

Soutenance de Thèse

Création et utilisation de vocabulaires visuels pour la catégorisation
d'images et la segmentation de classes d'objets

Diane Larlus

Équipe LEAR, INRIA Grenoble - INPG

Vendredi 28 novembre 2008

Composition du Jury :

- Pr. Roger Mohr, président
- Pr. Jean-Pierre Cocquerez, rapporteur
- Pr. Patrick Perez, rapporteur
- Pr. Matthieu Cord, examinateur
- Dr. Stéphane Herbin, examinateur
- Pr. Frédéric Jurie, directeur de thèse



Plan de la présentation

- 1 Contexte de l'étude
- 2 Représentation des images
- 3 Principales contributions de la thèse
- 4 Vocabulaires visuels discriminants
- 5 Segmentation d'objets
- 6 Conclusions et perspectives

Contexte de l'étude

Problématique :

- Interprétation des images

Problématique :

- Interprétation des images
- Reconnaissance des objets

Problématique :

- Interprétation des images
- Reconnaissance des objets
 - ▶ Reconnaissance des instances d'objets
 - ▶ Reconnaissance des catégories d'objets

Problématique :

- Interprétation des images
- Reconnaissance des objets
 - ▶ Reconnaissance des instances d'objets
 - ▶ Reconnaissance des catégories d'objets

Trois tâches associées à la reconnaissance

- catégorisation d'image



Mouton ?



Oui

Trois tâches associées à la reconnaissance

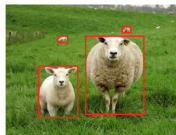
- catégorisation d'image



Mouton ?



- détection d'objets



Mouton ?

4

Trois tâches associées à la reconnaissance

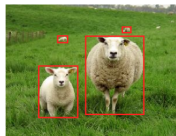
- catégorisation d'image



Mouton ?



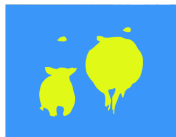
- détection d'objets



Mouton ?

4

- segmentation d'objets



Mouton ?



Difficulté de la reconnaissance

Les images appartenant à la même catégorie présentent de grandes variations d'apparences

- Variations d'échelle et de point de vue



Difficulté de la reconnaissance

Les images appartenant à la même catégorie présentent de grandes variations d'apparences

- Variations d'échelle et de point de vue
- Occultations



Difficulté de la reconnaissance

Les images appartenant à la même catégorie présentent de grandes variations d'apparences

- Variations d'échelle et de point de vue
- Occultations
- Changement d'illumination



Difficulté de la reconnaissance

Les images appartenant à la même catégorie présentent de grandes variations d'apparences

- Variations d'échelle et de point de vue
- Occultations
- Changement d'illumination
- Présence de fond encombré



Difficulté de la reconnaissance

Les images appartenant à la même catégorie présentent de grandes variations d'apparences

- Variations d'échelle et de point de vue
- Occultations
- Changement d'illumination
- Présence de fond encombré
- Variations intra-classe



Difficulté de la reconnaissance

- Construction d'un modèle commun : **catégorie vélo**



Difficulté de la reconnaissance

- Construction d'un modèle commun : catégorie vélo



- Comment choisir une représentation permettant de construire un tel modèle ?

Représentation des images

Représentation des images

- pixels peu informatifs



Représentation des images

- pixels peu informatifs
- représentations globales
- représentations locales

Représentation des images

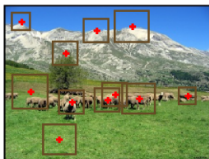
- pixels peu informatifs
- représentations globales
- représentations locales

- ▶ **représentation par sac-de-mots**

Csurka *et al*, ECCV workshop, 2004

la représentation par sac-de-mots

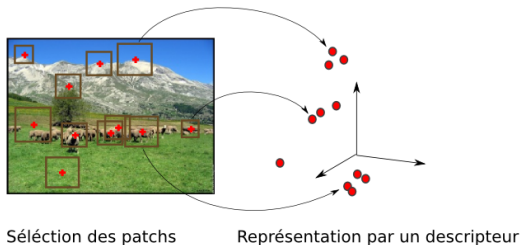
- Pour une image d'apprentissage



Sélection des patches

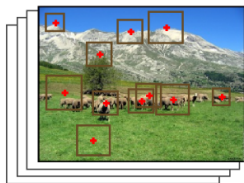
la représentation par sac-de-mots

- Pour une image d'apprentissage

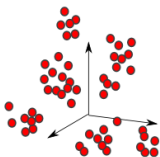


la représentation par sac-de-mots

- Pour toutes les images d'apprentissage



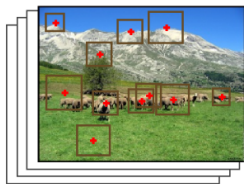
Sélection des patches



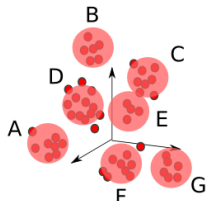
Représentation par un descripteur

la représentation par sac-de-mots

- Pour toutes les images d'apprentissage



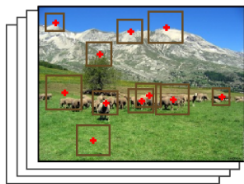
Sélection des patches



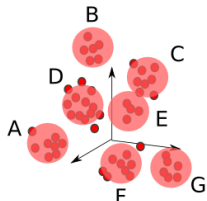
Création des mots visuels

la représentation par sac-de-mots

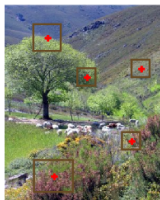
- Pour toute nouvelle image



Sélection des patches

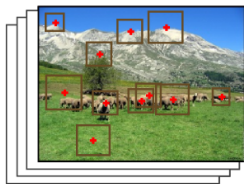


Création des mots visuels

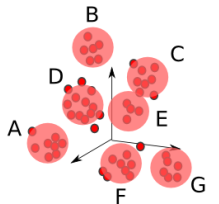


la représentation par sac-de-mots

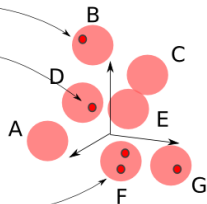
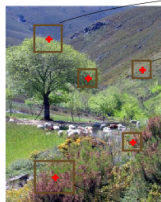
- Pour toute nouvelle image



Sélection des patches

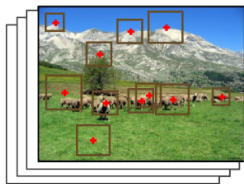


Création des mots visuels

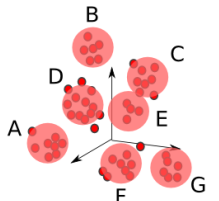


la représentation par sac-de-mots

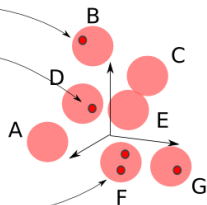
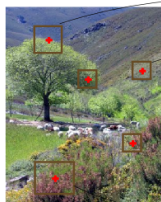
- Pour toute nouvelle image



Sélection des patches

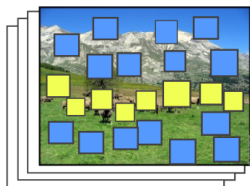


Création des mots visuels

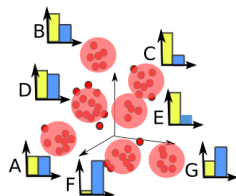


Classification des patches

- Acquisition de statistiques sur les mots visuels à partir d'images d'apprentissage segmentées.



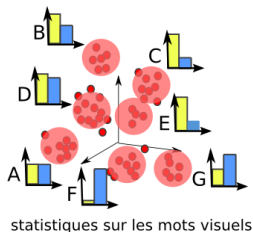
Vérité terrain : classe des patches



statistiques sur les mots visuels

Classification des patches

- Utilisation de ces statistiques pour prédire la classes des patches d'une nouvelle image

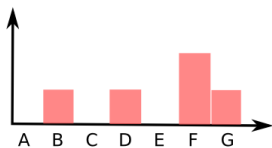
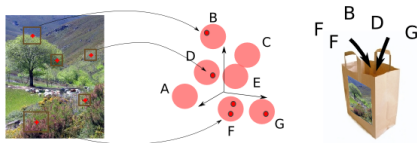


nouvelle image

- utilisée pour la segmentation d'objets

Classification des images

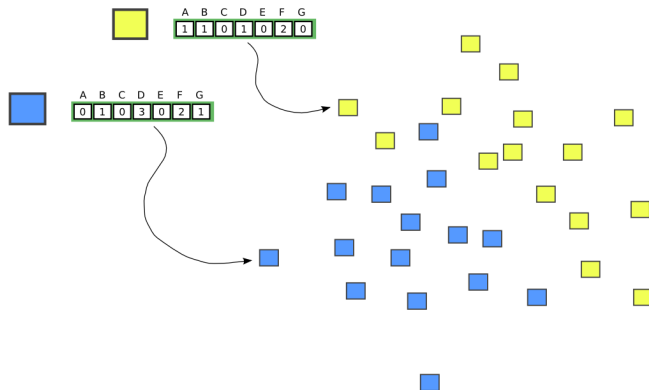
- Construction d'un vecteur de représentation pour l'image
 - ▶ **histogramme d'occurrences**



A	B	C	D	E	F	G
0	1	0	1	0	2	1

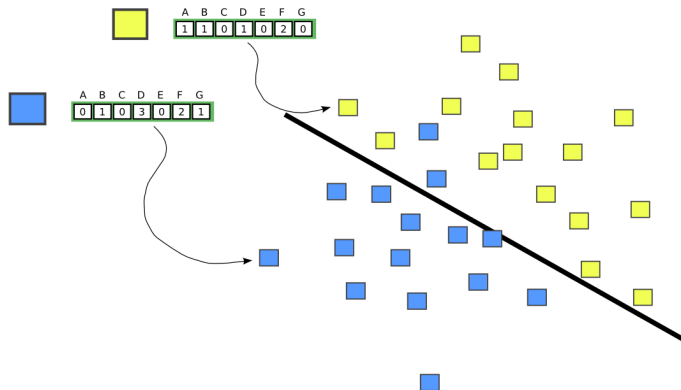
Classification des images

- Les images sont considérées dans ce nouvel espace de représentation



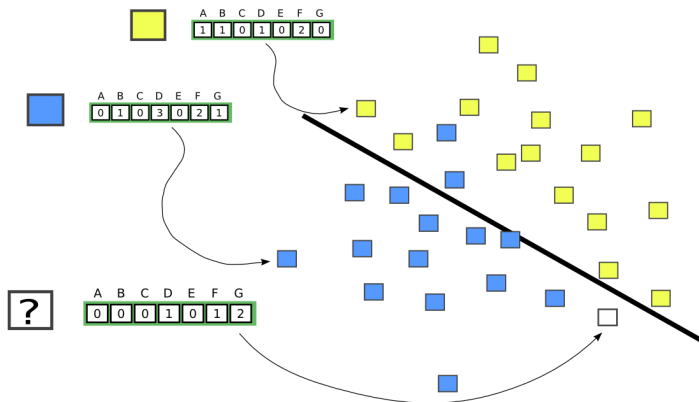
Classification des images

- Des règles de décision sont apprises ..



Classification des images

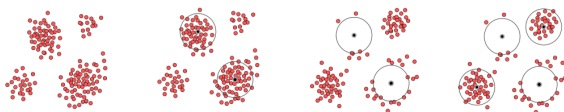
- .. et appliquées à la classification d'une nouvelle image



Contributions de la thèse

Une méthode efficace pour la quantification vectorielle de larges volumes de descripteurs visuels

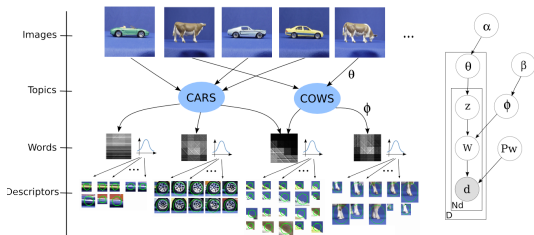
- Caractéristiques des données
 - ▶ Extraction d'un grand nombre de descripteurs par image
- Développement d'un algorithme de clustering adapté
 - ▶ algorithme à deux phases
 - ★ une phase d'échantillonnage biaisé
 - ★ une phase de clustering sur les échantillons choisis



Deuxième contribution

Vocabulaires discriminants pour la classification d'images

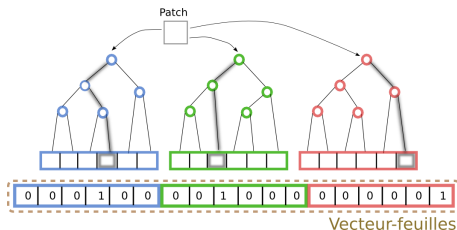
- Utilisation d'un modèle à variables latentes d'aspect
- Les informations de classe guident la quantification vectorielle
- Création de vocabulaires compacts et discriminants



Troisième contribution

Utilisation de forêts aléatoires pour la construction de cartes de saillance

- Classification des patches avec des arbres aléatoires
 - ▶ Construction d'une **carte de saillance**
 - ★ carte de probabilité sur la position des objets dans l'image
 - ▶ Guidage de l'échantillonnage des patches
- Vecteur de représentation par image, utilisé dans un classifieur SVM



Quatrième contribution

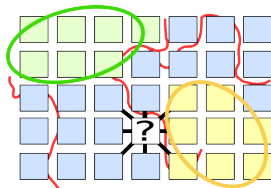
Un modèle LDA étendu pour la segmentation de catégories d'objets

- Recouvrement de l'image par un grand nombre de régions
- Estimation des classes présentes dans ces régions
- Échange de l'information et estimation conjointe
- Affectation d'un label à chaque patch
- Production de segmentations d'objets



Segmentation de catégories d'objets par combinaison d'un modèle par mots visuels et d'un champ de Markov

- Reconnaissance et localisation grossière des objets
 - ▶ un modèle par sac-de-mots
- Découpage fin de ces objets à partir de relations de voisinages locales et utilisation des contours de l'image
 - ▶ un champ de Markov



Détection d'objets appliquée à la recherche visuelle dans le cadre d'une application robotique

- Utilisation du modèle précédent pour la détection
 - ▶ instance d'objet particulière
 - ▶ informations stéréoscopiques
- Plateforme robotique humanoïde
 - ▶ application à la recherche visuelle d'objets

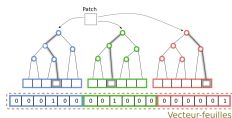


Résumé des contributions

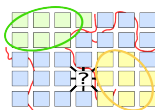
Une méthode efficace pour la quantification vectorielle de larges volumes de descripteurs visuels



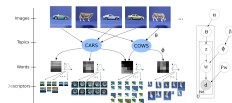
Utilisation de forêts aléatoires pour la construction de cartes de saillance



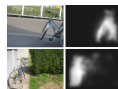
Segmentation de catégories d'objets par combinaison d'un modèle par mots visuels et d'un champ de Markov



Vocabulaires discriminants pour la classification d'images



Un modèle LDA étendu pour la segmentation de catégories d'objets

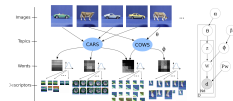


Détection d'objets appliquée à la recherche visuelle dans le cadre d'une application robotique



Résumé des contributions

Vocabulaires discriminants pour la classification d'images



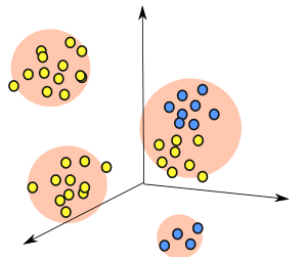
Segmentation de catégories d'objets par combinaison d'un modèle par mots visuels et d'un champ de Markov



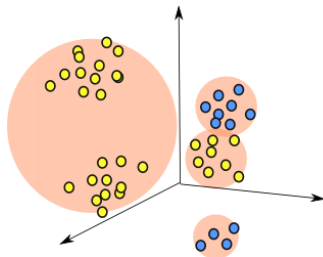
Vocabulaires discriminants pour la classification d'images

À propos du vocabulaire visuel

- Techniques de clustering classiques : l'information de classe n'est pas utilisée



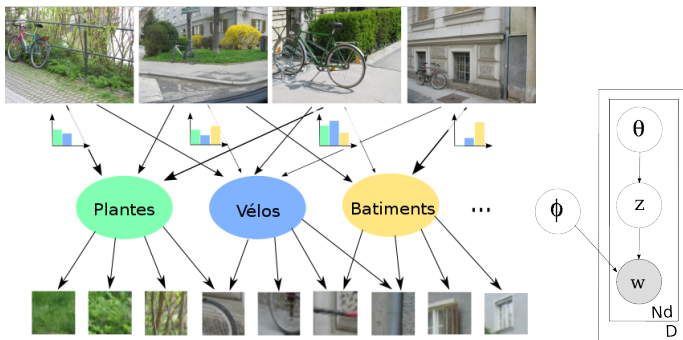
- Solution plus satisfaisante : précision des mots seulement lorsque cela est nécessaire



- Méthodes de création de vocabulaires discriminants :
 - ▶ Création d'un vocabulaire discriminant par fusion des mots d'un vocabulaire initial plus large :
Winn *et al*, ICCV 2006
 - ▶ Combinaison d'un vocabulaire universel et d'un vocabulaire spécifique aux classes :
Perronnin, PAMI 2008

Modèles à variables latentes d'aspect

- Les images sont représentées par des distributions sur des **variables latentes** nommées **topics**.



LDA, Blei *et al*, NIPS 2002

pLSA, Hofmann, ML 2001

Modèles à variables latentes d'aspect

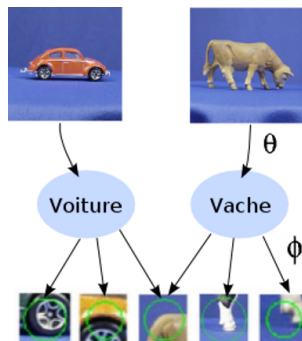
- Ces modèles à variables latentes ont été utilisés pour :
 - ▶ la classification supervisée d'objets
Fergus et al, ICCV05
 - ▶ la classification non-supervisée d'objets
Sivic et al, ICCV05
 - ▶ la classification de scènes
Quelhas et al, ICCV05
Fei-Fei et Perona, CVPR05

Modèles à variables latentes d'aspect

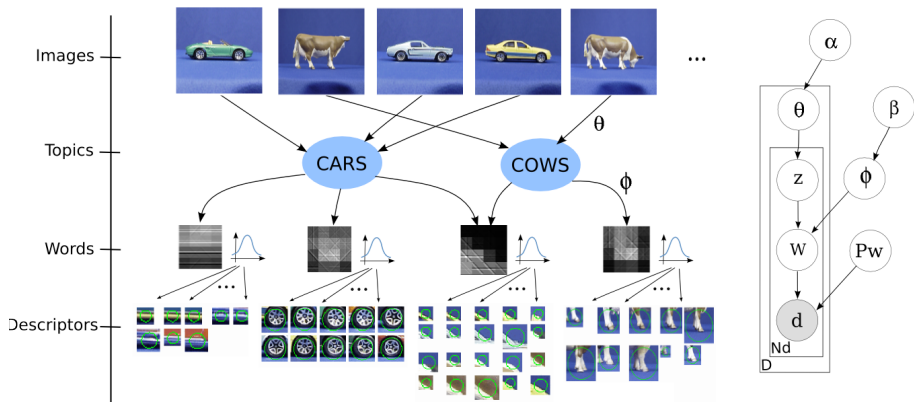
- Ces modèles à variables latentes ont été utilisés pour :
 - ▶ la classification supervisée d'objets
Fergus et al, ICCV05
 - ▶ la classification non-supervisée d'objets
Sivic et al, ICCV05
 - ▶ la classification de scènes
Quelhas et al, ICCV05
Fei-Fei et Perona, CVPR05
- Pour ces méthodes,
 - ▶ le vocabulaire est créé pendant une étape initiale
 - ▶ le modèle à variables latentes est construit ensuite

Modèles à variables latentes d'aspect

- Dans notre cas, les variables latentes sont directement liées aux catégories

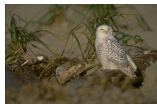
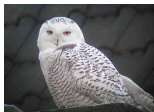
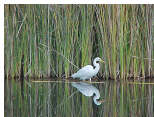


Modèle proposé



- Notre modèle possède une **couche supplémentaire**
 - ▶ responsable de la création des mots visuels
- Les mots sont générés par les topics, ils sont adaptés aux classes
- Les topics reposent sur des mots discriminants, ils sont plus précis

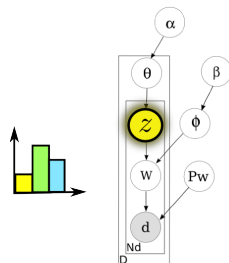
Base des oiseaux



Classification à partir des topics

- Autant de topics que de classes
- probabilité d'une classe = probabilité du topic correspondant

oiseaux	TOPIC-BAYES	
	LDA-VOC	STD-LDA
nb img labels	Moy	Moy
0	-	-
66	44.01%	-
198	55.97%	50.3%
300	60.68%	54.5%



Classification à partir des topics

- Autant de topics que de classes
- probabilité d'une classe = probabilité du topic correspondant

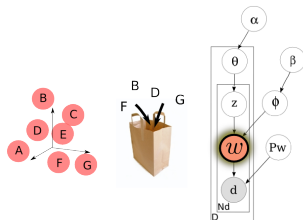
Observations

- sans supervision, les topics ne sont pas liés aux classes
- quelques images d'apprentissage corrigent ce problème
- l'estimation simultanée du vocabulaire et des topics est meilleure que celle du vocabulaire puis des topics

Classification à partir des mots visuels

- Création des mots visuels à l'aide du modèle
- Utilisation des mots dans le modèle sac-de-mots classique

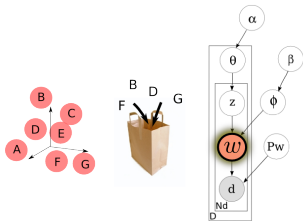
nb mots	oiseaux	
	LDA-VOC-BOF	KMEANS-BOF
200	74.6 %	65.33 %
500	85.1 %	76.58 %
1000	89.0 %	83.33 %
2000	90.9 %	86.17 %



Classification à partir des mots visuels

- Création des mots visuels à l'aide du modèle
- Utilisation des mots dans le modèle sac-de-mots classique

nb mots	oiseaux	
	LDA-VOC-BOF	KMEANS-BOF
200	74.6 %	65.33 %
500	85.1 %	76.58 %
1000	89.0 %	83.33 %
2000	90.9 %	86.17 %



Observations

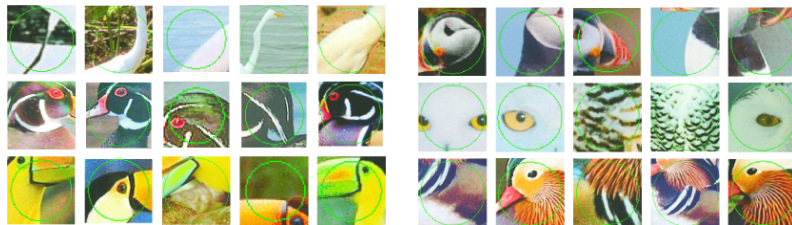
- la classification basée sur les mots est meilleure que celle sur les topics, pour un petit nombre de topics

Quelhas *et al*, ICCV05

- à taille de vocabulaire égale, les mots créés sont meilleurs

Mots visuels produits

- Pour chaque topic
 - ▶ 5 mots générés avec une forte probabilité



- le vocabulaire fait parti du modèle
 - ▶ conçu pour la méthode de classification (topics)
 - ▶ information de classes utilisée à l'échelle de l'image
 - ▶ utilisation simultanée de données labelisées et non labelisées
Lazebnik et al, PAMI 2008
 - ▶ complexité de l'estimation

- le vocabulaire fait parti du modèle
 - ▶ conçu pour la méthode de classification (topics)
 - ▶ information de classes utilisée à l'échelle de l'image
 - ▶ utilisation simultanée de données labelisées et non labelisées
Lazebnik et al, PAMI 2008
 - ▶ complexité de l'estimation
- modèle génératif moins efficace que le classifieur discriminatif
 - ▶ estimation des mots pour le classifieur discriminatif
Yang et al, CVPR 2008

- le vocabulaire fait parti du modèle
 - ▶ conçu pour la méthode de classification (topics)
 - ▶ information de classes utilisée à l'échelle de l'image
 - ▶ utilisation simultanée de données labélisées et non labélisées
Lazebnik et al, PAMI 2008
 - ▶ complexité de l'estimation
- modèle génératif moins efficace que le classifieur discriminatif
 - ▶ estimation des mots pour le classifieur discriminatif
Yang et al, CVPR 2008
- Variantes du modèle par sac-de-mots
 - ▶ noyaux de Fisher
Perronnin et al, CPVR 2007
 - ▶ arbres de décision aléatoires
Moosmann et al, PAMI 2008
 - ▶ "ferns"
Bosh et al, ICCV 2007

Segmentation de classes d'objets

Positionnement du problème

- Tâche de segmentation sans apprentissage



Positionnement du problème

- Tâche de segmentation sans apprentissage



- Tâche de segmentation avec modèle d'objet



Positionnement du problème

- Une méthode non supervisée seule ne peut résoudre le problème
- Nécessité d'introduire de la connaissance
 - ▶ Reconnaissance de l'objet
 - ▶ Localisation au niveau du pixel

Positionnement du problème

- Une méthode non supervisée seule ne peut résoudre le problème
- Nécessité d'introduire de la connaissance
 - ▶ Reconnaissance de l'objet
 - ▶ Localisation au niveau du pixel

Introduction d'un modèle des objets
dans le processus de segmentation

Positionnement du problème

- Une méthode non supervisée seule ne peut résoudre le problème
- Nécessité d'introduire de la connaissance
 - ▶ Reconnaissance de l'objet
 - ▶ Localisation au niveau du pixel

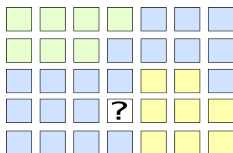
Introduction d'un modèle des objets dans le processus de segmentation

- Segmentation interactive
 - ▶ Définition interactive de la position des objets
GrabCut, Rother *et al*, Trans. Graph. 2004
- Segmentation avec modèle d'objets
 - ▶ Modèle de forme rigide
IMS, Leibe *et al*, BMVC 2003
LOCUS, Winn *et al*, ICCV 2005

Survol du modèle

Premier composant

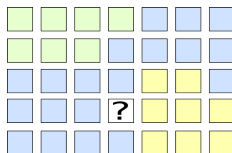
- L'image est vue comme une grille de patches
- Le **modèle génératif** affecte une classe à chaque patch



Survol du modèle

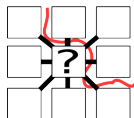
Premier composant

- L'image est vue comme une grille de patches
- Le **modèle génératif** affecte une classe à chaque patch



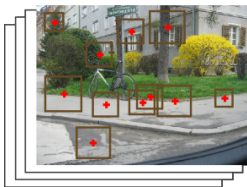
Deuxième composant

- L'ensemble de patches est vu comme un graphe
- Un **champ de Markov** est appliqué et prend en compte les contours de l'image

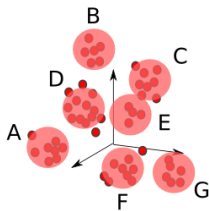


Phase d'apprentissage

- Modélisation des catégories à segmenter
 - ▶ liste d'images avec segmentations objet/fond connues
 - ▶ création du vocabulaire visuel



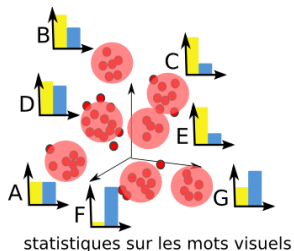
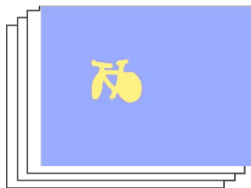
Sélection des patches



Création des mots visuels

Phase d'apprentissage

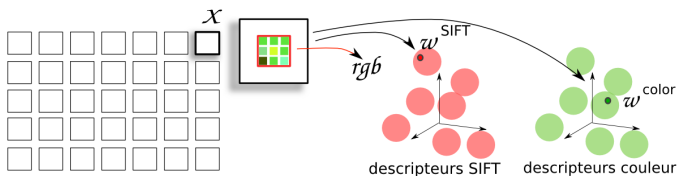
- Modélisation des catégories à segmenter
 - ▶ liste d'images avec segmentations objet/fond connues
 - ▶ création du vocabulaire visuel
 - ▶ statistiques sur les mots visuels



Description d'une image à segmenter

- Représentation d'un patch

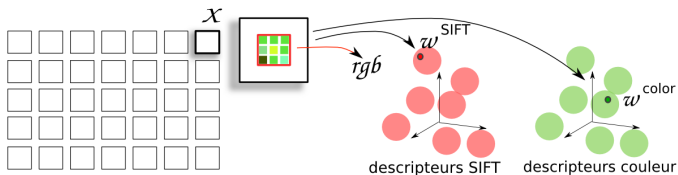
- ▶ par un mot visuel SIFT : w_i^{SIFT}
- ▶ par un mot visuel couleur : w_i^{color}
- ▶ par une position dans l'image : X_i
- ▶ par un vecteur RGB (moyenne des pixels du centre du patch) : rgb_i



Description d'une image à segmenter

- Représentation d'un patch

- ▶ par un mot visuel SIFT : w_i^{SIFT}
- ▶ par un mot visuel couleur : w_i^{color}
- ▶ par une position dans l'image : X_i
- ▶ par un vecteur RGB (moyenne des pixels du centre du patch) : rgb_i



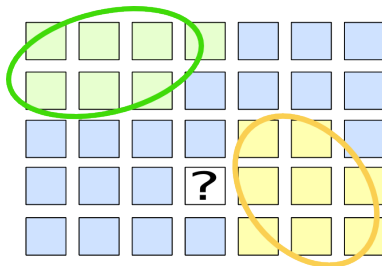
- Extraction d'une image gradient



Premier composant

Les patches sont regroupés en régions de forme elliptique ou **blobs** de classe d'objets, ainsi qu'en une région de fond.

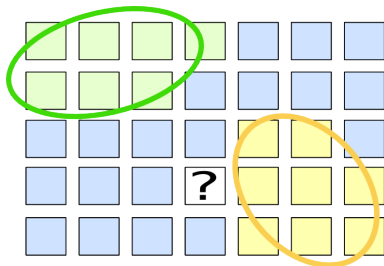
- ces blobs sont décrits par des variables
 - ▶ de position et de forme
 - ▶ de catégorie
 - ▶ de couleur



Premier composant

Les patches sont regroupés en régions de forme elliptique ou **blobs** de classe d'objets, ainsi qu'en une région de fond.

- ces blobs sont décrits par des variables
 - ▶ de position et de forme
 - ▶ de catégorie
 - ▶ de couleur
- les patches sont produits par un **processus génératif** à partir des variables définissant les blobs



Modèle génératif

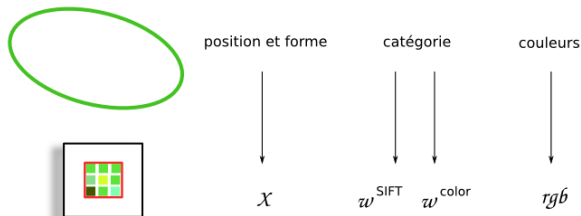
Modèle génératif d'une image à segmenter

- pour chaque patch de l'image :
 - ▶ sélection d'un blob
 - ★ choix d'un **blob existant** avec une probabilité proportionnelle à sa population
 - ★ création d'un **nouveau blob** avec une proba fixe

Modèle génératif

Modèle génératif d'une image à segmenter

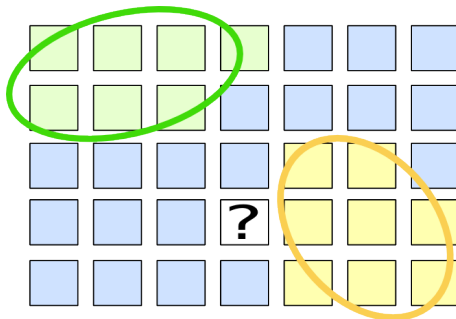
- pour chaque patch de l'image :
 - ▶ sélection d'un blob
 - ★ choix d'un blob existant avec une probabilité proportionnelle à sa population
 - ★ création d'un nouveau blob avec une proba fixe
 - ▶ à partir de ce blob, échantillonnage des caractéristiques du patch sachant le blob



Résumé du modèle génératif

Pour segmenter une image :

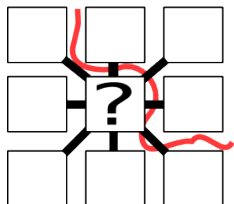
- trouver le modèle de blobs qui explique au mieux les observations
- trouver les affectations des patches aux blobs



Deuxième composant

Hypothèse de cohérence spatiale

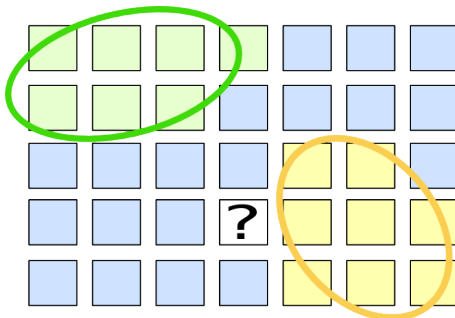
- **régularisation** du modèle précédent
 - ▶ les patches voisins appartiennent de préférence à la même région
 - ▶ la séparation entre les objets se fait le plus souvent le long des contours



Modèle complet

Combinaison des composants

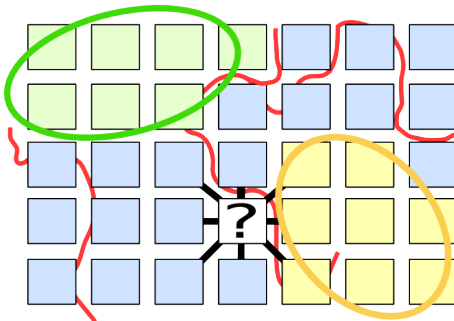
- modèle génératif : génération des patchs par les blobs



Modèle complet

Combinaison des composants

- modèle génératif : génération des patchs par les blobs
- champs de Markov sur les labels d'affectation aux blobs



Combinaison du modèle génératif avec le MRF de cohérence des labels par une formulation énergétique :

$$p(b, \mathcal{P}) \propto \exp(-E)$$

$$E = U + \gamma \sum_{i,j \in \mathcal{C}} V_{i,j}$$

- U est un potentiel qui encode le modèle génératif
- $V_{i,j}$ représente les contraintes entre les patches \mathcal{P}_i et \mathcal{P}_j voisins
- γ est un paramètre qui contrôle l'importance des termes

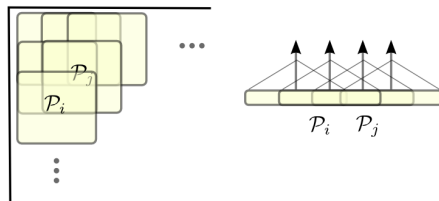
GrabCut, Rother *et al*, Trans. Graph. 2004

- Estimation d'un modèle pour **chaque image à segmenter** :
 - ▶ estimation du nombre de blobs et de leurs paramètres
 - ▶ estimation des affectations des patches aux blobs
- Estimation par échantillonnage de Gibbs
 - ▶ génération de valeurs successives pour les paramètres à estimer

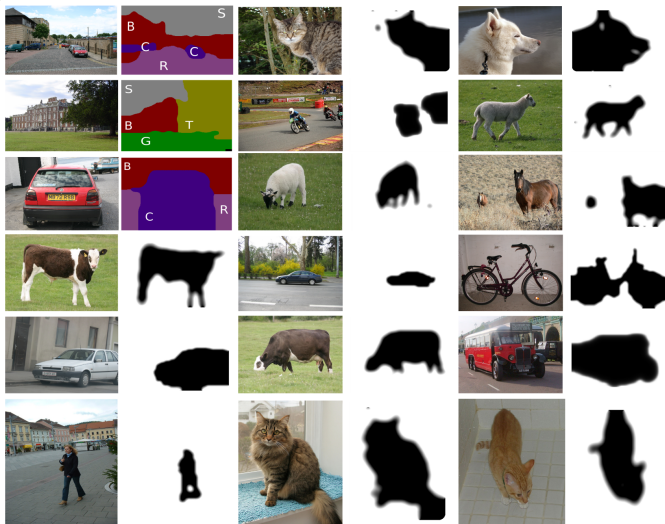
- Estimation d'un modèle pour chaque image à segmenter :
 - ▶ estimation du nombre de blobs et de leurs paramètres
 - ▶ estimation des affectations des patches aux blobs
- Estimation par **échantillonnage de Gibbs**
 - ▶ génération de valeurs successives pour les paramètres à estimer

Des patches aux pixels

- Le modèle prédit des labels à l'échelle des patches
 - ▶ prédiction des labels au niveau du pixel
 - ▶ en considérant tous les patches contenant ce pixel



Résultats obtenus



- Étude du comportement de la méthode :
 - ▶ Rôle des différents composants du modèle

- Étude du comportement de la méthode :
 - ▶ Rôle des différents composants du modèle
 - ▶ Dépendance aux données d'apprentissage

- Étude du comportement de la méthode :
 - ▶ Rôle des différents composants du modèle
 - ▶ Dépendance aux données d'apprentissage
 - ▶ Influence de l'initialisation

- Étude du comportement de la méthode :
 - ▶ Rôle des différents composants du modèle
 - ▶ Dépendance aux données d'apprentissage
 - ▶ Influence de l'initialisation

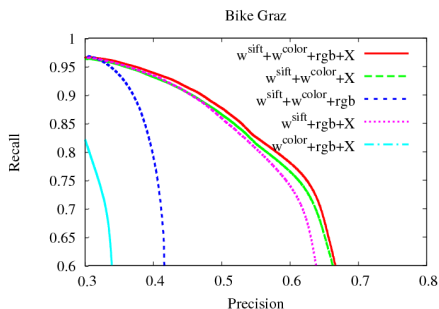
- La segmentation évaluée comme une classification au niveau du pixel

- Étude du comportement de la méthode :
 - ▶ Rôle des **différents composants** du modèle

- Expériences sur la base TU-Graz02



Influence des composants

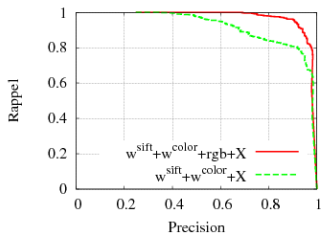


Observations :

- Les deux vocabulaires participent à la reconnaissance
- SIFT est plus critique
- La régularisation est indispensable
- influence plus visuelle que quantitative de la composante RGB

Influence des composants

Rôle du composant couleur *rgb*



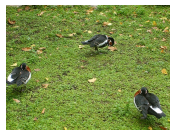
- Base d'images, découpée en
 - ▶ ~ 5000 images d'apprentissage
 - ▶ ~ 5000 images de tests
- Contient 20 classes d'objets
- Compétition de segmentation
 - ▶ classification multiclasse à l'échelle du pixel



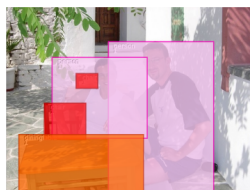
- Étude du comportement de la méthode :
 - ▶ Rôle des différents composants du modèle
 - ▶ Dépendance aux données d'apprentissage
 - ▶ Influence de l'initialisation

Images d'apprentissage disponibles

- ~ 400 images segmentées



- ~ 5000 images avec boîtes englobantes



Dépendance aux données d'apprentissage

- comparaison de différents ensembles d'apprentissage
 - ▶ **FT** ("Full Training") : \sim 5000 images
 - ▶ **ST** ("Small Training") : \sim 400 images

	moyenne
FT	37.16
ST	31.41
Brooks	8.5
TKK	30.4

Pascal VOC 2007, Everingham *et al*

Dépendance aux données d'apprentissage

	moyenne
FT	37.16
ST	31.41
Brooks	8.5
TKK	30.4

- comparaison de différents ensembles d'apprentissage
 - ▶ FT ("Full Training") : \sim 5000 images
 - ▶ ST ("Small Training") : \sim 400 images
- Observations
 - ▶ Amélioration de la moyenne des classifications
 - ★ également pour la majorité des classes
 - ▶ Meilleurs résultats que les soumissions à la compétition

Pascal VOC 2007, Everingham *et al*

Génération de nouvelles images d'apprentissage

- Augmentation de l'ensemble d'images segmentées



- ▶ les boites englobantes donnent la position et la classe des blobs
- ▶ application de la méthode sur ces images
- ▶ génération de nouvelles segmentations

Discussion (1)

- Structure du modèle

- ▶ évidences à l'échelle d'un nœud
- ▶ contraintes de régularisation très locales
- ▶ contraintes à plus grande échelle
 - ★ globales
 - Verbeek et Triggs, CVPR 07
 - Csurka et Perronnin, BMVC 08
 - ★ régions provenant d'une pré-segmentation
 - Russell *et al*, CVPR 06
 - Yang *et al*, CVPR 07
 - ★ taille de l'objet détecté
- ▶ utilisation d'évidences à plus grande échelle
 - ★ problème des classes visuellement proches

Discussion (2)

- Modèle entièrement génératif

- ▶ ajout d'une classe supplémentaire simple
- ▶ utilisation de classifieur discriminatif

Textonboost, Shotton *et al*, ECCV 06

Csurka et Perronnin, BMVC 08

Discussion (2)

- Modèle entièrement génératif

- ▶ ajout d'une classe supplémentaire simple
- ▶ utilisation de classifieur discriminatif

Textonboost, Shotton *et al*, ECCV 06

Csurka et Perronnin, BMVC 08

- Robustesse du modèle

- ▶ changement d'illumination

SIFT, Lowe, IJCV 04

Van de Weijer et Schmid, ECCV 06

- ▶ changement d'échelle

Discussion (2)

- Modèle entièrement génératif

- ▶ ajout d'une classe supplémentaire simple
- ▶ utilisation de classifieur discriminatif

Textonboost, Shotton *et al*, ECCV 06

Csurka et Perronnin, BMVC 08

- Robustesse du modèle

- ▶ changement d'illumination

SIFT, Lowe, IJCV 04

Van de Weijer et Schmid, ECCV 06

- ▶ changement d'échelle
- ▶ modèle d'apparence souple

- ★ robustesse intrinsèque aux occultations
- ★ traitement des classes à la géométrie faible
- ★ multi-vue
- ★ plus de faux positifs
- ★ problème avec les occultations entre objets de la même classe

Winn et Shotton, CVPR 06

Conclusions et perspectives

- Contributions de la thèse
 - ▶ Création de vocabulaires visuels pour la représentation des images
 - ★ Extraction dense de descripteurs
 - ★ Utilisation des informations de classe

- Contributions de la thèse
 - ▶ Création de vocabulaires visuels pour la représentation des images
 - ★ Extraction dense de descripteurs
 - ★ Utilisation des informations de classe
 - ▶ Utilisation des vocabulaires visuels pour la localisation des objets
 - ★ Localisation pour guider l'échantillonnage
 - ★ Modèles génératifs de mots visuels pour la segmentation

- Pistes à poursuivre
 - ▶ utilisation de représentations adaptées à la tâche considérée
 - ★ représentation fortement liée au classifieur

- Pistes à poursuivre
 - ▶ utilisation de représentations adaptées à la tâche considérée
 - ★ représentation fortement liée au classifieur
 - ▶ ajout de connaissance sous forme de modèle des catégories d'objet
 - ★ segmentation sémantique des images

- Pistes à poursuivre
 - ▶ utilisation de représentations adaptées à la tâche considérée
 - ★ représentation fortement liée au classifieur
 - ▶ ajout de connaissance sous forme de modèle des catégories d'objet
 - ★ segmentation sémantique des images
 - ▶ modélisation de chacune des différentes instances d'objet
 - ★ amélioration de la segmentation
 - ★ utilisation de descripteurs à l'échelle de l'objet

- Pistes à poursuivre
 - ▶ utilisation de représentations adaptées à la tâche considérée
 - ★ représentation fortement liée au classifieur
 - ▶ ajout de connaissance sous forme de modèle des catégories d'objet
 - ★ segmentation sémantique des images
 - ▶ modélisation de chacune des différentes instances d'objet
 - ★ amélioration de la segmentation
 - ★ utilisation de descripteurs à l'échelle de l'objet
 - ▶ interaction entre les instances, modélisation du contexte
 - ★ description sémantique de l'image

- Question du passage à l'échelle
 - ▶ Vocabulaires visuels dépendants des images d'apprentissage
 - ▶ Classifieurs à réapprendre pour un ensemble de classes
 - ▶ Problème de la structure en classes
 - ★ inclusion, chevauchement entre les classes
Sivic et al, CVPR 08
 - ★ annotations coûteuses
Malisiewicz et Efros, CVPR 08

Soutenance de Thèse

Création et utilisation de vocabulaires visuels pour la catégorisation d'images et la segmentation de classes d'objets

Diane Larlus

Équipe LEAR, INRIA Grenoble - INPG

Vendredi 28 novembre 2008

Composition du Jury :

- Pr. Roger Mohr, président
- Pr. Jean-Pierre Cocquerez, rapporteur
- Pr. Patrick Perez, rapporteur
- Pr. Matthieu Cord, examinateur
- Dr. Stéphane Herbin, examinateur
- Pr. Frédéric Jurie, directeur de thèse

