Proposition d'un système de reconnaissance structurelle probabiliste de caractères

Denis Arrivault, Noël Richard, Philippe Bouyer

Proposition d'un système de reconnaissance structurelle probabiliste de caractères

Denis Arrivault[1,2] – Noël Richard[1] – Philippe Bouyer[2]

[1] Laboratoire SIC
SP2MI, Boulevard Marie et Pierre Curie 86962 FUTUROSCOPE CEDEX

[2] RC-SOFT
Domaine de la Combe - BP39 16710 Saint-Yriex

denis.arrivault@sic.sp2mi.univ-poitiers.fr

Résumé : *La reconnaissance de caractères ou de symboles s'appuie actuellement sur de nombreux descripteurs statistiques ou géométriques. Cependant, l'accroissement du nombre d'objets différents à traiter pose le problème de l'interaction avec l'utilisateur et la capacité de celui-ci à caractériser cette complexité. La représentation structurelle est une alternative intéressante car plus proche de l'appréhension visuelle humaine. Elle offre une approche plus intuitive de la reconnaissance permettant d'aborder des caractères difficiles à discriminer par des approches classiques. Nous nous intéressons donc, dans cet article, aux systèmes de reconnaissance de la structure des caractères et des symboles. Le travail étend la proposition de* WONG ET YOU *d'une reconnaissance basée sur une modélisation par graphe aléatoire à variables discrètes. L'utilisation d'attributs vectoriels plus complets impose le passage aux graphes aléatoires à variables continues.*

Mots-clés : Reconnaissance structurelle, graphe d'attributs, graphe aléatoire.

1 Introduction

Partant de l'idée de HEBB ([HEB 49]) que toute forme est appréhendée par la perception visuelle humaine, non dans sa globalité, mais par parties, et que l'organisation et les positions spatiales relatives de ces parties jouent un rôle important dans l'apprentissage et la reconnaissance, les descriptions structurelles se basent sur une décomposition des formes en éléments simples. Appliquée à la reconnaissance de caractères, la recherche d'une entité « géométriquement simple » se comprend comme la volonté de ne pas faire entrer dans les structures utilisées des attributs géométriques trop complexes (comme les attributs statistiques par exemple). Puisque les descripteurs structurels reposent sur l'idée que les éléments simples sont interprétables par l'Homme, cette démarche est justifiée.

La reconnaissance de structures est un problème assez délicat qui devient vite très coûteux en temps de calcul. Par rapport aux méthodes statistiques numériques, la reconnaissance se fait en deux étapes. Il faut d'une part mesurer la similarité des structures et d'autre part mesurer la similarité des composantes primitives et ce bien souvent avec des formalismes différents. De nombreuses méthodes peuvent être dérivées suivant le type de structures utilisées et la nature des caractéristiques associées aux primitives. Les deux principales familles sont les approches *syntaxiques* qui se décrivent de manière *stochastique* avec les modèles markoviens et les approches à base de mise en correspondance de graphes ou approches *graphiques*. Ces dernières reposent sur des techniques *combinatoires* pour mesurer la ressemblance entre graphes. Pour donner plus de souplesse aux modèles et permettre la reconnaissance d'informations bruitées, des approches floues ([PER 02]) et bayésiennes ([CHR 95]) ont été dérivées. Utilisant des graphes d'attributs, le schéma de reconnaissance présenté dans cet article, se place dans la catégorie des méthodes graphiques.

Une chaîne de reconnaissance structurelle de caractères est typiquement constituée de trois étapes élémentaires. La première, appelée étape de décomposition, s'attache à extraire les segments primitifs des caractères. La seconde, appelée étape de modélisation, consiste à modéliser la structure du caractère à l'aide d'outils de description, les graphes d'attributs dans notre cas. Enfin la troisième et dernière étape dite d'interprétation va chercher à étiqueter une description inconnue par rapport aux descriptions d'une base d'apprentissage.

Nous présentons, dans cet article, un système de reconnaissance structurelle probabiliste de caractères. Il utilise une décomposition basée sur le squelette que nous exposons dans la première partie. La modélisation des caractères se fait à l'aide du formalisme des graphes d'attributs ([TSA 83]) suivant deux structurations au choix présentées dans la deuxième partie. L'interprétation utilise une généralisation du concept de graphe aléatoire ([WON 87]) et est détaillée dans la troisième partie. Enfin, nous terminons cet exposé par la présentation de quelques résultats obtenus sur différentes bases de caractères.

2 Etape de décomposition

En reconnaissance de caractères, les formes étant bien souvent minces, il est plus facile d'utiliser le squelette (ou axe médian) comme résumé interprétable de la courbe. Le découpage du squelette suivant certains points dits singuliers permet alors d'extraire les primitives du caractère.

[GON 92] [DEU 72] [ZHA 84] [ZHA 94]

FIGURE 1 – Squelettes obtenus à partir du caractère 'Beta' manuscrit.

2.1 Amincissement

Il existe globalement quatre approches algorithmiques pour le calcul de l'axe médian d'une forme binaire : les méthodes d'amincissement topologique, les méthodes par cartes de distances, les méthodes par évolution de courbe et les approches géométriques basées sur le diagramme de Voronoï. Les méthodes d'amincissement topologique sont rapides et donnent des squelettes minces. C'est pourquoi elles sont souvent privilégiées en reconnaissance de caractères.

Pour notre application, les contraintes de sélection ont été la construction d'un squelette mince, parfaitement homotopique (qui possède exactement le même nombre de composantes connexes et de trous que la forme initiale), positionné au centre de la forme et stable (n'érodant pas trop les extrémités). Après une étude comparative de différents algorithmes de la littérature (figure 1), notre choix s'est porté sur celui qui répondait le mieux aux critères précédents et qui, de plus, est strictement 8-connexe et rapide. Il s'agit de l'algorithme de ZHANG ET WANG ([ZHA 94]).

2.2 Extraction des segments primitifs

Une fois le squelette d'un caractère obtenu, il nous faut en extraire les segments primitifs. Un segment primitif est un élément « géométriquement le plus simple possible » constitutif d'un graphème. Le graphème étant maintenant schématisé par un squelette, une définition formelle du segment primitif peut être la suivante :

Définition 1 (Formalisme de la notion de segment primitif) *Nous appelons segment primitif tout ensemble de points strictement 8-connexes ayant au plus deux 8-voisins et formant une courbe de concavité stricte dans le plan image.*

L'objectif d'une telle définition est bien de permettre une description géométrique des segments primitifs avec des attributs simples comme la courbure ou la longueur. Or ces grandeurs pourraient être ambigües dès lors que la condition de concavité ne serait pas respectée.

A partir de la définition 1, le squelette devient un ensemble de segments primitifs connectés entre eux par des points d'intersection (dont le nombre de 8-voisins est supérieur à 2) ou des points d'inflexion. La décomposition d'un squelette en segments primitifs se résume à l'extraction des points de fin (ayant un seul 8-voisin), des points d'intersection et des points d'inflexion qui vont composer les extrémités des segments primitifs. Les points d'inflexion se calculent primitive par primitive suivant la méthode présentée dans [ARR 05].

3 Etape de Modélisation

Les graphes d'attributs ont été introduits en reconnaissance de formes par TSAI ET FU ([TSA 83]). Par définition, les nœuds des graphes représentent les primitives tandis que les arêtes expriment leurs relations. Chaque nœud du graphe prend ses attributs dans l'ensemble $Z = \{z_i | i = 1, ..., I\}$ où chaque attribut z_i va prendre ses valeurs dans l'ensemble $S_i = \{s_{ij} | j = 1, ..., J_i\}$. L'ensemble $L_v = \{(z_i, s_{ij}) | i = 1, ..., I; j = 1, ..., J_i\}$ est l'ensemble des couples de valeurs possibles pour les primitives. Une primitive valide est alors un sous-ensemble de L_v dans lequel chaque attribut ne peut apparaître qu'une seule fois. Si Π est l'ensemble de toutes ces primitives valides, chaque nœud est représenté par un élément de Π.

De manière similaire, pour les arêtes, l'ensemble des attributs possibles est appelé $F = \{f_i | i = 1, ..., I'\}$ dans lequel chaque f_i peut prendre ses valeurs dans $T_i = \{t_{ij} | j = 1, ..., J'_i\}$. L'ensemble $L_a = \{(f_i, t_{ij}) | i = 1, ..., I'; j = 1, ..., J'_i\}$ est l'ensemble des couples de valeurs possibles pour les relations. Une relation valide est alors un sous-ensemble de L_a dans lequel chaque attribut ne peut apparaître qu'une seule fois. L'ensemble de toutes les relations valides est noté Θ.

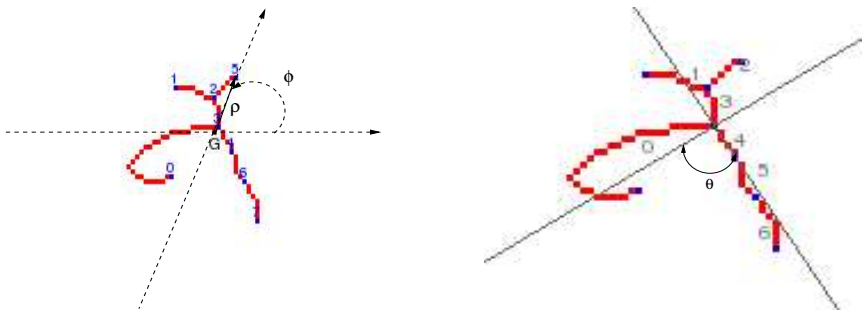
Un graphe d'attributs se définit formellement de la manière suivante :

Définition 2 (Graphe d'attributs ou GA) *un graphe d'attribut Ga sur $L = \{L_v, L_a\}$ avec une structure graphique $G = (N, A)$, est une paire ordonnée (V, E) où $V = (N, \sigma)$ est appelé un* ensemble de nœuds attribués *et $E = (A, \delta)$ est appelé un* ensemble d'arêtes attribuées. *Les applications $\sigma : N \rightarrow \Pi$ et $\delta : A \rightarrow \Theta$ sont appelées respectivement* interpréteurs *de nœuds et d'arêtes.*

Notre système utilise la structuration par graphe d'attributs avec deux simplifications de notation. En premier lieu, nos attributs sont continus à valeurs réelles sur $[0, 1]$ ce qui implique que $S_i = T_i = [0, 1]$. Ensuite, il est souvent intéressant d'exprimer toutes les valeurs des attributs d'un même élément par un vecteur. Nous noterons $\boldsymbol{\alpha}_a$ le vecteur des valeurs des attributs du nœud a et $\boldsymbol{\beta}_{ab}$ le vecteur des valeurs des attributs de l'arête (a, b). A partir de la décomposition du caractère en primitives, deux structurations à base de graphes d'attributs sont possibles.

3.1 Structure directe

La structuration que nous avons appelée directe, considère les primitives comme étant les points singuliers du squelette. Ainsi les nœuds des graphes représentent les points singuliers et les arêtes les segments primitifs qui relient ces



(a) (b)

FIGURE 2 – Attributs géométriques ; (a) liés aux nœuds des graphes directs - nœud 5 (ρ, ϕ) ; (b) liés aux arêtes des graphes d'arêtes - relation entre les nœuds 0 et 4 (θ)

points. Pour les nœuds, les attributs résument le positionnement spatial, pour les arêtes ils correspondent aux informations de forme des segments primitifs.

Les deux attributs associés aux nœuds sont les coordonnées polaires par rapport à un repère intrinsèque au caractère (notés ρ et ϕ, voir figure 2 (a)). Nous avons pris le centre de gravité du squelette comme centre du repère. Pour l'axe de référence angulaire, le choix s'est porté de façon empirique sur l'horizontale de l'image du caractère plutôt que sur l'axe principal ou de moindres carrés. L'erreur générée par ces derniers étant beaucoup plus dommageable à la reconnaissance que la perte de l'invariance par rotation qu'implique l'utilisation de l'axe horizontal. Le rayon et l'angle polaire sont normalisés respectivement par le rayon maximal sur l'image et 2π afin d'obtenir deux attributs à valeurs dans $[0, 1]$. Cette normalisation permet de rendre la description invariante par changement d'échelle.

Pour les relations, les attributs décrivent la longueur relative, lr, du segment et sa rectitude (« straightness » [PAL 86]), St, données pour l'arête reliant les nœuds a et b par :

$$lr_{ab} = \frac{\widehat{ab}}{\sum_{(c,d)\in A} \widehat{cd}} \qquad St_{ab} = \frac{ab}{\widehat{ab}} \qquad (1)$$

où $\widehat{ab}$ représente la longueur d'arc de l'arête (a, b) et ab désigne la longueur du segment droit reliant le point a et le point b.

La longueur relative comme la rectitude sont à valeurs dans $[0, 1]$. Ainsi, un nœud a est associé à un vecteur d'attributs $\boldsymbol{\alpha}_a = (\rho, \phi)^\perp$ et une arête (a, b) au vecteur $\boldsymbol{\beta}_{ab} = (lr_{ab}, St_{ab})^\perp$.

3.2 Structure d'arêtes

Dans les graphes que nous avons appelés d'arêtes, les nœuds représentent les relations des graphes directs et gardent donc les mêmes attributs à savoir lr et St. La notation, elle, change et $\boldsymbol{\alpha}_a = (lr_a, St_a)^\perp$. Les relations d'adjacence sont données par la présence ou l'abscence d'arête entre sommets, un attribut de positionnement angulaire relatif est associé à chaque arête. Le positionnement angulaire relatif (noté θ, figure 2 (b)) est établi par l'angle que forment les deux axes principaux des segments. Cet angle est non orienté et normalisé par π afin que l'attribut soit à valeur dans $[0, 1]$. L'avantage du choix de l'axe principal est qu'il varie moins que la droite des sommets ou de moindre carrés.

L'axe principal d'un segment primitif correspond au vecteur propre principal de l'ensemble de ses points. Ainsi pour une arête (a, b), $\boldsymbol{\beta}_{ab} = (\theta_{ab})$.

4 Etape d'interprétation

L'étape d'interprétation se fait en deux temps. Nous allons, dans un premier temps, modéliser chaque classe de la base d'apprentissage à l'aide d'une extension continue des graphes aléatoires. Cette modélisation se fait par appariement des tous les graphes d'attributs constitutifs d'une classe avec un modèle. Puis, dans un deuxième temps, nous effectuons une reconnaissance par appariement entre un graphe d'attributs inconnu et les graphes aléatoires de la base.

Le type d'appariement utilisé dans les deux phases de l'interprétation dépend du choix du modèle. La recherche de graphes « représentatifs » a été abordée par JIANG ET COLL. ([JIA 01]) qui proposent le graphe médian généralisé d'un ensemble de graphes. Il s'agit du graphe le plus proche de tous les autres graphes de la classe au sens d'une fonction de distance inter-graphe. Malheureusement cette distance passe par un appariement dans le cas de graphe d'attributs et la recherche du graphe médian généralisé nécessite l'appariement deux à deux de tous les graphes d'une classe. Cette recherche, en plus d'être extrêmement coûteuse pour les bases de grande taille, introduit une imprécision supplémentaire due aux erreurs d'appariement. Nous proposons d'utiliser plutôt la structure la moins contraignante pour chaque classe, c'est-à-dire le graphe de plus petit ordre (possédant le moins de nœuds). L'avantage de ce choix est qu'il est rapide et ne dépend pas de la qualité de l'algorithme d'appariement retenu. De plus, si la base a été correctement indexée, le graphe de plus petit ordre représente bien la structure la plus simple décrivant le caractère. Il convient alors de baser l'appariement sur la recherche d'un isomorphisme approximatif de sous-graphe entre cette structure et la structure inconnue.

4.1 Appariement

L'appariement réalisé par le système est de deux sortes :

– nous apparions les graphes d'attributs d'une classe avec le graphe modèle pour construire les graphes aléatoires dont la structure graphique est la même que celle

du graphe modèle. En retenant le graphe de plus petit ordre comme modèle de classe,
– nous apparions les graphes d'attributs inconnus avec les graphes aléatoires des classes de la base.

Dans les deux cas il s'agit de reconnaître tout ou partie du graphe modèle dans un graphe d'attribut donné. Nous recherchons donc un isomorphisme approximatif de sous-graphe. Pour les graphes non-valués, ce problème se résoud par la recherche de cliques maximale dans le graphe d'association. Pour les graphes d'attributs, il faut généraliser la notion de graphe d'association en lui associant un poids par élément correspondant à une mesure de ressemblance entre les deux éléments des graphes d'attributs qu'il représente. Plusieurs algorithmes permettent de trouver une solution locale à ce problème avec une complexité combinatoire raisonnable ([CHI 92], [GOL 96]). Nous avons retenu celui de GOLD ET RANGARAJAN ([GOL 96]) connu pour sa rapidité et sa précision. Le problème est alors de trouver une mesure de ressemblance entre les éléments de deux graphes d'attributs pour le premier type d'appariement et entre un élément d'un graphe d'attribut et celui d'un graphe aléatoire pour le deuxième type d'appariement.

4.2 Construction d'un graphe aléatoire

L'apprentissage consite à apparier tous les graphes d'attributs d'une même classe avec le graphe modèle. Pour réaliser l'appariement de deux graphes d'attributs, il nous faut définir une mesure de similarité entre les nœuds et les arêtes prenant en compte la similarité des attributs. D'une manière générale, toute similarité définie dans un cadre statistique entre vecteurs numériques est une candidate valable. Comme tous les attributs sont à valeurs dans $[0, 1]$, une similarité basée sur une distance de Minkowski sans pondération (ou « city block ») permet de mettre en évidence les faibles différences :

$$s(\mathbf{x}, \mathbf{y}) = 1 - \sum_{i}^{n} \frac{|\mathbf{x}(i) - \mathbf{y}(i)|}{n} \tag{2}$$

où $\mathbf{x}$ et $\mathbf{y}$ représentent les vecteurs des valeurs des attributs à comparer ($\boldsymbol{\alpha}_i$ ou $\boldsymbol{\beta}_{ij}$), $n = I$ ou I' selon le type d'élément considéré. Pour la comparaison d'attributs périodiques comme c'est le cas avec les mesures angulaires, il faut utiliser une différence périodique définie de la façon suivante :

$$\Delta(x, y) = \begin{cases} x - y & \text{si } |x - y| \leq \frac{T}{2} \\ x - y - T \times \text{sign}(x - y) & \text{si } |x - y| > \frac{T}{2} \end{cases} \tag{3}$$

où (x, y) sont deux réels périodiques de période T et $\text{sign}(x-y)$ le signe de leur différence. Dans notre cas $T = 1$.

Une fois tous les graphes d'une classe appariés avec le graphe modèle, chaque classe est représentée par un graphe de structure identique au graphe modèle et dont les éléments sont associés à une série d'observations, attributs des graphes de la classe.

Il est possible d'associer à chaque élément du graphe modèle d'une classe une loi de distribution de probabilités multivariée. Chaque classe est alors représentée par un graphe dont le formalisme est analogue à celui des graphes aléatoires définis par WONG ET YOU ([WON 85], [WON 87]) où les éléments (nœuds et arêtes) sont des variables aléatoires. L'adoption de ce formalisme nous permet d'utiliser certains résultats de ces articles pour aborder la reconnaissance.

4.3 Reconnaissance

La principale différence entre la structure proposée par WONG ET YOU et l'extension proposée ici est le passage des variables aléatoires discrètes dont les probabilités sont calculées de manière fréquentielle à des variables aléatoires continues. A chaque variable aléatoire du graphe est associée une loi de probabilité estimée à partir des observations de la base d'apprentissage. Les lois sont supposées empiriquement gaussiennes diagonales et les espérances comme les variances sont estimées statistiquement.

Ainsi, les nœuds α_i des graphes d'attributs modèles sont tels que $\alpha_i \rightsquigarrow \mathcal{N}(\boldsymbol{\mu}_{\alpha_i}, \Sigma_{\alpha_i})$ et les arêtes $\beta_i = (\alpha_l, \alpha_m)$ sont telles que $\beta_i \rightsquigarrow \mathcal{N}(\boldsymbol{\mu}_{\beta_i}, \Sigma_{\beta_i})$.

Soit $\{R_k = (W_k, B_k) | k = 1 \dots c\}$ les graphes aléatoires modèles des classes C_k de la base d'apprentissage et $Ga = (N, A, \sigma, \delta)$ un graphe d'attributs inconnu à reconnaître. La reconnaissance se fait en deux étapes. La première consiste à trouver un isomorphisme approximatif de sous-graphe entre le graphe inconnu et chacun des graphes aléatoire. La seconde étape est celle de la décision.

4.3.1 Appariement

Avec le même formalisme que pour l'étape d'interprétation, la recherche de l'isomorphisme approximatif de sous-graphe passe par la construction d'un graphe d'association généralisé entre le graphe d'attributs et chacun des graphes aléatoires. Les poids associés aux éléments du graphe d'association généralisé peuvent être calculés de plusieurs manières. L'approche la plus simple, utilisée pour les résultats présentés dans cet article, est d'apparier G_a avec les graphes d'attributs moyens issus des R_k qui possèdent la structure graphique des modèles et dont les attributs sont les moyennes des observations. Ignorer les variances des distributions présente l'avantage de minimiser l'influence des approximations d'appariement. Les variances sont ensuite réintégrées dans la phase de décision.

4.3.2 Décision

La phase de décision utilise un résultat donné par WONG ET YOU sur les graphes aléatoires. Il repose sur trois hypothèses :

1. les nœuds aléatoires d'un graphe aléatoire sont mutuellement indépendants ;
2. les arêtes aléatoires d'un graphe aléatoire sont indépendantes de tous les nœuds du graphe exceptés les deux qui forment ses extrémités ;
3. les distributions conditionnelles des arêtes aléatoires relatives aux nœuds aléatoires d'un graphe aléatoire sont mutuellement indépendantes.

Sous ces hypothèses, WONG ET YOU expriment la probabilité qu'un graphe $G = (N, A)$ soit la réalisation d'un graphe aléatoire $R = (W, B)$ par l'isomorphisme de sous-

graphe ϕ de la façon suivante :

$$p(G,\phi) = \prod_{\alpha_i \in W} Pr\{\alpha_i = \phi^{-1}(\alpha_i)\} \times \prod_{\substack{\beta_i=(\alpha_l,\alpha_m)\in B \\ \alpha_l = \phi^{-1}(\alpha_l),\ \alpha_m = \phi^{-1}(\alpha_m)\}}} Pr\left\{\beta = \phi^{-1}(\beta))\right| \quad (4)$$

Pour un graphe aléatoire à variables continues, cette probabilité n'a pas de sens. Nous proposons d'utiliser la probabilité de dispersion de α_i par rapport à a basée sur la distance de Mahalanobis. Nous rappelons sa définition :

Définition 3 Probabilité de dispersion *Soit X une variable aléatoire sur $\mathbb{R}^2$ de loi de distribution normale $\mathcal{N}(\boldsymbol{\mu}, \Sigma)$. Soit $\boldsymbol{\alpha}$ un point de $\mathbb{R}^2$. La probabilité de dispersion de X par rapport à $\boldsymbol{\alpha}$ est la probabilité qu'une observation $\boldsymbol{y}$ de X soit située en dehors de l'ellipse de dispersion de niveau p_Δ où Δ est la distance de Mahalanobis entre $\boldsymbol{\mu}$ et $\boldsymbol{\alpha}$ ($\Delta = d_M(\boldsymbol{\alpha}, \boldsymbol{\mu})$). Elle est donnée par :*

$$P(d_M(\boldsymbol{y}, \boldsymbol{\mu}) \geq \Delta) = 1 - p_\Delta = e^{\frac{-\Delta}{2}} \quad (5)$$

La demonstration de l'équation 5 en dimension 2 est donnée dans la thèse de DANG ([DAN 98]). En dimension 1 le résultat est classique et correspond aux tables de normalité. Nos attributs étant de dimension 1 ou 2, nous pouvons utiliser ce résultat et écrire :

$$P(d_M(\boldsymbol{y}, \boldsymbol{\mu}_{\alpha_i}) \geq P(d_M(\boldsymbol{\alpha}_a, \boldsymbol{\mu}_{\alpha_i}))) = \exp\left(-\tfrac{1}{2}(\boldsymbol{\alpha}_a - \boldsymbol{\mu}_{\alpha_i})^\perp {\Sigma_{\alpha_i}}^{-1}(\boldsymbol{\alpha}_a - \boldsymbol{\mu}_{\alpha_i})\right) \quad (6)$$

En pondérant cette probabilité par le rapport π_{α_i} entre le nombre d'observations associées au nœud α_i et le nombre total d'observations associées à R_k (égal au nombre de caractères de la classe dans la base), nous obtenons une mesure homogène à la probabilité discrète utilisée par WONG ET YOU que a soit une réalisation de α_i ($Pr(\alpha_i = \boldsymbol{\alpha}_a)$). Pour rester dans le formalisme WONG ET YOU, nous avons conservé cette notation :

$$Pr(\alpha_i = \boldsymbol{\alpha}_a) = \pi_{\alpha_i} \times \exp\left(-\tfrac{1}{2}(\boldsymbol{\alpha}_a - \boldsymbol{\mu}_{\alpha_i})^\perp {\Sigma_{\alpha_i}}^{-1}(\boldsymbol{\alpha}_a - \boldsymbol{\mu}_{\alpha_i})\right) \quad (7)$$

De façon similaire, si $\beta_i = (\alpha_l, \alpha_m)$ est une arête de R_k et (a, b) une arête de G_a (représentée par son vecteur d'attribut $\boldsymbol{\beta}_{ab}$) , la probabilité $Pr(\beta_i = (\boldsymbol{\beta}_{ab}))$ s'écrit :

$$Pr(\beta_i = \boldsymbol{\beta}_{ab}) = \pi_{\beta_i} \times \exp\left(-\tfrac{1}{2}(\boldsymbol{\beta}_{ab} - \boldsymbol{\mu}_{\beta_i})^\perp {\Sigma_{\beta_i}}^{-1}(\boldsymbol{\beta}_{ab} - \boldsymbol{\mu}_{\beta_i})\right) \quad (8)$$

avec π_{β_i} représentant le nombre d'observations associées à l'arête β_i sur le nombre total d'observations associées à R_k (égal au nombre de caractères de la classe dans la base).

Ainsi, par analogie avec le formalisme de WONG ET YOU, la probabilité que $Ga = (N, A, \sigma, \delta)$ soit une réalisation de $R_k = (W_k, B_k)$ par ϕ s'exprime :

$$p(Ga,\phi) = \prod_{\alpha_i \in W} Pr\{\alpha_i = \phi^{-1}(\alpha_i)\} \times \prod_{\substack{\beta_i=(\alpha_l,\alpha_m)\in B \\ \alpha_l = \phi^{-1}(\alpha_l),\ \alpha_m = \phi^{-1}(\alpha_m)\}}} Pr\left\{\beta = \phi^{-1}(\beta))\right| \quad (9)$$

Le problème de cette formulation est qu'elle ne tient pas compte des nœuds et arêtes de Ga qui n'ont pas d'image par ϕ^{-1}. WONG ET YOU considèrent, à juste titre, ces cas comme impossibles dans leur formalisme et attribuent alors une probabilité nulle à la réalisation Ga. Dans notre cas, le choix des graphes modèles nous oblige à en tenir compte puisque l'appariement est basé sur la recherche d'un sous-graphe. Il nous faut donc pondérer la probabilité par le nombre d'éléments de Ga non appariés. Plus la structure inconnu est différentes en ordre de la structure modèle et plus la ressemblance doit tendre vers 0. Nous avons donc multiplié empiriquement $p(Ga, \phi)$ par un facteur égal à $0, 1^{Na}$ où Na est le nombre d'éléments non appariés.

Une fois $p(Ga, \phi)$ calculée pour chaque R_k, la décision est prise par la règle du maximum de vraisemblance :

$$Ga \in C_i \Leftrightarrow i = \arg\max_k(\pi_k \times p(Ga, \phi)) \quad (10)$$

où $\pi_k = \frac{m_k}{m}$ est le rapport du nombre de caractères de la classe k sur le nombre total d'éléments dans la base.

5 Résultats

Les résultats présentés ici ont été obtenus avec une base de caractères grecs manuscrits possèdant 47 classes pour un total de 18 236 caractères. Les caractères ont été écrits sur des pages blanches avec un feutre noir par quatre scripteurs différents. Quatre milles caractères ont été tirés au hasard dans la base pour former les bases de tests. Ces 4 000 caractères de test ont été divisés en quatre groupes de 1 000 individus. Pour chaque groupe, un taux d'erreur de classification a été calculé par rapport à une base d'apprentissage composée des 3 000 individus des trois autres groupes de test plus le reste de la base. La table 1 donne la moyenne des erreurs sur les quatre groupes ainsi que l'écart type et ce pour les deux types de structures présentées (graphe direct ou d'arêtes). Toutes les phases d'appariement ont été réalisées avec l'algorithme de GOLD ET RANGARAJAN ([GOL 96]).

Type de graphe	direct	d'arêtes
Erreur Moyenne	47,4%	63,8%
Ecart Type	0,9%	1,9%

TABLE 1 – Résultats de reconnaissance

Ces résultats sont bruts et n'utilisent aucune optimisation sur les paramètres d'appariement (nous avons utilisé les paramètres donnés à titre d'exemple dans l'article de GOLD ET RANGARAJAN). Après analyse, il s'avère que la principale limitation de la méthode réside dans le choix du modèle gaussien pour les lois de probabilités des variables aléatoires. Le modèle n'est que rarement valide. Celà est imputable aux erreurs d'appariement induites par le choix d'une solution locale par l'algorithme.

D'une manière générale, et celà est valable sur quasiment toutes les bases testées, les structures directes donnent de meilleurs résultats que les structures d'arêtes. Là encore les erreurs proviennent de l'appariement, la structure d'arêtes y est plus sensible que la structure directe.

6 Conclusion

L'objectif de ce travail était de proposer un système complet de reconnaissance de caractère avec apprentissage basé sur une description structurelle. Nous avons présenté les choix scientifiques opérés sur différentes étapes crutiales dans un tel système : la squelettisation, la structuration, les mesures de comparaison ainsi que le processus d'apprentissage. Le dernier point délicat à résoudre reste l'appariement. Le problème n'est pas dans le choix d'un algorithme, la littérature en propose de très bons et bien adaptés, mais dans l'adaptation des structures à ce type de traitement. Il nous paraît, en effet, essentiel de reprendre cette étape en hiérarchisant son déroulement et en transformant itérativement, par exemple, les graphes à l'aide d'opérations simples.

Les résultats, bien que médiocres, permettent d'ores et déjà de souligner l'avantage des structures directes par rapport aux structures d'arêtes. De plus les informations apportées par ce type de structures sont très différentes des informations statistiques classiques et permettent des coopérations intéressantes ([ARR 04]).

Références

[ARR 04] ARRIVAULT D., RICHARD N., FERNANDEZ-MALOIGNE C., , , BOUYER P., Collaboration entre approches statistique et structurelle pour la reconnaissance de caractères anciens, *8ème Colloque International Francophone sur l'Ecrit et le Document, La Rochelle, France*, 2004, pp. 197-202.

[ARR 05] ARRIVAULT D., RICHARD N., , BOUYER P., A Fuzzy Hierarchical Attributed Graph Approach for Handwritten Hieroglyphs Description, *11th International Conference on Computer Analysis of Images and Patterns, Versailles, France*, 2005, page 748.

[CHI 92] CHIPMAN L., RANGANATH H., A Fuzzy Relaxation Algorithm for Matching Imperfectly Segmented Images to Models, *IEEE Southeastcon'92*, vol. 1, Avril 1992, pp. 128-136.

[CHR 95] CHRISTMAS W., KITTLER J., , PETROU M., Structural Matching in Computer Vision Using Probabilistic Relaxation, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 17, n° 8, 1995, pp. 749-764.

[DAN 98] DANG V., Classification de données spatiales : modèles probabilistes et critères de partitionnement, PhD thesis, Université Technologique de Compiègne, 1998.

[DEU 72] DEUTSCH E., Thinning Algorithms on Rectangular, Hexagonal, and Triangular Arrays, *Communications of the ACM*, vol. 15, n° 9, 1972, pp. 827-837.

[GOL 96] GOLD S., RANGARAJAN A., A Graduated Assignment Algorithm for Graph Matching, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 18, n° 4, 1996, pp. 377-388.

[GON 92] GONZALEZ R., WOODS R., *Digital Image Processing*, Addison Wesley, 1992.

[HEB 49] HEBB D., *The Organization of Behavior : A Neuropsychological Theory*, Wiley, 1949.

[JIA 01] JIANG X., MUNGER A., , BUNKE H., On Median Graphs : Properties, Algorithms, and Applications, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 23, n° 10, 2001, pp. 1144-1151.

[PAL 86] PAL S., DUTTA-MAJUMDER D., *Fuzzy mathematical approach to pattern recognition*, Halsted Press, New York, NY, USA, 1986.

[PER 02] PERCHANT A., BLOCH I., Fuzzy morphisms between graphs, *Fuzzy Sets and Systems*, vol. 128, 2002, pp. 149-168.

[TSA 83] TSAI W., FU K., Subgraph Error-Correcting Isomorphisms for Syntatic Pattern Recognition, *IEEE Trans. Systems, Man and Cybernetics*, vol. 13, n° 1, 1983, pp. 48-62.

[WON 85] WONG A., YOU M., Entropy and Distance of Random Graphs with Application to Structural pattern Recognition, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 7, n° 5, 1985, pp. 599-609.

[WON 87] WONG A., Structural Pattern Recognition : A Random Graph Approach, *Pattern Recognition Theory and Application*, vol. 87, 1987, pp. 323-345.

[ZHA 84] ZHANG T., SUEN C., A Fast Parallel Algorithm for Thinning Digital Patterns, *Communications of the ACM*, vol. 27, n° 3, 1984, pp. 236-240.

[ZHA 94] ZHANG Y., WANG P., A New Parallel Thinning Methodology, *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 8, 1994, pp. 999-1011.