordonnée des points tridimensionnels du modèle $\mathbb{X}_{MNT}$ obtenues avec $(\mu_u, \mu_v)_c = (0.33, 0.33)$ pixels pour la composante *East* (haut), *North* (milieu) et *Up* (bas). Gauche : propagation analytique de la variance (méthode GUM), échelle linéaire. Droite : propagation géométrique de la précision (méthode GDOP), échelle logarithmique. Résultats élargis (écart-type élargi expliquant 95% des réalisations) et en mètre.

	GUM		GDOP	
	$\mu = 0.1$	$\mu = 0.33$	$\mu = 0.1$	$\mu = 0.33$
Est	0.21	0.22	0.09	0.28
Nord	0.21	0.22	0.09	0.28
Up	0.20	0.20	0.06	0.20

TABLE 7.1 – Résultats obtenus avec les méthodes GUM et GDOP pour l'estimation de l'incertitude du vecteur vélocité métrique (flot de scène) pour une incertitude pixellique entre 0.1 pixel et 0.3 pixel. Résultats élargis (écart-type élargi expliquant 95% des réalisations) et en mètres.

stéréoscopique robuste. Tout d'abord, une détection et une description invariantes à la rotation ainsi qu'à la taille des points détectés de coins de Harris multi-échelles affinement adaptés, permet une mise en correspondance de couples stéréoscopiques. Les mesures épipolaires suivi d'une phase de calibration en aval nous ont ensuite permis de modéliser et de cartographier la scène observée par le système d'observation stéréoscopique sur un repère géo-référencé et d'y projeter, par triangulation, les mesures de vélocité pixellique en tout point. Afin de minimiser l'impact des erreurs liées aux ambiguïtés projectives et à l'incertitude des mesures 3D, nous proposons de modéliser la scène avec un modèle numérique de terrain unique en supposant que la scène ne subit pas de déformation importante au cours du temps (écoulement surfacique sans perte d'épaisseur perceptible). Ce modèle est estimé à partir d'un couple stéréoscopique disposant d'une bonne visibilité et dont la surface reconstruite minimise la distance point par point sur un ensemble de couples stéréoscopique de test. Le modèle retenu dispose d'une erreur de reprojection résiduelle moyenne égale à 3.10^{-4} pixel. Sous cette hypothèse le calcul du flot de scène métrique consiste alors à projeter les coordonnées de départ et d'arrivée en tout point via une via l'algorithme DLT de triangulation linéaire dont la solution est unique.

L'incertitude élargie liée à la mesure de flot de scène retenue est celle obtenue par la méthode GDOP et est estimée à 0.18m sur les composantes "*Est*" et "*Nord*" et à 0.13m sur la composante "*Up*" au niveau de la zone glaciaire marquée sur la figure 7.11 (moyenne des deux bornes). L'incertitude finale est enfin composée avec l'erreur-type expérimentale observée sur la moyenne de la vélocité résiduelle en zone fixes comme suit : Ces résultats montrent malheureusement que, pour une telle distance d'observation et

	Incertitude GDOP	Incertitude expérimentale	Incertitude composée
Est	0.18	0.8	0.82
Nord	0.18	0.8	0.82
Up	0.13	0.19	0.23

TABLE 7.2 – Estimation de l'incertitude composée et élargie avec un facteur 2 (niveau de confiance à 95%) pour les mesures de vélocité métrique en zone glaciaire proche, en mètres.

quoique l'erreur expérimentale reste centrée en 0, les séries temporelles de suivi glaciaire construites à partir de ces données risquent d'être entachées d'une erreur aléatoire ainsi

que de valeurs aberrantes trop importantes pour que les résultats soient exploitables et ce particulièrement pour un glacier à vélocité d'écoulement de l'ordre d'une dizaine de centimètres par jour. Ceci souligne encore une fois la nécessité de est donc de régulariser de telles mesures.

7.7 Perspectives

Étant donné que la scène observée est continue (pas d'objets déconnectés sur la scène), il serait intéressant d'affiner le modèle numérique de terrain obtenu en omettant les points tridimensionnels isolés dont la distance aux voisins les plus proches dépasse un seuil à fixer. Cela permettrait de générer des cartes cartographiques plus informatives (par exemple, la reprojection des pixels se situant entre la caméra et le début du glacier sur les cartes présentées dans ce chapitre seraient omises) et surtout d'optimiser le temps de calcul lors de la construction d'un ensemble de mesures bidate pour les traitements temporels.

Chapitre 8

Proposition de traitements en séries temporelles

Sommaire du chapitre

8.1 Introduction **133**

8.2 Données temporelles **134**

8.2.1 Données de modélisation temporelle de l'écoulement glaciaire 134

8.2.2 Mesures d'écoulement glaciaire non régularisées 135

8.3 Traitements temporels **140**

8.3.1 Méthode proposée 140

8.3.1.1 Fermeture temporelle 141

8.3.1.2 Résolution avec pénalisation de la variation totale 151

8.3.2 Paramétrisation, résolution et résultats 161

8.3.2.1 Paramétrisation et résolution 161

8.3.2.2 Résultats sur l'année 2018 167

8.4 Conclusion **167**

8.5 Perspectives **171**

8.1 Introduction

Dans ce chapitre, nous rentrerons plus en détail dans les méthodes adoptées afin de construire et traiter des séries temporelles de vélocité pixellique ou métrique à partir des mesures bidates monoscopiques (chapitre 6) et/ou stéréoscopiques (chapitre 7). Cette étape, dernière brique de la chaîne de traitements proposée et notre principale contribution, consiste en la régularisation des données de suivi glaciaire par ajustement à un modèle de cohérence temporelle augmenté de contraintes. Nous verrons tout d'abord des détails sur les séries temporelles non traitées et les expériences sur données simulées qui nous ont permis de poser les paramètres du modèle de fermeture temporelle à inverser. Nous verrons ensuite les différents résultats de régularisation générés sur les données acquises sur le site du glacier d'Argentière pour l'année 2018.

8.2 Données temporelles

8.2.1 Données de modélisation temporelle de l'écoulement glaciaire

Les données temporelles dont il est question dans ce chapitre font référence à l'ensemble des mesures bidates, i.e. la vélocité ou le déplacement du glacier. Afin de représenter le mouvement glaciaire, un formalisme Eulérien a été adopté. Comme stipulé dans le chapitre 4, celui-ci convient à la description d'un mouvement géodésique, sous-pixellique et calculé d'une manière dense par flot optique et est, de surcroît, plus facile à mettre en oeuvre qu'un formalisme Lagrangien. À chaque coordonnée du plan image choisi comme référence (gauche dans notre cas) sur lequel une mesure de flot optique a été calculée ou de flot de scène (scene-flow) re-projetée, il existe alors une série temporelle (données $\mathbf{t}$) que nous voulons régulariser à l'aide de l'ensemble des données multi-temporelles $\mathrm{A}^{\mathbf{2}}_{\mathbf{n}}\mathbf{t}$ et d'indicateurs sur la qualité des données impliquées et des mesures effectuées. Comme évoqué lors de la caractérisation des données temporelles issues de time-lapses (section 4.2), nous préconisons ce type de méthodes de par le fait que chaque mesure bidate dispose d'une définition théorique redondante. Pour rappel, celle-ci est définie comme la possibilité de modéliser une mesure de vélocité pixellique ou métrique entre deux dates comme une combinaison d'une ou plusieurs mesures de vélocité entre les deux dates impliquées. Quand les mesures sont effectuées après l'acquisition de l'ensemble des données à traiter, un modèle théorique peut donc permettre de prendre en compte cette redondance d'une manière pondérée afin de débruiter les données de suivi.

Deux différentes quantités peuvent être considérées afin de suivre l'évolution du déplacement glaciaire via un formalisme Eulérien :

— les séries temporelles de vélocité définies par :

$$\mathbf{F}_{\mathrm{LF}}(t_i) = F_{i,i+1} \qquad \forall i \in \{0, .., N-1\},$$

— les séries temporelles de position définies par :

$$\mathbf{F}_{\mathrm{CM}}(t_i) = F_{a,i} \qquad \forall i \in \{a+1, ..., a+N\},$$

où LF est l'acronyme de "*Leap-Frog*", qui désigne une approche en saute-mouton, et CM l'acronyme de "*Common-Master*", qui désigne une approche avec image maîtresse d'indice a. En l'absence de mesures manquantes, ces deux quantités sont en réalité équivalentes. En effet, il existe une bijection entre les deux types de séries temporelles puisque la deuxième quantité est la somme de Riemann de la première, telle que :

$$\mathbf{F}_{\mathrm{CM}}(t_k) = \sum_{i=a}^{k-1} \mathbf{F}_{\mathrm{LF}}(t_i). \tag{8.1}$$

8.2.2 Mesures d'écoulement glaciaire non régularisées

Les séries temporelles de vélocité à pas de temps journalier $\mathbf{F}_{\mathrm{LF}}$ obtenues via les briques fondamentales de la chaîne de traitements proposée sont illustrées sur les figures 8.1, 8.2 et 8.3. Elle s'appuient sur des vélocités moyennées spatialement au niveau du front du glacier (45°57'N, 6°58'E) entre les différentes images classées "glacier visible" (résultat de classification de la figure 6.3).

Les cartes de vélocité telles que celles illustrées sur la figure 7.11 se sont avérées assez reproductibles entre un couple d'images correctement classées "glacier visible" d'aspect photométrique et climatique similaire. Nous voyons cependant que, malgré une médiane mobile sur 10 jours assez cohérente, i.e. conforme avec un écoulement dans la direction Nord-Ouest et vers le bas (composantes "*East*" et "*Up*" négative et composante "*Nord*" positive), il existe un nombre assez élevé de valeurs aberrantes dans les séries temporelles de suivi journalier. Conformément aux mesures d'incertitude calculées dans le chapitre 7, ce bruit de mesure peut être expliqué à 95% par des incertitudes de ±0.82 m pour les composantes "*East*" et "*North*" et de ± 0.23 m pour la composante "*Up*".

Ces aberrations, qu'elles soient dues à une valeur manquante, à un bruit de mesure aléatoire dans les plages d'incertitude estimée ou à un biais systématique sur toute la zone spatiale (par exemple causé par une couche de brume non détectée), peuvent être détectable automatiquement via un des indicateurs suivants :

— l'indice de visibilité moyen entre les deux images impliquées,
— l'amplitude de la vélocité résiduelle moyenne en zones fixes,
— la variance de la vélocité résiduelle en zones fixes qui, comme déjà stipulé, indique un bruit de mesure par opposition à un vrai déplacement (fonte de glace, occlusions, etc.) en ces zones, etc.
— une des erreurs de symétrie temporelle (bidate, multi-date, etc.) telles que celles introduites dans le chapitre 4.

Par exemple, l'erreur de fermeture temporelle bidate illustrée sur les figures 8.1, 8.2 et 8.3 ("*Forward-Backward temporal-closure error on glacier*" dans la légende) et 8.4 pour une date i est définie comme suit :

$$\epsilon_{\mathrm{LF}_{\Delta t=1}} = |\mathbf{F}_{i,i+1} + \mathbf{F}_{i+1,i}|. \tag{8.2}$$

On peut voir sur la figure 8.4 un exemple de cartes d'erreurs de fermetures temporelles en pixels. En zone glaciaire, celles-ci indiquent des erreurs pixelliques horizontales et verticales situées entre 0.1 et 0.2 pixel et qui se situent bien dans la plage d'incertitude pixellique admise dans le chapitre 7 lors de l'estimation de l'incertitude métrique. Ces erreurs sont légèrement plus importantes sur les zones mobiles et sur les zones à haute altitude plus souvent bruitées ou occultées par des phénomènes météorologiques et par-

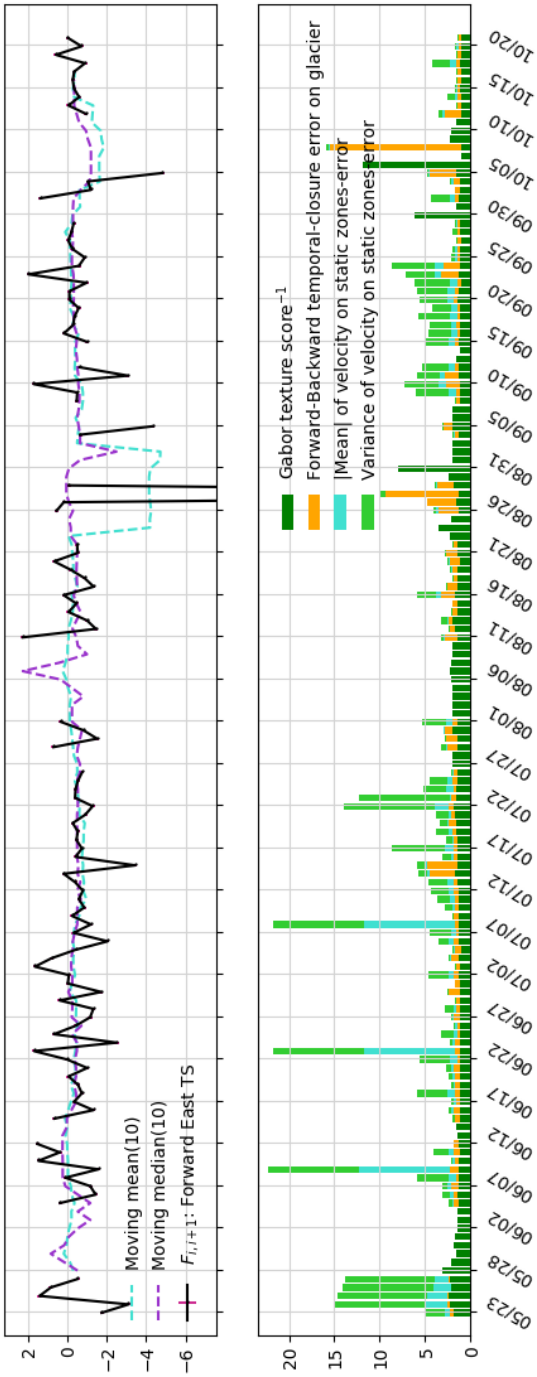


FIGURE 8.1 – Série temporelle de vélocités à pas de temps journalier $\mathbf{F}_{\mathrm{LF}_E}$ (en mètre) obtenue en zone frontale du glacier d'Argentière : Haut : Composante "*East*" dont l'incertitude élargie a été estimée à 0.82 m et ses moyenne (violet) et médiane (bleu) mobiles sur 10 jours. Bas : Certains indicateurs (poids) associés aux données et aux mesures bidates : score de visibilité moyen (vert foncé), erreur de fermeture temporelle bidate en mètre (orange), vélocité résiduelle moyenne en zones fixes en mètre (bleu) et variance de la vélocité résiduelle moyenne en zones fixes en mètre carré (vert clair).

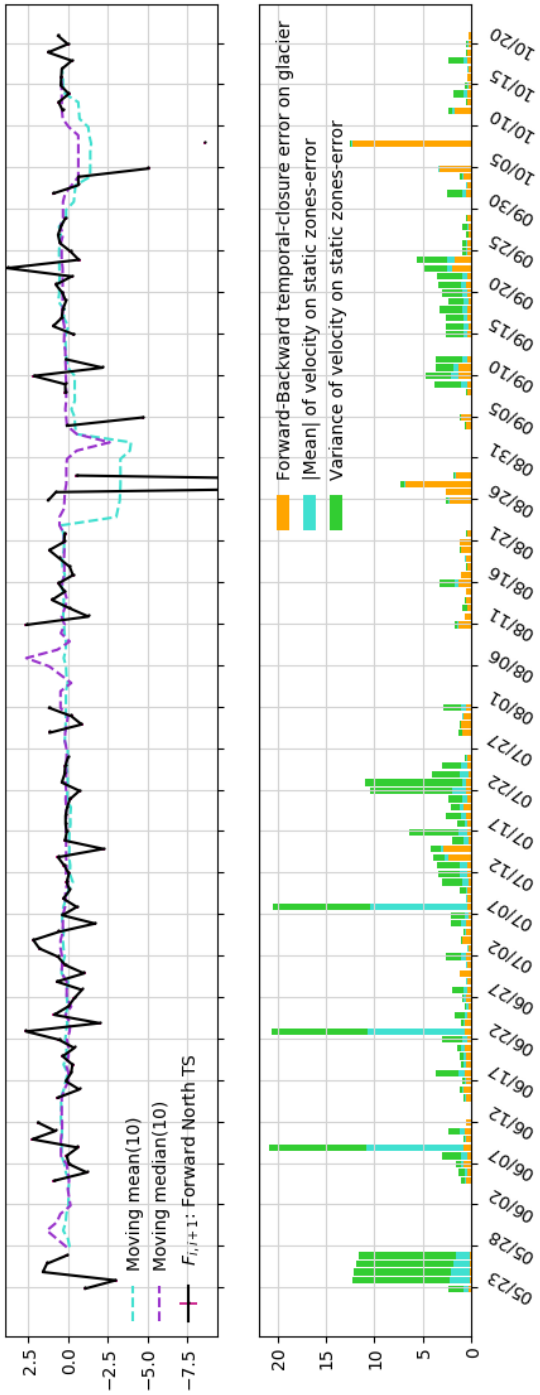


FIGURE 8.2 – Série temporelle de vélocités à pas de temps journalier $\mathbf{F}_{\mathrm{LF}_N}$ (en mètre) obtenue en zone frontale du glacier d'Argentière : Haut : Composante "*North*" dont l'incertitude élargie a été estimée à 0.82m et ses moyenne (violet) et médiane (bleu) mobiles sur 10 jours. Bas : Certains indicateurs (poids) associés aux mesures bidates : erreur de fermeture temporelle bidate en mètre (orange), vélocité résiduelle moyenne en zones fixes en mètre (bleu) et variance de vélocité résiduelle moyenne en zones fixes en mètre carré (vert clair).

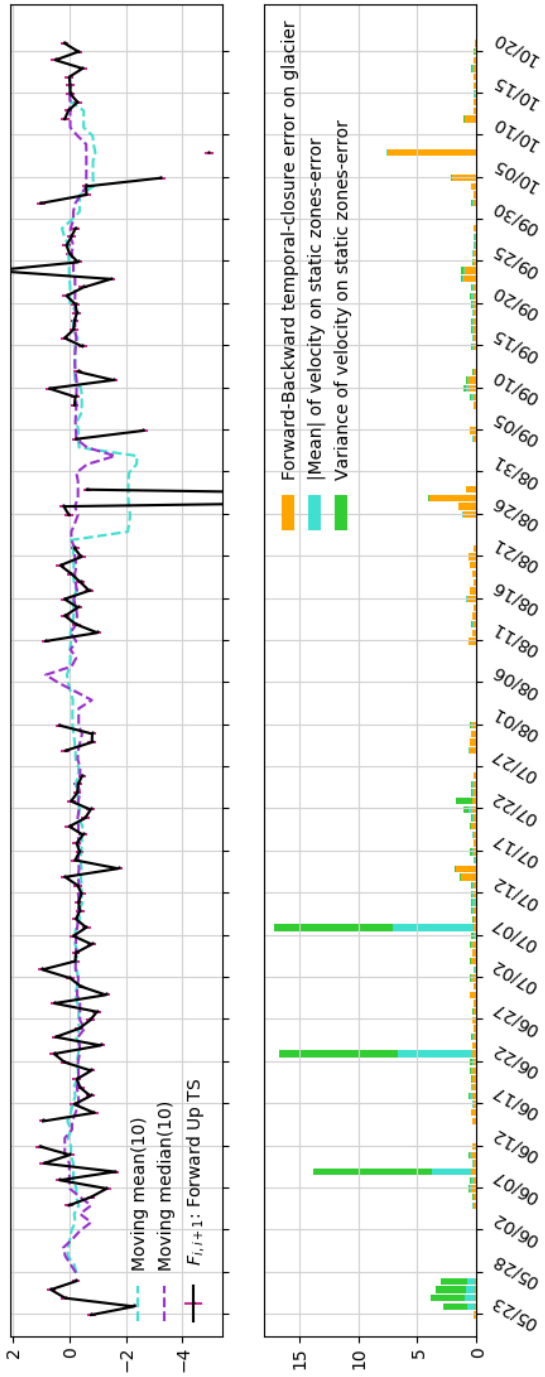


FIGURE 8.3 – Série temporelle de vélocités à pas de temps journalier $\mathbf{F}_{\mathrm{LF}_U}$ (en mètre) obtenue en zone frontale du glacier d'Argentière : Haut : Composante "*Up*" dont l'incertitude est estimée à 0.23m et ses moyenne (violet) et médiane (bleu) mobiles sur 10 jours. Bas : Certains indicateurs (poids) associés aux mesures bidates : erreur de fermeture temporelle bidate en mètre (orange), vélocité résiduelle moyenne en zones fixes en mètre (bleu) et variance de vélocité résiduelle moyenne en zones fixes en mètre carré (vert clair).

ticulièrement la brume. Au niveau du ciel, le flot optique n'est pas capable de capturer le déplacement et on peut observer une divergence de la mesure d'erreur de fermeture temporelle. Sur le front glaciaire et sur certaines zones fixes, une fermeture temporelle non satisfaite peut aussi être indicatrice d'une chute de Séracs ou d'une fonte de glace. La mesure bidate capture en effet difficilement la vélocité lorsque l'objet observé disparaît de l'une des images.

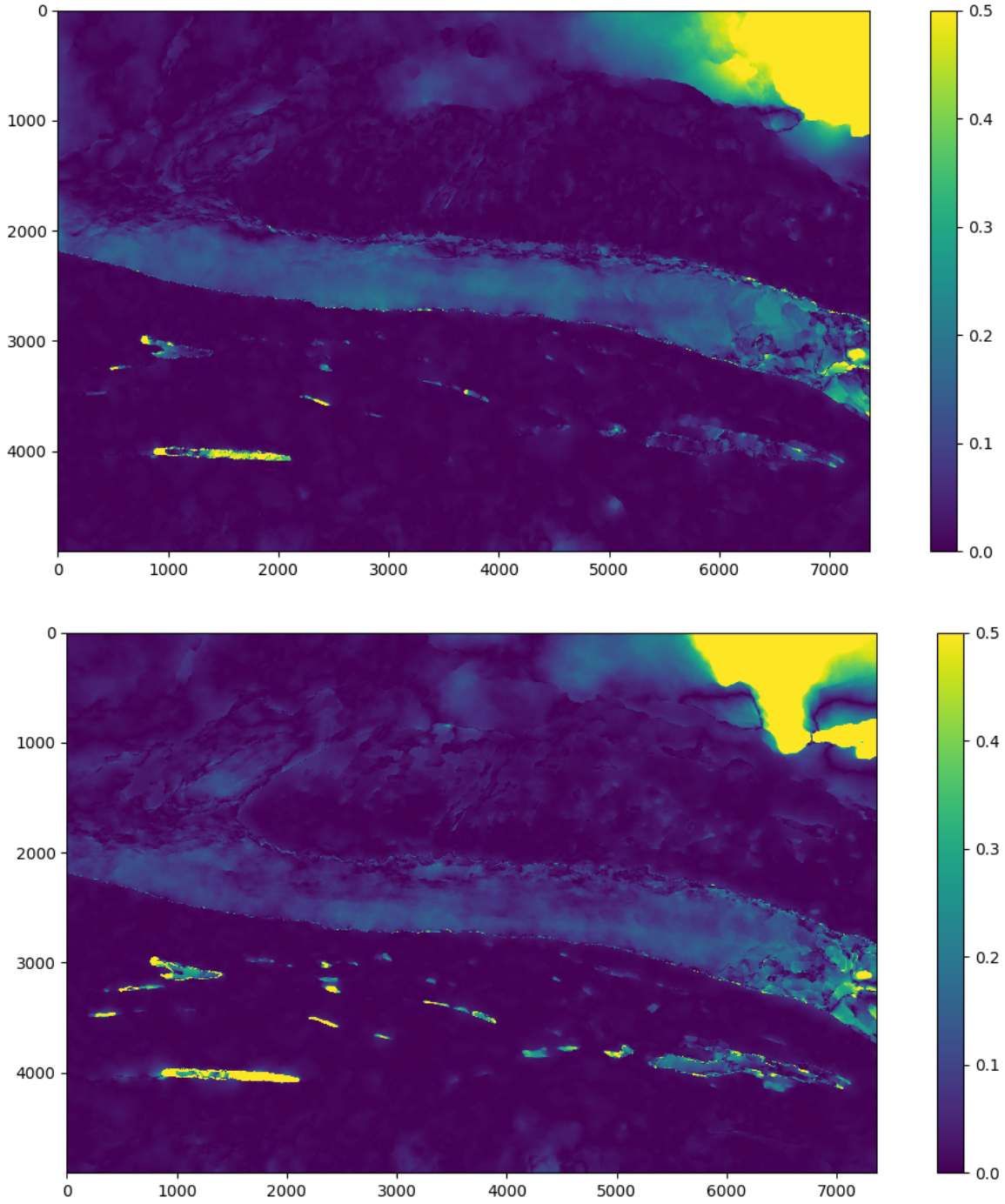


FIGURE 8.4 – Exemples de cartes d'erreurs de fermetures temporelles bidate en pixel. Haut : composante horizontale. Bas : composante verticale.

8.3 Traitements temporels

Si nous estimons que le bruit d'une mesure de vélocité est absolu, le signal a plus de chance d'être noyé dans le bruit avec un pas de temps court. De plus, nous notons que les images d'aspect similaire ne sont pas forcement acquises à un jour d'intervalle. Afin d'exploiter au mieux les données multi-temporelles dont nous disposons, nous avons donc noté impératif de :

— maximiser le nombre d'observations impliquant une même image tout en prenant en compte la redondance dans la définition des mesures,

— recréer, à partir de ces mesures, des séries temporelles de vélocité $\mathbf{F}_{\mathrm{LF}}$ ou de position $\mathbf{F}_{\mathrm{CM}}$ en accordant plus d'importance aux mesures de flot de scène impliquant de meilleurs indices de qualité.

8.3.1 Méthode proposée

La méthode proposée repose sur la construction d'une fonction objectif englobant :

— une fonction de coût linéaire (et donc convexe) en fonction de l'erreur de fermeture temporelle telle que celle utilisée par la méthode SBAS (Small BAseline Subset),

— une pondération prenant en compte des mesures #n 2D associées à chaque image, à savoir la moyenne d'indices de visibilité calculés en pré-traitements, ainsi que des mesures #n (2D+t) associées à chaque mesure de vélocité bidate, à savoir les déplacements résiduels en zones fixes et sa variance,

— une contrainte d'inégalité exprimant l'avancement du glacier dans la direction Nord-Ouest et vers le bas,

— un terme de pénalisation de la variation totale des séries temporelles de suivi,

à minimiser en tout point, sur toutes les composantes de déplacement : x et y si les séries temporelles sont pixelliques (monoscopiques) et "*East*", "*North*", "*Up*" si elles sont métriques (stéréoscopiques). La fonction objectif globale est modélisée de la manière suivante :

$$\hat{\mathbf{F}}_\tau = \underset{\mathbf{F}_\tau}{argmin}\ ||W_0\mathbf{F}_0 - W_0\Phi_\tau(\mathbf{F}_\tau)||_L + \gamma||\nabla\mathbf{F}_\tau||^2_{L'},$$

$$\text{sujette à : } \begin{cases} \mathbf{F}_\tau \geq 0 & \text{si composante N,} \\ \mathbf{F}_\tau \leq 0 & \text{si composante E ou U,} \end{cases} \tag{8.3}$$

ou

$$\text{sujette à } \quad \mathbf{F}_\tau \geq 0 \quad \text{si composante x ou y,}$$

avec Φ_τ l'application régissant les différentes contraintes de cohérence temporelle entre les différents arrangements de dates où une image a été retenue et γ la pondération donnée au terme de pénalisation de la variation totale de la solution. $\mathbf{F}_0 \in \mathbb{D}_{\Phi_\tau}$ $\forall \tau$, est le vecteur formé d'une des composantes de l'ensemble des données $\mathrm{A}^{\mathbf{2}}_{\mathbf{n}}\mathbf{t}$ observées telles que définies dans les tableaux 1 puis 5.2 et W_0 est la matrice de précision qui permettra de pondérer et éventuellement décorréler ces observations. Enfin, le vecteur $\mathbf{F}_\tau = \mathbf{F}_{\mathrm{LF}}$ ou $\mathbf{F}_{\mathrm{CM}}$ constitue la mesurande, ou variable de décision, que l'on cherche à estimer par inversion de la fonctionnelle globale sus-définie, respectivement avec un modèle de fermeture temporelle de type "Leap-Frog" (en saute-mouton) ou "Common-Master" (avec image maîtresse).

8.3.1.1 Fermeture temporelle

Les opérateurs de fermeture temporelle envisagés reposent sur les contraintes inspirées de la méthode SBAS et ses variantes (introduites dans le chapitre 4) que, pour rappel, nous avons redéfinies comme suit :

— les contraintes du type "Leap-Frog" :

$$\begin{gathered}\forall\, a \in \{1..n\}, \forall\, \Delta t \in \{1..n-a\} \text{ et } \forall\, ds \in \{1..a-1\} : \\ -\mathbf{F}_{a,a+\Delta t} + \sum_{i=a}^{a+\Delta t-1} \mathbf{F}_{i,i+1} = \epsilon_{\mathrm{LF}}, \\ \mathbf{F}_{a,a-ds} + \sum_{i=a-1}^{1} -\mathbf{F}_{i,i+1} = \epsilon_{\mathrm{LF}},\end{gathered} \tag{8.4}$$

— les contraintes du type "Common-Master" :

$$\begin{gathered}\forall\, a \in \{1..n\}, \forall i \in \{a..n\}, \forall \Delta t \in \{1..n-i\} : \\ \mathbf{F}_{i,i+\Delta t} + \mathbf{F}_{a,i} - \mathbf{F}_{a,i+\Delta t} = \epsilon_{\mathrm{CM}}.\end{gathered} \tag{8.5}$$

Ces contraintes sont particulièrement intéressantes, car celles-ci sont modélisables via des applications linéaires Φ_τ et sont adaptées au formalisme Eulérien. En effet, comme dans la méthode SBAS, l'ensemble des contraintes du type "Leap-Frog" et "Common-Master" considérées entre les différentes images d'un time-lapse de N dates peuvent être formulées sous forme de multiplication matricielle avec une matrice ternaire, formée uniquement de 0, de 1 et de -1, équivalente à une matrice d'adjacence qui lierait les entrées de $\mathbf{F}_\tau$ à celles de $\mathbf{F}_0$. Les applications linéaires Φ_{LF} et Φ_{CM} associées aux contraintes du type "Leap-Frog" et "Common-Master" prennent respectivement en entrée $\mathbf{F}_{\mathrm{LF}} \in \mathbb{M}_{\mathrm{LF}}$ et $\mathbf{F}_{\mathrm{CM}} \in \mathbb{M}_{\mathrm{CM}}$ et possèdent le même espace d'arrivée $\mathbb{D}_{\Phi_\tau} = \mathbb{R}^{(N-m)(N-m-1)}$, avec $m = N - n$ le nombre d'images manquantes dans le time-lapse traité.

Dans ce qui suit, nous étudierons le conditionnement des deux problèmes de minimisation qui découlent de ces deux applications avec un cas illustratif.

Cas illustratif Soit un time-lapse formée de $N = 19$ images sur lequel nous voulons générer des séries temporelles de suivi constituées de 18 estimations de vélocité (modèle "Leap-Frog") ou de position (modèle "Common-Master") bidates en tous pixels triangulés. Sur le graphe de la figure 8.5 on peut voir que les observations qui serviront à leur régularisation seront constituées de 342 mesures de flot de scène illustrées par des flèches (oranges) à double sens.

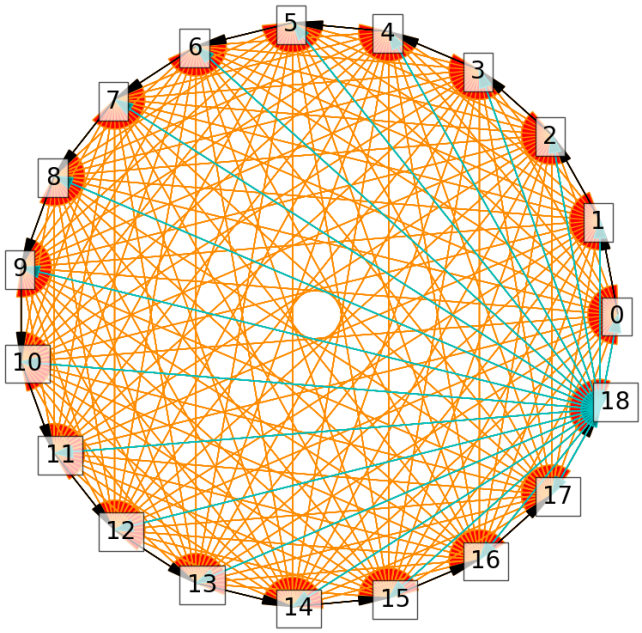


FIGURE 8.5 – Graphes d'indices pour un time-lapse à 19 dates. 342 flèches orange (et vertes) à double sens désignent les observations bidates et 18 flèches noires désignent les données temporelles de suivi de la vélocité ou de position bidate recherchées.

Les deux applications linéaires Φ_{LF} et Φ_{CM} sont alors définies comme une multiplication matricielle avec les matrices ternaires, respectivement, comme suit :

$$\begin{pmatrix} F_{0,1} \\ F_{0,2} \\ \vdots \\ F_{0,18} \\ F_{1,0} \\ F_{1,2} \\ \vdots \\ F_{18,16} \\ F_{18,17} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ & & & \vdots & & & & & \\ 1 & 1 & 1 & \cdots & 1 & 1 & 1 & 1 & 1 \\ -1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ & & & \vdots & & & & & \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} F_{0,1} \\ F_{1,2} \\ F_{2,3} \\ \vdots \\ F_{16,17} \\ F_{17,18} \end{pmatrix}, \tag{8.6}$$

et

$$\begin{pmatrix} F_{0,1} \\ F_{0,2} \\ \vdots \\ F_{0,18} \\ F_{1,0} \\ F_{1,2} \\ \vdots \\ F_{18,16} \\ F_{18,17} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ & & & \vdots & & & & & \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ & & & \vdots & & & & & \\ 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} F_{0,1} \\ F_{0,2} \\ F_{0,3} \\ \vdots \\ F_{0,16} \\ F_{0,17} \\ F_{0,18} \end{pmatrix}. \tag{8.7}$$

Les deux matrices d'adjacence associées aux équations 8.6 et 8.7 sont illustrées sur la figure 8.6 et leurs valeurs singulières sur la figure 8.7.

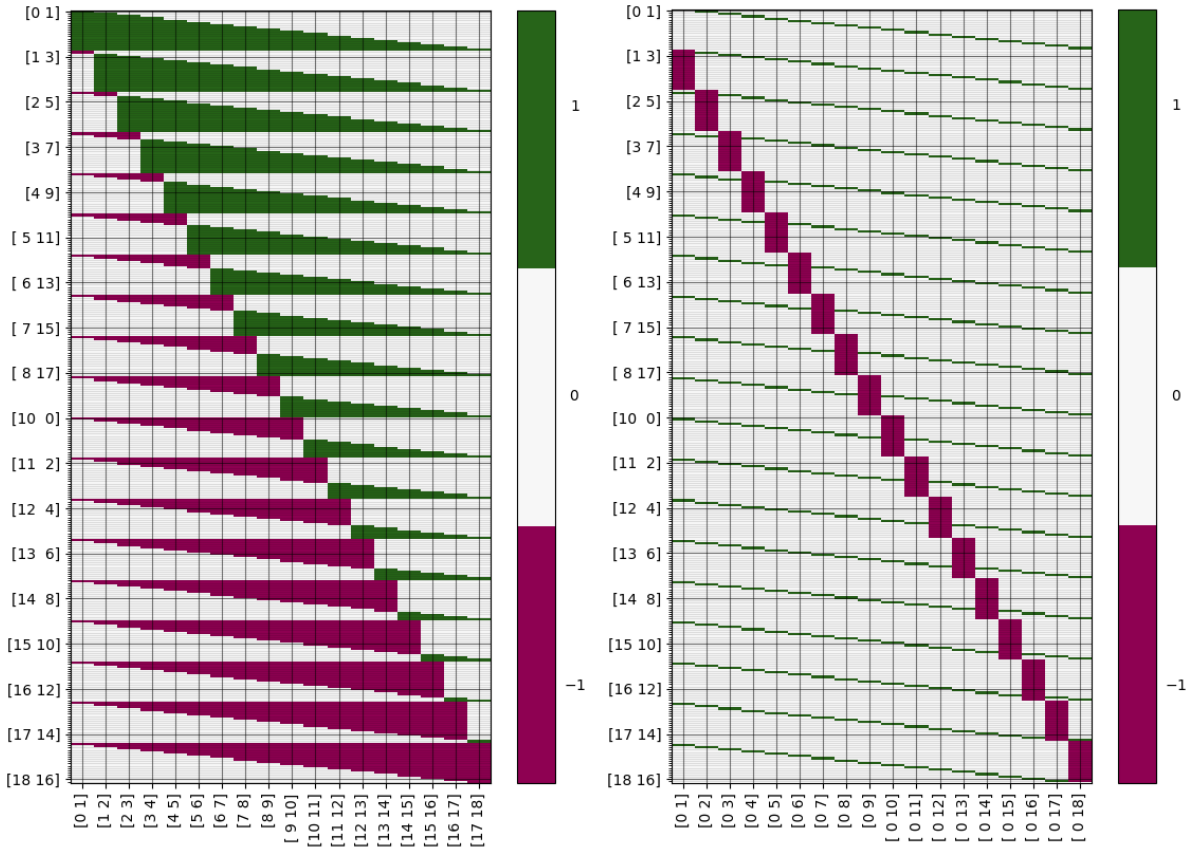


FIGURE 8.6 – Matrices de tailles 342×18 modélisant les applications Φ_{LF} et Φ_{CM} pour un time-lapse à 19 dates. L'ordre des lignes correspond à l'ordre des entrées dans le vecteur $\mathbf{F}_0$ tel qu'illustré dans les systèmes d'équations 8.6 et 8.7. Seulement une ligne sur 20 est possède une légende pour plus de lisibilité.

D'une manière générale, en l'absence de données manquantes, les matrices obtenues seront sur-déterminées, car le nombre de contraintes dépasse la dimension de $\mathbf{F}_\tau$. Puisque les deux opérateurs linéaires sont à rang plein, il est d'usage (tel que la version originale de la méthode SBAS), de minimiser la norme $L = 2$ de l'erreur d'ajustement au modèle via la pseudo-inverse analytique ou encore une décomposition en valeurs singulières de

la matrice ternaire. Le coefficient de conditionnement associé au problème de moindres carrés, tel que défini dans la section 1.2.1, est respectivement égal à $12,06$ et $4,35$ ce qui suggère que la mesurande ne sera pas trop affectée par un bruit aléatoire faible dans les observations.

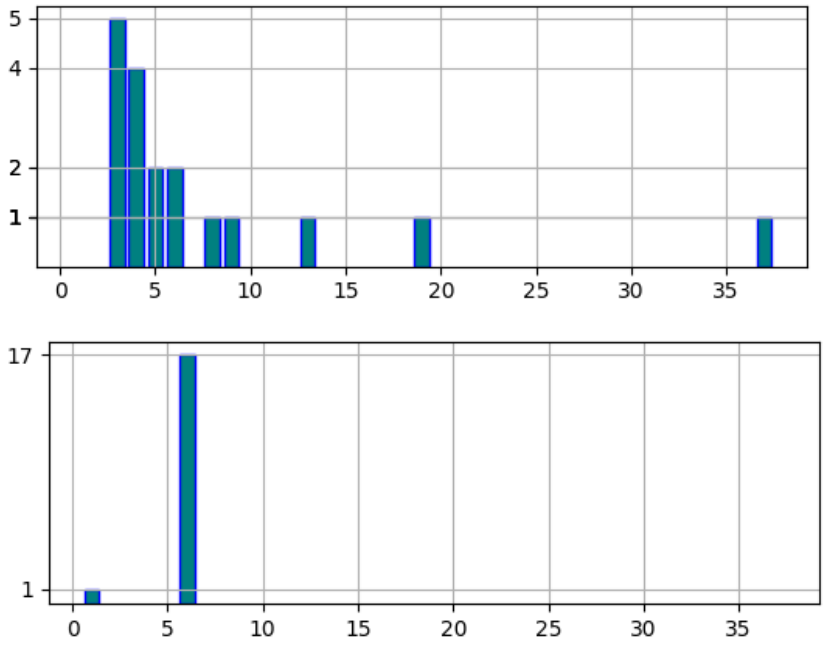


FIGURE 8.7 – Histogrammes de valeurs singulières associées aux applications linéaires Φ_{LF} (haut) et Φ_{CM} (bas) d'un time-lapse à 19 dates.

Décorrélation temporelle En considérant la décorrélation des images dans le time-lapse avec le temps, les contraintes de fermeture sur une boucle temporelle de taille supérieure à une portée donnée doivent être omises. En effet, la distance temporelle entre les images à l'intérieur de la zone de fermeture temporelle est idéalement suffisamment longue pour augmenter le signal (le déplacement du glacier), mais assez courte pour préserver les caractéristiques de suivi. Cette décorrélation (mise en évidence sur la figure 6.4) existe, comme nous l'avons déjà mentionné, à cause des changements photométriques et climatiques ; réflectance du glacier, fonte et chute de neige, neige balayée par le vent, etc. D'autre part, de bonnes caractéristiques de corrélation apparaissent de temps à autre aux mêmes endroits et avec des formes similaires sans pour autant désigner la même matière physique. Par exemple, les crevasses s'ouvrent régulièrement en amont d'une chute de séracs. Pour ces raisons, et comme le glacier d'Argentière est un glacier tempéré avec une vitesse relativement élevée, nous avons considéré, empiriquement, une portée temporelle maximale fixée à ± 5 jours. On obtient alors les matrices d'adjacence illustrées sur les figures 8.8 dont les valeurs singulières sont illustrées sur la figure 8.9 (coefficients de conditionnement respectifs à 5 and 7.79).

Nous constatons que la réduction de la portée temporelle maximale ne fait pas diver-

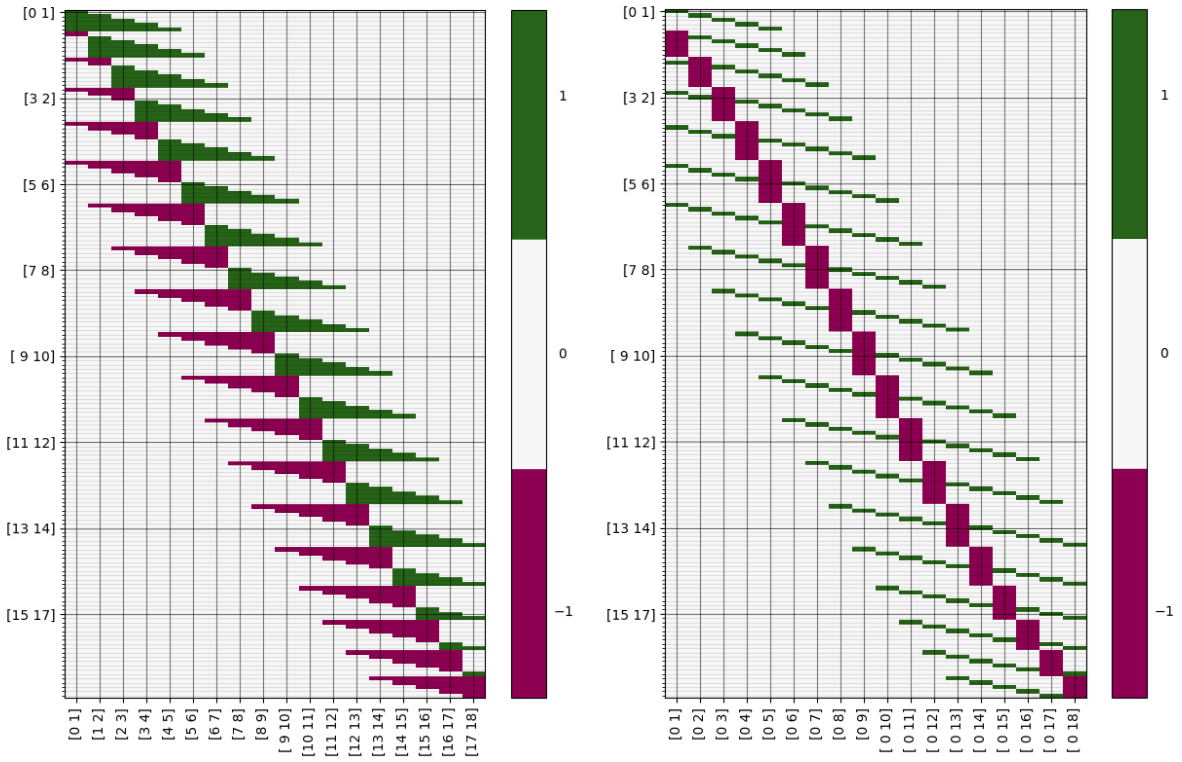


FIGURE 8.8 – Matrices de tailles 160×18 modélisant les applications Φ_{LF} et Φ_{CM} pour un time-lapse à 19 dates et une portée temporelle maximale de ± 5 jours. L'ordre des lignes correspond à l'ordre des entrées dans le vecteur $\mathbf{F}_0$.

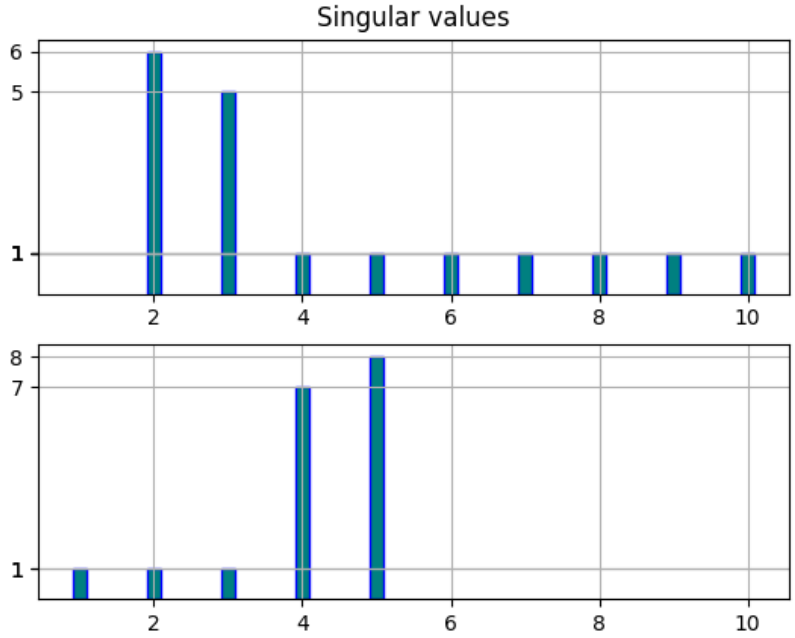


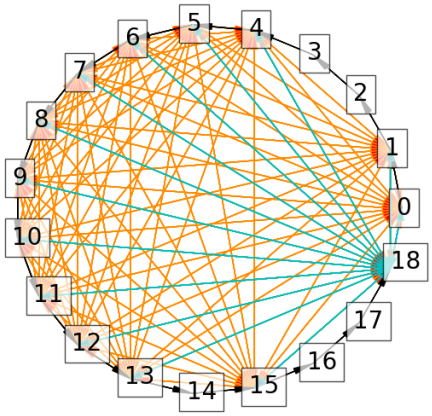
FIGURE 8.9 – Histogrammes de valeurs singulières associées aux applications linéaires Φ_{LF} (haut) et Φ_{CM} (bas) d'un time-lapse à 19 dates après réduction de la portée temporelle maximale.

ger le coefficient de conditionnement de l'opérateur fermeture temporelle. D'autre part, hormis le fait d'éviter les erreurs de décorrélation, cette réduction apporte aussi des avantages informatiques et analytiques remarquables : le nombre d'observations à calculer est considérablement réduit et les matrices à inverser ne possèdent plus de lignes ni de colonnes non-éparses, ce qui allège la résolution de l'équation normale impliquant le calcul de $\Phi_\tau^T\Phi_\tau$ ou $\Phi_\tau\Phi_\tau^T$ qui seraient systématiquement denses dans le cas contraire.

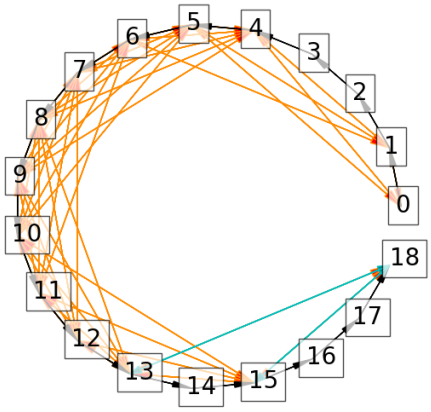
Observations manquantes En considérant maintenant que les observations liées à des images ayant été rejetées en pré-traitement sont manquantes, les contraintes (lignes) incluant leurs index doivent être omises. Ceci crée des colonnes nulles dans les matrices Φ_{LF} et Φ_{CM} et réduit leur rang. Si on considère par exemple que les images 2, 3, 14, 16 et 17 n'ont pas été classifiées comme contenant de l'information utile, nous obtenons alors les graphes d'indices de la figure 8.10 et des opérateurs linéaires avec un rang de 13 (sur 18 colonnes).

Dans ces conditions, Φ_{LF} et Φ_{CM} restent sur-déterminées et possèdent un coefficient de conditionnement infini (de par leur rang incomplet), l'espace $\mathbb{D}_{\Phi_\tau}$ rétréci et, sans un pré-conditionnement des espaces $\mathbb{M}_{\mathrm{LF}}$ and $\mathbb{M}_{\mathrm{CM}}$ via une matrice de pondération, des contraintes supplémentaires et un terme de régularisation dans la fonction objectif globale, les minima retrouvés peuvent ne pas être cohérents.

Afin de tester ce cas de figure, les séries temporelles $\mathbf{F}_{\mathrm{LF}}$ et $\mathbf{F}_{\mathrm{CM}}$ de la figure 8.11 ont



(a) Portée temporelle maximale : 182 observations



(b) Portée temporelle réduite à $\pm$ 5 jours : 94 observations

FIGURE 8.10 – Graphes d'index pour le traitement d'un time-lapse à $N = 19$ dates et $n = 14$ images retenues (images manquantes : 2, 3, 16, 14, 17).

été générées par la simulation d'observations multi-temporelles $\mathbf{F}_0 \in \mathrm{A}^{\mathbf{2}}_{\mathbf{14}}\mathbf{t}$ en supposant une distribution aléatoire normale (Gaussienne) des mesures de vélocité avec une mise à l'échelle par multiplication avec la portée temporelle factice, suivi d'une inversion du modèle par décomposition en valeurs singulières. Nous constatons alors que les valeurs correspondant aux observations manquantes sont complétées par une constante dans $\hat{\mathbf{F}}_{\mathrm{LF}}$ et par des estimations nulles dans $\hat{\mathbf{F}}_{\mathrm{CM}}$, respectivement résultats de l'inversion des opérateurs Φ_{LF} et Φ_{CM}. Nous notons aussi que la relation entre les séries temporelles de suivi de vélocité $\hat{\mathbf{F}}_{\mathrm{LF}}$ et de position $\hat{\mathbf{F}}_{\mathrm{CM}}$ n'est plus une bijection (somme de Riemmann).

Pour illustrer davantage la manière dont les solutions pourraient être affectées par les observations manquantes, nous avons simulé un problème où seules les images 0, 4, 10 et 18 sur les 19 images du time-lapse factice ont été retenues. En réalité, ce cas de figure extrême peut arriver par exemple suite un événement neigeux prolongé faussant la phase de classification de la visibilité intra time-lapse. Pour m images rejetées en prétraitement sur N et pour une portée temporelle maximale, nous avons $2Nm - m(m+1)$ contraintes de moins, auquel cas le sous-système impliqué sera sous-déterminé. Le rang des opérateurs linéaires dépendra alors de m et de l'adjacence des indices manquants. Les matrices d'adjacence obtenues pour ce cas de figure sont illustrées sur la figure 8.12.

Dans cet exemple, le sous-système sous-déterminé est de rang incomplet égal à 3 (sur 12 lignes). Il existe alors une infinité de solutions satisfaisant une erreur résiduelle de fermeture temporelle ϵ_{LF} et ϵ_{CM} nulles. Cet ensemble de solutions forme une variété affine de dimension $12 - 3 = 9$ qui constitue le noyau de Φ_τ. Le minimum avec la norme la moins importante est alors retenu via décomposition en valeurs singulières tronquée au

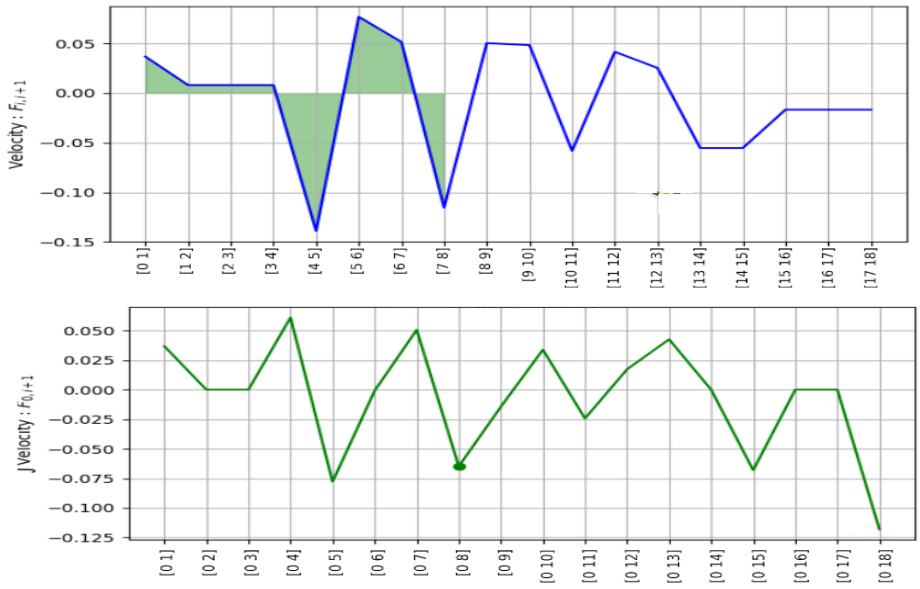


FIGURE 8.11 – Séries temporelles simulées de vélocités "Leap-Frog" (haut) et de positions "Common-Master" (bas). La valeur de $\hat{\mathbf{F}}_{\mathrm{CM}_{0,8}}$ n'est pas égale à l'aire de $\hat{\mathbf{F}}_{\mathrm{LF}}$ entre les 8 premiers index à cause des observations manquantes sur les dates 2 et 3.

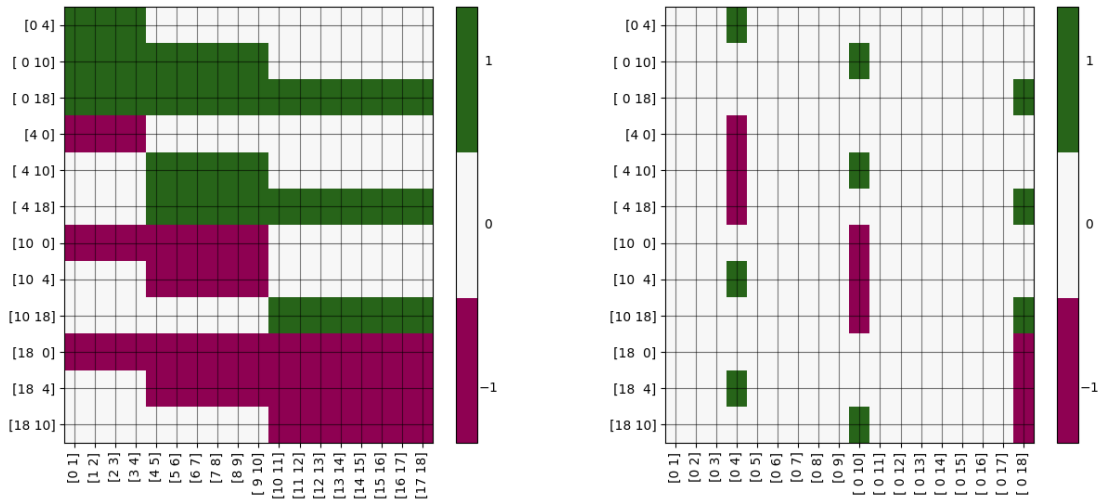


FIGURE 8.12 – Matrices de tailles 12×18 modélisant les applications Φ_{LF} et Φ_{CM} pour un time-lapse à 19 dates avec 15 observations manquantes et une portée temporelle limitée à $\pm$ 5 : cas sous-déterminé.

rang (tSVD). Le problème résolu est alors le suivant :

$$\hat{\mathbf{F}}_\tau = \underset{\mathbf{F}_\tau}{argmin}\ ||\mathbf{F}_\tau||_2^2,$$

sujette à :

$$||\mathbf{F}_0 - \Phi_\tau \mathbf{F}_\tau||_2^2 = 0.$$

Les minima obtenus sont illustrés sur la figure 8.13.

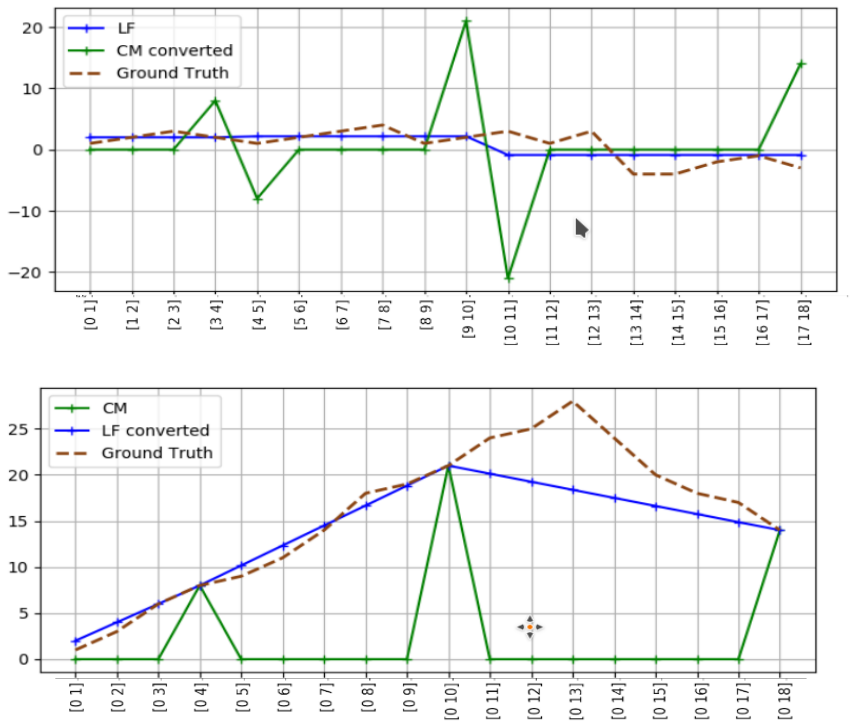


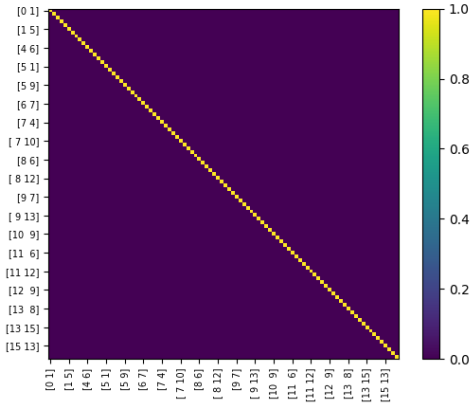
FIGURE 8.13 – Minima de l'inversion des opérateurs Φ_{LF} (bleu) et Φ_{CM} (vert) dans le cas sous-déterminé. Haut : Séries de suivi de vélocité, bas : Séries de suivi de position.

Cette expérience a montré que dans ce cas de figure, la solution de la minimisation de ϵ_{LF} donne un résultat plus réaliste, car plus proche de la vérité terrain qui a servi à simuler le problème.

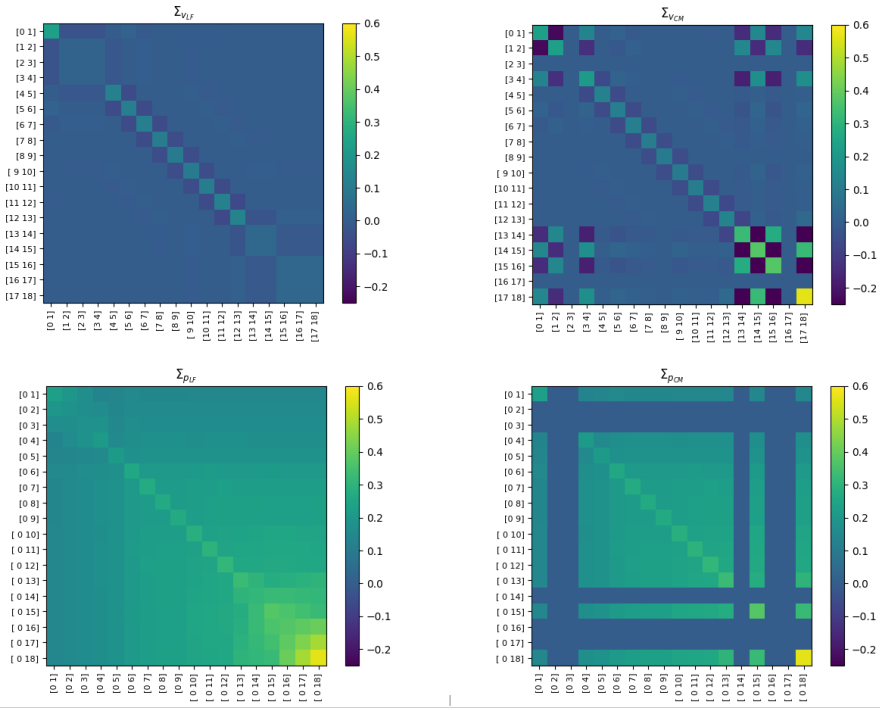
Propagation d'incertitude Soit Σ_0 la matrice de covariance décrivant la variance, et donc de l'incertitude, dans les mesures multi-temporelles $\mathbf{F}_0$. Nous supposons que celle-ci possède des termes de covariance nuls puisque les mesures de flot optique sont faites d'une manière indépendante (bruit de mesure indépendant). Une estimation de la propagation de cette incertitude à travers l'inversion des modèles linéaires de cohérence temporelle peut être réalisée comme suit :

$$\Sigma_\tau = \Phi_\tau^\dagger\ \Sigma_0\ \Phi_\tau^{\dagger^T}, \tag{8.8}$$

avec $\Phi_\tau^\dagger$ la pseudo-inverse de Φ_τ. Si nous considérons que les observations factices de l'exemple de la figure 8.10 (cas sur-déterminé donc $\Phi_\tau^\dagger = (\Phi_\tau^T \Phi_\tau)^{-1}\Phi_\tau^T$), disposant d'une portée temporelle maximale égale à $\pm$ 5 jours, sont mesurées avec une incertitude (écart-type) égale à 1 (mètre ou pixel), nous obtenons les matrices de covariances des mesurandes illustrées sur la figure 8.14.



(a) Matrice de covariance simulée associée à $\mathbf{F}_0$



(b) Matrice de covariance des mesurandes $\hat{\mathbf{F}}_{\mathrm{LF}}$ (haut) et $\hat{\mathbf{F}}_{\mathrm{CM}}$ (bas) obtenues par inversion du modèle Φ_{LF} (gauche) et Φ_{CM} (droite).

FIGURE 8.14 – Illustration de la propagation d'une incertitude unitaire à travers l'inversion des deux modèles de fermeture temporelle Φ_{LF} et Φ_{CM} proposés.

Ces résultats suggèrent que les séries temporelles de vélocité sont moins affectées par

l'accumulation d'incertitude dans les observations multi-temporelles que les séries temporelles de position, et ce, que le résultat soit issu de l'inversion de Φ_{LF} ou de celle Φ_{CM}. Cette conclusion peut aussi être retrouvée avec un raisonnement sur la relation entre les deux types de mesurandes $\mathbf{F}_\tau$, $\tau = \mathrm{LF}, \mathrm{CM}$. D'autre part, nous notons, une fois de plus, que les deux opérateurs n'ont pas le même comportement face aux données manquantes.

Pondération Pondérer la minimisation permet de prendre en compte les différents indices de qualité monodates et bidates estimées à travers la chaîne de traitements et qui rendent compte de l'hétéroscédasticité dans les mesures. Ainsi, à chaque valeur dans les vecteurs F_0 est associée une valeur diagonale dans W_0 comme suit :

$$\begin{aligned} W_{0_{i,j}} &\propto \frac{T_{G_{i,j}} CC_{i,j}}{|F_\epsilon|_{i,j}} \\ &= (T_G(i) + T_G(j)) \cdot \frac{1}{|F_\epsilon|_{i,j}} \cdot CC_{i,j}, \end{aligned} \tag{8.9}$$

où $T_G(i)$ est l'indice de visibilité attribué à l'image i en pré-traitement, $|F_\epsilon|_{i,j}$ est la vélocité résiduelle en zones fixes en rive gauche proche glacier et $CC_{i,j}$ est le score de corrélation entre l'image i et l'image j égal à la somme de toutes les similitudes de patch le long de leur chemin de "backtracking" lors de la phase DeepMatching (étape de corrélation précédant le calcul du flot optique introduite dans la section 2.4.1.1 puis 6.3.2).

Les poids attribués sont théoriquement proportionnels à la précision de chaque mesure et vont nous permettre d'un côté d'écarter les mesures où les cartes de vélocité biaisées, et d'un autre côté, de se concentrer sur les couples d'images disposant d'un meilleur indice de visibilité moyen et d'une meilleure corrélation DM.

8.3.1.2 Résolution avec pénalisation de la variation totale

Afin d'avoir une fonction de coût qui rend compte des cas de figure où nous sommes en présence de sous-systèmes sous-déterminés dans notre système linéaire globale surdéterminé, nous avons trouvé judicieux de rajouter un terme de régularisation afin que la fonction objectif possède une solution unique. Ce terme permettrait aussi d'écarter les solutions trop incohérentes, par exemple décrivant un mouvement fortement non-stationnaire. Un terme de pénalisation de la norme de la solution serait une contrainte trop forte, particulièrement pour $\hat{\mathbf{F}}_{\mathrm{CM}}$. Nous avons donc opté pour un terme de pénalisation de l'erreur quadratique de la variation totale telle que la fonction globale s'écrit comme suit :

$$\hat{\mathbf{F}}_\tau = \underset{\mathbf{F}_\tau}{argmin}\ ||W_0\mathbf{F}_0 - W_0\Phi_\tau(\mathbf{F}_\tau)||_2^2 + \gamma||\nabla\mathbf{F}_\tau||_2^2. \tag{8.10}$$

Dans ce qui suit, nous allons exposer les considérations théoriques et les observations

empiriques qui nous ont permis de résoudre cette fonction objectif, dite moindre carrées amortie (damped least squares), sur les jeux de données de tests obtenus avec les modèles $M1$,$M2$, $M3$ et $M4$ introduits dans la section 6.3.2, ainsi qu'avec les observations simulées du cas illustratif de la section 8.3.1.1. Nous testerons les différentes formulations et poids proposés, le comportement de l'ajustement au modèle en présence d'observations bruitées, biaisées ou manquantes, etc.

Observations théoriques Le problème 8.10, après pondération et amortissement, doit être considéré comme une approximation stable du problème éventuellement mal posé initial ($\mathbf{F}_0 = \Phi_\tau \mathbf{F}_\tau$) et garde de ce fait les mêmes propriétés de conditionnement (sous/sur détermination et convexité) que celui-ci [174]. Ce type de formulations augmentées est néanmoins plus efficacement résolu via un algorithme itératif. Deux types d'algorithmes de résolution numériques (*solvers*) ont été envisagés :

1. algorithmes où une approximation au premier ordre de la fonction objectif globale permet l'utilisation d'algorithmes du type SIMPLEX, spécialisées dans la l'optimisation linéaire sous contraintes d'inégalités,
2. algorithmes de résolution quadratique amortie qui prennent en compte des contraintes d'inégalité lors de la descente vers le minimum global.

De par les approximations locales autour de l'optimum (optimisation séquentielle), ces deux méthodes permettent de minimiser une fonction objectif globale convexe, qu'elle soit sous ou sur-déterminée, linéaires ou quadratiques, ainsi que des fonctions de coûts non-convexes lisses, i.e. approchables locales par un développement limité d'ordre 1 ou 2. Celles-ci vont permettre, en outre, d'inclure, dans une phase ultérieure, les contraintes d'inégalité traduisant le sens d'écoulement du glacier. Le premier type de résolution envisagé correspond à la méthode, nommée COByLA (Constrained Optimization By Linear Approximations), proposée par Powell [175], et dont l'implémentation Fortran [1] a été reprise par le projet Scipy [2]. Avec ces algorithmes de type SIMPLEX, une région convexe est d'abord définie comme l'intersection de tous les demi-espaces induits par les inégalités, et la fonction affine est ensuite minimisée dans le polyèdre convexe ainsi délimitée. Le deuxième type de résolution envisagé correspond à une implémentation [3] d'un algorithme d'optimisations quadratiques successives, méthode à direction de descente (méthode de quasi-Newton) prenant en compte des contraintes d'inégalité, proposée par Kraft [176].

Observations empiriques Dans ce qui suit seront comparées les solutions issues des deux algorithmes de résolution sus-introduits ainsi qu'une résolution du problème initial par

1. `https://zhangzk.net/software.html`
2. `https://docs.scipy.org/doc/scipy/reference/optimize.minimize-cobyla.html`
3. `https://docs.scipy.org/doc/scipy/reference/optimize.minimize-slsqp.html`

décomposition en valeurs singulières tronquée (tSVD) sans l'injection de contraintes d'inégalités contraignant le sens du déplacement des pixels. Les résultats présentés sont en termes de l'erreur spatio-temporelle quadratique moyenne définie par :

$$\epsilon_{\mathrm{LF}_{sim}} = \frac{1}{5N_\Omega} \sum_\Omega \sum_{t=0}^{5} ||\tilde{\mathbf{F}}_{t,t+1} - \hat{\mathbf{F}}_{\mathrm{LF}_{t,t+1}}||_2^2,$$

pour les résultats en séries temporelles de vélocité, et en terme de :

$$\epsilon_{\mathrm{CM}_{sim}} = \frac{1}{5N_\Omega} \sum_\Omega \sum_{t=0}^{5} ||\tilde{\mathbf{F}}_{0,t+1} - \hat{\mathbf{F}}_{\mathrm{CM}_{0,t+1}}||_2^2,$$

pour les résultats en séries temporelles de position. Ω étant le domaine spatial constitué de $N_\Omega = 500 \times 1000$ pixels de la zone glaciaire des sous-imagettes des time-lapses simulés et $\tilde{F}$ étant la vérité terrain nous ayant permise de créer les jeux de données.

Observations empiriques avec des transformations non-bruitées Cette première série de tests va nous permettre de comparer les deux formulations de l'opérateur fermeture temporelle et techniques de résolution avec les différentes combinaisons de poids suivantes :

$$W_0 = T_{G_{i,j}} CC_{i,j},$$
$$W_{0'} = \frac{1}{|F_\epsilon|_{i,j}},$$
$$W_{0''} = \frac{T_{G_{i,j}} CC_{i,j}}{|F_{\epsilon_{i,j}}|}.$$

La précision d'affichage des erreurs obtenus a été augmentée pour rendre compte de l'amélioration apportée par l'ajustement au modèle de cohérence temporelle sans contrainte (et sans amortissement pour les solutions obtenues par tSVD). Les résul-

Jeu de données	DMDF	DMDF+tSVD				DMDF+iteratif
		$\varnothing$	W_0	$W_{0'}$	$W_{0''}$	min
$M1$	0.01646	0.00171	0.00130	0.00115	0.00117	0.00278
$M2$	0.08366	0.02005	0.16890	0.01590	0.01583	0.01457
$M3$	0.10486	0.01737	0.01254	0.01097	0.01115	0.01022
$M4$	0.04363	0.02834	0.02970	0.02970	0.02834	0.02109

(a) $\epsilon_{\mathrm{LF}_{sim}}$

TABLE 8.1 – Erreurs spatio-temporelles quadratiques moyennes obtenues avec minimisation d'une fonction objectif en termes de Φ_{LF} sur les cartes de vélocité calculées par la méthode proposée (DMDF).

Jeu de données	DMDF	DMDF+tSVD				DMDF+itératif
		$\varnothing$	W_0	W_0'	W_0''	min
$M1$	0.05385	0.00339	0.00220	0.00135	0.00214	0.00442
$M2$	0.38395	0.06115	0.03311	0.03149	0.02862	0.03450
$M3$	0.42601	0.05623	0.02675	0.02763	0.02188	0.01850
$M4$	0.09780	0.01349	0.01973	0.01748	0.02058	0.01117

(a) $\epsilon_{\mathrm{CM}_{sim}}$

TABLE 8.2 – Erreurs spatio-temporelles quadratiques moyennes obtenues avec minimisation d'une fonction objectif en termes de Φ_{CM} sur les cartes de vélocité calculées par la méthode proposée (DMDF).

Jeu de données	CC	CC+itératif
$M1$	0.11597	0.07570
$M2$	0.63350	0.37497
$M3$	2.43873	1.19262
$M4$	1.71092	0.30109

(a) $\epsilon_{\mathrm{LF}_{sim}}$

Jeu de données	CC	CC+itératif
$M1$	0.25562	0.19200
$M2$	2.07360	0.43204
$M3$	6.91797	3.83019
$M4$	3.97034	0.51492

(b) $\epsilon_{\mathrm{CM}_{sim}}$

TABLE 8.3 – Erreurs spatio-temporelles quadratiques moyennes obtenues avec minimisation d'une fonction objectif en termes de Φ_{LF} (gauche) et Φ_{CM} (droite) sur les cartes de vélocité calculées par corrélation-croisée centrée normalisée (CC).

tats sont reportés dans les tableaux 8.1, 8.2 et 8.3. Sur tous les jeux de données, et quelle que soit la méthode adoptée afin de calculer la vélocité des pixels, la régularisation réduit la distance Euclidienne à la vérité terrain qui nous a permis de construire les données simulées et, dans la plupart des expériences, particulièrement lorsque la pondération est effectuée via l'indicateur W_0' qui disqualifie les observations ayant une erreur résiduelle empirique en zones fixes proche glacier élevée . Dans ce cas de figure, la méthode itérative fait référence aux deux méthodes SLSQP et COByLA puisque celles-ci donnent sensiblement le même résultat. Ce résultat, de par l'ajout d'un terme d'amortissement, possède une erreur résiduelle moindre que celui obtenu par décomposition en valeurs singulières tronquée (tSVD). La résolution du problème avec une norme $L = 1$ et une norme $L = 2$ donnent aussi sensiblement le même résultat, avec une convergence plus lente lors l'utilisation d'une norme $L = 1$.

Afin de tester la tolérance de l'ajustement au degré de déformation non-rigide, il a été intéressant de quantifier $\epsilon_{\mathrm{LF}_{sim}}$ en fonction du degré de cisaillement. Les résultats sur le jeu de données $M2$ où le cisaillement augmente avec x, y et t nous ont permis de générer la figure 8.15. Naturellement, nous notons que l'erreur avant et après régularisation sont corrélées positivement et que lorsque toutes les observations sont bruitées, par exemple lorsque la CC ne parvient pas à détecter le mouvement pixellique (voir figure 6.15), l'ajustement ne parvient pas non plus à régulariser les données de suivi et l'erreur résiduelle reste élevée.

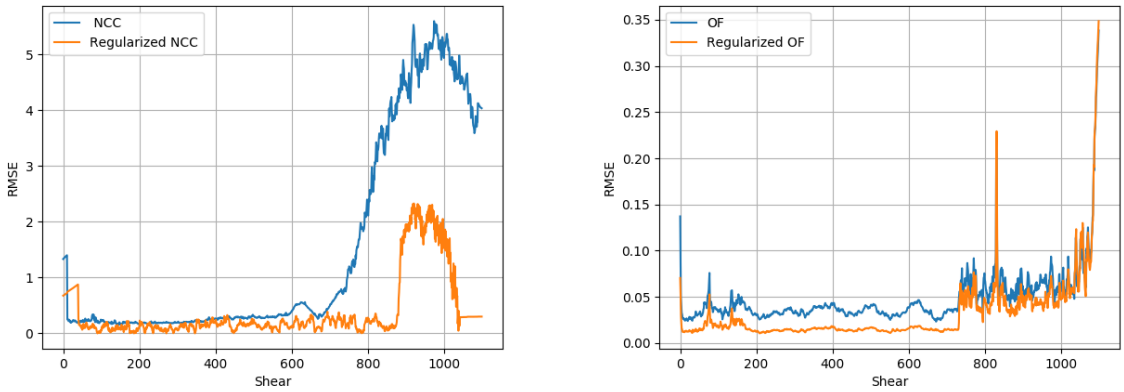


FIGURE 8.15 – $\epsilon_{\mathrm{LF}_{sim}}$ en fonction du degré de cisaillement avec une CC (haut) et une DMDF (bas) sur le jeu de données $M2$ avant (orange) et après (blue) régularisation par ajustement du modèle de cohérence temporel amorti.

Observations empiriques avec injection de bruit aléatoire Dans cette série d'expériences, un bruit additif gaussien ω centré en 0 et d'écart-type μ_ω a été est injecté spatialement dans les observations $\mathbf{F}_0$ afin de simuler une erreur de mesure absolue telle que :

$$F_{0_{i,j}} \leftarrow F_{0_{(i,j)}} + \omega_{i,j}.$$

Jeu de données	$\mu_\omega = 0.1$ pixel			$\mu_\omega = 1$ pixel			$\mu_\omega = 5$ pixel		
	DMDF $+\omega$	+tSVD	itératif	//	//	//	//	//	//
$M4$	0.13167	0.03364	0.03341	1.02832	0.53800	0.51000	5.02794	2.39173	1.38109

TABLE 8.4 – Erreurs spatio-temporelles quadratiques moyennes $\epsilon_{\text{LF}_{sim}}$ obtenues avec minimisation d'une fonction objectif en termes de Φ_{LF} sur les cartes de vélocité calculées par la méthode proposée (DMDF) sur lesquelles un bruit Gaussien absolu, centré et d'écart-type μ_ω a été injecté.

Jeu de données	$\mu_\omega = 0.1$ pixel			$\mu_\omega = 1$ pixel			$\mu_\omega = 5$ pixel		
	DMDF $+\omega$	+tSVD	itératif	//	//	//	//	//	//
$M4$	0.20621	6.91810	0.03761	1.10082	7.27980	0.54335	5.09987	5.09320	2.26979

TABLE 8.5 – Erreurs spatio-temporelles quadratiques moyennes $\epsilon_{\text{CM}_{sim}}$ obtenues avec minimisation d'une fonction objectif en termes de Φ_{CM} sur les cartes de vélocité calculées par la méthode proposée (DMDF) sur lesquelles un bruit Gaussien absolu, centré et d'écart-type μ_ω a été injecté.

Les résultats sont reportés sur les tableaux 8.4 and 8.5. L'ajustement au modèle de cohérence temporelle s'est avéré très efficace pour débruiter un bruit de mesure blanc additif. Dans la plupart des cas, l'erreur dans les observations est diminuée de plus de 50% par régularisation, et ce particulièrement quand l'écart-type du bruit est élevé. Nous apercevons néanmoins quelques valeurs divergentes dans les solutions obtenues via tSVD (en rouge dans le tableau 8.5), ce qui met en évidence l'importance du terme d'amortissement pour stabiliser l'optimisation. L'effet spatial et temporel de ce débruitage temporel est mis en évidence sur la figure 8.16 où nous pouvons voir que le bruit a complètement disparu des zones fixes et fortement diminué sur les zones en déplacement.

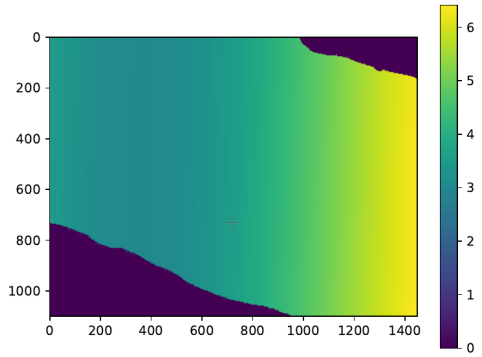
Observations empiriques dans le cas de données manquantes Le rejet factice de 1, 2 ou encore 2 images adjacentes sur les time-lapses simulés a permis de générer les résultats présentés dans les tableaux 8.6 et 8.7. Dans le cas où le problème est réduit à un problème

Jeu de données	1/6 observations manquantes			2/6 observations manquantes			2/6 observations manquantes adjacentes		
	DMDF	tSVD	SLSQP	//	//	//	//	//	//
$M4$	0.03864+nod	0.53062	0.01667	0.04214+nod	0.79308	0.59328	0.03873+nod	1.66724	0.63470

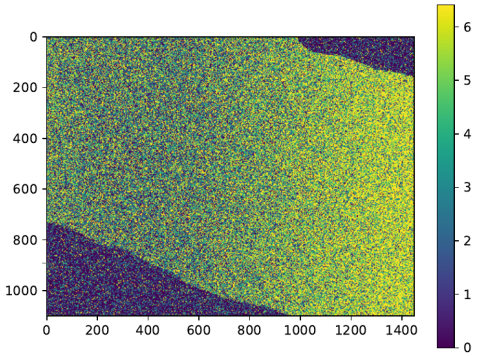
TABLE 8.6 – Erreurs spatio-temporelles quadratiques moyennes $\epsilon_{\text{LF}_{sim}}$ obtenues avec minimisation d'une fonction objectif en termes de Φ_{LF} sur les cartes de vélocité calculées par la méthode proposée (DMDF) avec rejet d'images. nod = no data.

	Φ_{CM}								
Jeu de données	1/6 observations manquantes			2/6 observations manquantes			2/6 observations manquantes adjacentes		
	DMDF	tSVD	SLSQP	//	//	//	//	//	//
$M4$	0.12232+nod	4.22397	1.95198	0.09683+nod	6.52000	3.65185	0.13430+nod	6.50164	2.51473

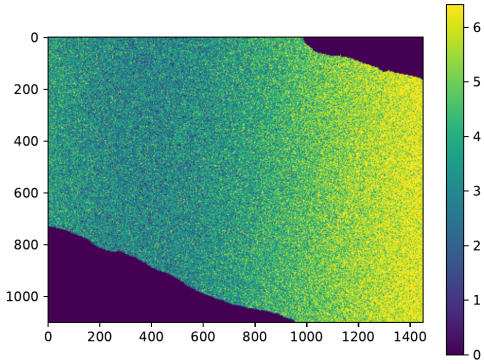
TABLE 8.7 – Erreurs spatio-temporelles quadratiques moyennes $\epsilon_{\text{CM}_{sim}}$ obtenues avec minimisation d'une fonction objectif en termes de Φ_{CM} sur les cartes de vélocité calculées par la méthode proposée (DMDF) avec rejet d'images. nod = no data.



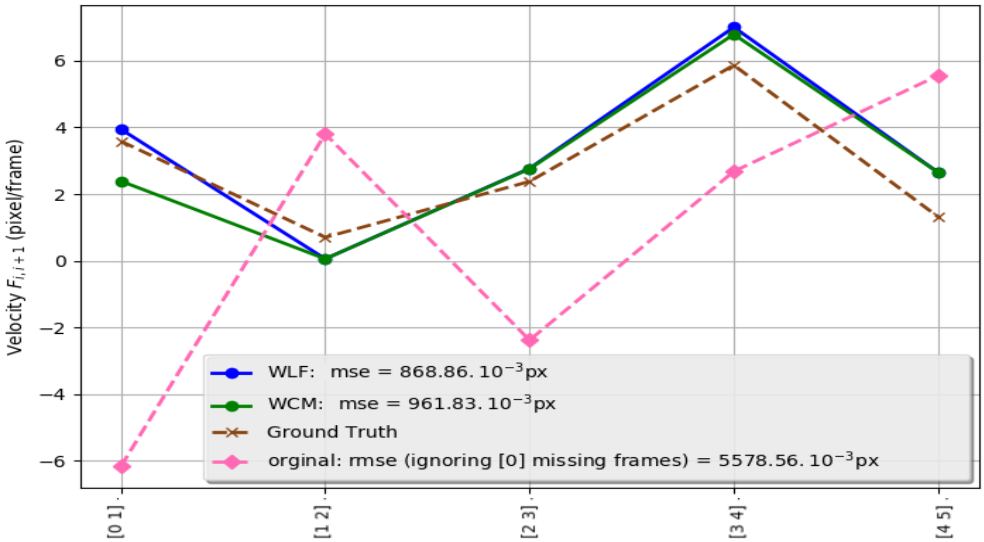
(a) $\tilde{\mathbf{F}}_{0,1}$: Vérité terrain



(b) $\mathbf{F}_{0,1}$: observations altérées par un bruit spatial aléatoire d'écart-type $\mu_\omega = 5$ pixels



(c) $\hat{\mathbf{F}}_{\mathrm{LF}_{0,1}}$: solutions obtenues après ajustement au modèle Φ_{LF} amorti

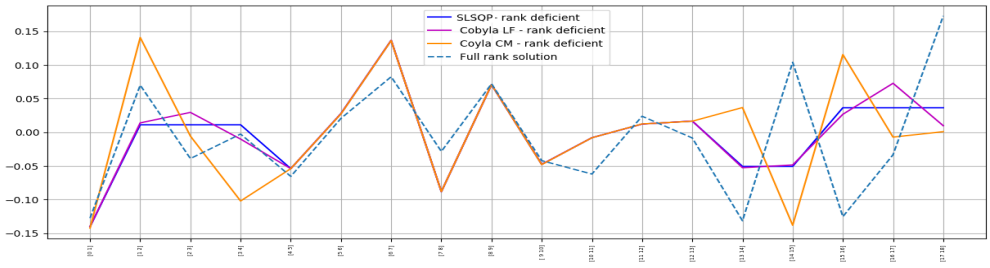


(d) Séries temporelles de suivi de vélocité obtenues par inversion de Φ_{LF} (bleu) et Φ_{CM} et conversion (en vert) sur le pixel central du jeu de données $M4$.

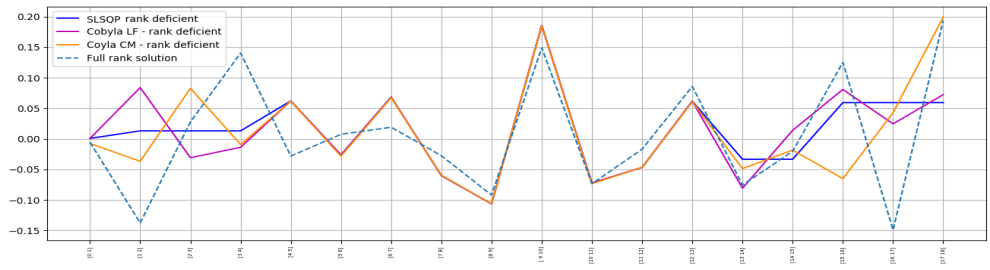
FIGURE 8.16 – Mise en évidence du débruitage d'un bruit blanc Gaussien d'écart-type μ_0 = 5 pixels sur le jeu de données $M4$ par régularisation.

à rang incomplet par rejet d'images en phase de pré-traitements, nous constatons tout d'abord que, tout comme mis en évidence dans la section 8.3.1.1, le comblement de mesures manquantes dans les séries temporelles de vélocité et de position dépend de l'opérateur de fermeture temporelle inversé. Ainsi, une minimisation quadratique (tSVD et SLSQP) de l'erreur résiduelle de l'ajustement à Φ_{LF} entraîne des valeurs constantes de suivi de vélocité (valeurs linéaires pour le suivi de position) sur les index où une des dates est manquante et celle de l'ajustement à Φ_{CM} remplis les trous temporels avec des valeurs nulles (comme mis en évidence dans le résultat de la figure 8.13). D'autre part, la minimisation linéaire via le *solver* COByLA a permis, au prix d'un temps d'exécution double, de combler les trous temporels avec des valeurs non-constantes et non-nulles, tout en générant le même résultat sur les autres index bidates que le *solver* SLSQP.

Des tests supplémentaires par simulation d'observations selon les deux configurations de l'exemple illustratif de la section 8.3.1.1 n'ont cependant pas pu nous permettre de justifier l'utilisation de cette méthode afin de combler les valeurs manquantes. Nous voyons sur plusieurs exemples tels ceux de la figure 8.17 que les séries temporelles obtenues ne se rapprochent pas plus de la solution obtenue par inversion du problème à rang plein qu'avec la méthode SLSQP, plus rapide.



(a) Exemple 1



(b) Exemple 2

FIGURE 8.17 – $\hat{\mathbf{F}}_{\mathrm{LF}}$ obtenues par simulation d'observations sur un time-lapse à 19 dates et 5 images manquantes (voir figure 8.10b). Bleu : solution obtenue par inversion du modèle Φ_{LF} amorti par le *solver* SLSQP. Rouge : solution obtenue par inversion du modèle Φ_{LF} amorti par le *solver* COByLA. Orange : solution obtenue par inversion du modèle Φ_{CM} amorti par le *solver* COByLA (et conversion). Pointillés : solution obtenue par inversion des observations avant rejet d'images (problème à rang entier).

Observations empiriques sur un biais systématique dans les observations Dans cette série d'expériences, du bruit a été injecté dans une région spatiale des observations $\mathbf{F}_{0_{i,j}}$ et $\mathbf{F}_{0_{j,i}}$ qui incluent une image commune I_i. L'objectif est de simuler un biais induit par les données dans les cartes de vélocité impliquées. Ce type de bruit systématique est modélisé par une distribution Gaussienne avec un écart-type $\mu = 1$ et une valeur moyenne moy qu'on fera varier. Sur les données réelles, ce cas de figure peut se produire par exemple en présence d'artefacts occultant la scène observée (voir figure 5.9) sur une image I_i. Les résultats obtenus avec différentes valeurs moyennes moy et avec les deux formulations de la fermeture temporelle sont résumés sur les tableaux 8.8 et 8.9. La minimisation de

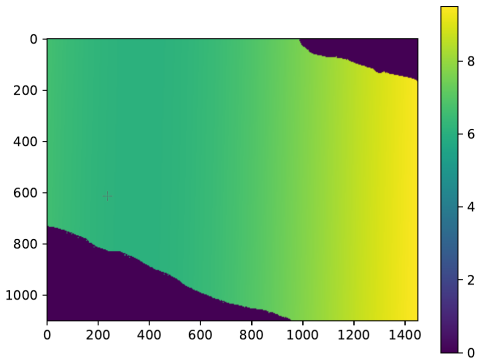
	$\epsilon_{\text{LF}_{sim}}$								
Jeu de données	moy = 5 pix			moy = 10 pix			moy = 20 pix		
	DMDF+ biais	tSVD	SLSQP	//	//	//	//	//	//
$M4$	2.25672	0.07824	0.07824	4.49264	1.31144	0.07824	8.96470	1.95816	1.95662

TABLE 8.8 – Erreur spatio-temporelle ϵ_{LF} avant et après une régularisation via le modèle Φ_{LF} avec différentes amplitudes de biais associées à l'image 1.

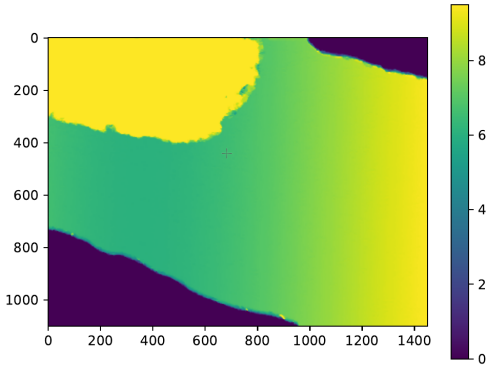
	$\epsilon_{\text{CM}_{sim}}$								
Jeu de données	moy = 5 pix			moy = 10 pix			moy = 20 pix		
	DMDF + biais	tSVD	SLSQP	//	//	//	//	//	//
$M4$	2.25912	6.94773	0.09298	4.49385	5.98537	1.66244	8.96531	8.65959	1.97657

TABLE 8.9 – Erreur spatio-temporelle ϵ_{CM} avant et après une régularisation via le modèle Φ_{CM} avec différentes amplitudes de biais associées à l'image 1.

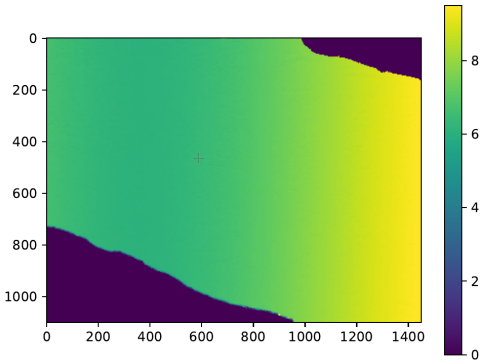
la fonction objectif proposée parvient à éliminer la quasi-totalité du bruit systématique injecté. L'utilisation d'un terme d'amortissement empêche en outre la solution de diverger (valeurs divergentes en rouge dans le tableau 8.9).



(a) $\tilde{\mathbf{F}}_{0,3}$: Vérité terrain



(b) $\hat{\mathbf{F}}_{\mathrm{LF}_{0,3}}$: solutions obtenues après régularisation



(c) $\mathbf{F}_{0,1}$: observations altérées par un bruit systèmatique d'écart-type 1 et de moyenne moy = 5 pixels

FIGURE 8.18 – Mise en évidence du débruitage d'un bruit systématique d'écart-type μ = 1 pixel et de moyenne moy = 5 pixels par régularisation.

8.3.2 Paramétrisation, résolution et résultats

8.3.2.1 Paramétrisation et résolution

Dans ce qui suit, nous présenterons certaines expériences qui nous ont permis de paramétriser et d'étudier le comportement de la fonction objectif globale quand elle est minimisée afin de régulariser l'ensemble des vraies mesures #n (3D+t) obtenues pour le suivi du glacier d'Argentière. Ceux-ci seront alors présentés sous forme de séries temporelles régularisées de vélocité métrique à pas de temps journalier, moyennées sur une zone spatiale de 100×100 pixels. Comme une vérité terrain n'est pas disponible, le comportement attendu ainsi que les séries temporelles obtenues en zones fixes sont un indicateur sur la qualité des résultats.

Tout d'abord, sur la figure 8.19 est illustrée l'influence de la contrainte d'inégalité contraignant le sens d'écoulement du glacier. Nous pouvons voir que celle-ci n'a pas pour effet de tronquer les estimations, mais que même des valeurs satisfaisant initialement (avant injection de la contrainte) un sens d'écoulement cohérent peuvent changer de valeur. En zones fixes, les séries temporelles se rapprochent d'une vélocité constamment nulle.

Sur la figure 8.20 sont illustrés les poids calculés sur les observations bidates multitemporelles (pour la composante "*North*") et les différentes normalisations envisagées. Nous notons alors que prendre le logarithme des poids calculés permet d'alléger un éventuel effet notoire dans la disparité observée sur ces poids.

Sur la figure 8.21 est illustrée la différence entre les solutions obtenues via l'utilisation des deux formulations "Leap-Frog" (LF) et "Common-Master" (CM). Ces résultats confirment que les résultats obtenus par inversion de la contrainte de fermeture temporelle du type CM sont entachés d'une plus grande dispersion que ceux obtenus par inversion de la contrainte de fermeture temporelle de type LF à proximité des données manquantes (comme mis en évidence sur la figure 8.14). Les résultats restent sensiblement similaires et notamment en zones fixes.

Sur la figure 8.22, nous voyons la différence entre les différents *solvers* itératifs envisagés et la solution directe obtenue par décomposition en valeurs singulières (SVD). Sans le terme d'amortissement, la solution obtenue via l'optimiseur quadratique SLSQP est égale à la solution obtenue par décomposition directe. Le comportement aléatoire des solutions obtenues via le *solver* COByLA est mis en évidence sur les résultats obtenus sur les zones fixes.

Enfin, la dernière expérience effectuée sur les vraies données de suivi porte sur l'effet du poids accordé au terme de minimisation de la variation totale par rapport au terme d'attache à la contrainte de fermeture temporelle. Nous pouvons voir sur la figure 8.23 l'effet régularisant de ce terme. La velocité résiduelle en zones fixes se rapproche notamment de 0. Le poids est retenu est égal à $\lambda = 0.005$.

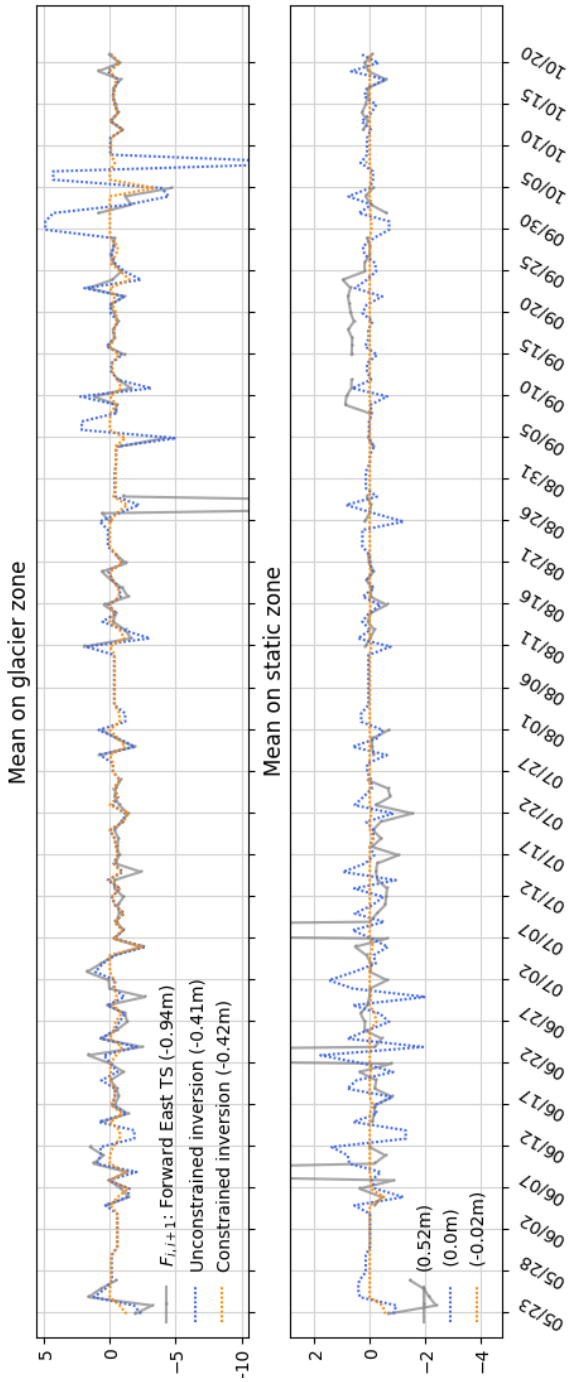


FIGURE 8.19 – Composante "*East*" des séries temporelles à pas de temps journalier en sortie du *solver* SLSQP en zone frontale du glacier (haut) et en zones statiques (bas). Gris : séries temporelles non traitées. Bleu : résultat d'une optimisation Leap-Frog (LF) pondérée, non amortie et non-contrainte. Orange : rajout de la contrainte d'inégalité (négativité pour la composante "*East*").

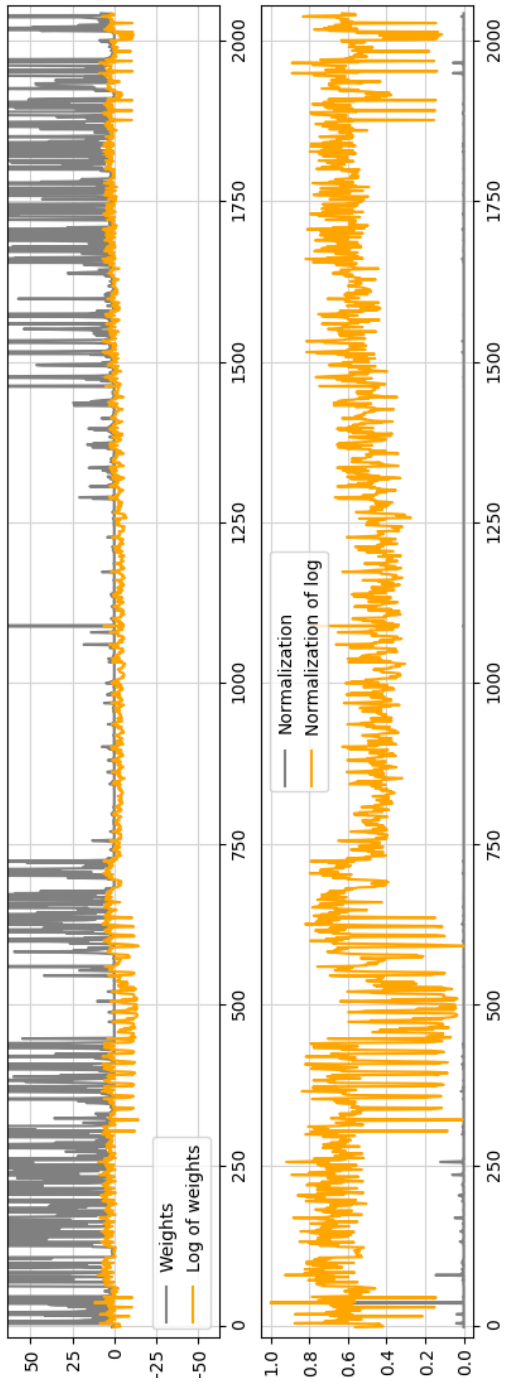


FIGURE 8.20 – Poids retenus $W_{0'}$ (gris) et leurs logarithmes (orange) des différentes mesures bidates (composante "*North*") avant (haut) et après (bas) normalisation par la portée temporelle.

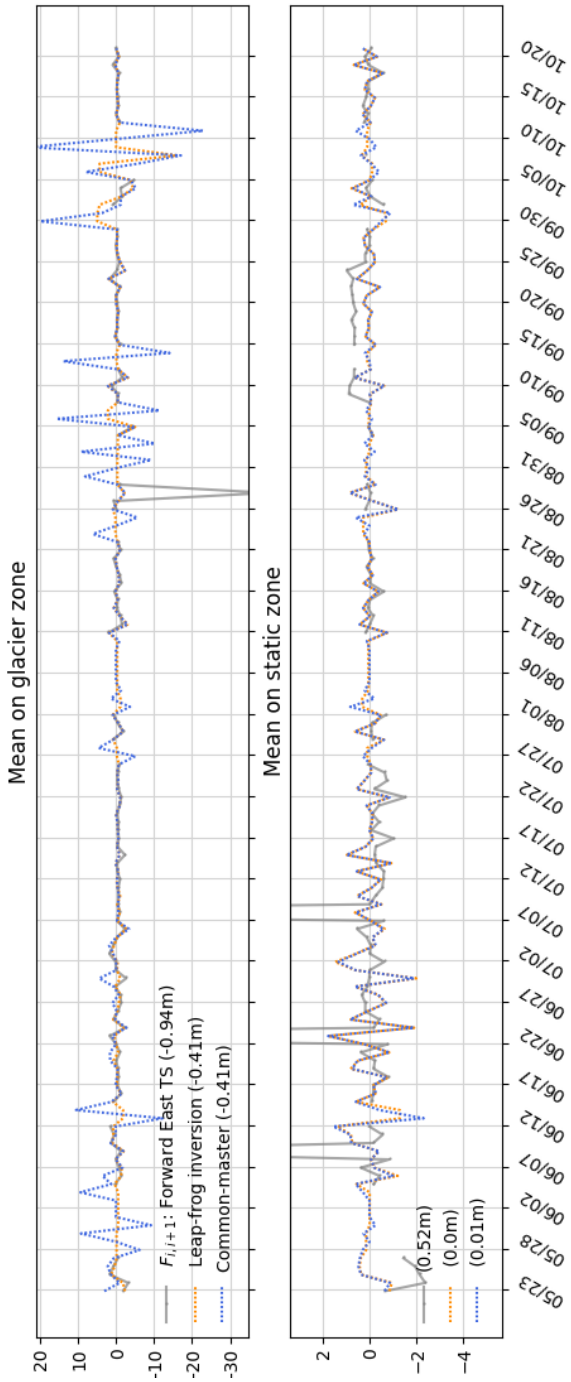


FIGURE 8.21 – Composante "*East*" des séries temporelles à pas de temps journalier en sortie du *solver* SLSQP en zone frontale du glacier (haut) et en zones statiques (bas). Gris : séries temporelles non traitées. Bleu : résultat d'une optimisation Leap-Frog (LF) pondérée, non amortie et non-contrainte. Orange : résultat d'une optimisation Common-Master (CM) pondérée, non amortie, non-contrainte et convertie de position à vélocité (VTS).

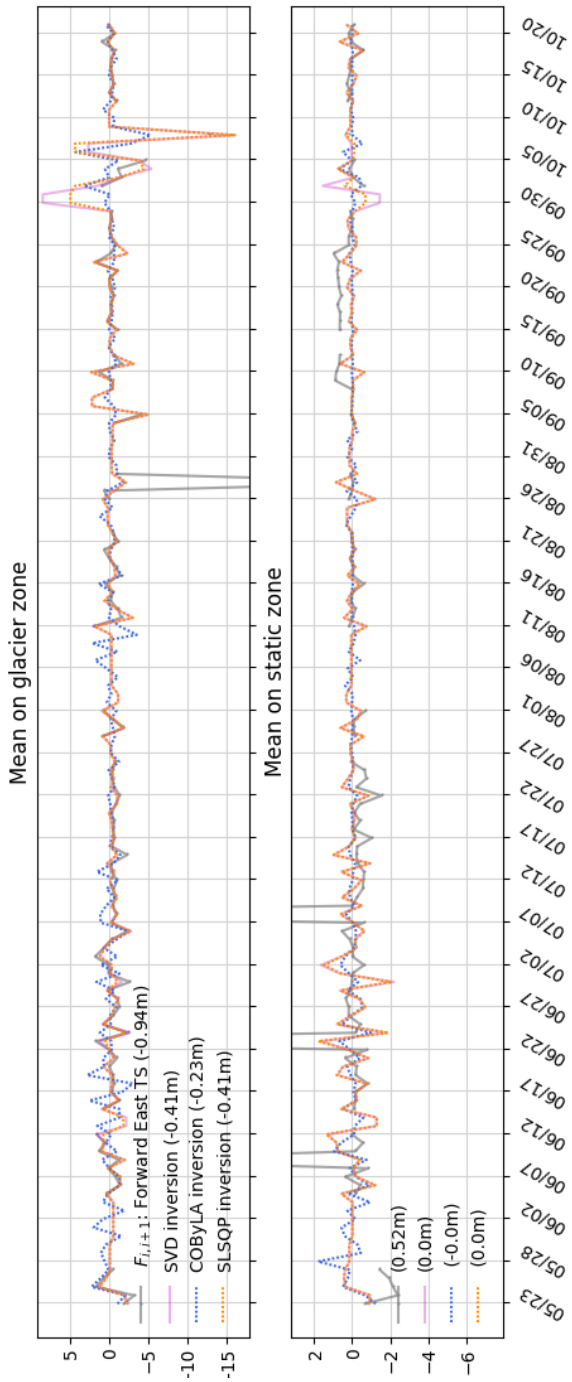


FIGURE 8.22 – Composante "*East*" des séries temporelles à pas de temps journalier en sortie d'optimisation en zone frontale du glacier (haut) et en zones statiques (bas). Gris : séries temporelles non traitées. Bleu : résultat d'une optimisation Leap-Frog (LF) pondérée, non amortie et non-contrainte obtenue via le *solver* COByLA. Orange : Bleu : résultat d'une optimisation Leap-Frog (LF) pondérée, non amortie et non-contrainte obtenue via le *solver* SLSQP. Bleu : résultat d'une optimisation Leap-Frog (LF) pondérée, non amortie et non-contrainte obtenue via décomposition en valeurs singulières (SVD).

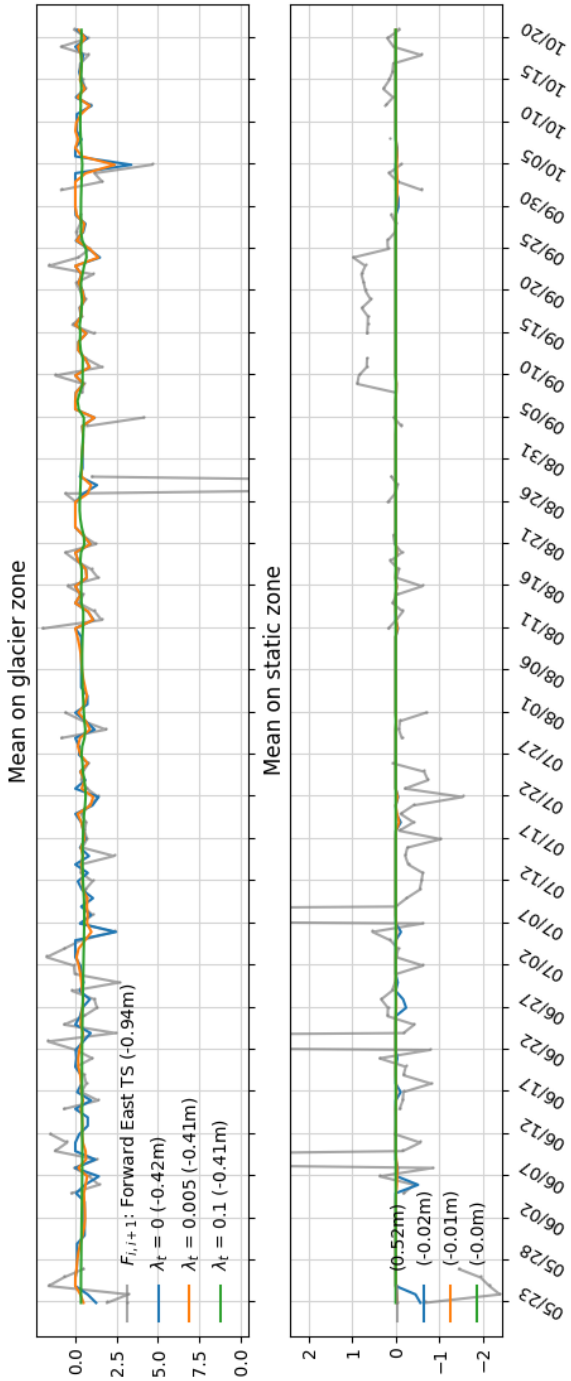


FIGURE 8.23 – Composante "*East*" des séries temporelles à pas de temps journalier en sortie d'une régularisation avec contrainte de négativité via le solver SLSQP en zone frontale du glacier (haut) et en zones statiques (bas) avec modulation du poids accordé à la pénalisation de la variation totale.

8.3.2.2 Résultats sur l'année 2018

Pour finir, les données de suivi obtenues via la méthode retenue sont illustrées sur les figures 8.24, 8.25 et 8.26 respectivement pour les composantes *East*, *North* et *Up* des champs de flot de scène métriques moyennés au niveau de la zone marquée par I sur la figure 7.11.

La régularisation temporelle nous a permis de fournir des estimations de vélocité métrique régularisée en toute date. Celle-ci a en outre considérablement réduit les vélocités résiduelles moyennes en zones statiques. On note que celles-ci étaient moins élevées que les estimations de l'incertitude présentées dans le tableau 7.2 avant régularisation (0.82m pour la composante *East*, 0.82m pour la composante *North* et 0.23m pour la composante *Up*), est quasi réduites à zéro après régularisation.

8.4 Conclusion

Dans ce chapitre, nous avons exposé la méthode proposée afin de traiter les séries temporelles de suivi de la vélocité métrique du glacier d'Argentière. Nous avons cherché à exploiter la multi-temporalité des mesures bidates issues de time-lapses. Cette multi-temporalité permet en effet de créer de la redondance théorique exploitable afin d'ajuster un modèle décrivant linéairement la symétrie dans la définition des mesures. Cette technique est inspirée de la méthode interférométrique SBAS (Small BAseline Subset) et notre contribution consiste en son étude théorique et empirique dans le cas particulier de time-lapses optiques pour l'observation de milieux naturels dont les mesures souffrent d'un bruit de mesure absolu (ne dépend pas de la portée temporelle) aléatoire (centré en zéro) et de la présence d'une quantité importante de valeurs aberrantes. Pour ce faire, le problème a été pondéré via des indicateurs de confiance monodates et bidates, un terme de régularisation a été ajouté afin d'écarter les éventuels minima souffrant d'une non-stationnarité importante en contradiction avec la nature tempérée du glacier et enfin la résolution a été contrainte par le sens d'écoulement Nord-Ouest et vers le bas. Afin de résoudre la fonctionnelle ainsi formulée, nous avons eu recours au solver quadratique itératif SLSQP.

La vélocité estimée est en moyenne plus élevée que les estimations annuelles sur les années précédentes (tableau 5.1). Ceci peut être expliqué par une hausse de la température moyenne annuelle dans les Alpes ainsi que par la fluctuation importante de la température intra-annuelle entraînant beaucoup d'alternance entre épisodes neigeux, pluvieux et chauds, et ce particulièrement pendant la deuxième moitié de l'année.

Comme il a déjà été mentionné, la régularisation étant effectuée indépendemment en tout pixel reprojeté sur le plan image, peut alors être parallélisée. Les expériences effectuées nous ont en effet montré un gain important en terme de performances informatiques.

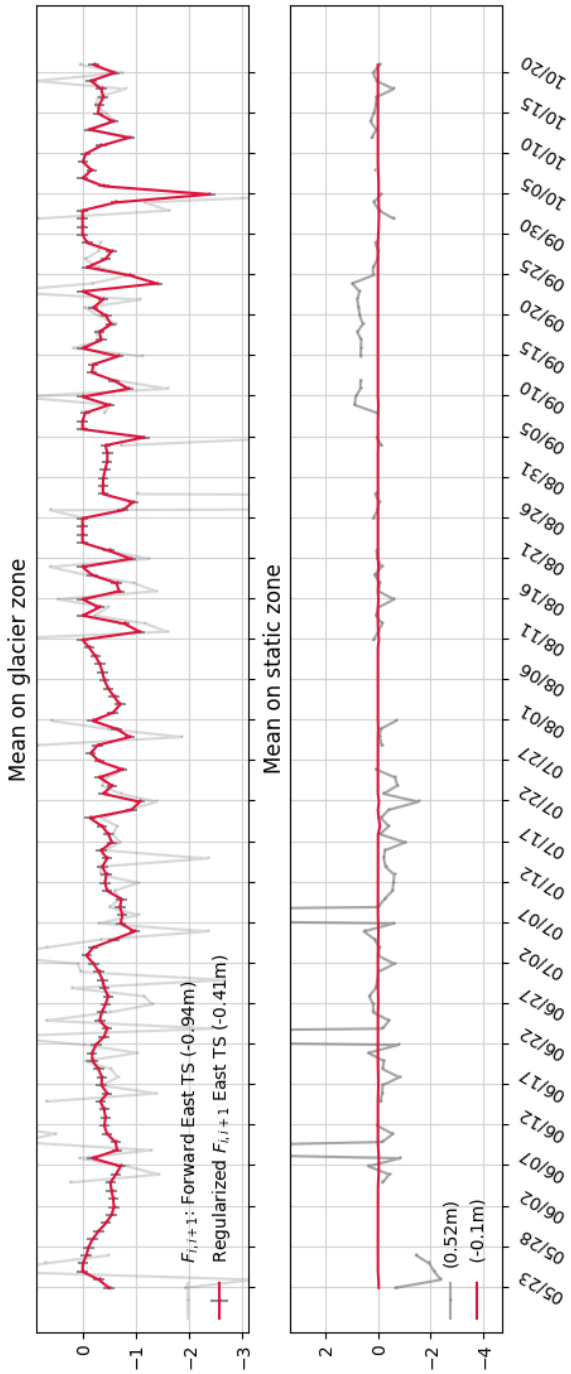


FIGURE 8.24 – Données de suivi de la composante *East* de la zone frontale du glacier d'Argentière sur les données retenues de l'année 2018 en mètres.

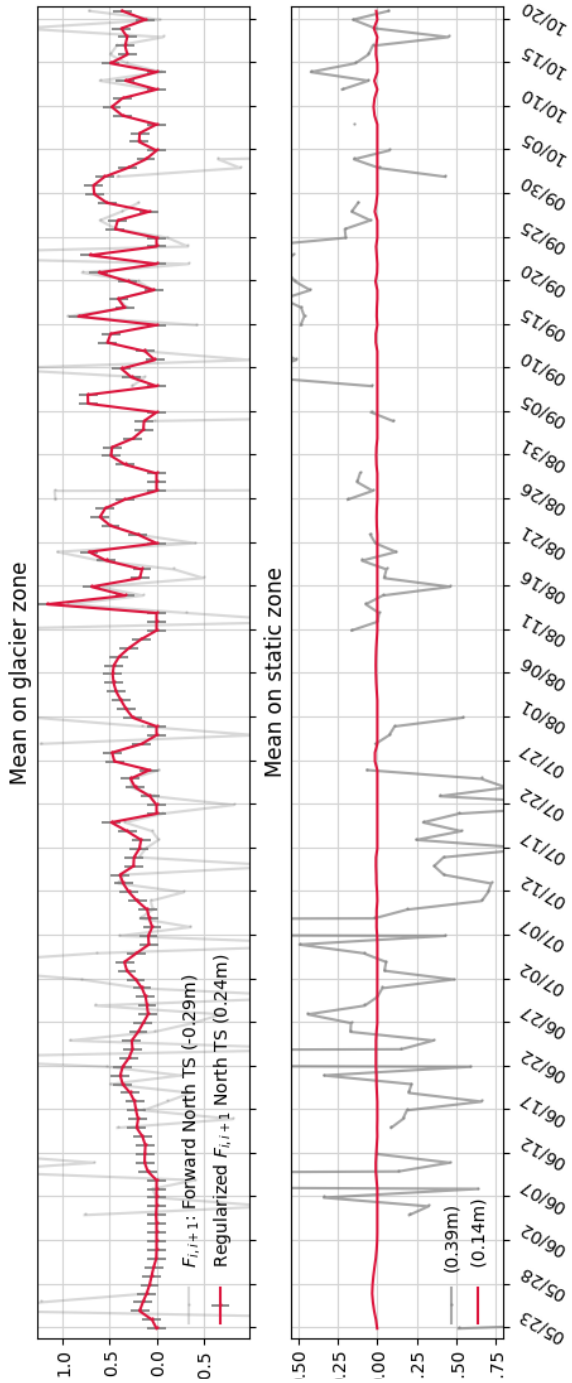


FIGURE 8.25 – Données de suivi de la composante *North* de la zone frontale du glacier d'Argentière sur les données retenues de l'année 2018 en mètres.

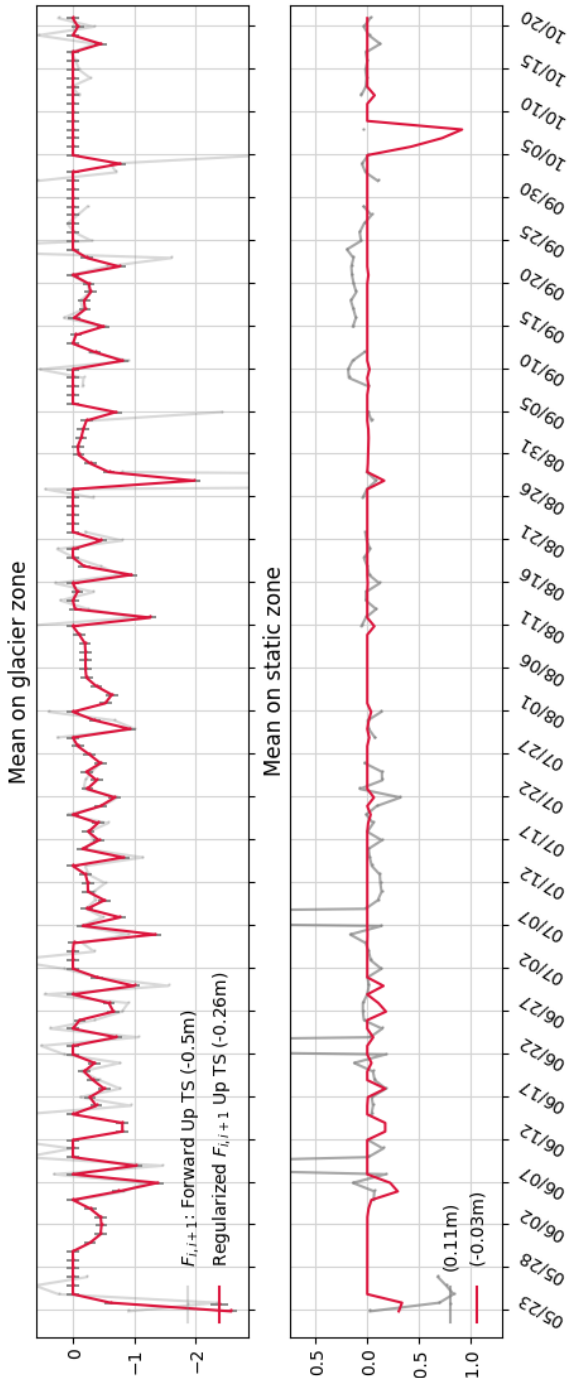


FIGURE 8.26 – Données de suivi de la composante *Up* de la zone frontale du glacier d'Argentière sur les données retenues de l'année 2018 en mètres.

8.5 Perspectives

Tout d'abord, il est important de noter qu'il n'a malheureusement pas été possible de générer les résultats spatiaux denses (tels une carte médiane sur les séries temporelles régularisées) sur les données acquises sur l'année 2018, et ce malgré un effort considérable fourni dans l'optimisation informatique de la chaîne de traitements proposée. Ceci est principalement lié à la taille des données et nous en concluons qu'il est nécessaire de déployer plus de moyens afin de rendre cette solution viable et de pouvoir la maintenir.

En outre, un axe de recherche intéressant serait d'étudier la faisabilité d'une implémentation incrémentale pour la régularisation de séries temporelles grâce à l'accumulation de nouvelles observations.

L'étude de la covariance des erreurs des mesures bidate permettrait d'autre part de rajouter des termes non-diagonaux dans la matrice de pondération des observations W_0.

Pou finir, l'incorporation de contraintes épipolaires spatiales dans le terme d'attache aux données pourrait permettre de maximiser jointement la cohérence stéréoscopique et temporelle sur l'ensemble des mesures bruitées effectuées sur un time-lapse stéréoscopique de milieux naturels.

Conclusion générale et perspectives

Conclusion

Dans cette thèse, nous avons étudié et traité la problématique de surveillance continue et métrique de milieux naturels extrêmes tels que les glaciers alpins via un système d'observation optique stéréoscopique dans la continuité des travaux existants. L'objectif était de proposer une chaîne de traitements dédiée aux time-lapses stéréoscopiques en tenant automatiquement compte de la qualité des acquisitions et des mesures, de l'incertitude liée à la profondeur d'observation, des ambiguïtés systématiques ainsi que de nos connaissances à priori sur la scène observée lors de la génération des séries temporelles de suivi de vélocité d'écoulement. Les limitations majeures rencontrées résident dans la sensibilité des mesures bidates aux données et cela particulièrement en considérant un pas de temps fixe (par exemple journalier).

Afin de générer des résultats cohérents, nous nous sommes, tout d'abord, positionnés d'emblée dans un contexte de caméras non finement calibrées et d'images entachées d'aléas atmosphériques et climatiques importants (brume, variations photométriques, manque de texture, occlusions, etc.) afin d'implémenter une chaîne de traitements avec des piliers robustes et complémentaires de sélection de l'information pertinente, de co-recalage des données, de calcul de la vélocité pixellique en tout point et enfin de construction d'un modèle numérique de terrain fidèle qui nous permet de remonter à des séries temporelles à 3 composantes géo-référencées East, North et Up.

Nous avons ensuite proposé de régulariser, en amont, les mesures ainsi obtenues par la construction d'un modèle de cohérence intra time-lapse prenant en compte des mesures de vélocité bidates à différentes échelles temporelles qui nous permet, d'une part, d'enrichir notre base d'observations afin de distinguer le bruit de la mesure, et d'autre part, d'imposer des contraintes de sens d'écoulement et de stationnarité des séries temporelles. Différents jeux de données simulés nous ont permis de tester la robustesse et la précision des calculs bidates ainsi que les hypothèses théoriques liées au modèle de cohérence temporelle et sa résolution.

Les résultats expérimentaux ainsi que l'étude des différentes stratégies possibles selon le type de données utilisées jointement nous permettent alors de conclure qu'il est important de déployer des techniques de mesure de la vélocité pixellique bidate possédant des propriétés d'invariance robustes et adaptées à la mesurande qui est dans notre cas un déplacement non-rigide, biaisé par un mouvement rotatif de la caméra et probablement sujet à une importante quantité de bruit de mesure systématique, aléatoire et de valeurs

aberrantes à cause de la nature des données. Pour alléger le coût de ce type d'approches, une parallélisation est possible puisque les traitements dans notre contexte sont effectués après acquisition de l'ensemble des données et non image par image comme dans les applications de vision temps réel.

Nous avons en l'occurrence proposé l'utilisation des techniques variationnelles de flot optique qui, après une initialisation et une paramétrisation adéquates, fournissent un résultat plus dense, plus lisse et plus adapté à un déplacement non paramétrique qu'une corrélation croisée. La considération d'une image maîtresse glissante pendant le co-recalage des données intra time-lapse et la construction d'un graphe de dates à portée temporelle maximale entre lesquelles effectuer des mesures pendant la phase de régularisation permettent en outre de prendre en compte la décorrélation photométrique des images. De plus, la construction d'indicateurs de qualité des mesures monodates (visibilité et dégradation des acquisitions) et bidates (cohérence d'un champ de déplacement) permet de ne propager que l'information cohérente au travers des différentes briques de traitements.

Enfin, il est important de disposer de mesures GPS bien réparties au sol et sur un, voire deux, des plans image et de procéder à un ajustement des faisceaux afin de modéliser la topographie de la scène avec avec un minimum d'ambiguïté projective et un maximum de cohérence stéréoscopique. Nous avons également déduit qu'un modèle 3D unique permet de minimiser les erreurs de projection métriques. Celui-ci devrait alors minimiser la dispersion de reproductibilité sur un ensemble de couples stéréoscopiques témoins disposant de bons indicateurs monodates.

Perspectives

Hormis les propositions d'amélioration et d'extension énoncées à la fin de chaque chapitre, cette thèse nous a permis d'ouvrir de nombreuses perspectives plus générales quant à la réflexion autour du développement d'un outil de traitements automatiques de time-lapses stéréoscopiques de mouvements géodésiques.

Tout d'abord, nous notons qu'il est important de s'intéresser davantage aux techniques de débruitage en télédétection qui, à l'instar de la méthode SBAS (Small BAseline Subset), sont plus adaptées à l'échelle des données, à un déplacement dense et non paramétrique et ne souffrent pas des limitations engendrées par les contraintes de rapidité. D'un autre autre coté, il peut s'avérer avantageux de les coupler avec des méthodes robustes, tels les descripteurs invariants, les méthodes de flot optique ou encore le DeepLearning, qui sont en plein essor.

Dans ce sens, le développement des techniques de restauration et de rehaussement d'images pourraient être exploitées pendant la phase de préparation des données à traiter. Des descripteurs monodates dédiés pourraient être développés par exemple par la prise en compte de contraintes stéréoscopiques et temporelles dans la phase de mise en

correspondance bidates monoscopiques.

Plus spécifiquement, pour des applications de surveillance glaciaire, des contraintes rhéologiques d'écoulement pour le suivi temporel pourraient permettre d'introduire de la spatialité dans les traitements temporels bi ou tri-dimensionnels. Le flot de scène peut aussi être résolu ou contraint via une modélisation rhéologique tridimensionnelle différentielle dont la solution peut être obtenue via une résolution numérique avec les techniques en éléments finis, par exemple via un logiciel de simulation d'écoulement glaciaire tel que Elmer/Ice.

Un axe de recherche intéressant impliquerait aussi une modélisation temporelle qui permettrait de combler les observations manquantes et d'effectuer des prévisions, par exemple via des méthodes du type filtre de Kalman.

L'idée d'exploitation de la redondance des mesures bidates par la construction d'un graphe de dates régi par une matrice d'adjacence peut être explorée davantage par étude des techniques développées en théorie des graphes.

D'un autre côté, les réflexions menées ainsi que les outils proposés au cours de cette étude peuvent facilement être transférés à d'autres applications de surveillance en milieux non contrôlés et ainsi contribuer à une meilleure compréhension d'autres phénomènes géophysiques dynamiques tels que les glissements de terrain ou encore les éruptions volcaniques. Une telle application permettrait en effet de révéler des dynamiques que nous n'aurions jamais pu observer à l'œil nu et pourrait en outre être déployée sur une échelle de temps différente, par exemple de l'ordre de l'heure ou du mois.

On the Glacier (Nietzsche 1884)

At noon, when
The summer first rises into the mountains,
The boy with the weary, burning eyes :
There, too, he speaks,
But we only *s*ee his speech.
His breath billows like an invalid's breath billows
In feverous night.
Icy mountain and fir and spring
Reply to him as well,
But we only see the reply.
For, as a greeting, the torrent
Drops down from the rocks
And stands like a white trembling pillar,
Longingly there.
And the fir looks even more somber and faithful
Than it usually looks
And between ice and deadly gray stone
Suddenly light breaks through — —
I have already seen such a light : this I realize. —
...

Bibliographie

[1] E. Berthier et al., « Surface motion of mountain glaciers derived from satellite optical imagery », *in* : *Remote Sensing of Environment* 95.1 (2005), p. 14 –28, issn : 0034-4257, doi : https://doi.org/10.1016/j.rse.2004.11.005, url : http://www.sciencedirect.com/science/article/pii/S0034425704003463.

[2] Ha-Thai Pham et al., « Analyse de ”Time-Lapse” stéréo pour la mesure de déformation 3D, application au suivi du glacier d'Argentière », *in* : *Reconnaissance de Formes et Intelligence Artificielle (RFIA) 2014*, France, juin 2014, url : https://hal.archives-ouvertes.fr/hal-00988890.

[3] Fabian Mathias Neyer, « Monitoring Rock Glaciers by Combining Photogrammetric and GNSS-Based Methods », thèse de doct., ETH Zürich, 2016.

[4] Anette Eltner et al., « Time lapse structure-from-motion photogrammetry for continuous geomorphic monitoring », *in* : *Earth Surface Processes and Landforms* (2017), doi : 10.1002/esp.4178.

[5] Richard Hartley et Andrew Zisserman, *Multiple View Geometry in Computer Vision*, 2e éd., New York, NY, USA : Cambridge University Press, 2003, isbn : 0521540518.

[6] Seungyong Lee et Sunghyun Cho, « Recent advances in image deblurring », *in* : nov. 2013, doi : 10.1145/2542266.2542272.

[7] Xue-Min Yang, Paul M. Davis et James H. Dieterich, « Deformation from inflation of a dipping finite prolate spheroid in an elastic half-space as a model for volcanic stressing », *in* : *Journal of Geophysical Research : Solid Earth* 93.B5 (1988), p. 4249–4257, issn : 2156-2202, doi : 10.1029/JB093iB05p04249, url : http://dx.doi.org/10.1029/JB093iB05p04249.

[8] Kui Zhang et al., « Interferometric phase reconstruction using simplified coherence network », *in* : *ISPRS Journal of Photogrammetry and Remote Sensing* 119 (sept. 2016), p. 1–9, doi : 10.1016/j.isprsjprs.2016.05.002.

[9] Albert Tarantola, « A strategy for nonlinear elastic inversion of seismic reflection data », *in* : *GEOPHYSICS* 51.10 (1986), p. 1893–1903, doi : 10.1190/1.1442046, eprint : https://doi.org/10.1190/1.1442046, url : https://doi.org/10.1190/1.1442046.

[10] Jorge Nocedal et Stephen J. Wright, *Numerical Optimization*, second, New York, NY, USA : Springer, 2006.

[11] Y. Saad, *Iterative Methods for Sparse Linear Systems*, Second, Society for Industrial et Applied Mathematics, 2003, doi : 10.1137/1.9780898718003, eprint : https://epubs.siam.org/doi/pdf/10.1137/1.9780898718003, url : https://epubs.siam.org/doi/abs/10.1137/1.9780898718003.

[12] Kenneth LEVENBERG, « A method for the solution of certain non-linear problems in least squares », *in* : *Quarterly of Applied Mathematics* 2.2 (1944), p. 164–168, issn : 0033569X, 15524485, url : http://www.jstor.org/stable/43633451.

[13] D. Marquardt, « An Algorithm for Least-Squares Estimation of Nonlinear Parameters », *in* : *Journal of the Society for Industrial and Applied Mathematics* 11.2 (1963), p. 431–441, doi : 10.1137/0111030, eprint : https://doi.org/10.1137/0111030, url : https://doi.org/10.1137/0111030.

[14] Chavent, G. et Kunisch, K., « Regularization of linear least squares problems by total bounded variation », *in* : *ESAIM : COCV* 2 (1997), p. 359–376, doi : 10.1051/cocv:1997113, url : https://doi.org/10.1051/cocv:1997113.

[15] C Ferrero, *Vocabulaire international des termes fondamentaux et généraux de métrologie*, avr. 2009, doi : 10.13140/RG.2.1.3771.4320.

[16] H. Bayram, J. V. Hook et V. Isler, « Gathering Bearing Data for Target Localization », *in* : *IEEE Robotics and Automation Letters* 1.1 (jan. 2016), p. 369–374, issn : 2377-3766, doi : 10.1109/LRA.2016.2521387.

[17] Alonzo Kelly, « Precision dilution in triangulation based mobile robot position estimation », *in* : *Intelligent Autonomous Systems* 8 (jan. 2003), issn : 1586034146.

[18] S. G. Narasimhan et S. K. Nayar, « Chromatic framework for vision in bad weather », *in* : *Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No.PR00662)*, t. 1, juin 2000, 598–605 vol.1, doi : 10.1109/CVPR.2000.855874.

[19] Srinivasa G. Narasimhan et Shree K. Nayar, « Vision and the Atmosphere », *in* : *International Journal of Computer Vision* 48.3 (juil. 2002), p. 233–254, issn : 1573-1405, doi : 10.1023/A:1016328200723, url : https://doi.org/10.1023/A:1016328200723.

[20] S. G. Narasimhan et S. K. Nayar, « Contrast restoration of weather degraded images », *in* : *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25.6 (juin 2003), p. 713–724, issn : 0162-8828, doi : 10.1109/TPAMI.2003.1201821.

[21] John Y. Chiang, Ying-Ching Chen et Yung-Fu Chen, « Underwater Image Enhancement : Using Wavelength Compensation and Image Dehazing (WCID) », *in* : août 2011, p. 372–383, doi : 10.1007/978-3-642-23687-7_34.

[22] Wei-Ta Chu, Xiang-You Zheng et Ding-Shiuan Ding, « Camera as weather sensor : Estimating weather information from single images », *in* : *Journal of Visual Communication and Image Representation* 46 (2017), p. 233 –249, issn : 1047-3203, doi : https://doi.org/10.1016/j.jvcir.2017.04.002, url : http://www.sciencedirect.com/science/article/pii/S1047320317300901.

[23] D. Berman, T. Treibitz et S. Avidan, « Non-Local Image Dehazing », *in* : *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, doi : 10.1109/CVPR.2016.185.

[24] Ko Nishino, Louis Kratz et Stephen Lombardi, « Bayesian Defogging », *in* : *International Journal of Computer Vision* 98.3 (juil. 2012), p. 263–278, issn : 1573-1405, doi : 10.1007/s11263-011-0508-1, url : https://doi.org/10.1007/s11263-011-0508-1.

[25] Kaiming He, Jian Sun et Xiaoou Tang, « Single Image Haze Removal Using Dark Channel Prior », *in* : *IEEE transactions on pattern analysis and machine intelligence* 33 (août 2010), doi : 10.1109/TPAMI.2010.168.

[26] J. Tarel et N. Hautière, « Fast visibility restoration from a single color or gray level image », *in* : *2009 IEEE 12th International Conference on Computer Vision*, sept. 2009, p. 2201–2208, doi : 10.1109/ICCV.2009.5459251.

[27] R. T. Tan, « Visibility in bad weather from a single image », *in* : *2008 IEEE Conference on Computer Vision and Pattern Recognition*, juin 2008, p. 1–8, doi : 10.1109/CVPR.2008.4587643.

[28] Raanan Fattal, « Single Image Dehazing », *in* : *ACM Trans. Graph.* 27.3 (août 2008), 72 :1–72 :9, issn : 0730-0301, doi : 10.1145/1360612.1360671, url : http://doi.acm.org/10.1145/1360612.1360671.

[29] Jue Wang et Michael Cohen, « Image and Video Matting : A Survey », *in* : *Foundations and Trends in Computer Graphics and Vision* 3 (jan. 2007), p. 97–175, doi : 10.1561/0600000019.

[30] G. W. Larson, H. Rushmeier et C. Piatko, « A visibility matching tone reproduction operator for high dynamic range scenes », *in* : *IEEE Transactions on Visualization and Computer Graphics* 3.4 (oct. 1997), p. 291–306, issn : 1077-2626, doi : 10.1109/2945.646233.

[31] Pruthvi Venkatesh Gatti et Krishna Teja Velugubantla, « Contrast Enhancement of Colour Images using Transform Based Gamma Correction and Histogram Equalization », *in* : *International Journal of Computing* (2017), issn : 23125381.

[32] R. M. Haralick, K. Shanmugam et I. Dinstein, « Textural Features for Image Classification », *in* : *IEEE Transactions on Systems, Man, and Cybernetics* SMC-3.6 (nov. 1973), p. 610–621, issn : 0018-9472, doi : 10.1109/TSMC.1973.4309314.

[33] Carole Delenne et al., « Textural approaches for vineyard detection and characterization using very high spatial resolution remote sensing data », *in* : *International Journal of Remote Sensing* 29.4 (fév. 2008), p. 1153–1167, doi : 10.1080/01431160701311259, url : https://hal.archives-ouvertes.fr/hal-01196900.

[34] Yashon Ouma, Josaphat Sri Sumantyo et R Tateishi, « Analysis of co-occurrence and discrete wavelet transform textures for differentiation of forest and non-forest vegetation in very-high-resolution optical-sensor imagery », *in* : *International Journal of Remote Sensing - INT J REMOTE SENS* 29 (juin 2008), p. 3417–3456, doi : 10.1080/01431160701601782.

[35] Martin Herold, Xiao Hang Liu et Keith Clarke, « Spatial Metrics and Image Texture for Mapping Urban Land Use », *in* : *Photogrammetric Engineering and Remote Sensing* 69 (sept. 2003), p. 991–1001, doi : 10.14358/PERS.69.9.991.

[36] T. Ojala, M. Pietikainen et T. Maenpaa, « Multiresolution gray-scale and rotation invariant texture classification with local binary patterns », *in* : *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24.7 (juil. 2002), p. 971–987, issn : 0162-8828, doi : 10.1109/TPAMI.2002.1017623.

[37] N. Dalal et B. Triggs, « Histograms of oriented gradients for human detection », *in* : *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, t. 1, juin 2005, 886–893 vol. 1, doi : 10.1109/CVPR.2005.177.

[38] Mallat Stéphane, « CHAPTER 4 - Time Meets Frequency », *in* : *A Wavelet Tour of Signal Processing (Third Edition)*, sous la dir. de Mallat Stéphane, Third Edition, Boston : Academic Press, 2009, p. 89 –153, isbn : 978-0-12-374370-1, doi : https://doi.org/10.1016/B978-0-12-374370-1.00008-2, url : http://www.sciencedirect.com/science/article/pii/B9780123743701000082.

[39] Mallat Stéphane, « CHAPTER 7 - Wavelet Bases », *in* : *A Wavelet Tour of Signal Processing (Third Edition)*, sous la dir. de Mallat Stéphane, Third Edition, Boston : Academic Press, 2009, p. 263 –376, isbn : 978-0-12-374370-1, doi : https://doi.org/10.1016/B978-0-12-374370-1.00011-2, url : http://www.sciencedirect.com/science/article/pii/B9780123743701000112.

[40] and D. M. Healy, « Contrast enhancement via multiscale gradient transformation », *in* : *Proceedings of 1st International Conference on Image Processing*, t. 2, nov. 1994, 482–486 vol.2, doi : 10.1109/ICIP.1994.413617.

[41] C. Lu et al., « Two-Class Weather Classification », *in* : *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39.12 (déc. 2017), p. 2510–2524, issn : 0162-8828, doi : 10.1109/TPAMI.2016.2640295.

[42] Cosmin Ancuti, Codruta O Ancuti et Radu Timofte, « NTIRE 2018 Challenge on Image Dehazing : Methods and Results », *in* : *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2018, p. 891–901.

[43] J. Mairal et al., « Discriminative learned dictionaries for local image analysis », *in* : *2008 IEEE Conference on Computer Vision and Pattern Recognition*, juin 2008, p. 1–8, doi : 10.1109/CVPR.2008.4587652.

[44] Harry Pratt et al., « FCNN : Fourier Convolutional Neural Networks », *in* : *Machine Learning and Knowledge Discovery in Databases*, sous la dir. de Michelangelo Ceci et al., Cham : Springer International Publishing, 2017, p. 786–798, isbn : 978-3-319-71249-9.

[45] S. Luan et al., « Gabor Convolutional Networks », *in* : *IEEE Transactions on Image Processing* 27.9 (sept. 2018), p. 4357–4366, issn : 1057-7149, doi : 10.1109/TIP.2018.2835143.

[46] Travis Williams et Robert Li, « Wavelet Pooling for Convolutional Neural Networks », *in* : *International Conference on Learning Representations*, 2018, url : https://openreview.net/forum?id=rkhlb8lCZ.

[47] P. Viola et W. M. Wells, « Alignment by maximization of mutual information », *in* : *Proceedings of IEEE International Conference on Computer Vision*, juin 1995, p. 16–23, doi : 10.1109/ICCV.1995.466930.

[48] Flavien Vernier et al., « Fast correlation technique for glacier flow monitoring by digital camera and space-borne SAR images », *in* : *EURASIP Journal on Image and Video Processing* 2011.1 (sept. 2011), p. 11, issn : 1687-5281, doi : 10.1186/1687-5281-2011-11, url : https://doi.org/10.1186/1687-5281-2011-11.

[49] G. Wolberg et S. Zokai, « Robust image registration using log-polar transform », *in* : *Proceedings 2000 International Conference on Image Processing (Cat. No.00CH37101)*, t. 1, sept. 2000, 493–496 vol.1, doi : 10.1109/ICIP.2000.901003.

[50] Philippe Weinzaepfel et al., « DeepFlow : Large displacement optical flow with deep matching », *in* : *IEEE Intenational Conference on Computer Vision (ICCV)*, Sydney, Australia, déc. 2013, url : http://hal.inria.fr/hal-00873592.

[51] M. H. Sidram et Nagappa. U. Bhajantri, « Exploration of Normalized Cross Correlation to Track the Object through Various Template Updating Techniques », *in* : (2015).

[52] Berthold K.P. Horn et Brian G. Schunck, « Determining optical flow », *in* : *Artificial Intelligence* 17.1 (1981), p. 185 –203, issn : 0004-3702, doi : https://doi.org/10.1016/0004-3702(81)90024-2, url : http://www.sciencedirect.com/science/article/pii/0004370281900242.

[53] Bruce D. Lucas et Takeo Kanade, « An Iterative Image Registration Technique with an Application to Stereo Vision (IJCAI) », *in* : t. 81, avr. 1981.

[54] Luis Alvarez et al., « Symmetrical Dense Optical Flow Estimation with Occlusions Detection », *in* : *International Journal of Computer Vision* 75.3 (déc. 2007), p. 371–385, issn : 1573-1405, doi : 10.1007/s11263-007-0041-4, url : https://doi.org/10.1007/s11263-007-0041-4.

[55] Thomas J. Burns et al., « Discrete, spatiotemporal, wavelet multiresolution analysis method for computing optical flow », *in* : *Optical Engineering* 33 (juil. 1994), p. 2236–2247, doi : 10.1117/12.172250.

[56] B. K.P. Horn, *Shape from shading : A method for obtaining the shape of a smooth opaque object from one view*, rapp. tech., Cambridge, MA, USA, 1970.

[57] D. Sun, S. Roth et M. J. Black, « Secrets of optical flow estimation and their principles », *in* : *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, juin 2010, p. 2432–2439, doi : 10.1109/CVPR.2010.5539939.

[58] Denis Fortun, Patrick Bouthemy et Charles Kervrann, « Optical flow modeling and computation : a survey », *in* : *Computer Vision and Image Understanding* 134 (mai 2015), p. 21, url : https://hal.inria.fr/hal-01104081.

[59] Deqing Sun, Stefan Roth et Michael J. Black, « A Quantitative Analysis of Current Practices in Optical Flow Estimation and the Principles behind Them », *in* : *International Journal of Computer Vision (IJCV)* 106.2 (2014), p. 115–137.

[60] Tianshu Liu et Lixin Shen, « Fluid flow and optical flow », *in* : *Journal of Fluid Mechanics* 614 (nov. 2008), p. 253 –291, doi : 10.1017/S0022112008003273.

[61] Anette Eltner et al., « Time lapse structure-from-motion photogrammetry for continuous geomorphic monitoring », *in* : *Earth Surface Processes and Landforms* 42.14 (2017), p. 2240–2253, doi : 10.1002/esp.4178, eprint : https://onlinelibrary.wiley.com/doi/pdf/10.1002/esp.4178, url : https://onlinelibrary.wiley.com/doi/abs/10.1002/esp.4178.

[62] K. R. Namuduri, R. Mehrotra et N. Ranganathan, « Edge detection models based on Gabor filters », *in* : *Proceedings., 11th IAPR International Conference on Pattern Recognition. Vol. III. Conference C : Image, Speech and Signal Analysis,* août 1992, p. 729–732, doi : 10.1109/ICPR.1992.202090.

[63] Radu Horaud, Thomas Skordas et Françoise Veillon, « Finding Geometric and Relational Structures in An Image », *in* : *1st European Conference on Computer Vision (ECCV'90)*, sous la dir. d'Olivier Faugeras, t. 427/1990, Lecture Notes in Computer Science (LNCS), Antibes, France : Springer-Verlag, avr. 1990, p. 374–384, doi : 10.1007/BFb0014886, url : https://hal.inria.fr/inria-00589995.

[64] Erez Shilat, Michael Werman et Yoram Gdalyahu, « Ridge's Corner Detection and Correspondence », *in* : *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1997, p. 976–981.

[65] Chris Harris et Mike Stephens, « A combined corner and edge detector », *in* : *In Proc. of Fourth Alvey Vision Conference*, 1988, p. 147–151.

[66] Lin Zhang, Lei Zhang et David Zhang, « A Multi-scale Bilateral Structure Tensor Based Corner Detector », *in* : *Computer Vision – ACCV 2009*, sous la dir. d'Hongbin Zha, Rin-ichiro Taniguchi et Stephen Maybank, Berlin, Heidelberg : Springer Berlin Heidelberg, 2010, p. 618–627, isbn : 978-3-642-12304-7.

[67] Lionel Gueguen et Martino Pesaresi, « Multi scale Harris corner detector based on Differential Morphological Decomposition », *in* : *Pattern Recognition Letters* 32 (oct. 2011), p. 1714–1719, doi : `10.1016/j.patrec.2011.07.021`.

[68] Krystian Mikolajczyk et Cordelia Schmid, « An affine invariant interest point detector », *in* : *7th European Conference on Computer Vision (ECCV '02)*, sous la dir. d'Anders Heyden et al., t. 2350, Lecture Notes in Computer Science (LNCS), Copenhagen, Denmark : Springer-Verlag, mai 2002, p. 128–142, doi : `10.1007/3-540-47969-4\_9`, url : `https://hal.inria.fr/inria-00548252`.

[69] Jianbo Shi et Tomasi, « Good features to track », *in* : *1994 Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1994, p. 593–600.

[70] Tony Lindeberg, « Discrete Scale-Space Theory and the Scale-Space Primal Sketch », thèse de doct., mai 1991.

[71] David Lowe, « Distinctive Image Features from Scale-Invariant Keypoints », *in* : *International Journal of Computer Vision* 60 (nov. 2004), p. 91–, doi : `10.1023/B:VISI.0000029664.99615.94`.

[72] Yan Ke et R. Sukthankar, « PCA-SIFT : a more distinctive representation for local image descriptors », *in* : *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.* T. 2, juin 2004, p. II–II, doi : `10.1109/CVPR.2004.1315206`.

[73] Herbert Bay, Tinne Tuytelaars et Luc Van Gool, « SURF : Speeded Up Robust Features », *in* : *Computer Vision – ECCV 2006*, sous la dir. d'Aleš Leonardis, Horst Bischof et Axel Pinz, Berlin, Heidelberg : Springer Berlin Heidelberg, 2006, p. 404–417, isbn : 978-3-540-33833-8.

[74] Tony Lindeberg, « Feature Detection with Automatic Scale Selection », *in* : *International Journal of Computer Vision* 30.2 (nov. 1998), p. 79–116, issn : 1573-1405, doi : `10.1023/A:1008045108935`, url : `https://doi.org/10.1023/A:1008045108935`.

[75] Tony Lindeberg et Jonas Gårding, « Shape-adapted smoothing in estimation of 3-D shape cues from affine deformations of local 2-D brightness structure », *in* : *Image and Vision Computing* 15.6 (1997), p. 415 –434, issn : 0262-8856, doi : https://doi.org/10.1016/S0262-8856(97)01144-X, url : http://www.sciencedirect.com/science/article/pii/S026288569701144X.

[76] K. Mikolajczyk et C. Schmid, « A performance evaluation of local descriptors », *in* : *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27.10 (oct. 2005), p. 1615–1630, issn : 0162-8828, doi : 10.1109/TPAMI.2005.188.

[77] S. Periaswamy et H. Farid, « Elastic registration in the presence of intensity variations », *in* : *IEEE Transactions on Medical Imaging* 22.7 (juil. 2003), p. 865–874, issn : 0278-0062, doi : 10.1109/TMI.2003.815069.

[78] Petros Maragos Chiou-Shann Fuh, « Motion displacement estimation using an affine model for image matching », *in* : *Optical Engineering* 30.7 (1991), p. 881 –887 –7, doi : 10.1117/12.55885, url : https://doi.org/10.1117/12.55885.

[79] Cordelia Schmid, Roger Mohr et Christian Bauckhage, « Evaluation of Interest Point Detectors », *in* : *International Journal of Computer Vision* 37.2 (juin 2000), p. 151–172, issn : 1573-1405, doi : 10.1023/A:1008199403446, url : https://doi.org/10.1023/A:1008199403446.

[80] K. Mikolajczyk et al., « A Comparison of Affine Region Detectors », *in* : *International Journal of Computer Vision* 65.1 (nov. 2005), p. 43–72, issn : 1573-1405, doi : 10.1007/s11263-005-3848-x, url : https://doi.org/10.1007/s11263-005-3848-x.

[81] Marius Muja et David G. Lowe, « Fast approximate nearest neighbors with automatic algorithm configuration », *in* : *In VISAPP International Conference on Computer Vision Theory and Applications*, 2009, p. 331–340.

[82] Connelly Barnes et al., « PatchMatch : A Randomized Correspondence Algorithm for Structural Image Editing », *in* : *ACM Transactions on Graphics (Proc. SIGGRAPH)* 28.3 (août 2009).

[83] O. Chum et J. Matas, « Matching with PROSAC - progressive sample consensus », *in* : *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, t. 1, juin 2005, 220–226 vol. 1, doi : 10.1109/CVPR.2005.221.

[84] Frédéric Cao et al., « A unified framework for detecting groups and application to shape recognition », *in* : *Journal of Mathematical Imaging and Vision* 27.2 (2007), p. 91–119, doi : 10.1007/s10851-006-9176-0, url : https://hal.inria.fr/inria-00104255.

[85] Vittorio Ferrari, Tinne Tuytelaars et Luc Van Gool, « Simultaneous Object Recognition and Segmentation from Single or Multiple Model Views », *in* : *International Journal of Computer Vision* 67.2 (avr. 2006), p. 159–188, issn : 1573-1405, doi : 10.1007/s11263-005-3964-7, url : https://doi.org/10.1007/s11263-005-3964-7.

[86] I. G. Karybali et al., « Efficient image registration with subpixel accuracy », *in* : *2006 14th European Signal Processing Conference*, sept. 2006, p. 1–5.

[87] Bruce D. Lucas et Takeo Kanade, « An Iterative Image Registration Technique with an Application to Stereo Vision (IJCAI) », *in* : t. 81, avr. 1981.

[88] Simon Baker et Iain Matthews, « Lucas-Kanade 20 Years On : A Unifying Framework », *in* : *International Journal of Computer Vision* 56.3 (fév. 2004), p. 221–255, issn : 1573-1405, doi : 10.1023/B:VISI.0000011205.11775.fd, url : https://doi.org/10.1023/B:VISI.0000011205.11775.fd.

[89] A. Bartoli, « Groupwise Geometric and Photometric Direct Image Registration », *in* : *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30.12 (déc. 2008), p. 2098–2108, issn : 0162-8828, doi : 10.1109/TPAMI.2008.22.

[90] G. D. Evangelidis et E. Z. Psarakis, « Parametric Image Alignment Using Enhanced Correlation Coefficient Maximization », *in* : *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30.10 (oct. 2008), p. 1858–1865, issn : 0162-8828, doi : 10.1109/TPAMI.2008.113.

[91] L. Bao, Q. Yang et H. Jin, « Fast Edge-Preserving PatchMatch for Large Displacement Optical Flow », *in* : *IEEE Transactions on Image Processing* 23.12 (2014), p. 4996–5006, doi : 10.1109/TIP.2014.2359374.

[92] Panasonic Corporation, *Lumix digital camera know-hows*, 2019, url : https://av.jpn.support.panasonic.com/support/global/cs/dsc/knowhow/knowhow12.html.

[93] Jonathan Carrivick, Mark Smith et Duncan J. Quincey, *Structure from Motion in the Geosciences*, juil. 2016, isbn : ISBN : 978-1-118-89584-9.

[94] Robert Woodham, « Photometric Method for Determining Surface Orientation from Multiple Images », *in* : *Optical Engineering* 19 (jan. 1992), doi : 10.1117/12.7972479.

[95] Andrew P. Witkin, « Recovering surface shape and orientation from texture », *in* : *Artificial Intelligence* 17.1 (1981), p. 17 –45, issn : 0004-3702, doi : https://doi.org/10.1016/0004-3702(81)90019-9, url : http://www.sciencedirect.com/science/article/pii/0004370281900199.

[96] Faraj Alhwarin et al., « Optimized KinectFusion Algorithm for 3D Scanning Applications », *in* : jan. 2018, doi : 10.5220/0006594700500057.

[97] A. Criminisi, I. Reid et A. Zisserman, « Single View Metrology », *in* : *International Journal of Computer Vision* 40.2 (nov. 2000), p. 123–148, issn : 1573-1405, doi : 10.1023/A:1026598000963, url : https://doi.org/10.1023/A:1026598000963.

[98] A. Bartoli et al., « Shape-from-Template », *in* : *IEEE Transactions on Pattern Analysis and Machine Intelligence* 37.10 (oct. 2015), p. 2099–2118, issn : 0162-8828, doi : 10.1109/TPAMI.2015.2392759.

[99] Richard I. Hartley et Peter Sturm, « Triangulation », *in* : *Computer Vision and Image Understanding* 68.2 (1997), p. 146 –157, issn : 1077-3142, doi : https://doi.org/10.1006/cviu.1997.0547, url : http://www.sciencedirect.com/science/article/pii/S1077314297905476.

[100] Quang-Tuan Luong, « Fundamental matrix and self-calibration in computer vision. », Theses, Université Paris Sud - Paris XI, déc. 1992, url : https://tel.archives-ouvertes.fr/tel-00549134.

[101] Anders Heyden et Kalle Åström, « Minimal Conditions on Intrinsic Parameters for Euclidean Reconstruction », *in* : *ACCV*, 1998.

[102] H.C. Longuet-Higgins, « A computer algorithm for reconstructing a scene from two projections », *in* : *Readings in Computer Vision*, sous la dir. de Martin A. Fischler et Oscar Firschein, San Francisco (CA) : Morgan Kaufmann, 1987, p. 61 –62, isbn : 978-0-08-051581-6, doi : https://doi.org/10.1016/B978-0-08-051581-6.50012-X, url : http://www.sciencedirect.com/science/article/pii/B978008051581650012X.

[103] S. Rusinkiewicz et M. Levoy, « Efficient variants of the ICP algorithm », *in* : *Proceedings Third International Conference on 3-D Digital Imaging and Modeling*, mai 2001, p. 145–152, doi : 10.1109/IM.2001.924423.

[104] Berthold K. P. Horn, « Closed-form solution of absolute orientation using unit quaternions », *in* : *J. Opt. Soc. Am. A* 4.4 (1987), p. 629–642, doi : 10.1364/JOSAA.4.000629, url : http://josaa.osa.org/abstract.cfm?URI=josaa-4-4-629.

[105] Stephen Fox et Damian Lyons, « An approach to stereo-point cloud registration using image homographies », *in* : 8301 (jan. 2012), p. 7–, doi : 10.1117/12.908968.

[106] Sundar Vedula et al., « Three-dimensional scene flow », *in* : *IEEE transactions on pattern analysis and machine intelligence* 27 (avr. 1999), p. 475–80, doi : 10.1109/TPAMI.2005.63.

[107] Ludovic Ravanel, Philip Deline et Stéphane Jaillet, « Quatre années de suivi de la morphodynamique des parois rocheuses du massif du Mont Blanc par laserscanning terrestre », *in* : *Images et modèles 3D en milieux naturels*, Collection EDYTEM n° 12, 2011, p. 69–76, url : https://hal.archives-ouvertes.fr/halsde-00604299.

[108] Haixing He et al., « 3D displacement retrieval on a scaled model of mountain slope by virtual multi-view photogrammetry », *in* : *2nd Virtual Geoscience Conference*, Bergen, Norway, 2016, URL : https://hal-sde.archives-ouvertes.fr/hal-01880618.

[109] Andreas Wedel et al., « Stereoscopic Scene Flow Computation for 3D Motion Understanding », *in* : *International Journal of Computer Vision* 95.1 (oct. 2011), p. 29–51, ISSN : 1573-1405, DOI : 10.1007/s11263-010-0404-0, URL : https://doi.org/10.1007/s11263-010-0404-0.

[110] René Schuster et al., « Combining Stereo Disparity and Optical Flow for Basic Scene Flow », *in* : jan. 2018, p. 90–101, ISBN : 978-3-658-21299-5, DOI : 10.1007/978-3-658-21300-8_8.

[111] Clemens Rabe et al., « Dense, Robust, and Accurate Motion Field Estimation from Stereo Image Sequences in Real-Time », *in* : *Computer Vision – ECCV 2010*, sous la dir. de Kostas Daniilidis, Petros Maragos et Nikos Paragios, Berlin, Heidelberg : Springer Berlin Heidelberg, 2010, p. 582–595, ISBN : 978-3-642-15561-1.

[112] Y. Zhang et C. Kambhamettu, « Integrated 3D scene flow and structure recovery from multiview image sequences », *in* : *Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No.PR00662)*, t. 2, juin 2000, 674–681 vol.2, DOI : 10.1109/CVPR.2000.854939.

[113] T. Basha, Y. Moses et N. Kiryati, « Multi-view scene flow estimation : A view centered variational approach », *in* : *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, juin 2010, p. 1506–1513, DOI : 10.1109/CVPR.2010.5539791.

[114] Nikolaus Mayer et al., « A Large Dataset to Train Convolutional Networks for Disparity, Optical Flow, and Scene Flow Estimation », *in* : *CoRR* abs/1512.02134 (2015), arXiv : 1512.02134, URL : http://arxiv.org/abs/1512.02134.

[115] Uwe Franke et al., « 6D-Vision : Fusion of Stereo and Motion for Robust Environment Perception », *in* : t. 3663, août 2005, p. 216–223, DOI : 10.1007/11550518_27.

[116] Xingyu Liu, Charles Ruizhongtai Qi et Leonidas J. Guibas, « Learning Scene Flow in 3D Point Clouds », *in* : *CoRR* abs/1806.01411 (2018).

[117] T. Basha et al., « Structure and motion from scene registration », *in* : *2012 IEEE Conference on Computer Vision and Pattern Recognition*, juin 2012, p. 1426–1433, DOI : 10.1109/CVPR.2012.6247830.

[118] J. Neumann et Y. Aloimonos, « Spatio-temporal stereo using multi-resolution subdivision surfaces », *in* : *Proceedings IEEE Workshop on Stereo and Multi-Baseline Vision (SMBV 2001)*, déc. 2001, p. 103–108, DOI : 10.1109/SMBV.2001.988768.

[119] Mariano Jaimez et al., « A primal-dual framework for real-time dense RGB-D scene flow », *in* : t. 2015, mai 2015, p. 98–104, doi : 10.1109/ICRA.2015.7138986.

[120] D. Sun, E. B. Sudderth et H. Pfister, « Layered RGBD scene flow estimation », *in* : *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, juin 2015, p. 548–556, doi : 10.1109/CVPR.2015.7298653.

[121] Zike Yan et Xuezhi Xiang, « Scene Flow Estimation : A Survey », *in* : *CoRR* abs/1612.02590 (2016).

[122] R. I. Hartley, « In defence of the 8-point algorithm », *in* : *Proceedings of IEEE International Conference on Computer Vision*, juin 1995, p. 1064–1070, doi : 10.1109/ICCV.1995.466816.

[123] David Nistér, « An Efficient Solution to the Five-Point Relative Pose Problem », *in* : *IEEE Trans. Pattern Anal. Mach. Intell.* 26.6 (juin 2004), p. 756–777, issn : 0162-8828, doi : 10.1109/TPAMI.2004.17, url : https://doi.org/10.1109/TPAMI.2004.17.

[124] Gang Li et S. W. Zucker, « Surface Geometric Constraints for Stereo in Belief Propagation », *in* : *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, t. 2, juin 2006, p. 2355–2362, doi : 10.1109/CVPR.2006.299.

[125] Bill Triggs et al., « Bundle Adjustment — A Modern Synthesis », *in* : *Vision Algorithms : Theory and Practice*, sous la dir. de Bill Triggs, Andrew Zisserman et Richard Szeliski, Berlin, Heidelberg : Springer Berlin Heidelberg, 2000, p. 298–372, isbn : 978-3-540-44480-0.

[126] W. Chojnacki et M. J. Brooks, « Revisiting Hartley's normalized eight-point algorithm », *in* : *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25.9 (2003), p. 1172–1177, issn : 0162-8828, doi : 10.1109/TPAMI.2003.1227992.

[127] A. Eltner et al., « Image-based surface reconstruction in geomorphometry – merits, limits and developments », *in* : *Earth Surface Dynamics* 4.2 (2016), p. 359–389, doi : 10.5194/esurf-4-359-2016, url : https://www.earth-surf-dynam.net/4/359/2016/.

[128] D. E. Clark et S. Ivekovic, « The Cramer-Rao Lower Bound for 3-D state estimation from rectified stereo cameras », *in* : *2010 13th International Conference on Information Fusion*, juil. 2010, p. 1–8, doi : 10.1109/ICIF.2010.5712095.

[129] Chienchung Chang, S. Chatterjee et P. R. Kube, « A quantization error analysis for convergent stereo », *in* : *Proceedings of 1st International Conference on Image Processing*, t. 2, nov. 1994, 735–739 vol.2, doi : 10.1109/ICIP.1994.413668.

[130] Jianyang Liu et Youfu Li, « Performance analysis of 3-D shape measurement algorithm with a short baseline projector-camera system », *in* : t. 1, déc. 2013, p. 1444–1449, isbn : 978-1-4799-2744-9, doi : 10.1109/ROBIO.2013.6739669.

[131] David Gallup et al., « Variable Baseline/Resolution Stereo », *in* : juin 2008, doi : 10.1109/CVPR.2008.4587671.

[132] Gustavo Olague et Roger Mohr, « Optimal camera placement for accurate reconstruction », *in* : *Pattern Recognition* 35.4 (2002), p. 927 –944, issn : 0031-3203, doi : https://doi.org/10.1016/S0031-3203(01)00076-0, url : http://www.sciencedirect.com/science/article/pii/S0031320301000760.

[133] G. Di Leo, C. Liguori et A. Paolillo, « Propagation of uncertainty through stereo triangulation », *in* : *2010 IEEE Instrumentation Measurement Technology Conference Proceedings*, mai 2010, p. 12–17, doi : 10.1109/IMTC.2010.5488057.

[134] W.W. Immerzeel et al., « High-resolution monitoring of Himalayan glacier dynamics using unmanned aerial vehicles », *in* : *Remote Sensing of Environment* 150 (2014), p. 93 –103, issn : 0034-4257, doi : https://doi.org/10.1016/j.rse.2014.04.025, url : http://www.sciencedirect.com/science/article/pii/S003442571400176X.

[135] Cheng Peng et Volkan Isler, « View Selection with Geometric Uncertainty Modeling », *in* : juin 2018, doi : 10.15607/RSS.2018.XIV.025.

[136] Cédric Demonceaux, « Habilitation à Diriger des Recherches : Étude de caméras sphériques : du traitement des images aux applications en robotique », *in* : (nov. 2012).

[137] Hoang Viet Tuan Nguyen et al., « A Pattern-Based Method For Handling Confidence Measures While Mining Satellite Displacement Field Time Series : Application to Greenland Ice Sheet and Alpine Glaciers », *in* : *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* PP (oct. 2018), p. 1–13, doi : 10.1109/JSTARS.2018.2874499.

[138] D. L. Danilov, « Principal Components in Time Series Forecast », *in* : *Journal of Computational and Graphical Statistics* 6.1 (1997), p. 112–121, issn : 10618600, url : http://www.jstor.org/stable/1390727.

[139] Yacine Aït-Sahalia et Dacheng Xiu, « Principal Component Analysis of High-Frequency Data », *in* : *Journal of the American Statistical Association* 114.525 (2019), p. 287–303, doi : 10.1080/01621459.2017.1401542, eprint : https://doi.org/10.1080/01621459.2017.1401542, url : https://doi.org/10.1080/01621459.2017.1401542.

[140] Y. Yan et al., « Mexico City Subsidence Measured by InSAR Time Series : Joint Analysis Using PS and SBAS Approaches », *in* : *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 5.4 (août 2012), p. 1312–1326, issn : 1939-1404, doi : 10.1109/JSTARS.2012.2191146.

[141] Ludmila I. Kuncheva et Juan J. Rodríguez, « An Experimental Study on Rotation Forest Ensembles », *in* : *Multiple Classifier Systems : 7th International Workshop, MCS 2007, Prague, Czech Republic, May 23-25, 2007. Proceedings*, Berlin, Heidelberg : Springer Berlin Heidelberg, 2007, p. 459–468, isbn : 978-3-540-72523-7, doi : 10.1007/978-3-540-72523-7_46, url : https://doi.org/10.1007/978-3-540-72523-7_46.

[142] Durdu Oemer Faruk, « A hybrid neural network and ARIMA model for water quality time series prediction », *in* : *Eng. Appl. of AI* 23 (juin 2010), p. 586–594.

[143] Sheng-Tun Li et al., « A vector forecasting model for fuzzy time series », *in* : *Applied Soft Computing* 11 (avr. 2011), p. 3125–3134.

[144] Marcus Schulz et Michael Matthies, « Artificial neural networks for modeling time series of beach litter in the southern North Sea », *in* : *Marine Environmental Research* 98.Supplement C (2014), p. 14 –20, issn : 0141-1136, doi : https://doi.org/10.1016/j.marenvres.2014.03.014, url : http://www.sciencedirect.com/science/article/pii/S0141113614000634.

[145] John Cristian Borges Gamboa, « Deep Learning for Time-Series Analysis », *in* : *CoRR* abs/1701.01887 (2017), arXiv : 1701.01887, url : http://arxiv.org/abs/1701.01887.

[146] M. Proesmans et al., « Determination of optical flow and its discontinuities using non-linear diffusion », *in* : *Computer Vision — ECCV '94*, sous la dir. de Jan-Olof Eklundh, Berlin, Heidelberg : Springer Berlin Heidelberg, 1994, p. 294–304, isbn : 978-3-540-48400-4.

[147] Alain Trouvé et Laurent Younes, « Diffeomorphic Matching Problems in One Dimension : Designing and Minimizing Matching Functionals. », *in* : t. 1842, jan. 2000, p. 573–587.

[148] Yunmei Chen et Xiaojing Ye, « Inverse Consistent Deformable Image Registration », *in* : *The Legacy of Alladi Ramakrishnan in the Mathematical Sciences*, sous la dir. de Krishnaswami Alladi, John R. Klauder et Calyampudi R. Rao, New York, NY : Springer New York, 2010, p. 419–440, isbn : 978-1-4419-6263-8, doi : 10.1007/978-1-4419-6263-8_26, url : https://doi.org/10.1007/978-1-4419-6263-8_26.

[149] John Ashburner et Gerard Ridgway, « Symmetric Diffeomorphic Modeling of Longitudinal Structural MRI », *in* : *Frontiers in Neuroscience* 6 (2013), p. 197, issn : 1662-453X, doi : 10.3389/fnins.2012.00197, url : https://www.frontiersin.org/article/10.3389/fnins.2012.00197.

[150] Iman Aganj et al., « Mid-Space-Independent Deformable Image Registration », *in* : *NeuroImage* 152 (fév. 2017), doi : 10.1016/j.neuroimage.2017.02.055.

[151] S. Volz et al., « Modeling temporal coherence for optical flow », *in* : *2011 International Conference on Computer Vision*, nov. 2011, p. 1116–1123, doi : 10.1109/ICCV.2011.6126359.

[152] P. Berardino et al., « A new algorithm for surface deformation monitoring based on small baseline differential SAR interferograms », *in* : *IEEE Transactions on Geoscience and Remote Sensing* 40.11 (nov. 2002), p. 2375–2383, issn : 0196-2892, doi : 10.1109/TGRS.2002.803792.

[153] T. R. Lauknes, H. A. Zebker et Y. Larsen, « InSAR Deformation Time Series Using an L_1 -Norm Small-Baseline Approach », *in* : *IEEE Transactions on Geoscience and Remote Sensing* 49.1 (jan. 2011), p. 536–546, issn : 0196-2892, doi : 10.1109/TGRS.2010.2051951.

[154] Kanika Goel et Nico Adam, « An advanced algorithm for deformation estimation in non-urban areas », *in* : *ISPRS Journal of Photogrammetry and Remote Sensing* 73 (2012), Innovative Applications of SAR Interferometry from modern Satellite Sensors, p. 100 – 110, issn : 0924-2716, doi : https://doi.org/10.1016/j.isprsjprs.2012.06.001, url : http://www.sciencedirect.com/science/article/pii/S0924271612001013.

[155] Sergey V. Samsonov et al., « Multidimensional Small Baseline Subset (MSBAS) for volcano monitoring in two dimensions : Opportunities and challenges. Case study Piton de la Fournaise volcano », *in* : *Journal of Volcanology and Geothermal Research* (avr. 2017), doi : 10.1016/j.jvolgeores.2017.04.017.

[156] Riccardo Lanari et al., « An Overview of the Small BAseline Subset Algorithm : A DInSAR Technique for Surface Deformation Analysis », *in* : *Deformation and Gravity Change : Indicators of Isostasy, Tectonics, Volcanism, and Climate Change*, sous la dir. de Detlef Wolf et José Fernández, Basel : Birkhäuser Basel, 2007, p. 637–661, isbn : 978-3-7643-8417-3.

[157] Marie-Pierre Doin et al., « Presentation of the small baseline NSBAS processing chain on a case example : The Etna deformation monitoring from 2003 to 2010 using Envisat data », *in* : sept. 2011.

[158] CHRISTIAN VINCENT et LUC MOREAU, « Sliding velocity fluctuations and subglacial hydrology over the last two decades on Argentière glacier, Mont Blanc area », *in* : *Journal of Glaciology* 62.235 (2016), 805–815, doi : 10.1017/jog.2016.35.

[159] C. Vincent et al., « Glacier thickening and decay analysis from 50 years of glaciological observations performed on Glacier d'Argentière, Mont Blanc area, France », *in* : *Annals of Glaciology* 50.50 (2009), 73–79, doi : 10.3189/172756409787769500.

[160] Fanny Ponton et al., « Observation of the Argentière Glacier Flow Variability from 2009 to 2011 by TerraSAR-X and GPS Displacement Measurements », *in* : *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 7 (2014), p. 3274–3284.

[161] C. Vincent et al., « Glacier thickening and decay analysis from 50 years of glaciological observations performed on Glacier d'Argentière, Mont Blanc area, France », *in* : *Annals of Glaciology* 50.50 (2009), 73–79, doi : 10.3189/172756409787769500.

[162] Lionel Benoit et al., « Multi-method monitoring of Glacier d'Argentière dynamics », *in* : *Annals of Glaciology* 56.70 (2015), p. 118–128, doi : 10.3189/2015AoG70A985, url : https://hal.archives-ouvertes.fr/hal-01154174.

[163] E. Trouve et al., « Combining Airborne Photographs and Spaceborne SAR Data to Monitor Temperate Glaciers : Potentials and Limits », *in* : *IEEE Transactions on Geoscience and Remote Sensing* 45.4 (2007), p. 905–924, issn : 0196-2892, doi : 10.1109/TGRS.2006.890554.

[164] Antoine Rabatel et al., « Estimation of Glacier Thickness From Surface Mass Balance and Ice Flow Velocities : A Case Study on Argentière Glacier, France », *in* : *Frontiers in Earth Science* 6 (2018), p. 112, issn : 2296-6463, url : https://www.frontiersin.org/article/10.3389/feart.2018.00112.

[165] C. Vincent et al., « A Nonlinear Statistical Model for Extracting a Climatic Signal From Glacier Mass Balance Measurements », *in* : *Journal of Geophysical Research : Earth Surface* 123.9 (), p. 2228–2242, doi : 10.1029/2018JF004702, eprint : https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2018JF004702, url : https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018JF004702.

[166] Lionel Benoit et al., « Multi-method monitoring of Glacier d'Argentière dynamics », *in* : *Annals of Glaciology* 56.70 (2015), p. 118–128, doi : 10.3189/2015AoG70A985, url : https://hal.archives-ouvertes.fr/hal-01154174.

[167] Michael Zemp et al., « Historically unprecedented global glacier decline in the early 21st century », *in* : *Journal of Glaciology* 61.228 (2015), 745–762, doi : 10.3189/2015JoG15J017.

[168] Bernard Guy, Martin Daigneault et Gérard Thomas, « Reflections on the formation of Forbes ogives : The instability of fusion of dirty ice », *in* : *Comptes Rendus Geoscience* 334 (déc. 2002), p. 1061–1070, doi : 10.1016/S1631-0713(02)01855-2.

[169] Louis Lliboutry et Louis Reynaud, « "Global Dynamics" of a Temperate Valley Glacier, Mer De Glace, and Past Velocities Deduced from Forbes' Bands », *in* : *Journal of Glaciology* 27 (jan. 1981), p. 207–226, doi : 10.3189/s0022143000015367.

[170] Tavi Murray et al., « Dynamics of glacier calving at the ungrounded margin of Helheim Glacier, southeast Greenland », *in* : *Journal of Geophysical Research : Earth Surface* 120.6 (), p. 964–982, doi : 10.1002/2015JF003531, eprint : https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2015JF003531, url : https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015JF003531.

[171] Robert Vivian, « Le glacier d'Argentière », *in* : *Revue de Géographie Alpine* 55 (1967), p. 579–582, url : https://www.persee.fr/doc/rga_0035-1121_1967_num_55_3_3335.

[172] E. Schwalbe, R. Koschitzki et H.-G. Maas, « Recognition of drainage tunnels during glacier lake outburst events from terrestrial image sequences », *in* : *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLI-B8 (2016), p. 537–543, doi : 10.5194/isprs-archives-XLI-B8-537-2016, url : https://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XLI-B8/537/2016/.

[173] Xavier Bodin et al., « L'apport d'appareils photo reflex autonomes pour le suivi quasi-continu de la dynamique des glaciers rocheux. Exemple dans les Alpes et dans les Andes. », *in* : *Monitoring en milieu naturel : retours d'expériences en terrains difficiles.* Sous la dir. d'Emmanuel Malet et Laurent Astrade, 19, Collection EDYTEM, 2017, pp. 47–54, url : https://hal-sde.archives-ouvertes.fr/hal-02071517.

[174] Ring, Wolfgang, « Structural Properties of Solutions to Total Variation Regularization Problems », *in* : *ESAIM : M2AN* 34.4 (2000), p. 799–810, doi : 10.1051/m2an:2000104, url : https://doi.org/10.1051/m2an:2000104.

[175] M. J. D. Powell, « A Direct Search Optimization Method That Models the Objective and Constraint Functions by Linear Interpolation », *in* : *Advances in Optimization and Numerical Analysis*, sous la dir. de Susana Gomez et Jean-Pierre Hennart, Dordrecht : Springer Netherlands, 1994, p. 51–67, isbn : 978-94-015-8330-5, doi : 10.1007/978-94-015-8330-5_4, url : https://doi.org/10.1007/978-94-015-8330-5_4.

[176] Donald H. Kraft, « A **software package** for sequential quadratic programming », *in* : 1988, url : http://degenerateconic.com/wp-content/uploads/2018/03/DFVLR_FB_88_28.pdf.