



SEGMENTATION D'IMAGES COULEUR : APPLICATION A LA QUANTIFICATION DE L'IMMUNOMARQUAGE.

**Belhomme P.¹, Elmoataz A.^{1,2}, Herlin P.³
Revenu M.¹, Bloyet D.¹, Mandard A.M.³**

¹ GREYC - URA CNRS 1526 - ISMRA 6 Boulevard Maréchal JUIN, F-14050 CAEN Cedex

² LUSAC - EIC Site Universitaire, B.P. 78, F-50130 OCTEVILLE

³ Laboratoire d'Anatomie Pathologique Centre F. BACLESSE, route de Lion/mer, F-14021 CAEN Cedex

RESUME

L'analyse de marqueurs immunohistochimiques sur des coupes histologiques est d'un grand intérêt en pathologie tumorale pour le diagnostic, l'évaluation du pronostic et la thérapeutique. Leur quantification pose cependant un problème en raison de la complexité des images liée à la structure du tissu et à la qualité de la coloration. Les facteurs majeurs qui limitent actuellement l'automatisation complète de la mesure sont la segmentation des noyaux sur les images difficiles et la détection des zones marquées par analyse couleur. Le but de ce travail consiste à proposer une stratégie générale permettant la quantification automatisée du rapport d'immunomarquage nucléaire. Un exemple d'application sur des images couleur de cancers du sein est présenté.

INTRODUCTION

L'immunohistochimie constitue un outil de première importance en histopathologie tumorale. Une large gamme d'anticorps permet au pathologiste de localiser à l'échelle tissulaire et cellulaire des marqueurs spécifiques de différenciation (récepteurs d'oestrogènes, de progestérone,...) et de prolifération (Ki67, PCNA, BrdU,...). L'évaluation quantitative de la réaction peut en outre fournir des indications précieuses quant au pronostic de certaines tumeurs et leur réponse à la thérapeutique [1]. Si l'analyse d'images offre la possibilité de quantifier ces réactions "in situ", les méthodes mises en oeuvre à l'heure actuelle ne sont pas standardisées [2]. Une automatisation complète de la mesure suppose en effet le développement de techniques de segmentation adaptées tant à l'échelle tissulaire (identification automatique des territoires tumoraux) que cellulaire (identification automatique des organites marqués). Notre travail a donc porté sur l'élaboration d'une stratégie de segmentation automatisée d'images couleur de coupes histologiques de cancers du sein permettant la quantification du rapport d'immunomarquage. Cette stratégie se base sur le concept de segmentation par régions et fait largement appel aux opérateurs issus de la morphologie mathématique. Pour tirer parti des informations couleur disponibles, nous avons été amenés à développer une extension de l'algorithme de la ligne de partage des eaux par files d'attente prioritaires [3] permettant d'aboutir à un opérateur de croissance de régions multi-critères à parcours optimal de l'image.

ABSTRACT

The analysis of immunohistochemical markers in histological sections is of great importance for diagnosis, prognosis, evaluation and tumour treatment. Nevertheless, the complexity of images, due to tissue structure and staining quality, makes their quantification uneasy. Until now, the main factors limiting a complete automation of the measurement are the segmentation of nuclei in complex images and the detection of stained areas through color analysis. The aim of this work is to propose a generalized strategy allowing an automated quantitation of the nuclear immunostaining ratio. An example of its application to breast cancer images is also presented.

MATERIEL

Un prélèvement de tissu tumoral de cancer du sein est fragmenté en coupes histologiques de 5 µm d'épaisseur. Ces coupes sont placées sur une lame de verre afin d'être mises en présence d'un marqueur immunohistochimique (c'est à dire un anti-corps spécifique de la protéine que l'on cherche à quantifier dans la tumeur) révélé à la peroxydase-diaminobenzidine. Elles subissent alors une contre coloration à l'hématoxiline de Harris. Le premier couple de réactifs révèle un colorant lié à l'anti-corps. Il entraîne une coloration brune au niveau des sites marqués positivement. Le second réactif, non-spécifique, donne une couleur bleue aux profils nucléaires. Les images, observées au microscope optique (OLYMPUS BH2), sont enregistrées sur trois canaux RGB par une caméra 3-CCD JVC KY-F30 connectée à une carte d'acquisition DataCell S2200 implantée dans une station SUN SPARC-Classic sous UNIX. Elles ont une taille de 768x576 avec une dynamique de 256 niveaux de gris par composante. Deux grossissements différents sont utilisés pour traiter le problème au niveau cellulaire puis tissulaire. Un premier (X132) permet l'étude du marquage intra-nucléaire (cf. fig. 1). Un deuxième (X33) permet de distinguer dans l'image les lobules tumoraux (cf. fig. 2) qui sont en fait des regroupements de noyaux caractérisés par une taille importante et une faible distance inter-noyaux. L'objectif de l'analyse consiste, dans un premier temps, à produire le rapport des surfaces de profils marqués positivement sur la surface totale des profils, puis de limiter cette mesure aux seuls compartiments tumoraux.



Les algorithmes développés dans ce travail ont été écrits en C standard puis intégrés sous forme de nouvelles commandes dans le logiciel de traitement d'images Visilog 4.1.4 de la société NOESIS.

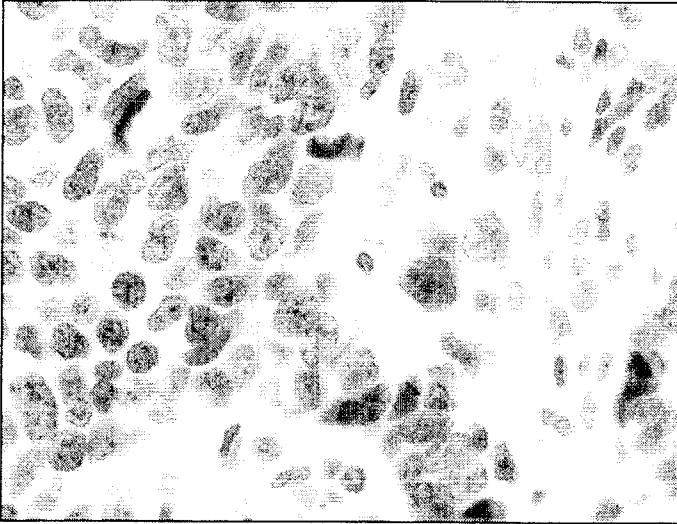


fig. 1 : coupe histologique de cancer du sein (X132)

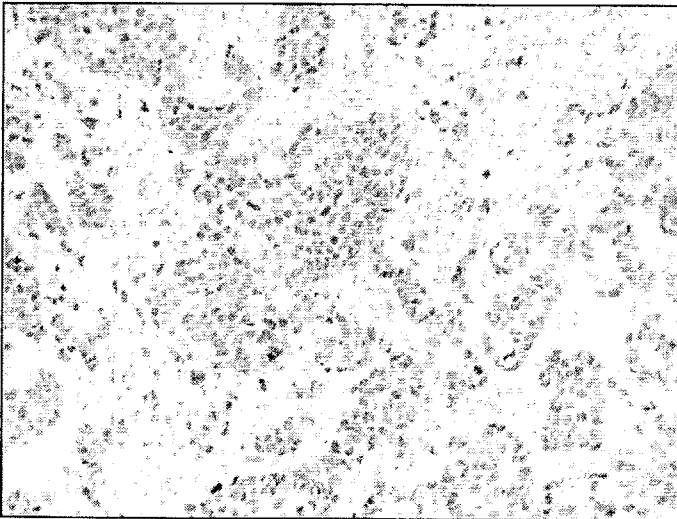


fig. 2 : coupe histologique de cancer du sein (X33)

METHODES

La recherche du rapport d'immunomarquage intra-tumoral nécessite quatre étapes d'analyse :

① Segmentation des profils nucléaires dans les images observées aux grossissements (X132) et (X33).

② Analyse couleur des régions formées dans l'image (X132) afin de détecter la présence d'un éventuel marquage immunohistochimique.

③ Recherche du masque binaire recouvrant les lobules tumoraux dans l'image à faible grossissement (X33).

④ Limitation des mesures de surfaces calculées à l'étape 2 aux lobules extraits à l'étape 3.

① Segmentation des profils nucléaires.

Les techniques de segmentation basées uniquement sur la détection de contours sont peu performantes compte tenu de la complexité des images de coupes histologiques. Celles basées

sur la détection de régions par croissance ou par fusion-division n'intègrent pas suffisamment l'information contour, pourtant de première importance dans les profils nucléaires. Nous avons donc opté dans un premier temps pour la segmentation par morphologie mathématique [4] qui est basée sur la recherche de la ligne de partage des eaux (LPE) d'une fonction potentielle. Son principe tient en deux étapes :

1- Simplification et extraction de primitives au moyen de procédures de recherche de marqueurs d'objets faisant appel à des connaissances sur leurs caractéristiques et sur les outils de traitement d'images.

2- Recherche de la LPE d'une "fonction potentielle" par immersion de son sous-graphe à partir de trous générés à l'emplacement des marqueurs. Cette immersion est simulée au moyen de files d'attente prioritaires [3]. L'algorithme résultant peut être vu alors comme un opérateur de croissance de régions à parcours optimal selon un seul critère de priorité qui est ici la valeur de la fonction en chaque point.

Dans la pratique on choisit souvent le module d'une image gradient puisque ses lignes de crêtes sont supposées coïncider avec les bords des objets à détecter. Toutefois, lorsque l'on est en présence d'images moins contrastées ou bruitées, ce seul critère se révèle insuffisant car il ne considère que des informations locales. Nous proposons donc de bâtir l'ordre de priorité des points sur des critères tenant compte de données locales sur les points et des statistiques des régions voisines qui vont les agréger. Nous présentons ici le principe général pour deux critères intégrant l'information contour et région :

Initialisation :

Labellisation de tous les marqueurs avec calcul de leurs moyennes (niveaux de gris ou couleur). Ces marqueurs constituent les régions qui devront croître.

Tout point p voisin d'un marqueur est placé dans une file prioritaire avec une priorité égale à :

$$C(p, R) = (1 - \alpha) \cdot \|\bar{I}_R - I(p)\| + \alpha \cdot \|\nabla I(p)\| \quad (1)$$

$I(p)$: valeur de l'image au point p . Cette valeur peut être vectorielle dans le cas d'une image couleur.

\bar{I}_R : moyenne de I sur la région R voisine de p .

$\|\nabla I(p)\|$: module du gradient au point p .

Ils reçoivent alors une étiquette particulière signifiant qu'ils ont été examinés.

Croissance :

Les files sont balayées par ordre croissant jusqu'à ce que l'on en trouve une non-vide. Chaque fois qu'un point p est retiré d'une file, il reçoit la même étiquette que la première région R rencontrée dans son voisinage. La moyenne de R est mise à jour, puis tous les voisins non labellisés de p (points p') sont à leur tour enfilés selon une nouvelle priorité $C(p', R)$. Lorsque toutes les files sont vides, le processus est terminé. Cet algorithme est appliqué sur les deux images (X132) et (X33). Dans le premier cas on prend en compte les informations couleur dans la formation des critères $C(p, R)$. Dans le deuxième cas on ne travaille que sur la composante

verte (composante chromatique neutre vis à vis de l'immunomarquage). Dans les deux cas les marqueurs d'objets sont obtenus par recherche des minima locaux de la composante verte lissée. L'intensité du lissage dépend de la taille moyenne des objets, donc du grossissement. Les minima dont la valeur est inférieure à un seuil caractérisant l'intensité du fond sont éliminés. Ce seuil, déterminé automatiquement [5], est utilisé pour en déduire un marqueur du fond.

② Classification des régions.

Un profil nucléaire marqué est par définition une région de l'image originelle possédant des pixels de couleur marron en nombre significatif. La proportion minimale de ces points (seuil de négativité S_{negat}) est une information définie à l'avance par l'expert pathologiste. Par opposition, un profil non-marqué correspond à une région de l'espace n'appartenant pas au fond de l'image et présentant des pixels marrons en proportion inférieure à S_{negat} . Il est donc nécessaire de distinguer, dans l'image à fort grossissement, les trois classes de points (marquée, non-marquée et fond) par analyse des composantes d'un espace couleur quelconque (RGB, HLS,...). Une étude a montré que l'utilisation des trois composantes RGB permettait une bonne discrimination des classes simplement en retenant les couples R-B ou G-B. On peut ainsi voir sur la figure 3 l'histogramme de l'image de la figure 1 dans l'espace R-B avec, en surimpression, les nuages de points correspondant à des zones sélectionnées manuellement dans chaque catégorie de pixels. Pour permettre la classification des points appartenant aux zones intermédiaires on a alors retenu comme référence l'espace de la figure 4 que l'on peut qualifier de "Zone d'intérêt générale" des catégories de pixels.

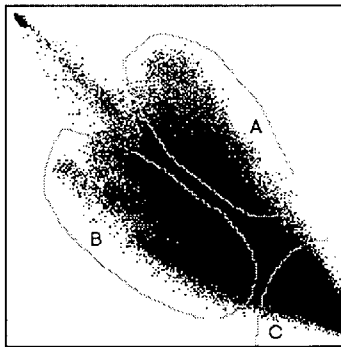


fig. 3 : Distribution des pixels dans l'espace R-B et zones d'intérêt des catégories.
A : zone des pixels "marqués".
B : zone des pixels "non-marqués".
C : zone des pixels du fond de l'image.

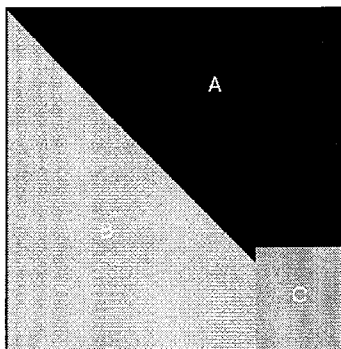


fig. 4 : Positions retenues pour les catégories de pixels dans l'espace R-B.

③ Recherche du masque des lobules.

Une fois les profils nucléaires segmentés dans l'image (X33), on construit le diagramme de Voronoï [6] à partir des germes constitués par les barycentres des régions ; on en déduit le graphe de Delaunay correspondant. Ses noeuds sont ensuite valués par la surface des objets et ses arcs par la distance euclidienne entre objets voisins. Ce graphe est segmenté par seuillage sur le poids des arcs. On obtient alors des sous-graphes traduisant les regroupements par proximité (cf. figure 5). On élimine toutes les composantes dont la surface totale est inférieure à une valeur pré-établie expérimentalement. Les régions correspondant aux noeuds restants définissent ainsi un masque binaire recouvrant les lobules tumoraux de la coupe histologique.

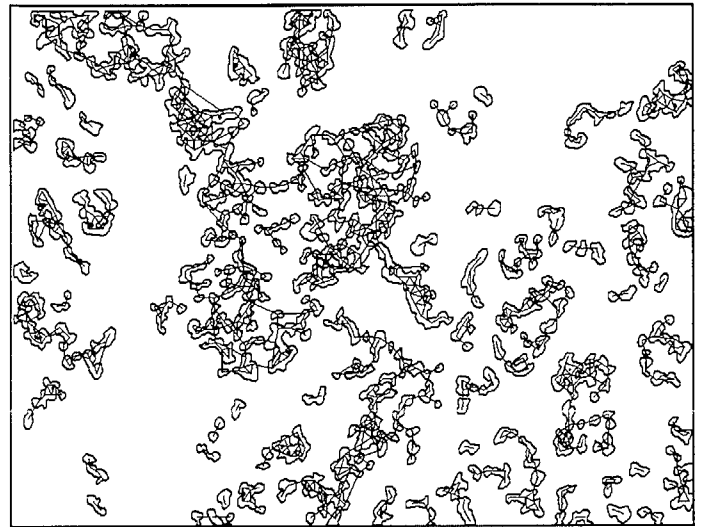


fig. 5 : Extraction des lobules par graphe de Delaunay sur l'image (X33).

④ Limitation des mesures.

Pour accéder au rapport d'immunomarquage, les surfaces calculées à l'étape 2 doivent être limitées au masque binaire des lobules. On doit donc recalcr entre elles les images observées aux deux grossissements différents afin de superposer les profils nucléaires identiques. Etant donné que l'on ne dispose pas à l'heure actuelle d'un microscope équipé d'une platine motorisée, ce recalcr doit être effectué mathématiquement. Pour diminuer le temps de traitement, nous avons adopté la technique du calcul de la distance de Hausdorff entre deux ensembles binaires de points constitués par les minima locaux de la composante verte. Les points issus du fond de l'image sont éliminés de la même manière que pour les marqueurs d'objets. L'ensemble des minima provenant de l'image à fort grossissement est obtenu après réduction de celle-ci (d'où le choix d'un multiple de 2 entre les grossissements utilisés).

RESULTATS

Les critères utilisés dans la segmentation de l'image (X132) prennent en compte l'information locale du gradient et l'information globale des niveaux couleur par l'intermédiaire d'une pondération analogue à celle de l'équation (1). La valeur du paramètre α est fixée à 0.7 et le vecteur I est composé des trois composantes RGB. La figure 6 montre le résultat de la segmentation obtenue superposée à l'image de la figure 1. La



position des marqueurs d'objets est également indiquée. Le seuil de négativité S'_{negat} retenu par l'expert pathologiste est fixé à 10 %. Le rapport d'immunomarquage brut, c'est à dire sur toute l'image (X132), est de 70 %. Après limitation de la mesure aux seuls compartiments tumoraux cette valeur passe à 81 %.

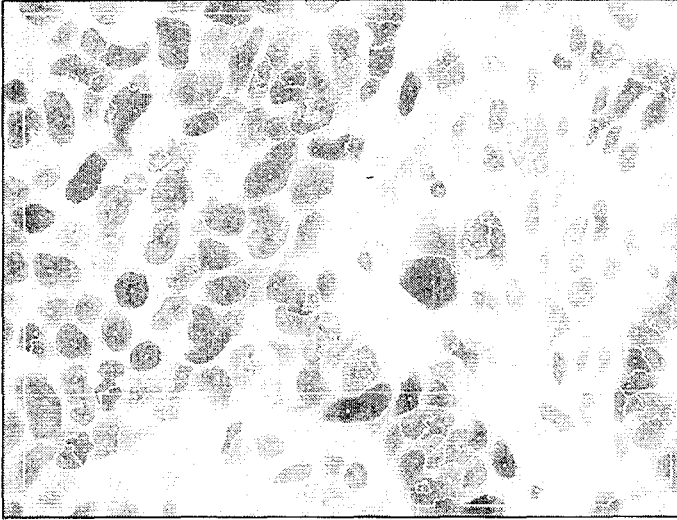


fig. 6 : Résultat de la segmentation de l'image (X132).

DISCUSSION

La résolution du problème par une approche orientée régions mettant en oeuvre des outils de la morphologie mathématique (dont une extension de la ligne de partage des eaux vers un opérateur de croissance multi-critères) et la limitation des mesures aux seuls compartiments tumoraux, offrent au pathologiste une technique de quantification des immunomarquages automatisée donc potentiellement indépendante du choix de la zone de mesure. La méthode de classification des régions, basée sur la comparaison des caractéristiques couleur des pixels avec un espace de référence, permet en outre d'adapter cette technique à de nombreux protocoles de coloration immunohistochimique (en changeant simplement l'espace de comparaison). Dès lors, si des champs de mesure régulièrement espacés sont sélectionnés sur les coupes histologiques, et si l'on respecte un échantillonnage aléatoire des coupes dans le prélèvement tumoral, on peut par des règles de stéréologie [7] en déduire la proportion volumique de l'immunomarquage nucléaire. Enfin, l'utilisation d'un graphe de Delaunay pour la recherche du masque binaire des lobules tumoraux permet d'accéder à des paramètres de distribution des profils nucléaires dans les zones actives du cancer (taille, nombre, distance inter-profil, et surtout distribution de l'immunomarquage).

REMERCIEMENTS

Ce travail a été réalisé dans le cadre du partenariat des laboratoires des auteurs au sein du *Pôle Traitement et Analyse d'Images de Basse-Normandie*.

REFERENCES

- [1] Mc Guire W.L., Clark G.M. (1992), *New England Journal of Medicine*, no. 236, pp. 1756-1761
- [2] Weinsberg D.S. (1994), *Image Analysis : a primer for Pathologists*, Marchevsky A.M. and Bartels P.H. eds., Raven Press Ltd, New-York, chap. 8, pp. 235-260
- [3] Beucher S., Meyer F. (1992), *Mathematical Morphology in image processing*. Marcel Dekker inc., New York, chap. 12, pp. 433-481
- [4] Beucher S. (1990), *Segmentation d'images et Morphologie Mathématique*. PhD thesis, Ecole Nat. Sup. des Mines de Paris
- [5] Glasbey C.A., (1993), *An analysis of histogram-based thresholding algorithms*, CVGIP graphical models and image processing, vol. 55, no. 6, pp. 532-537
- [6] Preparata F.P., Shamos M.I., (1985), *Computational geometry : an introduction*, Spinger Verlag
- [7] Gundersen H.J.G., Jensen E.B., (1985), *Stereological estimation of the volume-weighted mean volume of arbitrary particles observed on random sections*, Journal of Microscopy, no. 138, pp. 127-142