

Traitement automatique de "time lapse" : application à la surveillance de glaciers alpins,

Automatic processing of time lapse : application to the monitoring of alpine glaciers.

H. Hadhri¹

F. Vernier¹

A. Atto¹

E. Trouvé¹

¹Laboratoire d'Informatique, Systèmes, Traitement de l'Information et de la Connaissance
Polytech Annecy-Chambéry, Université Savoie Mont Blanc.

Hela.Hadhri, Flavien.Vernier, Emmanuel.Trouve, Abdourrahmane.Atto @univ-smb.fr

Résumé

L'imagerie optique bénéficie d'une maturité méthodologique importante et elle est utilisée avec succès dans une large gamme d'applications en vision par ordinateur. En télédétection, ces techniques sont de plus en plus utilisées afin de suivre l'évolution et la dynamique d'objets géophysiques (glaciers, parois rocheuses...). Dans ce contexte particulier impacté par les conditions climatiques, avec des acquisitions hebdomadaires, voire plus fréquentes, des traitements spécifiques sont nécessaires afin d'extraire et analyser l'information pertinente de la masse d'images acquises. Cet article se focalise ainsi sur le traitement automatique de séries d'images "time lapse" à des fins de surveillance d'objets géophysiques. Nous illustrerons les méthodes proposées par des résultats obtenus sur un glacier situé dans le massif du Mont-Blanc.

Mots Clef

Séries temporelles d'images, Suivi d'objets, Traitement automatique, Glaciers.

Abstract

Optical imaging reached significant methodological maturity and is used with some success in a wide range of computer vision applications. In remote sensing, these techniques are increasingly used in order to monitor the evolution and dynamics of geophysical objects (glaciers, rock walls ...). In this particular context impacted by climatic conditions, with weekly or more frequent acquisitions, specific image processing is necessary in order to extract and analyze the relevant information from the mass of images acquired and to track the evolution of these objects. This article focuses on the automatic processing of time lapse series for the monitoring of geophysical objects. We will illustrate our result with a follow-up of a glacier located in the Mont Blanc massif.¹

1. This work was supported by PHOENIX ANR-15-CE23-0012 grant of the French National Agency of Research.

Keywords

Time lapse, Image time series, Object tracking, Automatic processing, Glaciers.

1 Introduction

L'étude de la morphologie, des déformations et de l'évolution d'objets géophysiques est une étape préalable essentielle à toute analyse poussée d'un environnement naturel. L'estimation des déplacements des glaciers permet notamment de mieux cerner leur géomorphologie et leur dynamique et ainsi de mieux appréhender les conséquences locales (dans le temps) du réchauffement climatique. En effet, quelques glaciers sont en train de survivre à la hausse de température globale de la planète (environ 0,85 °C les 100 dernières années) et afin d'établir un modèle scientifique formel de cette menace, les techniques de télédétection sont depuis une quarantaine d'années abondamment déployées afin de surveiller les masses glacières de notre planète et cela afin de caractériser, valider et contraindre des modèles de déformation liés à la glaciologie.

La photographie aérienne [4] et la photogrammétrie terrestre [5] se sont en effet avérées efficaces pour quantifier l'écoulement de glaciers. L'analyse de "time lapse" de glaciers sur des échelles annuelles à décennales a été menée, par exemple, par [6] sur des données Landsat ou [7] par photographie aérienne et altimétrie laser. Ces techniques de télédétection ont l'avantage de bénéficier d'une haute couverture spatiale et temporelle, mais il s'agit souvent de déployer des moyens au sol, tel que coins réflecteurs ou mesures GPS. Une observation automatisée, de plus haute résolution au sol et à une fréquence plus grande, de l'ordre de l'heure ou du jour, nécessite d'adapter les techniques d'imagerie optique au contexte d'observation de milieux naturels. L'auteur de [8] décrit la difficulté du calcul d'une vitesse absolue d'objets géophysiques à partir de séries d'images optiques dans un contexte similaire au notre. Cette difficulté réside essentiellement dans la détection et le maintien de points de contrôle (image) au cours du temps et dans la décorrélation des mouvements induits par l'in-

stabilité de la caméras. Plus récemment, [9] propose une technique supervisée permettant le calcul d'un champs de vitesses, pour la quantification de la dynamique de glaciers en Alaska, par corrélation croisée des différentes images de la série, suivie d'un raffinement itératif par relaxation probabiliste des champs calculés. Dans [3], les auteurs proposent une approche permettant un traitement de séries chronologiques d'images de glaciers qui exploite la redondance d'information dans une même série. En effet, l'analyse de surfaces par imagerie optique quoique bénéficiant d'une maturité méthodologique importante et ayant été utilisée avec succès dans une large gamme d'applications en vision par ordinateur, généralement dans des contextes industriels [21] ou médicaux [20], souffre d'une grande dépendance aux conditions d'acquisition et à la nature des changements observés. Les contraintes avec les images de milieux naturels, et plus particulièrement celles occasionnées par notre installation expérimentale pour l'observation de glaciers alpins, peuvent être résumées comme suit :

- De hautes définition et résolution au sol, pour la couverture de la zone observée, nécessitant des algorithmes optimisés en temps de calcul.
- L'absence de texture (glace, neige) sur certaines régions de la scène engendrant un manque d'unicité des éventuels points de contrôle.
- La dégradation des images par des phénomènes non contrôlables et non prévisibles tels que neige, humidité, taux d'ensoleillement, gouttes d'eau sur l'écran, nuages, ombres selon la saison et l'heure, etc.
- La distorsion des images due à la variation extrême de températures affectant les composants internes de la caméra.
- Un mouvement aléatoire de la caméra dû à la dilatation et à la contraction des composants externes à la caméra (boîtier et pied du boîtier) liés aux conditions météorologiques.
- Objet à observer parfois occulté et subissant des déformations discontinues et non rigides. Conditions qui violent les hypothèses méthodologiques (occultations) et géométriques (non-rigidité) de la plupart des algorithmes de vision par ordinateur.

Nous proposons donc une technique qui permet de traiter des séries temporelles d'images "time lapse" acquises en milieu naturel, en faisant face à ces contraintes. Cette chaîne permet de quantifier le mouvement du glacier de façon *automatique et adaptée* aux contraintes citées avec comme seule donnée d'entrée un "time lapse" entier et un masque binaire grossier de l'image maîtresse (zones mobiles / zones supposées fixes). Une automatisation optimale nécessitant une identification efficace des valeurs aberrantes à tous les stades de calcul, cette chaîne donne accès à tous les résultats intermédiaires ainsi qu'aux erreurs et scores accumulés pendant chaque phase de traitement. Cet article s'organise de la manière suivante. Tout d'abord, nous exposerons une à une, les différentes étapes de la

chaîne de traitement "time lapse" proposée. Nous présenterons ensuite un exemple de résultat issu du traitement sur un site particulier, le site du glacier d'Argentière, afin d'évaluer les performances en termes de précision et de cohérence par rapport à une vérité terrain que détiennent les glaciologues. Nous discuterons ensuite brièvement des étapes à suivre pour compléter la chaîne par un traitement "time lapse *stereo*" qui nous permettra d'avoir des mesures métriques comme données de sortie.

2 Mesure de déplacement par "time lapse"

2.1 Prétraitements

En milieu naturel, les conditions météorologiques (vent, température, irradiance...) influent grandement sur l'appareil photo, son support et sur les images acquises. Il en résulte deux principaux problèmes qui nécessitent des prétraitements avant d'utiliser le "time lapse" pour une estimation des changements sur la scène observée. Une première étape consiste à rejeter de façon automatique les images ne contenant pas d'information pertinente ou intégrant dans la série une source d'erreur trop importante et faussant ainsi toute la suite de la chaîne de traitement.

Sélection de l'information pertinente. Le choix des images à rejeter peut être fait grâce aux informations texturales contenues dans chaque image. Pour y parvenir, nous avons eu recours à la construction d'histogrammes de gradients, une version simplifiée des descripteurs HOG [2], comme suit :

1. Normalisation globale de l'image
2. Calcul du gradient de l'image en x et y
3. Construction de l'histogramme de gradients
4. Aplatissement de l'histogramme en vecteur puis en un score par image

La première phase consiste en la normalisation globale de l'image. Cette étape a été conçue afin de réduire l'influence des effets d'illumination ce qui nous permet de fixer un même seuil (score minimal accepté) à toutes les images d'une même série quel que soit le taux d'illumination à l'instant t . En pratique, nous utilisons la compression gamma (power law), en calculant la racine carrée ou le log de chaque canal de couleur. La force de texture de l'image est typiquement proportionnelle à l'illumination de surface locale, de sorte que cette compression contribue à réduire les effets des variations d'illumination locales (dues principalement aux ombres et au taux d'ensoleillement).

La deuxième phase consiste à calculer les gradients de premier ordre. La manière la plus simple pour obtenir une image gradient est d'appliquer un filtre dérivatif centré dans les directions horizontales et verticales ($[-1, 0, 1]$ et $[-1, 0, 1]^T$) sur chaque canal RGB et de retenir pour chaque pixel les gradients de plus grande norme. En pratique, un filtre de Canny a été utilisé. Celui-ci capture le

contour, la silhouette et une certaine information de texture, tout en fournissant une résistance supplémentaire aux variations d'illumination. Le canal de couleur localement dominant est utilisé, ce qui fournit une invariance à la couleur étant donné que nos images sont prises dans des saisons différentes. Pour plus de robustesse contre le changement d'échelle une représentation d'espace échelle peut être utilisée, par exemple par la construction d'une pyramide de différence de Gaussiennes.

La phase de construction de l'histogramme se fait comme suit : chaque pixel vote pour une classe de l'histogramme, en fonction de l'*orientation* du gradient à ce point. Le vote du pixel est pondéré par l'intensité du gradient en ce point. Le score est alors la somme des éléments du vecteur concaténé des composantes des histogrammes par blocs normalisés.

Notons que le fait d'avoir choisi un histogramme encodant les orientations locales et non seulement la texture vient du fait que nous disposons de séries stéréos que nous utiliserons ultérieurement et de ce fait nous tenons à ce que la phase d'élimination d'images non pertinentes soit commune aux deux séries droite et gauche, et donc que le seuil à fixer soit invariant au changement de perspective.

L'élimination des images inutiles a été faite par un test de détection de valeurs aberrantes sur ces scores. Le test de Chauvenet stipule que si $n = N \cdot \text{erfc}(\frac{|X_i - \bar{X}|}{\sigma_X}) < 0.5$ alors X_i est une donnée aberrante, où N est le nombre de mesures, 0.5 est la probabilité d'une mesure d'être en dehors de $2 \sigma_X$ autour de \bar{X} , erfc est la fonction d'erreur complémentaire dénotant ici la probabilité d'avoir d'autres mesures aussi déviées que X_i et par conséquent n dénote le nombre de mesures qui devraient avoir le même écart que X_i .

Recalage des zones fixes sur toute la série. Afin de pouvoir combiner / comparer les informations respectives à toutes les images de la série, il faut que celles-ci soient d'abord recalées, *ie* que la position d'une région sur l'image (montagnes, glaciers, etc) soit invariante dans les temps.

Cette étape vise donc à compenser l'effet du mouvement aléatoire de la caméra. Une observation qualitative des déplacements entre les images de la série sans recalage préalable nous a permis de voir que le mouvement n'est pas une simple translation, mais que la caméra effectue en plus un mouvement de rotation qui introduit des changements de perspective dans les images. La transformation géométrique la plus à même de modéliser ce changement de perspective est l'homographie, qui modélise la projection perspective d'un plan dans un repère 3D, et prend donc en compte le fait que la caméra peut aussi "se pencher" contrairement à une transformation affine qui aurait modélisé un changement de perspective, mais dans un même plan 2D.

Cette étape de prétraitement vise donc à retrouver une transformation géométrique *globale* qui fait en sorte que seuls les déplacements dus à la dynamique du glacier sub-

sistent. Cela a été fait comme suit :

1. Masquage binaire des zones mobiles par un même masque donné en donnée d'entrée.
2. Conversion en niveaux de gris.
3. Calcul itératif des matrices de déformation perspective entre, d'une part, toutes les images de la série et d'autre part, la première image valide avec [10].
4. Application de la transformation géométrique estimée et interpolation.

L'algorithme utilisé dans l'étape 4 citée ci-dessus permet l'estimation des 8 paramètres de transformation homographique selon une version légèrement modifiée de l'algorithme [12]. Ceux-ci proposent, après une initialisation des paramètres de transformation soit par l'unité soit par une estimation grossière préalable, le raffinement itératif des paramètres de la projection perspective par une technique d'optimisation en descente de gradient.

La technique adoptée est celle par composition et inversion. Pour recaler une image I' sur une image I cette technique vise à retrouver, à chaque itération, la transformation géométrique $W(x, y; p)$, où $p^T = (p_1, p_2, \dots, p_8)^T$ est un ensemble de paramètres, qui feraient que $I(W(x, y; p)) \approx I'(x, y)$, et où (x, y) représentent les coordonnées des pixels dans une image. Par exemple, pour une simple translation :

$$W(x, y; p) = \begin{pmatrix} x + p_1 \\ y + p_2 \end{pmatrix}. \quad (1)$$

Dans notre cas,

$$W(x, y; p) = \frac{1}{1 + p_7x + p_8y} \begin{pmatrix} (1 + p_1)x + p_3y + p_5 \\ p_2x + (1 + p_4)y + p_6 \end{pmatrix}, \quad (2)$$

ce qui peut être fait par simple multiplication matricielle en utilisant les coordonnées homogènes. En d'autres termes, à chaque itération, l'algorithme calcule un incrément des paramètres à retrouver² :

$$\Delta p = -H^{-1} \sum_{x,y} [\Delta I \frac{\partial W}{\partial p}]^T [I(W(p)) - I'], \quad (3)$$

où ΔI est le gradient de l'image I évalué à $W(p)$, $I(W(p))$ est calculée par composition, inversion et interpolation, $\frac{\partial W}{\partial p}$ est la jacobienne de la matrice de transformation et H est la matrice Hessienne de la fonction objectif à minimiser :

$$H = \frac{1}{2} \sum_{x,y} [I(W(p)) + \Delta I \frac{\partial W}{\partial p} \Delta p - I'], \quad (4)$$

retrouvée par un développement de Taylor d'ordre 1 de $I(W(p + \Delta p))$.

L'algorithme [10] utilise une version un peu modifiée du coefficient de corrélation original nommé "Enhanced Correlation Coefficient (ECC)".

2. Pour plus de simplicité visuelle, on notera dans la suite $W(p)$ au lieu de $W(x, y; p)$ et I au lieu de $I(x, y)$.

Nous avons fixé les conditions d'arrêt de la boucle d'optimisation comme suit : le nombre d'itérations maximal est fixé à 200 et le seuil de l'incrément minimal du coefficient de corrélation entre deux itérations à $1e-10$, ce qui constitue des conditions assez stricts, mais puisque nous n'avons pas de soucis d'exécution en temps réel cela ne pose pas de problème.

2.2 Traitements

Une fois les images pertinentes sélectionnées, indexées (afin d'identifier les dates manquantes) et recalées, le traitement proprement dit consiste à mesurer la dynamique du glacier, isolée idéalement de tout artefact. Pour ce faire nous avons tout d'abord effectué une mise en correspondance dense [15] qui nous a permis d'initialiser un calcul de flot optique [16].

Mise en correspondance. L'algorithme utilisé pour la mise en correspondance s'appelle DeepMatching et est inspiré des "Deep Convolutional Network" dans le sens où il utilise une architecture similaire au quad-tree pour calculer la corrélation de patches de plus en plus grands. Celui-ci comporte deux principales étapes :

Construction d'une pyramide de corrélations à plusieurs échelles La base de la pyramide étant l'image de plus haute résolution (celle d'origine : I), à chaque voisinage R de 4×4 pixels, dits patches atomiques, est construite une carte de corrélation avec l'image I' :

$$\text{sim}(R, R') = \frac{1}{16} \sum_{i=0}^3 \sum_{j=0}^3 R_{i,j}^t R'_{i,j}. \quad (5)$$

La description R utilisée dans [15] est inspirée du descripteur de points de contrôle SIFT [14] invariant au changement d'échelle grâce à la représentation en espace échelle et à la rotation grâce à la construction d'un histogramme de gradient orienté, comme proposé par [2], mais avec quelques extensions, notamment le plafonnement de la carte des gradients avec une fonction sigmoïde.

La propagation aux niveaux supérieurs de la pyramide de corrélation est ensuite faite par agrégation itérative. Cette agrégation consiste en un max-pooling, sous-échantillonnage, le calcul d'une moyenne décalée suivit d'une rectification non linéaire (par application d'une puissance à la carte de corrélation obtenue afin de s'assurer de la propagation des valeurs importantes aux niveaux supérieurs). La figure 1 illustre la propagation d'un niveau où un patch atomique a une taille de 4×4 pixels (plus haute résolution) à un niveau où le patch (4×4 pixels) couvre virtuellement une zone sur l'image d'origine de 8×8 pixels.

Extraction des correspondances À partir du sommet de la pyramide, et étant donné que chaque patch a été créé par agrégation de 4 patches du niveau inférieur, il suffit de défaire récursivement les opérations qui ont permis sa construction : Figure 2. Autrement dit, il s'agit d'estimer,

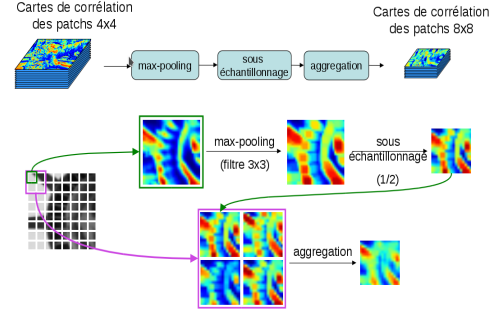


FIGURE 1 – Propagation de la pyramide de corrélation aux niveaux supérieurs [15].

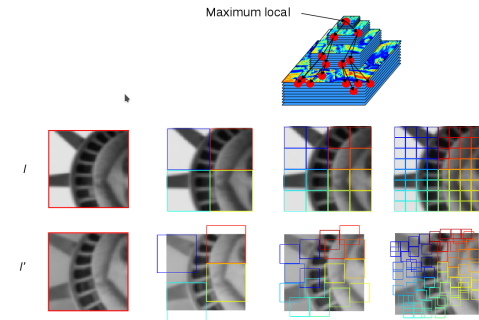


FIGURE 2 – "Backtracking" des correspondances [15].

pour chaque paire de patches homologues au sommet de la pyramide, le mouvement des patches atomiques (à la base de la pyramide) par optimisation de la position des descripteurs de voisinages de plus en plus petits³.

Nous obtenons ainsi une mise en correspondance quasi-dense, invariante à l'échelle et à la rotation.

A noter que, lors de l'implémentation de l'algorithme original, la fenêtre de recherche pour l'appariement de patches par corrélation croisée a été limitée à 150 pixels puisque nous supposons que nos déplacements ne peuvent pas dépasser cette portée et qu'il est donc intéressant de limiter la complexité de la méthode et le temps d'exécution de l'algorithme.

Flot optique. Il s'en suit le calcul du flot optique initialisé par ce matching dense. Par définition, le flot optique est le mouvement réel des objets projeté sur l'image.

Le calcul de ce flot part de l'hypothèse qu'un pixel (x, y) dans une image prise à un instant t conserve la même intensité $I(x, y)$ quand il se déplace de $(\Delta x, \Delta y)$ sur l'image I' prise à l'instant $t + \Delta t$:

$$I(x, y; t) = I(x + \Delta x, y + \Delta y, t + \Delta t). \quad (6)$$

En supposant que le mouvement n'est pas très grand, un

3. Une explication approfondie et des détails algorithmiques sont donnés dans l'article [15].

développement en série de Taylor nous permet d'obtenir :

$$I(x + \Delta x, y + \Delta y, t + \Delta t) = I(x, y, t) + \frac{\partial I}{\partial x} \Delta x + \frac{\partial I}{\partial y} \Delta y + \frac{\partial I}{\partial t} \Delta t + \epsilon. \quad (7)$$

Il s'ensuit que :

$$\frac{\partial I}{\partial x} \frac{\Delta x}{\Delta t} + \frac{\partial I}{\partial y} \frac{\Delta y}{\Delta t} + \frac{\partial I}{\partial t} \frac{\Delta t}{\Delta t} = 0, \quad (8)$$

où $\frac{\Delta x}{\Delta t} = V_x$ et $\frac{\Delta y}{\Delta t} = V_y$ sont les vitesses du mouvement à retrouver ou flot optique et où $\frac{\partial I}{\partial x}$, $\frac{\partial I}{\partial y}$ et $\frac{\partial I}{\partial t}$ sont les dérivées de l'image $I(x, y, t)$ respectivement dans les directions x , y et t notés I_x , I_y et I_t .

$$I_x V_x + I_y V_y = -I_t \Leftrightarrow \nabla I^T \cdot \vec{V} = -I_t \quad (9)$$

L'équation 9, appelé contrainte du mouvement apparent, est l'équation à résoudre de tout algorithme de calcul de flot optique.

Cette hypothèse est cependant forte, car elle suppose que les variations temporelles de luminance sont uniquement dues aux mouvements réels de l'objet observé, de plus, cette équation est insuffisante pour l'obtention d'une solution unique. Les méthodes récentes de calcul de flot optique considèrent donc un terme de régularisation afin de contraindre l'ensemble des solutions.

Nous avons considéré les formulations variationnelles du flot optique où cette contrainte est traduite en un problème d'optimisation minimisant une énergie basée sur l'équation 9 et où le terme de régularisation change selon le contexte et la méthode.

Ce terme de régularisation, qui sert à particulariser les solutions, peut être un terme de lissage, c'est-à-dire un terme prenant en compte le gradient ou le Laplacien (ou ordre supérieur) du champ de vitesse. L'une des premières formulations [11] est donc une minimisation d'une énergie :

$$E = \iint [(I_x V_x + I_y V_y + I_t)^2 + \alpha^2 (\|\nabla V_x\|^2 + \|\nabla V_y\|^2)] dx dy \quad (10)$$

où α est une constante de régularisation.

L'algorithme implémenté [16] est un algorithme d'optimisation à poids variés inspiré de [17] et il minimise une énergie qui prend en plus en compte un terme lié à une phase de mise en correspondance préalablement effectuée. Cette énergie est de la forme :

$$E = \iint E_{lum} + \alpha E_{liss} + \beta E_{matches} dx dy \quad (11)$$

2.3 Analyse des déplacements

Tracking. Une première analyse intuitive de nos champs de déplacement est le "tracking" ou suivi d'un pixel (ou un voisinage de pixels) donné (à sélectionner par l'utilisateur) afin de quantifier son évolution dans le temps. Après sélection du pixel en question, le suivi se fait sur les déplacements selon les axes x et y , sur l'amplitude et sur l'angle

de déplacement. Cela permet de juger de la cohérence des champs calculés, de détecter un éventuel événement survenant dans la région suivie et de collecter les observations qui permettront ultérieurement de poser un modèle sur la dynamique des différentes régions du glacier. Ce modèle peut être additif (composante saisonnière indépendante de la tendance), multiplicatif (la composante saisonnière proportionnelle à la tendance) ou mixte et son analyse peut être paramétrique (tendance à modèle linéaire, polynomial, exponentiel, etc.) ou non-paramétrique (moyenne mobile ou différentiation).

Moyennage. Des cartes des déplacements moyens sur toute la série ou entre deux dates choisies par l'utilisateur sont aussi construites afin de pouvoir évaluer qualitativement les champs calculés. Les estimations statistiques des déplacements selon les axes x et y et celui du module se fait selon les formules usuelles. Cependant, afin d'effectuer un moyennage temporel (pour avoir une estimation de la direction principale sur toute la série) et spatial (afin de réduire la taille de notre carte) des mesures d'angles du flot dense calculé, nous avons été amenés à utiliser les formules statistiques des variables aléatoires circulaires. On définit alors les estimateurs circulaires –moyenne empirique, biais sur l'estimation, variance et écart-type – de $\phi \in [-\pi, \pi]$ respectivement de la manière suivante :

$$\hat{m}_c(\phi) = \arg\left(\frac{1}{L} \sum_{l=1}^L \exp^{i\phi^{(l)}}\right), \quad (12)$$

$$\hat{b}_c(\phi) = \arg(\exp^{i(\hat{m}_c(\phi) - \phi_0)}), \quad (13)$$

$$\hat{\sigma}_c(\phi) = \sqrt{2\left(1 - \frac{1}{L} \left| \sum_{l=1}^L \exp^{i\phi^{(l)}} \right| \right)}, \quad (14)$$

$$\hat{var}_c(\phi) = \sigma_c(\hat{\phi})^2, \quad (15)$$

où $\arg(a)$ représente l'argument du nombre complexe a et $|a|$ son module.

Fermeture temporelle. Une évaluation de la cohérence des champs estimés peut aussi être faite par mesure de la fermeture temporelle sur une suite de dates. Cette contrainte peut s'exprimer de deux manières :

$$\vec{V}_{i,i+1} + \vec{V}_{i+1,i+2} + \vec{V}_{i+2,i+3} - \vec{V}_{i,i+3} = \vec{0} \quad (16)$$

$$\vec{V}_{i,j} + \vec{V}_{0,i} - \vec{V}_{0,j} = \vec{0}, \quad (17)$$

où $\vec{V}_{i,j}$ dénote les champs de vitesses \vec{V} entre deux images prises aux dates indexées i et j .

La première formulation permet de varier la portée de la fermeture temporelle. En l'occurrence, cette portée sur l'équation 16 est de 4, c'est à dire que l'évaluation est faite sur 4 dates adjacentes. Pour la deuxième formulation la portée est toujours maximale. Ces contraintes sont théoriquement satisfaites quelque soit la région de l'image

(régions supposées fixes (montagne) ou mobiles (glacier)) sauf sur les régions à dynamique chaotique. Nous avons donc construit une suite de cartes de fermeture temporelles selon la date centrale et le nombre de dates adjacentes voulues. Ces cartes se sont avérées très efficaces quant à la détection et au suivi d'événements survenant sur la scène observée (chute de Séracs par exemple).

Scores et erreurs. Pour une automatisation optimale de la chaîne de traitement, et afin d'identifier un éventuel niveau dans lequel a commencé une accumulation d'erreurs de calcul, plusieurs erreurs / scores ont été calculés :

- Pour chaque image, le score HOG avec lequel l'image a été retenue en tant que donnée d'entrée valide.
- Pour chaque image, la distance euclidienne, après recalage, avec la première image valide (seulement sur les zones fixes).
- Par paire d'images appariées, le score de corrélation (somme normalisée des similarités le long du chemin d'extraction des correspondances).
- Par paire d'images, calcul d'une homographie résiduelle H en utilisant les correspondances calculées pendant le traitement. Cette transformation résiduelle est estimée avec l'algorithme DLT (Direct Linear Transform) [19] qui retrouve les paramètres d'une homographie par résolution du problème sur-déterminé :

$$x_i = H.x'_i \simeq x'_i * H.x_i = 0 \quad (18)$$

où x_i et x'_i sont les coordonnées (homogènes) des points homologues. Il faut au moins 3 correspondances pour résoudre l'équation 18.

- Somme par paire d'images et moyenne sur toute la série des modules de déplacements en zones supposées fixes.
- Cartes de fermetures temporelles.

3 Expérimentation en milieu naturel

3.1 Données et site observé

Les résultats présentés dans cet article sont issus d'une série temporelle d'images du glacier d'Argentière, ce glacier se situe dans le massif du Mont Blanc (France). L'acquisition a été effectuée entre le 13/09/2013 et le 17/10/2013 à raison d'une image par jour (choix fait compte tenu de la dynamique assez lente du glacier en question). Pour cette acquisition, l'appareil photo utilisé est un Panasonic DMC-LX3 empaqueté pour résister aux conditions climatiques extrêmes du milieu alpin, alimenté par un panneau solaire et programmé pour acquérir de manière automatique les "time lapse". Placé à 2631m d'altitude et à environ 500m du glacier étudié, cet appareil photo permet d'observer une surface glaciaire d'environ $1km^2$. Les images du "time lapse" ont une taille de 3776×2520 (10Mp), qui permet, dans ces conditions d'installation et selon les caractéristiques du capteur, d'obtenir une taille de pixel au sol d'en-

viron 15cm en rive droite du glacier (rive la plus proche et dont les résultats sont présentés) et d'environ 30cm en rive gauche, mais cela n'est pas uniforme sur toute l'image vu l'obliquité de la scène par rapport au plan image.

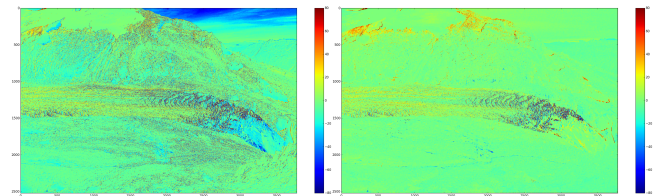
Cette série a été choisie car elle présente la plupart des problèmes possibles (images inutilisables, début de la série sans neige et fin de la série très enneigée, conditions météorologiques défavorables, images qui n'ont pas été correctement acquises, etc.). Elle illustre ainsi un panel non exhaustif des difficultés du traitement d'images acquises en milieu naturel. La figure 3 montre quelques exemples d'images de cette série.



FIGURE 3 – Exemple d'images traitées dans la série temporelle de la rive droite d'Argentière.

3.2 Résultats et analyse

Prétraitements. Dans un premier temps, afin d'illustrer l'impact des prétraitements sur la série d'images, la figure 4 présente la distance entre les deux premières images valides I_i et I_j de la série. Cette distance est calculée canal par canal en s'appuyant sur la simple différence : $I_i - I_j$ (valeurs négatives non saturées).



(a) Sans prétraitement.

(b) Avec prétraitement.

FIGURE 4 – Effet du prétraitement sur les déplacements calculés. Image de distance entre deux images de la même série. Le domaine de visualisation a été limité à $[-80, +80]$ et non $[-255, +255]$ et seulement sur le canal Vert de l'image RVB pour une meilleure visualisation de l'impact du recalage.

La figure 4a met clairement en évidence une quantité de changement plus importante que la figure 4b, notamment sur les zones de montagne qui sont censées être fixes avec peu de changement comparé à la zone glaciaire. La distance euclidienne $\|I_i - I_j\|$ est passée de 120 711.05 à 71 514.28 et de 82 483.21 à 40 515.67 en masquant le glacier.

Mesures des déplacements. L'impact de ces prétraitements a, comme prévu, aussi un effet direct sur la mesure de déplacements. L'image 5 met en évidence ce dernier sur une zone de mesure particulièrement complexe. Ce zoom correspond à une zone du glacier entièrement recouverte de roches, il n'est donc pas possible, sans connaissance

préalable, de déterminer où commence et où s'arrête le glacier. La figure 5a, pour laquelle la mesure de déplacement a

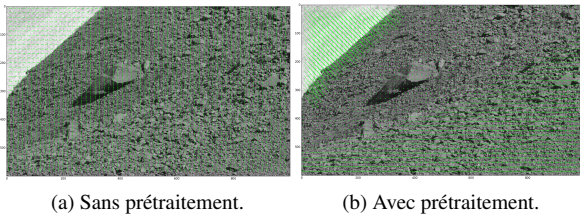
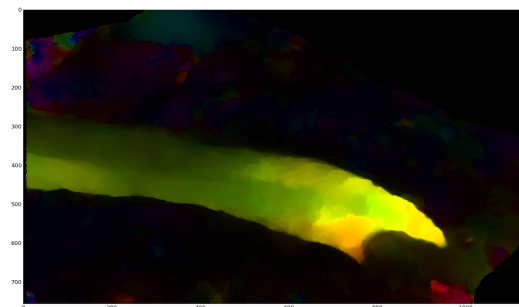


FIGURE 5 – Effet du prétraitement sur les déplacements calculés. Flot optique entre deux dates adjacentes (visualisation non dense par souci de visibilité).

été effectuée sans recalage, ne permet pas de déterminer la zone glacière et son déplacement. Ce dernier est noyé dans le déplacement induit par le mouvement de la caméra. A contrario, la figure 5b qui est le résultat incluant les prétraitements, fait ressortir une zone mobile sur la partie basse et droite de l'image qui correspond au déplacement du glacier. Ces résultats ont été validés, qualitativement, par des géo-morphologues du laboratoire EDYTEM (France). Notons que les grands déplacements qui apparaissent sur la partie haut gauche de la figure 5b correspondent à des erreurs de mesure induites par la partie enneigée en arrière plan.

En nous focalisant à présent sur la totalité de la série, la figure 6 présente le mouvement du glacier.



(a) Carte des directions principales sur toute la série pondérées par le module du flot optique.



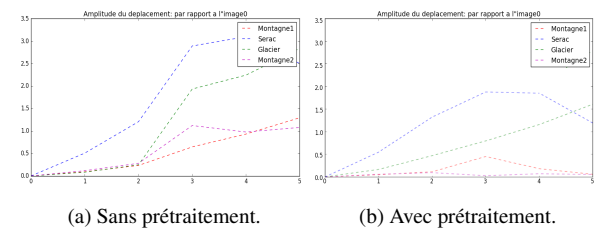
(b) Légende

FIGURE 6 – Directions principales moyennes : moyennage spatial et temporel sur toute la série. En bas, légende du codage couleur HSV : L'angle est codé sur le Hue et l'amplitude sur le Value.

Cette figure est obtenue par un moyennage spatial et temporel des mesures de déplacement. L'angle du flot optique a été considéré comme une variable aléatoire circulaire. Au niveau spatial, la moyenne est calculée sur un voisinage de

5x5 pixels, ceci permet de réduire la taille de la carte obtenue, mais un voisinage de taille différente peut être choisi par l'utilisateur de la chaîne de traitement. Au niveau temporel, la moyenne, ou direction principale, est calculée sur toute la portée temporelle de la période d'observation. La figure 6 illustre à la fois l'orientation et l'amplitude du déplacement grâce à une composition colorée. Les informations de déplacement – amplitude et orientation – sont codées dans l'espace colorimétrique Hue Saturation Value (HSV) (cf. figure 6b). L'angle de l'orientation, étant cyclique, est codé par la teinte H, selon l'angle correspondant du cercle colorimétrique, l'angle 0 correspondant à un déplacement vers le bas de l'image. L'amplitude en pixel – comprise entre 0 et Dmax – est projetée sur l'échelle de valeur V comprise entre 0 et 100%. Quant à la saturation, elle est fixée à 100%.

La teinte verte du glacier représente un déplacement vers la droite de l'image, avec une courbure (jaune) qui correspond à un virage d'écoulement du glacier au niveau de sa chute de sérac (chute de sérac de Lognan). On observe également sur cette zone une teinte plus soutenue, ceci correspond à une zone d'accélération du glacier suite au passage d'un éperon rocheux et à un changement de pente d'écoulement. Nous notons également que les zones hors du glacier observé ont un mouvement moyen quasi nul à l'exception de deux zones : Sous la chute de séracs, une tache rouge illustre un amas de glace alimenté par les séracs tombés depuis la chute. Elle ne correspond pas à un mouvement, mais à un changement chaotique de la surface en ce point. La seconde zone particulière se situe dans la partie haut gauche de la carte, elle apparaît de manière bleutée et homogène. Ceci correspond principalement à un glacier suspendu entièrement recouvert de neige.

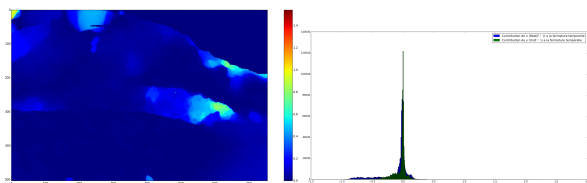


(a) Sans prétraitement. (b) Avec prétraitement.

FIGURE 7 – Évolution des modules des déplacements sur 5 dates. Montagne 1 : Région montagneuse (supposée fixe) en dessus du glacier. Montagne 2 : Région montagneuse (supposée fixe) en dessous du glacier. Glacier : Région lente du glacier (au milieu). Sérac : Région à dynamique chaotique du glacier (au bout).

La figure 7 montre que l'accumulation de déplacements avec la première date valide en zone glacière s'effectue d'une manière linéaire sur la période observée. En outre, les déplacements en zones fixes ont bien été compensés par les prétraitements.

La figure 8a montre que les régions obéissant le moins à une fermeture temporelle sur les 5 dates adjacentes à la



(a) Carte des fermetures temporelles. (b) Contribution des déplacements dx et dy sur la fermeture temporelle.

FIGURE 8 – Fermetures temporelles à la date indexée 7 sur 5 dates adjacentes : $\vec{V}_{56} + \vec{V}_{67} + \vec{V}_{79} + \vec{V}_{9,10} - \vec{V}_{5,10}$ (date 8 non traitée car rejetée en phase de prétraitements).

date indexée 7, sont les zones d'accélération du glacier, ainsi que celle du glacier suspendu en haut à droite (détecté grâce à la carte des déplacements et orientations moyens et confirmé par les géomorphologues) et certaines zones dont l'extrême enneigement (manque de texture) a faussé les estimations. Les chutes de séracs peuvent en effet être considérées comme la survenance d'un événement. Le taux d'erreur reste néanmoins faible sur toute l'image (maximum 1). A noter que le ciel (en haut à droite) a été binairement masqué. La figure 8b est un histogramme des composantes horizontale (x) et verticale (y) de la carte de fermetures temporelles montrant la contribution de chacun des deux axes. Les valeurs sont centrées autour de zéro et l'erreur la plus importante est induite par un déplacement horizontal.

La moyenne des déplacements résiduels en zones fixes sur toute la série est de $12,50 \cdot 10^{-4}$ en x et $10 \cdot 10^{-4}$ en y .

4 Conclusion et perspectives

Nous avons présenté une chaîne de traitement adaptée à l'observation de milieux naturels. Celle-ci prend en entrée un "time lapse", sélectionne l'information pertinente, effectue les traitements requis pour le calcul des déplacements image (de l'ordre du centième de pixel) et du flot optique de la scène observée pour enfin analyser les résultats obtenus en retrouvant, par exemple, les événements clés se produisant sur le site et en modélisant la dynamique des objets géophysiques observés.

Les traitements présentés constituent la première partie d'une chaîne stéréo plus complète permettant de quantifier le mouvement réel (3D métrique) de glaciers avec comme élément en bout de chaîne une expertise de la part de glaciologues permettant une modélisation précise de la dynamique des glaciers et une évaluation des risques liés au site étudié.

Références

[1] A. Dehecq, N. Gourmelen, E. Trouvé, Deriving large-scale glacier velocities from a complete satellite archive : Application to the Pamir-Karakoram-Himalaya. *Remote Sensing of Environment*, Elsevier, 162, 55-66, 2016.

[2] N. Dalal, B. Triggs, Histograms of Oriented Gradients for Human Detection, CVPR '05. *Proceedings of the 2005 IEEE Computer So-*

ciety Conference on Computer Vision and Pattern Recognition (CVPR'05), 2005.

- [3] H.T. Pham, H. He, F. Vernier, E. Trouvé, L. Benoit, L. Moreau, B. Girard, Analyse de "Time-Lapse" stéréo pour la mesure de déformation 3D, application au suivi du glacier d'Argentière *Reconnaissance de Formes et Intelligence Artificielle*, RFIA, 2014.
- [4] Motyka, R., M. Fahnestock, I.M. Howat, M. Truffer, H. Brecher and M. Luethi, Jakobshavn Isbrae, Greenland : DEMs, ortho- photos, surface velocities, and ice loss derived from photogrammetric reanalysis of July 1985 repeat aerial photography. *American Geophysical Union, Fall Meeting 2008*, abstract n : C13A-0563, 2008.
- [5] F. Vernier, R. Fallourd, Y. Yan, D. Rosu, E. Trouvé, J.M. Nicolas, J.M. Friedt, L. Moreau, Combining space-borne SAR data and digital camera images to monitor glacier flow by remote and proximal sensing, *IEEE International Geoscience and Remote Sensing Symposium*, jul., Vancouver, Canada, 2011.
- [6] I. Joughin, W. Abdalati, M.A. Fahnestock, Large fluctuations in speed on Greenland's Jakobshavn Isbrae glacier, *Nature*, 432(7017), 608-610, 2004.
- [7] B. Csatho, T. Schenk, C.J. van der Veen, W.B. Krabill, Intermittent thinning of Jakobshavn Isbrae, West Greenland, since the Little Ice Age, *J. Glaciol.*, 53(184), 131-144, 2008.
- [8] W.D. Harrison, K.A. Echelmeyer, D.M. Cosgrove, C.F. Raymond, The determination of glacier speed by time-lapse photography under unfavourable conditions, *J. Glaciol.*, 38(129), 257-265, 1992.
- [9] A.N. Evans, Glacier surface motion computation from digital image sequences, *IEEE Trans. Geosci. Remote Sens.*, 38(2), 1064-1072, 2000.
- [10] G. Evangelidis, E. Psarakis, Parametric Image Alignment Using Enhanced Correlation Coefficient Maximization, *IEEE Transactions on Pattern Analysis and Machine Intelligence, Institute of Electrical and Electronics Engineers*, 30(10), 1858-1865, 2008.
- [11] B.K.P. Horn, B.G. Schunck, Determining optical flow, *Artificial Intelligence*, 17, 185-203, 1981.
- [12] S. Baker, I. Matthews, *International Journal of Computer Vision* 56 :221, 2004.
- [13] B.K.P. Horn, B. G. Schunck, Determining OpticalFlow, *Artificial Intelligence*, 16(1-3), 185-203, 1981.
- [14] D.G. Lowe, Object recognition from local scale-invariant features, *Proceedings of the International Conference on Computer Vision*, 2, 1150-1157, 1999.
- [15] J. Revaud, P. Weinzaepfel, Z. Harchaoui, C. Schmid, DeepMatching : Hierarchical Deformable Dense Matching, *International Journal of Computer Vision (IJCV)*, 2016.
- [16] P. Weinzaepfel, J. Revaud, Z. Harchaoui, C. Schmid. DeepFlow : Large displacement optical ow with deep matching, *IEEE International Conference on Computer Vision (ICCV)*, 2013.
- [17] T. Brox, J. Malik, Large displacement optical flow : descriptor matching in variational motion estimation. *IEEE Trans. PAMI*, 2011.
- [18] H. Bay, T. Tuytelaars, L. Van Gool, SURF : Speeded Up Robust Features, *Computer Vision - ECCV 2006 : 9th European Conference on Computer Vision*, Graz, Austria, May 7-13, Proceedings, Part I, 2006.
- [19] I.E. Sutherland, Three-dimensional data input by tablet, *Proceedings of the IEEE*, 62 (4), 453-461, 1974.
- [20] M. Gallardo, T. Collins, A. Bartoli Using Shading and a 3D Template to Reconstruct Complex Surface Deformations *British Machine Vision Conference (BMVC)*, 2016.
- [21] Z. Liu, H. Ukida, P. Ramuhalli, K. Niel Integrated Imaging and Vision Techniques for Industrial Inspection : Advances and Applications. *Springer*, 2015. Print.