

Modèles graphiques probabilistes pour la reconnaissance de formes

Sabine Barrat

Équipe QGAR - LORIA

Composition du jury :

- Laurence LIKFORMAN-SULEM (Rapporteur)
- Philippe MULHEM (Rapporteur)
- Adam CICHON (Examineur)
- Pierre LOONIS (Examineur)
- Salvatore TABBONE (Directeur de thèse)
- Muriel VISANI (Examineur)
- Patrick NOURRISSIER (Invité)

Motivations

- essor d'Internet
- démocratisation du multimédia
- \Rightarrow surabondance d'information multimédia
- \Rightarrow difficultés à retrouver la bonne information
- **solution : indexer les images**

Indexation textuelle [*Liu 09c, Grangier 08*]

- manuelle : par des indexeurs humains
 - avantage : performante
 - inconvénients : lourde, subjective
- automatique : à partir de texte associé à l'image
 - avantages : peu coûteuse, rapide
 - inconvénient : moins performante que l'indexation manuelle
- fossé sensoriel : difficulté de rendre compte d'une image par des mots

Indexation visuelle [*Smeulders 00*]

- à base de caractéristiques de couleur, forme ou texture
 - avantages : moins subjective que le texte, performante sur certaines bases spécialisées
 - inconvénient : bases généralistes
- fossé sémantique : il n'y a pas *a priori* de relations entre caractéristiques visuelles et textuelles

Les problèmes de l'indexation visuelle et textuelle

		
balle ball sport tennis	ballon ballon ovale sport rugby	soleil planète chaleur été

Objectifs scientifiques

- développer des méthodes d'indexation efficaces
- élaborer des méthodes d'accès aux images :
 - classification
 - recherche

Solution envisagée

- combiner :
 - plusieurs caractéristiques visuelles
 - information visuelle et sémantique
- annoter automatiquement des bases d'images partiellement annotées

Plan

- 1 Modèles graphiques probabilistes
- 2 Contributions en reconnaissance et annotation d'images naturelles
- 3 Contributions en reconnaissance de symboles
- 4 Conclusion et perspectives

Plan

- 1 Modèles graphiques probabilistes
- 2 Contributions en reconnaissance et annotation d'images naturelles
- 3 Contributions en reconnaissance de symboles
- 4 Conclusion et perspectives

Comparaison des approches de l'état de l'art

méthode	critère	données	discret et continu	grande dimension	temps de calcul	inter- pré- tation
	manquantes					
Modèles probabilistes [Kotsiantis 07]	+	+	-	-	+	
SVM [Hsu 02]	-	+	+	-	-	
k -ppv [Keller 85]	-	-	+	-	+	
Réseaux de neurones [Bishop 95]	+	+	+	-	-	
Arbres et forêts aléatoires [Breiman 84]	+	+	-	+	+	

Propriétés utiles des représentations graphiques

- moyen simple de représentation de la structure des modèles probabilistes
- aperçu des propriétés du modèle
- calculs complexes exprimés en termes de manipulations graphiques

Modèles graphiques probabilistes

- famille de distributions de probabilités représentée par un graphe où :

nœuds variables aléatoires

arêtes relations de dépendance entre les variables

- deux familles de modèles graphiques :

graphes orientés réseaux Bayésiens (graphes sans cycle)

graphes non orientés champs de Markov

Principe des réseaux Bayésiens

- v_1, v_2, \dots, v_n : n variables aléatoires
- $Pa(v_i)$: ensemble des parents du nœud associé à la variable aléatoire v_i
- $P(v_1, v_2, \dots, v_n) = \prod_{i=1}^n P(v_i | Pa(v_i))$

Réseau Bayésien comme classificateur

- but : affecter $f_i = \{f_{i_1}, f_{i_2}, \dots, f_{i_n}\}$ à une classe $c_j \in \{c_1, c_2, \dots, c_k\}$
- règle : choisir la classe c_j qui maximise :

$$\begin{aligned} P(c_j | f_1, f_2, \dots, f_n) &= \frac{P(f_1, f_2, \dots, f_n, c_j)}{P(f_1, f_2, \dots, f_n)} \\ &= \frac{P(f_1, f_2, \dots, f_n | c_j) P(c_j)}{P(f_1, f_2, \dots, f_n)} \end{aligned}$$

où

$$P(f_1, f_2, \dots, f_n) = \sum_{j=1}^k P(f_1, f_2, \dots, f_n | c_j) P(c_j)$$

Le Naïve Bayes

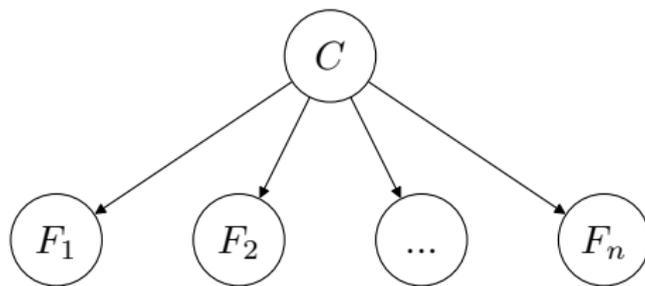


FIG.: Classificateur Bayésien naïf

Plan

- 1 Modèles graphiques probabilistes
- 2 Contributions en reconnaissance et annotation d'images naturelles
- 3 Contributions en reconnaissance de symboles
- 4 Conclusion et perspectives

Proposition

- combinaison de caractéristiques visuelles et textuelles dans un même réseau
- prise en compte des relations sémantiques entre mots-clés
- distributions de probabilités :

caractéristiques visuelles mélange de Gaussiennes multivariées

caractéristiques sémantiques loi de Bernoulli

- \Rightarrow modèle de mélange de lois de Bernoulli et de mélanges de Gaussiennes

Caractéristiques textuelles

- chaque variable KW
 - correspond à un terme du vocabulaire associé à la base
 - N : taille du vocabulaire
 - est une variable aléatoire de Bernoulli
 - peut prendre deux valeurs : 1 quand le terme correspondant peut faire partie de l'annotation de l'image, 0 sinon
- création de dépendances conditionnelles entre deux termes appartenant au même groupe sémantique

Dépendances conditionnelles entre termes

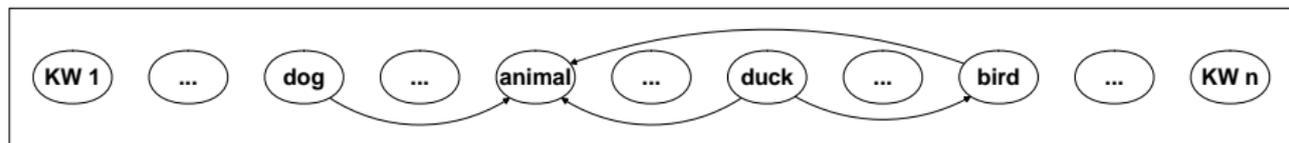


FIG.: Dépendances entre les termes du vocabulaire

Rappel sur les GMM

- F : échantillon d'apprentissage composé de m individus
 $f_{1_i}, \dots, f_{m_i}, \forall i \in \{1, \dots, n\}$
- n : dimension des signatures obtenues par concaténation des vecteurs caractéristiques issus du calcul des descripteurs
- division des m individus en k classes c_1, \dots, c_k
- G_1, \dots, G_g : g groupes dont chacun a une densité Gaussienne de paramètres
 - moyenne $\mu_l, \forall l \in \{1, \dots, g\}$
 - matrice de covariance \sum_l
- π_1, \dots, π_g : proportions des différents groupes

Modèle graphique de modèle de mélange de Gaussiennes

- GMM : modèle graphique à 3 nœuds :

Classe variable classe : $c_i, \forall i \in \{1, \dots, k\}$

Composante poids associés aux composantes Gaussiennes du mélange : $\pi_i, \forall i \in \{1, \dots, g\}$

Gaussienne chaque composante Gaussienne :
 $G_i, \forall i \in \{1, \dots, g\}$

Représentation graphique d'un GMM

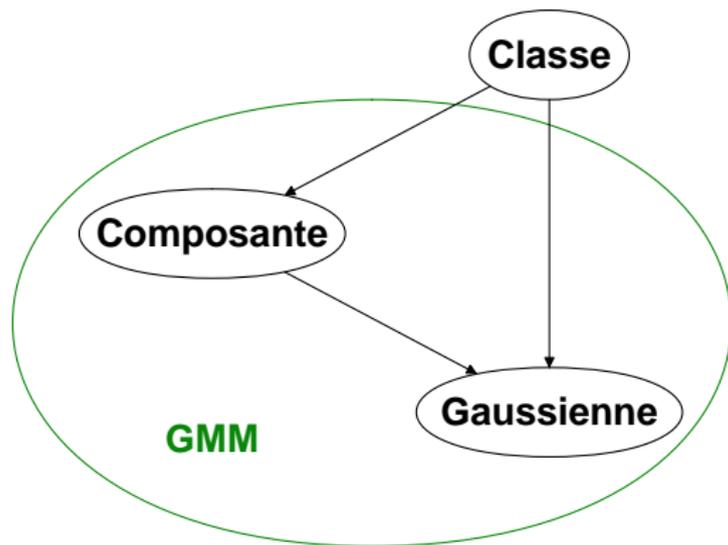


FIG.: GMMs représentés par un modèle graphique probabiliste

Modèle de mélange GM-B

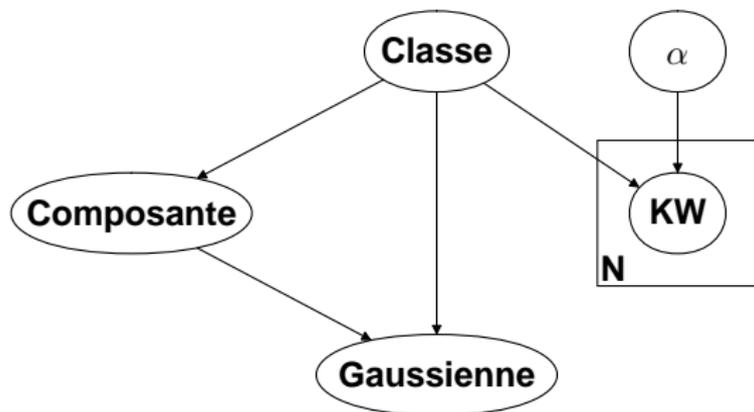


FIG.: Modèle de mélange GM-B

Zoom sur les nœuds *KW*

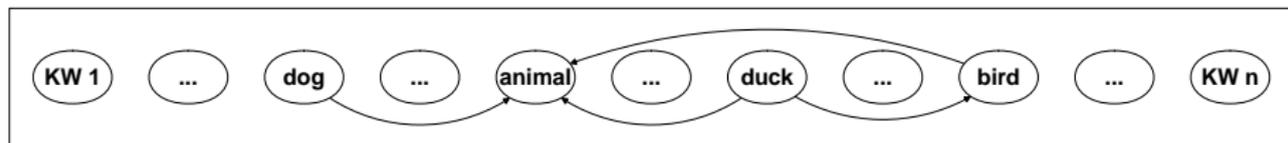


FIG.: Dépendances entre les termes du vocabulaire

Least Absolute Shrinkage and Selection Operator (LASSO) [*Tibshirani 96*]

- sélection de variables
- réduit
 - la dimension
 - la complexité du modèle et les temps de calculs
- améliore le taux de reconnaissance

Base de données et caractéristiques visuelles

- plus de 30000 issues de la librairie COREL [*Carneiro 07*]
 - 72% annotée par 4 mots-clés
 - 23% par 3 mots-clés
 - 4% par 2 mots-clés
 - 0,5% par 1 mots-clés
 - Vocabulaire : 1036 mots-clés
- 306 classes
- Caractéristiques visuelles
 - Histogramme de couleur [*Swain 91*]
 - \mathcal{R} -signature 1D [*Tabbone 02*]

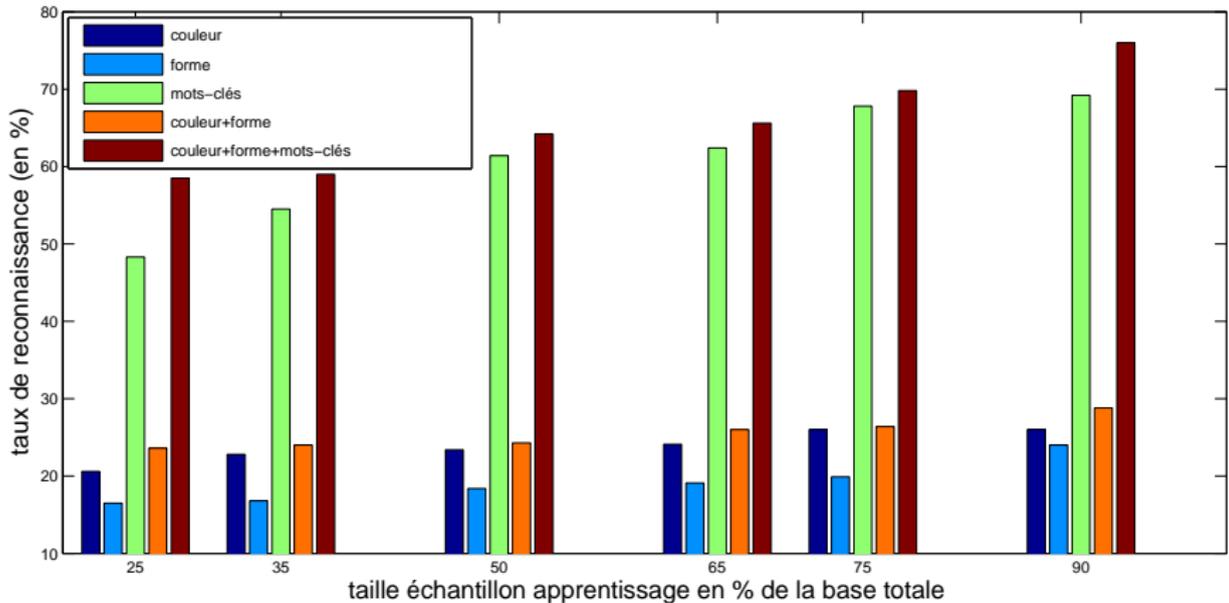
Exemples d'images

	water duck reflection flock		bird duck water close-up
Classe : Hong Kong		Classe : african birds	
	bird duck mallard baby		duck food cuisine meal
Classe : waterfowl		Classe : cuisine	

TAB.: Exemples d'images, avec leurs classes et éventuels mots-clés

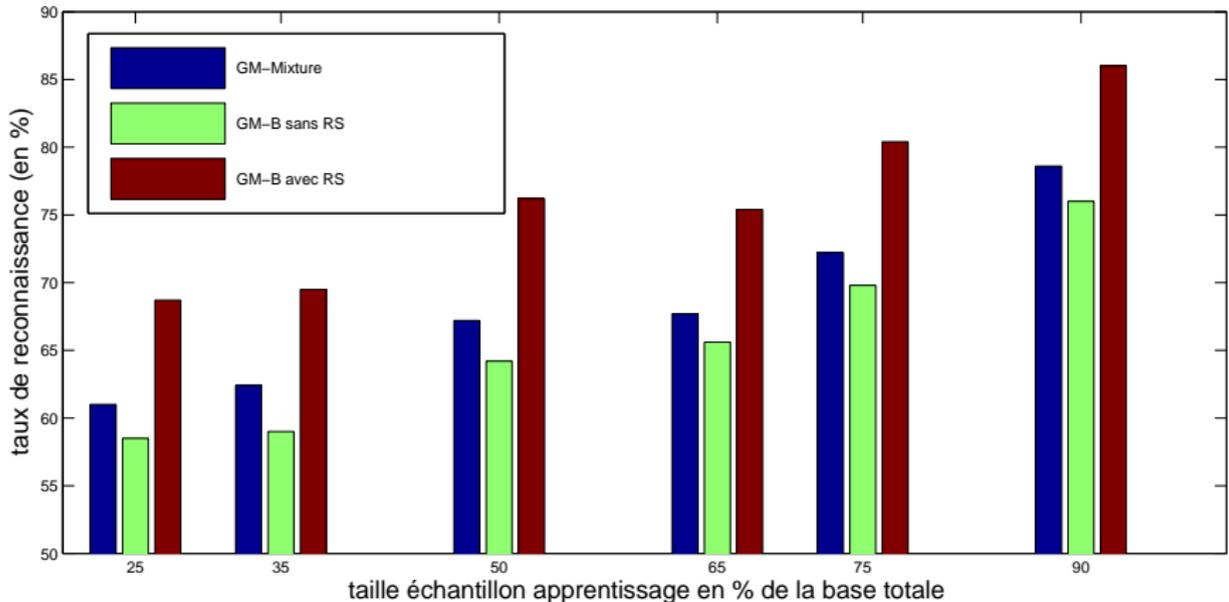
Impact de la combinaison de caractéristiques

Modèle GM-B sans RS

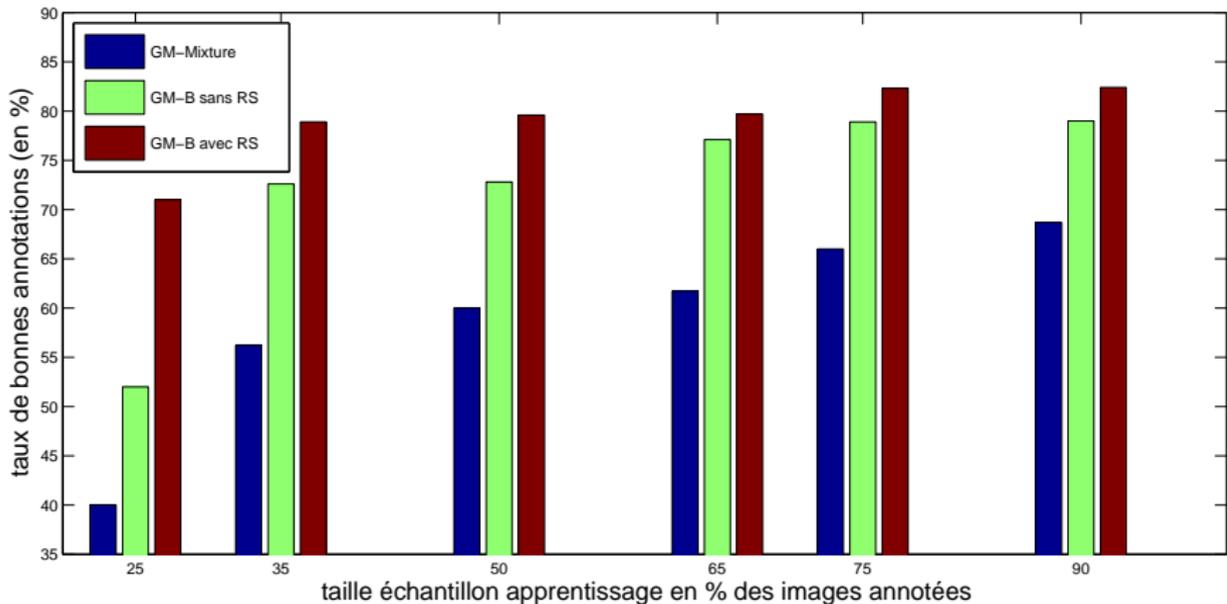


Classification : GM-B vs. GM-Mixture [Blei 03]

Caractéristiques : couleur + forme + mots-clés



Annotation : GM-B vs. GM-Mixture



Exemple d'images et de mots-clés associés

image	mots-clés initiaux	après annotation sans RS	après annotation avec RS
	bear black river	bear black river	bear black river water
	penguin water	penguin water iceberg	penguin water iceberg bird

TAB.: Exemples d'images et de mots-clés associés, avant et après extension d'annotations avec ou sans relations sémantiques, en utilisant le modèle GM-B

Plan

- 1 Modèles graphiques probabilistes
- 2 Contributions en reconnaissance et annotation d'images naturelles
- 3 Contributions en reconnaissance de symboles
- 4 Conclusion et perspectives

Proposition : combinaison de descripteurs et mesures de forme dans un même réseau

- descripteurs de forme : vecteurs caractéristiques de plus d'une composante
- mesures de forme : caractéristiques composées d'une seule valeur
 - discrétisées avec un seuil de discrétisation fixé à 0.5
 - deux valeurs possibles : 0 si la mesure de forme correspondante fournit une valeur inférieure à 0,5, 1 sinon

Distributions de probabilités

descripteurs de forme mélange de Gaussiennes multivariées

mesures de forme loi de Bernoulli

- modèle proposé : modèle de mélange de lois de Bernoulli et de mélanges de Gaussiennes

Modèle de mélange GM-B adapté

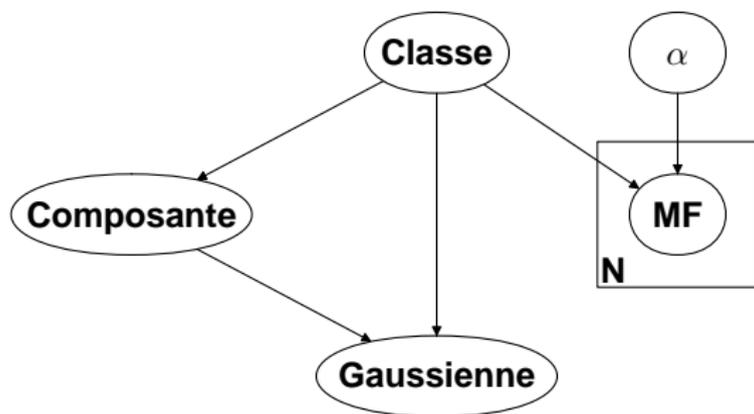


FIG.: Modèle de mélange de lois de Bernoulli et de mélanges de Gaussiennes, pour la combinaison de descripteurs

Caractéristiques de forme

- Descripteurs (vecteurs de valeurs continues)
 - Generic Fourier Descriptor (GFD) [*Zhang 02a*]
 - descripteur Zernike [*Kim 00*]
 - \mathcal{R} -signature 1D [*Tabbone 02*]
 - descripteur HRT [*Tabbone 08*]
- Mesures (valeurs discrètes)
 - compacité
 - rectangularité [*Rosin 99*]
 - degré d'ellipticité [*Teague 79*]

Modèles de la base GREC [Valveny04]

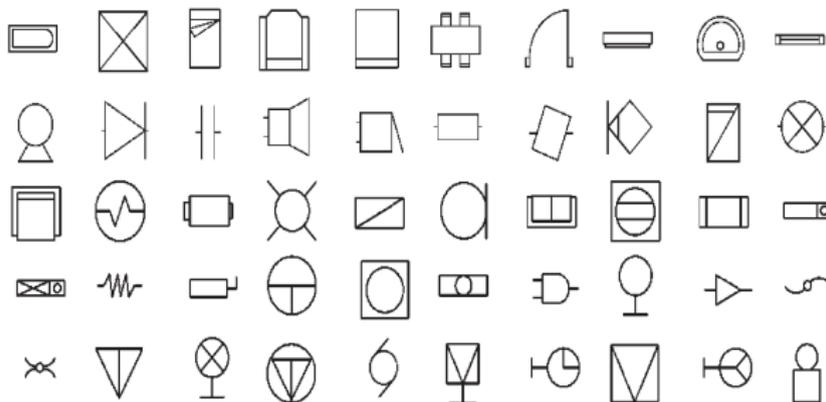


FIG.: 50 modèles de la base GREC.

Base A générée à partir de la base GREC

- 3600 images : 72 images bruitées et/ou ayant subi des rotations ou des zooms, par modèle

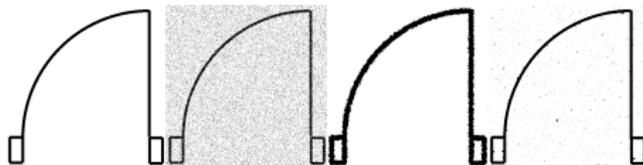


FIG.: Exemple d'un modèle sur lequel on a appliqué différents bruits

Impact de la combinaison de caractéristiques

G	Z	R	G+Z	G+R	Z+R	G+Z+R
99	100	46.1	100	99.3	100	100

TAB.: Taux de reconnaissance moyens (en %) du GM-B

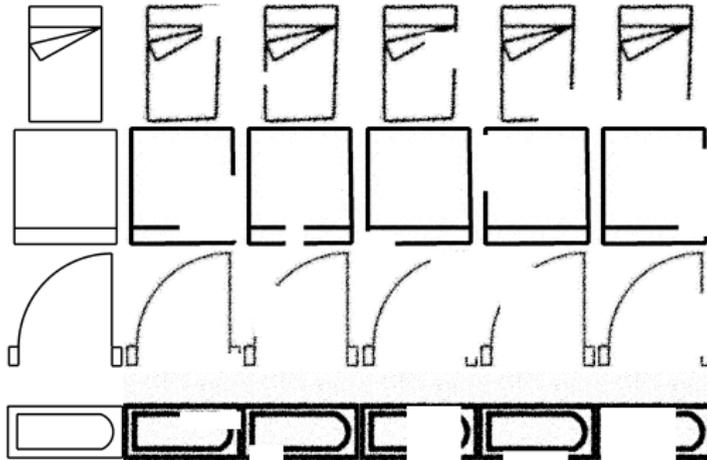
G : GFD

Z : descripteur Zernike

R : \mathcal{R} -signature 1D

Base B générée à partir de la base GREC

- 5400 images : 108 images bruitées et/ou occultées par modèle



Impact de la combinaison de caractéristiques

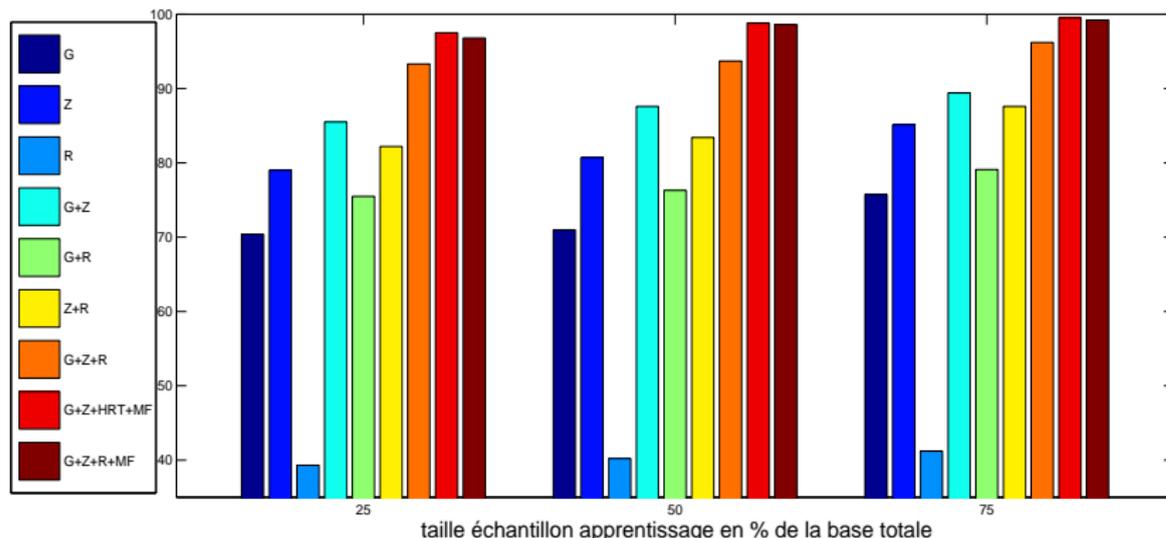
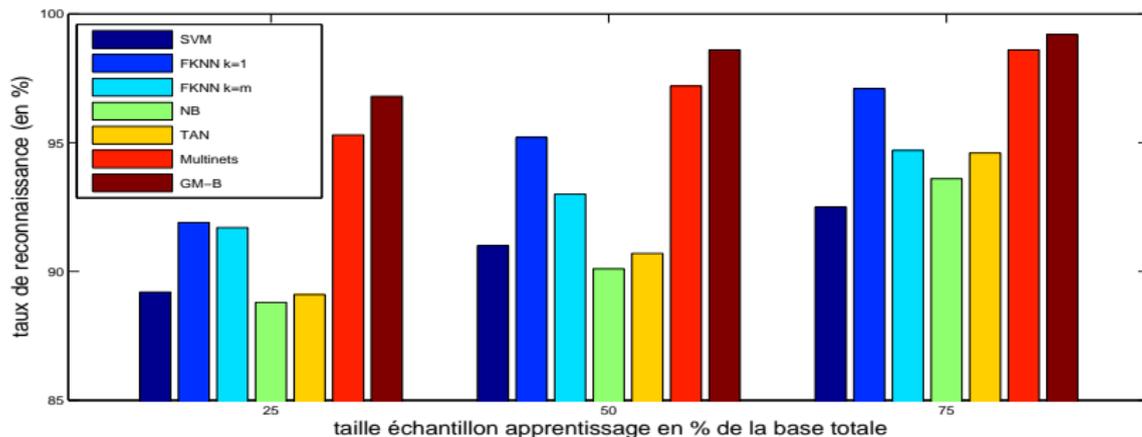


FIG.: Taux de reconnaissance moyens (en %) du modèle GM-B en fonction de la taille de l'échantillon d'apprentissage

GM-B vs. d'autres classificateurs

Caractéristiques : G+Z+R+MF



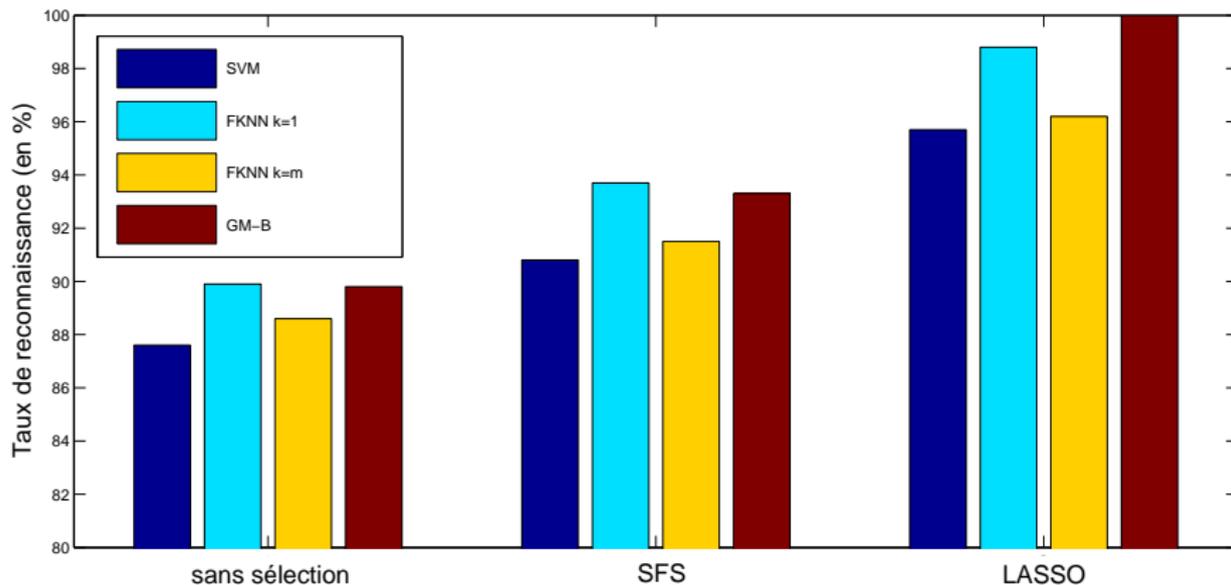
- SVM [*Chang 01*]
- FKNN [*Keller 85*]
- NB, TAN, Multinets [*Friedman 97b*]

Comparaison des temps de calcul

App	SVM		1-NN	m -NN	GM-B avec LASSO		GM-B sans SV	
	app	test			app	test	app	test
25%	4	5	40	41	58	78	2726	25608
50%	10	6	56	58	117	52	5696	17291
75%	19	4	42	45	168	24	8110	8453

TAB.: Temps CPU (en secondes), du SVM, du FKNN et du GM-B. Les temps sont donnés pour la classification de toutes les images test

Impact sur le taux de reconnaissance



Plan

- 1 Modèles graphiques probabilistes
- 2 Contributions en reconnaissance et annotation d'images naturelles
- 3 Contributions en reconnaissance de symboles
- 4 Conclusion et perspectives

Contributions

- proposition de modèles permettant de :
 - représenter
 - classer
 - annoter
 - rechercher des images

- proposition d'un nouveau descripteur de forme

Avantages de l'approche proposée

- combinaison de caractéristiques de différents types
 - amélioration du taux de reconnaissance
- souple pour l'utilisateur
 - requêtes à base d'image et/ou de mots-clés
 - annotation manuelle limitée
- application à des bases généralistes et spécialisées (symboles)
- compétitive avec des méthodes de l'état de l'art

À court terme

- ajouter d'autres descripteurs
- améliorer l'intégration du descripteur *HRT* dans nos modèles

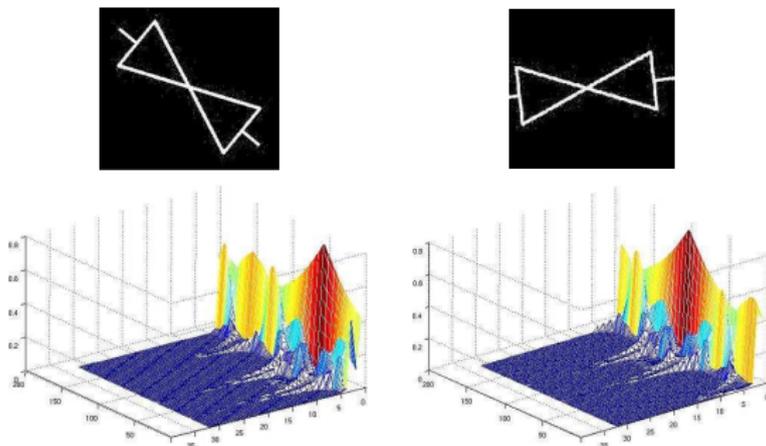


FIG.: Le descripteur *HRT* face à la rotation

À long terme : post-doctorat Osaka Prefecture University - Japon

- utiliser le système de recherche d'images existant du laboratoire japonais en amont de notre système d'annotation
 - intégration des caractéristiques textuelles dans le système de recherche
 - en amont : rechercher les images les plus proches d'une requête
 - utiliser l'ensemble d'images obtenu comme base d'apprentissage pour notre système d'annotation