

**RECALAGE D'IMAGES PAR ASSOCIATION DE PRIMITIVES,  
CONSTRUCTION ET ANALYSE D'HISTOGRAMMES  
MULTIDIMENSIONNELS**

Jean-Yves      DUFOUR  
Hugues        WALDBURGER

T.R.T. , Centre technique du PLESSIS-ROBINSON,  
5, avenue Réaumur, 92350 Le PLESSIS-ROBINSON

### RÉSUMÉ

Nous présentons une technique de recalage d'images mise au point dans le cadre de l'analyse de séquences d'images infrarouge. Nous nous sommes limités au cas où la transformation recherchée est la composée d'une rotation et d'une translation. L'extension de la complexité de la transformation est cependant tout à fait envisageable. La technique proposée est basée sur une analyse statistique de déplacements élémentaires. Dans un premier temps, nous abordons la détermination des déplacements élémentaires, réalisée par une mise en correspondance de listes de primitives extraites des deux images. Nous avons défini dans ce cadre un nouvel opérateur permettant de caractériser et de localiser ponctuellement les coins et les régions de petite taille dans une image. Ensuite nous présentons l'algorithme que nous avons conçu pour l'estimation des paramètres de la transformation. Enfin nous présentons les résultats obtenus sur des images infrarouges de scènes filmées en air-sol et en sol-sol.

### SUMMARY

A new method for image registration applied to infrared image sequences is presented. The proposed algorithm is restricted to a transformation composed of rotation and translation but it could be extended to more complex transformations. It is based on the statistical analysis of local displacements. First is presented the local displacements computation derived from the matching of features lists. In this context, a new operator dedicated to the detection of corners and small areas in images is introduced. Then we expose the algorithm used for geometrical transformation parameters estimation. Some results on air-to-ground and ground-to-ground images are included.

de 10 à 20 pixels). De plus, l'application visée impose la notion de réalisation temps-réel, ce qui entraîne des limitations au niveau des calculs et les mécanismes algorithmiques.

## 1 INTRODUCTION

Les travaux présentés ont été réalisés dans le cadre de la conception de systèmes de surveillance basés sur la détection de véhicules à partir d'images fournies par un senseur infrarouge (8-12  $\mu\text{m}$ ). La détection et l'analyse de mouvement d'objets peuvent s'avérer très utiles au sein de tels systèmes. Or, si leur application peut être simple lorsque la caméra est fixe, il n'en est pas de même lorsqu'elle est installée sur un support mobile. Il est alors nécessaire, avant l'application de tout autre traitement temporel sur une séquence d'images, d'estimer le déplacement du fond. Le recalage d'une image "courante" ICUR par rapport à une image "de référence" de la même séquence IREF est fondé sur une estimation des paramètres d'une transformation géométrique de forme connue. Cette estimation est réalisée par une analyse statistique de déplacements élémentaires déterminés en un certain nombre de points de IREF. La sélection de ces points est effectuée par l'application d'un critère portant sur des attributs calculés en chaque pixel et représentatifs de certaines propriétés de voisinage. Les contraintes que nous avons considérées tiennent essentiellement à la faible valeur de SNR caractéristique des images IR et à l'importance des déplacements élémentaires que l'on peut s'attendre à obtenir (de l'ordre

## 2 DETERMINATION DES DEPLACEMENTS ELEMENTAIRES

On peut envisager deux approches pour la détermination du déplacement d'un point donné.

La première consiste à déduire le déplacement de façon analytique à partir de propriétés de la séquence d'images, qui est alors considérée comme une fonction 3D au voisinage du point traité. De telles méthodes, dont les plus connues sont les méthodes différentielles [1], font appel à des techniques de filtrage 3D [2,3]; elles sont inadaptées à la résolution de notre problème car sensibles au bruit et surtout inefficaces dès lors que les déplacements excèdent quelques pixels.

La seconde approche consiste à rechercher au sein de l'image courante la position (ou éventuellement les positions) la plus probable du point de référence considéré.

Pour ce faire, on peut utiliser une mesure (comme la corrélation 2D) permettant de comparer le voisinage



du point de référence avec le voisinage de chaque point candidat dans l'image courante. Cependant, la complexité des calculs augmente très rapidement avec le déplacement maximal autorisé et avec la taille de la fenêtre de voisinage (qui doit être elle-même d'autant plus importante que le rapport signal à bruit de l'image est faible).

Nous avons donc opté pour une solution "discrète". Elle consiste à appliquer la technique de sélection de points de référence sur l'image courante puis à établir les correspondances entre les deux listes de points obtenues, sur la base de critères simples. Ces points sont représentatifs d'éléments caractéristiques de l'image habituellement désignés par le terme de "primitives". Les primitives que nous avons choisi de considérer sont respectivement les contours, les coins et les régions de petite taille (taches) de l'image traitée.

En général, les détecteurs de coins proposés dans la littérature sont fondés sur la recherche de points de contours présentant une variation maximale de direction de gradient [4]; les taches quant à elles sont détectées par des opérateurs mesurant le contraste local en chaque point de l'image analysée. Nous proposons dans cette partie un nouvel opérateur pour la localisation ponctuelle de ces deux types de primitives. Il est basé sur une analyse locale du champ de vecteurs formé par la dérivation de l'image convoluée avec une gaussienne.

## 2.1 DETECTEUR DE POINTS DE CONTOURS

La détection de points de contours est réalisée après convolution de l'image par un masque d'approximation de la fonction vectorielle "dérivée de gaussienne":

$$\vec{W}_\sigma = \vec{\nabla} W_\sigma = W_\sigma \cdot \vec{x} \cdot 1/\sigma^2,$$

où  $W_\sigma$  est une gaussienne bidimensionnelle isotrope d'écart-type  $\sigma$ . L'accroissement de la valeur de  $\sigma$  se traduit par une diminution de l'influence du bruit et par une déformation des contours réels de l'image. Notons que cette déformation n'est absolument pas gênante puisque nous ne cherchons pas les contours réels de l'image mais une primitive aussi stable que possible temporellement. Nous pouvons également remarquer le caractère incomplet des déplacements déterminés. En effet, un élément de contour peut être localement rectiligne et on ne peut donc estimer de son déplacement que la projection dans la direction perpendiculaire au contour.

## 2.2 DETECTEUR DE COINS ET DE TACHES

La déformation évoquée précédemment, dont on pourra trouver une analyse dans [5], constitue un élément essentiel pour la définition de l'attribut de caractérisation de coins et de taches. En effet, elle se traduit par un lissage qui transforme les contours anguleux en courbures locales: au voisinage d'une telle primitive, le champ de vecteurs gradient approximé présente alors une structure radiale, c'est-à-dire que les supports des vecteurs d'amplitude significative se croisent en un même point que l'on peut interpréter comme étant le centre de la tache ou le centre de courbure du coin. Notre opérateur a été conçu afin de permettre la détection de tels points. Il est fondé sur une mesure permettant de comparer des champs de directions:

$$M_P(\vec{V}_1, \vec{V}_2) = \int P.(\vec{V}_1, \vec{V}_2)^2 . d\vec{\mu} / \int P.d\vec{\mu}$$

avec  $d\vec{\mu} = F(\vec{x}) d\vec{x}$ , où  $F$  est une fonction "fenêtre" positive, d'intégrale égale à 1 et symétrique par rapport à l'origine et où  $P$  est une fonction à valeurs positives, que l'on peut interpréter comme une distribution spatiale de valeurs de vraisemblance.

Notons  $\vec{G}$  le champ de vecteurs gradient,  $G = |\vec{G}|$  et  $\vec{G}_u = \vec{G} / G$ . Nous définissons les champs de vecteurs suivants:

$$\vec{R}(\vec{x}) = \vec{x} / |\vec{x}|, \quad \vec{T} = \vec{R} \perp$$

$$\vec{D}_\alpha(\vec{x}) = (\cos \alpha, \sin \alpha)$$

Soit  $\vec{x}_0$  un point donné, nous noterons  $\text{Rad}(\vec{x}_0)$  la propriété:  $\vec{G}_u$  est radial au voisinage de  $\vec{x}_0$ . On peut associer une vraisemblance à  $\text{Rad}(\vec{x}_0)$  par:

$$P_1(\text{Rad}(\vec{x}_0)) = M_{G[\vec{x}_0]}(\vec{G}_u[\vec{x}_0], \vec{R}),$$

$$\text{en posant: } \vec{G}[\vec{x}_0](\vec{x}) = \vec{G}(\vec{x}_0 + \vec{x})$$

Cette mesure n'est pas satisfaisante car elle fournit des valeurs élevées pour des points situés au voisinage de contours rectilignes. Cependant, au voisinage d'un tel point, les vecteurs d'amplitude significative ont une direction identique. Il suffit donc, pour éviter ce problème, de retrancher de la vraisemblance  $P_1(\text{Rad}(\vec{x}_0))$  une vraisemblance d'avoir un champ à la fois radial et monodirectionnel, que nous avons définie en nous inspirant des probabilités:

$$P_1(\text{Rad,Dir}) = P_1(\text{Dir}|\text{Rad}) \cdot P_1(\text{Rad})$$

$$\text{avec: } P_1(\text{Dir}|\text{Rad}) = \text{Max}_\alpha (P_1(\text{Dir}_\alpha|\text{Rad}))$$

où  $\text{Dir}_\alpha(\vec{x}_0)$  désigne la propriété que le champ de vecteurs au voisinage de  $\vec{x}_0$  soit de direction uniforme  $\alpha$ . On pose alors:

$$P_1(\text{Dir}_\alpha(\vec{x}_0)|\text{Rad}(\vec{x}_0)) = M_{G[\vec{x}_0]}(\vec{R}, \vec{D}_\alpha)$$

On obtient finalement une nouvelle vraisemblance de la propriété  $\text{Rad}(\vec{x}_0)$  qui, après calcul, est égale à:

$$P_2(\text{Rad}) = (1 + \frac{\vec{K} * \vec{G}}{\vec{K} * G}) \cdot (1 - | \frac{\vec{K} * G}{\vec{K} * G} |)$$

où "\*" désigne le "produit scalaire de convolution":  $(f_x f_y) * (g_x g_y) = f_x * g_x + f_y * g_y$ ,  $\vec{G}$  est le champ de vecteur obtenu en multipliant par 2 la direction de  $\vec{G}$  et  $\vec{K}$  est le noyau de convolution défini en coordonnées polaires par:

$$K(\rho, \theta) = F(\rho), \quad \text{direction}(\vec{K}(\rho, \theta)) = 2\theta$$

Le calcul de  $P_2(\text{Rad})$  n'est pas trivial. Pour obtenir une mesure plus simple, nous avons utilisé la propriété complémentaire de  $\text{Rad}(\vec{x}_0)$  (notons la  $\text{Tan}(\vec{x}_0)$ ): "le champ de vecteur est tangentiel par rapport à  $\vec{x}_0$  au voisinage de ce dernier". Par un raisonnement analogue à celui utilisé pour définir  $P_2(\text{Rad})$ , nous définissons une vraisemblance  $P_2(\text{Tan})$  et nous adoptons comme mesure finale:

$$V(\text{Rad}) = (P_2(\text{Rad}) - P_2(\text{Tan})) / A$$

où  $A(\vec{x}_0)$ , égal à  $\int G[\vec{x}_0]$ ,  $F$ , est introduit pour tenir compte du contenu de  $\vec{G}$  au voisinage de  $\vec{x}_0$ . Après

calcul, on obtient la forme finale de notre opérateur:

$$V(\text{Rad}) = \vec{G} * \vec{K} - |G * \vec{K}|$$

### 3 ESTIMATION DES PARAMETRES

Nous ne pouvons décrire ici que les grandes lignes de l'algorithme, qui est développé plus en détail dans [6]. Notons  $\vec{u}$  le vecteur unitaire de la direction la plus représentée parmi les points de référence détectés avec l'opérateur GRADIENT (l'espace des directions étant quantifié),  $\vec{v} = \vec{u}^\perp$ ,  $(T_u, T_v)$  les composantes du vecteur translation dans le repère  $(\vec{u}, \vec{v})$ . Etant donnée la nature "incomplète" de l'information apportée par les points de contour, nous avons décomposé l'estimation des paramètres en deux étapes: l'estimation de  $\theta$  et  $T_u$ , puis celle de  $T_v$ .

#### 3.1 Estimation de $(\theta, T_u)$

Ils sont estimés en utilisant une liste  $(\vec{x}_i, \text{du}_i)$ ;  $\vec{x}_i$  est la position du point de référence, qui peut être un point de contour de vecteur normal  $\vec{u}$ , ou bien un centre de courbure.  $\text{du}_i$  est la projection du déplacement considéré sur l'axe porté par  $\vec{u}$ . Si le déplacement élémentaire correspondant a été correctement déterminé, alors le couple  $(\vec{x}_i, \text{du}_i)$  vérifie:

$$T_u = \text{du}_i - (\vec{x}_i \cdot \vec{v}) \cdot \theta \quad (1)$$

Autrement dit, le couple recherché est situé sur une droite  $D_i$  et, si la répartition des déplacements erronés est aléatoire, alors on peut déterminer ce couple en recherchant le point d'intersection le plus probable des droites  $D_i$  construites en considérant tous les éléments de la liste. Cette recherche peut être réalisée en quantifiant l'espace des valeurs possibles de  $(\theta, T_u)$ . Le principe de l'algorithme est alors le suivant:

Pour chaque droite  $D_i$  et pour chaque cellule de l'espace quantifié traversée par cette droite, on incrémente la cellule de même indice dans un histogramme 2D HIST2. La valeur recherchée est déterminée par l'indice de la valeur maximale de HIST2

On peut considérer la liste de couples  $(\text{du}_i, \vec{x}_i \cdot \vec{v})$  comme une liste de points appartenant à une image; l'expression (1) indique alors que la détermination de  $(\theta, T_u)$  est équivalente à la recherche des paramètres de la droite qui passe par tous les points déduits de déplacements non erronés. Nous pouvons alors constater que la démarche que nous avons employée s'apparente tout à fait à l'approche de HOUGH utilisée pour la reconnaissance de droites dans une image.

#### 3.2 Détermination de $T_v$

$T_v$  est estimé à partir d'une liste  $(\vec{x}_i, \vec{w}_i, \text{dw}_i)$ ;  $\vec{x}_i$  est la position du point de référence, qui peut être un point de contour de vecteur normal  $\vec{w}_i$  ( $\vec{w}_i \neq \vec{u}$ ) ou bien un centre de courbure, auquel cas  $w_i$  est fixé à  $\vec{v}$ ;  $\text{dw}_i$  est la projection du déplacement considéré sur l'axe porté par  $\vec{w}_i$ . Si le déplacement élémentaire a été correctement déterminé, alors le triplet  $(\vec{x}_i, \vec{w}_i, \text{dw}_i)$  vérifie:

$$T_v = (\text{dw}_i - \vec{x}_i \cdot \vec{w}_i) \cdot \theta - T_u \cdot c_i / s_i \quad (2)$$

$$\text{ou } (c_i, s_i) = (\vec{w}_i \cdot \vec{u}, \vec{w}_i \cdot \vec{v})$$

L'estimation de  $T_v$  est réalisée par la construction et l'analyse d'un histogramme 1D: pour chaque élément de la liste, on évalue et quantifie le terme droit de (2) puis on incrémente la cellule de l'historgramme ayant même indice que la valeur quantifiée. L'indice de la valeur recherchée est égal à l'indice du maximum de l'historgramme.

### 4 RESULTATS

Sur la page suivante, nous avons fait figurer quelques résultats concernant deux couples d'images. Les images du couple 1 ont été filmées à partir d'un hélicoptère volant à faible altitude. La scène représente un canal avec une autoroute en contre-bas et une route en peu plus en hauteur. L'image de référence est en fig. 1. La différence d'images après recalage est en fig. 2 et fait ressortir la présence de véhicules sur les deux routes. Les courbes en fig. 3A sont utilisées pour estimer  $\theta$  ( $C(m) = \text{Max} \{ \text{HIST2}[m, n] \}$ ),  $T_u$  ( $C(n) = \text{HIST2}[m^*, n]$ , où  $m^*$  est l'indice de  $\theta$ ) et  $T_v$  ( $C3(n) = \text{HIST1}[n]$ ). Le pas de quantification est de  $0.5^\circ$  pour  $\theta$  et de 1 pixel pour  $T_u$  et  $T_v$ . En ce qui concerne le second couple, les images ont été filmées par avion. Nous avons fait figurer l'image de référence (fig. 4), l'image différence sans recalage (fig. 5) et après recalage (fig. 6), et les courbes d'estimation (fig. 3B; pas de quantification:  $0.25^\circ$  et 0.5 pixel)

### 5 CONCLUSION

Nous avons effectué nos travaux dans le cadre de l'analyse de séquences d'images infrarouge mais nous n'avons retenu des caractéristiques de ces dernières que la pauvreté de leur contenu, en terme de rapport signal à bruit. Les primitives que nous avons choisies pour la détermination de déplacements élémentaires sont présentes dans la grande majorité des images que l'on peut avoir à traiter dans la pratique. Nous avons limité la complexité des transformations mais l'extension de notre algorithme d'estimation de paramètres à des modèles plus complexes est tout à fait envisageable.

Nous pensons donc que cette technique possède un caractère suffisamment général pour pouvoir être utilisée dans de nombreuses applications de recalage d'images.

### 6 BIBLIOGRAPHIE

- [1] NAGEL H.H.: *Image sequences - Ten (octal) years - from phenomenology towards a theoretical foundation*. Proc. IEEE ICPR 1986 (Paris)
- [2] ADELSON E.H. & BERGEN J.R.: *The extraction of spatio-temporal energy in human and machine vision*. Proc. IEEE Workshop on Motion 1986
- [3] BREMOND O. & WALDBURGER H.: *Détection de mouvement par filtrage spatio-temporel*. GRETSI 89
- [4] SHAH M.A. & JAIN R.: *Detecting time-varying corners*. Proc IEEE ICPR 1984 (Montreal)
- [5] BERGHOLM F.: *Edge focusing*, IEEE Trans. on PAMI, vol. PAMI-9, n° 6, nov. 1987
- [6] DUFOUR J.Y. *Contribution algorithmique à la conception d'un système intégré d'analyse temps-reel de scènes dynamiques: Reconnaissance d'objet et analyse de mouvement dans une séquence d'images*. Thèse de doctorat, Université de Caen. 1988

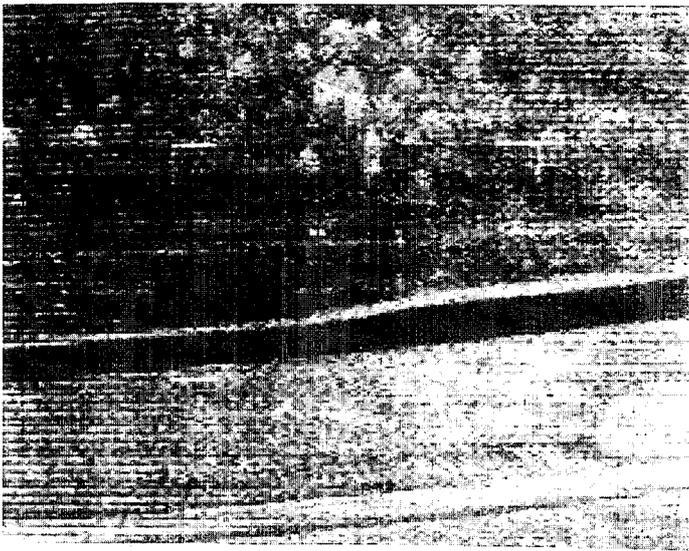


figure 1: couple 1, image de référence



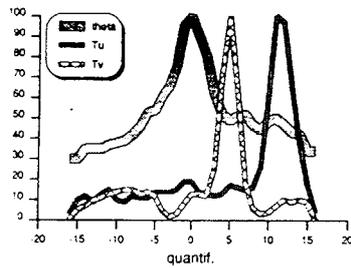
figure 4: couple 2, image de référence



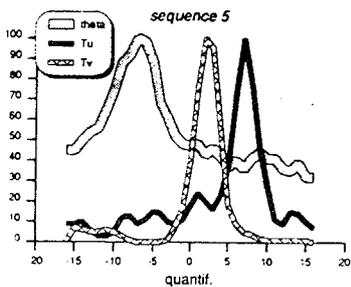
figure 2: couple 1, différence après recalage



figure 5: couple 2, différence avant recalage



A: Couple 1



B: Couple 2

figure 3 : courbes de détection

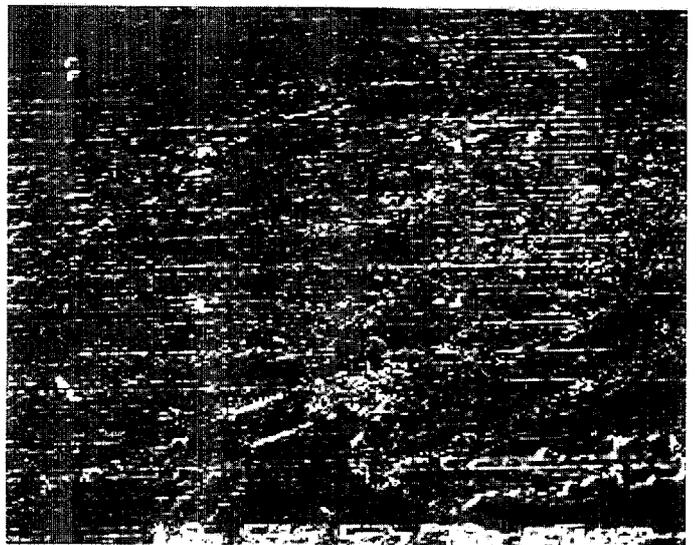


figure 6: couple 2, différence après recalage