

LIP **UPMC**
PARIS UNIVERSITAS

Utilisation de concepts visuels et de la diversité visuelle pour améliorer la recherche d'images

Sabrina Tollari, Marcin Detyniecki, Ali Fakeri Tabrizi, Christophe Marsala, Massih-Reza Amini, Patrick Gallinari

Université Pierre et Marie Curie – Paris 6
UMR CNRS 7606 - LIP6

CORIA 2009, le 5 mai 2009

Projet AVEIR ANR-06-MDCA-002

1

Plan

- Problématique : améliorer la recherche d'images basée sur le texte en utilisant :
 - Des concepts visuels
 - La diversification visuelle des résultats
- Contexte : campagne d'évaluation ImageCLEF
- Propositions :
 - Détecter les concepts visuels par forêt d'arbres de décision flous
 - Utiliser les concepts visuels pour filtrer les résultats de la recherche d'images par le texte en utilisant WordNet
 - Améliorer la diversité des résultats pertinents par clustering visuel

2

Problématique : améliorer la recherche d'images basée sur le texte en utilisant des concepts visuels

3

Détection de concepts visuels dans des images généralistes

- Proposition :
 - Apprendre des forêts d'arbres de décision flous pour détecter les concepts visuels
 - Pour chaque image et pour chaque concept, on obtient un score indiquant la présence du concept dans l'image
 - Analyser les cooccurrences de concepts pour déterminer les relations entre les concepts
 - Découvertes de relations d'exclusion et d'implication entre concepts à partir de la matrice de cooccurrences
 - Utiliser ces relations pour améliorer la détection des concepts
 - Applications de règles pour filtrer les scores des arbres

4

Tâche « Visual Concept Detection Task » (VCDT) de ImageCLEF 2008

outdoor person day vegetation animal

concepts ?

- 17 concepts visuels en partie hiérarchisés
- 2k images d'apprentissage :
 - chaque image d'apprentissage est associée à plusieurs concepts
- 1k images de test
- ⇒ Problème multi-classes multi-étiquettes avec classes hiérarchisées
- ⇒ Annotation automatique d'images

5

Forêt d'arbres de décision flou

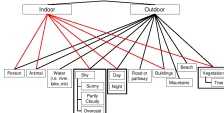
- Pour chaque concept :
 - une forêt de 50 arbres flous* (critère entropique flou et agrégation par somme) est apprise à partir des images d'apprentissage
 - Un seuil t est fixé pour prendre la décision d'annoter ou pas une image de test par ce concept

* Logiciel Salammbô, Christophe Marsala, Apprentissage inductif en présence de données imprécises : Construction et utilisation d'arbres de décision flous, thèse de doctorat de l'université Paris 6, 1998

6

VC DT

Comment découvrir les relations ?

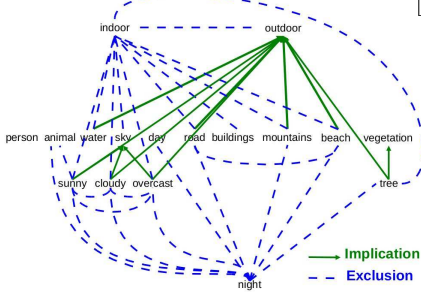


- Les arbres de décisions apprennent chacun des concepts indépendamment, or les concepts sont dépendants entre eux
- Relation d'exclusion :
 - A partir de la matrice de cooccurrence, nous déterminons les concepts qui ne sont jamais co-occurents
- Relation d'implication (ou de nécessité) :
 - Nous construisons une matrice : présence de A versus absence de B, pour tous couples de concepts (A,B)
 - Quand dans la matrice, on n'a jamais : A et non(B) alors A implique B
 - car $\text{non}(A \text{ et non}(B)) \rightarrow \text{non}(A) \text{ ou } B \rightarrow A \Rightarrow B$
 - Exemple : on n'a jamais *Tree* et non(*Vegetation*)

7

VC DT

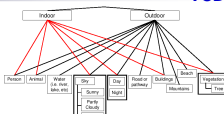
Relations découvertes automatiquement



8

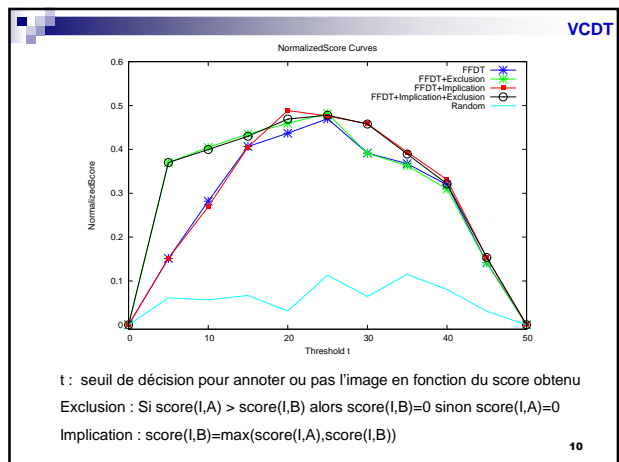
VC DT

Comment utiliser ces relations ?



- Exclusion :
 - $\text{COOC}(A,B)=0$
 - Si $\text{score}(I,A) > \text{score}(I,B)$ alors $\text{score}(I,B)=0$ sinon $\text{score}(I,A)=0$
 - Exemple :
 - Avant : outdoor=0.8 et indoor=0.4
 - Après : outdoor=0.8 et indoor=0
- Implication :
 - $\text{COOC}(A,\text{non } B)=0$
 - $\text{score}(I,B)=\max(\text{score}(I,A),\text{score}(I,B))$
 - Exemple :
 - Avant : cloudy=0.8 et sky=0.4
 - Après : cloudy=0.8 et sky=0.8

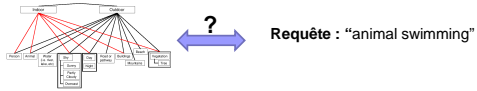
9



ImageCLEFphoto

Utilisation des concepts visuels pour améliorer la recherche d'images

- Comment utiliser les résultats de détection de concepts visuels pour améliorer la recherche textuelle d'images ?



Requête : "animal swimming"

- Proposition :
 - Filtrer à l'aide des scores des arbres de décision les résultats obtenus par la recherche textuelle en fonction de la présence des concepts dans :
 - Les mots de la requête (filtrage directe)
 - Les mots de la requête élargis à l'aide de la relation de synonymie de WordNet (filtrage WN)


11

ImageCLEFphoto

Exemple de filtrage directe

Requête : "seals near water"

Filtrage en fonction des scores des arbres de décision du concept « water »



12

ImageCLEFphoto

Exemple de filtrage en utilisant WordNet

- Requête : "animal swimming"
 - Animal: organism, plankton, mascot, fungus, ...
 - Swimming: bathe, diving, floating, surfing, **water** sport, ...
- Filtrage en fonction des scores des arbres de décision du concept « animal » et du concept « water »



Rang 1



Rang 4



Rang 11

13

ImageCLEFphoto

Résultats obtenus lors de ImageCLEFphoto 2008

- ImageCLEFphoto 2008 : 20 000 images avec du texte associé, 39 requêtes multimédia
- Par filtrage directe, 11 requêtes modifiées et 7 concepts utilisés
- Par le filtrage utilisant WordNet, 25 requêtes modifiées et 9 concepts utilisés

Texte	Méthode de filtrage		Les 39 requêtes		Les 11 requêtes filtrées	
	Directe	WN	P20 (gain %)	Nombre de requêtes	P20 (gain %)	
Modèle de langues			0.185 (-)	11	0.041 (-)	
	X		0.195(+6)	11	0.077(+88)	
	X	X	0.176(-5)	25	0.134 (-9)	
TF-IDF			0.250 (-)	11	0.155 (-)	
	X		0.269(+8)	11	0.223(+44)	
	X	X	0.258(+4)	25	0.226(+8)	

14

Diversité

Problématique : Amélioration de la diversité des résultats



15

Diversité

Diversification des résultats

- Exemples de requête dans ImageCLEFphoto 2008 :
 - Requête : « animal swimming » Critère de diversité : « animal »
 - Requête : « church with more than two towers » Critère : « city »
 - Autres exemples de critère : city, country, venue, landmark, location, tourist attraction, animal, bird, weather, vehicule, sport, group composition...

16

Diversité

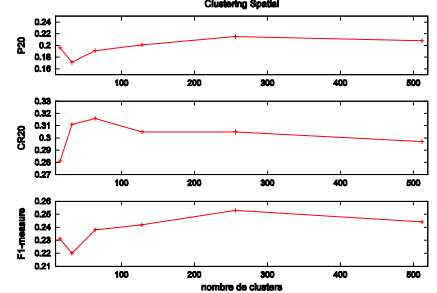
Diversification visuelle et aléatoire

- Diversification visuelle : DIVVISU
 - Binarisation de l'histogramme Hue de HSV global pour obtenir des clusters visuels
 - (0.2 0.7 0.1 0.6 0.6 0.9 0.1 0.6) → binarisation → 01011101
 - Avantage : méthode très rapide
 - Filtrage des images pour que n'apparaissent jamais deux images appartenant au même cluster dans les 20 premiers résultats
- Diversification aléatoire : DIVALEA
 - Permutation aléatoire sur les a premiers résultats

17

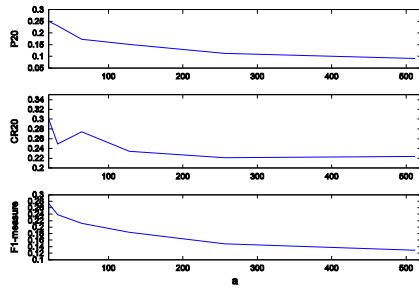
Diversité

Diversification visuelle par clustering spatial : DIVVISU

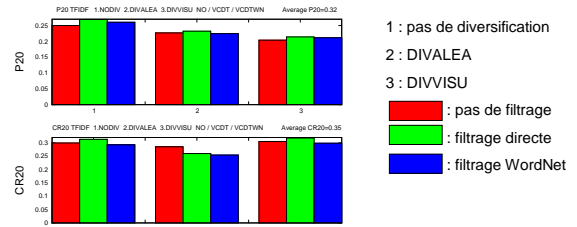


18

Diversification aléatoire : DIVALEA



ImageCLEFphoto 2008 : Résultats sur la diversité



- 1 : pas de diversification
- 2 : DIVALEA
- 3 : DIVISU
- Red : pas de filtrage
- Green : filtrage directe
- Blue : filtrage WordNet

Les diversifications aléatoires et visuelles donnent des précisions à 20 (P20) plus faibles, mais la diversification visuelle donne un cluster recall (CR20) légèrement meilleur

Conclusion

- Utiliser les concepts visuels pour améliorer la recherche d'image est une piste intéressante, mais la difficulté est de trouver la correspondance entre la requête textuelle et les concepts visuels
- L'utilisation des cooccurrences n'a pas été concluante dans notre cas, peut-être à cause du manque de données d'apprentissage
- La diversification visuelle basée sur le clustering spatial permet de diversifier les résultats rapidement, mais n'est pas adapté pour tous les critères de diversité
- Perspectives : construire un modèle prenant en compte les concepts dans les images requêtes

MERCI DE VOTRE ATTENTION

QUESTIONS ?

