
Bouclage de pertinence négatif pour la recherche des images à base de descripteurs de sous-bandes d'ondelettes

Abir Gallas, Walid Barhoumi, Ezzeddine Zagrouba

Laboratoire RIADI Équipe de recherche systèmes intelligents en imagerie et vision artificielle (SIIVA)

Institut Supérieur d'Informatique, Université de Tunis el Manar

2 rue Abou Rayhane Bayrouni, 2080 Ariana, Tunisie

abirgallas@yahoo.fr, walid.barhoumi@laposte.net, ezzeddine.zagrouba@fsm.rnu.tn

RÉSUMÉ. La recherche des images par le contenu dans des grandes bases généralistes d'images est une tâche fastidieuse vu le taux d'hétérogénéité assez élevé même au sein d'une seule classe de la base ainsi que la grande dimension de l'espace des descripteurs relatifs aux images. Pour cela, nous proposons de guider la recherche par des descripteurs de bas-niveau de taille réduite à base de sous-bandes d'ondelettes relatives aux régions les plus significatives dans chaque image après une phase de segmentation floue. De plus, nous proposons une technique de bouclage de pertinence négatif sur les poids relatifs aux régions. Les expérimentations et l'étude comparative avec des approches similaires prouvent la robustesse de l'approche proposée en termes d'apport sémantique offert par l'utilisation des sous-bandes d'ondelettes ainsi que par le bouclage de pertinence négatif.

ABSTRACT. Content-based image retrieval in large image data sets is a tiresome task considering the high rate of heterogeneity even within the same class as well as the high dimensionality of the descriptors space. For that, we propose to guide the research by low level descriptors of reduced size based on sub-bands of wavelet relating to the most significant regions in each image after a fuzzy segmentation step. Moreover, we propose a negative relevance feedback technique based on region weights. Experiments and comparative study with other similar approaches prove the robustness of the proposed approach in terms of semantic contribution thanks to the use of the wavelet sub-band and the negative relevance feedback.

MOTS-CLÉS: recherche des images par le contenu, bouclage de pertinence négatif, transformation d'ondelettes, sous-bande HH, pondération des régions.

KEYWORDS: region-based image retrieval, negative relevance feedback, wavelet transformation, HH wavelet sub-band, region weights.

DOI:10.3166/TS.29.157-177 © 2012 Lavoisier

Extended abstract

During the two last decades, image retrieval in large heterogeneous datasets has attracted many researchers' attention. The very first attempts at image retrieval were exclusively based on textual annotations. Although relatively simple and computationally efficient, text-based image retrieval poses several problems of subjectivity and feasibility. Thus, most of the recent image retrieval methods adopted the content-based approach. However, existing content-based image retrieval (CBIR) systems are underperforming in the case of huge bases and this is mainly due to the semantic gap between the description of an image in terms of low-level features and its high-level semantic content. In order to bridge this semantic gap, two solutions were proposed: region-based image retrieval (RBIR) and relevance feedback (RF). The first solution aims to avoid the global description of image content by using local descriptors for each object within the image. For the second solution, the user may have the opportunity to express more precisely his target by refining his request while commenting the already displayed results. The idea of improving retrieval results of CBIR systems by using the RF already exists. But, the combination of RF with RBIR is much less common. Most of the proposed solutions in this context are based on the color descriptors for the regions' characterization. However, in addition to the problem of choosing the right color space, it is often necessary to strengthen by other descriptors, what requires the reduction of the description space. For this, we opted to use the wavelets model on different resolutions, which allows the incorporation of the color variations within the objects silhouettes. Indeed, we propose to combine between a RBIR modeling, a negative relevance feedback and a description of the regions based only on the higher frequency sub-band of the second level of wavelet decomposition. In addition to the optimization of the indexing and retrieval costs, the restriction to the high resolution sub-band of level 2 is justified by the fact that this band comprises generally the totality of the useful information in each region. Thus, after coarsely segmenting an image into regions using a fuzzy technique, its index is composed of the spectral matrices of the HH_2 for all regions composing this image. Besides, the visual importance of each region is also expressed with two fuzzy factors in order to ensure robustness against inaccurate segmentation. In fact, a spatial disposition descriptor reflects the process of the human visual perception which attributes more importance to the most centralized regions. Additionally, each region is weighted according to its percentage of surface in the corresponding image, illustrating its relative importance according to the human visual system. Then, to compare two images, we used an efficient heuristic in order to obtain relevant results while keeping an acceptable computational time. Indeed, we adapted the traditional dynamic programming algorithm by integrating the notion of region weights. This procedure is guided by regions' weights while treating the regions of the query-image and the tested image in a descending order, with the purpose of adding a semantic dimension to ensure more logical comparison between the different images. Indeed, we start by comparing the spectral matrices of the treated sub-band (HH_2) for each couple of regions in the two compared images. Rates of similarity between the various matrices of all the regions are thereafter obtained and recorded

in the same order of the regions' weights. Then, in order to define the rate of total similarity, we calculate a weighted average by an additional factor supporting the principal regions of interest in the two compared images, and only images recording higher rates of similarity with the query image are displayed as the retrieval results. After displaying the first set of the retrieval results, the user has the possibility to express his judgment so guiding the next retrieval iteration in order to get closer to his target. In fact, we proposed a relevance feedback technique which belongs to the class of vector representation model. This choice is due to the simplicity of this model and to its consistency with the modeling of the index of an image in our case, which is presented as a set of descriptive vectors relatively to the set of regions composing the image. Moreover, the reduced size of the used description space in the proposed method avoids the necessity of reducing this space. Thus, the use of a vector representation model provides encouraging results even with only some negative examples, what is not the case for the other relevance feedback models (especially that based on learning). Indeed, we proposed a negative RF technique which is designed according to the region-based representation of the images with the wavelets sub-bands. This technique is based on the adaptation of the regions weights of the query-image according to the negative feedbacks of the user. These feedbacks aim to close the ideal target-query and to eliminate the false correspondences while trying to minimize the similarity rate between the objects of the query-image and those belonging to the selected negative examples. The principal objective of our region-based method of RF is to adjust the visual importance of the query-image regions. Thereafter, new retrieval iterations will be applied on the remainder of the database while adding additional spectral information. This is carried out thanks to the widening of the used wavelet band from HH_2 sub-band only towards HH_2 and HB_2 . It makes it possible to take into account more details relating to the shape of objects. Indeed, this second sub-band was not considered in the first retrieval iteration because it is likely to obtain several false matches, especially for natural objects. This is mainly due to the instability of the shape for these objects, especially in the case of large image database. However, during the negative RF, the irrelevant examples will be eliminated and only the images that keep eventually a semantic link with the query-image will be considered in the upcoming retrieval iterations. Thus, enlargement of the descriptive band of the regions during the negative RF permits to better characterize the target images while avoiding the risk of matching semantically different images which contain objects with similar shapes. A collection of experiments carried out on the challenging « Wang » database highlights the effectiveness of the proposed approach for region-based image retrieval, in terms of the Recall/Precision indicator. In particular, the recorded results prove that applying only one negative relevance feedback iteration improves considerably the retrieval performance. Moreover, the comparison with similar state-of-the-art systems (RBIR with wavelet descriptors) shows the improvement made by the proposed method. Indeed, comparatively to the standard SIMPLicity system, the Recall/Precision curve of our system indicates a considerable improvement of the quality of results, although we use a reduced descriptive vector compared to SIMPLicity which combines many descriptors (texture, color and wavelets) relatively to each 4×4 block.

1. Introduction

La quantité immense des images qui ne cesse d'évoluer surtout avec l'apparition d'une panoplie d'appareils d'acquisition simples et à coût raisonnable, explique le besoin des systèmes d'indexation et de recherche des images par le contenu visuel (CBIR, pour *Content-Based Image Retrieval*). Historiquement, ces systèmes ne présentaient pas la première solution pour rechercher efficacement des images. A l'origine, les systèmes de recherche d'images étaient à base de descripteurs textuels où chaque utilisateur insérait le mot-clé qu'il considérait le plus explicatif et représentatif de sa requête, et le système recherchait les images dont l'annotation textuelle était la plus similaire à celle donnée par l'utilisateur. Cette méthode paraît simple sauf qu'elle posait plusieurs problèmes notamment celui du fossé sémantique. Ainsi, dans le but de réduire cette distance entre ce que l'utilisateur désire avoir et ce que le système lui propose comme solutions, l'idée des CBIR s'avère plus prometteuse. Cependant, les systèmes CBIR actuels sont toujours peu performants en ce qui concerne la recherche des images dans des bases généralistes. Pour remédier à ce problème, deux solutions ont été principalement présentées : la recherche des images à base du contenu des régions (RBIR, pour *Region-Based Image Retrieval*) et le bouclage de pertinence (BP). La première solution vise à éviter l'aspect global de la description du contenu de l'image en se basant sur des descripteurs locaux pour chaque objet traité. Pour la deuxième solution, l'utilisateur peut avoir la possibilité d'exprimer plus précisément son besoin en raffinant sa requête par le bouclage de pertinence. En effet, l'idée d'améliorer les résultats de recherche des systèmes CBIR en utilisant le BP existe déjà. Mais, faire une combinaison entre le RBIR et le BP est beaucoup moins répandu (Barhoumi et *al.*, 2009). La majorité des solutions proposées dans ce contexte sont basées sur la caractérisation des régions par le descripteur couleur (Desalears et *al.*, 2004). Toutefois, en plus du problème de choix du bon espace couleur (entre RGB et Lab essentiellement) dans le cas des bases généralistes, il est souvent nécessaire de renforcer la caractérisation des régions par d'autres descripteurs, ce qui impose la réduction de l'espace de descripteurs (Huang et *al.*, 2010). Pour cela, nous avons opté pour l'utilisation des ondelettes afin de réaliser une modélisation sur différentes résolutions. En effet, nous proposons de combiner entre une modélisation RBIR, un BP négatif de modèle vectoriel et une description des régions uniquement par la sous-bande de haute fréquence du deuxième niveau des ondelettes.

La suite de cet article est organisée comme suit. La deuxième section synthétise l'état de l'art sur la combinaison entre les systèmes RBIR et les ondelettes. Nous proposons aussi une classification des méthodes de bouclage de pertinence utilisées dans le cadre du RBIR. La signature de l'image ainsi que les techniques proposées de comparaison d'images et de bouclage de pertinence négatif sont présentées dans la section 3. Une évaluation de l'approche interactive proposée ainsi qu'une étude comparative avec des approches similaires sont présentées dans la section 4. Enfin, des conclusions et des idées pour des futurs travaux sont récapitulées dans la section 5.

2. État de l'art

Des résultats prometteurs de recherche des images à base de transformations d'ondelettes ont été approuvés en se basant sur le calcul d'une distance entre les ondelettes représentant l'image-requête et celles des images indexées. En plus de l'intégration de l'information de texture, les ondelettes permettent de garder l'allure des couleurs d'une façon grossière sans tenir compte des petits changements visibles sur un nombre limité de pixels (Kao, 2008). Le vecteur descriptif d'une image dans la méthode proposée par Wang (Wang *et al.*, 2001) comporte une représentation des transformations des ondelettes appliquée sur des blocs de 4×4 pixels sur les composantes couleur de l'espace LUV. Hiremath *et al.* (2006) proposent une combinaison entre l'information spectrale, la texture et la couleur dans le cadre d'un RBIR. De même, Lamard *et al.* (2007) suggèrent un RBIR basé sur le calcul d'une distance pondérée entre les distributions des transformations des ondelettes de différentes images. Ces travaux utilisent les ondelettes pour caractériser l'image en globalité et non pas sur des régions séparées. D'autres travaux se sont focalisés sur l'étude fréquentielle au niveau local sur des points d'intérêt toujours en combinaison avec d'autres descripteurs de bas-niveau (Tian *et al.*, 2001). Tous ces travaux utilisent soit la totalité des différentes bandes spectrales ce qui est équivalent à un coût élevé de calcul ; soit les sous-bandes de basses fréquences décrivant l'allure des images et non les détails des régions séparées.

Le bouclage de pertinence est rarement modélisé avec des descripteurs fréquentiels combinés avec le poids de chaque région. En effet, les méthodes de bouclage de pertinence peuvent être classifiées en trois catégories : par modèle d'apprentissage, par modèle de représentation vectorielle et par modèle probabilistique. Pour la première classe, l'une des méthodes les plus utilisées est l'apprentissage par la méthode des séparateurs à vaste marge (SVM, pour *Support Vector Machine*) (Liu *et al.*, 2008). Bien que cette méthode offre une flexibilité au niveau des données traitées et une rapidité d'apprentissage et de test, il faut noter que le SVM ne fournit généralement pas des résultats pertinents à cause du nombre limité des échantillons (exemples positifs et négatifs) sur lesquels il opère. Pour cela, à chaque fois qu'il est proposé comme technique de bouclage de pertinence, il est hybridé avec d'autres méthodes remédiant à la défaillance du nombre réduit des exemples (Liu *et al.*, 2008). Nous citons également dans le même contexte la méthode d'apprentissage par amplification (*Boosting*) (Huang *et al.*, 2010). Comme son nom l'indique, cette méthode permet de renforcer les résultats de classification *via* une démarche spécifique. L'idée est de considérer l'ensemble des descripteurs de la collection traitée tout en accordant un classifieur à taux d'erreur élevé pour chaque descripteur. Par la suite, le classifieur qui minimise le taux d'erreur relatif est choisi avec une certaine pondération. Finalement, le classifieur adopté est une combinaison d'un ensemble de faibles classifieurs avec différents poids accordés à chacun. Toutefois, le problème de boosting est toujours la taille de l'information sur laquelle il opère : le nombre des descripteurs doit être très élevé pour garantir de bons résultats. De plus, cette méthode fonctionne mieux sur les descripteurs de l'image qui ne sont pas fortement corrélés avec la perception humaine (la couleur par exemple) (Deselaers *et al.*, 2004). La tendance actuelle consiste à réaliser

l'apprentissage par des algorithmes génétiques (Francisco da Silva *et al.*, 2011), des arbres de décision (Vannoorenberghe, 2004) ou même des colonies de fourmis (Patrik *et al.*, 2011) et des colonies d'abeilles (Zhang *et al.*, 2010). Ceci peut être vu comme un effet de mode mais les résultats obtenus dans certains cas argumentent ce choix (Bulò *et al.*, 2011).

La deuxième classe est celle du modèle vectoriel. Le principe de ce modèle est de focaliser sur la représentation de l'image sous la forme d'un vecteur descriptif. En effet, l'idée de l'apprentissage est toujours présente même avec ce modèle, sauf qu'il n'y a pas concrètement des classificateurs ou des machines d'apprentissage dédiées mais plutôt des mises à jour successives des données relatives à la représentation de la collection d'images. Deux variantes principales existent dans ce modèle : la première se base sur la pondération des descripteurs et la deuxième sur la pondération des images elles-mêmes. Pour la première variante, les systèmes de CBIR proposent l'intervention de l'utilisateur pour mettre à jour les descripteurs utilisés lors de la recherche. En effet, la recherche s'effectue généralement selon des critères de couleur, de texture et de forme. Chacun de ces critères est lié à une pondération pour que la recherche soit influencée le plus par le descripteur le plus pondéré. Après une première recherche, l'utilisateur peut redéfinir un nouvel ordre de pondération des descripteurs utilisés. Ceci peut être fait soit par la désélection de certains descripteurs (ce qui est équivalent à une pondération égale à zéro) ou par l'attribution d'un nouveau poids à chaque critère. Cette technique est aussi appelée « feature re-weighting » au niveau de laquelle on applique le mouvement du point-requête. Ceci consiste en un déplacement vers la zone des exemples positifs et un éloignement de la zone des exemples négatifs dans l'espace de descripteurs des images (Kim, Yu, 2008). La deuxième variante consiste à accorder un score à chaque image reflétant l'intérêt qu'elle représente pour l'utilisateur. Initialement, ce score est identique pour toutes les images de la collection. Par la suite, après la première recherche, la valeur de ce score est mise à jour pour traduire le degré de similarité avec l'image-requête ou exprimer l'ordre d'apparence lors de la recherche. La définition du niveau de pertinence par l'utilisateur affecte ce score pendant les recherches ultérieures. Une image sélectionnée comme pertinente peut regagner un poids additionnel favorisant son degré de similarité avec d'autres images et permettant ainsi de retrouver de nouvelles correspondances (Gallas *et al.*, 2010), alors que les images notées comme non pertinentes seront écartées des recherches postérieures avec une minimisation de score ou même une annulation (Kherfi, Ziou, 2006). La limite principale de ces deux variantes est le risque d'aboutir à des optimums locaux. En effet, la projection des images sur les axes vectoriels des descripteurs utilisés peut rassembler des images semblables selon l'apparence et non pas selon la sémantique. Donc, il devient difficile de retrouver d'autres images ayant la même sémantique lors de la phase de bouclage de pertinence. Pour cela, Hua *et al.* (2006) proposent une nouvelle structuration de la collection d'images sous la forme d'un arbre. Chaque nœud de cet arbre correspond à un cluster d'images similaires et à chaque niveau de l'arbre, le nœud du niveau supérieur se divise en un ensemble de nœuds représentant des clusters plus fins jusqu'à arriver aux feuilles qui sont des images uniques.

Pour le modèle probabilistique, les itérations de bouclage de pertinence sont des probabilités conditionnelles où il faut à chaque phase tenir compte des actions faites pendant les phases précédentes. L'information acquise antérieurement ne sert pas seulement à éliminer les fausses correspondances (surtout pour le BP négatif) mais aussi à définir de nouvelles métriques de calcul de similarité plus souples et plus convenables aux attentes de l'utilisateur. Kherfi et Ziou (2006), proposent un modèle bayésien qui soutient les classes d'images qui assignent une probabilité élevée aux images définies comme exemples positifs et pénalise les classes qui accordent une probabilité élevée d'appartenance aux exemples négatifs. Dans (Cox *et al.*, 2000), l'utilisateur choisit les exemples positifs parmi la collection d'images pertinentes, et parmi les images non pertinentes il sélectionne les exemples négatifs. Ensuite, les auteurs emploient un modèle bayésien dans lequel ils essayent d'estimer la distribution des images pertinentes tout en réduisant au minimum la probabilité de rechercher des images non pertinentes. Su *et al.* (2003), présentent une classe bayésienne pour les exemples positifs. Ceci sert à estimer la distribution gaussienne qui représente la classe des images cherchées, alors que les exemples négatifs sont employés pour modifier le rang des candidats recherchés. De même, un algorithme glouton d'approximation est défini par (Paredes *et al.*, 2000). L'idée de base de cette approche est d'utiliser la totalité de la collection d'images en accordant à chaque sous-ensemble de cette collection une étiquette. Par la suite, les probabilités calculées portent sur les informations données par ces étiquettes textuelles. Par rapport à la classe de modèle d'apprentissage, le nombre réduit des exemples présentés par l'utilisateur ne pose plus un problème pour le modèle probabilistique. Toutefois, le plus grand inconvénient de cette classe réside dans la complexité de la définition de la meilleure pondération qui permet d'augmenter la probabilité des exemples positifs tout en diminuant celle des exemples négatifs.

En ce qui nous concerne, nous proposons une approche qui se base sur la description des régions composant une image par les transformations des sous-bandes d'ondelettes de haut-niveau. De plus, nous appliquons une technique de BP faisant partie de la classe du modèle de représentation vectorielle et plus exactement de sa deuxième variante. Ce choix s'explique par la simplicité de ce modèle ainsi que sa conformité avec la modélisation de l'index de notre base d'images présenté sous la forme d'un ensemble de vecteurs descriptifs relatifs à chaque région. En outre, l'espace des descripteurs utilisé dans notre approche est réduit ce qui ne pose pas le problème de réduction de cet espace. Notons enfin que l'utilisation d'un modèle de représentation vectorielle permet d'obtenir des résultats encourageants même avec un nombre réduit d'exemples négatifs ce qui n'est pas le cas pour les autres modèles de bouclage de pertinence (surtout celui de l'apprentissage) (Borghesani *et al.*, 2012).

3. Méthode proposée

Le but principal de ce travail est de combiner la recherche d'images à base de régions avec un bouclage de pertinence interactif afin d'améliorer les performances de la recherche des images dans des collections hétérogènes. Toutes les images de la

collection sont segmentées en régions et des vecteurs descriptifs de l'ensemble des régions extraites sont définis. Ce processus est réalisé hors ligne et définit l'étape d'indexation de la base. Ensuite, le premier niveau d'interaction en ligne entre l'utilisateur et le système consiste en la présentation d'une image-requête (figure 1). Cette image est grossièrement segmentée en des régions floues illustrant les objets qui le composent et qui seront indexées par des descripteurs spectraux. Notons que nous avons évité la région-requête, que ce soit au niveau de la requête initiale ou au niveau des itérations de BP, afin de simplifier l'interaction avec l'utilisateur (Huang *et al.*, 2008).

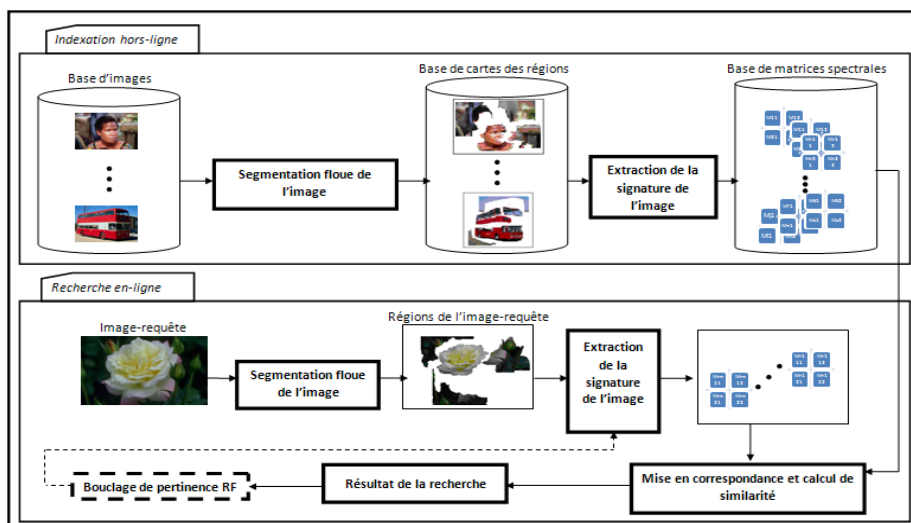


Figure 1. Illustration de la démarche appliquée par la méthode proposée pour la recherche des images

En effet, la technique la plus utilisée pour décrire une image ou une région sur le plan spectral est celle qui emploie les transformations d'ondelettes afin de réaliser une modélisation sur différentes résolutions (Zhang *et al.*, 2002). Ces transformations sont basées sur la structuration hiérarchique multiniveau qui est implémentée en utilisant une banque de filtrage. Cette modélisation en un ensemble de bandes fréquentielles peut être vue comme étant une segmentation, sauf que cette segmentation ne divise pas l'image en régions mais plutôt en hautes et basses fréquences. Une propriété importante de la transformation par ondelettes est le rapport entre les réponses d'impulsion des filtres passe-haut et les filtres passe-bas. Ils ne sont pas indépendants les uns des autres, mais sous la forme (pour chaque niveau n) : $h[n] = (-L)^n \cdot g[n]$, où $h[n]$ représente le filtre passe-haut, $g[n]$ est le filtre passe-bas et L est la longueur de filtre (en nombre de points). Dans notre cas, ces transformations d'ondelettes portent sur les régions de l'image. Ainsi, une image est représentée sous la forme d'un ensemble de nœuds tels que chaque nœud est marqué par les descripteurs de bas-

niveau d'une région particulière et il est pondéré selon l'importance visuelle de la région correspondante. Puis, l'ensemble des matrices spectrales, extrait en ligne de l'image-requête et représentant le descripteur de bas-niveau, est comparé avec les matrices, enregistrées hors ligne, des différentes régions de chaque image de la base en utilisant une heuristique de complexité minimale. En outre, nous avons intégré le bouclage de pertinence dans notre système RBIR afin d'ajouter de la sémantique à la recherche. En effet, notre méthode de bouclage de pertinence est conçue selon la caractéristique de la représentation à base de régions décrites par des sous-bandes d'ondelettes. Nous proposons un mécanisme interactif simple, mais efficace, qui se base sur l'adaptation des poids de régions de l'image-requête par les rétroactions négatives de l'utilisateur (figure 1). Ces rétroactions essaient de rapprocher la requête-cible idéale et éliminer les mauvaises correspondances tout en essayant de minimiser le taux de similarité entre les objets composant l'image-requête et celles des images sélectionnées comme exemples négatifs.

3.1. Signature d'une image

Fondée sur l'hypothèse que toute région pourrait être utile lors de l'évaluation de la similarité (Jing *et al.*, 2003), toutes les régions de chaque image sont considérées. De là, une liste triée de nœuds est définie relativement à chaque image, où chaque nœud représente une région floue grossièrement segmentée à base d'un algorithme flou de ligne de partage des eaux (Philipp-Foliguet *et al.*, 2005). Cet algorithme permet d'extraire un nombre réduit des régions représentant grossièrement les objets composant l'image. Chaque région est caractérisée ensuite par un ensemble de descripteurs de bas-niveau et par un poids évaluant son importance visuelle (Barhoumi *et al.*, 2009). D'une part, le vecteur descriptif de bas-niveau se réduit dans notre cas à l'information spectrale floue fournie par la sous-bande HH de niveau 2 des ondelettes appliquées sur les composantes couleurs de l'espace RGB de la région traitée R sous la forme d'une matrice M_{HH}^R . En particulier, puisque chaque pixel de l'image I en entrée admet un degré d'appartenance μ_R à la région R, la matrice M_{HH}^R est définie par (1) :

$$\forall (x, y) \in I, M_{HH}^R(x, y) = \mu_R(x, y) \cdot [H_x * [H_y * [H_x * [H_y * I(x, y)]_{\downarrow 2,1}]_{\downarrow 1,2}]_{\downarrow 2,1}]_{\downarrow 1,2} \quad (1)$$

où, * indique l'opérateur de convolution, $H = [H_x, H_y]^T$ est un filtre Canny passe-haut, ($\downarrow 2,1$) désigne l'échantillonnage vers le bas le long des lignes et ($\downarrow 1,2$) celui des colonnes. En effet, les tests réalisés pour retrouver l'espace couleur le plus adéquat aux images traitées, nous ont permis de déduire que chaque espace présente des points forts dans un certain contexte et pas dans l'autre. Par exemple, les espaces HSV et HSL ne sont pas uniformes donc l'utilisation de la distance euclidienne perd sa signification pour ces deux exemples. D'autres espaces couleurs tels que Lab, Luv, LCH et Ltø, qui sont des dérivations des espaces HSV et HSL mais uniformes, nécessitent une transformation à partir de l'espace RGB qui représente l'espace

couleur par défaut de la majorité des images numériques. De ce fait, nous risquons de surcharger notre système dans le cas où nous devons calculer ces transformations ou nous utilisons une combinaison d'espaces couleur afin de combler toutes les défaillances (Deselaers *et al.*, 2004). De là, nous avons choisi d'unifier les données et les descripteurs tout en prenant en compte l'information de multirésolution représentée par les ondelettes (Kao, 2008). De plus, afin de minimiser le coût d'indexation et de recherche, nous nous sommes limités à la sous-bande de haute résolution de niveau 2. Cette restriction est justifiée par le fait que cette bande comporte généralement la globalité de l'information utile de la région comme le montrent les résultats obtenus lors de la recherche. En effet, nous avons trouvé les mêmes résultats de recherche avec l'information entière des ondelettes mais en augmentant le temps de calcul par 85,31 % (cf. section 4). En effet, l'utilisation des bandes de basse fréquence, qui décrivent l'allure des objets détectés, peut provoquer de fausses correspondances, notamment pour les objets naturels (animal, plante, nourriture...). Ceci est principalement dû à l'instabilité de la forme pour ces objets surtout dans le cas de larges bases d'images. Cependant, lors du bouclage de pertinence négatif (cf. section 3.3), les exemples non pertinents pour l'utilisateur seront éliminés et seules les images qui gardent éventuellement une liaison sémantique avec l'image-requête seront considérées lors des recherches ultérieures. Ainsi, l'élargissement de la bande descriptive des régions lors du bouclage de pertinence négatif permet de mieux caractériser les images cibles tout en évitant le risque de faire correspondre des images sémantiquement différentes mais qui contiennent des objets avec des silhouettes similaires. D'autre part, l'importance visuelle de chaque région est aussi exprimée avec deux facteurs flous afin de minimiser la propagation de l'erreur et ceci en modélisant toujours l'incertitude d'affectation d'un pixel à une région unique. En effet, un facteur d'emplacement spatial reflète le processus de la perception visuelle humaine qui accorde plus d'importance à la région la plus localisée au centre. Ce degré de centralisation s'exprime sous la forme d'une distance entre le centre de la région et le centre G_I de l'image I traitée de dimension $N \times M$. Dans notre cas, le centre de chaque région R se définit comme étant son centre de gravité G_R . En effet, d'autres recherches ont défini un centre d'intérêt (Kim *et al.*, 2007) pour les régions composant une image. Mais, puisque dans la plupart des cas, ce point coïncide avec le centre de gravité nous avons opté à l'utilisation de ce dernier pour réduire le temps de calcul. En outre, chaque région R est pondérée selon son pourcentage de surface P_R dans l'image correspondante, illustrant son importance relativement au système visuel humain. Le poids w_R d'une région R est défini par la suite comme étant la moyenne entre le degré de centralisation et la portion P_R de la surface de l'image (2). Notons à ce niveau que les différentes valeurs utilisées pour la définition de la signature d'une image sont toutes floues et normalisées.

$$\forall R \in I, w_R = \frac{1}{2 \cdot N \cdot M} \left[\sum_{x=1}^N \sum_{y=1}^M \mu_R(x,y) \sqrt{(M \cdot (x_{G_R} - x_{G_I}))^2 + (N \cdot (y_{G_R} - y_{G_I}))^2} + \sum_{x=1}^N \sum_{y=1}^M \mu_R(x,y) \right] \quad (2)$$

3.2. Comparaison d'images

Nous présentons dans cette section notre approche de mise en correspondance et le principe de mesure de similarité. Nous utilisons des heuristiques afin de retrouver des résultats pertinents tout en gardant un temps de calcul acceptable. En effet, nous avons adapté l'algorithme classique de la programmation dynamique tout en intégrant la notion des poids des régions avec une complexité polynomiale égale à $\theta(n^4)$. Toute cette procédure est guidée par les pondérations des régions qui décident de la pertinence des objets contenus dans l'image. Ceci ajoute une dimension sémantique et assure une comparaison plus logique entre les différentes images. Notre approche de mise en correspondance consiste à traiter les régions de l'image-requête I_{Req} et de l'image testée I_{Test} dans un ordre décroissant de poids. En effet, les régions ayant les poids w_R les plus grands désignent généralement les objets d'intérêt que l'utilisateur cherche puisqu'elles se localisent au centre et occupent la plus grande partie de l'image. Nous commençons tout d'abord par comparer les matrices spectrales M_{HH}^R de la sous-bande traitée (HH₂) pour chaque région R . Un taux de similarité entre les différentes matrices de toutes les régions est obtenu par la suite et enregistré dans le même ordre d'ordonnement des régions (de la plus intéressante vers la moins intéressante). Puis, dans le but d'obtenir un taux de similarité totale, nous calculons une moyenne pondérée par un facteur additionnel P_{add} favorisant les régions d'intérêt principales R_1 et R'_1 dans les deux images à comparer (figure 2).








 Image I_{Req}	 R_1	 R_2				
	Poids de la région R_1 : w_{R_1}	Poids de la région R_2 : w_{R_2}				
	Matrice fréquentielle : $M_{HH}^{R_1}$	Matrice fréquentielle : $M_{HH}^{R_2}$				
 Image I_{Test}	 R'_1	 R'_2	 R'_3			
	Poids de la région R'_1 : $w_{R'_1}$	Poids de la région R'_2 : $w_{R'_2}$	Poids de la région R'_3 : $w_{R'_3}$			
	Matrice fréquentielle : $M_{HH}^{R'_1}$	Matrice fréquentielle : $M_{HH}^{R'_2}$	Matrice fréquentielle : $M_{HH}^{R'_3}$			
Liste de taux de similarité TS	TS (R_1, R'_1)	TS (R_1, R'_2)	TS (R_1, R'_3)	TS (R_2, R'_1)	TS (R_2, R'_2)	TS (R_2, R'_3)
Poids additionnel P_{add}	$\frac{(w_{R_1} + w_{R'_1})}{2}$	$\frac{(w_{R_1} + w_{R'_2})}{2}$	$\frac{(w_{R_1} + w_{R'_3})}{2}$	$\frac{(w_{R_2} + w_{R'_1})}{2}$	$\frac{(w_{R_2} + w_{R'_2})}{2}$	$\frac{(w_{R_2} + w_{R'_3})}{2}$

Figure 2. Extraction des taux de similarité entre les régions de deux images et préparation à la mise en correspondance

L'étape suivante consiste à garder les images ayant un taux de similarité de l'information spectrale supérieur à un seuil ϕ fixé au préalable. Dans le cas où deux régions ont deux poids initialement égaux, le fait de commencer par l'une ou l'autre

n'affecte pas le résultat final de mise en correspondance, en particulier pour les régions ayant les plus grands poids puisque ce sont elles qui guideront la recherche. Formellement, soit l'ensemble $\Omega = \{R_1, \dots, R_k\}$ des régions triées selon leur pondération dans un ordre décroissant et composant l'image-requête I_{Req} et l'ensemble $\Omega' = \{R'_1, \dots, R'_l\}$ des régions (triées de la même façon que Ω) composant l'image testée I_{Test} , la mesure de similarité entre I_{Req} et I_{Test} est donnée par (3) :

$$Sim(I_{Req}, I_{Test}) = \sum_{R_i \in \Omega} \sum_{R'_j \in \Omega'} \frac{w_{R_i} + w_{R'_j}}{2} \sqrt{(M_{HH}(R_i) - M_{HH}(R'_j))^2} \quad (3)$$

3.3. Le bouclage de pertinence

Un système de bouclage de pertinence efficace est celui qui présente à l'utilisateur ce qu'il cherche en peu d'itérations. L'objectif principal de notre schéma de bouclage de pertinence à base de régions est d'ajuster l'importance visuelle des régions de l'image-requête selon les rétroactions négatives des utilisateurs. En effet, si l'utilisateur est non satisfait par les images retrouvées, le bouclage de pertinence lui propose d'indiquer un ensemble d'images qu'il trouve non pertinentes dans la récupération déjà produite. La prétention de base derrière notre bouclage de pertinence à base de régions est que chacune des images déjà retrouvées devrait avoir au moins une région qui est fortement semblable, en termes de descripteurs de bas-niveau, à une (ou plusieurs) région(s) de l'image-requête. Néanmoins, il est habituel qu'une image soit affichée parce qu'elle a quelques régions qui sont fortement semblables à certaines régions peu importantes de l'image-requête. Les régions, qui peuvent être significatives mais ne représentent pas le centre de perception de l'utilisateur, correspondent dans la plupart des cas aux régions du fond (Kim *et al.*, 2007). Suite au choix par l'utilisateur des images I_{NP} appropriées (non pertinentes), l'approche de bouclage de pertinence identifie parmi les régions de l'image-requête, les moins importantes, relativement à l'intuition de l'utilisateur, afin de les éliminer pendant la prochaine itération de recherche. En effet, il s'avère plus facile pour l'utilisateur de décider directement à propos des images qu'il ne cherche pas à les obtenir plutôt que ce qu'il veut exactement avoir (Xu *et al.*, 2009). Ainsi, le bouclage négatif assure dans un nombre réduit d'itérations l'élimination des fausses correspondances. Cette élimination est basée sur la réduction du taux de similarité entre la liste des images non pertinentes I_{NP} et l'image-requête (4). En effet, au lieu d'augmenter l'importance des régions comparées par les poids additionnels utilisés au début (3), nous minimisons cette ressemblance potentielle en la divisant par le poids le plus pondérant ($P_{add} = w_{R_1} + w_{R'_1}$) (4), qui est celui accordé aux deux régions ayant les plus grands poids dans les deux images (R_1 et R'_1). Cette réduction assure que l'image choisie comme exemple négatif ne sera plus affichée parmi les résultats. Ceci s'explique par son taux de similarité minimal avec l'image-requête. En effet, si nous avons choisi de diviser par les poids du reste des régions nous risquons de ne pas abaisser suffisamment le taux de similarité.

$$Sim(I_{Req}, I_{NP}) = \frac{1}{w_{R_1} + w_{R'_1}} \left[\sum_{R_i \in \Omega} \sum_{R'_j \in \Omega'} \sqrt{(M_{HH}(R_i) - M_{HH}(R'_j))^2} \right] \quad (4)$$

Par la suite, nous lançons une recherche de nouvelles correspondances dans le reste de la base d'images. Ceci est réalisé grâce à l'élargissement de la bande des ondelettes utilisées comme descripteurs de comparaison de HH_2 seulement vers HH_2 et HB_2 (5). Ceci permet de prendre en compte plus de détails relatifs aux images traitées surtout ceux liés à l'allure des objets. Formellement, soit ζ l'ensemble des images de la base de test et ζ_{NP} l'ensemble des images non pertinentes sélectionnées par l'utilisateur, alors on a :

$$\forall I_{Test} \in \zeta / \zeta_{NP}, Sim(I_{Req}, I_{Test}) = \sum_{R_i \in \Omega} \sum_{R'_j \in \Omega'} \frac{w_{R_i} + w_{R'_j}}{2} \sqrt{(M_{HH}(R_i) - M_{HH}(R'_j))^2 + (M_{HB}(R_i) - M_{HB}(R'_j))^2} \quad (5)$$

où, $M_{HB}(x, y) = [H_x * [G_y * I(x, y)]_{1,2,1}]_{1,2}$ et G désigne un filtre Gaussien passe-bas.

4. Expérimentations et évaluations

Afin de valider l'approche proposée, nous l'avons appliquée sur des bases d'images généralistes. Le choix de la base d'images « Wang », utilisée dans le système SIMPLicity (Wang *et al.*, 2001) s'explique par le fait qu'elle représente un standard des bases généralistes pour évaluer les systèmes de recherche des images à base de contenu des régions (Deselaers *et al.*, 2004 ; Hsu, Li, 2005 ; Paredes *et al.*, 2008). Cette base, composée de 1 000 images divisées équitablement sur 10 classes étiquetées (« Afrique », « Plage », « Bâtiment », « Bus », « Dinosaur », « Eléphant », « Fleur », « Cheval », « Montagne » et « Nourriture »), présente plusieurs niveaux de difficulté vu son hétérogénéité même au sein d'une seule classe (Li *et al.*, 2008). Pour évaluer objectivement notre approche, nous proposons des courbes de Rappel/Précision sachant que la valeur du seuil ϕ pour tous les tests est égale à 50 %. Etant donné le résultat pour une requête avec une collection de documents, la mesure de Rappel/Précision basée sur la notion d'ensemble est la métrique la plus communément employée pour mesurer l'efficacité d'un système de recherche d'information. Le rappel (6) mesure la capacité de récupération d'un système pour présenter toutes les images d'une même classe. La précision (7) reflète la capacité du système d'afficher seulement les images pertinentes. Chaque requête peut être ainsi associée à une précision et à une valeur de rappel sur cette collection. Pour montrer l'efficacité de notre approche basée sur le descripteur bas-niveau d'ondelettes, le premier ensemble de courbes indique le taux de rappel en fonction de la précision des recherches effectuées avec la sous-bande HH_2 seulement et avec toute l'information d'ondelettes (figure 3). Il est clair que l'amélioration est trop faible aux dépens d'un temps de calcul très élevé.

$$\text{Rappel} = \frac{\text{Card}(\{\text{Images pertinentes}\} \cap \{\text{Images retrouvées}\})}{\text{Card}(\{\text{Images pertinentes}\})} \quad (6)$$

$$\text{Précision} = \frac{\text{Card}(\{\text{Images pertinentes}\} \cap \{\text{Images retrouvées}\})}{\text{Card}(\{\text{Images retrouvées}\})} \quad (7)$$

Le deuxième ensemble des courbes Rappel/Précision illustre les résultats produits pour les différentes classes de la base « Wang » (figure 4). Les résultats de ces courbes sont fortement sensibles au degré d'hétérogénéité de chaque classe. Prenons l'exemple de la classe « Dinosaur » ; cette classe est la plus homogène pour un système perceptuel humain par rapport aux autres classes. Pour cette classe, l'objet d'intérêt est toujours centralisé avec une taille importante et avec un fond quasiment unique pour les 100 images. Ceci explique le taux élevé de précision ainsi que du rappel. De plus, cette classe ne se compose pas de photographies mais plutôt de dessins. Ainsi, nous pouvons conclure que notre approche de mise en correspondance permet implicitement de faire une distinction entre ces deux types d'images (photographie vs. dessin) puisque les résultats affichés ne comportent que des images de cette classe (Rappel/Précision=1). De cette manière, il sera possible d'obtenir une classification raisonnable et préliminaire de l'ensemble des images de notre base sans avoir besoin d'utiliser un algorithme de classification dédié à cette tâche (Estevez, 2001). De même, la classe « Fleur » garde les mêmes caractéristiques que celle de « Dinosaur » (au niveau de centralisation et du fond uniforme). Pour les classes « Cheval », « Montagne », « Eléphant », « Plage » et « Bus » l'abaissement de taux de Rappel/Précision peut s'expliquer par l'augmentation du niveau d'hétérogénéité au sein de ces classes et de la différence élevée des apparences dans certains cas entre les mêmes objets de la même classe (figure 5). Concernant les classes « Afrique », « Bâtiment » et « Nourriture », elles sont très hétérogènes et ne présentent pas les mêmes objets dans les différentes images ce qui pose un problème d'appartenance à la classe dans le cas où nous ne considérons pas l'information apportée par la vérité terrain (figure 6).

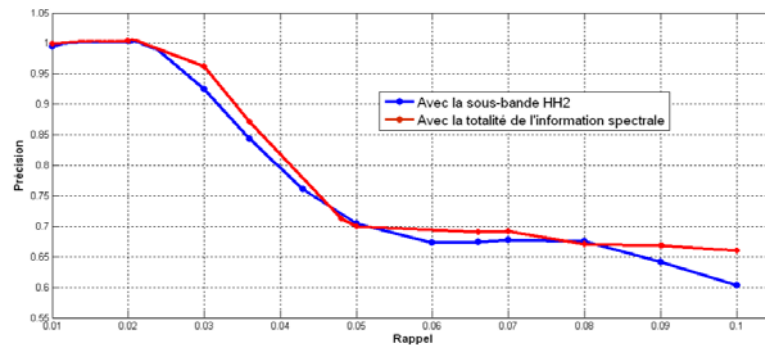


Figure 3. Courbes de Rappel/Précision comparant la recherche avec la sous-bande HH_2 et celle avec toutes les bandes du spectre

Pour évaluer l'apport de notre technique de bouclage de pertinence négatif, nous présentons le troisième ensemble de courbes Rappel/Précision (figure 7). Il est clairement remarquable que l'utilisation d'une seule itération de BP permet d'améliorer les résultats obtenus lors de la première recherche (Rappel/Précision ≈ 1). D'ailleurs, c'est ce que l'utilisateur souhaite avoir : un meilleur résultat avec le minimum d'itérations de BP. Ceci lui offre la possibilité de communiquer avec le système sans le déborder avec des tâches supplémentaires successives. En effet, les images obtenues au premier rang (généralement les quatre premières images) après la recherche répondent sémantiquement au besoin de l'utilisateur. Mais, l'intervention de ce dernier vise à améliorer le reste des résultats affichés. Nous déduisons ainsi que l'approche proposée même sans bouclage de pertinence garde un taux de Rappel/Précision élevé pour ce premier rang d'affichage (figure 8). D'ailleurs, il est clair que les deux courbes de la figure 7 sont confondues au niveau des premiers résultats.

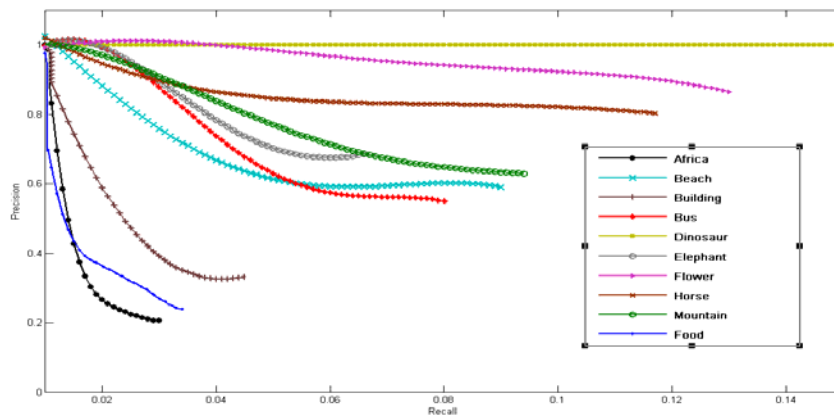


Figure 4. Courbes de Rappel/Précision des 10 classes de la base d'images de « Wang » (Wang et al., 2001)

Enfin, dans le but de comparer notre travail aux autres travaux présentés dans le même contexte (système RBIR avec des descripteurs d'ondelettes), nous avons recours à deux systèmes : le système SIMPLicity présenté par (Wang *et al.*, 2001) et le système présenté par (Hiremath *et al.*, 2006). Par rapport à SIMPLicity, la courbe de Rappel/Précision de notre système indique une amélioration considérable de la qualité des résultats (figure 9), bien que nous utilisons un vecteur descriptif réduit par rapport à SIMPLicity qui prend en compte les informations de texture, de l'espace couleur Lab et d'ondelettes. En effet, notre courbe de Rappel/Précision reste quasiment constante par rapport à celle de SIMPLicity qui décroît brusquement quand le nombre des résultats affichés augmente. Ceci peut être expliqué par la perte de l'information sémantique à cause de l'utilisation d'une description par ondelettes sur des blocs de l'image et non pas sur des objets. En ce qui concerne le second système (Hiremath *et al.*, 2006), quelques résultats de comparaison sont donnés dans

la figure 10. Ce système ne se limite pas également aux seuls descripteurs d'ondelettes mais il ajoute l'information de texture et de différents espaces couleur (RGB, HSV, YCbCr) avec une étape préalable de classification à base de texture. Pourtant, les résultats visuels prouvent une qualité meilleure de recherche par notre approche. Ce qui caractérise l'approche proposée par rapport à celles proposées par (Wang *et al.*, 2001) et (Hiremath *et al.*, 2006) est la modélisation au niveau des régions et l'application des ondelettes sur ces régions et pas sur des blocs qui ne correspondent pas forcément à des objets réels. De plus, notre travail est basé sur la notion de flou ce qui modélise l'incertitude et évite les erreurs potentiels d'appartenance aux régions exactes. En outre, nous ajoutons une phase d'interaction corrective de bouclage de pertinence ce qui permet d'améliorer encore les résultats. Ainsi, nous considérons notre approche comme plus simple que ces deux approches tout en étant plus performante vu les résultats produits sur les mêmes bases d'images.



Figure 5. Hétérogénéité des images de la base « Wang » : deux objets identiques dans la même classe (« Bus ») ayant des apparences très différentes

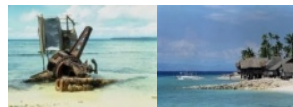


Figure 6. Deux scènes semblables illustrant deux classes distinctes (« Afrique », « Plage »)

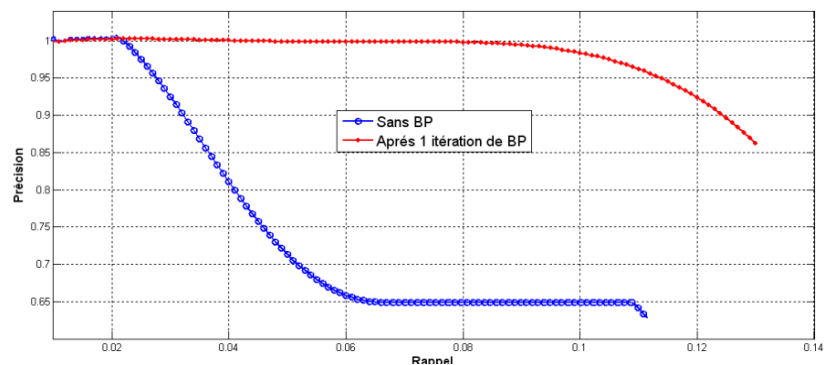


Figure 7. Courbes de Rappel/Précision illustrant l'amélioration des résultats de la recherche après le BP négatif



(a) Résultats de la recherche avant le BP, l'image-requête est encadrée par un trait uni et les exemples négatifs sont encadrés par un trait pointillé.



(b) Résultats de la recherche après l'application d'une itération de BP négatif.

Figure 8. Amélioration des résultats de la recherche après une itération de BP négatif sur une image-requête de la classe « Cheval »

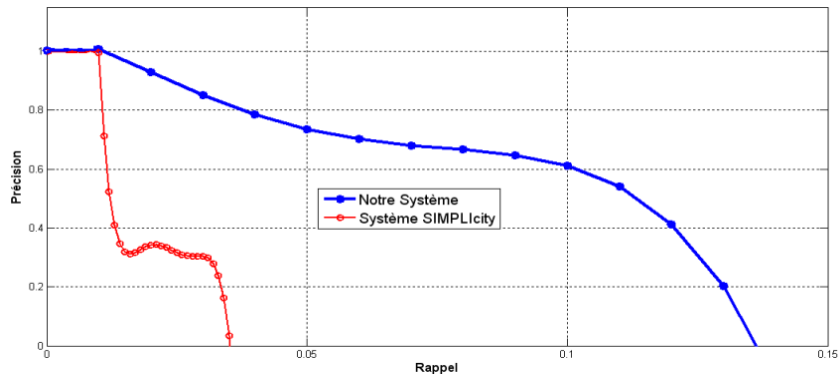


Figure 9. Courbes de Rappel/Précision illustrant la comparaison entre le système SIMPLicity et le nôtre



(a) Résultats de la recherche pour deux images-requêtes par notre système.



(b) Résultats de la recherche pour les mêmes images-requêtes par le système de (Hiremath et al., 2006).

Figure 10. Comparaison entre les résultats de la recherche de notre système et ceux de (Hiremath et al., 2006)

5. Conclusions et perspectives

Une approche fiable de recherche des images par le contenu de régions à base d'ondelettes ainsi qu'une classification des méthodes de bouclage de pertinence sont proposées dans cet article. Cette approche est basée sur une combinaison entre le RBIR, le bouclage de pertinence négatif et une description spectrale réduite. En particulier, un algorithme rapide de mise en correspondance et de mesure de similarité guidé par les poids des régions est présenté. Les expérimentations réalisées sur la base d'images de « Wang » mettent en évidence la robustesse de notre approche, qui permet l'obtention d'une classification préalable entre la photographie et le dessin (cas de « Dinosaur »). En outre, la comparaison avec les systèmes similaires (système RBIR avec des descripteurs d'ondelettes) (Wang *et al.*, 2001 ; Hiremath *et al.*, 2006) montre l'amélioration apportée par la méthode proposée. En guise de perspectives, nous essayons d'améliorer la définition du poids de chaque région pour ne pas négliger des objets sémantiquement importants dans l'image. En outre, vu le problème de gestion des grandes dimensionnalités (malédiction de dimensionnalité), nous cherchons à définir une structuration

approximative de l'espace de données. Ceci dans un but de préparer un index déjà préclassifié permettant d'afficher la page zéro de démarrage de recherche pour l'utilisateur. De plus, nous visons à définir un langage rationnel (Chen, 2010) donnant une grammaire de l'image. Ainsi, les itérations de bouclage de pertinence porteront sur une manipulation de l'automate caractérisant l'image traitée.

Remerciements

Les auteurs remercient vivement le Professeur James Z. Wang (College of Information Sciences and Technology – The Pennsylvania State University) pour avoir fourni généreusement sa base d'images « Wang » ainsi que le code source de son système SIMPLiCity.

Bibliographie

- Barhoumi W., Gallas A., Zagrouba E. (2009). Effective region-based relevance feedback for interactive content-based image retrieval. *New directions in intelligent interactive multimedia systems and services*, vol. 226, p. 177-187.
- Borghesani D., Grana C., Cucchiara R. (2012). Towards Artistic Collections Navigation Tools Based on Relevance Feedback. *Multimedia for cultural heritage*, vol. 247, p. 143-153.
- Bulò S.R., Rabbi M., Pelillo M. (2011). Content-based image retrieval with relevance feedback using random walks. *Pattern recognition*, vol. 44, n° 9, p. 2109-2122.
- Chen Z. (2010). Semantic Research on Content-Based Image Retrieval. *International Conference on Multimedia Technology*. Ningbo, 29-31 October 2010. p. 1-4.
- Cox I.J., Minka T.P., Papatomas T.V., Yianilos P.N. (2000). The Bayesian image retrieval system, PicHunter: Theory, implementation, and psychophysical experiments. *IEEE transactions on image processing*, vol. 9, n° 1, p. 20-37.
- Deselaers T., Keysers D., Ney H. (2004). Features for image retrieval: a quantitative comparison. *Notes in computer science*, vol. 3175, p. 228-236.
- Estevez D. (2001). *Dessin d'architecture et infographie, L'évolution contemporaine des pratiques graphiques*. Paris, France, CNRS Editions.
- Francisco da Silva S., Xavier Ribeiro M., Batista Neto J., Traina-Jr C., Traina A. (2011). Improving the ranking quality of medical image retrieval using a genetic feature selection method. *Decision support systems*, vol. 50, p. 1-11.
- Gallas A., Barhoumi W., Zagrouba E. (2010). Une approche rapide basée sur les régions floues pour une évaluation sémantique de la similarité d'images. *WIA 2010 Workshop en informatique et applications*. Tozeur, Tunisie, 12-14 février.
- Hiremath P.S., Shivashankar S., Pujari J.D. (2006). Wavelet based features for color texture classification with application to CBIR. *Journal of computer science and network security*, vol. 6, n° 9, p. 124-133.

- Hsu C.T., Li C.Y. (2005). Relevance feedback using generalized Bayesian framework with region-based optimization learning. *IEEE transactions on image processing*, vol. 14, n° 1, p. 1617-1631.
- Hua K.A., Yu N., Liu D. (2006). Query decomposition: a multiple neighborhood approach to relevance feedback processing in content-based image retrieval. *International conference on data engineering*. Atlanta, 3-8 Avril.
- Huang W., Gao Y., Chan K.L. (2008). A review of region-based image retrieval. *Sign process system*, vol. 59, p. 143-161.
- Jing F., Li M., Zhang L., Zhang H.J., Zhang B. (2003). Learning in region-based image retrieval. *Conference on image and video retrieval*. Urbana-Champaign, 24-25 Juillet, p. 206-215.
- Kao O. (2008). On parallel image retrieval with dynamically extracted features. *Parallel computing*, vol. 34, p. 700-709.
- Kherfi M.L., Ziou D. (2006). Relevance feedback for CBIR: a new approach based on probabilistic feature weighting with positive and negative examples. *IEEE transactions on image processing*, vol. 15, n° 4, p.1017-1030.
- Kim D.H., Yu S.H. (2008). A new region filtering and region weighting approach to relevance feedback in content-based image retrieval. *The journal of systems and software*, vol. 81, p. 1525-1538.
- Kim D.H., Song J.W., Lee J.H. (2007). A hybrid region weighting approach for relevance feedback in region-based image search on the web. *Lecture notes in computer science*, vol. 4362, p. 705-715.
- Lamard M., Cazuguel G., Quellec G., Bekri L., Roux C., Cochener B. (2007). Content based image retrieval based on wavelet transform coefficients distribution. *The annual international conference of the IEEE engineering in medicine and biology society*. Lyon, France, 23-26 Août.
- Li J., Wang J.Z. (2008). Real-time computerized annotation of pictures. *IEEE transactions on pattern analysis and machine intelligence*, vol. 30, n° 6, p. 985-1002.
- Liua R., Wanga Y., Babab T., Masumotob D., Nagatab S. (2008). SVM-based active feedback in image retrieval using clustering and unlabeled data. *Pattern recognition*, vol. 41, p. 2645-2655.
- Paredes R., Deselaers T., Vidal E. (2008). A probabilistic model for user relevance feedback on image retrieval. *The joint workshop on machine learning and multimodal interaction*. Utrecht, 8-10 Septembre, p. 260-271.
- Philipp-Foliguet S., Vieira M.B., Lekkat M., (2005). Image retrieval by fuzzy regions sets matching. *Information interaction intelligence*, vol. 5, n° 2, p. 9-39.
- Piatrik T., Izquierdo E. (2011). Multi-feature fusion in image clustering using ant-inspired methods. *Third orld congress on nature and biologically inspired computing*, Salamanca, 19-21 October, p. 377-382.

- Su Z., Zhang H.J., Li S., Ma S. (2003). Relevance feedback in content-based image retrieval: Bayesian framework feature sub-spaces and progressive learning. *IEEE transactions on image processing*, vol. 12, n° 8, p. 924-937.
- Tian Q., Sebe N., Lew M.S., Loupias E., Huang T.S. (2001). Content-based image retrieval using wavelet-based salient points. *Storage and retrieval for media databases*. San Jose, 24 Janvier, p. 425-436.
- Vannoorenbergh P. (2004). Bouclage de pertinence par arbres de décision crédibilistes en indexation d'images par le contenu. *Compression et représentation des signaux audiovisuels*, Lille, 25- 26 mai.
- Wang J.Z., Li J., Wiederhold G. (2001). SIMPLicity: Semantics-sensitive integrated matching for picture libraries. *IEEE Transactions on pattern analysis and machine intelligence*, vol. 23, n° 9, p. 947-963.
- Xu X., Lee D., Antani S.K., Long L.R., Archibald J.K. (2009). Using relevance feedback with short-term memory for content-based spine X-ray image retrieval. *Neurocomputing*, vol. 71, n° 10, p. 2259-2269.
- Zhang C., Ouyang D., Ning J. (2010). An artificial bee colony approach for clustering. *Expert systems with applications*, vol. 37, p. 4761-4767.
- Zhang Z., Toda H., Kawabata H. (2002). A new complex wavelet: RI-Spline wavelet and its application to signal processing. *The annual conference of society of instrument and control engineers*. Osaka, 5-7 Août, p. 2496-2501.

Abir Gallas est titulaire d'un diplôme d'ingénieur en Génie Logiciel et Systèmes d'Information de l'Institut Supérieur d'Informatique – Université de Tunis El Manar en 2008 et d'un diplôme de mastère de recherche en Génie Logiciel en 2009. Elle prépare actuellement une thèse de doctorat en informatique sur l'indexation et la recherche d'images par combinaison de RBIR et du bouclage de pertinence, au sein de l'équipe SIIVA du laboratoire RIADI.

Walid Barhoumi est titulaire d'un DEA en Informatique de l'Institut National Polytechnique de Toulouse (Juillet 1998) et d'un doctorat en Informatique de l'Ecole Nationale des Sciences de l'Informatique - Université de la Manouba (Janvier 2006). Il est actuellement Maître Assistant à l'Ecole Supérieure de Technologie et d'Informatique - Université de Carthage et membre de l'équipe SIIVA du laboratoire RIADI. Ses recherches portent sur l'indexation et la recherche de l'information multimédia, l'imagerie médicale et la vision par ordinateur.

Ezzeddine Zagrouba est titulaire d'un doctorat en Informatique de l'Institut National Polytechnique de Toulouse en 1994 et d'un HDR en Informatique de l'université de Tunis El Manar en 2005. Il est actuellement Professeur à l'Institut Supérieur d'Informatique - Université de Tunis El Manar et responsable de l'équipe SIIVA du laboratoire RIADI. Ses recherches portent sur l'indexation et la recherche de l'information multimédia, l'extraction des connaissances et la vision par ordinateur.

