

Stratégie de vote pour la mise en correspondance de lignes de niveaux

Cumulative decision strategy for level-line matching

Samia BOUCHAFA

Bertrand ZAVIDOVIQUE

Institut d'Electronique Fondamentale, UMR 8622 - CNRS

University Paris XI

91405 Orsay Cedex, France

samia.bouchafa@ief.u-psud.fr

Résumé

Une nouvelle stratégie de mise en correspondance cumulative inspirée de la transformée de Hough est proposée pour le recalage d'images en particulier et l'appariement d'images en général. Cette approche s'appuie sur des primitives image - les lignes de niveaux - particulièrement robustes aux variations de luminosité de la scène, elle ne nécessite aucune estimation initiale de la transformation recherchée et permet de lever les ambiguïtés d'appariement engendrées par les motifs répétitifs naturellement présents dans les images réelles. Le recalage en lui-même est réalisé à travers un processus cumulatif de mise en correspondance de segments basé sur un vote à plusieurs tours. Chaque tour fournit ainsi une estimation de la transformation recherchée que le tour suivant est chargé de raffiner si nécessaire. Dans l'application considérée ici, les transformations sont affines, incluant rotation, translation et échelle. L'approche est naturellement généralisable à d'autres types de transformations.

Mots Clef

Lignes de niveaux, mise en correspondance de primitives image, recalage, transformée de Hough.

Abstract

A new level-line registration technique is proposed for image transform estimation. This approach is robust towards contrast changes, does not require any prior estimate of the unknown transformation between images and tackles very challenging situations that usually lead to pairing ambiguities : repetitive patterns in the images. The registration by itself is performed through an efficient level-line cumulative matching based on a multi-stage primitive election procedure. Each stage provides a coarse estimate of the transformation that the next stage gets to refine. Even if we deal in this paper with similarity transform (rotation, scale and translation), our approach can be easily adapted to more general transformations.

Keywords

Level lines, matching, registration, Hough transform.

1 Introduction

De nombreux domaines d'application tels que la stéréovision, l'analyse du mouvement 2D/3D, la reconnaissance de formes ou la reconstitution de panoramas, reposent essentiellement sur l'appariement d'images. Les images considérées peuvent être acquises à partir de points de vues différents, à deux instants différents ou à partir de capteurs différents. Ces dernières décennies ont vu apparaître une grande variété de techniques de mise en correspondance, chacune adaptée à l'application visée. Le nombre croissant de ces nouvelles approches a eu pour conséquence la publication de synthèses très complètes permettant de comparer - selon des critères très divers - les techniques existantes [5], [18], [16], [26], [1]. Parmi les critères de distinction utilisés, nous citerons :

- La nature (fréquentielle ou spatiale) du domaine considéré. Les approches fréquentielles sont basées sur la corrélation des phases et exploitent des propriétés basiques de la transformée de Fourier (ou par ondelettes) [27]. Quant aux approches spatiales, elles s'appuient sur une extraction et une comparaison de primitives image.
- La nature de la transformation recherchée (euclidienne, affine, projective ou élastique). Cette transformation dépend en général de l'application considérée. Chaque transformation particulière implique des invariants adéquats, calculés à partir des primitives extraites : longueur, orientation, rapports de longueur, incidence, birapport, etc.
- Les primitives choisies. Les plus usuelles sont les points particuliers [21], les contours [23], [19], les surfaces [25], les régions [10], les lignes [30], ou les descripteurs de Fourier [15].

- La mesure de similarité entre primitives, qui dépend aussi du type de primitives utilisées. Les plus classiques sont : la corrélation croisée [28], la somme des différences en valeur absolue [4] ou la distance de Levenstein [12].
- L'espace de recherche et la stratégie d'appariement. De manière simplifiée, les stratégies sont basées sur une mise en correspondance de primitives qui est explicite (ex. recherche exhaustive, programmation dynamique [12], relaxation [29], transformée de Hough [3]), ou implicite (ex. erreur quadratique, minimisation [31], programmation linéaire [2]).

Notre étude est partie de trois constatations :

- D'abord, les images à recaler sont rarement acquises dans les mêmes conditions d'éclairément. Pourtant, la plupart des primitives utilisées sont sensibles aux variations de contraste. Elles sont soit directement dépendantes des niveaux de gris (points, régions), ou alors souffrent d'un manque de techniques réellement indépendantes du contraste (et/ou nécessitent des seuils difficiles à ajuster) capables de les extraire (contours) [7].
- Ensuite, les motifs répétitifs - pourtant fréquents dans les images naturelles - engendrent des ambiguïtés d'appariement qui sont rarement prises en considération dans les stratégies de mise en correspondance existantes.
- Enfin, dans les applications temps-réel, la robustesse du système complet tirerait profit d'un processus de décision progressif qui facilite un contrôle rapide par une première estimée des paramètres de la transformation recherchée, puis si nécessaire son raffinement.

Notre approche tente alors de fournir des éléments de réponse aux problèmes cités ci-dessus. Pour cela, nous avons choisi une classe de primitives adaptées : les lignes de niveaux, robustes vis-à-vis de variations de contraste [7], [20]. Soulignons que d'autres existent : Li et al. [17] exploitent par exemple une variable de chromaticité par nature indépendante de la luminosité. Nous avons utilisé les lignes de niveaux avec succès dans le passé en analyse du mouvement, en particulier pour des scènes d'extérieur, là où l'hypothèse d'invariance de la luminosité n'est pas vérifiée [6], [11]. De plus, elles sont particulièrement adaptées à une stratégie d'appariement basée sur un processus de vote. En effet, le nombre important de lignes de niveaux dans l'image y entraîne de fortes redondances locales et permet d'envisager un processus de mise en correspondance cumulatif où chaque ligne est appelée au vote à une phase donnée en fonction de sa fiabilité. Pour restreindre l'espace de vote, nous exploitons une technique de décision sur graphe bi-partite dite des mariages stables [13], [14]. Chaque primitive crée une liste de préférence des primitives dans l'autre image, classées des plus aux moins similaires en fonction de mesures ou

métriques adaptées. Ces préférences pondèrent le processus de décision qui se déroule en plusieurs étapes : les primitives les plus fiables participent d'abord au vote, les autres sont sollicitées ensuite pour confirmer ou infirmer le vote précédent. La transformation retenue, issue des couples de primitives, est celle qui récolte le maximum de votes (voir FIG. 1).

Notre article s'articule autour de 3 parties principales : d'abord, nous décrivons le processus d'extraction des lignes de niveaux qui fournit en sortie un ensemble de segments et leurs caractéristiques. Ensuite, nous définissons la notion de fiabilité de lignes de niveaux et la manière dont les segments issus de ces lignes sont triés par ordre de fiabilité et groupés pour former des primitives. Enfin, la dernière partie est consacrée à la description du processus d'appariement proprement dit à travers d'abord la construction des listes de préférence et ensuite l'estimation - basée sur un vote à plusieurs tours - de la transformation recherchée.

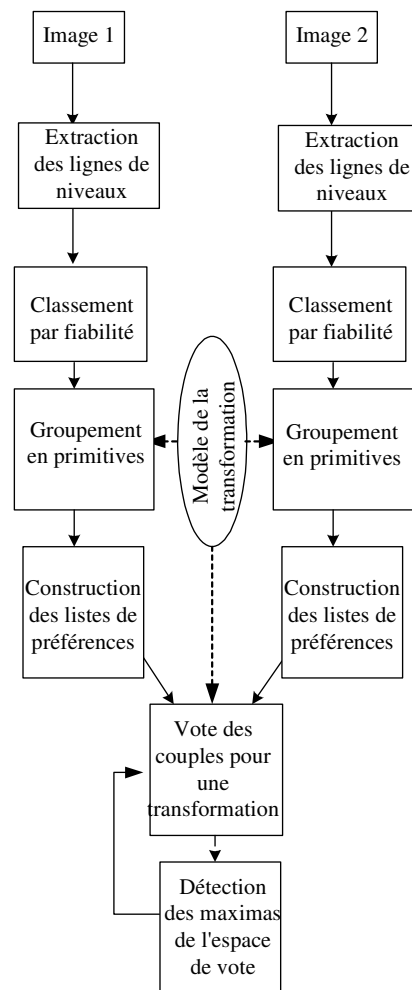


FIG. 1: Vue générale de notre approche.

2 Extraction des lignes de niveaux

Soit $I(\mathbf{p})$ l'intensité lumineuse du pixel de coordonnées $\mathbf{p}(x, y)$. L'ensemble de niveaux λ de l'image I est l'ensemble de tous les points ayant une luminosité supérieure ou égale à λ :

$$\mathbb{N}_\lambda^I = \{\mathbf{p} / I(\mathbf{p}) \geq \lambda\}.$$

Les ensembles de niveaux sont une représentation complète de l'image puisqu'il est possible de la reconstituer à partir des $\mathbb{N}_\lambda^I$:

$$I(\mathbf{p}) = \sup \{\lambda / I(\mathbf{p}) \geq \lambda\}.$$

La frontière d'un ensemble de niveaux est appelée "ligne de niveaux". La représentation par lignes de niveaux a été étudiée en détails par V. Caselles, J.M. Morel et J. Froment [7], [9]. Notons en particulier sa propriété d'invariance par rapport aux changements de contraste et sa commutation avec une transformation affine (translation, rotation, zoom). Caselles propose alors de considérer les lignes de niveaux comme les atomes de base à partir desquels devraient s'appuyer tout algorithme d'analyse d'images.

L'extraction des lignes de niveaux est souvent réalisée - à partir des ensembles de niveaux associés - en effectuant simplement une série de seuillages. L'approche que nous proposons permet une exploitation immédiate et efficace du résultat en fournissant directement des groupes rectilignes de lignes de niveaux superposées. En effet, nous exploitons une propriété élémentaire des ensembles de niveaux que traduit leur inclusion :

$$\forall \lambda > \mu \quad \mathbb{N}_\lambda^I \subset \mathbb{N}_\mu^I,$$

les lignes de niveaux peuvent se superposer mais jamais se croiser. Nous proposons alors un simple processus récursif de suivi de lignes par groupe (se superposant) qui s'arrête lorsque ce dernier cesse d'être rectiligne. On peut vérifier que la complexité totale n'est pas supérieure à celle d'une détection et codage de lignes à base de multi-seuils. De plus, notre processus n'utilise à aucun moment les niveaux de gris directement mais seulement leur ordre relatif, ce qui le rend particulièrement robuste aux variations de contraste. Il démarre en chaque point¹ $\mathbf{p}_o(x, y)$, détermine lequel parmi ses 4 voisins $\mathbf{p}_o(x, y + 1)$, $\mathbf{p}_o(x - 1, y)$, $\mathbf{p}_o(x, y - 1)$ ou $\mathbf{p}_o(x + 1, y)$ est le successeur, selon les 4 chemins possibles (voir FIG. 2). Chaque successeur sélectionné devient à son tour le point courant $\mathbf{p}_k$ et le processus se répète jusqu'à ce qu'un des critères d'arrêt cesse d'être vérifié.

$\mathbf{p}_{i=1,4}^k$ est un successeur de $\mathbf{p}_k$, à l'étape k si les conditions suivantes sont vérifiées :

Condition 1 Au moins une ligne de niveau passe entre les pixels $\mathbf{q}$ et $\mathbf{r}$: $|I(\mathbf{r}_k) - I(\mathbf{q}_k)| \geq \text{seuil}$.

¹En réalité, les lignes de niveaux passent **entre** les pixels. Le référentiel est donc placé au niveau de l'espace inter pixel.

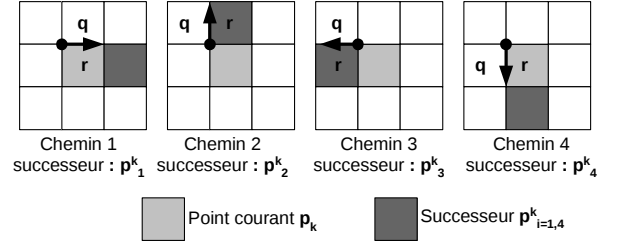


FIG. 2: Le point courant et ses 4 successeurs possibles en fonction des chemins 4-connexité.

Dans le but de prendre en compte les effets de quantifications, le seuil de différence de niveaux ci-dessus est fixé à 2 (pas de quantification + 1). Par conséquent, toutes les lignes sont extraites à ce stade sans aucune présélection. Il est à noter que le processus d'extraction ne nécessite à aucun moment de fixer des seuils arbitraires.

Condition 2 Les lignes de niveaux suivies se poursuivent effectivement :

$$[\min(I(\mathbf{q}_{k-1}), I(\mathbf{r}_{k-1})), \max(I(\mathbf{q}_{k-1}), I(\mathbf{r}_{k-1}))] \cap [\min(I(\mathbf{q}_k), I(\mathbf{r}_k)), \max(I(\mathbf{q}_k), I(\mathbf{r}_k))] \neq \emptyset$$

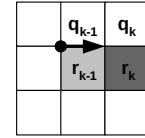


FIG. 3: Condition de test de la poursuite effective d'un groupe de lignes de niveaux.

Cette condition permet de garantir le suivi du même groupe de lignes (voir FIG. 3). Certaines lignes peuvent toutefois se séparer du groupe en empruntant des chemins différents. Le processus de suivi étant récursif, chaque nouveau chemin engendré est aussi exploré si la condition 2 est toujours vérifiée. Cette dernière garantit en fait la non confusion des lignes et le respect de la propriété d'inclusion des ensembles de niveaux.

Condition 3 L'intérieur (resp. extérieur) des ensembles de niveaux - associés au groupe de lignes de niveaux suivi - est toujours laissé du même côté.

Ce qui se traduit par : $I(\mathbf{r}_k)$ est toujours plus grand (resp. plus petit) que $I(\mathbf{q}_k)$ pour le même groupe de lignes de niveaux. Cette condition évite ainsi les retours en arrière.

Condition 4 Les lignes de niveaux suivies sont toujours rectilignes.

Cette condition est testée en calculant la distance maximale entre les points déjà parcourus et la corde formée par le point initial et le point courant (voir FIG. 4). Cette distance ne doit pas dépasser un seuil. Les points déjà parcourus ne sont bien sûr pas stockés. Seul le code de Freeman 4-connexité correspondant au chemin parcouru entre

deux points successifs est gardé à chaque étape. L'origine du référentiel - pour le calcul de l'équation de droite - est le point initial p_o .

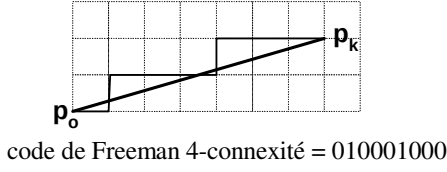


FIG. 4: La condition 4 permet de vérifier que les lignes suivies sont toujours rectilignes en calculant la distance entre chaque point parcouru et la droite formée par le point initial et le point courant.

Le processus itératif s'arrête lorsque l'une des conditions ci-dessus n'est plus vérifiée. Il est possible alors d'associer au point initial p_o :

- le milieu $\mathbf{p} = \frac{\mathbf{p}_o + \mathbf{p}_k}{2}$;
- l'orientation approchée de ce groupe de lignes rectilignes $\theta_{\overrightarrow{p_o p_k}}$;
- le contraste moyen de part et d'autre des lignes suivies : $c = \frac{\sum |I(\mathbf{r}_k) - I(\mathbf{q}_k)|}{k}$;
- et enfin la longueur : $l = \|\overrightarrow{p_o p_k}\|$.

A partir de ces informations, il est possible de grouper tous les points appartenant effectivement au même *segment* et d'obtenir ainsi un ensemble de *segments* $\{L_I^i\}_{i=1,\eta}$ associés à l'image I avec leurs caractéristiques :

$$\vec{L}_I = [\mathbf{p}_i \quad \theta_i \quad l_i \quad c_i]^T.$$

Définition 5 Un "segment" désigne un groupe de lignes de niveaux rectilignes obtenu à partir de la procédure de suivi.

3 Classement des segments par fiabilité

Les lignes de niveaux n'ont pas toutes la même importance visuelle et ne coïncident pas toujours avec les contours que nous percevons à l'oeil nu. Certaines, en effet, séparent des ensembles de niveaux en réalité très proches en terme de niveaux de gris. Par ailleurs, les *segments* de faible longueur sont probablement associés à des ensembles de niveaux de taille réduite générés par du bruit. Il existe plusieurs moyens de sélectionner les lignes de niveaux qui correspondent le mieux aux contours perçus. Citons à ce propos les travaux de Jacques Froment [9] ou ceux de N. Paragios et R. Deriche [24]. Nous avons choisi, dans le cadre de l'application que nous traitons, de tirer profit des différences perceptuelles entre les lignes extraites afin de définir une notion de fiabilité basée sur la longueur l du *segment* et le contraste moyen c de part et d'autre. Les *segments* jugés peu fiables ne seront pas pour autant complètement écartés du processus de décision : ils

y contribuent avec toutefois un plus faible crédit. Afin de préparer cette étape, nous avons choisi de classer les *segments* $\{L_I^i\}_{i=1,\eta}$ par ordre décroissant (des plus fiables aux moins fiables) du produit $l \times c$ (voir FIG. 5).

4 Groupement des segments

Le groupement des *segments* est une étape qui dépend du modèle de la transformation recherchée. En effet, chaque *segment* fournit un seul point (le milieu de la ligne, donc 2 coordonnées). Ce point n'est pas suffisant pour estimer une transformation affine à 4 paramètres (dx, dy, α, ψ) (où dx et dy sont les translations selon x et y respectivement, ψ le facteur d'échelle et α la rotation plane). En théorie, il est possible de prendre en compte les extrémités de chaque *segment* (4 coordonnées) afin d'estimer cette transformation. En pratique (bruit), il est plus intéressant, en terme de stabilité et de robustesse, de s'appuyer sur deux points, provenant de deux *segments* différents (les milieux), plutôt que sur deux points provenant d'un seul *segment* (les extrémités). L'estimée de la transformation - même si la mise en correspondance des lignes est parfaite - repose entièrement sur les coordonnées des points appariés fournis. Ceux-ci doivent correspondre avec la meilleure précision à la position réelle des *segments* dans l'image.

Par ailleurs, la mise en correspondance nécessite de choisir au préalable des caractéristiques associées aux primitives choisies qui soient invariantes par rapport à la transformation recherchée [22]. Nous savons que la longueur et l'orientation d'un *segment* ne le sont pas par transformation affine. En revanche, le rapport des longueurs et la différence des orientations sont effectivement invariants.

Pour toutes ces raisons, nous décidons de grouper les *segments* par couple appelé "angle" pour éviter la confusion avec les couples inter images à appairer plus tard afin de déterminer la transformation. Les deux points milieu serviront à l'estimation directe de la transformation. Le rapport des longueurs et la différence d'orientations entre les 2 *segments* serviront de caractéristiques invariantes de la primitive. Si une transformation plus complexe était recherchée la primitive *angle* ferait place à un triplet nommé "zig-zag" ou "coin", etc. afin de fournir le nombre de paramètres adéquat.

Définition 6 Une "primitive" désigne un ensemble de p *segments*.

Définition 7 Un "angle" désigne un ensemble de $p = 2$ *segments*.

Définition 8 Un "couple" de primitives u et v désigne deux primitives issues de deux images différentes.

Soit deux *segments* $\vec{L}_I^i$ et $\vec{L}_I^j$ qui forment une primitive *angle* : $\vec{C}_I^{i,j} = [\mathbf{p}_i \quad \mathbf{p}_j \quad \Delta\theta \quad \ell \quad c_i \quad c_j]^T$ où : $\Delta\theta = \theta_i - \theta_j$, $\ell = \frac{l_i}{l_j}$.

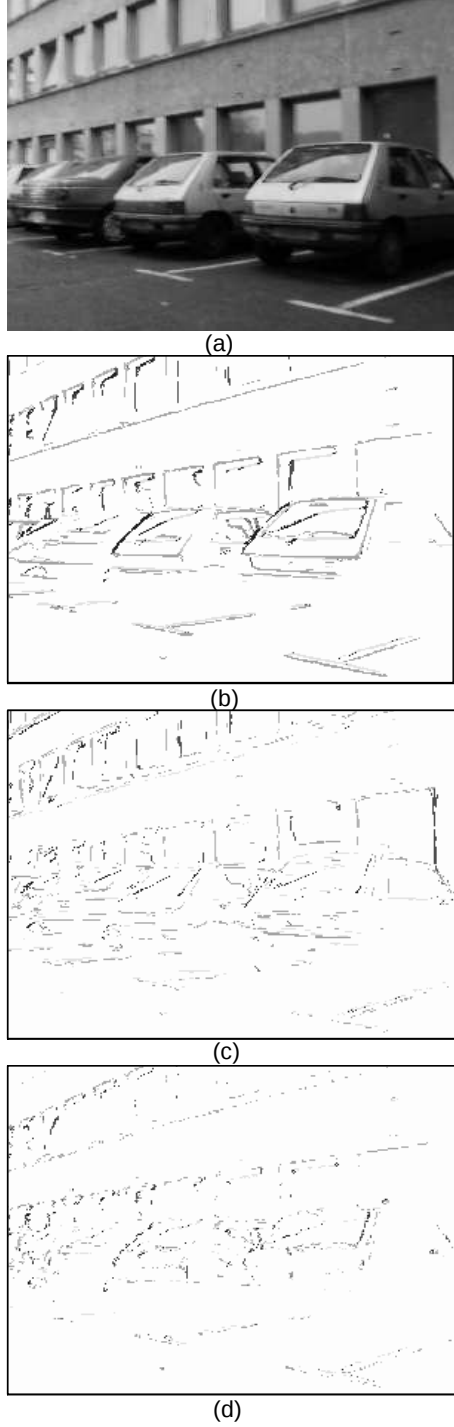


FIG. 5: Exemple de classement des segments des plus fiables aux moins fiables. a) image originale extraite d'une séquence prise par une caméra embarquée (Parking IEF). b) Les $\frac{\eta}{4}$ premiers segments qui correspondent aux plus fiables. Le nombre total de segments = η . Les segments sont retracés sur l'image en utilisant les caractéristiques trouvées (longueurs, orientation, point milieu). c) Les $\frac{\eta}{4}$ suivants (moins fiables). d) Les $\frac{\eta}{4}$ suivants (encore moins fiables). etc.

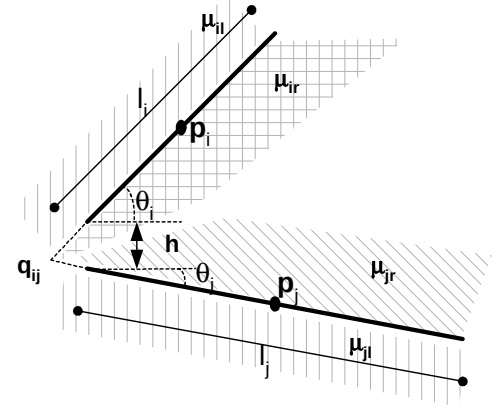


FIG. 6: Le couplage des segments pour former une primitive "angle" s'appuie sur les conditions de non collinéarité et de proximité.

Cette primitive est considérée valide ssi deux conditions sont vérifiées :

1. les *segments* ne sont pas colinéaires
2. et ils sont suffisamment proches en terme de distance (voir FIG. 6).

De la même manière que pour les *segments*, les primitives *angle* sont classées par fiabilité en fonction de la fiabilité des *segments* qui les composent².

5 Mise en correspondance

La notion de fiabilité précédemment introduite nous amène à envisager une stratégie de vote à plusieurs tours où chaque tour permet à une nouvelle catégorie de votants d'exprimer leur opinion, en fonction de leur fiabilité. Nous nous sommes inspirés de la technique des mariages stables [13], [14], [32] afin de construire pour chaque primitive une liste de préférence des $\kappa_{\max}$ -plus proches primitives trouvées dans l'autre image. Le *couple* inter images [primitive, correspondant potentiel] permet alors d'obtenir une estimée et donc un vote pour une transformation donnée. Ce vote aura plus ou moins de crédit en fonction de la fiabilité des primitives votant et de la position du candidat dans la liste de préférence. Détaillons ci-dessous cette procédure.

5.1 Construction des listes de préférence à longueur variable

Chaque primitive $u \in \{\vec{C}_I^k\}_{k=1, N}$ de l'image I construit une liste de préférence triée de primitives parmi les primitives de l'image J , des plus similaires aux moins similaires, à l'aide d'une mesure de distance basée sur le rap-

²Pour des raisons d'efficacité en temps d'exécution, le couplage en primitive angle repose sur un algorithme rapide - que nous ne développerons pas ici - combinant les *segments* à partir de la liste de *segments* déjà triée par fiabilité.

port des longueurs, la différence des orientations et le contraste moyen de part et d'autres des lignes. La distance $dist$ entre u et $v \in \{\vec{C}_{j,k}^k\}_{k=1,M}$ est définie simplement par :

$$dist(u, v) = k_1 |\Delta\theta_u - \Delta\theta_v| + k_2 |\ell_u - \ell_v| + k_3 (|c_{iu} - c_{iv}| + |c_{ju} - c_{jv}|).$$

k_1, k_2 et k_3 permettent d'ajuster l'importance relative de chaque caractéristique et :

$$x = [\mathbf{p}_{ix} \quad \mathbf{p}_{jx} \quad \Delta\theta_x \quad \ell_x \quad c_{ix} \quad c_{jx}]^T.$$

Il est à noter que $\sum k_i = 1$ et $k_3 < k_{i=1,2}$ car le contraste n'est pas une caractéristique invariante mais reste néanmoins à considérer avec un plus faible poids. Bien sûr, chaque différence en valeur absolue est normalisée par la valeur maximale prise afin d'obtenir une distance finale comprise entre 0 et 1 (de la plus à la moins ressemblante). Chaque primitive construit sa liste des κ -plus proches ($\kappa < \kappa_{\max}$) primitives en utilisant la distance définie ci-dessus. Il est important de noter que la taille de la liste de préférence est variable d'une primitive à l'autre. Les nouveaux candidats sont insérés dans cette liste jusqu'à ce que la distance obtenue soit trop élevée. En fait, le choix de listes de préférences variables est nécessaire afin d'éviter l'élimination arbitraire d'un candidat en raison de tailles de listes fixées *a priori* s'avérant inadaptées *a posteriori*. Dans la pratique, nous avons fixé prophylactiquement $\kappa_{\max} = 10$, qui n'a jamais été atteint. Adapter cette borne automatiquement, si cela s'avérait nécessaire, en fonction de la distribution des *segments* (longueurs, orientations), ne présente pas de difficulté.

5.2 Processus de vote

A chaque tour de vote, une nouvelle catégorie de *couples* exprime son opinion en fonction de sa fiabilité. Cette stratégie a deux avantages :

1. La disponibilité à chaque tour d'une approximation même grossière de la transformation recherchée peut être - même si elle n'est pas très précise - utilisée dans certaines applications nécessitant un résultat immédiat.
2. En raison des motifs répétitifs présents dans les images naturelles, conduisant fatalement à des ambiguïtés d'appariement, le processus qui consiste à inviter au vote des primitives moins fiables - mais souvent plus nombreuses - donne une nouvelle chance à l'émergence de la solution exacte.

Les *couples* sont classés en autant de catégories que de tours de votes en divisant simplement la liste de primitives. Au premier tour, la première catégorie vote. Si un pic unique apparaît dans l'espace de vote alors une transformation obtient une majorité absolue : le processus de vote peut donc s'arrêter. Sinon, un second tour peut démarrer autorisant une nouvelle population à voter : les primitives ayant une fiabilité moindre. S pics dans l'espace de

vote sont sélectionnés. Les valeurs cumulatives des *maxima* sélectionnés sont amplifiées pour le tour suivant afin de privilégier les décisions des primitives les plus fiables. Le vote s'arrête lorsqu'il n'y a plus de primitives ou lorsqu'un pic émerge, récoltant ainsi la majorité absolue.

L'espace de vote V est à 4 dimensions (2 translations, 1 rotation et 1 échelle). $V(dx, dy, \alpha, \psi)$ représente la valeur cumulative pour la transformation (dx, dy, α, ψ) . Les translations peuvent varier entre $[-\mathbf{dx}_{max}, \mathbf{dx}_{max}]$ et $[-\mathbf{dy}_{max}, \mathbf{dy}_{max}]$ respectivement. L'origine - correspondant à la translation nulle - est placée au milieu de l'espace V . La rotation peut varier entre $[\alpha_{\min}, \alpha_{\max}] = [0^\circ, 360^\circ]$. Enfin, l'échelle varie entre $[\psi_{\min}, \psi_{\max}]$.

Chaque couple de primitives (u, v) vote - si le couple se préfère mutuellement³ - en fonction de la position de v dans la liste de préférence de u pour une transformation donnée. La transformation (4 inconnues : dx, dy, α et ψ) est estimée en résolvant simplement un système de 4 équations (chaque point fournit 2 équations, une pour chaque coordonnée x et y) :

$$\mathbf{p}_{ku} = \begin{bmatrix} \psi & 0 \\ 0 & \psi \end{bmatrix} \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix} \mathbf{p}_{kv} + \begin{bmatrix} dx \\ dy \end{bmatrix}$$

Avec $k = i$ ou j . Posons :

$$\begin{cases} U = \psi \cos(\alpha) \\ V = \psi \sin(\alpha) \\ Z = dx \\ T = dy \end{cases}$$

Le système précédent s'écrit alors :

$$\begin{bmatrix} x_{p_{iv}} \\ y_{p_{iv}} \\ x_{p_{jv}} \\ y_{p_{jv}} \end{bmatrix} = \begin{bmatrix} x_{p_{iu}} & y_{p_{iu}} & 1 & 0 \\ y_{p_{iu}} & x_{p_{iu}} & 0 & 1 \\ x_{p_{ju}} & -y_{p_{ju}} & 1 & 0 \\ y_{p_{ju}} & x_{p_{ju}} & 0 & 1 \end{bmatrix} \begin{bmatrix} U \\ V \\ Z \\ T \end{bmatrix} = M \begin{bmatrix} U \\ V \\ Z \\ T \end{bmatrix}$$

Où :

$$\begin{bmatrix} U \\ V \\ Z \\ T \end{bmatrix} = M^{-1} \begin{bmatrix} x_{p_{iv}} \\ y_{p_{iv}} \\ x_{p_{jv}} \\ y_{p_{jv}} \end{bmatrix}$$

et x, y sont les coordonnées du point milieu. Ce qui donne :

$$\begin{cases} dx = Z \\ dy = T \\ \alpha = \arctg\left(\frac{V}{U}\right) \\ \psi = \frac{U}{\cos(\alpha)} = \frac{V}{\sin(\alpha)} \end{cases}$$

³Préférence mutuelle = "u est dans la liste de préférence de v et v est dans la liste de préférence de u".

ALGORITHME :

INITIALISATION

$\{\text{maxima sélectionnés}\} = \{dx^s, dy^s, \alpha^s, \psi^s\}$

$dx^s = \{-\mathbf{dx}_{max}, \mathbf{dx}_{max}\},$

$dy^s = \{-\mathbf{dy}_{max}, \mathbf{dy}_{max}\},$

$\alpha^s = \{\alpha_{min}, \alpha_{max}\},$

$\psi^s = \{\psi_{min}, \psi_{max}\}.$

POUR $t = W_{max}$ JUSQU'A 1 FAIRE

$\forall (dx, dy, \alpha, \psi), \mathbf{V}_t[dx, dy, \alpha, \psi] = 0$

$\mathbf{S} = t$

POUR tout $\mathbf{u} \in \{\vec{C}_I^k\}_{k=1, N}$ avec $\mathbf{u}_w < t$ FAIRE

POUR $pos = 1, \kappa$ FAIRE

Soit $\mathbf{v}_{pos} = \mathcal{I}_u[pos]$

$(\mathbf{dx}, \mathbf{dy}, \alpha, \psi)$ = résoudre pour $(\mathbf{p}_{iu}, \mathbf{p}_{ju}, \mathbf{p}_{iv}, \mathbf{p}_{jv})$

$\mathbf{V}_t[dx, dy, \alpha, \psi] = \mathbf{V}_t[dx, dy, \alpha, \psi] + \Delta V$

FIN

FIN

$\{dx^s, dy^s, \alpha^s, \psi^s\} = \mathbf{S} - \max(\mathbf{V}_t[dx, dy, \alpha, \psi])$

$\mathbf{V}_t[\{dx^s, dy^s, \alpha^s, \psi^s\}] = \mathbf{V}_t[\{dx^s, dy^s, \alpha^s, \psi^s\}] \times \mathbf{K}$

$\mathbf{V}_{t+1} = \mathbf{V}_t$

Stopper si un seul pic apparaît dans l'espace de vote.

FIN

La contribution ΔV d'un vote pour une transformation donnée (dx, dy, α, ψ) entre 2 primitives u et v est calculée en tenant compte des considérations ci-dessous. Soit $\Delta V = a_1 \times a_2 \times a_3 \times a_4$, où les paramètres a_1, a_2, a_3, a_4 sont ajustés comme suit :

1. Un correspondant potentiel v_{pos} aura une plus grande contribution s'il se trouve en début de liste de préférence et symétriquement en u . a_1 est alors inversement proportionnel à une fonction des positions pos_v (resp. pos_u) de la primitive v_{pos_v} (resp. u_{pos_u}). Choisissons⁴ :

$$a_1 = \frac{1}{\sqrt{pos_v \times pos_u}} ; a_1 \in \left\{ \frac{1}{\kappa}, \dots, 1 \right\} \\ / \sup(pos_u, pos_v) \leq \kappa < \kappa_{max}$$

2. A un tour de vote donné, les primitives de catégorie inférieure ou égale à t ($t < W_{max}$ est l'itération courante du processus de vote) sont autorisées à s'associer à d'autres primitives de la seconde image. Au tour $i, t = i$ et les votants doivent avoir une priorité au moins de i . Les primitives de priorité supérieure vont voter avec plus de poids. a_2 est alors proportionnel à la fiabilité de la primitive :

$$a_2 = \frac{1}{w_u}; a_2 = \left\{ \frac{1}{W_{max}}, \dots, 1 \right\}$$

3. Les primitives les plus similaires (distance faible) votent avec un poids plus grand :

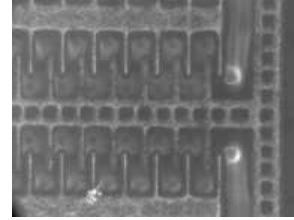
$$a_3 = (1 - \text{dist}(\mathbf{u}, \mathbf{v})); a_3 \in [0 \dots 1]$$

4. Enfin, puisque tous les $a_i, i = 1, 3$ appartiennent à l'intervalle $[0, 1]$, nous pouvons utiliser un facteur zoom a_4 afin d'obtenir des incréments entiers.

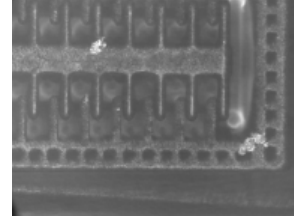
⁴Remarque : d'autres fonctions symétriques ou non de u et v peuvent être testées en regard de l'application.

6 Résultats

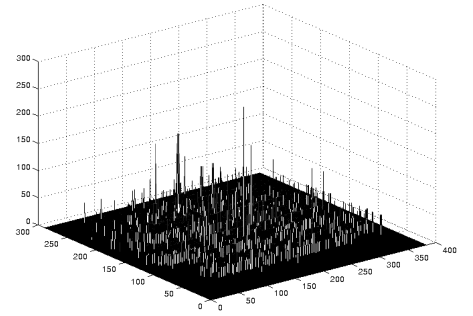
Nous avons comparé notre approche avec 2 techniques classiques : la première est basée sur une corrélation de phases (utilisant la TFR), la seconde sur une minimisation de distance entre intensités (méthode d'optimisation de Levenberg-Marquardt). Aucune n'a donné les résultats escomptés en particulier sur les images contenant de nombreuses répétitions. Nous avons aussi testé une démonstration web en ligne basée sur une approche différentielle [8]. Là encore, la transformation obtenue n'est pas correcte (voir FIG. 7.f). Nous avons sélectionné ci-dessous des résultats obtenus avec des images particulièrement difficiles à traiter en raison des variations de contraste et des nombreuses répétitions créant des ambiguïtés.



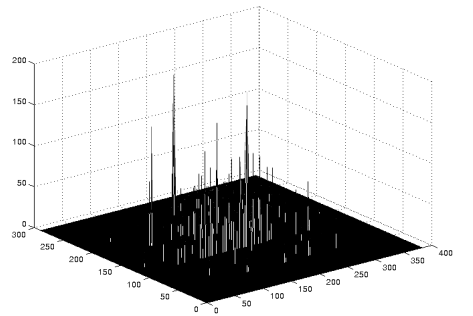
(a)



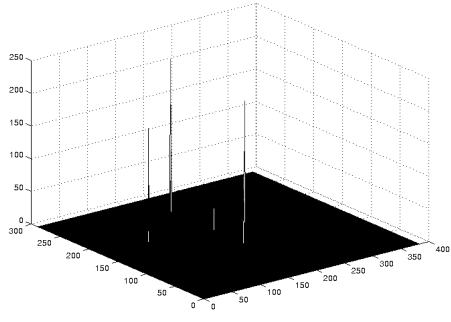
(b)



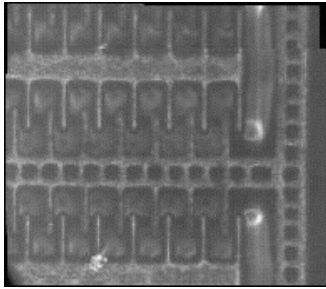
(c)



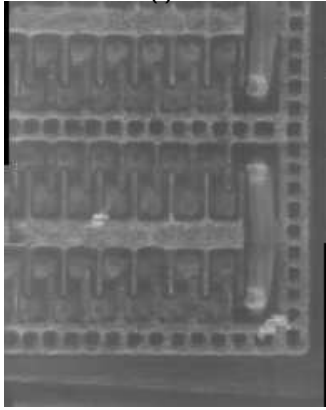
(d)



(e)

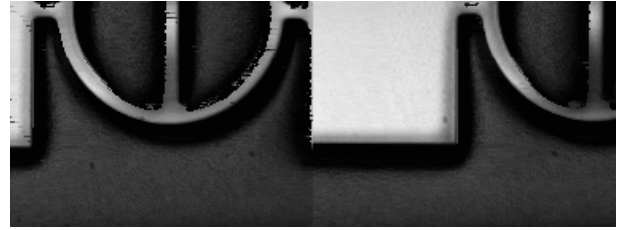


(f)

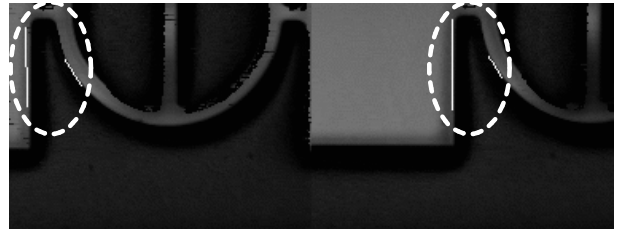


(g)

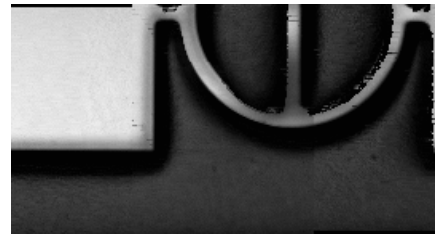
FIG. 7: (a), (b) Images obtenues par un microscope électronique à balayage de franges qui correspondent à des profils partiels d'un peigne électrostatique en cuivre d'un micro gyromètre. (c) à (e) Les espaces de vote en translation pour 3 tours de vote. (f) Résultats obtenus en utilisant une approche classique différentielle basée sur une minimisation de distance d'intensités (utilisant la méthode Levenberg-Marquardt pour l'optimisation). En raison de la présence de motifs répétitifs dans les images, la transformation obtenue n'est pas correcte. (g) La transformation correcte estimée (translation de -98, 3, rotation = 0, échelle = 1) malgré les motifs répétitifs. Ce résultat est très satisfaisant pour le recalage d'images de micro-dispositifs obtenues par un microscope électronique dont le champ est limité.



(a)



(b)



(c)

FIG. 8: (a) Images à recaler, mises côte à côte, prises par un microscope électronique. Elles correspondent aux 2 profils partiels d'un anneau en aluminium. (b) Exemple de couples se préférant mutuellement. La dynamique des images a été réduite de 2 afin de mettre en évidence les couples choisis, surlignés en blanc et entourés en pointillé. (c) Raccord correct obtenu : ici une translation de (-2, 78) entre les deux images.



(a)



(b)



(c)

FIG. 9: (a) image 1. (b) image 2. (c) Raccord obtenu : la dynamique de l'image 1 a été réduite pour transparence. Après transformation, elle est superposée sur l'image 2. La transformation trouvée et effectuée est : $dx = 5$, $dy = 3$, $teta = 9$, échelle = 1, 0096. Le recalage est satisfaisant pour les besoins de commande dans ce type d'applications.

7 Conclusion

Nous avons proposé ici une technique de mise en correspondance entre images, illustrée par une application au recalage d'image. Cette technique est indépendante des variations de contraste et ne nécessite aucun *a priori* ni aucune estimée de la transformation recherchée. L'approche permet de résoudre quelques problèmes spécifiques au système d'acquisition et aux images dans cette application, mais que l'on retrouve souvent par ailleurs : les variations de luminosité et les motifs répétitifs. Nous avons donc choisi des primitives particulières ayant des propriétés d'invariance et de robustesse intéressantes : les lignes de niveaux. Nous avons ensuite mis en place un processus d'extraction de ces lignes indépendant de la valeur des niveaux de gris et exploitant seulement leur relation d'ordre. Le recalage en lui-même est réalisé selon une stratégie efficace de mise en correspondance basée sur une vote à plusieurs tours. Il consiste à doter chaque primitive d'une liste de préférence, i.e. une liste de correspondants potentiels, les plus similaires selon une métrique et des mesures spécifiques adaptées. La transformation gagnante cumule le maximum de votes. La multiplication des tours a pour but de prendre en compte les maxima locaux – correspondant à la réponse de motifs répétés – pouvant apparaître dans l'espace de vote. Les primitives sont alors invitées à voter en fonction de leur fiabilité à chaque phase des élections. Les résultats obtenus, tant sur des images en microscopie que sur des séquences "parking", sont conformes à la réalité des transformations (lorsqu'elle est connue). Ceci nous permet d'envisager la prochaine étape : traiter des transformations plus complètes incluant par exemple la perspective.

Références

- [1] P. Anandan. Video Registration and Motion Estimation : A Retrospective. Conference on Recent Advances in 3-D Digital Imaging and Modeling, Montreal, Canada, Juin 2001.
- [2] H.S. Baird. Model-Based Image Matching Using Location. MIT Press, 1985.
- [3] D.H. Ballard. Generalizing the Hough Transform to Detect Arbitrary Shapes. Pattern Recognition journal, vol. 13, n°2, 1981, pp 111-122.
- [4] D.I. Barnea, H.F. Silverman. A Class of Algorithms for Fast Digital Image Registration. IEEE Trans. Computers, vol. 21, n°2, feb. 1972, pp 179-186.
- [5] L.G. Brown. A survey of image registration techniques. ACM computing surveys, 24(4), 325–376, 1992
- [6] S. Bouchafa. Détection de mouvement insensible aux variations de contraste. Application à la détection de comportements anormaux dans le métro par traitement d'images. Thèse de doctorat de l'université Paris 6, déc. 1998.

- [7] V. Caselles, B. Coll, J. Morel. Topographic maps & local contrast change in natural images. IJCV, sept. 1999
- [8] F. Dmitry, C. Kenney, B.S. Manjunath, L.M.G. Fonseca. On line registration demo. <http://nayana.ece.ucsb.edu/registration/>, University of California Santa Barbara, 2002.
- [9] J. Froment. A compact and multiscale image model based on level sets. *Scape Space Theories in Computer Vision*, pp 152-163. *Lecture Notes in Computer Science* 1682, 1999. *Proc. of Sec. Int. Conf. Scale-Space '99*.
- [10] A. Goshtasby, G.C. Stockman, C.V. Page. A region-based approach to digital image registration with sub-pixel accuracy. *IEEE Trans. Geosci. Remote sensing* 24, n03, 1986, pp 390-399.
- [11] F. Guichard, S. Bouchafa, D. Aubert. A change detector based on level sets. *ISMM2000*, Palo Alto, june 2000.
- [12] Y. Le Guilloux. A matching algorithm for horizontal motion, application to tracking. *IEEE Proc. 8th Inter. Conf. on Pattern Recognition*, Paris, oct. 1986.
- [13] D. Gusfield, R.W. Irwing. *The stable marriage problem - structure and analysis*, MIT Press, 1989.
- [14] S. Khuller, S. G. Mitchell, V. V. Vazirani. On-Line Algorithms for Weighted Bipartite Matching and Stable Marriages. *Theoretical Computer Science* 127(2), pp 255-267, 1994.
- [15] P. Kuhl, C. Giardina. Elliptic Fourier Features of a Closed Contour, *CGIP*, vol. 18, 1982, pp 236-258.
- [16] H. Lester, S. R. Arridge. A survey of hierarchical non-linear medical image registration. *Pattern Recognition* 32(1), 1999. Special issue on Image Registration. In press.
- [17] Z.N. Li, Z. Tauber and M.S. Drew. Locale-based object search under illumination change using chromaticity voting and elastic correlation. *IEEE Conf. on Multimedia and Expo*, 2000.
- [18] J.B.A. Maintz, M.A. Viergever. *A Survey of medical Image Registration. Medical Image Analysis*, Oxford University Press, Volume 2, 1998.
- [19] G. G. Medioni, R. Nevatia. Matching Images Using Linear Features. *PAMI*, vol. 6, n°6, nov. 1984, pp 675-685.
- [20] P. Monasse, F. Guichard. Fast computation of a contrast-invariant image representation. *IEEE Trans. on Image Proc.*, 9(5), 860–872.
- [21] H. Moravec. *Rover Visual Obstacle Avoidance*. IJ-CAI, Vancouver, 1981, pp 785-790.
- [22] L. Mundy and A. Zisserman. *Geometric Invariance in Computer Vision*, MIT Press, Cambridge, 1992.
- [23] M. L. Nack. Rectification and registration of digital images and the effect of cloud detection. *Proc. Machine Processing of Remotely sensed data*, 1977, pp 12-23.
- [24] N. Paragios, R. Deriche. A PDE-based level set approach for detection and tracking of moving objects. *Proc. of ICCV'98*, pp 1139-1145, 1998.
- [25] C.A. Pelizzari, G.T.Y. Chen, D.R. Spelbring, R.R. Weichselbaum, C.T. Chen. Accurate three dimensional registration of CT, PET and/or MR images of the brain. *J. of Computer Assisted Tomography*, 13, january/February 1989, pp 20-26.
- [26] G.P. Penney, J. Weese, J.A. Little, P. Desmedt, D.L.G. Hill, D. J. Hawkes. A comparison of similarity measures for use in 2D/3D medical image registration. *IEEE trans. Med. Image* 17(4), pp 586-595, 1998.
- [27] B.S. Reddy, B.N. Chatterji. An FFT-based technique for translation, rotation and scale-invariant image registration. *IEEE Trans. on Image Proc.*, 5(8), 1996.
- [28] A. Rosenfeld, A.C. Kak. *Digital Picture Processing*. Academic Press, 1982.
- [29] L.G. Shapiro, R.M. Haralick. Matching relational structures using discrete relaxation. *Syntactic and Structural Pattern Analysis Theory and Applications*, World Scientific Pub. Co., Teaneck, NJ, 1990.
- [30] G.C. Stockman, S. Kopstein, S. Benett. Matching Images to Models for Registration and Object Detection via Clustering. *PAMI*, vol. 4, n°3, may 1982, pp 229-241.
- [31] R. Szeliski. Image mosaicing for tele-reality applications. In *Proc. of IEEE Workshop on applications of computer vision*, 1994.
- [32] K. Zemirli, G. Seetharaman, B. Zavidovique. Stable matching or selective junction points grouping. *IEEE JCIS 2000*, Atlantic city N.J, feb. 2000