

Recherche d'image par la sémantique

R. Oulad Haj Thami¹, M. Daoudi², Y. El Mansouri³

¹UFR R&T, ENSIAS, Université Mohamed V-Souissi,

²ENIC/INT, rue Marconi, Cité Scientifique, Villeneuve d'Ascq, France

³EMSI, Ecole Marocaine des Sciences de l'Ingénieur, Casa, Maroc

B.P. 713, Rabat-Agdal, Maroc

E-mail : oulad@ensias.um5souissi.ac.ma

Résumé

Dans cet article nous proposons un modèle générique pour une modélisation du contenu de l'image. Le modèle est basé sur plusieurs niveaux/vues. Nous distinguons trois types de vues : les vues concepts, relations et interprétation. Le nombre, la nature et le contenu de chaque vue sont adaptables et peuvent varier selon le type d'application. Le but étant la recherche d'image par la sémantique. Le modèle opérationnel est basé sur les graphes conceptuels emboîtés que nous avons étendus. Le modèle est implémenté et opérationnel et les résultats obtenus sont très encourageants.

Mots clés

Bases de données multimédia, modélisation du contenu de l'image, recherche d'image par la sémantique.

1 Introduction

La modélisation du contenu des images pour la recherche par la sémantique est l'un des problèmes majeurs des systèmes d'informations multimédia. En effet, en plus des attributs visuels le modèle doit être générique pour répondre à une large classe d'applications en prenant en compte la sémantique des différents types d'images.

Dans cet article, nous présentons un modèle générique pour représenter le contenu de l'image. Notre modèle est basé sur plusieurs niveaux. Chaque niveau concerne une classe d'informations. Le nombre de niveaux et leurs contenus sont adaptables et peuvent être réorganisés selon la perception de l'expert de son application et de type de requêtes sur les images.

Nous commençons par présenter les concepts de base de notre modèle et donnons après un exemple d'un objet multimédia dans selon notre modèle. A la fin, un exemple de requête est donnée et une conclusion pour tracer les perspectives de notre travail.

2 Modèle de données

On considère une base d'images $B_i = \{I_1, I_2, \dots, I_n\}$ avec les $I_{1 \leq i \leq n}$ des images couleurs ou aux niveaux de gris. Chaque image peut être décomposée en plusieurs objets. Soit $O = \{o_1, o_2, \dots, o_m\}$ l'ensemble des objets extraits des images de B_i . Chaque objet $o_{1 \leq i \leq m}$ de O peut être également décomposé en d'autres objets. Il est à noter que $B_i \subseteq O$. En effet, chaque image est considérée aussi comme un objet, appelé objet racine

Soit F l'ensemble des informations utiles dans tous les objets et, soit $\Phi = \{\varphi_1, \varphi_2, \dots, \varphi_p\}$ une classification de l'ensemble des informations de F . Cette classification Φ peut être vue comme l'organisation des données dans les bases de données classiques, comme le relationnel par exemple ou les bases de données orientées objets. L'expert organise les informations, auxquelles il s'intéresse dans le contenu des images, selon ses besoins et sa perception de ce contenu.

2.1 Taxinomie des attributs d'une image

Dans cette partie, nous donnons une taxinomie des attributs et caractéristiques du contenu d'une image. Ces attributs sont divisés entre deux classes : les concepts et les relations. Les concepts constituent les attributs associés à la description du contenu de l'image ou les informations caractérisant ce contenu. Par contre les relations forment le vocabulaire nécessaire à la description des liens qui peuvent exister entre les concepts. Notre taxinomie est basée sur la propriété de calculabilité automatique ou non. Nous distinguons quatre classes fondamentales : les concepts concrets, les concepts abstraits, les relations concrètes et les relations abstraites. Nous donnons, dans ce qui suit, la définition de chacun de ces éléments.

Définition 1 : Concept

Un concept est une notion générale, désignée par un mot, qui reflète un ensemble de caractères communs à un ensemble d'objets.

Définition 2 : Concept concret

Constitue une notion calculable qui ne nécessite pas une interprétation particulière.

Par exemple le contour d'un objet dans l'image, la couleur ou la distribution de la couleur dans une image. Ces concepts peuvent être calculés automatiquement ou d'une façon semi-automatique.

Définition 3 : Concept abstrait

C'est un concept générique qui dépend de la culture ou de la vision d'une personne.

Par exemple, le concept Homme ou Voiture.

Définition 4 : Relation

C'est une idée particulière qui désigne tout lien par lequel on peut lier un concept à un autre ou d'autres concepts.

Définition 5 : Relation Concrète

Elle constitue une relation calculable ne nécessitant pas une interprétation particulière de la part d'un observateur.

Par exemple les relations spatiales comme “à coté” ou “à gauche”.

Définition 6 : Relation Abstraite

C’est une relation nécessitant une interprétation particulière et dépend en général de la culture et du contexte de l’observateur.

Par exemple “parler à”, “aimer” ou “haïr”.

2.2 Le support : hiérarchie de concepts et hiérarchie de relations

La classification Φ est munie d’une hiérarchie de concepts T_c et d’une hiérarchie de relations T_r . Les deux hiérarchies sont composées respectivement de concepts et de relations abstraits. T_c et T_r constituent le vocabulaire utilisé pour la description de l’interprétation du contenu des images.

Les concepts et les relations permettent d’exprimer les connaissances liées à chacune des classes de Φ . Les concepts décrivent la nature des objets. Les relations étant des liants entre les différents concepts.

2.3 Vue ou niveau de modélisation

Une vue ou un niveau de modélisation correspond à un type de contenu informatif défini sur un ensemble d’objets multimédia. Autrement, une vue V_i est donnée par :

$$V_i = (\varphi_i, C_i, R_i, t_c, t_r, T_c, T_r, B_i).$$

Avec $\varphi_i \in \Phi$, C_i et R_i respectivement l’ensemble des concepts et de relations concrets et abstraits utilisés dans V_i , t_c et t_r le treillis de concepts et de relations définis, respectivement, sur les concepts concrets de C_i et les relations concrètes de R_i . T_c et T_r les hiérarchies globales de concepts et de relations abstraits définis ci-dessus. B_i est un ensemble de contraintes exprimées sous forme de (c_i, r, c_j) avec $c_i, c_j \in T_c \cup t_c$, et $r \in T_r \cup t_r$ précisant que, le concept c_i peut être lié par une relation r avec le concept c_j .

Définition 7 : vue relations

Elle est donnée par $V_i = (\varphi_i, C_i, R_i, \emptyset, t_r, T_c, \emptyset, B_i)$ avec $C_i \subseteq T_c$ et $\forall r \in R_i, r \in t_r$.

Définition 8 : Vue concepts

Elle est donnée par $V_i = (\varphi_i, C_i, \emptyset, t_c, \emptyset, \emptyset, \emptyset, B_i)$ avec $\forall c \in C_i, c \in t_c$ et les éléments de B_i de la forme $(c_i, \emptyset, \emptyset)$ avec $c_i \in t_c$.

Définition 9 : Vue interprétation

Ce niveau est donné par $V_i = (\varphi_i, C_i, R_i, \emptyset, \emptyset, T_c, T_r, B_i)$ avec $C_i \subseteq T_c, R_i \subseteq T_r$ et B_i l’ensemble des graphes syntaxiquement corrects de la forme (c_i, r, c_j) avec $c_i \in T_c, c_j \in T_c$ et $r \in T_r$.

Ce niveau est destiné principalement pour la description de la sémantique de l’image ou de l’objet multimédia. Son contenu dépend de l’interprétation de l’expert ou de l’observateur.

A chacune des classes de Φ est associée une vue. Soit Φ_v l’ensemble de vues définies sur Φ : $\Phi_v = \{V_1, V_2, \dots, V_p\}$.

2.4 Le descripteur de l’objet Multimédia

Chaque Objet est décrit selon un certain nombre de vues définies et organisées dans un treillis appelé le descripteur de l’objet multimédia (figure 1). Il est en quelque sorte le schéma de la base de données pour la description d’un objet multimédia. Il est composé des treillis généraux T_c et T_r , ainsi que de l’ensemble des vues utilisées pour la description de l’objet multimédia. Chaque vue est définie par son treillis de concepts t_c et son treillis de relations t_r selon le type de la vue.

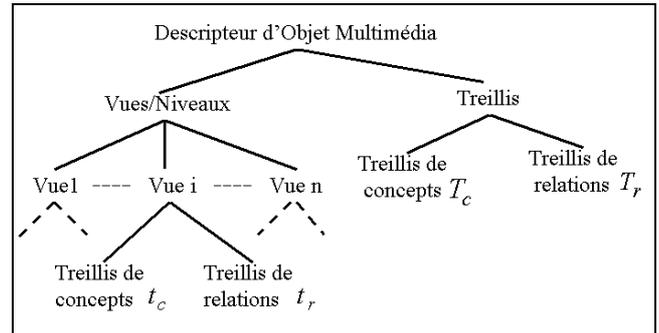


Figure 1 : Descripteur d’objet multimédia

3 Le modèle opérationnel

La description des objets de O , par rapport à Φ_v , est réalisée grâce aux graphes conceptuels emboîtés. C’est une représentation visuelle des connaissances descriptives des objets ou de leurs contenus [1]. On définit une fonction Π de O dans G^* , où G^* l’ensemble des graphes conceptuels définis sur les concepts de $T^* = T_c \cup t_c^i$, avec t_c^i le treillis de concepts concrets défini dans $V_i \in \Phi_v, 1 \leq i \leq p$, et les relations de $R^* = T_r \cup t_r^i$, avec t_r^i le treillis de relations concrètes défini dans $V_i \in \Phi_v, 1 \leq i \leq p$. Un graphe conceptuel emboîté $G = (C, R, U)$ est un multigraphe tel qu’il peut exister plusieurs arêtes entre deux sommets, non orienté, biparti et non nécessairement connexe. Ici, $C \in T^*$ et $C \neq \emptyset$ le concept sommet du graphe, $R \in R^*$ est la relation sommet. U est l’ensemble des arrêtes du graphe. L’emboîtement permet d’affecter comme valeur à un concept un graphe conceptuel également emboîté. Les graphes conceptuels permettent une représentation très efficace et hiérarchique des connaissances. Les opérations définies sur les graphes conceptuels permettent de faire des inférences sur ces connaissances [2]. Contrairement au modèle de base, nous avons étendu le modèle des graphes conceptuels emboîtés pour les adapter à nos besoins. Les graphes conceptuels emboîtés utilisés dans notre modèle sont typés et on peut leur attacher, dans le cas de concept ou de relation concrets, un programme pour calculer sa valeur automatiquement [3] [4].

A un objet o de O , Π associe sa description selon les vues de Φ_v avec, $\Pi\Phi_v(o)=(\Pi V_1(o),\Pi V_2(o),\dots,\Pi V_p(o))$. Chaque $\Pi V_i(o)=G_{V_i}$ où G_{V_i} est un graphe conceptuel emboîté [15][16]. De même $\Pi\Phi_v(o)=(G_{V_1},G_{V_2},\dots,G_{V_p})$.

4 Objet multimédia

Un objet multimédia O_m est défini par $O_m=(o,I,L_c,L_r,L_o,G)$ avec $o\in O$ l'objet dans l'image, I l'image contenant l'objet o , L_c et L_r respectivement l'ensemble de concepts et de relations utilisés pour la description de l'objet o . G est un graphe conceptuel

emboîté décrivant le contenu de l'objet o , avec $G=\Pi\Phi_v(o)$. Il est à noter que, la description de o peut porter uniquement sur un sous-ensemble de vues de Φ_v . Dans ce cas, la description de o sur les autres vues de Φ_v est nulle.

Dans la Figure 2, nous présentons un objet multimédia et sa description réalisée avec des graphes conceptuels emboîtés. L'image est décrite par les quatre vues : vue physique, vue structurelle, vue spatiale et la vue sémantique. La vue structurelle est composée de trois objets désignés par les concepts "Gardon", "Woman" et "Child".

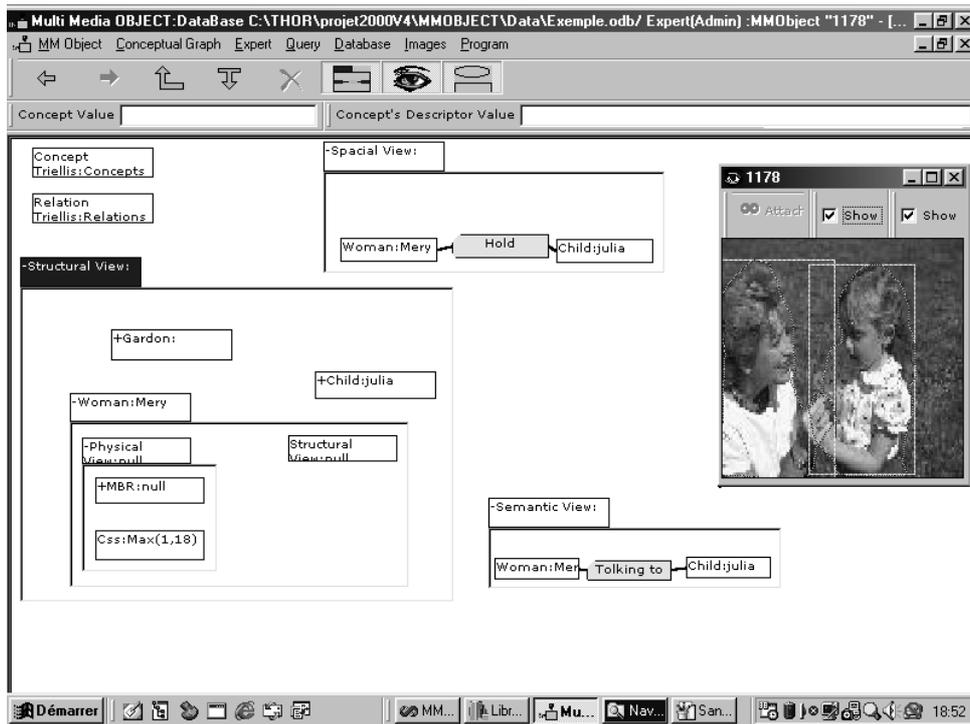


Figure 2 : Exemple d'objet multimédia et de sa description

5 Comparaison de concepts et de relations

Soit $T^*=T_c \cup t_c^i$, avec t_c^i le treillis de concepts concrets défini dans $V_i \in \Phi_v, 1 \leq i \leq p$, l'ensemble de concepts utilisés. Et, soit $R^*=T_r \cup t_r^i$, avec t_r^i le treillis de relations concrètes définies dans $V_i \in \Phi_v, 1 \leq i \leq p$, l'ensemble de relations utilisées. Les deux types de hiérarchies, concepts et relations, sont munies, respectivement, d'une distance d_c et d_r donnant la similarité entre deux concepts ou deux relations. Les deux distances sont définies de la manière suivante : $d_c : T^* \times T^* \rightarrow \mathfrak{R}$ avec :

$$\begin{cases} d_c(c_i, c_i) = d_c^i(c_i, c_i) & \forall c_i \in T^* \\ d_c(c_i, c_j) = d_c^j & \forall (c_i, c_j) \in T^* \times T^* \text{ et } c_j \text{ fils direct de } c_i \\ d_c(c_i, c_j) = \sum_{k=i}^{j-1} d_c(c_k, c_{k+1}) & \exists \zeta(c_i, c_{i+1}, \dots, c_{j-1}, c_j) \in T^* \\ d_c(c_i, c_j) = \infty & \text{autrement} \end{cases}$$

où \mathfrak{R} est l'ensemble des réels, ∞ désigne l'infini et $\zeta(c_i, c_{i+1}, \dots, c_{j-1}, c_j)$ le chemin continu de c_i à c_j de sorte que $\forall i \leq k \leq j-1$, c_{k+1} est fils direct de c_k . $d_c^i(c_i, c_i)$ est une distance quelconque permettant de comparer les valeurs de deux concepts de même nature. En général, $d_c^i(c_i, c_i) = 0 \forall c_i \in T_c$. De la même manière, on définit la distance entre les relations par $d_r : R^* \times R^* \rightarrow \mathfrak{R}$.

6 Calcul de requêtes

Soit $c \in T^*$ un concept. Un concept $c' \in T^*$ est considéré similaire à c si et seulement si $d_c(c, c') \leq s$ avec s un seuil de similarité pour le concept c . Nous définissons aussi l'ensemble $c^{s,*}$, l'ensemble des concepts similaires à c par $c^{s,*} = \{c' \in T^* / d_c(c, c') \leq s\}$. De la même manière, l'ensemble des relations similaires à une relation r , est $r^{s,*} = \{r' \in R^* / d_r(r, r') \leq s\}$. Soit $G=(C,R,U)$ un graphe

conceptuel et soit $G^{s,*}$ l'ensemble des graphes similaires à G par rapport à un seuil de similarité s . $G^{s,*}$ est donné par $G^{s,*}=(C^{s,*},R^{s,*},U^{s,*})$ avec $C^{s,*} = \bigcup_{c \in T_c} C^{s,*}$, $R^{s,*} = \bigcup_{r \in T_r} R^{s,*}$ et $U=U^{s,*}$. Soit $Q=(G_{v_1}, G_{v_2}, \dots, G_{v_p}, S)$ une requête, où G_{v_i} , avec $V_i \in \Phi_v$, les graphes conceptuels requêtes associés à chaque vue de la requête et S l'ensemble des seuils de similarité pour chaque type de concept ou de relation. L'ensemble des réponses pertinentes à Q est $R_Q=\{I \in B_i / \exists o \in O, o \subseteq I \text{ et } \Pi_{\Phi_v}(o) \in (G_{v_1}^{s,*}, G_{v_2}^{s,*}, \dots, G_{v_p}^{s,*})\}$. La comparaison de deux concepts ou relations de même nature se fait par rapport au seuil approprié défini dans S . Dans la figure 3, nous donnons un exemple de requête. Elle est exprimée sur les connaissances associées à l'image et sur la forme du concept "Humain". Elle est définie sur la vue structurelle, qui doit être composée de deux concepts "Humain" et "Child", et sur la vue spatiale. Le Concept "Humain" dans la vue structurelle a été développée au niveau de sa vue physique et, son concept contour a pour

valeur le contour donnée dans la Figure 3. On voit sur le coté droit, les images réponses à cette requête. Le test est réalisé sur une base d'images composée de 89 images issues de la base d'image de l'université de Berkley. Les réponses sont pertinentes et prometteuses. Pour le moment, une requête ne peut être définie que par rapport aux connaissances associées à l'image ou sur la forme des contours des objets qui apparaissent dans l'image. La recherche sur la forme des objets présents dans l'image est réalisée par la technique proposée dans [1].

7 Conclusion

Notre modèle est opérationnel. Une plate forme a été développée pour implémenter le model. La persistance des objets de la base de données est assurée par le SGBDOO ObjectStore. La plate forme est opérationnelle et les résultats obtenus sont très encourageants. Actuellement, nous travaillons sur l'extraction, par des algorithmes de segmentation, des objets pertinents d'une image en vue d'automatiser cette tâche.

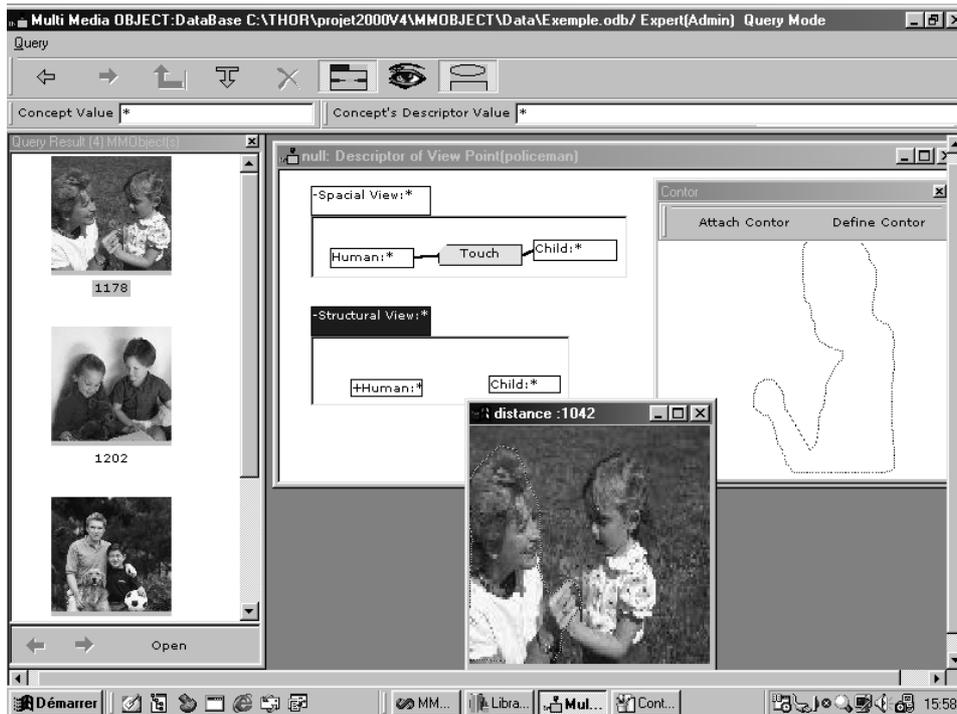


Figure 3 : Exemple de requête et de réponse

Références

[1] M. Daoudi, S. Matusiak, "Visual Image Retrieval by Multiscale Description of User Sketches", Journal of Visual Languages and Computing, Vol. 11, pp. 287-301, 2000.
 [2] M. Mugnier, M. Chein, "To Represent Knowledge and to Reason With Graphs", Research Report LIRMM, CNRS and Montpellier II university, 1995.

[3] R. Oulad Haj Thami, B.Frih, "Modeling Multi-Viewpoints of the Image Contents", MCSEAI 98, pp. 187-201, Tunis, Tunisia.
 [4] R. Oulad Haj Thami, B. Frih, J. Kouloumdjian, M. Rachik, "Content Image Modeling by Nested Conceptual Graphs", ISIVIC'2000, Vol. II, pp 296-309, April 17-20, 2000, Rabat.