

Conception d'un algorithme robuste pour une mise en correspondance précise de coupes s eries

Application en histologie de fibres musculaires

Patrice Buche et David Mauron

INA-Informatique/INRA BIA

16, rue Claude Bernard 75231 Paris C dex 5, E-mail : Patrice.Buche@inapg.inra.fr

RESUME

Ce papier pr sente une m thode de mise en correspondance d'images de coupes s eries de fibres musculaires. Au vu de l'application, l'accent a  t  mis sur la pr cision de la technique de superposition mise en oeuvre et sur la qualit  des r sultats obtenus par la m thode semi-automatique d'obtention de points homologues entre les images   recal r.

ABSTRACT

In this paper, we propose a method to superimpose serial cuttings sets of muscle fibres. A special effort has been done to enhance the accuracy of the superimposition technic and to improve the results quality of the semi-automatic method employed to obtain homologous points between the images to match.

1. Le logiciel RACINE r alise la mise en correspondance de coupes s eries.

Il est utilis  dans plusieurs laboratoires de l'INRA pour caract riser de mani re quantitative des fibres musculaires de provenances vari es (porc, poulet, truite, boeuf, mouton, lapin, ...) vues en section sur un jeu de coupes s eries. Cette caract risation s'effectue en deux  tapes (cf fig. 1):

- l'individualisation des fibres est r alis e une seule fois sur une coloration de r f rence (cf fig. 2),
- les fibres sont suivies sur les autres colorations (cf fig. 2) du jeu afin de les typer et de mesurer diff rents param tres histochimiques (par estimation de la luminance moyenne dans la fibre).

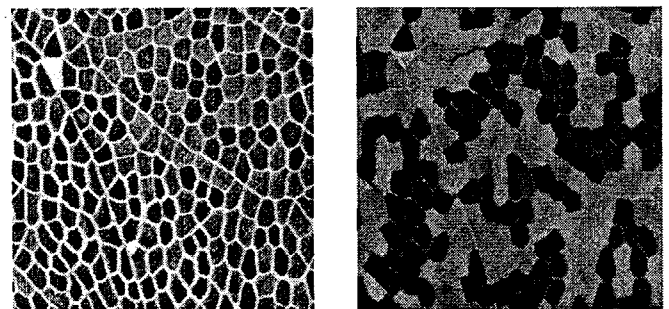


Figure 2 : Un exemple de coloration de r f rence (  gauche) et de coloration de typage (  droite)

Dans la premi re version du logiciel, nous avons pour hypoth se de d part que l'intensit  de la coloration  tait homog ne   l'int rieur de la fibre. La luminance moyenne  tait calcul e dans une fen tre de petite taille (en g n ral 7x7 pixels) centr e sur la position estim e du barycentre. Le mod le g om trique   employer ne n cessitait pas une tr s grande pr cision. Nous avons montr  dans Buche *et al.* (1991) qu'un mod le polyn mial de degr  3  tait suffisant pour r aliser le suivi des fibres malgr  les probl mes de d formation entre les coupes d'origines diverses (effet coloration, apparition de cellules entre deux coupes, ...). Dans l' tape d'individualisation des fibres, le logiciel d termine la position du barycentre de chaque fibre sur la coloration de r f rence. Dans l' tape du suivi des fibres, on estime la position du barycentre dans les autres

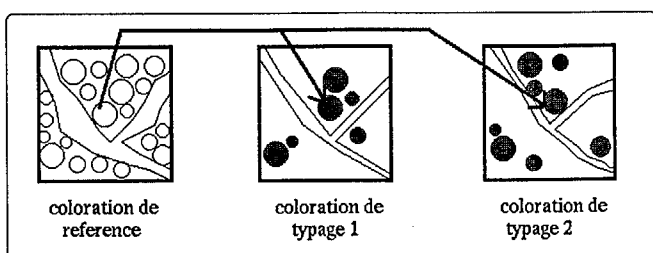


Figure 1 : Suivi sur l'ensemble des colorations du jeu de coupes d'une fibre individualis e sur la coloration de r f rence

Une premi re version du logiciel a  t  d crite dans Buche *et al.* (1991). L'utilisation en routine de cette version et l' volution des besoins nous ont amen   



colorations en utilisant le modèle polynômial. La précision obtenue avec ce modèle est de **4 à 7 pixels** autour de la position effective alors que le diamètre moyen des fibres est de 20 pixels. On était donc sur en moyenne de tomber à l'intérieur de la fibre.

Les utilisateurs nous ont alors demandé de pouvoir traiter de nouvelles colorations de typage sur lesquelles le colorant ne prend pas de manière homogène dans la fibre mais préférentiellement sur le bord. Il est alors nécessaire d'utiliser un modèle de déformation le plus précis possible afin d'éviter de lire la luminance en dehors de la fibre. Nous proposons dans la section suivante l'utilisation d'un modèle de déformation à base de courbes splines. Les points de contrôle utilisés pour calculer les paramètres du modèle de déformation sont obtenus par un algorithme de recherche de structures identiques dans les deux images à apparier. Un processus de relaxation stochastique inspiré de Barnard *et al* (1980) permet d'évaluer itérativement la pertinence des points de contrôle apparierés. Pour cela, on analyse la cohérence de ces appariements avec les appariements de forte probabilité calculés dans le voisinage spatial. Dans la première version de notre algorithme, nous ne pouvions pas aller jusqu'à stabilisation du processus de relaxation car nous obtenions un nombre de couples de points de contrôle erronés difficiles à éliminer automatiquement. Nous proposons dans la dernière section un critère de recherche amélioré pour le processus de relaxation et une méthode de filtrage des bons couples de points plus robuste qui nous permet d'obtenir plus de points de contrôle.

2. Un modèle de déformation plus précis

Le polynôme de degré 3 n'offrant pas des degrés de liberté suffisants pour approximer des distorsions locales importantes, nous avons testé un nouveau modèle de déformation à base de courbes splines inspiré de Bookstein (1989), l'objectif étant d'obtenir une grande précision d'ajustement (au pixel près). Notons S_1 et S_2 les ensembles de points qui se correspondent sur les deux images à apparier :

$$S_1 = \{ (X_{1i}, Y_{1i}), i \in (1, \dots, n) \} \quad (1)$$

$$S_2 = \{ (X_{2i}, Y_{2i}), i \in (1, \dots, n) \} \quad (2)$$

où (X_{1i}, Y_{1i}) est le point de l'image 1 qui correspond au point (X_{2i}, Y_{2i}) dans l'image 2. Nous appelons S'_2 l'ensemble des points dans l'image 2 résultant de l'application du modèle de déformation sur les points de S_1 :

$$S'_2 = \{ (X'_{2i}, Y'_{2i}), i \in (1, \dots, n) \} \quad (3)$$

avec

$$X'_{2i} = f(X_{1i}, Y_{1i}) \quad (4)$$

$$Y'_{2i} = g(X_{1i}, Y_{1i}) \quad (5)$$

Le modèle spline plaque mince s'inspire de la physique. Il consiste à utiliser un modèle de déformation d'une plaque mince dans l'espace subissant des contraintes en plusieurs endroits. Cette modélisation a la propriété de définir la configuration de la plaque qui minimise son énergie de courbure. L'idée est alors d'utiliser ce modèle pour résoudre un problème d'interpolation bidimensionnel : on cherche à déterminer la fonction qui permet de passer de l'ensemble de points S_1 à l'ensemble S_2 en minimisant « l'énergie de courbure » qui aurait été requise si les écarts entre les points homologues de S_1 et de S_2 s'étaient exercés à la normale du plan de l'image vue comme une plaque mince plutôt que dans le plan lui-même. Les fonctions f et g s'expriment comme des transformations de type spline plaque mince. L'équation (4) s'écrit de la manière suivante :

$$X_{2i} = a_1 + a_2 X_{1i} + a_3 Y_{1i} + \sum_{j=1}^n w_j U(|(X_{1j}, Y_{1j}) - (X_{1i}, Y_{1i})|) \quad (6)$$

où $U(r) = r^2 \log r^2$ et $|(X_{1j}, Y_{1j}) - (X_{1i}, Y_{1i})|$ est la distance euclidienne entre les couples de points (X_{1i}, Y_{1i}) et (X_{1j}, Y_{1j}) .

Les coefficients de l'équation (6) sont estimés par la résolution du système :

$$(7) \begin{cases} X'_{2i} = X_{2i}, \forall i \in (1, \dots, n) \\ Y'_{2i} = Y_{2i}, \forall i \in (1, \dots, n) \\ \sum_{i=1}^n w_i = 0 \\ \sum_{i=1}^n X_{1i} w_i = 0 \\ \sum_{i=1}^n Y_{1i} w_i = 0 \end{cases}$$

La spécificité des courbes splines plaque mince nous assure (cf système d'équations (7)) que l'erreur de reconstruction aux points de contrôle (points homologues de l'image à recaler) est nulle. Nous avons vérifié expérimentalement que le modèle spline plaque mince permet d'obtenir une superposition de bien plus grande précision que le modèle polynômial (cf fig. 4 et 5).

Cependant, dans la pratique, nous constatons que le modèle spline ne remplace pas de manière systématique le modèle polynômial. En présence d'une image à recaler présentant une ou deux déformations locales par rapport à la coloration de référence, le modèle polynômial reste supérieur car plus rapide à calculer et nécessitant en moyenne moins de points de contrôle. Dans le cas de déformations locales plus importantes en nombre et en intensité, le modèle spline se révèle supérieur au modèle polynômial (cf fig. 4 et 5). Son utilisation implique la plupart du temps une saisie interactive de couples de points supplémentaires pour bien prendre en compte toutes les déformations. En conséquence, les deux modèles coexistent dans la nouvelle version de RACINE. L'utilisateur emploie le modèle spline lorsqu'il rencontre de fortes distorsions locales qui ne peuvent pas être prises en compte par le modèle polynômial.

3. Une méthode de calcul des couples de points de contrôle plus robuste et plus efficace

- L'algorithme de relaxation réévalue à chaque itération une probabilité d'appariement à chaque couple de points de contrôle potentiellement homologues. En fin d'algorithme, seuls les couples associés à une probabilité supérieure à un seuil donné sont gardés. Les probabilités calculées traduisent entre autre le fait que, si dans un voisinage donné, on constate plusieurs appariements de forte probabilité dans une direction, la probabilité d'appariement dans cette direction est renforcée. Cependant, dans le cas inverse (incompatibilité avec le voisinage), la probabilité n'est pas réduite. Nous avons donc modifié l'algorithme en défavorisant les appariements incompatibles avec le voisinage. Les différences entre les appariements improbables et probables apparaissent plus nettement. Cela nous permet d'aller jusqu'à la stabilisation du processus et d'obtenir plus de points de contrôle.

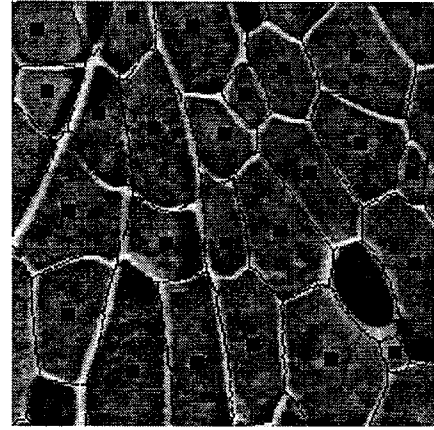


Fig. 4 : Superposition du réseau interfibre en noir (extrait sur la coloration de référence) sur une coloration de typage avec le modèle spline plaque mince

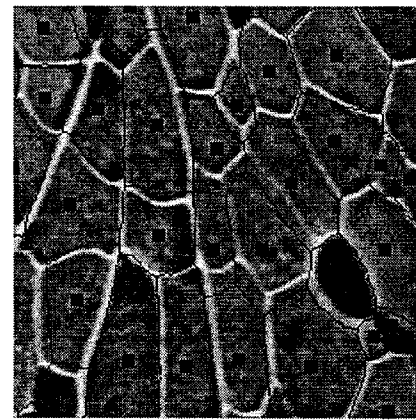


Fig. 5 : Superposition du réseau interfibre en noir (extrait sur la coloration de référence) sur une coloration de typage avec le modèle polynômial. Le résultat est moins bon que celui obtenu avec le modèle spline plaque mince (cf fig. 4) en utilisant pourtant le même ensemble de points de contrôle

- Dans la première version de l'algorithme, nous éliminions itérativement les couples de points de contrôle calculés par la méthode de relaxation ayant un résidu aberrant (différence entre la position réelle des points de contrôle et celle estimée par un modèle polynômial de degré 3). Nous utilisons pour évaluer les coefficients du polynôme la méthode des moindres carrés. Cette méthode n'est plus efficace dès qu'il y a plus de deux ou trois couples de points de contrôle erronés. Pour filtrer efficacement les couples de points de contrôle calculés lors du processus de relaxation, nous avons opté pour la méthode des **moindres valeurs médianes** inspiré de *Rosenfeld et al. (1991)*. Cette technique de **régression robuste** détermine dans un premier temps, un modèle polynômial de degré 3 minimisant la valeur **médiane** des carrés des résidus et élimine ensuite les couples dont le résidu s'écarte trop de



ce modèle (écart-type robuste). Les résultats s'avèrent satisfaisants pour notre application car cette technique reste fiable pour un taux de couples erronés pouvant aller jusqu'à 50 % du nombre total de couples de points; ce qui couvre la grande majorité des cas que nous traitons.

● Ces modifications ayant fortement diminué le risque de conserver des appariements erronés, nous avons pu augmenter le **nombre d'itérations** dans l'algorithme de relaxation (de 4 à 10). Par conséquent, le nombre de couples de points de contrôle obtenus en fin du processus de mise en correspondance s'est substantiellement accru.

4. Validation expérimentale de la méthode de mise en correspondance

Nous avons évalué le temps de traitement (cf tab 1) et le taux d'erreur (cf tab 2) de la méthode de mise en correspondance sur deux lots de données. Le premier est constitué de 20 jeux de coupes sériées de muscle de poulet (représentant au total 3763 fibres). Chaque jeu comporte une coloration de référence et une coloration de typage. Le second lot est constitué de 20 jeux de coupes sériées de muscle de porc (représentant au total 3252 fibres). On remarquera que les taux d'erreur et de points homologues ajoutés sont nettement plus faibles pour le lot de données « porc » comparativement au lot de données « poulet ». Cette différence se reflète évidemment sur les temps d'interaction nécessaire avec le système pour ajouter des points de contrôle. Cette différence s'explique par le fait que l'algorithme de recherche de points homologues se base principalement sur la présence de transitions entre îlots (groupes de fibres connexes de même couleur, cf fig. 2) . Or, les délimitations entre les fibres appartenant à un même îlot sont mieux marquées sur les colorations de typage du lot de données « porc » traité que sur celles du lot de données « poulet ». Plus de 80 % des points de contrôle utilisés pour recalibrer les images sont obtenus par la procédure automatique. Les taux de points homologues erronés sont très faibles. Ces résultats montrent que notre méthode de recherche semi-automatique de points homologues est une aide efficace pour l'opérateur. En effet, elle détermine la grande majorité des points de contrôle utilisés pour recalibrer les images avec un taux d'erreur résiduel.

	correction du recalage	
Poulet	5'58"	
Porc	1'25"	2'24"

Tableau 1 : Temps de correction interactive de la mise en correspondance pour un jeu d'images à appairer. Les deux colonnes pour le porc correspondent aux deux coupes à recalibrer.

	Taux d'erreur sur les points homologues		Taux de points homologues ajoutés	
Poulet	1,91 %		16,41 %	
Porc	0,12 %	0,63 %	2,45 %	2,35 %

Tableau 2 : Taux moyen d'erreur de la méthode de recherche de points homologues et taux moyen de points homologues ajoutés interactivement pour obtenir une bonne superposition

5. Implantation

Le logiciel est écrit en langage C (plus de 30 000 lignes). Il tourne sur station de travail UNIX sous l'interface graphique X-Window. Les images traitées sont numérisées au format 512 x 512 pixels codés sur 256 niveaux de gris.

6. Bibliographie

- Barnard, ST, Thompson WB (1980)*. Disparity analysis of images. IEEE PAMI vol PAMI-2, no 4, p 333-340.
- Buche, P, Camillerapp, J (1991)*. Serial cutting matching : an application to muscle fiber characterization. Eurographics'91, Vienna : 329-340.
- Bookstein, F (1989)*. Principal warps: thin-plates splines and the decomposition of deformations. IEEE Transactions on patterns analysis and machine intelligence, vol II, No 6.
- Rosenfeld, A, Meer, P, Doron M (1991)*. Robust regression methods for computer vision : a review. International Journal of Computer Vision, 6:1 : 59-70.