

Indexation et recherche visuo-textuelle des bases de données images

Hadi Akram

LRSD, Département
d'informatique, Faculté des
Sciences, Université Ferhat
Abbas-Sétif 1, Algérie.
akram_ha92@yahoo.fr

Hadi Fairouz

LRSD, Département
d'informatique, Faculté des
Sciences, Université Ferhat
Abbas-Sétif 1, Algérie.
Info_had@yahoo.fr

Boukerram Abdallah

LRSD, Département
d'informatique, Faculté des
Sciences, Université Ferhat
Abbas-Sétif 1, Algérie.
boukerram@hotmail.com

Résumé— La recherche d'images par le contenu est toujours considérée comme une tâche difficile. A cet effet, on utilise des informations textuelles, tel que le nom de l'image, pour tenter de comprendre le sens de l'image. Cependant, le nom de l'image n'est toujours pas en rapport avec celle-ci. Pour améliorer la recherche d'images, la combinaison des informations textuelles et visuelles semble être une méthode prometteuse.

Ce travail porte sur l'étude des différentes méthodes d'indexation de bases de données images, pour proposer ensuite, un système d'indexation visuo-textuelle de bases de données images.

Mots clés- Traitement d'images; indexation textuelle ; indexation visuelle ; indexation visuo-textuelle ; base de donnée image ; distances métriques.

I. INTRODUCTION

L'image est porteuse de deux types d'informations :

- ✓ d'une part, *son contenu* : il est appelé niveau *syntaxique* (ou graphique) qui donne des renseignements sur la scène que représente l'image.
- ✓ d'autre part, son interprétation : elle se nomme niveau *sémantique*, c'est la phase d'interprétation de l'image qui varie d'une personne à une autre en fonction des connaissances de chacun et du contexte d'observation.

Des bases de données ont alors été construites afin de stocker toutes ces images. Au fil des années, les images de natures très diverses ont gagné en importance occupant parfois jusqu'à plusieurs téraoctets de mémoire. Afin de pouvoir consulter et retrouver au mieux les images, de nombreux chercheurs s'intéressent de près cette discipline qu'est, la recherche d'images. C'est un problème ouvert qui met à l'épreuve des thématiques de recherche diverses :

- ✓ Les *bases de données* pour le stockage et l'indexation des images,
- ✓ Le *traitement d'images* pour l'extraction de paramètres discriminants.
- ✓ Les *statistiques*, les *distances métriques* pour l'évaluation de distances entre images.

- ✓ La *classification* pour le classement des images les unes par rapport aux autres.
- ✓ L'aspect *réseau* pour la gestion des bases de données à distance en tenant compte des problèmes de sécurité et de transmission de l'information qui en découlent.

Le but des systèmes de tri automatique (indexation) est de permettre à un utilisateur de trouver, dans des bases d'images, toutes celles qui sont semblables à l'image qui l'intéresse. Un programme d'indexation se conçoit comme un système qui prend en entrée une image de référence et qui retourne un critère de similarité entre l'image de référence et toutes les images de la base. Ceci permet de trier ces images, de la plus similaire à la moins similaire.

On distingue, deux approches principales: l'une emploie des annotations textuelles manuelles et l'autre emploie des descripteurs extraits automatiquement à partir des images. La première approche, basée sur l'annotation textuelle manuelle d'images, est aujourd'hui la plus employée. Mais l'indexation de ces images représente une tâche longue et répétitive pour l'humain, surtout avec les bases d'images qui deviennent de plus en plus grandes, où l'on remarque beaucoup d'erreurs d'indexation, liées au fait que le texte ne correspond pas toujours à l'image.

Pour surmonter les défaites des systèmes de recherche d'images basées sur l'annotation textuelle manuelle, le système de recherche des images basé essentiellement sur le contenu est proposé. Cette deuxième approche a deux directions :

- recherche d'image basée sur le contenu symbolique
- recherche d'image basée sur le contenu sémantique.

Les méthodes symboliques emploient des descripteurs extraits automatiquement de l'image telle que la couleur, la texture, les formes.... Quoiqu' il est difficile de trouver les descripteurs puissants pour représenter des images : à cet effet, des méthodes de recherche sémantique sont apparues pour améliorer le résultat de recherche.

Ce travail présente, dans une première partie, le principe d'indexation textuel de base de données images pour présenter ensuite, l'indexation et recherche d'images par le contenu visuel. En fin, nous développons l'implémentation du système mis en œuvre suivi de la combinaison des deux méthodes décrites précédemment. Une conclusion reprenant les points forts de ce travail et les perspectives futures terminent ce papier.

II. INDEXATION ET RECHERCHE TEXTUELLE D'IMAGES

L'indexation des images par des mots, bien qu'incontournable, reste une gageure en raison de la multiplicité des entités présentes dans l'image. Dès les premières tentatives de description d'une photographie, nous nous trouvons confrontés à l'infinité des éléments qui la composent. Il semble que la charge sémantique d'un mot et la charge sémantique d'une photographie soient incomparables. Définir une image à partir d'un seul terme s'apparente à une véritable réduction de l'image, d'autant plus que le terme qui a un sens pour moi, ne sera pas forcément celui partagé par la majorité. En effet, une image est le support d'une multitude d'interprétations possibles.

Le domaine de recherche d'indexation d'images est de plus en plus actif. La problématique de la recherche d'images se résume en la recherche de mots, et ce, en se basant sur les attributs textuels des images tels que le nom des fichiers par exemple. Cette approche nécessite une entrée manuelle des mots définissant l'image (légende) et ne peut donc être appliquée au flux, toujours croissant, d'arrivée de nouvelles images.

Dans la première génération des systèmes d'indexation, les images sont représentées par des termes sémantiques de type mots-clés comme par exemple, Google, Yahoo.

A. Indexation textuelle manuelle

L'indexation textuelle manuelle d'images est le plus souvent réalisée par un documentaliste appelé iconographe. Son rôle est de classer et d'indexer les images en les associant à des catégories et à des groupes de mots, souvent extraits d'un thésaurus, permettant de retrouver facilement les images. Son travail est très utile pour les agences de presse, les centres de documentation, les musées... Du fait de l'accroissement du nombre de photographies personnelles, ce travail est aussi souvent réalisé par les utilisateurs qui souhaitent décrire leurs images personnelles [1].

L'indexation textuelle se fait en termes d'objets, de contenu et de structure. La sémantique elle-même n'est pas inscrite dans l'image, mais se trouve ailleurs. Il faut donc rechercher ces sources extérieures pour avoir accès aux clés de décodage sémantique de l'image.

Dans une image, la sémantique dépend de deux principes: - du niveau de connaissances et de la perception que possède l'observateur,

- de l'objectif poursuivi par l'utilisateur de cette image lorsqu'il la regarde.

B. Domaines d'application

Les domaines potentiels d'application de l'indexation sont nombreux [2] :

Domaines	Applications
Aérospatiale	Analyse des images satellites
Médecine	Indexation des images représentant des caractères pathologiques
Télécommunications	Archivage et codage d'images par leur contenu
Art	Etude et archivage de styles artistiques Création de musées en ligne
Education	Création d'encyclopédie d'illustrations
Médias	Recherche d'illustration d'un événement ou d'une personne
Surveillance	Identification de personne
Authentification	Recherche et vérification des contrefaçons
Design	Styles et textures dans l'industrie
Photographie	Recherche des photos adéquates pour un article
Audiovisuel	Recherche d'un plan cinématographique

C. Systèmes de recherche existants

Voici une liste non-exhaustive de logiciels existants et présentés sur Internet. Cela montre bien l'intérêt industriel que suscite le problème de l'indexation de grandes bases de données multimédia. Il est très difficile de comparer les résultats de ces produits, pour plusieurs raisons : ils utilisent tous des bases d'images différentes. Ces bases ne sont pas connues dans leur intégralité par l'utilisateur, ce qui rend impossible le calcul d'une quelconque efficacité ou précision. Nous ne pouvons pas savoir si les bases ne sont pas déjà pré-triées, car souvent les images similaires se suivent par leur numéro de série [2].

- ✓ *QBIC (Query by Image Content)*
- ✓ *Virage*
- ✓ *RetrievalWare*
- ✓ *Photobook*
- ✓ *NeTra*
- ✓ *CIIR (Center for Intelligent Info Retrieval)*

III. INDEXATION ET RECHERCHE D'IMAGES PAR LE CONTENU VISUEL

L'archivage des images et des séquences vidéos aussi bien pour les chaînes de télévision, les journaux, les musées, voire même pour les moteurs de recherche sur Internet ne se fait qu'au prix d'une étape d'annotation manuelle à l'aide de mots-clés. Cette indexation représente une tâche longue et répétitive pour l'humain, surtout avec les bases d'images qui deviennent aujourd'hui de plus en plus grandes. De plus, cette tâche est très subjective à la culture, à la connaissance et aux sentiments de chaque personne. Le besoin de méthodes d'indexation et de

recherche directement basé sur le contenu de l'image n'est donc plus à démontrer. Le premier prototype de système fut proposé en 1970 et attirera l'attention de la communauté des chercheurs. Les premiers systèmes d'indexation d'images par le contenu sont créés au milieu des années 90 pour des bases de données fournies, spécialisées et pour la plupart fermées.

Les systèmes d'indexation et recherche d'images par le contenu (CBIR) (*Content-Based Image Retrieval systems*) permettent de rechercher les images d'une base d'images en fonction de leurs caractéristiques visuelles. Ces caractéristiques, encore appelées caractéristiques de bas-niveau sont la couleur, la texture, la forme et tout autre caractéristique de l'image qu'on peut imaginer.

A. Architecture d'un système d'indexation et recherche d'images

Un système de recherche d'images par le contenu visuel comporte une phase *hors ligne* d'indexation de la base d'images et une phase *en ligne* de recherche à proprement dit [3].

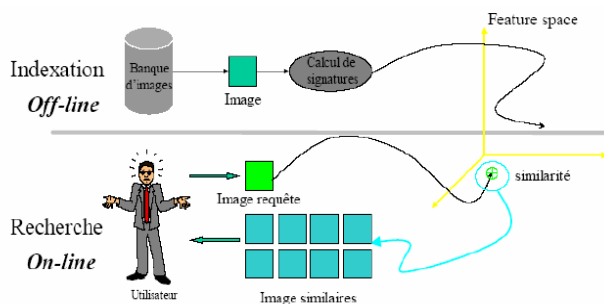


Figure 1: l'architecture du système de recherche d'images par le contenu [4].

- **Indexation de la base d'images**

Ces systèmes s'exécutent en deux étapes : l'étape d'indexation et l'étape de recherche.

- ✓ **Indexation** : des caractéristiques sont automatiquement extraites à partir de l'image et stockées dans un vecteur numérique appelé descripteur visuel.
- ✓ **Recherche** : le système en fonction de la requête de l'utilisateur, lui renvoie les résultats correspondant la plupart sous forme d'une galerie d'images ordonnées en fonction de la similarité entre leur descripteur visuel et celui de l'image requête en utilisant une mesure de distance.

B. Les bases d'images

On peut distinguer deux catégories de bases d'images :

- les bases d'images spécialisées qui contiennent des images de même type : images médicales, images géographiques, portraits de personnes...
- les bases d'images généralistes qui contiennent des images de natures variées : photographies personnelles, images des encyclopédies, de Corel, du web...

Les stratégies de recherche d'images pour ces deux catégories de base sont très différentes. Pour la première, on connaît *à priori* le type d'images que l'on peut y rencontrer, ainsi que le type de recherche que l'on va y mener. Cette connaissance *à priori* permet de développer des techniques d'indexation et de recherche très efficaces. Pour la seconde catégorie, par contre, on ne sait pas ce que contiennent les images, et on ne sait pas sémantiquement, ce que recherche l'utilisateur [1].

C. Types de requêtes

Il existe 3 façons de faire une requête dans un système d'indexation et recherche des images :

- ✓ Une requête par mots clés.
- ✓ Une requête par esquisse.
- ✓ Une requête par exemple : l'utilisateur fournit un exemple, le système en extrait des descripteurs visuels, puis recherche les images ayant des descripteurs similaires. La figure 2 donne une démonstration pour les trois façons.

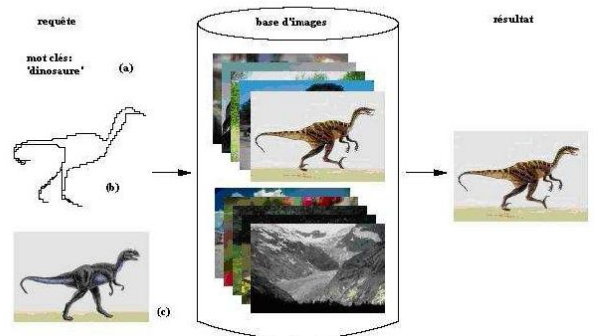


Figure 2 : Trois façons pour donner une requête dans CBIR [5].

D. Les Caractéristiques symboliques

Une image contient plusieurs caractéristiques symboliques.

1. Couleurs

La couleur est en effet très souvent la première caractéristique qui est employée pour la recherche d'images. Elle est la caractéristique la plus accessible puisqu'elle est directement codée au niveau de chaque pixel [1].

La technique classique de calcul des zones de couleur est celle de l'histogramme. Le principe est de déterminer la couleur présente dans l'image et la proportion de la surface de l'image qu'elle remplit. C'est ce qu'on appelle un histogramme de couleur. Les histogrammes de couleurs peuvent être construits dans plusieurs plages de couleurs, RVB (Rouge, Vert, Bleu) ou toute autre plage de couleurs de toute dimension. Un histogramme de couleurs est produit en découpant d'abord les couleurs de l'image dans un certain nombre de cases, puis en comptant le nombre de pixels dans chaque case. Bien évidemment, un tel procédé de description est assez simpliste mais il est suffisant pour rechercher des images de coucher de soleil dans une collection d'images de villes ou maritimes... Cette technique est très utilisée, les

histogrammes sont faciles et rapides à calculer, robustes à la manipulation des images comme la rotation et la translation [6].

Malgré cela l'utilisation d'histogrammes pour l'indexation et la recherche d'images pose des problèmes :

- ✓ Ils sont de grandes tailles, par conséquent il est difficile de créer une indexation rapide et efficace en les utilisant tels quels.
- ✓ Ils ne possèdent pas d'informations spatiales sur les positions des couleurs.
- ✓ Ils sont sensibles à de petits changements de luminosité, ce qui est problématique pour comparer des images similaires, mais acquises dans des conditions différentes.
- ✓ Ils sont inutilisables pour la comparaison partielle des images (objet particulier dans une image), puisque calculés globalement sur toute l'image.

2. Textures

Il n'existe pas de définition pertinente de la texture. Cependant, une définition de sens commun est la suivante : la texture est la répétition d'éléments de base construits à partir de pixels qui respectent un certain ordre. Le sable, l'eau, l'herbe, la peau sont autant d'exemples de textures. L'aléatoire joue un rôle particulier dans les textures.

Au niveau sémantique, les textures peuvent apporter une information précieuse. En effet, elles permettent de différencier des parties d'images dont les descripteurs de couleurs sont identiques. Par exemple, le ciel (texture unie ou nuageuse) et la mer (texture des vagues) [1].

3. Formes

L'utilisation de caractéristiques ou descripteurs de forme (*shape*) n'a de sens que sur une image segmentée. Pour extraire les descripteurs d'une forme, la première chose à faire est de définir sa fonction caractéristique. En général, elle est représentée sous la forme d'un masque dans lequel chaque pixel est représenté par le numéro de la région à laquelle il appartient. C'est à partir de cette fonction, que sont calculés la plupart des descripteurs de formes, soit à partir de la région entière, soit à partir des contours seulement [1].

IV. IMPLEMENTATION

La figure suivante schématise l'architecture générale de INDEX, que nous avons développé en C++ :

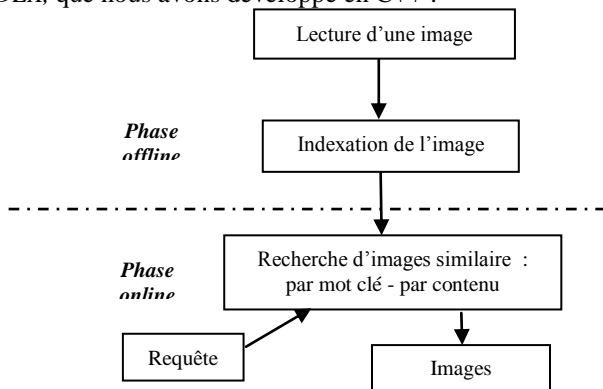


Figure 3 : Architecture générale de « INDEX »

A. Le type de base d'images utilisées

Dans une base de données, il existe une multitude de types d'images (IRM, microscopie médicale, dessin, etc.) Pour limiter notre étude, nous ne considérons dans ce travail que les bases d'images généralistes qui contiennent des images de nature variée. Nous allons nous intéresser principalement aux images fixes.

Nous avons utilisé la base de données relationnelle MySQL pour la création des différentes bases de données. C'est une base de données *opensource* qui offre de nombreuses fonctionnalités pour un système de gestion de bases de données relationnelles. Cette base de données est gratuite.

B. Méthodes d'indexation

1. Indexation par mot clé :

Nous avons donné pour chaque image de notre base, un ou plusieurs mots clés qui interprètent l'image.

Nous avons utilisé six (06) catégories qui sont :

- Animal
- Plante
- Maison
- Transport
- Nature
- Insecte

Cette phase d'indexation est lente, puisque c'est une phase manuelle. Elle demande beaucoup de concentration, et une bonne connaissance de la langue.

La figure 4 représente le résultat de la recherche obtenue avec le mot clé voiture.

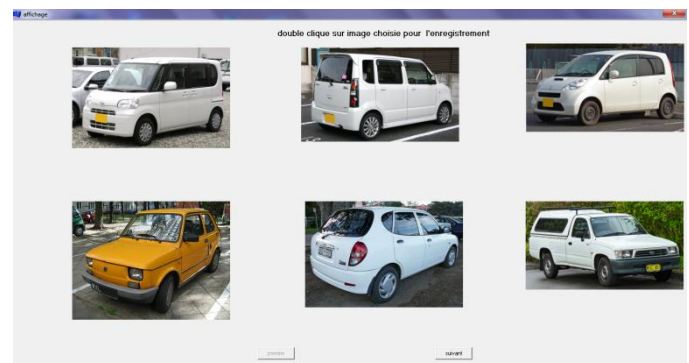


Figure 4 : capture d'écran N°1 de notre système (résultat de la recherche par mot clé = voiture)

2. Indexation par le contenu :

a) Indexation par la Couleur :

La couleur est l'une des composantes les plus utilisées en indexation d'image. Elle est indépendante de la taille de l'image et de son orientation. Dans notre système, nous utilisons l'espace couleur : RVB. Nous avons choisi cet espace parce que RVB est couramment utilisé.

✓ *Couleur – corrélation d’histogrammes :*

Les histogrammes sont résistants à un certain nombre de transformations sur l’image. Ils sont invariants aux rotations, aux translations, et aux changements d’échelle. Mais malgré ça, il faut dire que les histogrammes sont sensibles aux changements d’illumination et aux conditions d’éclairage.

Nous avons utilisé la technique de corrélation d’histogrammes. Avec cette méthode, on calcule une certaine distance de similarité pour mesurer si deux histogrammes sont « proches » l’un de l’autre.

Soit h_1 et h_2 deux histogrammes de même taille N , la formule utilisée pour la corrélation d’histogramme est la suivante :

$$d(h_1, h_2) = \frac{\sum_{i=1}^N \bar{h}_1(i) \bar{h}_2(i)}{\sqrt{\sum_{i=1}^N \bar{h}_1(i)^2 \sum_{i=1}^N \bar{h}_2(i)^2}}$$

$$\text{Ou } \bar{h}(i) = h(i) - \frac{1}{N} \sum_{i=1}^N h(i)$$

b) *Algorithme utilisé :*

L’algorithme utilisé pour la recherche des images similaires est le suivant :

- Etape 1 :** Lecture de l’image requête.
- Etape 2 :** Calcul de l’histogramme de trois couleurs R,V,B de l’image requête.
- Etape 3 :** Calcul de la distance par une fonction de similarité avec l’image requête.
- Etape 4 :** Tri des distances pour chaque image calculée.
- Etape 5 :** Affichage des N premiers résultats.

La figure 5, représente le résultat de la recherche par contenu. Les images les plus ressemblantes à l’image requête sont affichées dans les premières fenêtres d’affichage.

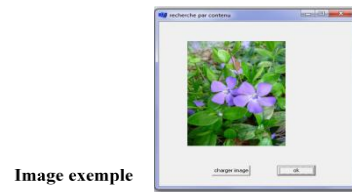


Image exemple

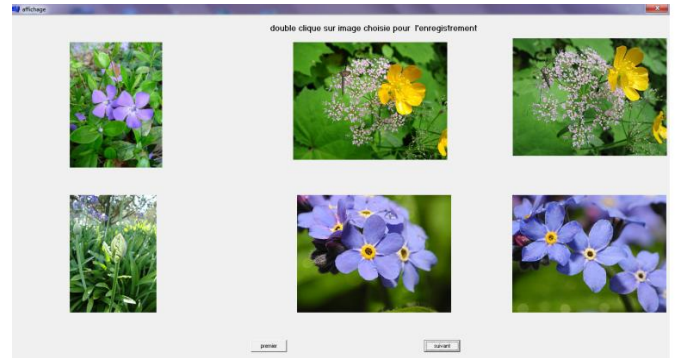


Figure 5 : capture d’écran N°2 de notre système (résultat de la recherche par contenu)

3. *Indexation par mot clé & par contenu (visuo-textuelle):*

Ce qui concerne l’indexation par mot clé & par contenu, c’est une combinaison des deux premières méthodes. Tout d’abord, on recherche les images par mot clé, et puis par contenu. C’est-à-dire, la première méthode de recherche nous donne un résultat qui contient un ensemble d’images, et la deuxième méthode de recherche nous donne un autre résultat qui contient aussi un ensemble d’images. Une fois les deux résultats sont trouvés, on calcule l’intersection des deux ensembles, ce qui nous donne un nouvel ensemble d’images.

La figure 6 représente le résultat de la recherche visuo-textuelle. D’après les expériences qu’on a fait, on a remarqué, que les résultats de la requête visuo-textuelle sont meilleurs que ceux de la requête textuelle seule, ou ceux de la requête visuelle seule.



Figure 6 : capture d'écran N°3 de notre logiciel (résultat de la recherche visuo-textuelle)

VI.CONCLUSION

Dans ce travail, nous avons étudié un ensemble de méthodes d'indexation et de recherche d'images qui sont l'indexation et recherche textuelle d'images et l'indexation et

recherche d'images par le contenu visuel. Nous avons montré l'intérêt d'un système de recherche visuo-textuelle d'images.

En perspective, d'autres développements comme extensions à ce travail sont envisageables:

- L'extraction d'autres descripteurs (caractéristiques) visuels de l'image comme par exemple la forme, la texture...etc. Et de les combiner afin d'obtenir une meilleure description de l'image améliorant ainsi la qualité de recherche future.
- Nous envisageons d'utiliser d'autres bases pour des applications professionnelles (médicales, visages,...).

REFERENCES

- [1] S. Tollari. "Indexation et recherche d'images par fusion d'informations textuelles et visuelles". Thèse de doctorat, Université du Sud Toulon-Var, 2006.
- [2] M. Vissac et J. Dugelay. "Un panorama sur l'indexation d'images fixes". Journal marocain d'automatique, d'informatique et de traitement du signal - Mars 2000
- [3] V. Gouet-Brunet. "Recherche par contenu visuel dans les grandes collections d'images". 2005/9/28
- [4] VU Ngoc Son. "Indexation et recherche d'images par le contenu". Rapport final (TIPE), promotion 11
- [5] Lê thi lan. "Indexation et recherche d'images par le contenu". Mémoire de master, hanoi 2005
- [6] A. Segonds. "Indexation visuelle et recherche d'images sur le Web : Enjeux et problèmes". Mémoire de Master II : Histoire visuelle, 2009.