

Utilisation de la structure pour la recherche d'images dans des documents semi-structurés

Mouna Torjmen

SIG-RFI, IRIT
118, Route Narbonne, 31062 Toulouse, France
torjmen@irit.fr

Résumé. Nous proposons dans cet article une méthode de recherche d'images dans les documents semi-structurés en exploitant le contexte textuel et structurel de ces documents.

1 Introduction

La recherche d'images dans des collections multimédia (texte + images) peut se faire par deux méthodes : (1) la recherche d'images par contenu, c'est à dire en utilisant les caractéristiques de bas niveau des images, et (2) la recherche d'images par contexte. Les approches de recherche contextuelle utilisent le contenu textuel entourant les images pour leur recherche. Dans la littérature, ces approches ont montré leur grande efficacité, et ceci apparaît surtout dans le domaine du WEB, où la plupart des moteurs de recherche images tel que GoogleImage l'utilise. Nous nous proposons d'appliquer la recherche contextuelle d'images aux documents semi-structurés, en ajoutant le contexte structurel des images à leur contexte textuel.

2 Traitement des requêtes images avec le contexte structurel

Les documents semi-structurés tels que XML peuvent être représentés sous forme d'arbre dans lequel la racine est le document, les noeuds internes sont les noeuds représentant les éléments et les noeuds feuilles sont les noeuds contenant des valeurs (texte, images...). En se basant sur cette représentation des documents, l'idée est de calculer une pertinence sémantique des images en fonction du contexte textuel et structurel (distance à l'image, niveau de profondeur dans l'arbre,...). Nous utilisons pour cela les scores de pertinence des noeuds de l'arbre, préalablement calculés par le système XFIRM, qui est destiné à la recherche de fragments XML et basé sur une méthode de propagation de pertinence (Sauvagnat, 2005).

Une fois les noeuds pertinents définis par XFIRM (noeuds ayant un score non nul), nous proposons d'utiliser ceux les plus proches de l'image pour participer au calcul de son score de pertinence. Pour cela, nous considérons trois sources d'évidence : noeuds descendants, noeuds frères et noeuds ascendants. Le choix des noeuds descendants est justifié par le fait qu'ils peuvent contenir des informations spécifiques aux images comme la *légende*. L'utilisation des noeuds frères est due au fait qu'ils ont plus de chances que les autres noeuds de partager le même thème que l'image. Et enfin, l'utilisation des noeuds ascendants permet de prendre en

compte tout le contexte du document. La formule suivante est utilisée pour le calcul de score final de chaque image :

$$W_{im} = p1.W_d^{im} + p2.W_f^{im} + p3.W_a^{im} \quad (1)$$

Avec : W_d^{im} est le score de l'image en fonction des noeuds descendants, W_f^{im} est le score de l'image en fonction des noeuds frères, W_a^{im} est le score de l'image en fonction des noeuds ascendants, et $p1$, $p2$ et $p3$ sont des paramètres utilisés pour varier l'importance des différentes sources d'évidence avec $p1+p2+p3=1$. La pertinence des noeuds non-images est évaluée par le système XFIRM.

- **Evaluation des noeuds descendants** : Le score total de ces noeuds est calculé en fonction du nombre total de noeuds descendants jugés pertinents, du nombre total de noeuds descendants non pertinents (ces deux paramètres permettent d'introduire la quantité de pertinence et non-pertinence) et du score de chaque noeud descendant jugé pertinent.
- **Evaluation des noeuds frères** : Le score total de ces noeuds est calculé en fonction du score de chaque noeud frère jugé pertinent, de la distance entre l'image et chaque noeud frère : plus un noeud frère est proche, plus il a de chances de partager l'image le même thème, du nombre de noeuds frères pertinents et du nombre de noeuds frères non pertinents afin d'introduire la quantité de pertinence et non-pertinence.
- **Evaluation des noeuds ascendants** : Le score total de ces noeuds est calculé en fonction du score de chaque noeud ascendant jugé pertinent, et de la distance entre le noeud image et son ascendant : plus un noeud ascendant est loin du noeud image, moins il participe à sa pertinence.

3 Conclusion

Nous avons proposé une méthode pour la recherche contextuelle d'images dans des documents semi-structurés. Des premières expérimentations sont faites dans le cadre de la campagne d'évaluation INEX Multimedia 2006 (Hlaoua et al., 2006) et ont montré l'importance de la prise en considération de la structure pour la recherche d'images dans les documents semi-structurés.

Références

- Hlaoua, L., M. Torjmen, K. Pinel-Sauvagnat, et M. Boughanem (2006). Xfirm at inex 2006 : Ad-hoc, relevance feedback and multimedia tracks. In *NEX, Dagstuhl, Allemagne*.
- Sauvagnat, K. (2005). *Modèle flexible pour la recherche d'information dans des corpus de documents semi-structurés*. Thèse de doctorat, Université Paul Sabatier.

Summary

We propose in this article a method for image retrieval in semi-structured documents that exploit textual and structural context of these documents.