

Connaissances a priori, ontologies et multimédia

Céline Hudelot

Ecole Centrale Paris, MAS

Plan

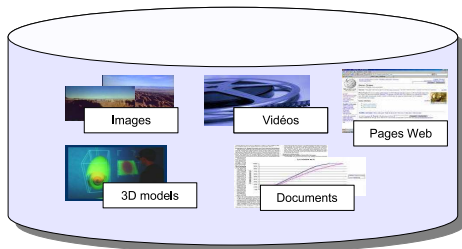
- 1 Présentation équipe MASDO
- 2 Formalisme de représentation de la connaissance
 - Logique de description floue morphologique
 - Illustration en imagerie cérébrale
 - Conclusion
- 3 Intégration de connaissance a priori dans l'analyse
- 4 Fusion d'information multimédia
- 5 Conclusion

Travaux de l'équipe MASDO

Contexte

Explosion des masses de données multimédia :

- Grand public : Flickr, YouTube, ...
- Domaines plus spécifiques : imagerie médicale, modèles 3D en CAO, données de simulations, ...



Travaux de l'équipe MASDO : hypothèse forte

Sémantique de cette image ?

- *Un objet blanc sur un fond vert*
- *Un insecte*
- *Une aleurode sur une feuille de rosier*



Conséquences

- La sémantique n'est pas dans l'image elle même.
- L'interprétation dépend de connaissances a priori.
- => Apport du **génie ontologique**.

Travaux de l'équipe MASDO

Axes de recherches

Représentation et intégration de connaissances et de raisonnements pour guider et améliorer la recherche sémantique d'informations.

Trouver des solutions pour la réduction du **fossé sémantique**.

- Au niveau des formalismes de représentation de la connaissance: extensions de la logique de description.
- Utilisation conjointe de ressources sémantiques et de ressources visuelles.
- Au niveau de la fusion *multi-niveaux* d'information.

Application aux **données multimédias** (images, vidéos, textes) ou **données scientifiques** (images médicales, données biologiques, données scientifiques).

Formalismes de représentation de la connaissance

Proposition

- Extension des logiques de descriptions pour combiner raisonnement quantitatif et qualitatif
- Application aux relations spatiales
- Travail en coopération avec I. Bloch (Telecom ParisTech) et J. Atif (Université de Guyane).

Raisonnement spatial et formalisation des relations spatiales

De nombreux travaux en intelligence artificielle :

- Principalement pour les relations topologiques (logique formelle, ex: méréologie)
- Ontologie de l'information spatiale (*Ontospace*)
- Inférence

Mais, en interprétation d'images

- Représentation (**Semi-quantitative**) (\Rightarrow évaluation numérique) pour une utilisation opérationnelle.
- Nécessité de prendre en compte l'imprécision et la subjectivité de l'information spatiale.

Problèmes

- Nécessité d'un formalisme combinant raisonnement spatial **qualitatif** et **quantitatif**.
- **Fossé sémantique** en interprétation d'images : *Manque de concordance entre les informations perceptuelles que l'on peut extraire de l'image et les concepts sémantiques.* [Smeulders,00]

Logiques de description

- Famille de logiques formelles pour représenter des informations structurées.
- Sémantique bien définie.
- Définies par un ensemble de concepts et des constructeurs de rôles.
- Compactes, expressives et bases du langage d'ontologie OWL.

Constructeur	Syntaxe	Exemple	Sémantique
Concept atomique	A	Humain	$A^I \subseteq \Delta^I$
individu	a	Lea	$a^I \in \Delta^I$
Top	\top	Thing	$\top^I = \Delta^I$
Bottom	\perp	Nothing	$\perp^I = \emptyset^I$
rôle atomique	r	has-age	$R^I \subseteq \Delta^I \times \Delta^I$
conjonction	$C \sqcap D$	Human \sqcap Male	$C^I \cap D^I$
disjonction	$C \sqcup D$	Male \sqcup Female	$C^I \cup D^I$
négation	$\neg C$	\neg Human	$\Delta^I \setminus C^I$
restriction existentielle	$\exists r.C$	\exists has-child.Girl	$\{x \in \Delta^I \mid \exists y \in \Delta^I : (x, y) \in R^I \wedge y \in C^I\}$
restriction universelle	$\forall r.C$	\forall has-child.Human	$\{x \in \Delta^I \mid \forall y \in \Delta^I : (x, y) \in R^I \Rightarrow y \in C^I\}$
restriction de valeur	$\exists r.\{a\}$	\exists has-child.{Lea}	$\{x \in \Delta^I \mid \exists y \in \Delta^I : (x, y) \in R^I \Rightarrow y = a^I\}$
restriction de nombre	$(\geq nR)$	$(\geq 3$ has-child)	$\{x \in \Delta^I \mid \{y \mid (x, y) \in R^I\} \geq n\}$
	$(\leq nR)$	$(\leq 1$ has-mother)	$\{x \in \Delta^I \mid \{y \mid (x, y) \in R^I\} \leq n\}$
Subsomption	$C \sqsubseteq D$	Man \sqsubseteq Human	$C^I \subseteq D^I$
définition de concept	$C \equiv D$	Father \equiv Man \sqcap \exists has-child.Human	$C^I = D^I$
assertion de concept	$a : C$	John:Man	$a^I \in C^I$
assertion de rôle	$(a, b) : R$	(John,Helen):has-child	$(a^I, b^I) \in R^I$

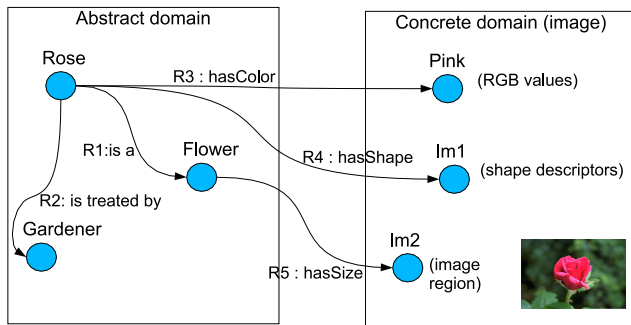
Domaines concrets

- Permettent l'intégration *des propriétés concrètes et quantitatives* (entiers, chaînes de caractères,...) d'objets réels avec de la connaissance conceptuelle [Baader,91].
- Un couple (Δ_D, Φ_D) où Δ_D est un ensemble et Φ_D un ensemble de noms de prédicats sur Δ_D . Chaque nom de prédicat P est associé à une arité n et à un prédicat n -aire $P^D \subseteq \Delta_D^n$

Exemples

- Domaine concret \mathcal{N}
 - ▶ domaine : entiers non négatifs
 - ▶ prédicats : \leq (prédicat binaire), $\leq n$ (prédicat unaire)
 - ▶ $\text{Person} \sqcap \exists \text{age} . \leq 20$ définit une personne dont l'âge est inférieur à 20
- Domaine concret \mathcal{AL} : intervalles de Allen
 - ▶ domaine : intervalles
 - ▶ prédicats : construits à l'aide des relations de Allen

Ontologies, domaines concrets et fossé sémantique



Idée

Chaque concept du domaine d'application est lié à sa représentation dans le domaine de l'image.

Incertitude et imprécision dans les ontologies

Domaines concrets flous

- Un couple (Δ_D, Φ_D) où Δ_D est le domaine d'interprétation et Φ_D un ensemble de prédicats flous d avec une arité n et une interprétation $d^D : \Delta_D^n \rightarrow [0, 1]$, qui est une relation floue n -aire sur Δ_D .

Exemple

Prédicat sur les entiers naturels :

- Jeune: $\mathbb{N} \rightarrow [0, 1]$ qui représente le degré de jeunesse
- Peut être représenté par une fonction d'appartenance trapezoidale par exemple

Notre contribution

Précédents travaux

- Une ontologie générique de relations spatiales
- Représentations floues des relations spatiales dans le domaine de l'image
- Utilisation des domaines concrets pour réduire le fossé sémantique
- Les liens sont construits explicitement

Objectifs

- Un nouveau formalisme qui approfondit l'exploitation des domaines concrets
- Utilisation combinée des logiques de descriptions avec des domaines concrets flous et de la **morphologie mathématique**

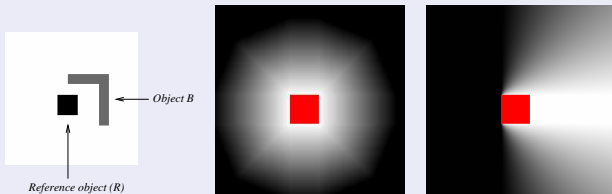
Représentations floues de relations spatiales

La représentation dépend de :

- La classe de la relation
- Le type de raisonnement

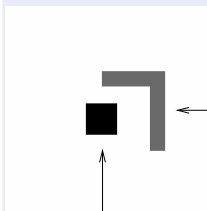
Deux questions pour la reconnaissance :

- Etant donné deux objets (flous), quel est le degré avec lequel une relation est satisfaite entre ces deux objets ?
- Etant donné un objet de référence, quelle est la région de l'espace dans laquelle une relation à cet objet est satisfaite (à un certain degré) ?



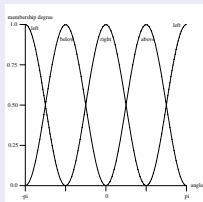
Représentation floue de relations directionnelles

Définie par dilatation floue

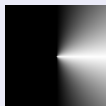


Reference object (R)

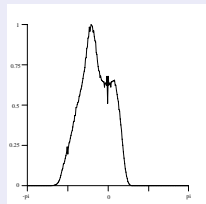
Object B



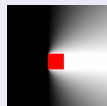
Ensembles flous



A droite de



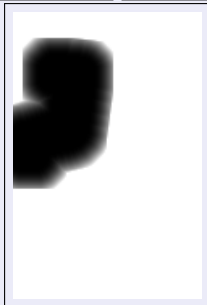
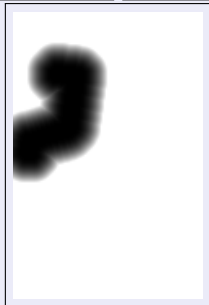
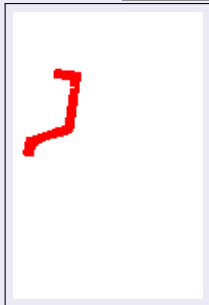
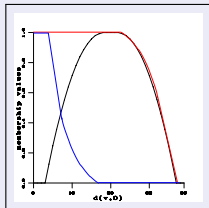
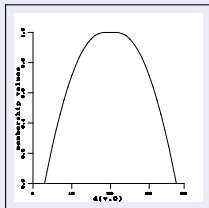
Histogramme d'angles



A droite de R

Représentation floue de relations de distance

Définie par dilatation floue



Cadre formel

Instanciation de la logique de description $\mathcal{ALCF}(\mathcal{D})$ avec le domaine concret

$\mathcal{S} = (\Delta_{\mathcal{S}}, \Phi_{\mathcal{S}})$

- $\Delta_{\mathcal{S}} = \mathcal{S}$ est un espace 2D ou 3D (espace de l'image, \mathbb{Z}^2 ou \mathbb{Z}^3)
- $\Phi_{\mathcal{S}}$ contient :
 - ▶ les prédicats unaires $\perp_{\mathcal{S}}$ et $\top_{\mathcal{S}}$ (\emptyset et $\Delta_{\mathcal{S}}$)
 - ▶ deux prédicats unaires flous μ et ν avec :
 - ★ pour un objet spatial X , $\mu(X) = \mu_X$ et $\forall x \in \mathcal{S}$, $\mu_X(x)$ est le degré avec lequel x appartient à la représentation spatiale de l'objet X dans l'image
 - ★ pour une relation spatiale R , $\nu(R) = \nu_R$ et ν_R représente l'élément structurant flou défini sur \mathcal{S} , i.e. la représentation floue spatiale de la relation spatiale R dans l'espace de l'image
 - ▶ deux prédicats flous binaires δ et ε avec
 - ★ $\delta_{\nu_R}^{\mu_X}$ dilatation floue :

$$\forall x \in \mathcal{S}, \delta_{\nu_R}(\mu_X)(x) = \delta_{\nu_R}^{\mu_X}(x) = \sup_{y \in \mathcal{S}} t(\nu_R(x - y), \mu_X(y))$$
 ($t = t$ -norme)
 - ★ $\varepsilon_{\nu_R}^{\mu_X}$ érosion floue :

$$\forall x \in \mathcal{S}, \varepsilon_{\nu_R}(\mu_X)(x) = \varepsilon_{\nu_R}^{\mu_X}(x) = \inf_{y \in \mathcal{S}} T(1 - \nu_R(y - x), \mu_X(y))$$
 ($T = t$ -conorme)

Inférence spatiale

Prédication d'existence : $\exists f.P$

- Toute relation spatiale abstraite et tout objet spatial abstrait sont associés à leur représentation floue dans le domaine concret par **hasFR**
- $\exists \text{ hasFR}.\mu_X$ restreint la région concrète de l'objet X à μ_X
- $\exists \text{ hasFR}.\nu_R$ restreint la région concrète de la relation R à ν_R
- $\exists \text{ hasFR}.\delta_{\nu_R}^{\mu_X}$ restreint la région concrète de la relation spatiale R_X à l'ensemble flou spatial obtenu par dilatation de μ_X par ν_R

Inférence spatiale

Prédicat de formation de rôle (Haarslev,99) : $\exists (u) (v) . P$

Concept R_X défini par :

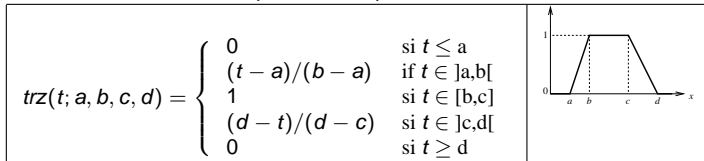
- $R_X \doteq \text{SpatialRelation} \sqcap \exists \text{ hasRO}.X \sqsubseteq \text{SpatialRelationWith}$
- $R_X \equiv \text{SpatialRelation} \sqcap \exists (\text{hasFR}, \text{hasRO}. \text{hasFR}). \lambda$ où λ est un prédicat binaire défini par les opérateurs de morphologie mathématique floue δ et ε
- $(\text{hasFR}, \text{hasRO}. \text{hasFR}). \delta \equiv \text{hasFR}. \delta_{\nu R}^{\mu X}$

Objets spatiaux qui satisfont une relation spatiale R par rapport à un objet de référence X

- $C \doteq \text{SpatialObject} \sqcap \text{hasSR}.R_X$
- axiomes : $C \sqsubseteq \exists \text{relationTo}.X$ et $C \sqsubseteq \text{SpatiallyRelatedObject}$

Exemples : relations de distance

- Fonction en trapèze $trz(x; a, b, c, d)$ définissant la sémantique de “proche de” :



- Élément structurant $\nu_{CLOSE.TO}$
- Close_to \doteq DistanceRelation $\sqcap \exists hasFR. \nu_{CLOSE.TO}$
- Close_to_X \doteq DistanceRelation $\sqcap \exists hasFR. \delta_{\nu_{CLOSE.TO}}^{\mu_X}$



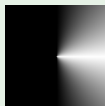
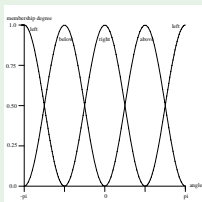
Exemples : relations directionnelles

Relation $In_Direction_α$

- Élément structurant flou $\forall x \in S, \nu_{IN_DIRECTION_α}(x) = f_α(|\angle(Ox, \vec{i}) - α|) \pmod{2π}$ déduit de l'ensemble flou $f_α(\theta)$ sur l'axe des angles
- $In_Direction_α_of_X \doteq DirectionalRelation \sqcap \exists hasFR. \delta_{\nu_{IN_DIRECTION_α}}^{\mu_X}$

Exemple : Right_of

Déduit de $f_0(\theta) = \cos^2(\theta)$ si $\theta \in [-\pi/2, \pi/2]$ et 0 sinon



Exemples : relations topologiques

Adjacence

- Exprimée comme une relation de distance, avec une sémantique de *très proche*
- \Rightarrow entre dans le même cadre, avec le même formalisme

Frontières

- externe : $\text{ExtB_X} = \text{SpatialObject} \sqcap \exists \text{hasFR}. (\delta_{\nu_0}^{\mu_X} \setminus \mu_X)$ où ν_0 est un élément structurant élémentaire
- interne : $\text{IntB_X} = \text{SpatialObject} \sqcap \exists \text{hasFR}. (\mu_X \setminus \varepsilon_{\nu_0}^{\mu_X})$

Relations méréotopologiques

- Partie : $P(X,Y) \doteq X \sqsubseteq Y$
- Recouvrement partiel : $\text{PO}(X,Y) \doteq (X \sqcap Y \neq \perp)$
- Déconnexion : $\text{DC}(X,Y) \doteq (X \sqcap Y = \perp)$
- Connexion externe : $\text{EC}(X,Y) \doteq (X \sqcap Y = \perp) \sqcap X_ADJ_TO_Y$
- Partie propre : $\text{PP}(X,Y) \doteq (X \sqsubseteq Y) \sqcap (\text{ExtB_X} \sqsubseteq Y)$;
- Partie propre tangentielle: $\text{TPP}(X,Y) \doteq (X \sqsubseteq Y) \sqcap (\text{ExtB_X} \sqcap \neg Y \neq \perp)$

Propriétés (1)

Déduites de propriétés de morphologie mathématique

Commutativité de la dilatation avec le sup :

- $\delta_\nu(\mu_{X_1}) \vee \delta_\nu(\mu_{X_2}) = \delta_\nu(\mu_{X_1} \vee \mu_{X_2})$
- $\delta_{\nu_1}(\mu_X) \vee \delta_{\nu_2}(\mu_X) = \delta_{\nu_1 \vee \nu_2}(\mu_X)$
- $\Rightarrow R_{X_1} \sqcup R_{X_2} \equiv R_{(X_1 \sqcup X_2)}$ et $R1_{X_1} \sqcup R2_{X_2} \equiv R12_{X_1 \sqcup X_2}$ où $R12$ a pour représentation floue $\nu_1 \vee \nu_2$

Relation partielle avec l'inf :

- $\delta_\nu(\mu_{X_1}) \wedge \delta_\nu(\mu_{X_2}) \geq \delta_\nu(\mu_{X_1} \wedge \mu_{X_2})$
- $R_{(X_1 \sqcap X_2)} \sqsubseteq R_{X_1} \sqcap R_{X_2}$

Propriétés (2)

Croissance :

- $\mu_{X_1} \leq \mu_{X_2} \Rightarrow \forall \nu \in \mathcal{F}, \delta_\nu(\mu_{X_1}) \leq \delta_\nu(\mu_{X_2})$
- $\nu_1 \leq \nu_2 \Rightarrow \forall \mu_X \in \mathcal{F}, \delta_{\nu_1}(\mu_X) \leq \delta_{\nu_2}(\mu_X)$
- $\Rightarrow X_1 \sqsubseteq X_2 \Rightarrow \forall R, R.X_1 \sqsubseteq R.X_2$ et $R_1 \sqsubseteq R_2 \Rightarrow \forall X, R_1.X \sqsubseteq R_2.X$

Itérativité :

- $\delta_{\nu_1}(\delta_{\nu_2}(\mu_X)) = \delta_{\delta_{\nu_1}(\nu_2)}(\mu_X)$
- $\Rightarrow R_1.(R_2.X) \equiv (R_1.R_2).X$, où $R_1.R_2$ est la relation ayant pour représentation floue $\delta_{\nu_1}(\nu_2)$

Propriétés (3)

Extensivité :

- $\nu(O) = 1 \Rightarrow \forall \mu_X \in \mathcal{F}, \mu_X \leq \delta_\nu(\mu_X)$
- $\Rightarrow X \sqsubseteq R_X$ pour toute relation définie par une dilatation avec un élément structurant contenant l'origine de \mathcal{S} (avec un degré d'appartenance de 1)

Dualité :

- Avec les définitions retenues pour la dilatation floue et l'érosion floue, on a : $\varepsilon_\nu(\mu_X) = 1 - \delta_\nu(1 - \mu_X)$
- \Rightarrow liens entre certaines relations
- exemple : $\text{IntB}_X = \mu_X \setminus \varepsilon_{\nu_0}^{\mu_X} = \mu_X \wedge \delta_{\nu_0}^{1-\mu_X} = \delta_{\nu_0}^{1-\mu_X} \setminus (1 - \mu_X)$, donc $\text{IntB}_X \equiv \text{ExtB}_{\neg X}$

Connaissances en imagerie cérébrale

Concepts :

- **cerveau** : partie du système nerveux central, localisé dans la tête
- **noyau caudé** : un des noyaux gris centraux, situé dans le télencéphale
- **gliome** : tumeur du système nerveux central provenant de cellules gliales
- ...

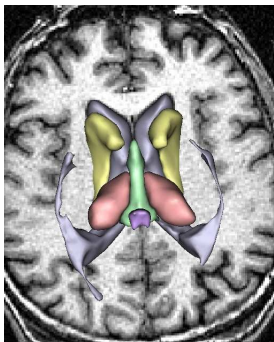
Organisation spatiale :

- le **noyau caudé gauche** est **dans** l'**hémisphère gauche**
- il est **proche** du **ventricule latéral**
- il est **extérieur au (à gauche du) ventricule latéral gauche**
- il est **au-dessus** du **thalamus**
- ...

Reconnaissance séquentielle de structures anatomiques

Objectif

Reconnaissance séquentielle de structures anatomiques en utilisant **des informations spatiales** (relations par rapport à des structures déjà reconnues)

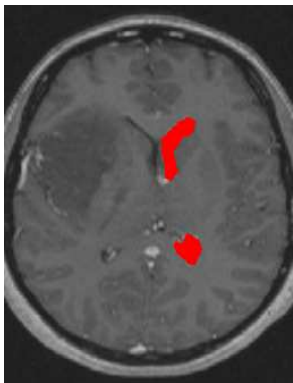


Description de connaissances anatomiques

Tbox :

- $\text{AnatomicalStructure} \sqsubseteq \text{SpatialObject}$
- $\text{GN} \sqsubseteq \text{AnatomicalStructure}$
- $\text{RLV} \doteq \text{AnatomicalStructure} \sqcap \exists \text{hasFR}.\mu_{RLV}$
- $\text{LLV} \doteq \text{AnatomicalStructure} \sqcap \exists \text{hasFR}.\mu_{LLV}$
- $\text{LV} \equiv \text{RLV} \sqcup \text{LLV}$
- $\text{LV} \equiv \text{RLV} \sqcup \text{LLV}$
- $\text{Right_of} \doteq \text{DirectionalRelation} \sqcap \exists \text{hasFR}.\nu_{IN_DIRECTION.O}$
- $\text{Close_to} \doteq \text{DistanceRelation} \sqcap \exists \text{hasFR}.\nu_{CLOSE.TO}$
- $\text{Right_of_RLV} \doteq \text{DirectionalRelation} \sqcap \exists \text{hasReferentObject.RLV} \sqcap \exists \text{hasFR}.\delta_{\nu_{IN_DIRECTION.O}}^{\mu_{RLV}}$
- $\text{Close_To_RLV} \doteq \text{DistanceRelation} \sqcap \exists \text{hasReferentObject.RLV} \sqcap \exists \text{hasFR}.\delta_{\nu_{CLOSE.TO}}^{\mu_{RLV}}$
- $\text{RCN} \doteq \text{GN} \sqcap \exists \text{hasSR} . (\text{Right_of_RLV} \sqcap \text{Close_To_RLV})$
- $\text{CN} \doteq \text{GN} \sqcap \exists \text{hasSR} . (\text{Close_To_LV})$
- $\text{CN} \equiv \text{RCN} \sqcup \text{LCN}$

Exemple



Abox :

- c_1 : RLV , (c_1, μ_{S_1}) : hasFR
- r_1 : Right_of,
 $(r_1, \nu_{IN_DIRECTION_0})$: hasFR
- r_2 : Close_to, (r_2, ν_{CLOSE_TO}) :
hasFR

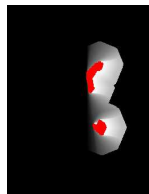
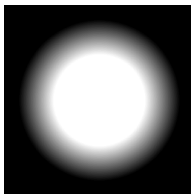
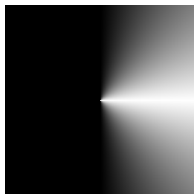
Exemple 1

Objectif

- Déterminer des contraintes spatiales dans le domaine de l'image que doit vérifier une instance C_2 du noyau caudé gauche
- \Rightarrow Contraintes dans le domaine concret garantissant la satisfiabilité des assertions C_2 : RCN , (C_2, μ_{S_2}) : hasFR

Résultats obtenus par inférence en utilisant les propriétés

$$\mu_{S_2} \subseteq \delta_{\nu_{IN_DIRECTION_0}}^{\mu_{I_1}} \cap \delta_{\nu_{CLOSE_TO}}^{\mu_{I_1}}$$



Exemple 2

Objectif

- Trouver toutes les instances du noyau caudé dans l'image
- $CN \doteq GN \sqcap \exists \text{ hasSR.}(\text{Right_of_LV} \sqcup \text{Left_of_LV})$

Propriété de disjonction

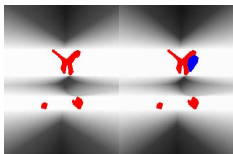
$\text{Right_of_LV} \sqcup \text{Left_of_LV} \equiv \text{SpatialRelation} \sqcap \exists \text{ hasFR.} \delta_{\nu_{\text{RIGHT_OF}} \vee \nu_{\text{LEFT_OF}}}^{\mu_{LV}}$

Résultats

L'espace de recherche pour les noyaux caudés est calculé selon :

$\delta_{\nu_{\text{RIGHT_OF}} \vee \nu_{\text{LEFT_OF}}}(\mu_{LV})$

ou de manière équivalente : $\delta_{\nu_{\text{RIGHT_OF}}}(\mu_{LV}) \vee \delta_{\nu_{\text{LEFT_OF}}}(\mu_{LV})$



Conclusion

- Ontologie de relations spatiales
- Représentations floues d'objets et de relations dans un domaine concret (l'espace de l'image)
- Extension d'une logique de description avec des opérateurs morphologiques \Rightarrow plus d'expressivité, plus d'outils de raisonnement et d'inférence
- \Rightarrow Nouveau cadre de raisonnement sur des informations spatiales structurelles
- Illustration en imagerie cérébrale

Perspectives et questions ouvertes

- Prendre en compte les cas pathologiques
- Autres applications : imagerie satellitaire
- Généralisation à d'autres concepts que les relations spatiales.

Annotation sémantique d'images

Travaux de thèse de Nicolas James

Contexte

- Limites des approches de recherche d'images basés sur le texte : subjectivité des annotations
- Limites des approches de recherche par le contenu : fossé sémantique
- Une alternative : annotation automatique d'images par des concepts sémantiques. Une littérature abondante :
 - ▶ Apprentissage statistique et classification : [5, 2]
 - ▶ Approches probabilistes : [8, 3, 1, 7]
 - ▶ Approches basées sur des graphes : [10]
 - ▶ ...

Annotation sémantique d'images

De nombreux challenges



(a) beach people sand desert snow



(b) beach people sand desert snow

Des annotations imprécises, souvent incomplètes et non **sémantiques** :

- Besoin d'**améliorer** et de **sémantiser** les annotations.

Annotation sémantique d'images

De nombreux challenges



Image issue de la base de test
du Efficient Content-Based Retrieval
Group de l'Université de Washington.

<http://www.cs.washington.edu/research/imagetdatabase>

Annotation automatique	
Mot-clés	Nombre de sens dans WordNet
Fern	1
Street	5
Pole	10
Sign	11
House	12
Tree trunks	1
Bench	7
People	4
Tree	3

Problème des sens multiples : besoin de **désambiguïser**

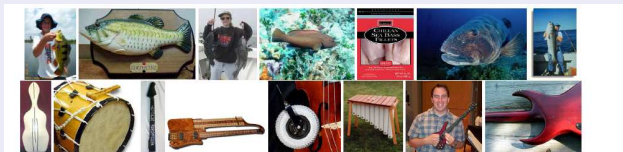
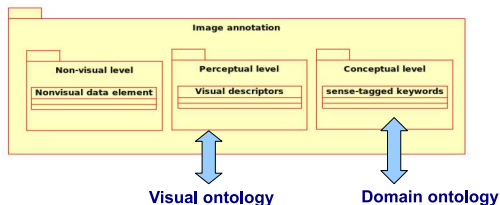


Figure 2: *Which fish or instruments are BASS?* Image sense annotation is more vague and subjective than in text.

Objectifs

- Utilisation conjointe de connaissances a priori et de bases d'apprentissage visuelles pour l'annotation d'images :
 - ▶ Des annotations **sémantiques** (relations sémantiques entre les différents mots-clés)
 - ▶ Des annotations **multi-niveaux** (ontologie de descripteurs et de concepts visuels : par exemple MPEG-7)



Approche globale



document



Système d'annotation automatique
(de type Content-Based)



bag-of-tags

$$\begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}$$


bases de connaissances (ontologies, thésaurus)



metadata

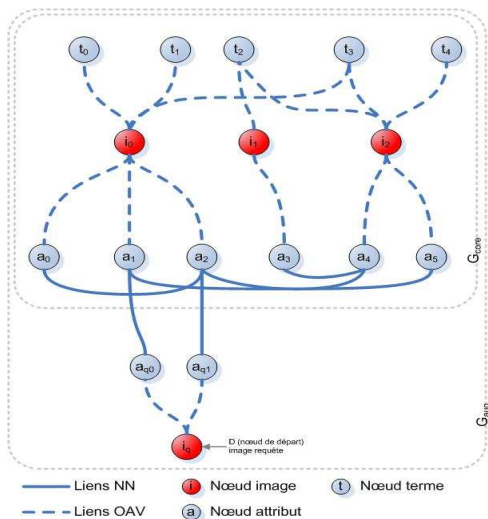


désambiguïsation,
expansion sémantique



Annotation automatique

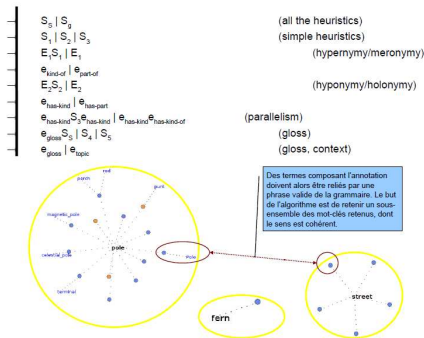
Basée sur un algorithme de marche aléatoire dans un graphe ([10])



Enrichissement sémantique de l'annotation

Adaptation d'un algorithme de désambiguïsation pour le texte (SSI, Structural Semantic Interconnexion) et adaptation pour l'annotation d'images :

- Adaptation de la grammaire de manière explicite et par apprentissage dans une base d'images : pondération des relations sémantiques à prendre en compte
- Ajout d'un a priori visuel dans l'étape de désambiguïsation (contexte visuel: Label Me).



Perspectives

- Propagation d'annotations sémantiques dans le graphe.
- Prise en compte des relations spatiales pour une annotation *structurelle et spatiale*.
- Recherche d'images en se basant sur les ontologies : requêtes complexes plus proches de l'intention de l'utilisateur
- Prise en compte des relations sémantiques dans le graphe et pour la marche aléatoire

Contexte

Thèse de Bassem Makni, en coopération avec le CEA LIST (équipe LIC2M) Début : Janvier 2009

Principe de la fusion

- Combinaison d'informations hétérogènes issues de plusieurs sources afin d'améliorer la prise de décision (d'après I. Bloch et H. Maitre)
- Plusieurs approches :
 - ▶ Approches probabilistes
 - ▶ Approches basées sur la théorie de l'évidence
 - ▶ Approches basées sur la logique floue et la théorie des possibilités

En indexation et recherche sémantique d'information multimédia

Deux approches

- **Fusion tardive :**
 - ▶ Autant de classifieurs que de vecteurs de caractéristiques
 - ▶ Un score par modalité
 - ▶ Fusion des scores (fonction explicite ou nouveau classifieur)
- **Fusion précoce :**
 - ▶ *Concaténation* des vecteurs de caractéristiques.
 - ▶ Classifieur unique

Fusion précoce

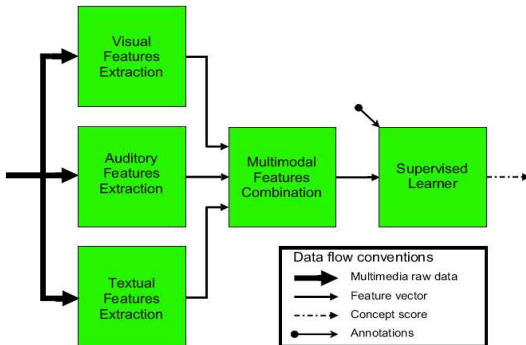


Figure: D'après Snoek, [11]

Fusion tardive

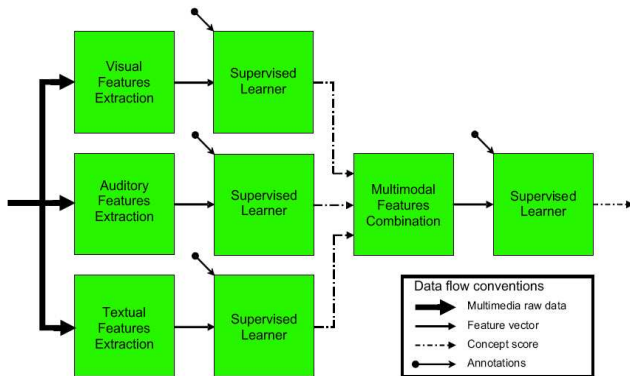


Figure: D'après Snoek, [11]

Directions envisagées

- Quel(s) modèle(s) de fusion ?
- Apport de connaissances a priori dans le processus de fusion
- Vers des ontologies multimédia perceptuellement ancrées :
 - ▶ Création de l'ancrage : apport des techniques d'apprentissage statistique
 - ▶ Ontologies dynamiques : maintenir et mettre à jour l'ancrage.

Une thématique à la mode

Approche *Web semantique*

Ontologies multimédia pour l'annotation multi-niveaux.

- FUSION (Jane Hunter) : ontologie MPEG-7 en RDFS ^a
- AceMedia : deux ontologies en RDFS : Visual Descriptor Ontology et Multimedia Structure Ontology ^b
- SmartWeb : une ontologie en OWL DL ^c
- BOEMIE : deux ontologies en OWL DL : Multimedia Content Ontology + Multimedia Descriptor Ontology ^d
- DS-MIRF : Core Multimedia ontology ^e
- COMM (Core Ontology for Multimedia) ^f

^a<http://metadata.net/sunago/prototypes.htm>

^b[://www.acemedia.org/aceMedia/results/Ontologies.html](http://www.acemedia.org/aceMedia/results/Ontologies.html)

^chttp://www.smartweb-project.org/ontology_en.html

^dhttp://www.boemie.org/public_docs

^e<http://comm.semanticweb.org/examples>

^f<http://comm.semanticweb.org/examples>

Une thématique à la mode

Approche *classification*

- LSCOM : Large Scale Concept Ontology for Multimedia : un ensemble de concepts (relations sémantiques faibles) avec des vidéos annotés ^a
- Label Me : un ensemble d'images annotés (lien avec les synsets Wordnet)^b
- Annotation multi-labels [13, 4] ...
- ...

^a<http://www.lsc.com.org/>

^b<http://people.csail.mit.edu/torr/alba/research/LabelMe/wordnet/>

Une thématique à la mode

Approche représentation et raisonnement formel

- Logiques de description floues [12, 9]
- ...

Mais,

Des mondes qui se *parlent peu*.

Références I



Kobus Barnard and Matthew Johnson.

Word sense disambiguation with pictures.

Artif. Intell., 167(1-2):13–30, 2005.



Gustavo Carneiro and Nuno Vasconcelos.

Formulating semantic image annotation as a supervised learning problem.

In *CVPR '05: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Volume 2*, pages 163–168, Washington, DC, USA, 2005. IEEE Computer Society.



P. Duygulu, K. Barnard, J.F.G. de Freitas, and D.A. Forsyth.

Object recognition as machine translation: Learning a lexicon for a fixed image vocabulary.

Lecture Notes in Computer Science, pages 97–112, 2002.

Références II



Y. Gao and J. Fan.

Incorporating concept ontology to enable probabilistic concept reasoning for multi-level image annotation.

In *Proceedings of the 8th ACM international workshop on Multimedia information retrieval*, pages 79–88. ACM New York, NY, USA, 2006.



Yuli Gao, Jianping Fan, Xiangyang Xue, and Ramesh Jain.

Automatic image annotation by incorporating feature hierarchy and boosting to scale up svm classifiers.

In *MULTIMEDIA '06: Proceedings of the 14th annual ACM international conference on Multimedia*, pages 901–910, New York, NY, USA, 2006. ACM.



N. Loeff, C.O. Alm, and D.A. Forsyth.

Discriminating image senses by clustering with multimodal features.

In *Proceedings of the COLING/ACL on Main conference poster sessions*, pages 547–554. Association for Computational Linguistics Morristown, NJ, USA, 2006.

Références III



Florent Monay and Daniel Gatica-Perez.

On image auto-annotation with latent space models.

In *MULTIMEDIA '03: Proceedings of the eleventh ACM international conference on Multimedia*, pages 275–278, New York, NY, USA, 2003. ACM.



Y. Mori, H. Takahashi, and R. Oka.

Image-to-word transformation based on dividing and vector quantizing images with words.

In *First International Workshop on Multimedia Intelligent Storage and Retrieval Management*, 1999.



Bernd Neumann and Ralf Möller.

On scene interpretation with description logics.

Image Vision Comput., 26(1):82–101, 2008.



Jia-Yu Pan, Hyung-Jeong Yang, Christos Faloutsos, and Pinar Duygulu.

Gcap: Graph-based automatic image captioning.

cvprw, 09:146, 2004.

Références IV

 C.G.M. Snoek, M. Worring, and A.W.M. Smeulders.

Early versus late fusion in semantic video analysis.

In *Proceedings of the 13th annual ACM International Conference on Multimedia*, pages 399–402. ACM New York, NY, USA, 2005.

 G. Stoilos, G. Stamou, V. Tzouvaras, J.Z. Pan, and I. Horrocks.

A fuzzy description logic for multimedia knowledge representation.

In *Proc. of the International Workshop on Multimedia and the Semantic Web*, pages 12–19, 2005.

 Anne-Marie Tousch, Stéphane Herbin, and Jean-Yves Audibert.

Semantic lattices for multiple annotation of images.

In *MIR '08: Proceeding of the 1st ACM international conference on Multimedia information retrieval*, pages 342–349, New York, NY, USA, 2008. ACM.