

Apprentissage progressif pour la reconnaissance de symboles dans les documents graphiques

Sabine Barrat, Salvatore Tabbone

► **To cite this version:**

Sabine Barrat, Salvatore Tabbone. Apprentissage progressif pour la reconnaissance de symboles dans les documents graphiques. 15e congrès francophone AFRIF-AFIA Reconnaissance des Formes et Intelligence Artificielle - RFIA 2006, Jan 2006, Tours, France. 2006. <inria-00118259>

HAL Id: inria-00118259

<https://hal.inria.fr/inria-00118259>

Submitted on 4 Dec 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Apprentissage progressif pour la reconnaissance de symboles dans les documents graphiques

S. Barrat

S. Tabbone

LORIA-UMR 7503
Université de Nancy 2
BP 239, 54506 Vandoeuvre-les-Nancy

email : {barrat,tabbone@loria.fr}

Résumé

Les méthodes actuelles de reconnaissance de symboles donnent de bons résultats quand il s'agit de reconnaître peu de symboles différents qui sont peu bruités et souvent déconnectés du graphique. Cependant, dans le cas d'applications réelles, les méthodes sont encore mal maîtrisées quand il s'agit de discriminer dans de grandes bases entre plusieurs centaines de symboles différents, souvent complexes et bruités et encapsulés dans les couches graphiques. Dans ce contexte, il est nécessaire de mettre en oeuvre des méthodes d'apprentissage. Nous présentons dans cet article une méthode d'apprentissage progressif pour la reconnaissance de symboles qui améliore son propre taux de reconnaissance au fur et à mesure que de nouveaux symboles sont reconnus dans les documents. Pour ce faire, nous proposons une nouvelle exploitation de l'analyse discriminante qui fournit des règles d'affectation à partir d'un échantillon d'apprentissage sur lequel les appartenances aux classes sont connues (apprentissage supervisé). Mais cette méthode ne se révèle efficace que si l'échantillon d'apprentissage et les données ultérieures sont observés dans les mêmes conditions. Or cette hypothèse est rarement vérifiée dans les conditions réelles. Pour pallier ce problème, nous avons adapté une approche récente d'analyse discriminante conditionnelle qui ajoute à chaque observation l'observation d'un vecteur aléatoire, représentatif des effets parasites observés dans l'analyse discriminante classique.

Mots Clef

Analyse discriminante, apprentissage progressif, reconnaissance de symboles, traitement de document graphiques.

Abstract

Current symbols recognition methods have good results when we want to recognize few different symbols with

low noise and often disconnected from the graphics. However, in real life, when we have to distinguish in large symbols databases between several hundred of different symbols, often complex and embedded in graphics methods are provide weak results. In this context it's necessary to carry out learning methods. We use a progressive learning method for symbols recognition which improve its own recognition rate when new symbols are recognized in documents. We propose a new exploitation of discriminant analysis which provide allocation rules from learning sample for which class belonging are known (supervised learning). However this method gives efficiency only if learning samples and data are defined in the same conditions but it is rarely the case in real conditions. In order to overcome this problem, we have adapted a recent approach of conditional discriminant analysis which add to each observation a random vector observation, representative of parasite effects observed in usual discriminant analysis.

Keywords

Discriminant analysis, conditional discriminant analysis, progressive learning, symbol recognition, graphics document analysis.

1 Introduction

Ces dernières décennies nous assistons à un fort engouement de la communauté scientifique pour la reconnaissance de symboles. La montée en puissance de l'Internet et le développement de capacités de stockage gigantesques a fortement contribué au foisonnement d'approches pour traiter ce problème de reconnaissance. De nombreux articles de synthèse ont été réalisés [1, 2, 3, 4, 13, 6, 7] ce qui souligne l'importance de cet axe de recherche en reconnaissance des formes et en analyse de documents. On peut distinguer deux grand types de méthodes de reconnaissance de symboles dans des

documents [8, 9, 10, 11, 12, 14, 5], chaque type utilisant un mode de représentation des symboles particulier : les méthodes dites statistiques, basées sur des transformées et autres caractéristiques mathématiques, et les méthodes structurelles, pour lesquelles les symboles sont représentés par une description vectorielle de leur forme. Cependant, il n'existe pas pour l'instant de méthode universelle à ce problème de reconnaissance car dans la réalité les symboles sont rarement déconnectés du document graphique ce qui pose un problème ouvert de segmentation car pour segmenter il faudrait avoir reconnu le symbole et réciproquement, pour reconnaître le symbole il faudrait au préalable l'extraire du document. Par ailleurs, les symboles sont généralement bruités et peuvent subir des déformations. Ainsi une méthode de reconnaissance doit être robuste au bruit et insensible à diverses transformations géométriques simples telles que la rotation, la translation et le zoom. De plus, les méthodes sont encore mal maîtrisées quand il s'agit de discriminer dans de grandes bases entre plusieurs centaines de symboles différents, souvent complexes et bruités et encapsulés dans les couches graphiques. Dans ce contexte il est nécessaire de mettre en oeuvre des méthodes d'apprentissage.

Cet article est dédié à une méthode d'apprentissage progressif. Ici nous n'abordons pas le problème d'un point vue structurelle[15, 16] mais statistique. Il existe beaucoup de travaux sur les méthodes de classification statistique et le lecteur intéressé peut consulter l'article de synthèse de A.K. Jain[4]. Nous avons adapté l'analyse discriminante à la reconnaissance de symboles. L'inconvénient de l'analyse discriminante est qu'elle repose sur certaines hypothèses qui ne sont pas toujours vérifiées à cause de la grande diversité des données réelles. Ainsi, elle peut donner des résultats erronés. On est alors amené à proposer et à utiliser des méthodes alternatives plus adaptées au cas de figure qui nous intéresse. Ainsi nous utilisons ici une variante de l'analyse discriminante décisionnelle usuelle : l'analyse discriminante conditionnelle (ADC) proposée dans l'article de A. Baccini[17]. Les tests conduits, dans cet article, sur des unités statistiques classiques ont donné des résultats encourageants avec une complexité peu élevée et il nous a donc semblé intéressant de l'adapter au problème de la reconnaissance de symboles.

Le reste de l'article est organisé comme suit : dans la section suivante nous rappelons le principe de base de l'analyse discriminante. Ensuite nous décrivons l'analyse discriminante conditionnelle (section 3). Enfin, nous montrons des résultats expérimentaux (section 4) et concluons en donnant des perspectives à notre méthode (section 5).

2 Analyse discriminante décisionnelle

L'analyse discriminante est une famille de techniques destinées à décrire et à classer des individus caractérisés par un nombre important de variables. L'origine de cette méthode remonte aux travaux de Fisher et de Mahalanobis dans les années trente. Son but est de trouver le sous-espace de projection qui sépare au mieux les observations (individus). Le problème est alors ramené à un test de Student, puisqu'il faut que les moyennes intra-groupes des groupes formés par la projection soient significativement différentes. L'idée est alors, qu'étant donné que les unités d'une même classe doivent se projeter sur une classe la plus compacte possible et que les différentes classes doivent être les plus séparées possibles, le rapport des variances inter-classes et intra-classes dans la projection qui sépare le mieux les classes est maximal. On va donc utiliser un test \mathcal{F} de Fisher et maximiser \mathcal{F} qui est le rapport des variances inter-classes et intra-classes.

2.1 Formulation de la méthode et notations

L'analyse discriminante décisionnelle a pour but l'explication d'une variable qualitative C par p variables quantitatives $X_{k,j}$, $k = 1, 2, \dots, p$. Les fonctions linéaires discriminantes sont des combinaisons linéaires de ces variables, séparant au mieux les classes. Disposant d'un nouvel individu sur lequel on a observé les $X_{k,j}$ ($j = 1, \dots, n$) mais pas C, il s'agit maintenant de décider de la modalité de C (ou de la classe correspondante) de ce nouvel individu. On parle donc de problème d'affectation. L'analyse discriminante fournit des règles de décision (ou d'affectation) à partir d'un échantillon d'apprentissage sur lequel les appartenances aux classes sont connues (apprentissage supervisé).

Soient :

- X_j , un vecteur représentatif d'une observation :

$$X_j = {}^t(X_{1,j}, X_{2,j}, \dots, X_{p,j}),$$

où p est le nombre de composantes du vecteur caractéristique.

- $X_{l,j}$, la j-ième composante de la l-ième classe :

$$X_{l,j} = {}^t(X_{1,l,j}, X_{2,l,j}, \dots, X_{p,l,j})$$

$\forall j = 1, 2, \dots, n_l$, où n_l est le nombre d'observations dans la classe l et $\forall l = 1, 2, \dots, s$, où s correspond au nombre de classes.

- \bar{X} , le barycentre global.
- \bar{X}_l , le barycentre de la classe l vecteur des moyennes de chaque variable dans la classe l.

- W la matrice des covariances intra-classes supposée identique dans chaque classe, de dimension $p \times p$, symétrique et régulière.
- Soit X un tableau de données, à $n = \sum_{l=1}^s n_l$ lignes (individus) et p colonnes (variables). Les n observations sont partitionnées en l groupes. On suppose que le j-ième groupe a une distribution $\mathcal{N}(\overline{X}_j, W)$ ¹
- Soit x une nouvelle observation (vecteur 1 ligne, p colonnes). On affecte cette nouvelle observation au groupe avec la probabilité a posteriori la plus grande. Supposons que la probabilité a priori d'appartenance à un groupe ne change pas ($p_1 = p_2 = \dots = p_s$). Ainsi, la probabilité a posteriori de la classe j s'écrit :

$$\frac{\exp(-\frac{1}{2}\|x - \overline{X}_j\|_{W^{-1}}^2)}{\sum_{i=1}^s \exp(-\frac{1}{2}\|x - \overline{X}_i\|_{W^{-1}}^2)}$$

- Pour pouvoir affecter cette nouvelle observation il faut maximiser cette probabilité a posteriori ou bien minimiser la quantité :

$$\|x - \overline{X}_j\|_{W^{-1}}^2 = {}^t(x - \overline{X}_j)W^{-1}(x - \overline{X}_j) \quad (1)$$

2.2 Estimation des paramètres

On estime \overline{X}_j , $\overline{\overline{X}}$ et W à partir de l'échantillon d'apprentissage pour lequel on connaît le groupe d'appartenance des individus. La matrice W est estimée par la matrice de covariances intragroupes empiriques et les \overline{X}_j sont estimés par les moyennes empiriques des groupes. En règle générale, la matrice W est inversible, ce qui nous permet de calculer les distances entre les individus supplémentaires et les centres de gravité de chaque groupe suivant la formule 1.

$$\overline{X}_j = {}^t(\overline{X}_{1,j}, \overline{X}_{2,j}, \dots, \overline{X}_{p,j}) \quad (2)$$

$\forall j = 1, 2, \dots, s$

où

$$\overline{X}_{i,j} = \frac{1}{n_j} \sum_{k=1}^{n_j} X_{i,j,k}$$

$\forall i = 1, 2, \dots, p$

$$\overline{\overline{X}} = {}^t(\overline{\overline{X}}_1, \overline{\overline{X}}_2, \dots, \overline{\overline{X}}_p)$$

où

$$\overline{\overline{X}}_i = \frac{1}{s} \sum_{j=1}^s n_j \overline{X}_{i,j} = \frac{1}{s} \sum_{j=1}^s \sum_{k=1}^{n_j} X_{i,j,k}$$

¹Cette hypothèse est considérée pour affirmer le caractère optimal de la technique de discrimination linéaire. Pour des observations non normales, ce caractère optimal n'est pas assuré, mais la méthode peut encore être justifiée si l'on s'en tient aux propriétés du second ordre, c'est à dire portant seulement sur les moyennes et les variances. Cette remarque reste vraie pour l'analyse discriminante conditionnelle.

$\forall i = 1, 2, \dots, p$

$$W = (W_{i,i'})_{i,i' \in 1,2,\dots,p}$$

où

$$W_{i,i'} = \frac{1}{n} \sum_{j=1}^s \sum_{k=1}^{n_j} (X_{i,j,k} - \overline{X}_{i,j})(X_{i',j,k} - \overline{X}_{i',j}) \quad (3)$$

où

$j \in 1, \dots, s$, s = nombre de classes

n_j = nombre d'éléments dans la classe j

$X_{i,j,k}$ = observation de la variable i sur le k-ième individu de la classe j

$X_{i',j,k}$ = observation de la variable i' sur le k-ième individu de la classe j.

En résumé, l'analyse discriminante décisionnelle permet d'affecter une unité statistique à un groupe sur la base de mesures réelles, selon une règle de décision obtenue à partir d'un échantillon d'apprentissage (cf formule 1). On suppose que chaque groupe est caractérisé par un vecteur moyen (cf. formule 2) qui est estimé au sein d'un échantillon d'apprentissage, pour lequel les appartenances aux groupes sont connues. Ainsi, on est capable d'estimer la variabilité à l'intérieur de chacun des groupes ainsi que la dispersion des mesures d'un groupe à l'autre (cf. formule 3). Ceci permet d'affecter de nouveaux individus dont on ne connaît pas le groupe d'appartenance.

Cette méthode est fiable seulement si les conditions des mesures restent invariantes c'est-à-dire si les données sont observées sous les mêmes conditions pendant et après l'apprentissage. Cependant, ceci est peu vraisemblable dans la plupart des situations réelles car les conditions de mesure dépendent très souvent de facteurs de variabilité importante, mais contrôlée, qu'on appellera facteurs de dérive. Dans ce cas l'échantillon d'apprentissage ne réussit pas à bien déterminer l'appartenance à un groupe d'un individu supplémentaire et l'apprentissage est donc partiel.

3 Analyse discriminante conditionnelle

Pour pallier ce problème de facteur de dérive (cf. section 2), nous nous sommes intéressés à une analyse convenablement modifiée qui complète l'apprentissage initial par un contrôle progressif des conditions d'utilisation. Dans des nombreuses situations réelles l'échantillon d'apprentissage et l'échantillon de validation ne sont pas observés dans les mêmes conditions. Dans ce cas, les mesures effectuées sur les individus dépendent des facteurs de dérive qu'il serait prudent de considérer. L'idée est d'ajouter à chaque observation X, l'observation d'un vecteur aléatoire Y, représentatif des

effets parasites ou facteurs de dérive. Cette méthode récente appelée analyse discriminante conditionnelle a été proposée par A. Baccini[17] dans un contexte différent.

3.1 Formulation de la méthode et notations

Supposons que les estimations des $\overline{X_j}, j \in 1, 2, \dots, s$ et de W ont été préalablement faites. On suppose en plus que l'on dispose d'un tableau Y à n lignes et h dimensions qui va tenir compte du phénomène de dérive. On se place sous les hypothèses suivantes :

1. La moyenne intra-classes de Y est \overline{Y} (la moyenne empirique), quelle que soit la classe j .
2. La variance de $Z = {}^t(X, Y)$ s'écrit :

$$W_Z = \begin{bmatrix} W_X & W_{XY} \\ W_{YX} & W_Y \end{bmatrix}$$

où $W_X = W$: matrice de covariances intra-classes empiriques de X

$W_{XY} = {}^t W_{YX}$: matrice des covariances de X et de Y

W_Y : matrice de covariances de Y supposée régulière La variance de Z est une variable intra-classes supposée identique pour chaque classe.

3. Z suit une loi normale.

Pour pouvoir considérer les facteurs de dérive, on va effectuer une analyse discriminante décisionnelle non plus sur X , mais sur Z , vecteur à $p+h$ dimensions. Ainsi, les changements effectués sont les suivants :

1. La matrice W est remplacée par la matrice W_Z
2. Les $\overline{X_j}$ sont remplacés par $\mu_j = {}^t(\overline{X_j}, \overline{Y})$

Posons $C = W_X - W_{XY}W_Y^{-1}W_{YX}$ qu'on suppose régulière. D'après le principe de l'analyse discriminante décisionnelle usuelle, on doit affecter un individu supplémentaire ${}^t(x, y)$ à la classe j qui minimise la quantité :

$$\begin{bmatrix} {}^t(x - \mu_j) \\ {}^t(y - \mu_j) \end{bmatrix} W_{Z^{-1}} [(x - \mu_j)(y - \mu_j)]$$

ce qui revient à minimiser l'expression :

$$\|(x - \overline{X_j}) - W_{XY}W_Y^{-1}(y - \overline{Y})\|_{C^{-1}}^2 \quad (4)$$

Si on se replace dans la métrique W^{-1} , il faut affecter une nouvelle observation ${}^t(x, y)$ à la classe j pour laquelle l'expression suivante est minimale :

$$\begin{aligned} \|(x - \overline{X_j}) - W_{XY}W_Y^{-1}(y - \overline{Y})\|_{C^{-1}}^2 = \\ {}^t((X - \overline{X_j}) - W_{XY}W_Y^{-1}(y - \overline{Y})) \cdot \\ (W_X - W_{XY}W_Y^{-1}W_{YX})^{-1} \cdot \\ ((X - \overline{X_j}) - W_{XY}W_Y^{-1}(y - \overline{Y})) \end{aligned}$$

Le point important dans cette analyse est la correction de x . Cependant, le remplacement de W^{-1} par C^{-1} améliore théoriquement l'analyse.

3.2 Estimation des paramètres

On suppose que les estimations des paramètres $W_X, \overline{X_j}$ et \overline{X} sont préalablement faites dans l'échantillon d'apprentissage comme on l'a montré au paragraphe 2.2.2.

On a :

$$W_{XY} = E[(X - \overline{X_j})^t(Y - \overline{Y})] = E[X^t(Y - \overline{Y})]$$

Donc on peut estimer W_{XY} par :

$$\frac{1}{N} \sum_{i=1}^N X_i^t(Y_i - \overline{Y})$$

où :

- N est le nombre de données pour lesquelles Y est disponible.
- On rappelle que Y est mesuré sur l'échantillon d'apprentissage et l'échantillon test.
- \overline{Y} est la moyenne empirique des Y_i .
- De même W_Y est estimé par la matrice des covariances empiriques de Y sur l'ensemble des observations où cette variable est disponible.
- C'est grâce à l'hypothèse 2 (§3.1) que W_{XY} est estimable au moyen de l'ensemble des données, même sans connaître le groupe de provenance de X .

4 Choix des paramètres : descripteurs utilisés

4.1 Vecteur X

Pour que l'analyse discriminante se fasse dans les meilleures conditions possibles, il faut au préalable choisir des variables pertinentes. Dans notre contexte, ces variables peuvent être obtenues en utilisant un ou plusieurs descripteurs qui vont nous permettre d'extraire des images des variables quantitatives.

Pour générer des variables pertinentes, qui vont permettre de bien discriminer les données, il nous faut des descripteurs les plus robustes possible au bruit, aux déformations, et si possible ayant des propriétés d'invariance à certaines transformations géométriques. En effet, ces propriétés d'invariance vont permettre de classer correctement un même symbole indépendamment de sa position et de sa taille dans le document.

Notre choix s'est donc porté sur la R-signature[18] qui se base sur la transformée de Radon pour représenter une image. La transformée de Radon est la projection d'une image dans un plan particulier. Cette projection possède des propriétés géométriques intéressantes qui font d'elle un bon descripteur de forme. Suivant ces propriétés géométriques, une signature de

la transformée est créée. Cette signature vérifie les propriétés d'invariance à certaines transformations géométriques, telles que la translation et le changement d'échelle (après normalisation). Par contre l'invariance à la rotation est restaurée par permutation cyclique de la signature ou directement à partir de sa transformée de Fourier. La figure 1 montre des exemples de R-signature pour un symbole qui a subi indépendamment un zoom et une rotation. Ainsi, concrètement, pour notre analyse discriminante, pour chaque symbole de l'échantillon d'apprentissage et de test, nous calculons sa R-signature.

4.2 Vecteur conditionnel Y

Le vecteur Y constitue le point clé de la méthode. Le choix de ses composantes est essentiel pour sa réussite, et il peut être envisagé de deux façons :

- On peut d'abord effectuer un choix "externe" en considérant les mesures d'un ou plusieurs indicateurs indépendants de la classe à laquelle appartient le symbole étudié et susceptibles d'être bien corrélés avec les effets parasites à maîtriser et donc bien représentatifs d'une éventuelle dérive.
- On peut aussi chercher à effectuer un choix "interne" en analysant convenablement les données (pendant et après apprentissage) afin de découvrir d'éventuelles combinaisons des mesures X qui semblent les plus caractéristiques d'une éventuelle dérive tout en étant indépendantes du groupe d'appartenance.

Ces deux types de construction de Y ne s'excluent pas, certaines coordonnées de Y pouvant être obtenues de la première façon, les autres étant obtenues de la seconde.

Quel que soit le type de construction choisi, les composantes doivent respecter ces deux contraintes, sans quoi le succès de l'analyse pourrait être compromis. Les variables choisies doivent être :

1. Représentatives d'effets parasites de l'analyse, de facteurs de dérive.
2. Indépendantes de la classe du symbole considéré.

Notre choix s'est porté sur le premier type. Il nous fallait trouver pour construire le vecteur Y une mesure représentative des effets parasites de l'analyse et qui ne soit pas prise en compte dans les descripteurs utilisés pour les vecteurs X . Pour ce faire, nous avons cherché à minimiser le résidu de la régression linéaire entre la R-signature du symbole à reconnaître et la R-signature de chaque modèle représentant de classe. Le vecteur Y choisi correspond au résidu minimal trouvé. Ce résidu est censé représenter l'information supplémentaire par rapport aux modèles qui est liée à la dégradation que les symboles subissent en condition réelle.

5 Résultats expérimentaux

Nous avons utilisé des symboles issus de la base GREC pour effectuer nos tests [20] Cette base (voir fig 1) a

été créée spécialement dans le cadre du concours de reconnaissance de symbole à GREC'2005.

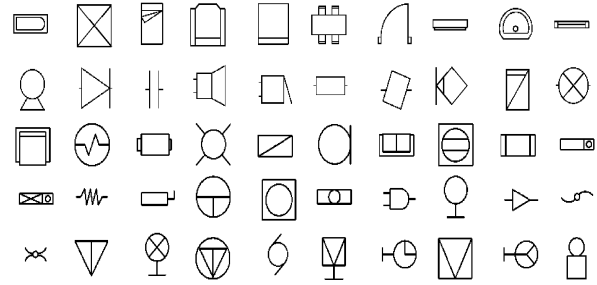


FIG. 1 – Base de symboles GREC.

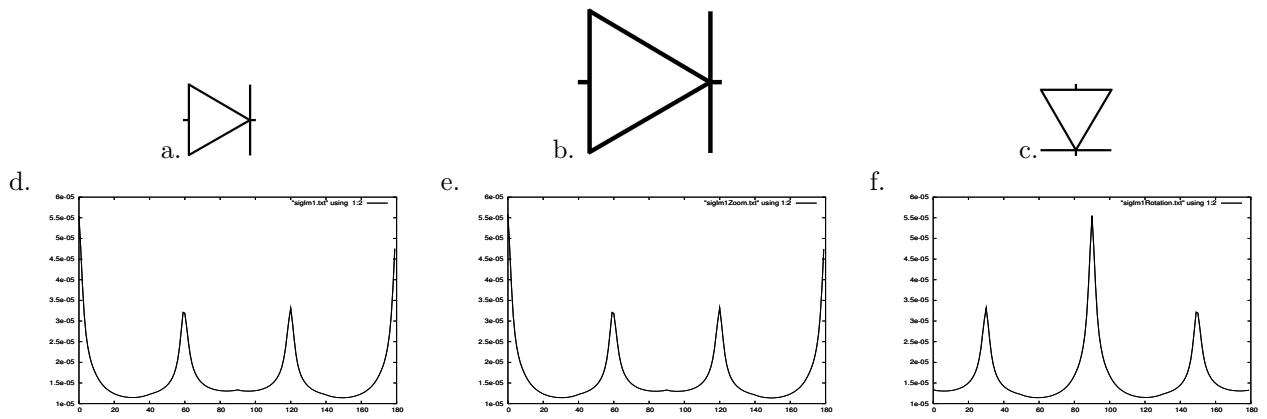
Les symboles de cette base sont issus principalement de deux domaines d'application, architecture et électronique, car ces symboles sont le plus largement utilisés par les équipes et représentent un grand nombre de formes différentes. Cent-cinquante modèles différents ont été utilisés, plusieurs d'entre eux ayant des formes similaires.

On dispose donc d'une base de 150 symboles parfaits. A partir de cette base, pour des raisons de commodité de traitement, nous avons pour l'instant extrait une sous-base de 50 modèles.

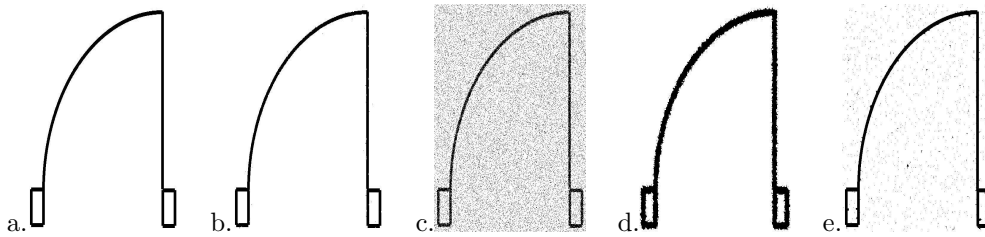
Sur ces modèles nous avons effectué des dégradations basées sur le modèle Kanungo [19]. Ces dégradations sont similaires au bruit obtenu quand un document est scanné, imprimé ou photocopie (bruits de type global et local et fermeture morphologique). Nous avons ainsi appliqué quatre sortes de bruits différents, à des intensités différentes, sur chacun des 50 modèles retenus, afin d'obtenir une base de travail de quatre classes de bruits contenant chacune 100 dégradations pour chaque modèle, de même type mais à une intensité différente. On dispose ainsi d'une base de 20000 symboles différents (4 classes de bruits \times 50 modèles \times 100 intensités de dégradation).

Le TAB.2 présente un modèle dégradé avec la même intensité mais de 4 façons différentes (4 classes de bruits).

Sur cette base nous avons défini plusieurs test d'apprentissage. Nous avons constitué plusieurs échantillons d'apprentissage constitués chacun d'une classe de bruit différente (l'échantillon d'apprentissage contient 5000 symboles : 100 déformations distinctes des 50 modèles) et un échantillon de test constitué de deux autres classes de bruits (l'échantillon de test contient 10000 symboles : 2 classes de bruits contenant chacune 100 déformations distinctes par modèles, on a 200 déformations distinctes des 50 modèles). Pour cet échantillon d'apprentissage nous avons calculé le taux de reconnaissance dans l'échantillon test à chaque



TAB. 1 – Exemples de R-signatures. a), b) et c) respectivement un symbole parfait, le même symbole avec un zoom X 2 et le symbole parfait tourné de 90 degré. d), e) et f) les R-signatures respectives.



TAB. 2 – Exemple d'un symbole "parfait" a) auquel on a appliqué 4 dégradations différentes de même intensité en b), c), d) et e).

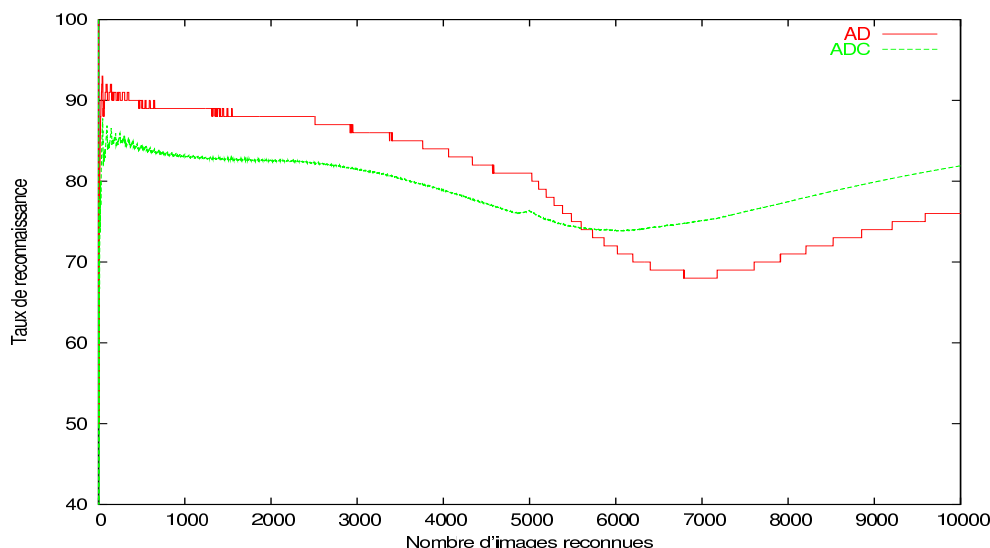
image reconnue, pour l'AD et pour l'ADC. Le TAB.3 présente les résultats d'un de ces tests. Les résultats ne commencent à être significatifs qu'à partir de 50 symboles reconnus environ. A partir de ce moment, nous remarquons qu'au début de l'apprentissage et jusqu'à la moitié, l'AD et l'ADC ont un comportement similaire : le taux de reconnaissance chute progressivement (de 92% à 74% pour l'AD et de 85% à 74% pour l'ADC) mais l'AD donne des résultats légèrement meilleurs que l'ADC (taux de reconnaissance meilleur d'environ 5 à 10 points). A partir de 5500 symboles reconnus environ, soit à un peu plus de la moitié de l'apprentissage, le taux de reconnaissance de l'AD continue de chuter jusqu'à 7000 symboles reconnus et remonte ensuite légèrement (jusqu'à 76%), contrairement à celui de l'ADC qui remonte progressivement jusqu'à la fin de l'apprentissage pour atteindre le taux de 82% et est ainsi meilleur que celui de l'AD dans la deuxième partie de l'apprentissage.

Ces résultats montrent que l'ADC a une meilleure résistance à l'échelle (en terme de quantités de symboles) que l'AD, or ces types d'approche, comme toutes les méthodes d'apprentissage, deviennent justement nécessaires lorsqu'on travaille sur de grandes bases. L'ADC semble donc être à privilégier par rapport à l'AD et donne de bons résultats quand on l'applique à notre problème de reconnaissance de symboles.

6 Conclusion et perspectives

Dans ce papier nous avons proposé une adaptation originale d'une méthode d'analyse discriminante conditionnelle. Les résultats obtenus sont prometteurs mais doivent être considéré comme préliminaires car ils doivent être étendu à de grandes bases de symboles pour voir l'apport effectif de la méthode. Notre choix pour la variable complémentaire Y s'est porté sur les résidus entre le modèle parfait et le symbole à reconnaître. Nous avons vu que cette variable complémentaire permet de prendre en compte les effets de dérive liés aux limites du descripteur choisi et de la méthode d'analyse discriminante.

L'intérêt est donc de tenir compte d'autres perturbations du symbole car il est difficile avec un descripteur unique de prendre en compte toutes les transformations possibles. Pour cette raison, nous souhaitons intégrer dans de futurs travaux des descripteurs qui prennent en compte certaines déformations affines du symbole et qui ne sont pas invariantes pour la R-signature. Par exemple, sur la figure 2 nous avons reproduit le symbole de la figure avec un zoom différents sur les abscisses et les ordonnées. Nous constatons que dans ce cas les R-signatures entre le symbole parfait et le symbole déformé sont différentes (cf Fig.2.b et TAB.1.d).



TAB. 3 – Evolution du taux de reconnaissance de l’ADC par rapport à celui de l’AD avec un apprentissage sur 5000 symboles (100 symboles par classe, 50 classes en tout) et une reconnaissance sur 10000 (200 symboles par classe, 50 classes en tout).

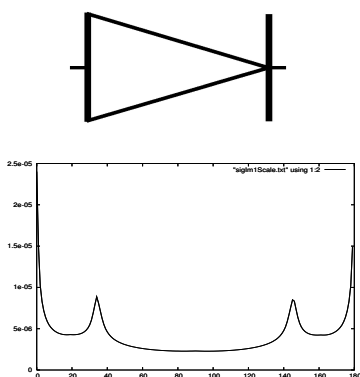


FIG. 2 – Exemple de \mathcal{R} -signature d’un modèle déformé (un zoom de 150% sur les x et de 300% sur les y par rapport au modèle parfait de la Fig.1.a).

Références

- [1] A.K. Chhabra, Graphic Symbol Recognition : An Overview, *Graphics Recognition : methods and applications (GREC’97)*, Nancy, France, 1997, pp. 68–79.
- [2] Cordella L. P., Vento M., Symbol recognition in documents : a collection of techniques?, *International Journal on Document Analysis and Recognition*, vol. 3(2), 2000, pp. 73–88.
- [3] Doermann D., The Indexing and Retrieval of Document Images : A Survey, *Computer Vision and Image Understanding*, vol. 70(3), 1998, pp. 287–298.
- [4] A.K. Jain, Duin R. P. W., Mao J., Statistical Pattern Recognition : A Review, *Image and Vision Computing*, vol. 22 (1), 2000, pp. 4–37.
- [5] Lladós J., Valveny E. and Sanchez G., Symbol recognition : current advances and perspectives, *Graphics Recognition : methods and applications (GREC’01)*, LNCS 2390, 2001, pp. 104–127.
- [6] Loncaric S., A Survey of Shape Analysis Techniques, *Pattern Recognition*, vol. 31(8), 1998, pp. 983–1001.
- [7] Zhang D., Lu G., Review of shape representation and description techniques, *Pattern Recognition*, vol. 37(1), 2004, pp. 1–19.
- [8] Adam S., Ogier J., Cariou C., Mullot R., Labiche J., Gardes J., Symbol and character recognition : application to engineering drawings, *International Journal on Document Analysis and Recognition*, vol. 3(2), 2001, pp. 89–101.
- [9] DeStefano C., Tortella F., Vento M., An entropy based method for extracting robust binary templates, *Machine Vision and Applications*, vol. 8, 1995, pp. 173–178.
- [10] Doermann D., Rivling E., Weiss I., Applying algebraic differential invariants for logo recognition, *Machine Vision and Applications*, vol. 9, 1996, pp. 73–86.
- [11] Fränti P., Mednonogov A., Kyrki V., Kälviäinen H., Content-based matching of line-drawing images using the Hough transform, *International*

Journal on Document Analysis and Recognition, vol. 3(2), 2000, pp. 117–124.

- [12] Ghorbel F., A complete invariant description for gray level images by harmonic analysis approach, *Pattern Recognition Letters*, vol. 15, 1994, pp. 1043–1051.
- [13] Lladós J., Martí E., Villanueva J., Symbol Recognition by Error-Tolerant Subgraph Matching Between Region Adjacency Graphs, *IEEE Transactions on PAMI*, vol. 23(10), 2001, pp. 1137–1143.
- [14] Yu B., Automatic Understanding of Symbol-Connected diagrams, *International Conference on Document Analysis and Recognition, Montréal*, vol. 2, 1995, pp. 803–806.
- [15] Messmer B., Bunke H., Automatic learning and recognition of graphical symbols in engineering drawings, *Graphics Recognition : methods and applications (GREC'95), LNCS 1072.*, 1996, pp. 123–134.
- [16] Mesmer B., Bunke H., A New Algorithm for Error-Tolerant Subgraph Isomorphism Detection, *IEEE Transactions on PAMI*, vol. 20(5), 1998, pp. 493–504.
- [17] Baccini, A. and Caussin, H. and Ruiz-Gazen, A. Apprentissage Progressif en Analyse Discriminante *Rev. Statistique Appliquée*, XLIX (4), 2001.
- [18] Tabbone, S., Wendling, L. Technical Symbols Recognition Using the Two-dimensional Radon Transform, *ICPR'02, Québec, Canada*, 2002.
- [19] Kanungo, T., Haralick, R.M., Baird, H.S., Stuezle, W. and Madigan, D. A Statistical, Nonparametric Methodology for Document Degradation Model Validation. *IEEE Transactions on PAMI*, 22(11) :1209-1223, 2000.
- [20] Valveny, E., Dosch, P. Symbol Recognition Contest : A Synthesis, *GREC 2003, LNCS 3088*, 2004.