

Suivi de Coeur Battant : Tracking 2D et 3D utilisant la texture

A. Noce¹

J. Triboulet

P. Poignet

¹ Laboratoire LIRMM, UMR CNRS 5506

Université Montpellier II
161 rue Ada, 34392 Montpellier cedex 5

{noce, triboulet, poignet}@lirmm.fr

Résumé

Dans cet article, nous nous intéressons au suivi d'un organe en mouvement, un cœur battant en imagerie endoscopique et à l'évaluation de ses mouvements dans l'image et dans l'espace. La tâche est complexe, car le cœur est un objet déformable et la nature de sa surface (réflexivité, forte homogénéité de certaines zones) pose plusieurs problèmes en terme de suivi. De plus l'aspect médical suppose de trouver une méthode à la fois robuste et précise. Nous proposons de faire un compromis entre les méthodes de pattern matching (précises) et le suivi de caractéristiques liées à la textures (plus robuste) afin d'obtenir des performances intéressantes.

Mots clefs

Suivi visuel, cœur battant, imagerie médicale, robotique médicale, texture, ESM.

1 Introduction

Lors d'une opération sur le cœur, typiquement un *pontage coronarien*, le chirurgien est amené à ouvrir la cage thoracique du patient et à installer un dispositif de circulation extra corporelle (CEC) afin de stopper l'organe le temps de l'opération. Cette façon de procéder, bien qu'éprouvée, présente des risques certains pour le malade. Ces risques peuvent être liés au traumatismes de la sternotomie, qui est une intervention très lourde. Ils peuvent également être liés au dispositif de CEC. Enfin la durée de l'opération (au moins 1 à 2 heures, parfois plus) est en elle même un facteur aggravant.

Une façon de contourner ces problèmes serait d'utiliser les techniques de *chirurgie mini-invasive* et d'opérer le cœur sans l'arrêter. Or, si ces méthodes existent, elle ne peuvent être associées à l'heure actuelle au sein d'une même opération. Pour pallier ce manque, nous travaillons au développement d'un robot médical associant les fonctions de manipulation des outils chirurgicaux (téléopération) et de compensation des mouvements cardiaques. C'est dans cette seconde phase qu'intervient la composante "vision" : nous souhaitons

évaluer les déplacements du cœur à travers les images provenant de la caméra endoscopique, pour fournir cette information à la partie commande du robot.

Ce papier se divise en trois parties. La première définit l'information de texture et montre comment nous l'utilisons. La seconde présente les algorithmes utilisés pour le suivi. Enfin la dernière partie expose les résultats et les difficultés rencontrées dans la reconstruction 3D de la position du cœur.

2 Principe et Utilité de l'Information de Texture

La texture est un concept naturellement intuitif, mais difficile à définir formellement. En première approximation, on pourrait définir la texture d'un objet comme l'impact visuel de la nature de sa surface. Cela se traduit dans l'image par une uniformité visuelle, liée à une répétition d'un motif ou à un "grain" particulier au niveau de la région d'intérêt (ROI).

En pratique on quantifie cette information sous forme d'attribut, statistiques pour la plupart (moyenne, moments, etc.) que l'on peut généralement lier à des caractéristiques visuelle comme la granularité ou le contraste. Certains attributs décrivent la régularité dans la répartition des pixels (attributs basés sur les matrices de cooccurrences, les matrices de longueurs des plages) ou le nombre de composantes dans l'image pour les approches multiéchelles et spectrales. En utilisant ce formalisme, raisonner sur la texture d'une région d'intérêt revient à passer de l'espace image à l'espace des attributs, comme on peut le faire en classification automatique par exemple. La Figure 5 illustre la constitution du vecteur caractéristique de la ROI, le vecteur de texture :

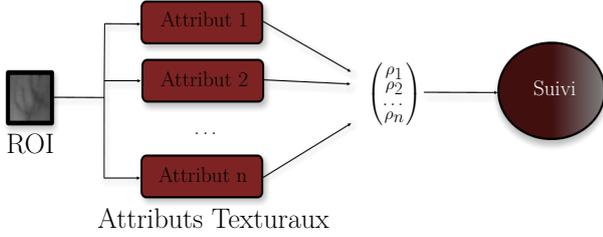


Figure 1 – Constitution du vecteur de texture

De nombreux attributs sont envisageables, nous avons mené une étude comparative de la littérature afin de limiter la redondance d'information par *Analyse en Composantes Principales* (ACP) et *Analyse Discriminante*, puis nous avons sélectionné les attributs par *Analyse de l'Entropie* afin de concentrer le maximum d'informations dans le minimum d'attributs. En effet, notre objectif est de viser le temps réel, contrairement aux systèmes experts médicaux qui utilisent habituellement l'information de texture à des fins de diagnostics, et il fallait donc limiter au maximum le nombre d'attributs à calculer. Cette étude a été détaillée dans [?], et a aboutit à la sélection de 8 attributs répertoriés dans le tableau 1 :

Tableau 1 – Attributs Textureux Sélectionnés

Méthode	Attribut
Matrice de Cooccurrence	Energy
	Contrast
	Cluster Shade
	Cluster Prominence
Matrice des Longueurs de Plages	Non-Uniformity
	Short Low Grey
	Level Run Emphasis
Statistiques	Skewness
	Kurtosis

Toujours dans un souci d'efficacité, les différentes régions d'intérêt sont comparées par des métriques dans l'espace des textures. Des tests sur des séquences d'images ont montré qu'une métrique euclidienne est suffisante pour comparer deux régions d'intérêt, la représentation par attributs textureux étant par nature plutôt robuste. En posant \mathcal{I} l'image de $n \times m$ pixels, \mathbf{p} et \mathbf{p}^* les pixels de deux ROI ($\mathcal{I}(\mathbf{p})$ représente alors les valeurs de niveaux de gris de la première ROI) et \mathcal{T} la transformation qui associe son vecteur attribut à une ROI, notre *distance entre textures* d_T s'écrit alors :

$$\begin{aligned}
 d_T(\mathcal{I}(\mathbf{p}), \mathcal{I}(\mathbf{p}^*)) &= \frac{1}{8} \|\mathcal{T}(\mathcal{I}(\mathbf{p})) - \mathcal{T}(\mathcal{I}(\mathbf{p}^*))\|_2 \quad (1) \\
 &= \frac{1}{8} \sum_{i=1}^8 (\rho_i - \rho_i^*)^2
 \end{aligned}$$

avec ρ_i et ρ_i^* composantes des vecteurs de texture associés aux deux ROI. Il est intéressant de noter que les attributs prennent des valeurs entre 0 et 1, tout comme la distance d_t . C'est une bonne chose en terme de calculs, ainsi que pour la fusion avec la distance entre motifs décrite en section 3.2.

3 Suivi du cœur

3.1 Background

Les méthodes de suivi d'objets dans l'image reposent généralement sur le choix d'un modèle pour l'objet à suivre, d'un modèle pour les déplacements de cet objet et sur des hypothèses concernant la nature de l'image. Dans la suite, nous considérons les cas :

- où l'objet suivi est un point ou plan à la surface du cœur,
- où le déplacement est une translation dans le cas du point, ou une homographie dans le cas d'un plan,
- où la séquence d'images est "consistante".

Cette dernière hypothèse signifie qu'on peut retrouver l'objet dans deux images consécutives, c.à.d trouver une transformation w (qui applique l'opérateur \mathbf{G} aux pixels) entre l'image \mathcal{I} et la suivante \mathcal{I}^* telle que :

$$\mathcal{I}^*(\mathbf{p}^*) = \mathcal{I}(w(\mathbf{G})(\mathbf{p}^*)) \quad (2)$$

avec \mathbf{p} et \mathbf{p}^* les pixels de l'objet dans les deux images.

3.2 Suivi de point

Le principe de suivi de point est facile à visualiser. En pratique on dispose d'un patch et on cherche sa position dans la nouvelle image. Une manière classique de procéder est de trouver une mesure de similarité comme la corrélation et de chercher le déplacement qui maximise (ou minimise suivant les cas) la fonction de coût ainsi définie :

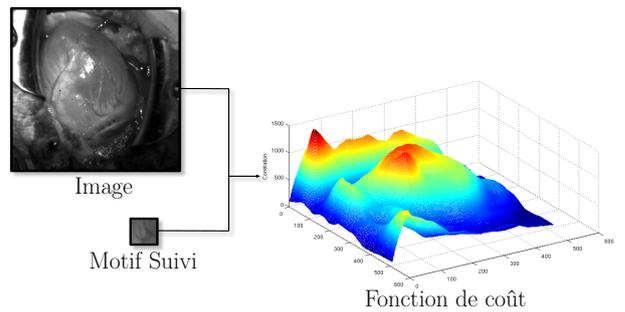


Figure 2 – Principe du suivi de point.

L'hypothèse de petits déplacements permet alors de limiter la zone de recherche et d'obtenir des performances intéressantes. Cependant le point faible de ces approches par mesure de similarité est une assez forte sensibilité au bruit et aux déformations du motif. L'idée que nous avons développée est alors d'utiliser la

distance de texture d_T définie plus haut, qui est moins sensible aux bruit et déformations (mais un tracking basé sur ce seul critère est plus approximatif) et de la combiner avec la distance entre motifs d (généralement par *Sum of Squared Differences*, SSD) dans une distance composite \mathcal{D} :

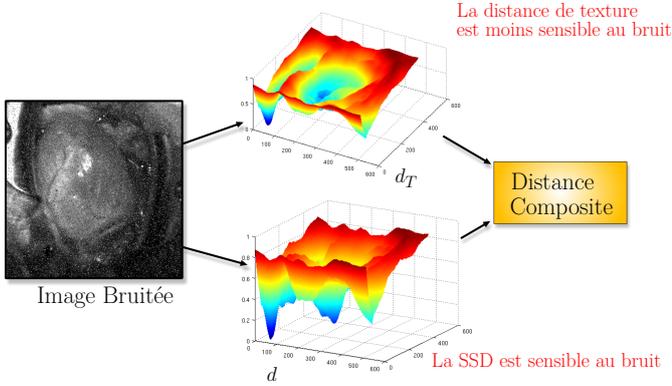


Figure 3 – Principe de la distance composite.

Le calcul de \mathcal{D} est une somme pondérée des deux distances décrites précédemment. Avec les notations de la section 3.1 on peut écrire :

$$\mathcal{D}(\mathcal{I}(\mathbf{p}), \mathcal{I}^*(\mathbf{p}^*)) = \sigma \cdot d(\mathcal{I}(\mathbf{p}), \mathcal{I}^*(\mathbf{p}^*)) + \gamma \cdot d_T(\mathcal{I}(\mathbf{p}), \mathcal{I}^*(\mathbf{p}^*)) \quad (3)$$

avec

$$\sigma + \gamma = 1 \quad (4)$$

pour garder \mathcal{D} dans $[0, 1]$.

La valeur σ représente alors l'influence de l'approche par comparaison de motifs et γ celle de l'analyse de texture. Ces paramètres peuvent être fixés en fonction de l'application visée, ou variables. Une stratégie offrant de bons résultats est d'évaluer d'abord la similitude des motifs et de favoriser cette approche si les images sont très proches, ou de favoriser l'analyse de texture sinon. On définit alors un σ fonction de pondération b :

$$b : [0, 1] \longrightarrow [0, 1] \\ d(\mathcal{I}(\mathbf{p}), \mathcal{I}^*(\mathbf{p}^*)) \longmapsto \sigma(\mathcal{I}(\mathbf{p}), \mathcal{I}^*(\mathbf{p}^*)) \quad (5)$$

décroissante pour que $\hat{\sigma}$ soit maximal lorsque la similarité est maximale. Les fonctions suivantes offrent de bons résultats, même si d'autres formules sont possibles :

$$x \longmapsto 1 - x \quad (6)$$

$$x \longmapsto \exp \frac{-x}{1 - dx} \quad (7)$$

Cette approche a été évaluée sur des séquences naturelles de cœur battant et comparée à des suivis à base de corrélation, SSD et flot optique. Notre approche se montre plus robuste, tout en ayant une bonne précision (appréciation visuelle) dans le suivi de la position de l'organe :

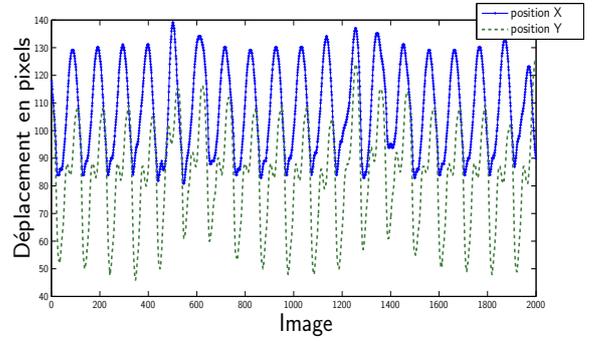


Figure 4 – Suivi du cœur dans l'image.

3.3 Suivi de Plan

L'idée est ici de suivre une portion plane à la surface du cœur. La planéité est évidemment un hypothèse, plus ou moins vérifiée suivant la région du cœur considérée et la taille de la ROI. L'algorithme que nous utilisons pour le suivi est l'algorithme ESM de [1]. Cet algorithme se base sur la distance SSD :

$$d_{SSD}(\mathcal{I}(\mathbf{p}), \mathcal{I}^*(\mathbf{p}^*)) = \|\mathcal{I}(\mathbf{p}) - \mathcal{I}^*(\mathbf{p}^*)\|_2 \quad (8)$$

$$= \|\mathcal{I}(\mathbf{p}) - \mathcal{I}^*(w(\mathbf{G})(\mathbf{p}))\|_2 \quad (9)$$

L'algorithme ESM permet alors de trouver itérativement les paramètres de \mathbf{G} qui minimisent (9) avec un bon taux de convergence. L'existence de cette solution étant assurée par l'hypothèse de consistance introduite en section 3.1.

Cette approche a été utilisé par [5] pour faire du suivi sur cœur battant en utilisant des marqueurs passifs disposés sur le cœur :



Figure 5 – Patches artificiels sur le cœur [5].

A partir de l'homographie obtenue par le suivi et les paramètres intrinsèques \mathbf{K} de la caméra supposés connus, on peut remonter à l'information de position (voir annexes de [6] et [2]).

Notre objectif est de réitérer la même opération sans utiliser de marqueur sur la surface du cœur. Pour cela nous avons repris l'algorithme ESM en limitant les variations de l'homographie pour éviter à l'algorithme de diverger et en choisissant des zones bien texturées en utilisant un détecteur de Harris. On obtient ainsi un suivi des mouvements cardiaques :

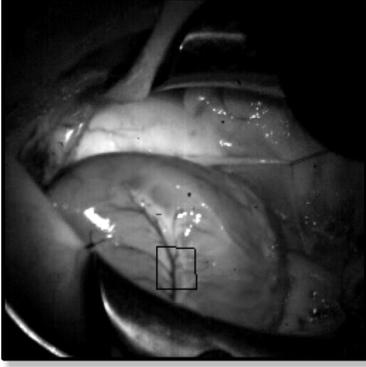


Figure 6 – Suivi d'homographie sur le cœur.

4 Du suivi vers le mouvement

4.1 Mouvement dans le plan

Si on considère un (ou des) points à la surface du cœur, on peut décomposer son mouvement 3D en un mouvement parallèle au plan image, plus une composante “ z ” qui est la distance point-caméra. Si on suppose z connu, on peut reconstituer les composantes restantes (notées x et y) du mouvement de l'objet très facilement à partir des coordonnées image u et v .

$$\begin{pmatrix} \Delta x \\ \Delta y \\ 1 \end{pmatrix} \propto z \mathbf{K}^{-1} \begin{pmatrix} \Delta u \\ \Delta v \\ 1 \end{pmatrix} \quad (10)$$

avec \mathbf{K} les paramètres intrinsèques de la caméra.

La question de la mesure de z est ouverte. On pourrait par exemple utiliser un laser un capteur de contact. Diverses solutions sont actuellement étudiées pour répondre à cette problématique.

Enfin il faut noter que des solutions purement automatiques, basées sur un boucle d'asservissement sont présentées dans [3, 4] pour l'asservissement 2D. Il s'agit alors de piloter le robot pour éliminer l'erreur dans l'image. Ces solutions impliquent une boucle d'asservissement très rapide (500MHz dans les deux cas) et ont le défaut de ne pas fournir d'information sur le mouvement de l'organe.

4.2 Mouvement 3D

On cherche à reconstruire l'information de profondeur à partir de la seule information visuelle. Nous proposons cette fois d'utiliser un suivi par homographie. Lorsque la dimension du patch est connue, comme dans [5], il est possible d'évaluer la distance objet-caméra. Notre objectif est de nous passer de patch, et de suivre une petite portion de cœur. Pour cela nous utilisons les relations d'homographie, donc la forme la plus générale est donnée dans [2] :

$$\mathbf{H}_{33} = \mathbf{K} \left(\mathbf{R}_{33} - \frac{\mathbf{T}_{31} \mathbf{n}_{31}^t}{z} \right) \mathbf{K}^{-1} \quad (11)$$

avec \mathbf{H}_{33} la matrice de l'homothétie, \mathbf{R}_{33} la matrice de la rotation, \mathbf{T}_{31} la matrice de la translation et \mathbf{n}_{31} le vecteur normal au plan. Nous supposons aussi connue la position initiale, ou du moins une approximation de cette position.

Cependant, les résultats actuels ne sont pas encore satisfaisants, et ce pour plusieurs raisons :

- le suivi n'est pas encore assez précis, ce qui amène à des mesures parfois incohérentes,
- nous supposons toujours que le plan initial est parallèle à la caméra, ce qui limite les zones que l'on peut suivre,
- la reconstruction du mouvement par l'homothétie a tendance à “dévier” peu à peu, comme le montre la Figure 7, où les cycles cardiaques ne se superposent pas.

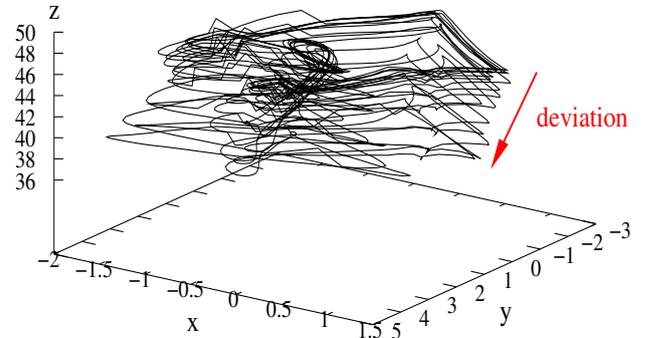


Figure 7 – Exemple de trajectoire - et déviation.

5 Conclusion

L'appréciation d'un mouvement dans une image est un exercice difficile, surtout avec une seule caméra. Le travail que nous avons effectué au niveau du suivi nous permet de récupérer une première information de manière robuste et avec une précision suffisante pour pouvoir envisager la phase de reconstruction du mouvement. Les divers problèmes présentés dans la section 4 montrent cependant la difficulté d'obtenir la précision recherchée. Nous évaluons actuellement d'autres possibilités, comme la stéréovision sur laquelle nous

travaillons ou des solutions d'asservissement 2D associées à une mesure de profondeur non basée sur la vision, comme cela a été évoqué dans la section 4.1, tout en continuant notre effort pour résoudre les problèmes soulevés.

Références

- [1] S. Benhimane and E. Malis. Real-time image-based tracking of planes using efficient second-order minimization. In *Proc. IEEE IROS*, pages 943–948, September 2004.
- [2] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2004.
- [3] Y. Nakamura, K. Kishi, and H. Kawakami. Heart-beat synchronization for robotic cardiac surgery. In *Proc. IEEE ICRA*, pages 2014–2019, Seoul, Korea, Mai 2001.
- [4] Ginhoux R., Gangloff J.A., de Mathelin M.F., Soller L., Sanchez M.M.A., and J. Marescaux. Beating heart tracking in robotic surgery using 500 hz visual servoing, model predictive control and an adaptive observer. In *Proc. IEEE OCRA'04*, pages 274–279, New Orleans, LA, USA, Feb 2004.
- [5] M. Sauvée, P. Poignet, J. Triboulet, E. Dombre, E. Malis, and R. Demaria. 3d heart motion estimation using endoscopic monocular vision system. In *IFAC MCBMS'06*, Reims, France, sep.
- [6] Z. Zhang. A flexible new technique for camera calibration. Technical report, Microsoft Research, 1998.