

École de technologie supérieure
Programme de maîtrise en :
 génie de la production automatisée
 génie électrique
 technologie des systèmes

Session : Automne 2009
Professeur : Mohamed Cheriet

Préalable : Aucun
Crédits : 3

SYS-821 RECONNAISSANCE DE FORMES ET INSPECTION

Liste de Projets

1. Affectation de projets

Chaque étudiant doit réaliser un projet de cours à travers lequel il montre sa compréhension de la matière du cours, en approfondissant un ou plusieurs concepts. En somme, il s'agit de réaliser un projet dans un cadre théorique/pratique ou théorique. Dans le premier cas, il s'agit d'implémenter une méthodologie bien identifiée. Dans le second cas, il s'agit de faire une revue de la littérature autour d'une problématique identifiée, et d'en faire une synthèse.

2. Choix de projets

i) Un projet en lien avec le sujet de recherche de l'étudiant

Un étudiant qui poursuit une thèse, au niveau de la maîtrise ou du doctorat, dans le domaine du cours, pourrait poursuivre son travail de recherche, en proposant la partie qu'il souhaiterait développer dans le cadre de ce cours. Il suffit de montrer la continuité de son projet avec sa recherche, et surtout de bien montrer le lien avec le cours.

Ou

ii) Un projet du choix de l'étudiant

Un étudiant qui souhaite proposer un projet de son choix, bien sûr, il le pourra, à condition de montrer le rapport avec le cours, et la faisabilité de réalisation.

Ou

iii) Un projet de la liste suivante.

Projet 1 (Synchromédia)

Extraction de caractéristiques pour l'écriture manuscrite en utilisant les moments de Zernike

Il s'agit d'implémenter et de tester les moments de Zernike pour l'extraction de caractéristiques. Le problème spécifique de l'écriture manuscrite sera traité. À cette fin, la base MNIST (60,000 exemples d'apprentissage et 10,000 exemples de texte) est disponible.

Référence:

Khotanzad, A. and Y.H. Hong, 1990. Invariant image recognition by Zernike moments. *IEEE TPAMI*, 12(5), pp 489-497.

Projet 2 (Synchromédia)

" Kernel-Based KNN " pour la reconnaissance d'écriture manuscrite

Il s'agit d'implémenter et d'évaluer la méthode de(s) plus proche(s) voisin(s) dans l'espace défini par la fonction noyau. Le problème spécifique de l'écriture manuscrite sera traité. À cette fin, la base MNIST (60,000 exemples d'apprentissage et 10,000 exemples de texte) est disponible.

Référence:

Bor-Chen Kuo; Jinn-Min Yang; Tian-Wei Sheu; Szu-Wei Yang, "Kernel-Based KNN and Gaussian Classifiers for Hyperspectral Image Classification," *Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008. IEEE International* , vol.2, no., pp.II-1006-II-1008, 7-11 July 2008

Projet 3 (Synchromédia)

Reconnaissance d'une chaîne de chiffres manuscrits

Il s'agit de programmer la segmentation d'une chaîne de chiffres manuscrits en utilisant un SVM pour valider la performance à partir du score de classification.

Référence:

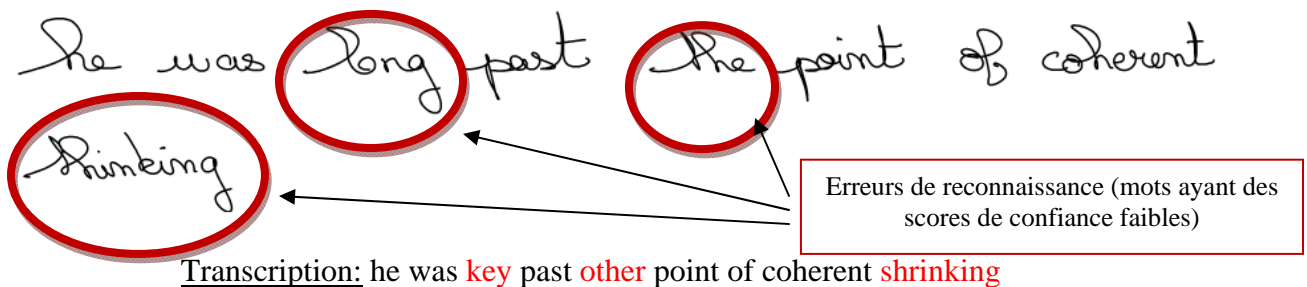
Casey, R.G.; Lecolinet, E., "Strategies in character segmentation: a survey," *Document Analysis and Recognition, 1995., Proceedings of the Third International Conference on* , vol.2, no., pp.1028-1033 vol.2, 14-16 Aug 1995.

Projet 4 (Sychromédia)

Fusion de scores hétérogènes pour le calcul d'indices de confiance appliqués à la reconnaissance d'écriture

Problématique

La reconnaissance d'écriture manuscrite consiste à retranscrire le texte qui a été écrit à partir de l'image de ce texte (appelée écriture hors-ligne comme, par exemple, pour des pages de documents scannés) ou de la suite des points du tracé dynamique de ce texte (appelée écriture en-ligne comme, par exemple, pour du texte écrit sur un PC tablette ou un tableau blanc interactif). Cette tâche de reconnaissance est très complexe et il est difficile d'obtenir une transcription entièrement correcte. Il est alors intéressant de pouvoir associer des mesures de confiance à la transcription finale afin d'identifier les parties de cette transcription pour lesquelles la confiance est élevée de celles pour lesquelles elle est faible : cela permet ensuite de ne considérer que les parties de la transcription ayant une confiance élevée mais aussi de reconsidérer celles qui ont une confiance faible en effectuant une nouvelle étape de reconnaissance en utilisant d'autres informations.



Objectif

Les indices de confiance se calculent généralement au niveau des mots du texte résultat de la reconnaissance. Différentes informations existent au niveau de chaque mot, des caractères le constituant ou encore des mots autour de lui dans la transcription. Ces différentes informations peuvent être de nature différente (probabiliste, discrète) et d'ordres de grandeur différents.

L'objectif est alors de calculer différentes caractéristiques au niveau des caractères et des mots des textes manuscrits (en considérant également les mots autour du mot considéré) et de calculer ensuite des scores de confiance au niveau des mots en proposant différentes méthodes pour fusionner ces caractéristiques.

Références

[1] J.F. Pitrelli and J. Subrahmonia and M.P. Perrone. *Confidence Modeling for handwriting Recognition: Algorithms and Applications*, International Journal on Document Analysis and Recognition (IJDAR), 8(1), pages 35-46, 2006.

Projet 5 (Synchromédia)

Skeleton-based shape clustering

In this project, the final goal is to **cluster** a set of **shapes** based on the distances between them. The shapes are binary images. They mainly composed on intersection of some lines and curves. Therefore, they can be considered as **network-like shapes**. In addition to form of each shape, its **orientation and position** with respect to a coordinate origin is considered as a feature in clustering. The main question is to find the best representation of the shapes, which then will be used to compute the distances. The preferred representation is supposed to be based on skeleton (skeleton, shock graph, etc). Not only the **topological**, but also the **geometrical** features of the connections should be considered in the representation. Possible ideas on correction of representation in order to capture **missed** features are also welcomed.

Projet 6 (Synchromédia)

Correcting a clustered set of objects

A set of objects is clustered using a set of **pre-assumed central nodes**. The first goal is to determine a **new** set of central nodes, which represents better the clusters rather than the initial set, using the clustered data. Secondly, after finding the new central nodes, correction of **miss-clustered** objects, without re-clustering all the objects is required.

Projet 7 (Synchromédia)

Horizontal sequence retrieval

A set of **shapes** are placed on a page. These shapes constitute a set of **objects**. A shape can be shared by two objects. An object consists of one or more shapes. The shapes/objects resemble a series of **lines** whose average distance is known *a priori*. The goal is to generate all possible **sequences** of objects on the lines.

Projet 8 (Synchronédia)

Une méthode statistique pour la découverte des ressources

De nouvelles technologies permettent le partage et l'utilisation, d'une façon sécuritaire, une grande collection de ressources, par exemple des ordinateurs, espaces de stockage, des logiciels etc.

Un utilisateur connecté à internet, et qui a besoin d'une ressource non disponible localement, peut soumettre au système une requête avec les caractéristiques désirées, un mécanisme de recherche prend en charge cette requête et retourne un ensemble de ressources candidates ; l'utilisateur n'a qu'à choisir celle qui répond au mieux à ses besoins.

Un tel service de base est nommé service de découverte des ressources (ressource mining) ; ce service doit faire face à plusieurs défis, à savoir la grande quantité de ressources disponible, la nature même des ressources qui sont hétérogène et dynamique.

Le but du projet est de proposer une méthode basée sur les statistiques permettant la découverte des ressources, et de comparer cette méthode avec celles qui existent.

Projet 9 (Synchromédia)

Objectif du projet: Étude de l'auto-similarité d'une image de document.

Données :

Soit l'image de document suivante « imgdoc.png », de taille $M \times M$ (M est dans \mathbb{Z}).

L'image montre trois classes d'objets, texte, arrière-plan, interférence (dégradation).

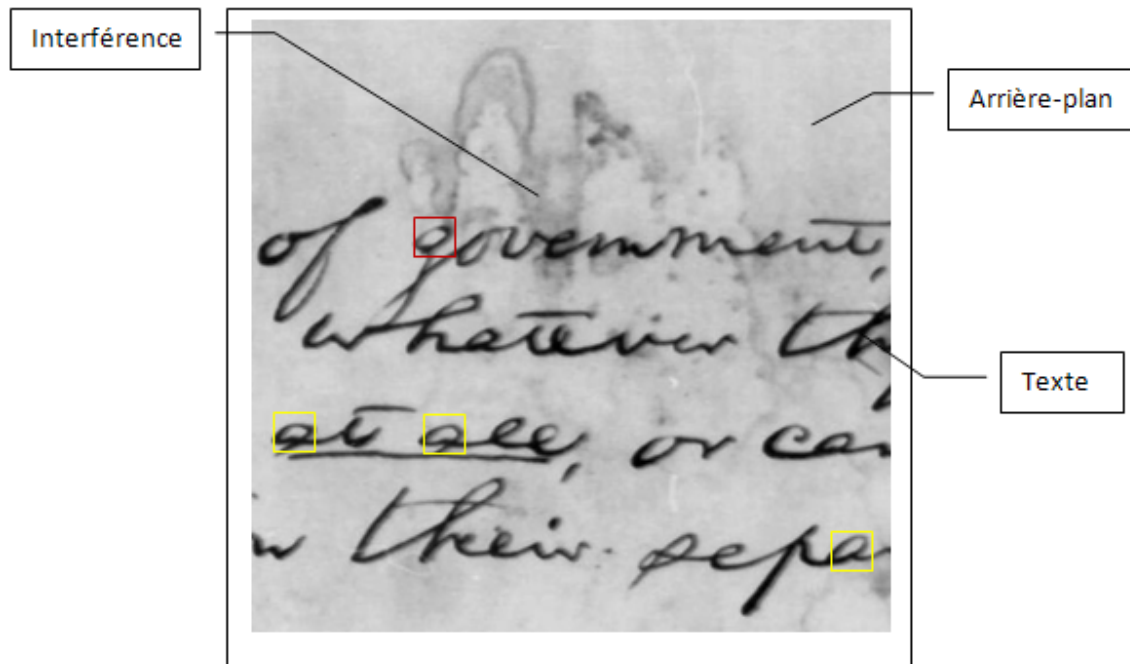


Figure 1, Image de document

Question 1 :

L'image peut être subdivisée en une grille (G), quadrillée recouvrant toute sa surface. Les éléments carrés constituant la grille sont appelés patchs et sont placés en adjacence horizontalement et verticalement l'un après l'autre sans chevauchement (voir figure 2 en bas). La taille de chaque patch est de $N \times N$, avec N un entier positif impair ($1 < N < M$).

Un patch est défini par un vecteur de caractéristiques VC à deux entrées

- i. μ , moyenne au sens des niveaux de gris des pixels appartenant à ce patch.
- ii. Σ , sa variance.

Il s'agit de classifier l'ensemble des patchs en trois classes différentes reflétant, les objets de l'image (texte, arrière-plan et interférences (dégradations)). La classification est basée sur le vecteur de caractéristique VC décrit plus haut. La méthode de classification par SVM est recommandée.

Question 2 :

Choisir un patch contenant un caractère ou une partie de caractère, par exemple celui sélectionné en rouge dans la figure (figure 1). Puis chercher les autres patches qui lui ressemblent en basant sur le vecteur de caractéristique VC décrit en (1.a).

Exemple, Sur la figure 1, les patches jaunes sont similaires approximativement au patch rouge.

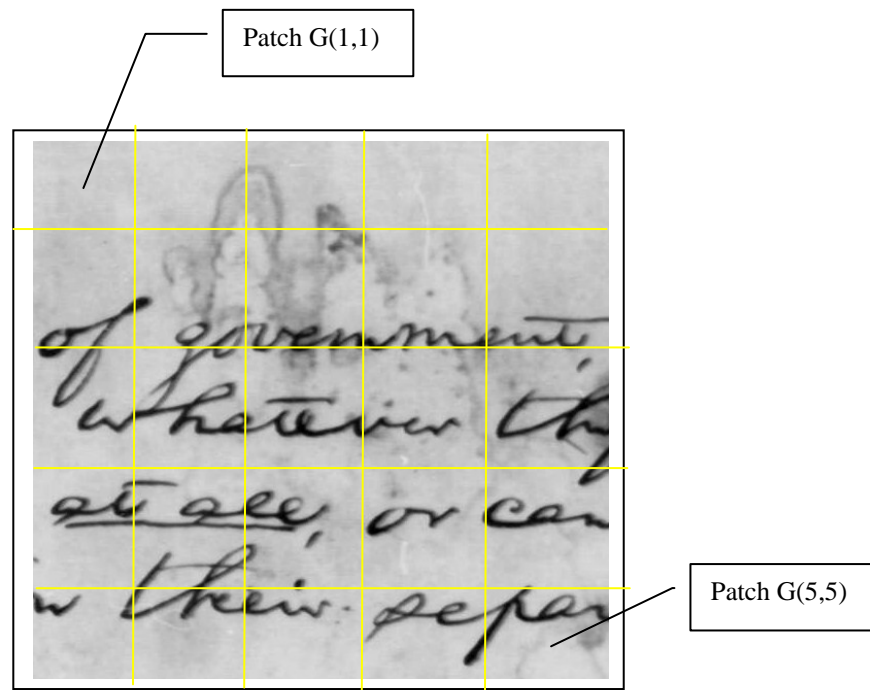


Figure 2. Exemple, Grille de patches carrés

Question 3 :

Le même travail sera demandé pour une image couleur en trois bandes Rouge, Vert et Bleue (Exemple imgdoc2.png).

SUITE DE PROJETS

- 1) Extraction et sélection des caractéristiques acoustiques. Application à la reconnaissance de caractères manuscrits.

Il s'agit d'étudier une nouvelle famille de caractéristiques de type acoustique. On transforme l'image binaire en densité d'un milieu hétérogène solide-fluide, et on calcule le vecteur des caractéristiques de la pression P induite par l'objet dans ce milieu visqueux.

Plus de détails chez le prof.

- 2) Caractéristiques de Gabor.

Il s'agit d'étudier la fonction de Gabor et son application à l'extraction de caractéristiques, connues sous le nom de caractéristiques de Gabor. L'étudiant doit étudier en premier la fonction de base de Gabor 1D, ensuite, la fonction 2D, et finalement les caractéristiques de Gabor 1D et 2D. Il y'a toute une panoplie de références sur la fonction de Gabor, mais on liste ci-dessous celles qui se rapportent aux caractéristiques de Gabor.

Référence :

Y. Ge, Q. Huo, Z-D. Feng, "Offline Recognition Of Handwritten Chinese Characters Using Gabor Features And CDHMM Modeling And MCE Training," in Proc. ICASSP-2002, Orlando, Florida

- 3) Étude de la variante *Modified Quadratic Discriminant Functions*. Application à la reconnaissance de caractères.

Référence :

Classificateur à partir du papier : F. Kimura, K. Takashina, S. Tsuruoka, and Y. Miyake, "Modified Quadratic Discriminant Functions and the Application to Chinese Character Recognition," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 9, No. 1, pp. 149-153, Jan. 1987.

- 4) Recouvrement de l'information dynamique à partir des images hors-ligne. Comparaison de deux méthodes.

Référence :

[1] Stefan Jager, Recovering Dynamic Information from Static, Handwritten Word Images, In ICDAR'97 & Ulm'2000.

[2] Y. Al-Ouali, M. Cheriet, and C.Y. Suen, Efficient Estimation of Pen Trajectory from off-line Handwritten Words, in Proc. ICPR'2002.

- 5) Critère du *Minimum Classification Error (MCE)* pour l'apprentissage des HMMs.

Il s'agit d'étudier le critère *Minimum Classification Error criterion (MCE)*, utilisé pour entraîner les HMMs. Il y'a des points particuliers que l'étudiant doit prendre en compte avant de s'attaquer à ce problème. Ces points sont mentionnés ci-dessous avec leur référence :

- Des notions de base de la reconnaissance de forme statistique, notamment le cadre théorique du classificateur de Bayes.
- Comprendre la différence entre une classification à contexte libre ou à contexte lié. Celle-ci est importante pour comprendre pourquoi on a besoin de classificateurs de type HMM au lieu d'autres techniques.
- Comprendre les notions de base des HMMs (discrets et continus).
- Comprendre l'idée de base de l'algorithme *Expectation-Maximisation (EM)* et par ricochet, la méthode de Baum-Welch pour l'estimation des paramètres.
- Finalement, de comprendre la méthode MCE, annoncée plus haut.

Références :

- [1] S. Theodoridis, K. Koutroubas, "*Pattern Recognition*," Ch. 9, Academic Press, 1999.
- [2] K. Fukunaga, "*Introduction To statistical Pattern Recognition*," Academic Press, 2nd Edition, 1990.
- [3] L. Rabiner, B-H. Juang, "*Fundamentals Of Speech Recognition*," Ch. 6, Printice Hall, 1993.
- [4] B-H Juang, W. Chou, "*Minimum Classification Error Rate Methods For Speech Recognition*," IEEE Trans. Speech And Audio Processing, Vol. 5, No. 3, pp. 257-265, May 1997.
- [5] B-H Juang, S. Katagiri, "*Discriminative Learning For Minimum Error Classification*," IEEE Trans. Signal Processing, Vol. 40, No. 12, pp. 3043-3054, Dec. 1992.

6) Kernel Methods and Feature Vector Selection.

Dans ce projet l'étudiant doit faire une revue sur la théorie des noyaux et de l'extension de certains algorithmes classiques comme l'analyse linéaire discriminante (LDA) à l'analyse discriminante généralisée (GDA). Il poursuit aussi son étude sur l'extraction des données pertinentes dans l'espace des caractéristiques, comment ces données sont ensuite projetées dans le sous-espace des vecteurs sélectionnés, où les algorithmes classiques sont utilisés. On doit montrer que les noyaux de type GDA ou Kernel Principal Component Analysis peuvent être facilement obtenus.

Références:

- Baudat G., Anouar F., "Generalized discriminant analysis using a kernel approach", *Neural Computation*, 12, pp.2385-2404, 2000.
- Baudat G., Anouar F., "Kernel-based Methods and Function Approximation", International Joint Conference on Neural Network, IJCNN 2001.

Pour plus de détails, consulter la section associée à la présentation des auteurs dans le cadre des séminaires-LIVIA: http://www.livia.etsmtl.ca/fr/events_fr.html

7) SVM pour la reconnaissance de formes. Revue de la littérature.

Il s'agit de faire une étude des techniques usuelles de reconnaissance de formes avec des SVMs (reconnaissance de visage, detection de visage, verification, reconnaissance de caractères, ...etc).

Références :

- [1] Byun, S.W. Lee. Applications of Support Vector Machines for Pattern Recognition: a Survey . LNCS, vol 2388.
- [2] J.C. Burges, A Tutorial on Support Vector Machines for Pattern Recognition (1998)
- [3] Colin Campbell. Kernel Methods: A Survey of Current Techniques.

8) Apprentissage des SVM : revue de la littérature .

Il s'agit d'étudier la panoplie de techniques d'apprentissage existantes des SVM : Algorithme SMO, algorithme de Keerthi, algorithme de Osuna, algorithme de Joachim, Lagrangian SVM, techniques d'optimisation non linéaires avec contraintes, techniques de programmation linéaires avec contraintes (LP optimization)

Références :

- [1] J.C. Burges, A Tutorial on Support Vector Machines for Pattern Recognition (1998)
- [2] C. Campbell, Kernel Methods: A Survey of Current Techniques.
- [3] J. Platt, *Fast Training of Support Vector Machines using Sequential Minimal Optimization*
- [4] O. Mangasarian. Lagrangian support Vector Machines. Machine Learning.

9) Noyaux de SVM pour la reconnaissance d'images de niveau de gris et de couleurs.

Il s'agit de faire une étude expérimentale de la validité des noyaux d'histogrammes et de noyaux de Hausdroff pour la classification ou le regroupement d'images de couleurs (RGB, HSV, HSI) et d'images de niveau de gris. Une étude comparative avec les noyaux RBF et KMOD est aussi demandée.

Ressources :

code matlab qui implémente le SVM pour la classification / regroupement. Ce dernier permet aussi de coder des nouveaux noyaux.

Références :

- [1] A. Barla, E. Franceschi, F. Odone, A. Verri. Image Kernels. LNCS, vol. 2388.
- [2] N.E. Ayat, M. Cheriet, C. Suen. KMOD: A two parameter SVM kernel for Pattern Recognition. ICPR 2002.

10) Méthodes hybrides de décision : combinaison des méthodes génératives probabilistes avec des méthodes discriminantes

Il s'agit de faire une étude théorique et/ou expérimentale de la validité de la validité de l'approche hybride en combinant les points forts des modèles génératifs probabilistes (i.e. HMMs, Fisher Discriminant Analysis, Naive Bayes) avec les modèles discriminants (i.e. SVM, MLP, Generative Additive Models)

Ressources :

La base de données MNIST est disponible sur le site <http://yann.lecun.com/exdb/mnist/>

Références :

- [1] Y. Rubenstein and T. Hastie, "Discriminative vs Informative Learning," Proc. Of Knowledge Discovery and Data mining (1997).
- [2] T. Jaakkola and D. Haussler, "Exploiting Generative Models in Discriminative Classifiers," In Advances in NIPS 11 (1998).
- [3] L. Quna and S. Bengio, Hybrid Generative-Discriminative Models for Speech and Speaker Recognition," IDIAP Research Report March (2002).
- [4] A. Ng and M. Jordan, On Generative vs. Discriminative Classifiers: A comparison of logistic Regression and Naïve Bayes," Proc. of Advances in NIPS 14 (2002).
- [5] G. Bouchard, The Trade-off Between Generative and Discriminative Classifiers, in Proc. Of Advances in NIPS 15 (2003).
- [6] L. Prevost, C. Michel-Sendis, A. Moises, L. Oudut and M. Millgram, "Combining Model-based and Discriminative Classifiers: Application to Handwritten Character Recognition," Proc. ICDAR'2003, 31-35.

Bon Choix & Bon Projet