

Une méthode globale de recalage inter-images appliquée aux séquences aériennes

Marie-Lise DUPLAQUET - Hubert CANTALLOUBE

ONERA

29 Avenue de la Division Leclerc 92320 Châtillon - FRANCE

RÉSUMÉ

Savoir mesurer la déformation entre deux images est important pour de nombreuses applications. Après le calcul des formules de déformation inter-images, nous présentons la méthode globale de recalage retenue. Les déformations locales (translations) sont mesurées par corrélation centrée normée, puis l'optimisation d'un coût global utilisant toute la surface de corrélation se fait par descente de simplexe. Cette méthode ne demande aucun pré-traitement et utilise trois paramètres robustes. Son utilisation est illustrée sur un couple d'images aériennes.

ABSTRACT

Measurement of deformation between two images is useful in many applications. After calculating equations of inter-images deformation, the global matching method is exposed. Local deformations (translations) are computed with normalized central correlation, and a global cost using the whole correlation surface is optimized by the downhill simplex method. This matching method requires no pre-treatment and uses only three robust parameters. It is then illustrated on a pair of aerial images.

I - Introduction

La navigation autonome d'un engin à l'aide d'une séquence d'images aériennes est généralement réalisée par détection et reconnaissance d'amers. Si les conditions de prise de vue (position de l'engin, axe de la caméra, vitesses) ne sont pas disponibles, savoir mesurer la déformation d'une image par rapport à la suivante permet de suivre des amers le long de la séquence, même s'ils sont momentanément cachés (par des fumées par exemple), ou difficiles à repérer. Le nombre de fausses détections est réduit, et la précision plus grande.

Une autre application est l'addition d'images recalées dans les séquences bruitées : en ajoutant de 3 à 5 images recalées, le niveau du signal augmente par rapport au bruit dû à la caméra. L'image résultat est souvent un peu floue mais plus facile à interpréter.

La troisième application est le repérage des petits objets en mouvement par soustraction d'images. Cela suppose que le recalage est très bon (moins d'un pixel d'erreur) et que les objets se déplacent de plus d'un pixel.

Nous présentons ici une méthode de recalage globale qui ne nécessite aucun pré-traitement et utilise seulement trois paramètres robustes. Nous la décrirons après avoir posé le problème et calculé les formules de déformation. Les résultats

des tests effectués ainsi que les applications possibles seront ensuite exposés.

II - Position du problème

Pour établir un modèle de déformation entre deux images successives de la séquence, nous avons considéré une caméra embarquée à bord d'un engin.

A l'instant t , l'engin est à la position (L_E, G_E, H_E) et l'axe de la caméra est repéré par les angles (θ, φ, ψ) . Le passage du repère (L, G, H) lié au sol au repère (X, Y, Z) lié à l'engin se calcule par la transformation suivante :

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = M(\theta, \varphi, \psi) \begin{bmatrix} L-L_E \\ G-G_E \\ H-H_E \end{bmatrix} \quad (1)$$

où M est la matrice combinant les rotations.

Le passage aux coordonnées images (x, y) est donné par les formules :

$$x = \frac{Xf}{Z} \quad \text{et} \quad y = \frac{Yf}{Z} \quad (2)$$

avec f distance focale de l'objectif.

En dérivant les équations 1 et 2, sachant que l'engin se déplace avec la vitesse v en translation et ω en rotation (voir le détail des



calculs dans [4]), on obtient les formules ci-dessous (calculées pour la première fois par Longuet-Higgins) qui donnent le déplacement 2D (dx,dy) en tout point (x,y) :

$$dx = \frac{-f}{Z} v_x + \frac{x}{Z} v_z + \frac{xy}{f} \omega_x - (f + \frac{x^2}{f}) \omega_y + y \omega_z \quad (3)$$

$$dy = \frac{-f}{Z} v_y + \frac{y}{Z} v_z - \frac{xy}{f} \omega_y + (f + \frac{y^2}{f}) \omega_x - x \omega_z$$

Les vitesses v et ω sont à estimer, mais les capteurs de bord en donnent une approximation.

Avec les équations (1) et (2), on élimine L et G et on obtient :

$$\frac{1}{Z} = Ax + By + C \quad (4)$$

A, B et C dépendent des angles de prises de vue et de l'altitude H-H_E. Pour un sol plan, ils sont constants sur toute l'image.

On reporte l'expression (4) dans les formules (3), et on élimine le degré de liberté entre (A,B,C) et (v_x,v_y,v_z) en posant C=1 (normalisation de la distance caméra-centre image). On obtient :

$$dx = (Ax + By + 1) (-fv_x + x v_z) + \frac{xy}{f} \omega_x - (f + \frac{x^2}{f}) \omega_y + y \omega_z \quad (5)$$

$$dy = (Ax + By + 1) (-fv_y + y v_z) - \frac{xy}{f} \omega_y + (f + \frac{y^2}{f}) \omega_x - x \omega_z$$

Ces formules sont non-linéaires en fonction des 8 paramètres à estimer (avec des termes en Av_x, Bv_x, ...). Après leur estimation, il est théoriquement possible de retrouver les angles de prise de vue :

$$\varphi = \text{Arctg}(\frac{A}{B}) \quad \theta = \text{Arctg}(\frac{1}{f \sqrt{A^2 + B^2}}) \quad (6)$$

Mais si le champ de la caméra est étroit, et pour des visées obliques, A et B sont très faibles, et leur estimation est trop imprécise pour permettre le calcul de φ et θ. Les formules (5) sont cependant utilisées pour calculer la déformation initiale à partir de conditions de prise de vue connues approximativement.

Il est possible de linéariser les équations (5) par un changement de variables adéquat pour obtenir un modèle de déformation plus simple, que nous utiliserons par la suite :

$$\begin{aligned} dx &= ax^2 + bxy + cx + dy + e \\ dy &= axy + by^2 + fx + gy + h \end{aligned} \quad (7)$$

Cette formulation ne nécessite plus la connaissance de la focale, et reste valable après un changement d'échelle sur un axe. On notera que les paramètres du deuxième ordre a et b sont les mêmes pour dx et dy. Les formules proposées dans [3] pourraient aussi être utilisées.

Ces formules restent valides tant que l'écart entre les deux images est suffisamment faible pour que les vitesses soient

supposées constantes et que le relief de la scène est négligeable devant la hauteur de prise de vue.

III - Méthode globale de recalage

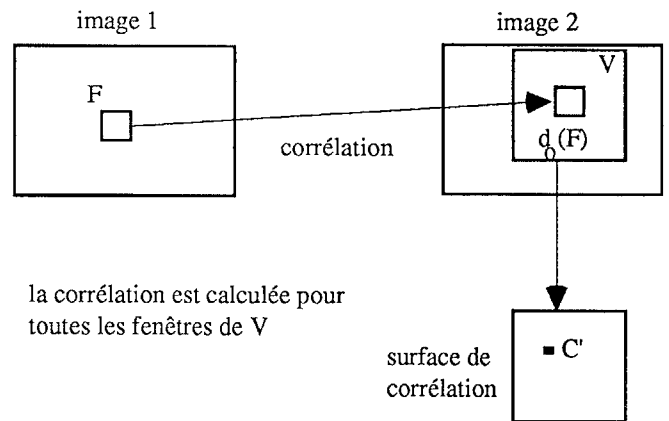
La déformation complexe entre les deux images (formules (5) ou (7)) est approximée localement par des translations, mesurées en un certain nombre de points de l'image. La mesure se fait par comparaison de petites fenêtres, supposées non déformées. Compte-tenu des sauts de dynamique entre images successives dans certaines de nos séquences, nous utilisons la corrélation centrée normée, mais des mesures plus rapides à calculer (par exemple la Somme des Différences en Valeurs Absolues) pourraient être envisagées (une comparaison de ces mesures est faite dans [1]).

Pour chaque fenêtre F₁ de centre (x,y), plusieurs positions de recalage (x+dx,y+dy), définissant une fenêtre F₂, sont testées sur l'image 2 dans un voisinage V autour d'une position prévisionnelle (x+d_{0x},y+d_{0y}). Pour chaque position, le coefficient de corrélation centré normé est calculé :

$$CC(F_1, F_2) = \frac{\sum_{i,j} (F_1(i,j) - \overline{F_1})(F_2(i,j) - \overline{F_2})}{\sqrt{\sum_{i,j} (F_1(i,j) - \overline{F_1})^2 \sum_{i,j} (F_2(i,j) - \overline{F_2})^2}}$$

$\overline{F_1}$ et $\overline{F_2}$ sont les intensités moyennes des fenêtres.

L'ensemble des coefficients calculés sur V forme une surface de corrélation.



la corrélation est calculée pour toutes les fenêtres de V

surface de corrélation

Dans les algorithmes classiques, l'étape suivante est la recherche du pic de corrélation (le maximum C' de la surface), qui donne une mesure (dx,dy) de la translation locale. Cette mesure est effectuée en plusieurs points de l'image, et la déformation est estimée à partir de cet ensemble de mesures (optimisation par moindres carrés). Pour obtenir une bonne estimation, les mesures doivent être précises. Il faut donc bien choisir les fenêtres qui donneront un pic de corrélation net.

Avec notre méthode globale, l'optimisation exploite toute la surface de corrélation, qui peut donc être quelconque. Les fenêtres à corrélérer sont alors prises régulièrement réparties dans l'image selon une grille G. On veut estimer les 8 paramètres (a,b,c,d,e,f,g,h) qui définissent la déformation D la plus adaptée. Nous cherchons à optimiser la fonction Q(D) suivante :

$$Q(D) = \sum_{F_1 \in G} CC (F_1(x,y) , F_2(x+dx,y+dy))$$

(dx,dy) sont calculés par les formules de déformation (7) pour la déformation D testée, c'est à dire que pour chaque fenêtre F₁ prise sur la première image, les formules de déformation font correspondre une fenêtre F₂, avec laquelle on calcule la corrélation (en pratique, cette valeur est pré-calculée dans le voisinage V).

Cette fonction n'étant pas dérivable, l'algorithme d'optimisation choisi est la "descente de simplexe" (voir sa description dans [2]), qui consiste à essayer simultanément plusieurs hypothèses autour de la solution pour réaliser un encadrement de plus en plus fin. La maximisation de Q(D) par la descente de simplexe va donner une déformation qui maximise les corrélations pour l'ensemble des fenêtres.

Description de l'algorithme complet :

1- choix d'une déformation initiale : les capteurs de bord peuvent permettre son calcul, sinon l'initialisation peut être manuelle ou reprendre la déformation estimée entre les deux images précédentes.

2- calcul des corrélations pour des fenêtres régulièrement réparties sur l'image 1 et celles dans un voisinage V des fenêtres correspondantes sur l'image 2 par la déformation initiale.

3- optimisation par la descente de simplexe : partir de la déformation initiale, et arrêter quand la fonction Q(D) ne s'améliore plus.

remarques :

- au début de l'optimisation, il est possible que certains décalages (dx,dy) soient situés hors du voisinage V utilisé pour le calcul des corrélations. Dans ce cas, on pénalise Q(D) en retranchant la distance à la fenêtre initiale :

$$\text{dist} = \sqrt{(dx - d_0x)^2 + (dy - d_0y)^2}$$

La maximisation donne ainsi rapidement une déformation telle que tous les décalages soient situés dans les bons voisinages.

- pour Q(D), on calcule les décalages (dx,dy) en certains points (x,y) de la première image. Ces décalages ne sont pas arrondis au pixel le plus proche : la valeur de corrélation retenue est l'interpolation des quatre valeurs des pixels touchés.

IV - Résultats

Nous avons appliqué cette méthode globale de recalage à diverses séquences d'images. Nous présentons ici les résultats sur des images 400x400 pixels, pour lesquelles la déformation est connue (la photo 1 montre l'image utilisée, la photo 2 visualise la déformation). La déformation initiale est une translation entrée manuellement, assez éloignée de la déformation réelle. Le tableau ci-dessous compare les paramètres réels et les paramètres estimés :

	réel	estimé
a	-1.97 E-5	-1.19 E-5
b	4.88 E-5	3.81 E-5
c	-0.00603	-0.00542
d	0.0170	0.0167
e	-5.352	-5.394
f	-0.0119	-0.0119
g	-0.0287	-0.0291
h	11.17	11.33

On remarque que les paramètres les plus mal estimés sont a et b, car ils sont très faibles, mais leur influence est largement inférieure au pixel, même en bord d'image. Par contre, les autres paramètres estimés sont très proches des paramètres réels. Pour h, le paramètre ayant le plus d'influence, l'erreur est de 0.16 pixel seulement. La soustraction des images recalées confirme la qualité du résultat obtenu.

Une méthode classique de recherche des maxima de corrélation pour des points sélectionnés par pré-traitement aurait été confrontée à plusieurs problèmes dans ces images :

- l'image comporte plusieurs grandes plages uniformes dans lesquelles on ne peut sélectionner aucune fenêtre donnant des pics de corrélation marqués,
- dans certaines zones régulières, la surface de corrélation présentera plusieurs pics, et le choix du pic maximum sera peut-être une erreur.

Grâce à la méthode globale, on utilise aussi les surfaces de corrélation peu marquées des régions homogènes, et on n'a pas à choisir parmi les pics de corrélation dans les zones texturées.

Réglage des paramètres

Cette méthode globale repose sur trois paramètres (pas d'échantillonnage, taille de la fenêtre, voisinage), qui sont robustes et relativement indépendant des images. Nous avons effectué des tests systématiques pour les régler. La qualité du recalage est mesurée par la corrélation centrée normée entre les images recalées. Les valeurs de corrélation obtenues sont d'environ 0.95 pour les bons recalages.



1- le pas d'échantillonnage intervient uniquement pour fixer le nombre de fenêtres à corrélérer : il dépend donc simplement de la taille de l'image. Il faut au moins 30 à 40 fenêtres pour assurer un bon recalage, voire plus si les fenêtres sont petites. Au-delà de 100 fenêtres, les calculs sont plus longs alors que la qualité n'augmente plus.

2- la taille des fenêtres à corrélérer : trop petite, la corrélation apporte peu d'information, trop grande, les fenêtres risquent d'être déformées, et les calculs sont plus longs. Les tests effectués donnent comme optimale la taille 11x11 pixels.

3- le voisinage sur lequel est calculée la corrélation dépend uniquement de la précision sur la déformation initiale (qui peut être calculée pour une séquence donnée). Un voisinage trop grand entraîne des calculs inutiles. Nous avons utilisé un voisinage 17x17 sur nos images car la déformation cherchée était assez éloignée de la translation utilisée comme initialisation (une dizaine de pixels d'écart).

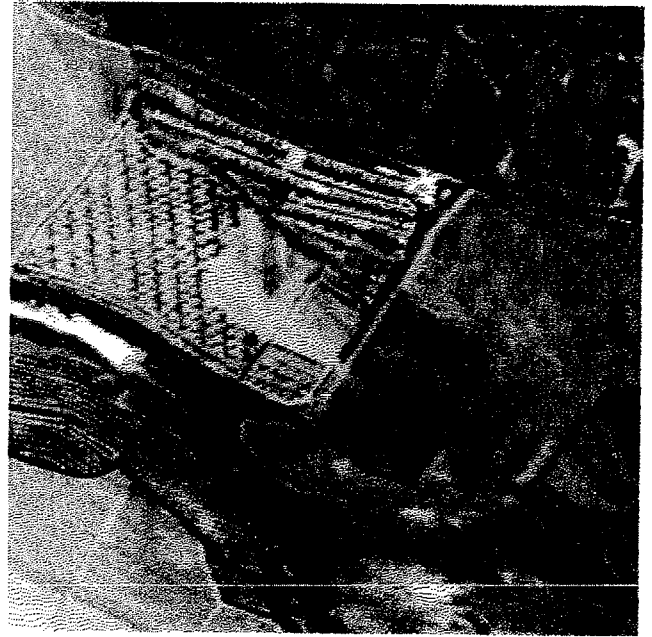


photo 1 : image aérienne prise près de Calvi (Corse). La deuxième image est obtenue en appliquant la déformation test.

V - Conclusion

Les résultats obtenus avec la méthode globale de recalage sont tout à fait satisfaisants visuellement, pour des temps de calcul raisonnables : les corrélations pourraient être cablées, et le simplexe converge en quelques centaines d'itérations. La méthode globale ne demande pas de pré-traitement et le réglage des trois paramètres a été optimisé.

Nous utilisons la méthode globale sur de nombreuses séquences d'images, pour réaliser diverses applications : suivi de points particuliers, amélioration d'image par addition de trames recalées, repérage de petits objets en mouvement.

Il reste maintenant à étudier l'influence du relief, et voir les limites de nos formules de déformation inter-images. Le principe de la méthode globale pourrait être repris pour des recalages différents avec d'autres formules de déformations.

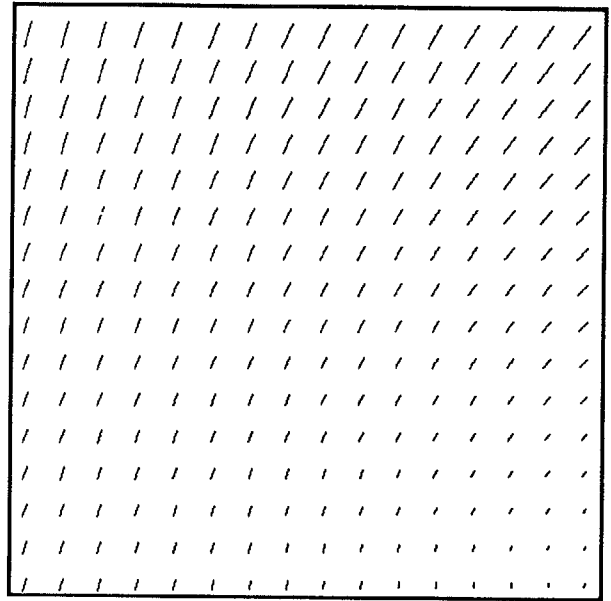


photo 2 : visualisation du champ de déformation.

BIBLIOGRAPHIE

- [1] **Vincent LECLERC**
"Recalage d'images angiographiques"
Thèse de Docteur Ingénieur - ENST - Octobre 1987
- [2] **W.K. Press, B.P. Flannery, S.A. Teukolsky, W.K. Vetterling**
"Numerical recipes in C - The Art of Scientific Computing"
Cambridge University Press
- [3] **J. Weng, N. Ahuja, T.S. Huang**
"Motion and Structure from Point Correspondances with Error Estimation : Planar Surfaces."
IEEE Transactions on Signal Processing, Vol 39, N° 12, Dec 91
- [4] **Ming XIE**
"Contribution à la vision dynamique : reconstruction d'objets 3D polyédriques par une camera mobile."
Thèse de l'université de Rennes - juin 89