

Proposition de sujet de thèse de doctorat
Allocation fléchée – Campagne 2009

***Vers un système interactif de structuration des index pour une
recherche par le contenu dans des grandes bases d'images***

On assiste depuis quelques années au développement de nombreuses bases d'images volumineuses. Ce développement est favorisé par l'essor et la démocratisation des dispositifs d'acquisition tels que les appareils photo numériques, mais aussi par les progrès technologiques et techniques réalisés pour le stockage, la transmission et la diffusion de documents multimedia. La recherche d'information dans ces grandes masses de données nécessite des méthodes efficaces et robustes pour indexer les images par leur contenu et structurer les bases d'images ainsi indexées [1].

Le schéma traditionnel de l'indexation et la recherche d'images par le contenu, illustré en Figure 1, est articulé autour de trois étapes.

Dans une première étape d'indexation (généralement menée hors-ligne), il s'agit d'extraire, pour chaque image de la base, une signature qui lui soit spécifique. Cette signature est généralement issue de caractéristiques basées sur la forme, la texture, la couleur... [2]).

La deuxième phase de structuration (généralement menée hors-ligne) vise à organiser les images dans l'espace des signatures de manière à faciliter, accélérer et améliorer les résultats de la recherche ultérieure. Il existe plusieurs familles de méthodes de structuration : par partitionnement des données ou de l'espace des signatures, par linéarisation de l'espace de description, par filtrage des données, par regroupement des données... Dans cette thèse, nous nous intéressons plus particulièrement aux méthodes de regroupement des données, basées sur l'utilisation d'une mesure de similarité entre les signatures extraites lors de la première étape. Il s'agit d'organiser l'espace des signatures en groupes d'objets homogènes.

Lors de la troisième étape de recherche par le contenu (généralement menée en-ligne), l'utilisateur soumet au système une requête sous la forme d'une image. Le système doit alors être capable de retourner à l'utilisateur les images les plus similaires à cette image requête, en utilisant une mesure de similarité entre les images dans l'espace des signatures structuré. Les images les plus similaires sont retournées par ordre de similarité décroissante.

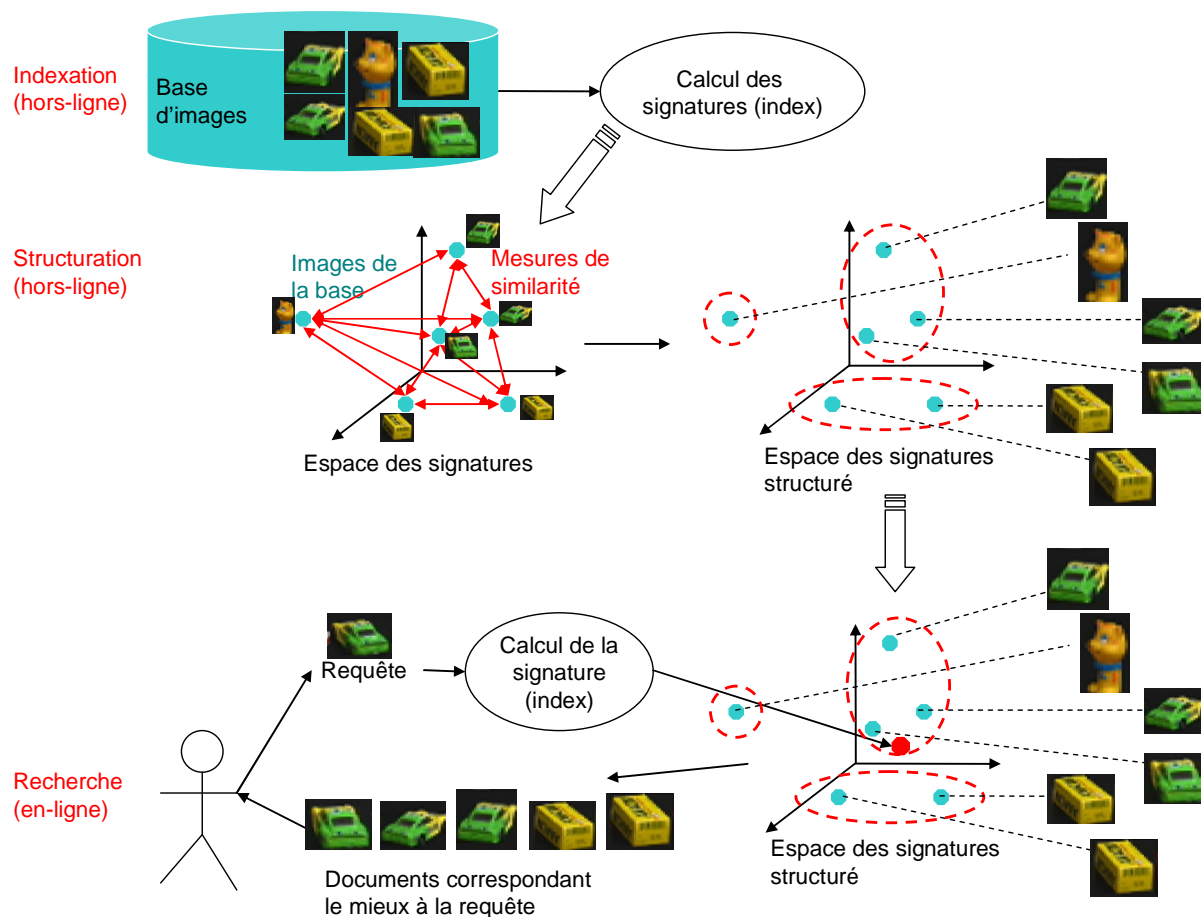


Figure 1 : Recherche par le contenu dans une base d'images basée sur l'indexation et la structuration des index

Dans le cadre de cette thèse, nous nous intéresserons plus particulièrement à l'étape de structuration de l'espace des signatures. On considère que la phase d'indexation a été préalablement menée à bien et que l'on dispose des signatures des images. Les signatures sont issues de caractéristiques d'un bas niveau sémantique (forme, texture, couleur...) extraites automatiquement depuis le contenu des images. Or, il peut exister ainsi un *fossé sémantique* entre certains concepts de haut niveau sémantique exprimés par l'utilisateur via l'image requête et ces caractéristiques de bas niveau.

Par exemple, dans la base d'images de la Figure 1, il y a des images de jouets (voiturette et poupée) et des images d'une boîte de thé. L'utilisateur, lorsqu'il soumet comme requête une image de voiturette, peut attendre du système qu'il lui retourne toutes les images de jouets (y compris la poupée). Alors que le système, s'il est basé uniquement sur des caractéristiques de bas niveau, aura tendance à renvoyer la boîte de thé avant la poupée, plus éloigné selon les caractéristiques de forme notamment.

L'approche envisagée dans le cadre de cette thèse, illustrée en Figure 2, consiste à impliquer l'utilisateur dans la phase de structuration de l'espace des signatures. Lors de cette phase de structuration, qui devient en-ligne, l'utilisateur peut être amené à modifier les groupes d'objets en fonction de ses attentes, en modifiant les étiquettes associées automatiquement à certains objets par exemple. Ce qui a pour effet de réduire le fossé sémantique, comme illustré en Figure 2.

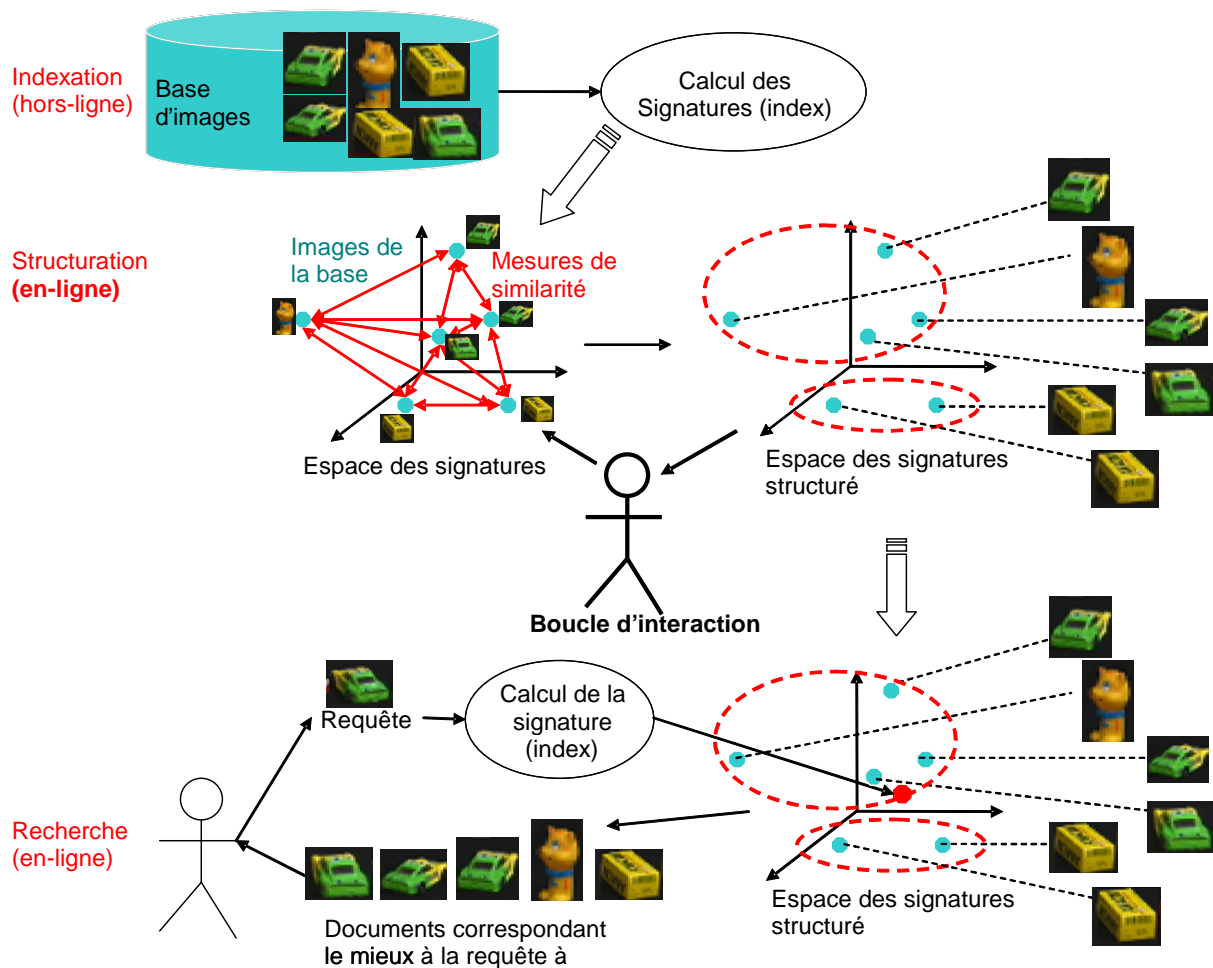


Figure 2 : Recherche par le contenu dans une base d'images basée sur l'indexation et la structuration des index avec interaction utilisateur pour la structuration

Le premier objectif de cette thèse sera de réfléchir aux modalités de l'intervention de l'utilisateur dans cette boucle d'interaction. En effet, ces modalités détermineront le contexte de l'étude :

- Si l'utilisateur intervient dès le début de la structuration en créant des classes d'objets et en étiquetant « à la main » au moins un exemple par classe, alors on se trouve dans un contexte de catégorisation (classification) semi-supervisée interactive. En effet, les classes sont définies mais seulement certains exemples sont étiquetés (apprentissage semi-supervisé).
- Si l'utilisateur n'intervient que plus tard dans la structuration, alors on se place dans un contexte de clustering interactif. On parle de clustering puisque les groupes ne sont pas prédéfinis : c'est le système qui découvre automatiquement les groupes d'objets (clusters). Dans ces deux cas, l'interactivité provient du fait que l'utilisateur pourra, à tout moment, modifier l'étiquette associée à certains exemples qui lui semblent mal classés selon ses critères de haut niveau sémantique.

Des travaux ont déjà été menés dans le cadre de la collaboration entre le L3I et le MSI concernant l'interactivité dans l'indexation et la recherche d'images par leur contenu [3]. Quelques recherches se sont intéressées à la problématique de clustering interactif [4,5]. Dans cette thèse, nous irons plus loin en nous intéressant également à la problématique de catégorisation semi-supervisée interactive.

Le second objectif de cette thèse est d'adapter des méthodes de clustering et de catégorisation usuelles à ce schéma interactif. Puisque, dans le cadre de ce nouveau schéma, l'étape de structuration devient en-ligne et par voie de conséquence les méthodes de structuration par regroupement des données doivent être le moins coûteuses possibles (hors phase d'apprentissage), nous proposons d'utiliser des méthodes basées sur des réseaux de neurones telles que les cartes de Kohonen (clustering) ou les réseaux de fonctions à Base Radiale (catégorisation). Dans un contexte de catégorisation, on peut également envisager l'utilisation de machines à vecteur de support. Un certain nombre de travaux ont déjà été menés dans le domaine de l'apprentissage incrémental pour la catégorisation [6]. Le but sera d'explorer ces travaux et de les étendre à notre contexte interactif.

Notre approche suppose a priori que les similarités entre signatures soient calculables. Nous nous restreindrons donc, dans un premier temps, à des signatures statistiques exprimées sous la forme d'un vecteur de caractéristiques et pour lesquelles les similarités sont directement calculables. Dans un second temps, nous pouvons envisager d'étendre ces travaux à des signatures exprimées sous la forme de graphes, en utilisant une méthode de type graph probing pour obtenir les similarités par exemple.

Enfin, un objectif à plus long terme sera de confronter les mesures de similarité entre caractéristiques de bas niveau calculées automatiquement par le système et les similarités perceptuelles données par l'utilisateur pour mieux en expliquer les analogies et les différences. Cette étude aura pour but d'améliorer la prise en compte conjointe des descripteurs images et des indications données par l'utilisateur dans une même méthode interactive.

- [1] A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta and R. Jain. Content-based image retrieval at the end of the early years. *IEEE Transactions on PAMI*, 22(12), pp. 1349-1380, 2000.
- [2] M. Nixon, A.S. Aguado. *Feature Extraction and Image Processing*. Academic Press Inc, 2nd Edition, 2007.
- [3] N.V. Nguyen, J.M. Ogier, S. Tabbone and A. Boucher. Text Retrieval Relevance Feedback Techniques for Bag of Words Model in CBIR. *International Conference on Machine Learning and Pattern Recognition (ICMLPR)*, Juillet 2009, Paris.
- [4] M. Koskela. *Interactive image retrieval using Self-Organizing Maps*. Doctoral thesis, Helsinki University of Technology, *Dissertations in Computer and Information Science*, Report D1, Espoo, Finland, 2003
- [5] M. Detyniecki and A. Nürnberger. Weighted Self-Organizing Maps: Incorporating User Feedback. *Proceedings of the Joint 13th International Conference on Artificial Neural Networks and 10th International Conference on Neural Information Processing (ICANN/ICONIP)*, pp. 883-890, 2003.
- [6] G. Cauwenberghs and T. Poggio. Incremental and decremental support vector machine learning. *Advances in Neural Information Processing Systems (NIPS 2000)*, volume 13. MIT Press, Cambridge, MA, 2001.