

UNIVERSITÉ DU QUÉBEC

MÉMOIRE PRÉSENTÉ À
L'UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES

COMME EXIGENCE PARTIELLE
DE LA MAÎTRISE EN MATHÉMATIQUES ET INFORMATIQUE
APPLIQUÉES

PAR
EMNA MEJRI

SÉLECTION AUTOMATIQUE DE ZOOM EN RECHERCHE D'IMAGES

OCTOBRE 2014

Université du Québec à Trois-Rivières

Service de la bibliothèque

Avertissement

L'auteur de ce mémoire ou de cette thèse a autorisé l'Université du Québec à Trois-Rivières à diffuser, à des fins non lucratives, une copie de son mémoire ou de sa thèse.

Cette diffusion n'entraîne pas une renonciation de la part de l'auteur à ses droits de propriété intellectuelle, incluant le droit d'auteur, sur ce mémoire ou cette thèse. Notamment, la reproduction ou la publication de la totalité ou d'une partie importante de ce mémoire ou de cette thèse requiert son autorisation.

Résumé

En recherche d'images par le contenu, la requête peut être une image au complet ou une partie d'une image (région d'intérêt). Nous allons appeler le premier mode « Recherche globale », et le second « Recherche locale ». Suite à une multitude d'expériences, nous avons constaté que la recherche globale donne de bons résultats pour certaines catégories d'images, alors qu'elle échoue complètement avec d'autres. La même chose peut être dite à propos de la recherche locale. Dans ce mémoire, nous proposons une solution à ce problème. Concrètement, nous proposons un algorithme qui sélectionne, en fonction de la situation qu'il commence par détecter, s'il faut procéder par recherche locale ou globale. La situation est décrite par le type d'image requête et le contenu de la BD, ce qui permet de prendre en compte les préférences de l'utilisateur. Notre algorithme commence par une étape d'apprentissage, suite à laquelle, il classe le reste de la BD en des classes locales et classes globales; pour finalement exploiter cela au niveau de la recherche. Les évaluations qui ont été faites sur une collection diversifiée d'images démontrent que notre algorithme contribue à améliorer sensiblement la précision des résultats et à mieux capter les besoins de l'utilisateur.

Remerciements

À mes directeurs de mémoire, Messieurs Kherfi Mohamed Lamine et Mesfioui Mhamed, je vous remercie vivement pour votre aide dans la réalisation de ce travail et je suis vraiment très honorée d'avoir collaboré avec vous. J'apprécie beaucoup votre modestie, votre disponibilité et vos grandes qualités scientifiques et humaines. J'espère que ce travail vous satisfera et reflétera ma haute gratitude.

J'exprime aussi ma reconnaissance à toute ma famille, à mes parents pour leur soutien et leur encouragement sans faille. Ce travail n'aurait pu aboutir sans l'aide et la patience de mon mari qui m'a énormément soutenu pendant ces années. Ces remerciements ne seraient pas complets sans une pensée pour mon petit fils SKANDER qui a constitué ma source de motivation et d'espoir.

Je remercie également tous les professeurs ainsi que tous les membres du département de mathématiques et d'informatique.

Table des matières

Résumé.....	ii
Remerciements.....	iii
Table des matières.....	iv
Liste des figures	vi
Chapitre 1 - Introduction.....	7
1.1 Mise en contexte.....	7
1.2 Recherche d'images.....	8
1.2.1 Principe de fonctionnement d'un système de recherche d'images.....	8
1.2.2 Types de moteurs de recherche d'images	9
1.3 État de l'art	11
1.3.1 Exemple des travaux en recherche globale :.....	11
1.3.2 Exemple des travaux en recherche Locale.....	13
1.4 Choix de Zoom.....	15
1.4.1 Introduction.....	15
1.4.2 Quelques notions.....	15

1.4.3	Détails de la problématique	17
1.4.4	Solution proposée.....	25
1.5	Conclusion.....	25
Chapitre 2 - Article.....		26
Chapitre 3 - Conclusion		44
Bibliographies		48

Liste des figures

Figure 1: Fonctionnement d'un moteur de recherche.	8
Figure 2: Illustration de la subjectivité de l'annotation.	9
Figure 3: Exemple de résultats de recherche par texte.....	10
Figure 4: Aigle	10
Figure 5: Illustration de segmentation en régions.	17
Figure 6: Chevaux	18
Figure 7: Roses.....	18
Figure 8: Exemple d'une requête où la recherche globale est plus appropriée.....	19
Figure 9: Exemple d'une requête où la recherche locale est requise.	19
Figure 10: Exemple d'une requête où la recherche locale est la plus appropriée.	20
Figure 11: Exemple d'une requête qui se prête plus à une recherche globale.	20
Figure 12: Exemple d'une requête où la recherche globale donne de bons résultats.....	22
Figure 13: Exemple d'une requête où la recherche globale n'est pas suffisante.	24

Chapitre 1 - Introduction

1.1 Mise en contexte

De nos jours, les collections électroniques d'images sont nombreuses et diverses. Ceci est dû à la progression des dispositifs de stockage, à l'augmentation de leur capacité ainsi qu'à la baisse des prix des dispositifs d'acquisition tels que les appareils photos, les caméras, les téléphones cellulaires, etc. En outre, le partage des images numériques est de plus en plus facile et rapide. En effet, une personne peut partager ses images personnelles avec ses amis à l'aide d'une clé USB, d'un disque dur externe, par courrier électronique, via une panoplie de réseaux sociaux tels que Facebook ou Twitter, ou bien avec Microsoft Skydrive. Cette énorme progression technologique a permis d'augmenter le nombre d'images sur le web, dans les bases de données personnelles et professionnelles.

Prenons l'exemple du nombre d'images partagées sur Facebook. Selon Planetoscope, toutes les secondes plus de 2 263 photos sont mises en ligne sur Facebook, soit 2 716 000 photos toutes les 20 minutes et plus de 71 milliards photos par an.

Cette disponibilité d'information a permis d'ouvrir un nouveau champ d'investigation qui s'intéresse à organiser cette quantité d'information de façon automatique et à localiser les images désirées en un temps raisonnable. D'où l'apparition des moteurs de recherche d'images, appelés également systèmes de recherche d'images. Plusieurs chercheurs, dont des chercheurs de notre équipe, se sont intéressés à cette question au cours des dernières années, ce qui a donné naissance à un certain nombre de moteurs. Finalement, notons que divers domaines s'intéressent à la recherche d'images, dont la médecine, l'architecture, la géographie et les sciences du territoire, la sécurité, l'édition, la numérisation des livres anciens, la mode, etc.

1.2 Recherche d'images

1.2.1 Principe de fonctionnement d'un système de recherche d'images

Le moteur donne la main à l'utilisateur pour qu'il entre sa requête. Ce dernier formule sa requête soit en saisissant un texte contenant un ensemble de mots clés qui décrivent ce qu'il cherche, soit en choisissant une ou plusieurs image(s) exemple parmi un échantillon d'images exposées. Ensuite, le moteur effectue sa recherche. Il compare cette requête avec les images de la base de données, puis retourne à l'utilisateur les images qui correspondent le plus à sa requête. Ceci est illustré dans Figure 1.

Pour pouvoir effectuer la recherche, le moteur procède comme suit :

- ✓ Extraction de caractéristiques.
- ✓ Calcul de la similarité (ex. distance) entre la requête et les images de la BD.
- ✓ Tri des images selon la similarité.
- ✓ Affichage des images résultantes

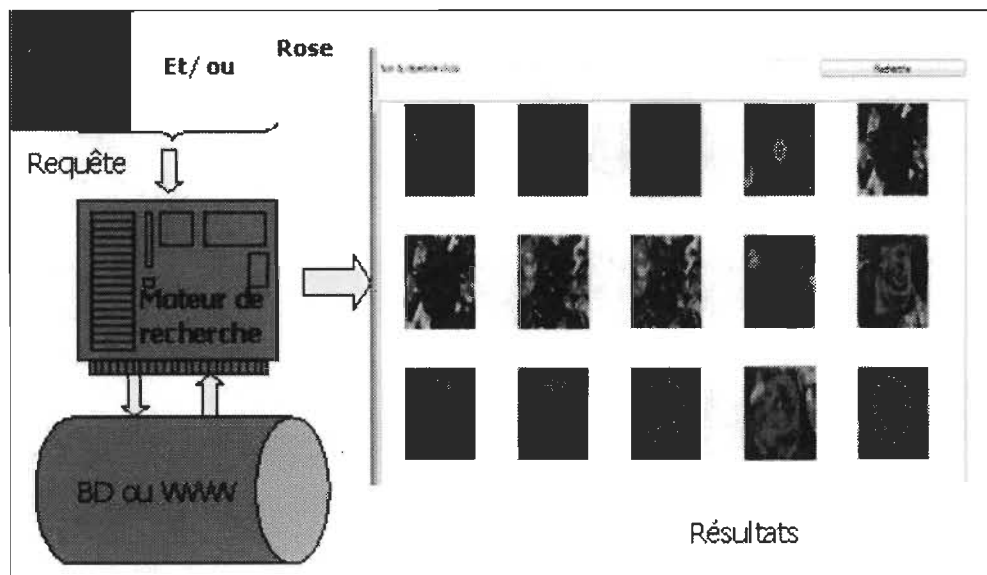


Figure 1: Fonctionnement d'un moteur de recherche.

1.2.2 Types de moteurs de recherche d'images

1.2.2.1 Moteurs basés sur le texte

Le moteur de recherche donne la main à l'utilisateur pour saisir une description textuelle de ce qu'il cherche. L'utilisateur peut alors saisir une liste de mots clés connectés avec des connecteurs logiques, ou bien il peut exprimer ses besoins d'une façon naturelle avec une phrase simple.

La pratique actuelle d'indexation des images repose en grande partie sur les descripteurs de textes ou codes de classification. En dépit de sa familiarité, la requête textuelle a des limites, dont on peut citer :

- ✓ La dépendance du texte vis-à-vis de la langue.
- ✓ L'impossibilité de l'utiliser quand il n'y a pas de texte qui accompagne les images.
- ✓ La subjectivité d'annotation des images. Exemple : L'image de la figure 2 peut être annotée avec des mots-clés différents par des utilisateurs différents, dépendamment de l'intérêt de chacun. Ainsi, elle peut être annotée avec : Manhattan, New York, Skyline, Gratte-ciel, Mer, Navire, Ciel, Pont, Arbres, Urbanisme.
- ✓ La subjectivité de texte figurant parfois sur l'image.



Figure 2: Illustration de la subjectivité de l'annotation.

- ✓ La recherche textuelle est incapable de répondre à certains besoins de l'utilisateur, surtout s'il cherche des images avec un contenu bien précis et des détails bien spécifiques. Voici un exemple :

Considérons le résultat de recherche par texte, que donne Google, pour le mot clé «aigle», présenté dans la figure 3.



Figure 3: Exemple de résultats de recherche par texte.

C'est vrai que la plupart des images résultantes contiennent des aigles. Cependant, si l'utilisateur avait besoins d'images qui rassemblent à celle de la figure 4, serait-il possible de décrire tous ses détails (la position, les couleurs, la montagne, le ciel, etc.) en utilisant du texte? La réponse est évidemment « Non ». Nous verrons plus tard que la recherche par image exemple permet de surmonter cet obstacle; effectivement, « Une image vaut mille mots ».



Figure 4: Aigle

1.2.2.2 Moteurs basés sur le contenu des images

Cette méthode est apparue pour résoudre les limitations de la recherche par le texte précédemment annoncées. Elle consiste à permettre à l'utilisateur de formuler sa requête en utilisant des images et à trouver des images de la base de données qui leurs ressemblent. Il s'agit d'une recherche d'images par le contenu ou CBIR (*Content-Based Image Retrieval*). Dans le cadre de notre recherche, nous nous sommes intéressés à la recherche par le contenu, et plus particulièrement à la «**Recherche par image exemple**».

1.3 État de l'art

Plusieurs chercheurs se sont intéressés à la recherche d'images par le contenu. Les premiers travaux se sont penchés principalement sur la recherche par image globale. Pour ce faire, ces travaux utilisent les signatures (caractéristiques) de tous les pixels de chaque image de la base de données d'images. Plus récemment, les recherches se sont orientées de plus en plus vers la recherche d'images par région d'intérêt. Cette technique se base sur les caractéristiques des régions qui composent l'image. Dans la section 1.3.1 qui suit, nous allons présenter quelques moteurs qui font la recherche par image globale. Ensuite, nous présenterons dans la section 1.3.2 des moteurs qui se sont intéressés à la recherche par régions d'intérêt.

1.3.1 Exemple des travaux en recherche globale :

Lin et al. [1] ont noté que certaines images sont distinguables par leurs caractéristiques couleur et texture, alors que d'autres sont identifiables par leurs caractéristiques couleurs et spatiales. Afin de résoudre le problème de la diversité des bases de données d'images, ils proposaient trois caractéristiques pour la recherche :

1. La matrice de co-occurrence des couleurs (CCM),
2. La différence entre les pixels des motifs (DBPSP),
3. L'histogramme de couleurs K-moyennes (CHKM).

Pour améliorer la détection d'images et pour simplifier le calcul, ils utilisent une technique de sélection préalable des caractéristiques.

Dans une autre étude, Saad et al. [2] utilisaient à leur tour l'information couleur et l'information forme pour la recherche d'images. Ils exploitaient l'histogramme des couleurs CBYR comme caractéristique globale et les descripteurs de Fourier comme caractéristique locale. Ils ont choisi comme mesure de similarité la distance de Bhattacharyya pour le descripteur de Fourier, tandis que pour des histogrammes CBYR, ils calculaient l'intersection des histogrammes. Ils se sont attaqués au problème des déformations géométriques et de bruit, mais le problème d'irrégularité des résultats persistait toujours.

Agarwal et al [3] proposent un algorithme pour la recherche d'images par le contenu basé sur la transformée en ondelettes discrète (DWT) et sur l'histogramme des bordures (EHD). Leur méthode consiste, en premier lieu, à décomposer l'image d'entrée en coefficients d'ondelettes. Ces coefficients d'ondelettes donnent des caractéristiques horizontales, verticales et diagonales de l'image. En second lieu, ils utilisent l'histogramme des bordures sur les coefficients d'ondelettes sélectionnés afin d'assembler les informations d'orientations des bordures dominantes.

Ainsi et après avoir calculé l'histogramme précédemment décrit, pour chaque image de la base de données, ils les rassemblent tous dans une base de données de caractéristiques.

Ils prétendent que la combinaison de DWT et des techniques des EHD augmente la performance du système de recherche d'image basé sur la forme et la texture. Mais ils se limitent à des bases de données ayant une certaine particularité. En effet, ils excluent les bases de données basées sur les descriptifs couleur. En outre, ils ne traitent pas les cas où il y a une multitude de bords.

Dans un autre travail, Shrivastava et al. [4] présentent une nouvelle technique de récupération des images similaires en deux étapes :

1. Les images sont d'abord extraites en fonction de leur similitude de point de vu couleurs,
2. La pertinence des images extraites est encore améliorée par leurs caractéristiques de texture et de forme.

Ils proposent une architecture à trois couches : où chaque étage de sortie est l'entrée pour la prochaine étape.

Première étape : extraction des caractéristiques couleurs de toute la BD + choix de la requête + calcul de la similarité + retour des résultats.

Deuxième étape : extraction des caractéristiques texture des images retournées à l'étape 1 + calcul de la similarité + retour des résultats.

Troisième étape : extraction des caractéristiques forme des images retournées à l'étape 2 + calcul de la similarité + retour des résultats.

Selon eux, cette méthode élimine la dépendance de la précision de la technique de segmentation utilisée par la réduction de l'intervalle de recherche à chaque étape. Elle réduit le problème de la dimensionnalité des vecteurs de caractéristiques ainsi que le fossé sémantique. Cependant, les résultats d'une étape donnée peuvent contenir des images de trop ou se limiter à un champ restreint d'images. Par conséquent, la recherche de l'étape suivante risque de retourner des images indésirables, alors que les bonnes images qui n'ont pas été retournées à l'étape précédente seront complètement écartées.

1.3.2 Exemple des travaux en recherche Locale

D'autres chercheurs se sont penchés sur la recherche par région d'intérêt.

Premchaiswad et al. [5] présentent une méthode de détection non supervisée pour déterminer automatiquement la région d'intérêt dans des scènes naturelles. Ils commencent

par construire un contour fermé qui définit la zone de la région d'intérêt, ensuite ils rapetissent le contour et déterminaient la région en utilisant la méthode des contours actifs. Cependant, une scène naturelle peut convenir mieux à une recherche globale. C'est souvent le cas, surtout si cette dernière contient plusieurs couleurs et ou plusieurs régions.

Vimina et al. [6] définissait les régions d'intérêt par la segmentation de l'image en partitions fixes, ils détectaient la carte de bord à l'aide des bords de filtre de Sobel, puis, ils remplissaient les trous dans la carte de bord, afin de fixer un seuil pour la détermination de la région voulue. Les caractéristiques de couleur et de texture de la région d'intérêt sont calculées à partir des histogrammes HSV et les niveaux de gris de la matrice de cooccurrence. La mesure de similarité utilisée est la distance euclidienne. Cette méthode n'est pas applicable pour les images complexes (ex. paysage). En effet, ces dernières ne possèdent pas de contours fermés à détecter, et par conséquent c'est presque impossible de les segmenter de cette manière.

Feng et al. [7] s'intéressent à leur tour à la recherche locale d'images par le contenu. Ils présentent une nouvelle méthode de retour de pertinence pour résoudre le problème de la CBIR locale.

Rudinac et al. [8] proposaient trois stratégies de recherche régionale. Ces auteurs affirment donc que la recherche globale donne des meilleurs résultats dans certaines situations et que la recherche locale orientée par l'utilisateur est plus appropriée dans d'autres situations.

Suite à une investigation poussée que nous avons menée, nous avons constaté que dans certains cas la recherche globale donne les meilleurs résultats, alors que dans d'autres c'est la recherche locale qui est la plus précise. On se trouve donc dans une situation où le choix entre recherche locale et globale s'impose. La sélection automatique est donc la meilleure solution.

1.4 Choix de Zoom

1.4.1 Introduction

Afin d'analyser la précision des résultats rapportés par les moteurs de recherche existants, il a fallu faire plusieurs expériences sur les moteurs de recherche d'images développés par notre équipe et ceux développés par d'autres chercheurs. Nous avons constaté que la précision de leurs résultats reste très discutable, dans le sens où :

- ils peuvent donner de très bons résultats avec une requête donnée comme ils peuvent présenter des résultats médiocres avec une autre requête,
- ils peuvent réussir à bien organiser une collection d'images donnée mais échouer complètement avec une autre collection,
- etc.

Suite à une analyse approfondie, nous avons découvert que cette irrégularité est essentiellement due au choix des caractéristiques, des mesures de similarité, du choix de la recherche locale versus la recherche globale (ou choix du zoom), etc.

Pour permettre au lecteur de bien comprendre la problématique, nous allons commencer par expliquer les notions de caractéristiques, de mesures de similarité, et de zoom de recherche; ensuite, nous parlerons de l'effet de la sélection de ces ingrédients sur la précision de la recherche.

1.4.2 Quelques notions

1.4.2.1 Notion de caractéristique (ou descripteur)

Une étape de prime importance dans la représentation du contenu des images est l'extraction des caractéristiques visuelles de ces dernières.

Ces caractéristiques peuvent être classifiées en celles qui décrivent :

- La couleur : incluant les histogrammes, les moments, etc.
- La texture : incluant la matrice de cooccurrence, le filtre de Gabor, l'autocovariance, etc.

- La forme : incluant les moments invariants, les descripteurs de Fourier, les points de contour, etc.

Les caractéristiques doivent être extraites *offline*, puis enregistrées dans un fichier créé à cet effet. Cette étape permettra d'accélérer énormément la recherche. Évidemment, la recherche reviendra tout simplement à comparer le vecteur descripteur de la requête et celle de chaque image de la BD.

1.4.2.2 Notion de mesure de similarité

La mesure de similarité est le critère sur lequel se base la comparaison entre le vecteur descripteur de la requête et le vecteur correspondant à chaque image de la BD. On peut citer la distance Euclidienne, la distance L1, ou encore la distance de Mahalanobis.

1.4.2.3 Notion de zoom de recherche

Parfois l'utilisateur est intéressé par l'image requête au complet et parfois, il est intéressé par une partie (région ou objet) de cette image seulement. D'où l'apparition de la notion de la recherche globale et de la recherche locale. Voici comment procède chacun de ces modes de recherche :

✓ Recherche globale :

- Le moteur commence par extraire un ensemble de caractéristiques de chaque image de la BD.
- Ensuite, la recherche consiste à comparer le vecteur descripteur de la requête et celui de chaque image de la BD.

✓ Recherche locale :

- Le moteur commence par segmenter chaque image de la BD en un ensemble de régions, comme illustré dans Figure 5.

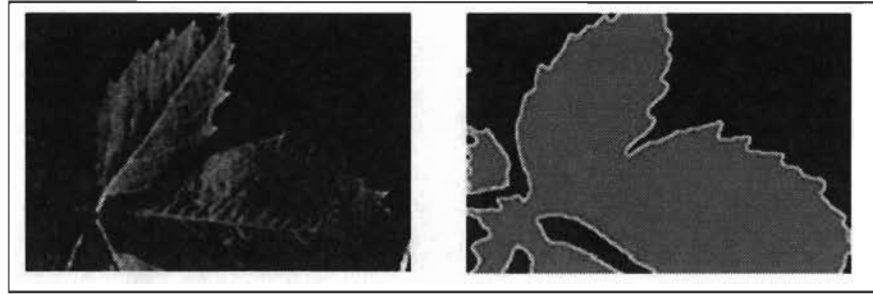


Figure 5: Illustration de segmentation en régions.

- Ensuite, il extrait le vecteur descripteur de chaque région.
- Finalement, la recherche consiste à comparer le vecteur descripteur de la région requête et celui de chaque région de chaque image de la BD (après avoir segmenté chaque image de la BD en régions).

1.4.3 Détails de la problématique

Comme nous l'avons dit auparavant, nous avons procédé à une analyse approfondie du problème, qui nous a révélé que les principales causes de l'irrégularité de la précision des moteurs de recherche d'images par le contenu sont : le choix des descripteurs, le choix des mesures de similarité, et le choix du zoom. Nous allons détailler ces points.

1.4.3.1 Le choix de descripteurs

L'efficacité des descripteurs ou caractéristiques peut être très variable. En effet, un descripteur peut fonctionner bien pour un ensemble d'images et ne pas donner de bons résultats pour d'autres ensembles. Par exemple, si on considère la couleur comme descripteur, les images de la classe Chevaux (Figure 6) se ressemblent, alors que celles de la classe Rose (Figure 7) ne se ressemblent pas. Ce descripteur est donc indispensable pour détecter la ressemblance entre les images de chevaux alors qu'il doit être utilisé avec précaution dans la classe des roses.



Figure 6: Chevaux

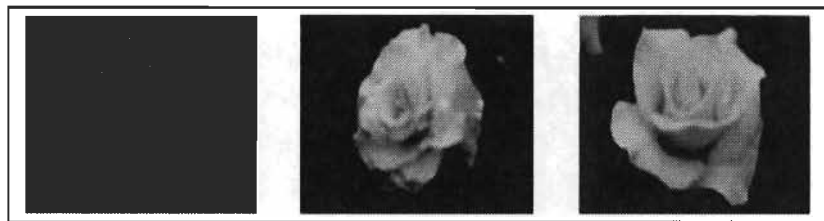


Figure 7: Roses

1.4.3.2 Le choix des mesures de similarité

Le choix des mesures de similarité est semblable au choix des caractéristiques. En effet, certaines mesures de similarité fonctionnent dans certaines situations, alors que d'autres mesures fonctionnent mieux dans d'autres situations. Il s'agit donc de détecter et utiliser les mesures qui conviennent à chaque situation.

1.4.3.3 Le choix du zoom

Notre travail dans ce mémoire se focalise justement sur le choix du zoom, donc nous allons expliquer cet élément en détail. Mentionnons d'abord que le zoom peut être choisi explicitement par l'utilisateur, comme il peut être choisi implicitement par le moteur. Dans ce qui suit, nous allons expliquer chacun de ces deux scénarios.

1.4.3.3.1 Choix explicite du zoom par l'utilisateur

Dans certaines situations, l'utilisateur s'intéresse à tout le contenu de son image requête, c'est-à-dire qu'il cherche des images contenant tous les objets et détails de son image requête, ex. l'image requête donnée dans la figure 8. Dans ce cas, les descripteurs globaux, ex. l'histogramme de la couleur de toute l'image, seront le bon choix. Dans d'autres

situations par contre, il ne s'intéresse qu'à une partie de l'image. La figure 9 par exemple illustre le cas d'un utilisateur qui cherche à trouver toutes les images contenant des pièces de monnaie de cette couleur et cette forme, et ce, quel que soit la couleur du fond. Dans ce cas, seuls les descripteurs locaux de la région qui l'intéresse (la pièce de monnaie) doivent être utilisés, ex. histogramme de la couleur de la région d'intérêt.



Figure 8: Exemple d'une requête où la recherche globale est plus appropriée.

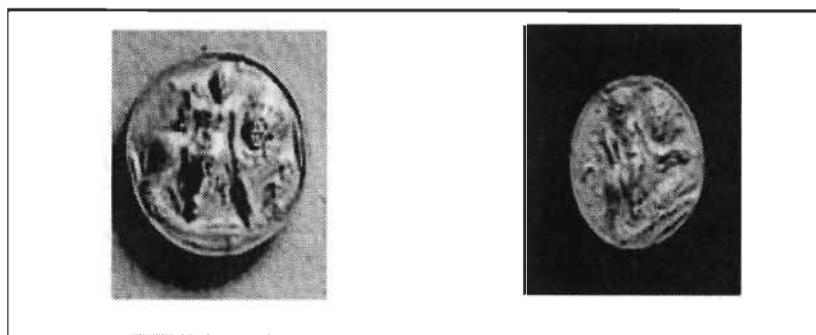


Figure 9: Exemple d'une requête où la recherche locale est requise.

1.4.3.3.2 Choix implicite du zoom par le moteur

Dans certaines situations, et bien que l'utilisateur ait choisi d'effectuer une recherche par l'image globale, il se peut que la recherche locale soit la plus appropriée.

Prenons comme requête l'exemple de la citrouille de Figure 10. On peut constater que les descripteurs locaux de la citrouille répondent mieux à cette requête que les descripteurs globaux de l'image au complet. En effet, bien qu'il ne l'ait pas précisé explicitement, le moteur doit supposer que l'utilisateur cherche des citrouilles, et ce, quel que soit le *'background'*, ce qui a du sens. Il faut donc que le moteur choisisse AUTOMATIQUEMENT les descripteurs locaux à la place des descripteurs globaux.

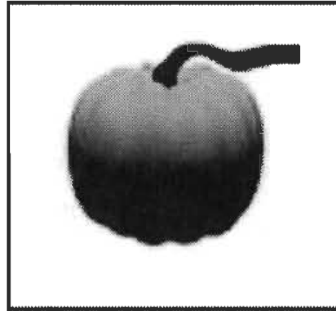


Figure 10: Exemple d'une requête où la recherche locale est la plus appropriée.

Prenons maintenant l'exemple de la requête de la figure 11, image plus complexe qui contient plusieurs objets, avec une disposition bien particulière, etc. Cette requête se prête donc plus à une recherche par les descripteurs globaux.

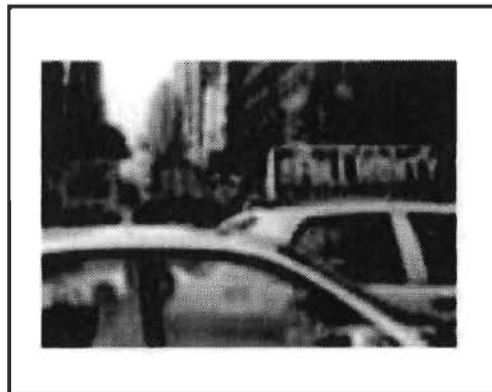


Figure 11: Exemple d'une requête qui se prête plus à une recherche globale.

Nous avons confirmé cette constatation expérimentalement, et ce, en effectuant des recherches avec un moteur de recherche qui a été développé par des membres de notre équipe de recherche, et qui n'utilise que la recherche par descripteurs globaux.

Dans l'exemple présenté à la Figure 12, nous avons choisi une image de papillon comme requête. Voir l'image encadrée en rouge dans la partie gauche de cette figure. Dans ce cas, nous avons obtenu les résultats présentés à la partie droite de la même figure. Nous constatons que les résultats de cette recherche sont parfaits, ce qui nous fait conclure que les images de cette famille se prêtent mieux à la recherche par descripteurs globaux.

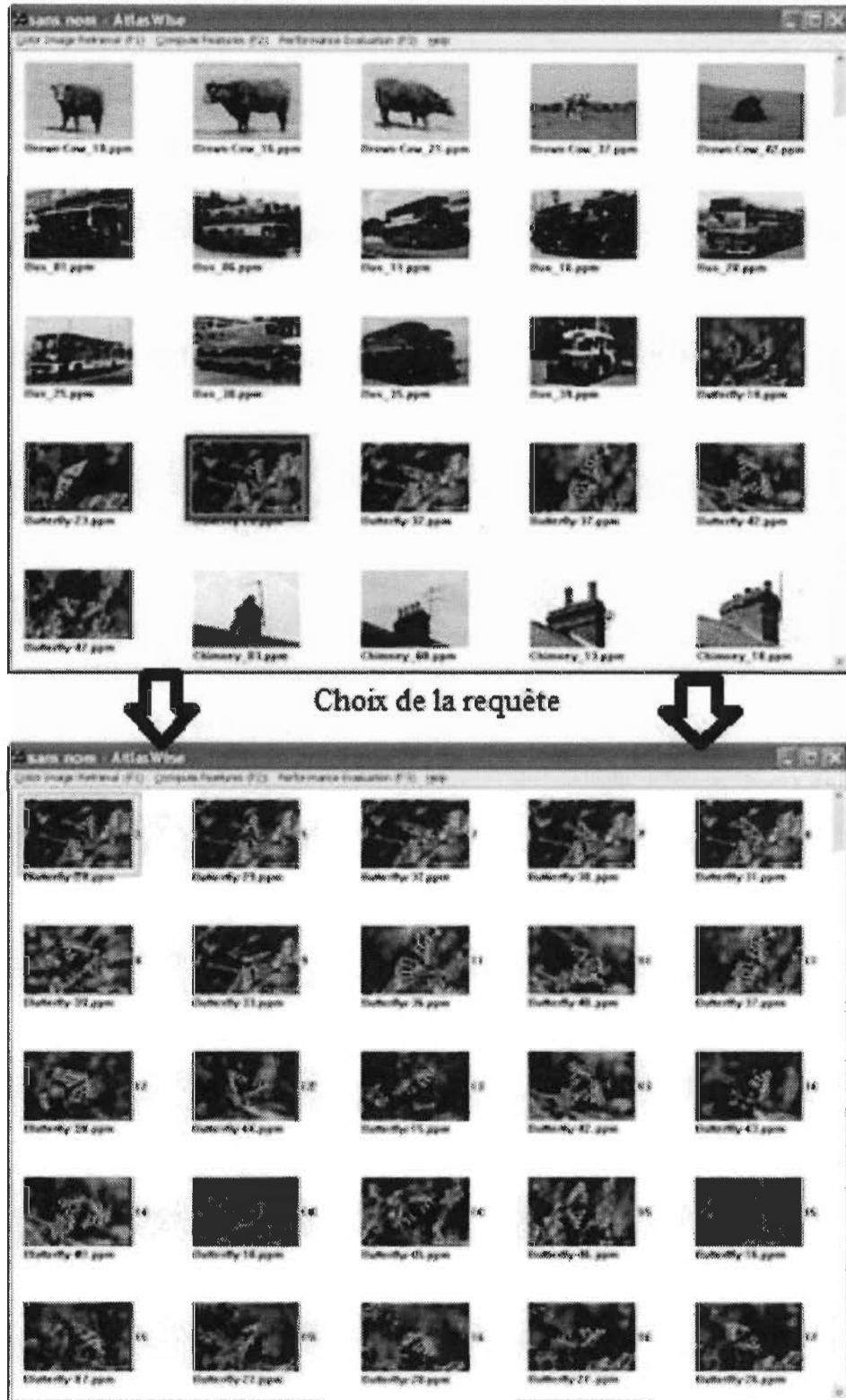


Figure 12: Exemple d'une requête où la recherche globale donne de bons résultats.

Dans l'exemple présenté à la figure 13, nous avons choisi une vache comme requête, voir l'image encadrée en rouge dans la partie gauche de cette figure. Nous avons obtenu les résultats présentés à la partie droite de la même figure. On constate que les résultats de cette recherche ne sont pas satisfaisants. En effet, le moteur ne se limite pas à la vache brune dans l'image requête, mais il considère également l'herbe et la verdure, puis il retourne des images à dominance verte. Ceci nous fait conclure que les images de cette famille se prêtent mieux à la recherche par descripteurs locaux.

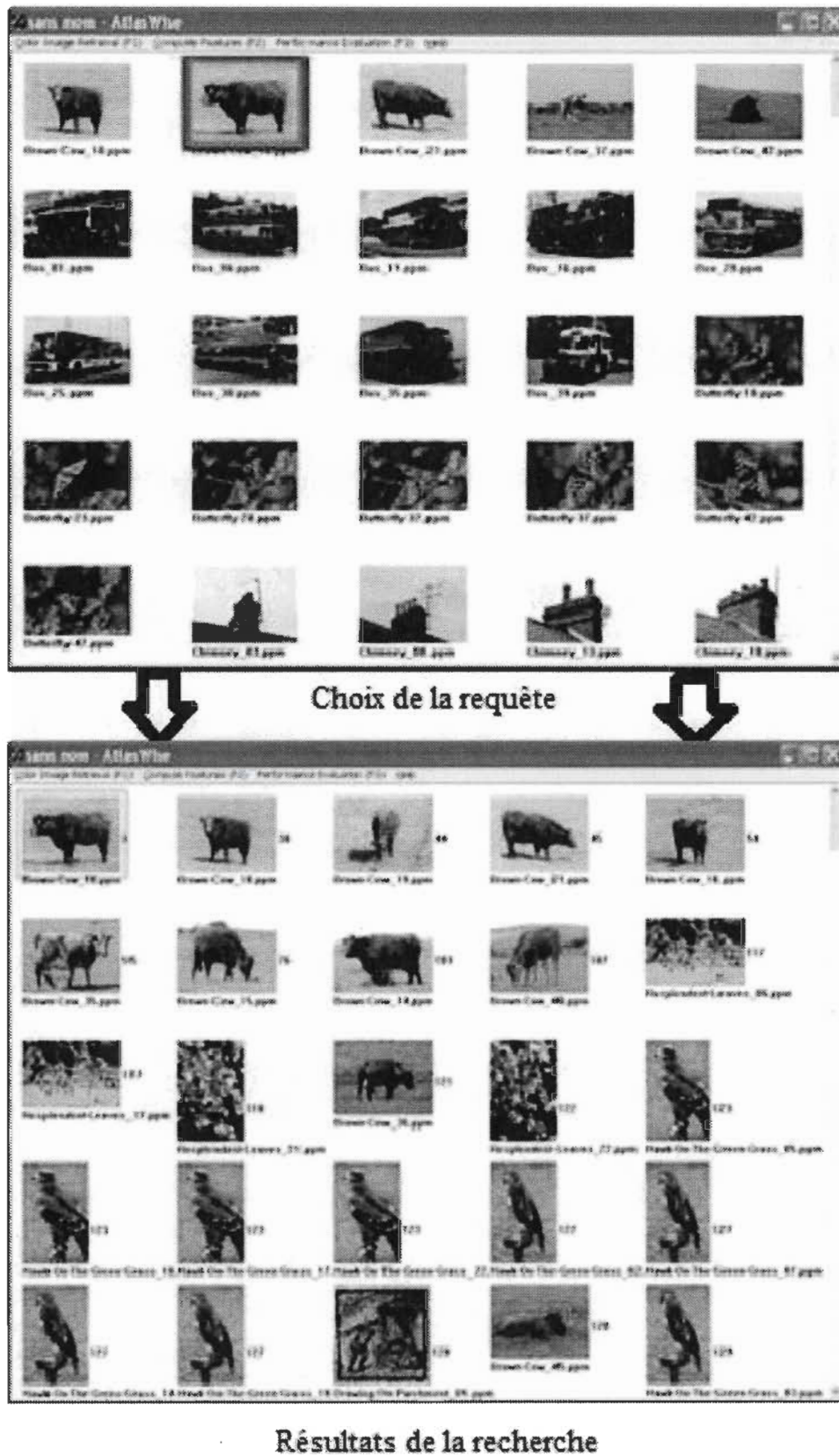


Figure 13: Exemple d'une requête où la recherche globale n'est pas suffisante.

1.4.4 *Solution proposée*

- a. Identifier la situation dans laquelle le moteur se trouve :
 - Le contenu de l'image requête.
 - Les préférences de l'utilisateur.
 - La constitution de la BD.
 - Etc.

- b. À la lumière de la situation, déterminer s'il faut utiliser la recherche locale ou bien globale.

1.5 **Conclusion**

Nous avons vu que la disponibilité d'information, due à l'augmentation considérable du nombre d'images que ce soit sur le Web ou dans les collections personnelles, a donné naissance aux moteurs de recherche d'images. Ensuite, nous avons constaté que la précision des moteurs existants est très variable. Dans notre investigation, nous mène à conclure, entre autres, que dans certaines situations, la recherche globale donne les meilleurs résultats, alors que dans d'autres situations, c'est la recherche locale qui est la plus appropriée.

Nous avons décidé donc d'explorer la piste de la configuration automatique du moteur. Nous avons commencé par développer un moteur qui permet aussi bien la recherche par image que la recherche par région. Ensuite, nous avons développé un mécanisme qui permet au moteur de choisir entre les deux modes, et ce, en fonction de la situation dans laquelle il se trouve. Les détails de notre moteur et notre mécanisme de sélection seront exposés dans le chapitre suivant.

Dans le chapitre 2, nous présentons notre article qui décrit ce travail.

Chapitre 2 - Article

Plusieurs moteurs de recherche d'images par le contenu ont été développés durant les dernières années, mais la précision des résultats qu'ils fournissent est très discutable. En plus, le même moteur peut donner de bons résultats avec une requête donnée mais apporter des résultats médiocres avec une autre requête; il peut réussir à bien organiser une collection d'images donnée mais échouer complètement avec une autre collection, etc.

L'un des facteurs principaux qui contribuent à cette irrégularité est le bon choix du Zoom. En effet, en faisant une recherche, l'utilisateur choisit une image requête. Mais la question qui se pose toujours est la suivante : est-ce que l'utilisateur est intéressé par l'image requête dans son intégralité (recherche **globale**) ou bien par une partie particulière de cette image requête (recherche **locale**)?

Dans l'article que nous présentons dans les pages qui suivent, nous allons nous intéresser à la sélection automatique entre recherche locale et recherche globale, ce que nous allons appeler « Sélection du zoom ». Nous y détaillerons la problématique, expliquerons l'idée de la solution préconisée et les détails de l'implémentation, pour finalement présenter une évaluation expérimentale de l'algorithme proposé. Les résultats obtenus illustrent bien la capacité de notre outil à améliorer la précision de la recherche afin de mieux répondre aux besoins de l'utilisateur.

Les pages qui suivent contiennent notre article intitulé « *An Algorithm for Automatic Zoom Selection in CBIR* ». Cet article décrit notre travail sur la sélection automatique du zoom en recherche d'images.

J'ai réalisé, validé et rédigé ce travail, sous la supervision du Professeur Mohammed Lamine Kherfi. En détails :

- J'ai fait une analyse de la problématique de sélection entre recherche par image globale et recherche par région d'intérêt.
- J'ai développé un nouvel algorithme pour la sélection automatique entre ces deux modes de recherche.
- J'ai effectué une évaluation expérimentale du travail développé.
- J'ai rédigé l'article qui décrit ce travail.

Cet article a été soumis à «Pattern Recognition Letters» le 12 juillet 2014.

An Algorithm for Automatic Zoom Selection in Content-Based Image Retrieval

Emna Mejri and Mohammed Lamine Kherfi

Emna.Mejri@uqtr.ca, Kherfi@uqtr.ca

LAMIA Laboratory, Université du Québec à Trois-Rivières, G9A 5H7, (QC) Canada

Abstract

In Content-Based Image Retrieval (CBIR), query may be an entire image or a part of this latter, i.e., region of interest (ROI). The first mode will be called "global search", and the second "local search". A multitude of experiments conducted with both categories of search revealed to us that global search yields good results for some kinds of images, while it completely fails with other kinds. The same can be said about local search. In this paper, we propose a solution to this problem. Concretely, the proposed algorithm selects, depending on the situation it detects beforehand, whether to proceed by local or global search. The situation is described with the query image type and the DB content, which allows taking user's preferences into account. Our algorithm starts with a learning step, after which, the remaining of the DB is clustered into local classes and global ones. Finally, this is used when performing retrieval. Experiments done on a diversified image DB show that the proposed algorithm helps significantly improving the accuracy of results and better meeting the user's needs.

1. Introduction

With the development achieved recently in the technologies of acquisition, storage and sharing of visual document, a large amount of such documents became available on the web and in all kinds of personal and professional databases. In order to take advantage of all this amount of information and make it easily accessible, it is vital to develop tools that automatically organise it, and help users locate the document they look for in a reasonable time. Many researchers have focused on this issue in recent years, which has given rise to a number of image search engines.

An image search engine relies on a number of "ingredients" such as features, similarity measures or query types: by region of interest (ROI) or by entire image. The importance of a given ingredient depends on the situation. A feature that yields good results with a given query may give poor results with another query. The same can be said about similarity measures. This is also true concerning zoom: in some situations, entire-image-based search is more suitable, while in other situations it is region-based search that best suits. This depends on several factors, including user preferences, DB content, and especially query image: a single object on a background, a complex scene, etc.

In this article, we will precisely focus on automatic zoom selection issue. The engine we developed starts by detecting the current situation. After that, it decides, in a way fully transparent to the user, whether to consider the entire chosen image as a query or a part of it only. Furthermore, in the second case, it automatically determines the part that is likely to interest the user.

The paper is organised as follows: Section 2 will give an overview of the state of the art. Then, in Section 3, we will show the importance of performing zoom selection in order to obtain precise results, and expose the problem in detail. Section 4 will explain the steps of the proposed algorithm, then in Section 5, an empirical evaluation of this algorithm is conducted. We terminate the article with a brief conclusion.

2. State of the art

Several researchers worked on the development of entire-image-based search engines in recent years. Examples include [1, 2, 3, 4]. In such engines, first, the user selects one or more query images. Second, the engine traverses the DB, comparing the features of the query with those of every image in the DB. Finally, the images closest to the query are returned.

Other researchers focused on region-based search. Among the engines developed, we can cite [5, 6, 7, 8]. When formulating query, rather than selecting an entire image, user chooses a part of it only. This part generally represents an object or region of interest. During retrieval, the engine attempts to locate images containing a region that resembles the query region.

A multitude of experiments carried out with both families of engines revealed to us that entire-image-based search is more suitable in some situations, whereas region-based search is more appropriate in other situations. However, despite the importance of the automatic selection between the two search modes as we will see in Section 3, little attention has been paid so far to this issue.

Indeed, we conducted a thorough investigation which revealed that the only ingredient researchers tried to select is image features. In this context, several techniques have emerged. For instance, axis reweighting technique, which has been used in [15, 16, 18], gives more importance to features in terms of which query images are concentrated, and less importance to other features. Some techniques attempt to perform feature selection in the case the query comprises both positive and negative example images [9, 10, 11, 12]. They give more importance to features that discriminate well between positive and negative examples and less importance to features that confound them. Other techniques that are worth mentioning include [13, 14, 17].

It is clear that the developing algorithms that automatically select between local and global search has not aroused enough interest. This will precisely be our concern in this paper.

3. Problem position

As we said in previous sections, several CBIR systems have emerged in recent years, some performing retrieval based on entire image while others use region of interest. We carried out a thorough experimentation with both kinds of systems, which revealed that none of them is suitable for all situations. Indeed, we found out that entire-image-based search is more appropriate in some situations like:

- Those where the query image is complex or contains a multitude of objects but none of them is salient compared to the others.
- Those where the query image contains a single object without background, e.g., a uniform color or a particular texture.

Images of Fig. 1 give an illustration.

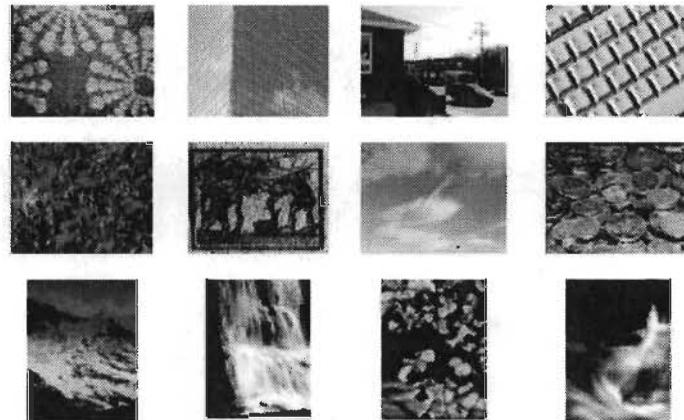


Fig. 1. Some sample images where global search is more suitable.

As for region-based search, it is more appropriate in other situations, like when the query image contains one or two salient objects on a background. For an illustration, see Fig. 2.

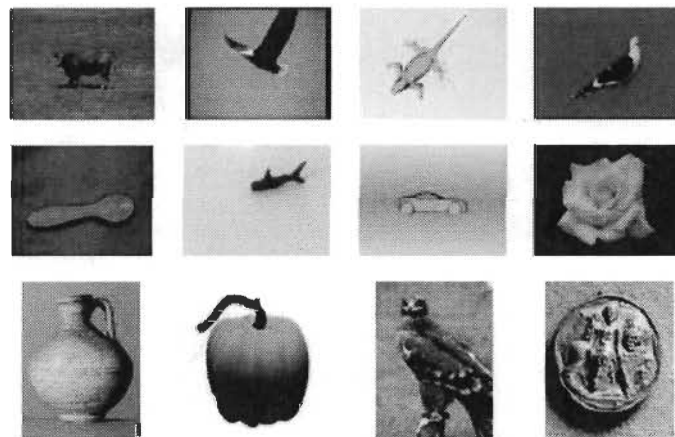


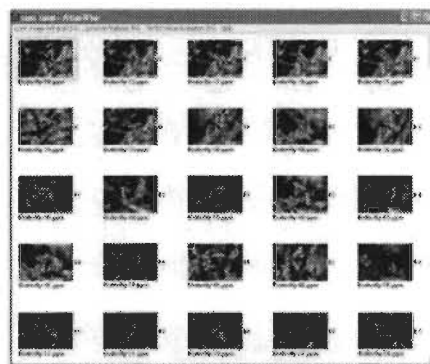
Fig. 2. Some example images where local search is more suitable.

Before continuing our discussion, let us first define two terms that will be used in subsequent sections:

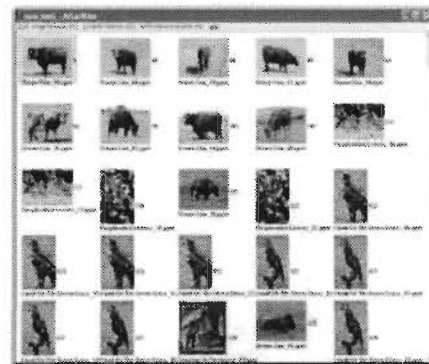
- **Global image:** an image that is more suited for global search, according to our expectations.
- **Local image:** an image that is more suited for local search, according to our expectations.

Now, and in order to illustrate the findings set out above, we give in Fig. 3 the results of four retrieval sessions we performed (Notice that the query image is the one returned at the first position):

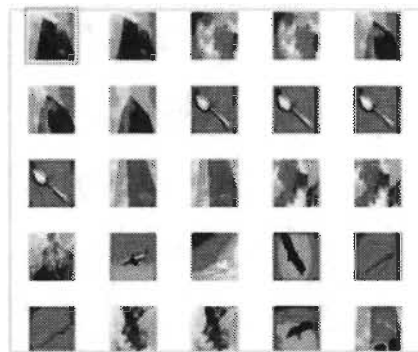
- In the first, the query was a global image, and the engine performed global search. We notice that the results are very good.
- In the second, the query was a local image, whereas the engine performed global search. We notice that the results are rather poor.
- In the third, the query is a global image, whereas the engine performed local search. We notice that the results are rather poor.
- In the fourth, the query is a local image, and the engine performed local search. We notice that the results are very good.



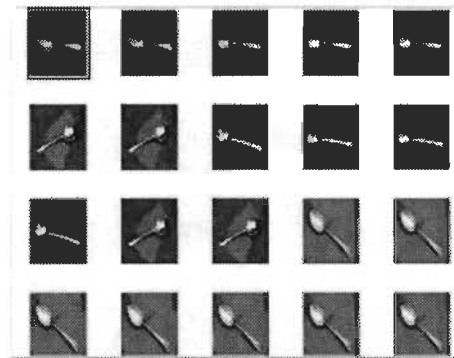
Global query + global features => good results



Local query + global features => bad results



Global query + local features => bad results



Local query + local features => good results

Fig. 3. The need to select the appropriate search mode depending on the query.

Thus, it is clear that, in order to be able to yield good results, the engine should select the right search mode: local or global, according to the situation. This latter strongly depends on the query image, as seen in the previous example. It also depends on the DB content: with a DB where each image contains one or two salient objects on a background, local search is more appropriate. Otherwise, global search is more suitable.

The solution proposed in this paper allows to automatically select between local and global search. This selection is based on the query image type and the DB content. Our selection algorithm has the advantage of understanding the user preferences without asking him explicitly. For example, even if he chooses an entire image as a query, if the algorithm discovers that this image contain a unique object on a background, it will assume that the user is looking for the object not the entire image, and a local search will be performed.

4. Proposed solution

The proposed solution consists in:

- Developing a search engine that allows both retrieval by entire image and by region.
- Providing the engine with a mechanism enabling it to decide, depending on the current situation, whether to conduct a local or a global search.

To do so, the tool we developed proceeds in 5 steps:

1. Segmenting the DB images.
2. Automatically subdividing the DB images into local images and global ones.
3. Classifying a small sample from the DB. This is done by a human administrator.
4. Automatically generalising the classification to the rest of the DB.
5. Making use of this when performing retrieval.

Those steps are detailed in the remainder of the Section. Furthermore, Fig. 4 represents a flowchart of the steps 1 to 4.

4.1. Segmenting the DB images

The first step of our algorithm consists in segmenting each image in our DB into regions. This segmentation is needed for at least two reasons:

1. The algorithm that subdivides the DB into global images and local ones relies on the number of regions by image.
2. Concerning images categorised as local, only the regions of interest should be taken into account and compared with each other.

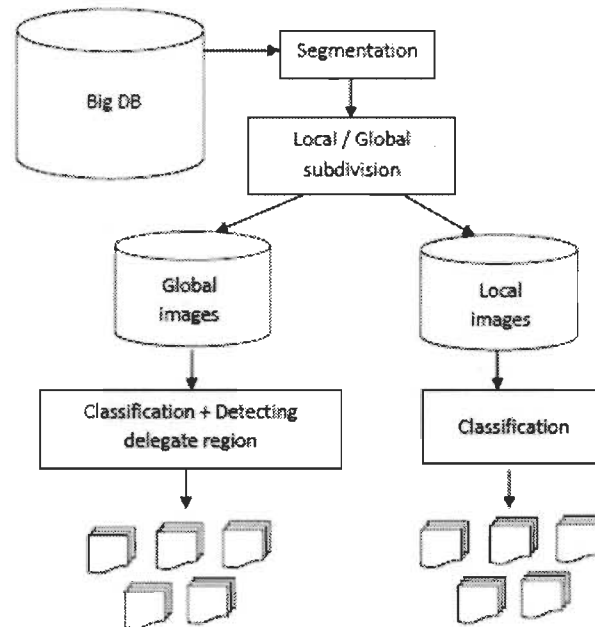


Fig. 4. Flowchart of the classification algorithm.

4.2. Subdividing the DB images into local and global

The aim of this step is to subdivide the DB images into two families: local and global. The following algorithm has been used to do so:

<p>Algorithm 1. Subdividing images into local and global.</p> <p>for each new image j to be categorized, decide whether it is local or global based on the following rules: <small>(see justification below)</small></p> <p style="padding-left: 2em;">If it comprises between 2 and 5 regions then It is considered to be local.</p> <p style="padding-left: 2em;">else (i.e., the number of regions is equal to 1 or greater than 5) It is considered to be global.</p> <p>end for</p>
--

The principle of this algorithm is justified by the following observation:

- Images that are better match local search are those with a small number of salient objects. Such images generally contain between 1 and 4 regions. By adding the background, it gives 2 to 5 regions.
- An image with a single and uniform region (a background, a part of an object, etc.) better suits global search. Similarly, an image with a multitude of regions (more than 5) is typically a complex image containing too much detail, a texture, or overlapping objects. It better suits global search too.

4.3. Classifying a small sample of the DB

The third step consists in asking a human user (or administrator) to classify a small sample of the DB, according to the following steps:

- Choosing, among the global images, a sample that represents well this family.
- Subdividing this sample into a set of classes (said global classes)
- Choosing, among the local images, a sample that represents well this family.
- Subdividing this sample into a set of classes (said local classes)
- Choosing a delegate region that represents each local image. This notion of delegate region is detailed below.

Notion of a delegate region:

When retrieval is based on entire image, it is this latter that is considered as a query. In contrast, when retrieval is based on a region of interest, it is not sufficient to specify the query image, the sought region must be specified too. It is this region that we call delegate region.

4.4. Generalising the classification to the rest of the DB

During this step, our tool classifies the remaining images of the big DB. It assigns each global image to the appropriate global class using Algorithm 2 below. It assigns each local image to the appropriate local class and automatically chooses the delegate region of this image. This is done using Algorithm 3 below.

<p>Algorithm 2. Assigning each global image to the appropriate global class</p> <pre> for each new global image i to be classified for each global class k_g Calculate the membership degree of i to k_g (See how in Note 1 below) end for Assign i to the class having obtained the highest membership degree. end for </pre>
<p>Note 1 : The membership of a global image i to a global class k_g, noted $Mem(i, k_g)$, is calculated as follows:</p> <ul style="list-style-type: none"> - For each image i belonging to k_g : perform a search using i as query. - Calculate the average rank of i among results. - $Mem(i, k_g) = 1 /$ the average rank of i.

<p>Algorithm 3. Assigning each local image to the appropriate local class, then determining its delegate region.</p> <pre> for each new local image i to be classified for each local class k_l : Calculate the membership degree of i to k_l (See how in Note 2 below) end for Assign i to the class having obtained the highest membership degree. Identify the delegate region of i. (See how in Note 3 below) end for </pre>
<p>Note 2 : The membership degree of a local image i to a local class k_l, noted $Mem(i, k_l)$, is calculated as follows:</p> <ul style="list-style-type: none"> - For each image belonging to k_l, perform a search using its delegate region as query. - Calculate the average rank of i among results. - $Mem(i, k_l) = 1 /$ the average rank of i.
<p>Note 3: The delegate region of an image is chosen as follows:</p> <ul style="list-style-type: none"> - In the case the image comprises only 2 regions, a background subtraction algorithm was applied. Then, the remaining region is considered as the delegate of its image. - In the case the image comprises more than 2 regions, the region having most frequently been retained in similarity calculation during the assignment of the image to a class is chosen to be the delegate. In other words, it is the region closest to regions that represent the images of class k_l.

4.5. Making use of the obtained index to perform retrieval

Now that the training and classification algorithms were exposed, Algorithm 4 explains how search is carried out depending on the situation:

<p>Algorithm 4. Retrieval.</p> <p>The user starts by choosing an image q as query, without specifying whether he is interested by a global or a local search.</p> <p>Our engine finds out the class from which comes the chosen image</p> <p>If it comes from a global class k_g, then Search is done within k_g and neighboring global classes, using the global features of q. else (i.e., it comes from a local class k_l) Search is done within k_l and neighboring local classes, using the local features of the delegate region of q. end if</p> <p>The engine returns results to the user.</p>
--

5. Experimental evaluation

In order to evaluate the proposed algorithm, we used a DB of 900 images divided into 30 classes of 30 images each. Our DB contains various images, with different colors and shapes. Obviously, it contains both local and global images. Fig. 5 gives an overview of the content of the used DB.

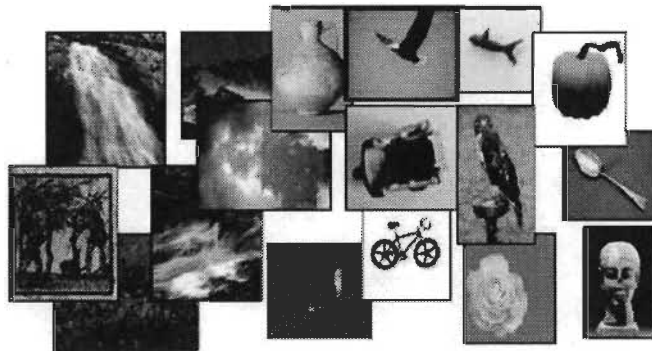


Fig. 5. Some sample images from the DB used.

Each image in the DB was first segmented. Second, our tool chose whether this image is local or global, based on the number of regions it embeds. Third, the administrator classified a sample of the global images and a sample of the local images. Fourth, classification was generalized to the rest of the DB. Finally, this classification has been exploited by the search module. The user starts by choosing an image to constitute a query. Then the search module automatically decides whether to opt for local search or global one.

Since the developed tool performs both classification and search, we carried out two experiments. The first was intended to measure its classification performance while the second was designed to measure its search performance.

5.1. First experiment: evaluating classification

This experiment aimed to evaluate the performance of our classification algorithm. For this purpose, we compared the classification it gave with a classification made by a human.

For each of the classes our algorithm generated, we measured two indicators:

Precision = Number of good images in the class / Total number of images assigned to this class.

Recall = Number of good images the tool put in this class / Total number of images that should have been put in this class.

Results are given in Table 1.

From Table 1, we notice that both Precision and Recall are appreciable high for most classes. Our algorithm achieved an average precision of 93.4% and an average recall of 92.8%. This clearly assesses its ability to automatically classify the DB images with a good accuracy.

TABLE I
PRECISION AND RECALL OF EACH CLASS

Class	Precision	Recall	Class	Precision	Recall
Building	100	86.67	Eagle	92.59	83.33
Waterfall	100	100	Bike	90.91	100
Photo-Frame	100	96.67	Trolley	100	90
Keyboard	100	83.33	Pumpkin	88.89	80
Poppy	100	86.67	Spoon	88.23	100
Dead-Leaves	100	100	Hawk	96	80
Straw	100	90	Onion	93.75	100
House	90.91	100	Bird	100	100
Money	81.08	100	Fish	85.71	100
Mosaic	75	100	White-Rose	100	93.33
Landscape	100	100	Monkey	83.33	100
Beach-At-Night	92.31	80	Statue	84.62	73.33
Bridge	96.55	93.33	Tomato	81.82	90
Scene	100	100	Mud	86.67	86.67
Waves	93.75	100	Walkman	100	90

A closer look into Table I shows that performance varies from class to class. Roughly, four categories of classes can be identified:

- Those with a high precision and a high recall, like "Waterfall".
- Those with a high precision but a moderate (yet acceptable) recall, like "Hawk".
- Those with a high recall but a moderate (yet acceptable) precision, like "Mosaic".
- Those with a moderate precision and a moderate recall, like "Mud".

After analysis, we found that classes with a moderate recall are those whose images are dispersed, i.e., do not resemble each other. Thus, it is quite natural that the algorithm was not able to identify them all. Nevertheless, it succeeded locating most of them, reaching a fairly acceptable recall.

The same can be said about precision. Indeed, classes having obtained a moderate precision are those that are close to other classes. Therefore, some of their images may be attracted by those classes and vice-versa. This explains the fact that such classes contain more noise, which negatively affects their precision. Nevertheless, our algorithm was able to reach a fairly good precision on the whole

5.2. Second experiment: evaluating retrieval

The aim of this experiment was to demonstrate that our automatic-selection algorithm performs better than a simple local search or a simple global search.

To do so, results of the three search modes (local, global and automatic) were compared. Search Precision has been used as a measure of performance. It is given by:

$$\text{Search Precision} = \frac{\text{Number of good images returned to the user}}{\text{Total number of images returned to the user}}$$

The experiment was conducted as follows:

- A multitude of retrieval sessions was launched. At each session, first an image is chosen. Then, global, local and automatic retrieval are done both with this image as query.
- The precision of each search mode was measured as a function of the number of results returned to the user.
- The average precision of each search mode was calculated.

Fig. 6 shows the curves of average precision as a function of scope $Pr=f(Sc)$ for the 3 search modes (local, global and automatic) where *Scope* (*Sc*) is the number of images returned to the user.

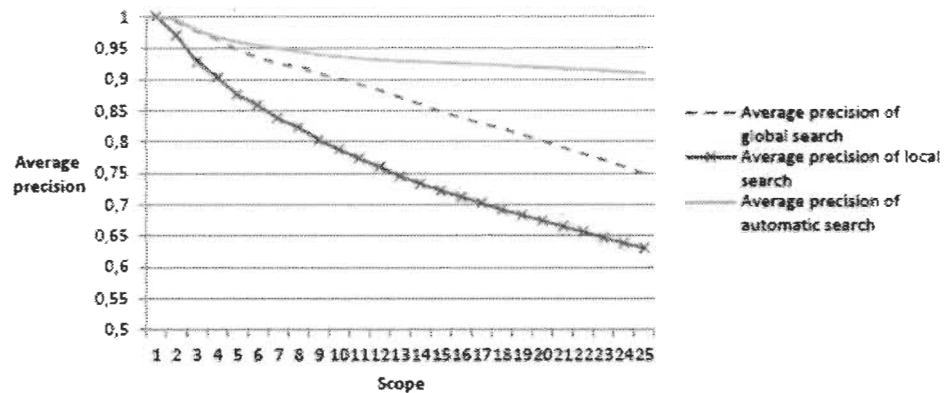


Fig. 6. Comparison of the average Search Precision for the three search modes: local, global and automatic

The first observation we can make from Fig. 6 is that the curve corresponding to automatic mode lies above the two other curves, for all values of the scope. This confirms that the proposed automatic zoom selection algorithm significantly helps improving results precision.

The second observation is that for automatic mode, the relevance of the retrieved images decreases less quickly with the number of retrieved images than when global mode or local mode are used. For a scope of 25 for example, the average precision is about 0.74 with global mode, doesn't exceed 0.63 with local mode, whereas it reaches 0.91 when automatic mode is used.

The third observation is that global search yields better results than local search, for almost all values of the scope. This can be explained by the fact that:

- It is not necessarily a bad thing to carry out global search with a query that better fits local search. Indeed, including the background helps sometimes finding good results, even if it is not the region sought by the user. Imagine for example a user who is looking for planes, and who chooses an image of a plane in the blue sky as query. Global search takes the entire-image features into account, including those of the sky. Therefore, images containing the sky will be retrieved, and there is precisely a good probability that such images contain planes too.
- On the opposite, when local search is adopted with an image that better fits global search, this may have a negative effect on the quality of results because some important parts of the image may be neglected.

6. Conclusion

The main objective of any image search engine is to retrieve results that meet user needs and expectations. However, understanding what this latter wants is not an easy task at all, for several reasons. Sometimes, the user is subjective or doesn't know what he exactly wants. Sometimes, it is hard for him to express his needs using the limited tools available. Sometimes, he doesn't find the right example image to be used as a query.

A careful study of those issues, especially the question of zoom, revealed that user is sometimes interested by the entire query image, while in other situations only a part of this image interests him. This led us to propose a novel algorithm for automatic zoom selection. Depending on the situation it

detects, our algorithm selects whether to perform a local or a global search. First, it partitions the DB into local and global classes. Then, the obtained classification is exploited during retrieval.

The developed engine was validated on a diverse DB. The results are very promising.

Acknowledgments

This research was partially funded by NSERC Canada through its Discovery Grant program.

References

- [1] C.H. Lin, R.T. Chen and Y.K. Chan, "A Smart Content-based Image Retrieval System Based on Color and Texture Feature," *Image and Vision Computing*, 2009 – Elsevier
- [2] M.H. Saad, H.I. Saleh, H. Konbor, M. Ashour, "Image Retrieval Based on Integration Between YCbCr-Color Histogram and Shape Feature," 7th International Conference on Computer Engineering (ICENCO), p.p 97-102, 2011.
- [3] E. Tiakas, D. Rafailidis, A. Dimou, P. Daras, "MSIDX: Multi-Sort Indexing for Efficient Content-based Image Search and Retrieval," *IEEE Transactions on Multimedia*, Issue: 99, p.p 1, 2013.
- [4] N. Shrivastava, V. Tyagi, "Multistage Content-based Image Retrieval," 6th International Conference on Software Engineering (CONSEG), p.p 1-4, 2012.
- [5] W. Premchaiswadi, A. Tungksathan, and N. Premchaiswadi, "Unsupervised Detection for Minimizing a Region of Interest around Distinct Object in Natural Images," *International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, pp.202 – 207, 1-3 Dec. 2010.
- [6] E.R.Vimina, K. Poulouse Jacob, "Image Retrieval using Colour and Texture Features of Regions of Interest," *International Conference on Information Retrieval & Knowledge Management (CAMP)*, pp. 240-243, 13-15 March, 2012.
- [7] S. Feng, C. Lang, and D. Xu, "Localized Content-based Image Retrieval using Saliency-based Graph Learning Framework," 10th IEEE International Conference on Signal Processing (ICSP), pp. 1029-1032, 24-28 Oct. 2010.
- [8] S. Rudinac, M. Uscumlic, M. Rudinac, G. Zajic and B. Reljin, "Global Image Search vs. Regional Search in CBIR Systems," *International Workshop on Image Analysis for Multimedia Interactive Services: IZAMIS*, pp.14, 6-8 June 2007.
- [9] M.L. Kherfi, D. Ziou and A. Bernardi, "Content-Based Image Retrieval Using Positive and Negative Examples," *Journal of Visual Communication and Image Representation*, 14(4):428-437, December 2003.

- [10] M.L. Kherfi and D. Ziou, "Image Retrieval Based on Feature Weighting and Relevance Feedback," *IEEE International Conference on Image Processing*, Singapore, October 2004.
- [11] M.L. Kherfi and D. Ziou, "Image Collection Modeling and its Application to Indexing, Browsing and Semantic Retrieval," *IEEE Transactions on Multimedia*, 9(4):893-900, June 2007.
- [12] M.L. Kherfi and D. Ziou, "Relevance Feedback for CBIR: A New Approach Based on Probabilistic Feature Weighting with Positive and Negative Examples," *IEEE Transactions on Image Processing*, 15(4): 1017-1030, April 2006.
- [13] N. Vasconcelos and M. Vasconcelos, "Scalable Discriminant Feature Selection for Image Retrieval and Recognition," *IEEE International Conference on Computer Vision and Pattern Recognition CVPR*, Washington, D.C., 2004.
- [14] X.S. Zhou, T.S. Huang, "Comparing Discriminating Transformations and SVM for Learning During Multimedia Retrieval," *ACM Multimedia Conference*, Ottawa, Canada, 2001.
- [15] Y. Ishikawa, R. Subramanya, and C. Faloutsos, "MindReader: Querying Databases through Multiple Examples," *24th International Conference on Very Large DataBases VLDB*, 1998.
- [16] Y. Rui and T. Huang, "Optimizing Learning in Image Retrieval," *IEEE International Conference on Computer Vision and Pattern Recognition CVPR*, Hilton Head, South Carolina, 2000.
- [17] Y. Liu, J.R. Kender, "Fast Video Segment Retrieval by Sort-Merge Feature Selection, Boundary Refinement, and Lazy Evaluation," *Computer Vision and Image Understanding*, 91(3), 2003.
- [18] Y. Rui, T. S. Huang, and S. Mehrotra, "Content-based Image Retrieval with Relevance Feedback in MARS," *Proceedings of IEEE International Conference on Image Processing*, Santa Barbara, CA, October 1997.

Chapitre 3 - Conclusion

Nous avons vu que l'augmentation considérable du nombre d'images, que ce soit sur le Web ou dans les collections personnelles, a motivé l'émergence des moteurs de recherche d'images. Ensuite, nous avons constaté que la précision des moteurs existants est très variable. L'investigation que nous avons menée nous a dévoilé que cette précision dépend énormément de la situation dans laquelle le moteur se trouve. Nous nous sommes focalisés sur la question du zoom, et nous avons découvert que dans certaines situations la recherche globale donne de meilleurs résultats, alors que dans d'autres situations, c'est la recherche locale qui est la plus appropriée.

Nous avons donc décidé d'explorer la piste de la configuration automatique du moteur. Nous avons commencé à développer un moteur qui permet aussi bien la recherche par image globale que la recherche par région. Ensuite, nous avons développé un mécanisme qui permet au moteur de choisir entre les deux modes, et ce, en fonction de la situation dans laquelle il se trouve. Notre mécanisme procède par :

1. Une catégorisation des images de la BD entre globales et locales.
2. Une classification des images de chaque catégorie en classes.
3. Ensuite, une fois que l'utilisateur soumet sa requête, l'outil choisit automatiquement entre recherche locale et globale.

Pour ce qui est de la catégorisation des images entre globales et locales, nous avons adopté le principe suivant :

- Si une image contient une seule région
→ une recherche par image globale est plus appropriée.
- Si une image contient un nombre restreint de régions qui de surcroît sont assez apparentes
→ une recherche par région d'intérêt est plus appropriée.
- Si une image est complexe et contient beaucoup de régions dont certaines en cachent d'autres

→ dans cette situation aussi une recherche par image globale est plus convenable.

Pour ce qui est de la classification des images (qu'elles soient locales ou globales) en classes, nous avons développé un algorithme dont les grandes étapes sont les suivantes :

1. Un administrateur commence par classifier un échantillon de notre BD en des classes.
2. Notre outil classifie le reste de la BD en affectant chaque nouvelle image à la classe dont les images lui sont les plus proches.

Finalement, notre algorithme de recherche décide s'il faut procéder par recherche locale ou globale, et ce en se basant sur les résultats des deux étapes précédentes. Nous avons procédé à une évaluation expérimentale des algorithmes que nous avons développés. Pour ce faire, nous avons utilisé une base de données de 900 images réparties sur 30 classes composées de 30 images chacune. En dépit de la variation du contenu visuel de cette dernière, notre moteur a été capable :

- 1) de repêcher les images appartenant à la même classe (**Classification**)
- 2) de retourner les images les plus similaires à la requête choisie par l'utilisateur (**Recherche**).

Ceci a été confirmé par les deux expériences que nous avons conduites :

La première a tenté d'évaluer les performances de notre algorithme de classification. Nous avons comparé la classification qu'il a donnée avec une classification faite par un humain. Notre algorithme a donné une précision moyenne de 93,4% et un rappel moyen de 92,8%, ce qui est très intéressant et démontre clairement la capacité de notre outil à bien classifier les images de la BD. Lors de la deuxième expérience, nous avons essayé de mesurer les performances de notre algorithme de sélection automatique entre recherche locale et globale. Nous l'avons comparé avec les deux algorithmes classiques : celui qui fait toujours la recherche par image globale, et celui qui fait toujours la recherche locale. Les résultats obtenus nous ont permis de constater que notre algorithme de sélection automatique obtient

de meilleurs résultats que les deux autres. Ceci démontre l'utilité et l'importance de la sélection automatique du zoom.

À travers le travail que nous avons accompli, nous pouvons dire que nous avons réussi à :

- Réduire le temps de calcul et par conséquent avoir un moteur plus rapide.
- Réduire l'effet des régions non voulues et qui faussent souvent les résultats.
- Réduire l'espace de recherche en le limitant aux classes les plus proches de la requête. Ceci nous a permis d'augmenter la probabilité de retrouver les images les plus pertinentes.
- Combiner les avantages de la recherche globale avec ceux de la recherche locale.
- Surmonter certaines faiblesses de la recherche locale et d'autres faiblesses de la recherche globale.

Ceci étant, plusieurs autres améliorations peuvent être apportées à notre travail, comme par exemple :

- Trouver des principes de sélection plus élaborés, puis les transformer en des algorithmes.
- Automatiser les tâches qui sont faites actuellement de façon manuelle.
- Intégrer la sélection des caractéristiques, la sélection des mesures de similarités, la sélection du zoom et la sélections de/des objet(s) d'intérêt(s) dans un même moteur; tout en attribuant des priorités ou des pourcentages pour chaque ingrédient selon la situation.
- Avoir un retour de performance (un jugement humain des résultats) suivi par une amélioration des résultats.
- Prendre en compte l'aspect sémantique des images, ce qui devrait permettre de s'approcher davantage des préférences de l'utilisateur et de mieux comprendre ce qu'il veut. L'aspect sémantique repose sur le sens de l'image, ce qu'elle décrit ce qu'elle représente, ce qu'elle contient (une ontologie de l'image).

Selon moi, quelques recherches faites dans le domaine de la science de l'information peuvent être utiles pour décrire les images telles que le thesaurus par exemple. Une bonne solution serait alors de combiner la recherche d'images par texte et la recherche d'images par le contenu visuel. Comme ça, on pourra bien satisfaire les besoins sémantiques et visuels de l'utilisateur.

Bibliographie

[1] C.H. Lin, R.T. Chen and Y.K. Chan, “A smart content-based image retrieval system based on color and texture feature”, *Image and Vision Computing*, Elsevier, Volume 27, Issue 6, p.p 658–665, 4 May 2009,.

[2] M.H. Saad, H.I. Saleh, H. Konbor, M. Ashour, “Image retrieval based on integration between YCbCr color histogram and shape feature”, *Computer Engineering Conference (ICENCO)*, 2011 Seventh International, p.p 97-102, Giza, 2011.

[3] S. Agarwal, A. K. Verma, P. Singh, “Content Based Image Retrieval using Discrete Wavelet Transform and Edge Histogram Descriptor”, *International Conference on Information Systems and Computer Networks (ISCON)*, p.p 19 – 23, Mathura, India, 9-10 March 2013.

[4] N. Shrivastava, V. Tyagi, “Multistage content-based image retrieval”, 2012 CSI Sixth International Conference on Software Engineering (CONSEG), p.p 1-4, Indore, 2012.

[5] W. Premchaiswadi, A. Tungksathan, and N. Premchaiswadi, “Unsupervised Detection for Minimizing a Region of Interest around Distinct Object in Natural Images”, *International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, pp.202 – 207, Sydney, NDW, 1-3 Dec. 2010.

[6] E. R.Vimina, K. Poulouse Jacob, “Image Retrieval using Colour and Texture Features of Regions of Interest ”, *International Conference on Information Retrieval & Knowledge Management (CAMP)*, pp. 240-243, Kuala Lumpur, 13-15 March. 2012.

[7] S. Feng, C. Lang, and D. Xu, “Localized content-based image retrieval using saliency based graph learning framework”, *IEEE 10th International Conference on Signal Processing (ICSP)*, pp. 1029-1032, Beijing, 24-28 Oct. 2010.

[8] S. Rudinac, M. Uscumlic, M. Rudinac, G. Zajic, B. Reljin, “Global Image Search vs. Regional Search in CBIR Systems”, *International Workshop on Image Analysis for Multimedia Interactive Services WIAMIS*, pp.14, Santorini, 6-8 June 2007.