

# Recherche textuelle et visuelle - Indexation par concepts

Chabane Djeraba<sup>1</sup> Fatma Bouali

<sup>1</sup> Université des Sciences et Technologies de Lille  
LIFL - UMR CNRS 8022- Bâtiment M3  
59655 Villeneuve d'Ascq Cédex - FRANCE  
Tél: 33 (0)3 20 33 55 64

{djeraba, bouali}@lifl.fr

## Résumé

*Les concepts utilisateurs permettent : - l'exploitation des descriptions visuelles et textuelles, - une intégration simple et naturelle des connaissances liées au domaine d'application, et leur réutilisation dans la définition d'autres concepts ou dans d'autres requêtes, ce qui évite d'avoir à gérer des connaissances à priori.*

## Mots clefs

Concepts, utilisateurs, textes, images.

## 1 Introduction

Le nombre impressionnant des prototypes et projets de recherche ([Hir 00], [Tan 01]) atteste de la popularité de l'indexation et de la recherche d'images par des moyens visuels ou textuels. Ce domaine ouvre la voie à de nombreuses recherches pleines de défis. En effet, en dépit de tous les efforts, sans précédents, dans le domaine de l'indexation et de la recherche d'images, tout reste à faire, notamment dans : - l'apprentissage des images utiles à la suite des requêtes., - la compréhension des vrais besoins des utilisateurs, - l'anticipation des besoins des utilisateurs [Mur 84], - la perception de la mémoire visuelle, - la sémiologie de l'image [Le 89], - la synergie entre les branches de l'indexation et les professionnels (bibliothécaire, agences de photos de presse, etc.) [Caw 92], - la sémantique dans l'indexation et la recherche [Che 97], - la viabilité des images, et - la normalisation des contenus [Ohl 82], [Get 00]).

Nous nous focaliserons dans le cadre du papier à l'introduction d'une plus grande sémantique dans l'indexation et la recherche [Che 97], et ceci à travers la combinaison de l'indexation visuelle et textuelle. Dans notre approche, nous plaçons l'utilisateur au centre du dispositif d'indexation et de recherche en lui permettant d'introduire ses propres concepts à la suite des requêtes. Ces concepts reposent sur la pondération des caractéristiques textuelles et visuelles qui composent le concept. La pondération est composée des opérateurs "Et", "Ou" et "Composition" [Dje 00]. Les opérateurs

définissent des concepts plus ou moins larges. Les différentes valeurs de pondération permettent de hiérarchiser les priorités entre les caractéristiques visuelles et textuelles qui composent le concept. Un concept peut être : - stocké dans la base de données, - réutilisé dans les requêtes, - réutilisé en tant que tel pour accéder rapidement aux images qui lui sont associées, - ou réutilisé dans la définition d'autres concepts.

## 2 Définition et représentation

Un concept est une description, faite par un utilisateur, de la structure d'une classe d'images recherchées. La structure est définie par une hiérarchie pondérée des caractéristiques visuelles et textuelles nécessaire à la description de la classe d'images associé au concept. Tout concept peut, à la demande de l'utilisateur, être stocké dans la base de données.

Un concept est exprimé aux travers de trois opérateurs, à savoir Et, Ou et Composition. Ces trois opérateurs inspirés de [Chu 96] permettent de préciser le rôle des caractéristiques dans le concept défini.

Soit le concept « fusée en phase de décollage ». Ce concept est décrit par des caractéristiques textuelles et visuelles. Les caractéristiques textuelles incluent les mots-clés « décollage », « fusée », « vol ». Les caractéristiques visuelles concernent la couleur et la forme. La caractéristique couleur comporte une proportion importante de couleur orange tandis que la forme estime celle d'une fusée. L'opérateur « composition » utilisé, pour décrire le concept, décrit le mieux le concept utilisateur recherché. La description visuelle du concept recherché, repose sur une combinaison des descriptions visuelles de deux régions qui expriment respectivement la couleur orange et la forme d'une fusée. Le choix de l'opérateur « Et » aurait signifié que la présence des deux descriptions visuelles était obligatoire pour définir une image d'une fusée en phase de décollage (ce qui donne un concept trop restreint), tandis que le choix de l'opérateur « Ou » aurait nécessité la présence d'un seul de ces deux descripteurs (ce qui donne un concept trop large). Les

opérateurs « Et » et « Ou » traduisent un appariement exact, contrairement à l'opérateur « Composition » qui traduit un appariement de similarité.

Chaque composant du concept est décrit par une valeur de pondération qui indique son importance dans le concept. Dans notre exemple, la description textuelle est privilégiée par rapport à la description visuelle, en lui affectant un poids de 60%. La description visuelle lui est associée un poids de 40%. Ces valeurs de pondération jouent un rôle primordial lors des opérations de mise en correspondance du concept et des caractéristiques des images de la base, et influencent les résultats obtenus.

La valeur de pondération est un réel compris entre 0 et 1 où le rapprochement vers la valeur 1 (respectivement 0) est synonyme d'une grande (respectivement faible) importance. Dans le cas des opérateurs « Et » et « Ou », le poids de l'attribut n'aurait aucune signification. Si un concept est formé par une seule description, alors la valeur de pondération est égale à 1, ce qui signifie une grande importance.

La description d'un concept peut faire intervenir d'autres concepts. Par exemple le nouveau concept « Fusée en phase de décollage à l'aube » est défini à partir des concepts « Fusée en phase de décollage » et « levée du jour ».

Si la description actuelle est fonctionnelle, il peut être envisagé de décrire un concept à travers des outils graphiques adaptés. Sur la base d'une telle description, la description fonctionnelle peut être générée automatiquement.

### 3 Mise en correspondance

Nous allons voir maintenant, comment est mis en correspondance le concept décrit par l'utilisateur et les caractéristiques des images de la base de données.

Si la description de la requête est syntaxiquement et sémantiquement correcte, le système construit la requête afin d'effectuer la mise en correspondance avec les images de la base de données. La mise en correspondance consiste à calculer le degré de similitude entre le concept décrit dans la requête et les images de la base de données. Le calcul se fait d'une manière ascendante de la base de la hiérarchie jusqu'au sommet. Au niveau de la base, il s'agit de calculer les distances entre les caractéristiques primitives du concept et celles des images de la base de données. Les distances utilisées sont conformes au schéma que nous avons présenté dans les sections précédentes. En ce qui concerne les mots clés, la distance est estimée à 1 si le terme (ou son synonyme) est trouvé dans l'image, à 0 sinon (l'intégration des acquis des systèmes de recherche d'information textuelle améliorerait certainement ces choix triviaux). Les distances ainsi calculées sont alors propagées au niveau supérieur. Cette propagation dépend du type de relation utilisée, comme le montre le schéma ci-dessous :

*Calcul de la similarité au niveau  $j+1$ , pour le composant  $i$  ( $C_{i,j+1}$ )*

*Pondération  $k, j-1$  : le poids associé au composant  $i$  du niveau  $j$  au rang  $k$*

*$C_{i, j+1}$  : la similarité associée au composant  $i$  du niveau  $j$*

- *Composant  $i$  = Composition-Of (Composition)*  

$$C_{i, j+1} = \sum_{k=1, m} (Pondération_{k, j-1} * C_{k, j-1})$$
 avec  $n \geq j \geq 1$  et  $j > 0$
- *Composant  $i$  = AND (Et)*  

$$C_{i, j+1} = \sum_{k=1, m} Min (C_{1, j-1}, C_{2, j-1}, C_{1, j-1} \dots, C_{k, j-1})$$
 avec  $n \geq j \geq 1$  et  $j > 0$
- *Composant  $i$  = OR (Ou)*  

$$C_{i, j+1} = \sum_{k=1, m} Max (C_{1, j-1}, C_{2, j-1}, C_{1, j-1} \dots, C_{k, j-1})$$
 avec  $n \geq j \geq 1$  et  $j > 0$

Par exemple, soit les distances entre les caractéristiques primitives (feuilles de la hiérarchie) du concept « Fusée en phase de décollage » et les caractéristiques d'une image (mots-clés, couleur, forme) : 0, 1, 1, 0.5 et 0.8. Niveau 1 : 0.6 (description textuelle), 0.5 description visuelle-couleur), 0.8 (description visuelle-couleur). Niveau 3 : 0.6 (description textuelle), 0.68 (description visuelle).  $0.68 = 0.5 * 40\% + 0.8 * 60\%$ . Niveau 3 : 0.62 (sommet de la hiérarchie).  $0.62 = 0.6 * 60\% + 0.68 * 40\%$ . La similarité au sommet de la hiérarchie correspond à la similarité estimée entre le concept et l'image considérée.

### 4 Conclusion

Les concepts utilisateurs permettent : - l'exploitation des descriptions visuelles et textuelles, - une intégration simple et naturelle des connaissances liées au domaine d'application, et leur réutilisation dans la définition d'autres concepts ou dans d'autres requêtes, ce qui évite d'avoir à gérer des connaissances a priori.

Cependant, l'introduction des concepts, malgré leurs avantages, souffrent de certaines limites qu'il faudrait bien pallier : - les concepts sont créés et stockés à fur et à mesure de la génération des requêtes, et non une fois pour toutes lors de la création de la base, ce qui réduit la rapidité de leur exploitation, - le nombre de concepts peut devenir très vite important. Ce qui induit de nouveaux problèmes tels que la gestion du volume, des redondances et des chevauchements entre concepts, etc. Ces problèmes réduisent la vision synthétique des concepts et indirectement leur réutilisation dans les requêtes et dans la description de nouveaux concepts. Notre objectif pour atténuer ces limites est de développer une vraie méthodologie d'indexation et de recherche visuelle.

### Références

[Caw 92] Cawkell A. E., "Selected aspects of image processing and management : review and future

- prospects". *Journal of Information Science* 18 :179-192, 1992.
- [Che 97] Chen H., Smith T. R., Larsgaard M., Hill L. L., Ramsey M., "A Geographic Knowledge Representation System for Multimedia Geospatial Retrieval and Analysis", *Int. J. on Digital Libraries* 1(2), pages 132 – 152, 1997.
- [Chu 96] Chua T. and al., "Using Domain Knowledge in Querying Image Databases", *Conference Multi Media Modeling, Towards The Information Superhighway*, France, Toulouse, 12-15 november 1996.
- [Dje 00] Djeraba C., Bouet M., Briand H., Khenchaf A., "Visual and Textual Content Based Indexing and Retrieval". *Int. J. on Digital Libraries* 2(4): pages 269-287, 2000.
- [Get 00] Getty Vocabulary Program, "Art & Architecture Thesaurus", Los Angeles: *J. Paul Getty Trust*, Vocabulary Program, 2000.
- [Hir 00] Hirata Kyoji, Mukherjea Sougata, Li W.-S., Hara Y., "Integration of Image Matching and Classification for Multimedia Navigation", *Multimedia Tools and Applications*, 11(3): 295-309, August 2000.
- [Le 89] Le Guern, Odile. "Images et bases de données", *Bulletin des Bibliothèques de France* 34, no 5, Pages 422 – 435, 1989.
- [Mur 84] Murry Rita, "Criteria for Subject Indexing of Pictures : an Introductory Survey", *Master dissertation of the University of Alberta*, 1984.
- [Ohl 82] Ohlgren, Thomas H., "Image analysis and indexing in North America : a survey", *Art Libraries Journal (summer)*, pages 51 - 60, 1982.
- [Tan 01] Tan Kian-Lee, Ooi B. C., Yee C. Y., "An Evaluation of Color-Spatial Retrieval Techniques for Large Image Databases", *Multimedia Tools and Applications*, 14(1): 55-78, May 2001.