

Recherche d'images médicales par leur contenu numérique dans le domaine compressé : comparaisons de signatures construites à partir de la quantification vectorielle et des normes JPEG.

Guy CAZUGUEL¹, Richard ORDOÑEZ¹, John PUENTES¹, Jean-Michel CAUVIN²,
Basel SOLAIMAN¹, Christian ROUX¹

LATIM – INSERM U650, Bat 2 bis, CHU Morvan, BREST, FRANCE

¹Département I.T.I., ENST de Bretagne, CS 83818 29238 BREST Cedex (FRANCE)

²DIM, CHU Morvan, BREST, FRANCE

Guy.Cazuguel@enst-bretagne.fr, JR.Ordonez@enst-bretagne.fr

Résumé

Nous nous intéressons à la recherche d'images médicales par leur contenu numérique dans le domaine compressé. Il s'agit d'apporter une aide au diagnostic au médecin, en lui proposant de retrouver dans une base d'images médicales des images visuellement voisines des images dont il dispose à la suite d'examens. Nous associons aux images des signatures basées sur des informations globales, sans segmentation, ni extraction de caractéristiques particulières des images. La recherche s'effectue en calculant, au sens d'une métrique donnée, la distance entre les signatures dans la base et la signature de l'image requête. Ce calcul nous fournit les images réponses, en dehors de toute signification sémantique. Nous avons choisi d'utiliser les informations issues de processus de compression pour construire ces signatures. Nous présentons des méthodes s'appuyant sur la quantification vectorielle et les normes JPEG et JPEG2000. Nous comparons ces méthodes sur des bases d'images médicales, en utilisant des critères quantitatifs.

Mots clés

images, recherche par le contenu, indexation

1 Introduction

En médecine, les praticiens disposent aujourd'hui, à côté des examens cliniques et biologiques, d'images de plus en plus nombreuses, issues de modalités variées (2D, 3D, séquences d'images), pour les aider dans leur diagnostic et leurs prises de décisions thérapeutiques. L'analyse de ces images est en premier lieu une comparaison avec des situations déjà rencontrées par le médecin. Grâce à la numérisation et aux possibilités actuelles d'archivage, la communauté médicale commence à disposer de bases d'images de plus en plus importantes, accessibles via ordinateur, et correspondant à des cas cliniques représentatifs. Pour utiliser ces bases à des fins diagnostiques, il faut les indexer. L'indexation peut être faite à partir de descripteurs textuels donnés par un expert, mais aussi envisagée par des procédures automatiques travaillant sur le contenu numérique de l'image, générant des signatures caractéristiques des images analysées. A partir d'une image requête, un

système de recherche d'image par le contenu pourra alors proposer aux médecins les images de la base ayant des signatures voisines de la signature de l'image requête, au sens d'une métrique donnée. Notre travail porte sur cet aspect de l'indexation et de la recherche par le contenu. Plus précisément, nous nous intéressons à des caractérisations globales des images, ne nécessitant pas de segmentation, ni extraction de caractéristiques particulières des images du type formes, contours, etc..

Pour construire les signatures, nous nous sommes intéressés aux méthodes de compression d'images. Les traitements mis en œuvre dans les processus de compression doivent préserver la qualité diagnostique des images médicales (pour un taux de compression maximal admissible). Les données compressées doivent donc contenir l'information pertinente et devraient présenter un intérêt pour l'indexation, à la fois du point de vue qualité des signatures extraites de ces données, que du point de vue réutilisation des données pour construire les signatures, sans passer par une étape de décompression.

Dans ce papier, nous rappelons et comparons les performances des signatures que nous avons définies, associées à des méthodes de compression : la quantification vectorielle et les normes JPEG-DCT (avec pertes) et JPEG2000.

2 Signatures d'images

Nous proposons trois types de signatures, associées respectivement à la compression par quantification vectorielle, la norme JPEG-DCT et la norme JPEG-2000. Pour chaque type de signature, nous avons étudié différentes métriques, et gardé les plus performantes pour les évaluations comparatives des méthodes.

2.1 Les signatures basées sur la quantification vectorielle

Nous avons proposé comme signature [1] une image de petite taille, *la carte d'activation*, ayant la même topologie que le dictionnaire-carte de Kohonen[4]. Chaque élément de cette carte bidimensionnelle contient le nombre d'occurrences du mot de code correspondant dans l'image comprimée. Cette signature donne une information sur le contenu en textures de l'image. Intuitivement, deux images semblables utilisent les mêmes mots de code dans une

proportion équivalente, d'où des signatures voisines. Chaque image dans la base est analysée en utilisant le dictionnaire adéquat pour le type d'images considérées, et une carte d'activation lui est associée. Une image requête est analysée suivant le même dictionnaire et sa carte d'activation est comparée aux cartes de la base. Nous avons expérimenté plusieurs types de métriques pour comparer ces cartes (distance Euclidienne, intersection, corrélation, etc.) et évalué leur efficacité de retrouvaille (voir §3). Les simulations ont donné des résultats très encourageants, mais des cartes identiques peuvent correspondre à des organisations spatiales différentes des mots de code, donc des images différentes. Pour ajouter de l'information spatiale à la signature, nous proposons l'utilisation des moments spatiaux associés à chaque composante de la carte d'activation. Plus précisément, nous utilisons le centre de gravité et le moment d'inertie des blocs représentés par chaque classe de l'histogramme d'occurrences (cf figure 1 ci-dessous).

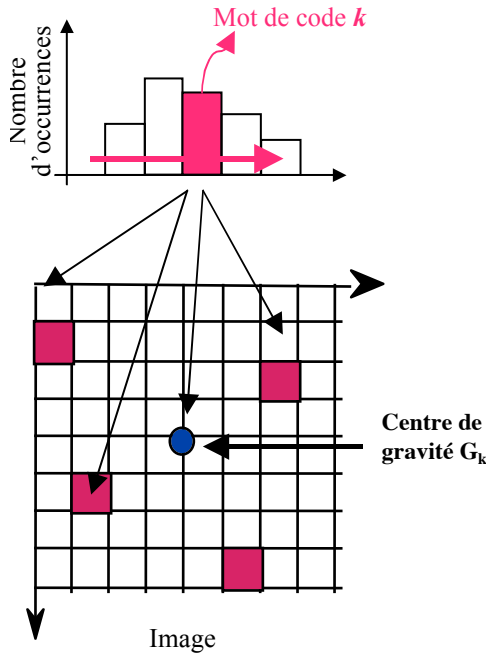


Figure 1. Ajout d'information spatiale : principe

En finale, nous ne conservons que les centres de gravité. Pour comparer les signatures, nous utilisons la distance euclidienne pour la carte d'activation (a) et une distance combinant carte et moments (b):

$$d_a(P, Q) = \sum_{i=1}^{256} \left\| H_i^P - H_i^Q \right\| \quad (a)$$

$$d_b(P, Q) = \sum_{i=1}^{256} \left\| H_i^P - H_i^Q \right\| \times \left\| \bar{G}_i^P - \bar{G}_i^Q \right\| \quad (b)$$

P et Q sont les images dont on compare les signatures. L'introduction des moments spatiaux améliore nettement les performances (voir graphique 3 à la fin de l'article)

2.2 Signatures basées sur JPEG-DCT

Dans un premier travail, nous avons étudié la signature proposée dans [5], qui utilise les relations mutuelles entre les coefficients DCT de régions non connexes dans les images. Mais les résultats obtenus n'étaient pas satisfaisants [6]. Nous avons donc développé une autre approche qui emploie les coefficients quantifiés et ordonnés générés par le standard pour produire une signature. Cette signature se compose de deux parties principales : une composante spectrale et une composante spatiale. La partie spectrale est l'histogramme bidimensionnel des fréquences d'occurrence des coefficients DCT. Un axe représente chacun des 64 coefficients DCT des blocs 8×8 transformés ; l'autre axe définit les intervalles de variation de chaque coefficient. Pour réduire la taille de la signature, nous avons fixé à 20 le nombre d'intervalles qui couvrent toutes les valeurs possibles des coefficients. La partie spatiale reprend l'idée développée en 2.1 : elle est constituée par les deux premiers moments spatiaux - centre de gravité et moment d'inertie - de tous les blocs associés à chaque classe de l'histogramme bidimensionnel des coefficients DCT (vecteurs de 64x20 valeurs au maximum). Pour la recherche d'images, nous comparons les signatures de l'image requête Q et de l'image candidate C de la base (utilisation de la norme L1). Dans ce premier travail, les coefficients de pondération λ sont égaux à 1.

$$d(P, Q) = \sum_{i=1}^{\text{DCT classes}} \sum_{j=1}^{20} \left[\lambda_1 \left| b_{i,j}^Q - b_{i,j}^C \right| + \lambda_2 \left| g_{i,j}^Q - g_{i,j}^C \right| + \lambda_3 \left| m_{i,j}^Q - m_{i,j}^C \right| \right]$$

3 Signatures basées sur JPEG-2000

La norme JPEG 2000 est une méthode à fort taux de compression, avec une optimisation de la distorsion et une qualité subjective de l'image supérieure à celle obtenue par les méthodes existantes (en particulier JPEG-DCT). L'encodeur JPEG-2000 s'appuie sur la Transformée en Ondelettes Discrètes (DWT). Le codage entropique dans l'encodeur JPEG-2000 utilise un algorithme de "modélisation de bits" : les coefficients des ondelettes sont d'abord quantifiés sous forme de mots de 14 bits, et regroupés en plans de bits (*bit-planes*) (voir figure 2). Le codage considère d'abord le plan de bits d'ordre le plus élevé. L'algorithme s'appuie sur une représentation hiérarchique des plans de bits, des bits les plus significatifs aux bits les moins significatifs des coefficients des ondelettes. Par conséquent, le flux binaire formé, avec la hiérarchie inhérente, peut être enregistré ou transféré à un débit binaire donné en maintenant les caractéristiques principales de l'image.

Pour générer une signature, nous partons du travail de Liu et al [3], qui utilisent ce principe de codage de JPEG-2000. Ils proposent comme signature (spectrale) un histogramme

bidimensionnel des "bits significatifs" pour chaque résolution. Le bit significatif correspond au premier bit à 1 dans la représentation hiérarchique des plans de bits des coefficients d'ondelettes.

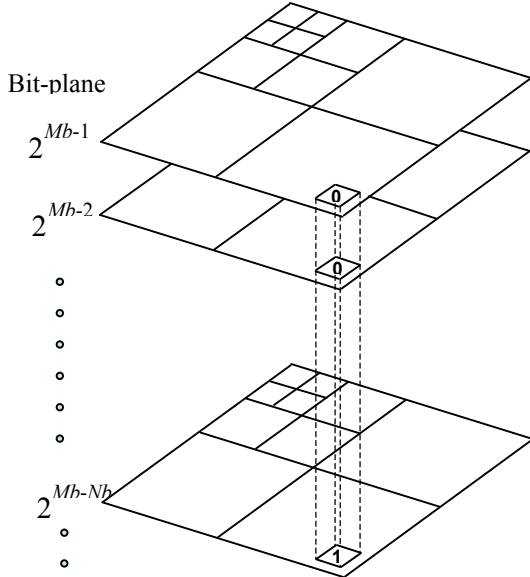


Figure 2 – Bit-planes formés à partir des coefficients des ondelettes quantifiés.

Dans la fig.2 par exemple, le bit le plus significatif du coefficient considéré se trouve dans le bit-plane 2^{Mb-Nb} . Les bits moins significatifs d'un coefficient ne sont pas pris en compte pour les calculs dans les plans correspondants. Cette signature est intéressante mais elle ne prend pas en compte l'information spatiale dans l'image. Pour cette raison, nous proposons deux variantes [2] : dans la première, nous calculons le nombre de bits significatifs dans chaque résolution et chaque direction données par la transformation en ondelettes ; dans la deuxième, nous introduisons de l'information spatiale en calculant les centres de gravité et les moments d'inertie associés à chaque plan de bits, en ne tenant compte pour un coefficient que du bit significatif. La recherche d'images s'effectue en comparant la signature de l'image requête Q aux signature des images de la base. La norme L1 a été utilisée comme mesure de similitude :

$$d(Q,C) = \sum_{l=1}^L \sum_{s=1}^S \left[\lambda_1 |b_{l,s}^Q - b_{l,s}^C| \right. \\ \left. + \lambda_2 |g_{i,j}^Q - g_{i,j}^C| + \lambda_3 |m_{i,j}^Q - m_{i,j}^C| \right]$$

Les $b_{i,j}$ sont les nombres de bits significatifs dans le plan i , à la résolution j , au sens défini ci-dessus, les $g_{i,j}$ sont les centres de gravité et les $m_{i,j}$ sont les moments d'inertie ; les paramètres λ_k , permettent d'ajuster la signature aux caractéristiques des images étudiées (λ_2 et λ_3 plus

importants pour des images présentant des contours que pour des images très texturées).

4 Méthodologie de comparaison

Pour comparer les signatures, nous avons utilisé l'indice d'efficacité de retrouvaille proposé par B.M. Methre et al [7], et le graphique précision – rappel utilisé dans la communauté de la recherche d'information. Ces deux méthodes présentent l'avantage de fournir des résultats quantitatifs, mais nécessitent de classer les images, ce qui n'est pas toujours simple en imagerie médicale.

Indice d'efficacité : nous considérons dans une base de K images, les images "similaires" (visuellement, ou appartenant à une même classe lorsque qu'une classification existe). Soit N_i le nombre d'images similaires à l'image i . Pour une image requête q , nous cherchons les $k = N_i + \tau$ images les plus proches. Soit n_q le nombre d'images correctes parmi ces images. L'indice d'efficacité, pourcentage global de bonnes réponses, est alors défini par :

$$\eta_r = \frac{\sum_{q=1}^K n_q}{\sum_{q=1}^K N_q}$$

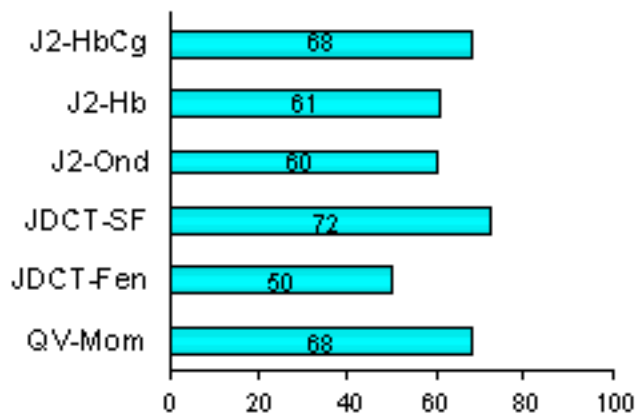
Graphiques précision-rappel : soit a le nombre d'images correctes (similaires à l'image requête) trouvées parmi les k images les plus proches de l'image requête, c le nombre d'images correctes non trouvées dans ces k images. La précision p est égale à a/k et le rappel r à $a/(a+c)$. Pour une valeur de k donnée, on calcule la moyenne des valeurs p et r sur une base d'images. On construit ensuite une courbe précision-rappel, paramétrée par k . Plus les valeurs de p et r sont voisines de 1, meilleure est la performance de la signature (nombre de bonnes propositions par rapport à l'image requête),

5 Résultats et conclusion

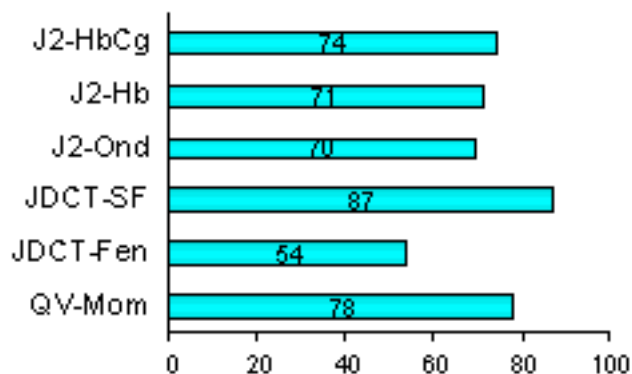
Les résultats sont résumés dans les graphiques 1 et 2 ci-dessous. En prenant en compte la taille des signatures (tableau 1), les signatures basées quantification vectorielle (QV-Mom) donnent les meilleurs résultats sur l'ensemble des bases, mais les signatures basées sur la norme JPEG 2000 (J2-HbCg) donnent des résultats pratiquement aussi intéressants, mais avec une taille de signature plus élevée. Dans notre problématique, ce sont cependant ces signatures qu'il faudra privilégier car elles s'appuient sur une norme générique, contrairement à la quantification vectorielle qui nécessite d'avoir des dictionnaires adaptés à chaque base, i.e. à chaque type d'image. Dans la poursuite de ce travail, nous étudions en particulier les possibilités de réduire la taille de cette signature. Enfin il faut noter que la quantification vectorielle pourrait cependant présenter un intérêt dans des applications différentes. Les mots des dictionnaires sont des éléments constitutifs d'images qu'on pourrait utiliser pour composer des requêtes, avec des organisations spatiales prédéfinies.

Références

- [1] J. R. Ordóñez, G. Cazuguel, J. Puentes, B. Solaiman, C. Roux. " L'utilisation des moments spatiaux pour la recherche d'images médicales par leur contenu dans le domaine compressé : application à la quantification vectorielle et le standard JPEG-DCT". TAIMA 2003, Hammameth, Tunisie, octobre 2003.
- [2] J. R. Ordóñez, G. Cazuguel, J. Puentes, B. Solaiman and C. Roux. " Indexation d'images médicales basée sur les informations spatiales et spectrales extraites de JPEG-2000". Compression et REprésentation des Signaux Audiovisuels (CORESA2003). Lyon, France, Janvier 2003
- [3] C. Liu et M. Mandal. Image Indexing in the JPEG2000 Framework. Dans *Proc. of the SPIE, Vol. 4210*. Pages 272-280, Boston – USA, 5-8 Nov. 2000.
- [4] T. Kohonen, *Self-Organization and associative Memory*, 3rd ed., Springer-Verlag, 1989.
- [5] M. Shneier, "Exploiting the JPEG compression scheme for image retrieval", *IEEE Trans. on Pattern Analysis*, vol. 18, no. 8, pp. 849-853, August 1996.
- [6] G. Cazuguel, J. R. Ordóñez, B. Solaiman, J.-M. Cauvin, C. Roux, "Indexation d'images médicales par leur contenu numérique : comparaison d'une approche textures et d'une approche dans le domaine transformé DCT", (*TAIMA'01*), Hammamet, Tunisie, Octobre 2001.
- [7] B. M. Methre, M. Kankanhalli, A. D. Narasimhalu et G. C. Man. Color matching for image retrieval. *Pattern Recognition Letters*, 16: 325-331, 1995



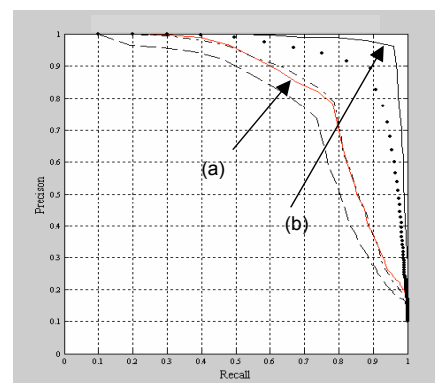
Graphique 1 : Efficacités moyennes sur l'ensemble des bases d'images ($\tau = 0$)



Graphique 2 : Indice d'efficacité sur la base d'échoendoscopies ($\tau = 0$)

Méthode	Taille de la signature (image de 2,01Mbits)
J2000-HbCg	2.73 kbits
J2000-Hb	0.91 kbits
J2000-Ond	5.4 kbits
JDCT-SF	13.3 kbits
JDCT-Fen	2.0 kbits
QV-Mom	1.92 kbits
J2000-HbCg	2.73 kbits

Tableau I



Graphique 3 : graphique précision – rappel pour les signatures basées QV sans (a) et avec moments spatiaux (b).