



RIENNE DÉMOCRATIQUE ET POPULAIRE
ENT SUPÉRIEUR ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITÉ LARBI BEN MØHIDI Ó OUM ELBOUAGHI
INSTITUT DES SCIENCES EXACTES

RECONNAISSANCE DES FORMES APPLIQUÉE AUX CARACTÈRES ARABES

Mémoire présenté par

Chergui Leila

Pour l'obtention du

DIPLOME DE MAGISTÈRE EN INFORMATIQUE

Option:

INTELLIGENCE ARTIFICIELLE

Devant la commission d'examen :

Mr. Mohamed Benmohamed (Rapporteur)
Professeur – Université de Constantine
Mr. Azzedine Bilami (Président de jury)
Maître de conférence ó Université de Batna
Mr. Salim Chikhi (Examineur)
Maître de conférence ó Université de Constantine
Mr. Allaoua Chaoui (Examineur)
Maître de conférence ó Université de Constantine

N° d'ordre :

Série :

Décembre 2008



*Your complimentary
use period has ended.
Thank you for using
PDF Complete.*

[Click Here to upgrade to
Unlimited Pages and Expanded Features](#)

RECONNAISSANCE DES FORMES APPLIQUÉE AUX CARACTÈRES ARABES

 *Your complimentary use period has ended. Thank you for using PDF Complete.*

[Click Here to upgrade to Unlimited Pages and Expanded Features](#)

LISTE DES ABRÉVIATIONS

- ART** : Théorie de la Résonance Adaptative (Adaptative Resonance Theory).
- HMM** : Chaîne de Markov Cachées (Hidden Markov Models).
- K-ppv** : K-plus proches voisins.
- NSHMM** : Chaîne de Markov Cachées Non Stationnaire (No-Stationary Hidden Markov Models).
- OCR** : Système de Reconnaissance d'Écriture (Optical Character Recognition).
- PHMM** : Chaîne de Markov Cachées Planaires (Planar Hidden Markov Models).
- PMC** : Perceptron Multi Couches.
- RF** : Reconnaissance de Formes.
- SVM** : Machine à Vecteur du Support (Separator Vaste Marge).
- TH** : Transformée de Hough.
- THG** : Transformée de Hough généralisée.

2 : ALGORITHME D'APPRENTISSAGE DES PMC_s

- 1- Initialiser l'ensemble des vecteurs de poids synaptiques W_{ij} à des valeurs aléatoires.
- 2- Appliquer un exemple d'apprentissage X_k à l'entrée du Perceptron Multi Couches et évaluer l'erreur quadratique commise par le réseau sur cet exemple :

$$E^{(k)} = \sum_{i=1}^{h_L} (y_{Li}^{(k)} - t_i^{(k)})^2 .$$

- 3- Corriger l'ensemble des poids synaptiques selon :

$$w_{L_{ij}}(\tau + 1) = w_{L_{ij}}(\tau) - \alpha \frac{\partial E^{(k)}}{\partial w_{L_{ij}}} .$$

Où α est le taux d'apprentissage, $\alpha > 0$, et où :

$$\frac{\partial E^{(k)}}{\partial w_{L_{ij}}} = \delta_{L_i}^{(k)} y_{l-L_j}^{(k)} .$$

Avec, un neurone de la couche de sortie :

$$\delta_{L,i}^{(k)} = \phi_{L,i} (W_{L,i}^T Y_{L-l}^{(k)}) (y_{L,i}^{(k)} - t_i^{(k)}) .$$

Et pour un neurone de la couche cachée :

$$\delta_{lj}^{(k)} = \phi_{lj} (W_{lj}^T Y_{L-l}^{(k)}) \sum_{q=1}^{h_{l+1}} w_{l+Lq} \delta_{l+Lq}^{(k)} .$$

Lorsque la fonction d'activation des neurones est la fonction sigmoïde, il vient

- pour un neurone de la couche de sortie :

$$\delta_{L_j}^{(k)} = y_{L_j}^{(k)} (1 - y_{L_j}^{(k)}) (y_{L_i}^{(k)} - t_i^{(k)}) .$$

- Et pour un neurone de la couche cachée :

$$\delta_{Li}^{(k)} = y_{Li}^{(k)} (1 - y_{Li}^{(k)}) \sum_{q=1}^{h_{l+1}} w_{l+Lqi} \delta_{l-Lq}^{(k)} .$$

- 4- Incrémenter τ et k, et reprendre de point 2. [44]

ALGORITHME DE RECHERCHE DES K-PPVs

1. Déterminer le plus petit hyperparallélépipède H délimitant l'espace contenant l'échantillon d'apprentissage.

$H = m$ paires de coordonnées (\min_i, \max_i) , $i = 1, 2, \dots, m$.

L'intervalle $[\min_i, \max_i]$ est découpé en λ_i intervalles adjacents formant une partition.

2. Calculer pour chaque cellule :

$C_x^0 =$ Cellule contenant l'observation \underline{x} .

$C_x^i (i > 0) =$ Domaine formé par la réunion de C_x^{i-1} et des autres cellules qui lui sont contiguës.

C_x^i contient $(2i + 1)^m$ cellules.

3. Déterminer $C_x^0; i \leftarrow 0$.

4. S'il y a une autre observation que \underline{x} dans C_x^i alors

Rechercher \underline{y}_0 le plus voisin de \underline{x} dans C_x^i .

Si $d(\underline{x}, \underline{y}_0) < d(\underline{x}, C_x^i)$ alors fin.

Sinon $i \leftarrow i + 1$, et aller à 4 (on cherche le PPV à l'extérieur de C_x^i).

5. S'il n'y a pas d'autres observations alors $i \leftarrow i + 1$ et aller à 4 (on élargit la zone de recherche). [62]

MOTS DE LA BASE IFN/ENIT

القبروان	سببلة	أريانة
القبروان الجنوبية	السخيرات	برج التومي
القبروان عقبة	سيدي إبراهيم الزهار	برج الطويل
الكبارة	سيدي سهيل	برج العامري
الكرمة	الشرايع	برج الوزير
المتبسطة	صحراوي	البساتين
المساعد	العجرد	البطان
مسيوتة	عين الخمايسية	التضامن 2
منزل المهيري	عين الحمادنة	تونس قرطاج
نصر الله	عين الدفلة	الجديدة
هندي زيتونة	العيون	الحبيبية
الوسلاتية	فريانة	حي التضامن
برنوسة	الفكة	حي الغزالة
بهرة	فوسانة	الدخيلة
بوجابر	قرعة الناظور	الذندان
تاجروين	قرعة النعام	دوار هيشر
تل الغزلان	القرين	رواد
الجريصة	القصرين	رياض الأندلس
جزة	القصرين الزهور	سبالة بن عمار
الدهماني	القصرين النور	السعيدة
زعفران	الكامور	سكرة
زيتونة	لحواش	سيدي ثابت
الساقية	ماجل بلعباس	سيدي علي الحطاب
السرس	هنشير العسل	شرفش
سيدي أحمد الصالح	وادي الميو	شواط
سيدي خيار	أولاد فرج الله	شويقي
سيدي مطير	بئر أحمد	طبرية
الطويرف	بئر الدولة	الفجة
عين الكرمة	الباطن	قلعة الأندلس
عين سنان	البرجي	قنطرة بنزرت
القصور	بن سالم	المرناقية 20 مارس
القلعة الخضباء	بوحجلة	المرناقية
قلعة سنان	جهينة	المسعدين
كاف ال	حاجب العيون	منوبة
الكاف الغربي	حفوز	المنزه 6
اللاس	حي ابن الجزائر	المنزه 8
المحاسن الشمالية	حي الحجام	المنصورة
محجوية	حي النصر	المنيهلة
منزل سالم	الخطراء	واد الليل
نبر	خيطة الواد	بزوكة
وادي السواني	الذريعات	بودرياس
وادي ملاق	رقادة	بوزقام

بن
بنبله
بني حسان
البهيرة
بوحجر
بوذر
بوعثمان
جمال
جمال خير الدين
حي البساتين
خنيص
زاوية قنطش
زرمدين
الساحلين
سكرين
سيدي بنور
سيدي عامر
شراحيل
الشرف
صيادة
طبلية
طوزة
الطيايرة
عميرة التوازره
عميرة الحجاج
عميرة الفحول
عميرة حاتم
غنلده
قصر هلال
قصر هلال الرياض
قصيبة المديوني
لمطة
المزاوغة
مزدور
مسجد عيسى
المكنين
المكنين الجديدة
مليشات
المنارة
منزل النور
منزل حرب
منزل حياة
منزل خير
منزل فارسي
منزل كامل
المنستير
المنستير الجمهورية
المنستير المحطة

الرقوبة الحمراء
زعفرانة
السييخة
سيدي سعد
سيدي مسعود
سيسب
الشيبيكة
الشراردي
عبيدة الشرقية
العلاء
العلم
العوامرية
عين الخزازية
عين بومرة
عين جلولة
غابات ال
قصر اللمسة
الغضابنة
قصور الساف
قصور الساق حشاد
كركر
كساسة
المحارزة 18
ملولش
منزل حشاد
المنصورة السواسي
المهدية
المهدية الجمهورية
المهدية الزهراء
المهدية هيبون
النفاتية
هبيرة
وادي باجة
باجة
تيرسق
تستور
توكابر
تيبار
حمام سيالة
دجبة
دقة
زهرة مدين
السخيرة
السلوقية
سيدي اسماعيل
سيدي فرج
سيدي مدين
سيدي نصر
طباية
العيادة

بوشبكة
بولحناش
بولعابة
تالة
تالة الجنوبية
تلايت
جدليان
حاسي الفريد
حي السرور
حيدرة
خمودة
خنفة الجازية
الذشرة
الرحايات
زاوية بن عقار
زلغان
سببية
الرسالة
الزهراء
الزهراء الحبيب
فوشانة
المحمدية
المدينة الجديدة
مرناق
المروج
مقرين
مقرين الرياض
مقرين شاكرا
نعسان
الياسمينات
أوتيك
أوتيك الجديدة
بازينة
باش حانبة
برج الشلوف
بنزرت
بنزرت باب ماطر
بنزرت بوقطفة
بنزرت حشاد
بني عطا
بوزراعة
تسكراية
تمرة
تينجة
جرزونة
الجواودة
جومين
الختمين
راس الجبل
رفراف

أولاد الشامخ
 أولاد جاب الله
 البرادعة
 بومرداس
 بو هلال الجنوبية
 التلالسة
 الجم
 الجماملة
 الحكايمة
 الرحيش
 الرشراشة
 رياض بو هلال
 زردة
 زلبة
 ساقية الخادم
 سلقطة
 السواسي
 سيدي زيد
 سيدي عساکر
 سيدي علوان
 الشابة
 الشحيمات
 شربان
 شهدة
 شبية
 توزر
 توزر المطار
 توزر شقراطسي
 حامة الجريد
 حزوة
 دغومس
 دقاش
 رأس الزراع
 سبعة آبار
 سبيكة الجريد
 شتاوة صحراوي
 الضفوية
 المحاسن
 نفطة
 بئر عمير
 بئر الأحمر
 بئر ثلاثين
 بني بركة
 بني مهيرة
 تطاوين
 تطاوين 7 نوفمبر
 تلالت

عين تونقة
 الغرم
 فدان السوق
 قبلاط
 القريعات
 القصر
 مجاز الباب
 المزارة
 المعقولة
 نفزة
 هاري ال
 واد الزرقة
 وشتاتة
 بئر الباي
 بئر القصعة
 برج السدرية
 بن عروس
 بن عروس الجنوبية
 بومهل
 بوقرنين
 جبل الرصاص
 حمام الأنف
 حمام الشط
 حي بوصفارة
 الخليدية
 رادس
 رادس المدينة
 رادس الملاحه
 تونس باب البنات
 تونس باب السويقة
 تونس تأمر
 تونس جباري
 تونس حشاد
 تونس سوق الجملة
 حلق الوادي
 حي ابن خلدون
 حي الإنطلاقة
 حي التحرير
 حي الخطراء
 حي الزهور
 حي المهرجان
 حي المهيري
 حي هلال
 خزندار
 الزرهوني
 الزرهوني فلة
 سيدي بوسعيد
 سيدي حسين
 سيدي داود
 سيدي فتح الله

رفراف الشاطئ
 الزواوين
 سجنان
 سونين
 سيدي عثمان
 سيدي علي الشباب
 العالية
 العزيز
 العوسجة
 عين غلال
 غار الملح
 غزالة
 القتمة
 القرية رأس الجبل
 اللواتة
 ماطر
 ماطر حشاد
 مثلين
 مثلين
 المصيدة
 منزل بورقيبة
 منزل بورقيبة النجاح
 منزل جميل
 منزل عبد الرحمان
 بلاد الحضرة
 بن فرج الله
 بو هلال
 تمغزة
 سوق الجمعة
 سوق السبت
 سيدي مسكين
 طبرقة
 عين النية
 عين الصباح
 عين دراهم
 غار الدماء
 فرنانة
 المرجي
 الملقى
 هذيل
 وادي المعادن
 وادي مليز
 ورغش
 بئر الشاوش
 بئر حليلة
 بئر مشاركة
 بئر مشاركة المحطة
 بنت سعيدان
 جبل الوسط
 جرادو

الزيرك
ذهبية
رأس الوادي
الرقبة
رمادة
الزهراء تطاوين
السعادة
شنني
شنني الجديدة
الصارم
غمار سن
الفرش
قرماسة
قصر أولاد سلطان
قصر الحدادة
قصر المرابطين
قصر المقابلة
قصر دباب
قصر عون
قطوفة
كرشاو
كمبوط
مزطورية
نكريف
وادي الخيل
وادي الغار
إبن سناء
باب سعدون المحطة
باردو
تونس البلفير
تونس الجمهورية
تونس الحفصية
تونس الشابي
تونس القباضة الأصلية
تونس المطار
تونس باب الخضراء
تونس باب المنارة
القلعة الكبرى
القنطاوي
الكنائس
كندار
مساكن
مساكن القبليّة
مساكن حي الجديد
المسعدين
منزل المحطة
منزل بالواعر
مور الدين

سلامبو
ال عمران
ال عمران الأعلى
قرطاج
قرطاج ياسمينية
قصر السعيد
الكبارية
الكرم
الكرم الغربي
مرسى الرياض
مرسى الصفصاف
المروج 2
المنار 2
المنزه 9
المنزه
الوردية
أولاد هلال
ببوش
بدرونة
بريرم
بلارجيا
بلطة
بن بشير
بني مطير
بوسالم
بوعوان
التباينية
جاب الله
جنتورة
جندوبة
جندوبة الشمالية
حكيم
حمام بورقوية
الحمدية
حي التطور
الدخايلية
السعادة
السمران
السنليل
الذويبات
الرضاع
الرقاب
رميلة
السبالة
السعيدة
السلامة
السوق الجديد
سيدي بوزيد
العبادات
الفايض

حمام الزربية
ذراع بن جودر
الزربية
زغوان
سمنجة
سيدي عويدات
صواف
عين البطوم
عين العسكر
الفحص
مقرن
الناصور
وادي الزيت
أكودة
البرجين
بني كلثوم
بوفيشة
حمام سوسة
خزامة الغربية
زاوية سوسة
السد الجوفي
سوسة
سوسة إبن خلدون
سوسة الرياض
سوسة الزهور
سوسة خزامة
سيدي الهاني
سيدي بوعلي
سيدي خليفة
شط مريم
الشقرانية
شباب
طرود سوسة
عين الرحمة
عين قارسي
عين مذاكر
قروسية ال
قصيبة سوسة
القلعة الصغرى
شعال
صبيح
صفاقس
صفاقس الجديدة
صفاقس المغرب العربي
صفاقس الميناء
صفاقس حشاد
العامرة
عجنقة
العطابة
عقارب

الأقباب
 برج المسعودي
 برقو
 بوجليدة
 بورويس
 بورويس فلاح
 بوسعدية
 بوعرادة
 الحبابسة الجنوبية
 حمام بياضة
 الخلصة
 الدخانية
 رأس الماء
 الروحية
 السفينة
 سليمانة
 سند حداد
 سيدي حمادة
 سيدي سعيد
 سيدي عيلد
 صدقة
 العروسة
 عين زريق
 القرية
 قعفرور
 القنطرة
 الكريب
 الكريب محطة
 كسرى
 كسرى العليا
 مكثر
 المنصورة
 حي الصلاح
 حمام سوسة الغرابي
 أم العظام
 أولاد حفوز
 أولاد منصر
 بئر الحفي
 برج الكرمة
 بن عون
 البوغة
 جلمة
 الحوامد
 حي أولاد بلهادي
 حي القوافل
 الخرشف
 الخشم

الفوني
 قيرار
 قصر المزاراة
 القلال
 لسودة
 الماش
 مبروكة
 مرفق الفريو
 المزونة
 المكارم
 المكناسي
 منزل بوزيان
 النصر
 النوايل
 الهيشرية
 أم شوشة
 أولاد بوسمير
 أولاد قاسم
 بئر صالح
 بئر علي بن خليفة
 البستان
 بليانة
 بوثدي
 بوجربوع
 تينة
 تينة الجديدة
 جبارنة
 الجبنيانة
 حاجب
 حزق
 الحنشة
 حي البحري
 حي الحبيب
 حي الخيري
 حي بورقية
 الخزانات
 الخليج
 دوالر اللواتة
 ذراع بن زياد
 رباع سيدي ظاهر
 ساقية الداير
 ساقية الزيت
 السخيرات
 السعدي
 السلام
 السلطنية
 سيدي حسن
 سيدي صالح
 سيدي عباس
 سيدي منصور

العوابد
 العين
 الغرابية
 غريبة
 قايد إمحمد
 القراطن
 قرقنة
 قرقور
 القندول
 اللوزة
 المحرص
 مركز العلوي
 مركز السبعي
 مركز الشحيحة
 مركز الصغار
 مركز العالية
 مركز اللجمي
 مركز بو عصيدة
 مركز درويش
 مركز سحنون
 مركز شاكرا
 مركز قصاص
 مركز كمون
 مركز والي
 مليئة
 منزل الهادي شاكرا
 النجاة
 نكتة
 النيقرو
 وادي الرمل
 وادي الشعبوني
 بشيمة
 البطاهرة
 بني زلطن
 بني غيلوف
 بوشمة
 بوعطوش
 تاشين
 تمزرت
 توجان
 الحامة
 الحامة الجنوبية
 حي الأمل
 دخيلة توجان
 الزارات
 زراوة الجديدة
 الزركين
 زريق
 سمياط
 سيدي التواتي

شنتي قابس
 طبلبو
 عرام
 العكاريت
 العالاية
 عين تونين
 عيون الزركين
 غنوش
 قابس
 قابس الجمهورية
 قابس المنارة
 قابس الميناء
 قابس الهداية
 قابس حشاد
 كنانة
 لفام
 مارث
 المدو
 مطرش
 مظماطة
 مظماطة الجديدة
 المطوية
 مطوية المية
 منزل الحبيب
 نحال
 وذرف
 أم الصمعة
 إستقطيمية
 بزمة
 بشري
 بشلي
 البلديات
 بوعبد الله
 تلمين
 تنديب
 جرسين
 جمنة
 جنورة
 دوز
 دوز الشرقي
 دوز عوينة
 ربطة ال
 رجم معتوق
 الرحمان
 زاوية العانس
 زعفران دوز
 سعيدان

أم العرايس المحطة
 أولاد بوسعد
 المتلوي المحطة
 برج المظيلة
 بلخير
 بو عمران
 جبل المظيلة
 حوال الواد
 دوالي قفصة
 الرديف
 الرديف المحطة
 الرقوبة
 زمرة الرديف
 زنوش
 السند
 سيدي إحمد زروق
 سيدي بوبكر
 سيدي عيش
 الشيبية
 عبد الصادق
 العليم
 العيايشة
 القطار
 قفصة
 قفصة المحطة
 قفصة المطار
 قفصة حي الشباب
 قفصة حي النور
 كاف دربي
 لالة
 لورطس
 ماجورة
 المتلوي
 متلوي المناجم
 متلوي ثالجة
 منزل ميمون
 نشيو
 حي السرور
 أجم
 أركو
 أم التمر
 أولاد عمر
 بن قردان
 بني خداش
 بني معقل
 بوغراة
 الجدارية
 جربة
 جربة المطار
 جرجيس

الرقوبة
 الرياض
 سد غيان
 سد ويكش
 السواني
 سويحل
 سيدي زايد
 سيدي محرز
 سيدي مخلوف
 شماخ
 الشهبانية
 الشوامخ
 العاشين
 الغرابط
 غيزن
 القرع
 قربيص
 القصر الجديد
 قصر الجيرة
 قلالة
 كوتين
 مثنائية
 المحبوبين
 مدين
 مدين الجديدة
 مدين المثابرة
 مزاية
 مليئة جربة
 المنطقة الحرة جرجيس
 الموانسة
 المي
 ميدون
 الهشام
 وادي الزبيب
 والغ
 ورجيجان
 الورسنية
 أزمو
 بئر بورقيبة
 بئر دراسن
 بئر مروة
 براكاة الساحل
 بلي
 بني خالد
 بني خيار
 بوجريدة
 بوشراي
 بوعرقوب
 بوكريم
 تازركة



PDF Complete

Your complimentary use period has ended.
Thank you for using PDF Complete.

[Click Here to upgrade to Unlimited Pages and Expanded Features](#)

صور
قبلي
قبلي بياز
القلعة
ليماقس
المنصورة
نقة
نويل
أم العرايس
زاوية الجديدي
سليمان
سيدي داود
الصمعة
فرتونة
قربة حشاد
قرنبالية
قليبية الشرقية
المعمورة
منزل تميم
نابل
الهورية

الجرف
جلال
جميلة
حاسي الجربي
حاسي عمر
حلق الجمل
حمادي القبلي
دار جربة
رأس جدير
ربانة
الرجاء
رينين
السقالية
سيدي الظاهر
صاحب الجبل
غرداية
قربة
قرعة ساسي
قليبية
المرايسة
منزل بوزلفة
الميدة
نيانو
الوديان

تازقران
تاقلون
تركي
حمام الجبلي
حمام الغزاز
الحمامات
خنقة الحجاج
دار شعبان الشاطيء
دار شعبان الفهري
دار علوش
ديار الحجاج
الرحمة
زاوية المقاييز
سيدي الجديدي
الشريفات
عين تبرنق
فندق الجديد
قربص
قرنبالية الزهور
لبنة
ملول
منزل حر
نابل ثامر
وادي الخطف

ملخص

حدث تطور كبير في أنظمة التعرف الآلي على الكتابة اللاتينية و الصينية و في المقابل فإن التعرف على الكتابة العربية لم يتعد بعد أطواره الأولى بالرغم من ازدياد الأبحاث في هذا المجال مقارنة بالسنوات الماضية و يرجع ذلك إلى الصعوبات التي تختص بها اللغة العربية و كذا نقص وسائل التأكد و التصحيح.

تعالج هذه المذكرة مختلف المراحل التي يتكون منها نظام التعرف على كلمات اللغة العربية المكتوبة. سنتطرق في المحور الأول إلى موضوع التعرف على الكتابة العربية حيث سنعرض نبذة تاريخية عن مختلف الأبحاث التي أنجزت في هذا المجال و نصف أهم الخصائص و الصعوبات المتعلقة بالكتابة العربية و نذكر مختلف الطرق المستعملة. أما المحور الثاني فهو مخصص للتعريف بالمصنفات المستخدمة للتعرف على اللغة العربية المتمثلة في الشبكات العصبية الاصطناعية, شبكات Bayes, SVM, حلقات Markov المخفية, K-ppv. بينما ترك المحور الثالث لوصف مختلف مراحل التعرف على الكتابة العربية متضمنا المعالجة الأولية, تجزئة الكلمات, استخراج السمات, التصنيف و المعالجة النهائية. المحور الرابع و الأخير يشرح مراحل تطبيق النظام المقترح REMA على IFN/ENIT متطرقا لمختلف التقنيات و الخوارزميات المستعملة في كل مرحلة مثبتا نجاح هذا النظام الذي حقق نتائج جيدة و مشجعة.

BIBLIOGRAPHIE

- [1] Abd, M.A. and Paschos, G. "Effective Arabic Character Recognition using Support Vector Machines", *Springer on Pattern recognition*, pp. 7-11, 2007.
- [2] Abe, S. *Support Vector Machines for Pattern Classification*, Springer, USA, 2005.
- [3] Aburas, A and Rehiel, M.A. "Off-Line Omni-Style Handwritten Arabic Character Recognition System based on Wavelet Compression", *CS&IT ARISER*, (3), pp. 123-135, 2007.
- [4] Al-Badr, B and Haralick, R.M. "Segmentation Free Word Recognition with Application to Arabic", *IEEE proceeding of ICDAR*, pp. 49-77, 1995.
- [5] Alder, M. *An introduction to Pattern Recognition*, Heaven for books.com, 2001.
- [6] Alherbish, J, Ammar, R.A and Abdallah, M. "Arabic Character Recognition in a Multi-Processing Environment", *In Proceeding. IEEE Symposium on Computers and Communications*, Alexandria, Egypt, pp. 286-292, 1997.
- [7] Al-Imami, S and Usher, M. "On-line Recognition of Handwritten Arabic Characters", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, pp. 704-710, July 1990.
- [8] Allevard, M.T. *Représentation Symbolique de la Configuration de la Main: Application à la Reconnaissance de Signes et au Contrôle d'un Robot Mobile*, thèse, université de Savoie, le 07 juillet 2005.
- [9] Al-Ma'adeed, S., Elimam, D and Higgens, C. "Off-line Recognition of Handwritten Arabic Words using Multiple Hidden Markov Models", *Knowledge-Based Systems*, vol. 17, pp. 75-79, 2004.
- [10] Almuallim, H and Yamaguchi, S. "A Method of Recognition of Arabic Cursive Handwriting", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, n°5, pp. 715-722, 1987.

- oud, S.A and Qahwaji, R.S. "Recognition of Off-line Arabic Text using Hidden Markov Models", *Signal processing*, vol. 88, pp. 2902-2912, 2008.
- [12] Al-Ohali, Y. "Handwritten Word Recognition: Application to Arabic Check Processing", Thesis, Concordia University, Montreal, Quebec, Canada, February 2002.
- [13] Al-Qahtani, S.A and Khorsheed, M.S. "An Omni-Font HTK-Based Arabic Recognition System", *In Proceeding. 8th IASTED International Conference on Artificial Intelligence and Soft Computing*, Spain, 2004.
- [14] Al-Rashaideh, H. "Preprocessing for Arabic Word Handwriting Recognition", *Information Transmission in Computer Networks*, (6), pp. 11-19, 2006.
- [15] Al-Yousefi, H and Udpa, S.S. "Recognition of Arabic Characters", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, pp. 853-857, 1992.
- [16] Amin, A and Al-Fedaghi, S. "Machine Recognition of Printed Arabic Text utilizing Natural Language Morphology", *International journal Manuscript- Machine Studies*, vol. 35, pp. 769-788, 1991.
- [17] Amin, A et Masini, G. "Deux Méthodes de Reconnaissance de Mots pour l'écriture Arabe Manuscrite", *Actes RFIA'85, 5^{ème} congrès en Reconnaissance des Formes et Intelligence Artificielle*, Tome 2, pp. 837-848, 1985.
- [18] Amin, A and Murshed, N. "Off-line Recognition of Printed Arabic Words through Global Features and Neural Networks", *4 international workshop on Document Analysis Systems DAS'2000*, Rio de Janeiro, Brazil, pp. 267-277, 2000.
- [19] Ayat, N.E. "Sélection de Modèle Automatique des Machines à Vecteurs de Support: Application à la Reconnaissance d'images de Chiffres Manuscrits", Thèse de doctorat, université de Québec, Montréal, Canada, le 20 Janvier 2004.
- [20] Azizi, Sari, N.T. Souici-Meslati, L et Sellami, M. "Une Architecture de Combinaison Floue de Classifieurs Neuronaux pour la Reconnaissance des Mots Arabes Manuscrits", *CIFED'2002, Colloque International*

et le Document, pp. 89-96, Hammamet, Tunisie,

- [21] Badie, K and Shimura, M. "Machine Recognition of Arabic Handprinted Scripts", *the transaction of the IECE of Japan*, vol. E65, n°2, February 1982.
- [22] Bayley, R.R and Srinath, M. "Orthogonal Moments Features for use with Parametric and Non-Parametric Classifiers", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol 18, pp. 389-399, 1996.
- [23] Ben Amara, N et Belaid, A. "Une Méthode Stochastique pour la Reconnaissance de L'écriture Arabe Imprimée", *ICPR*, (1), pp. 121-126, 1999.
- [24] Ben Amara, N, Belaid, A et Ellouz, N. "Utilisation des Modèles Markoviens en Reconnaissance de L'Écriture Arabe", *In Proceeding. Colloque International Francophone sur l'Écrit et le Document*, Lyon, France, 1998.
- [25] Ben Amor N and Essoukhri, N. "Combining a Hybrid Approach for Features Selection and Hidden Markov Models in Multi-font Arabic Character Recognition", *in Proceeding. IEEE Conference. Document Image Analysis for Libraries*, 2006.
- [26] Benouareth, A, Ennaji, A et Sellami, M. "Utilisation des HMMs de Durée d'Etat Explicite pour la Reconnaissance des Mots Arabes Manuscrits", 2003.
- [27] Bezdek, J.C. Keller, J, Krisnapuram, R and Pal, N.R. *Fuzzy Models and Algorithms for Pattern Recognition and Image Processing*, Springer, New York, 2005.
- [28] Biadsy, F, El-Sana, J and Habash, N. "Online Arabic Handwritten Recognition using Hidden Markov Models", *Pattern Recognition*, (3), pp. 65-70, 2004.
- [29] Burrow, P. *Arabic Handwriting Recognition*, University of Edinburgh, 2004.
- [30] Camelin, N. *Stratégies Robustes de Compréhension de la Parole Basées sur des Méthodes de Classification Automatique*, Thèse de doctorat, université d'Avignon, le 26 novembre 2007.



- et Modélisation de Connaissances Linguistiques
l'écriture Manuscrite En-ligne, Thèse de doctorat,
institut national des sciences appliqués de Rennes, le 13 décembre 2005.
- [32] Chatelain, C. Extraction de Séquences Numériques dans des Documents
Manuscrits Quelconques, thèse, université de Rouen, le 05 décembre 2006.
- [33] Cheriet, M, Kharma, M, Liu, C.L and Suen, C.Y. *Character Recognition
Systems*, Wiley, 2007.
- [34] Cheung, A, Bennamoun, M and Bergmann, N.W. "An Arabic Optical
Character Recognition System using Recognition-based Segmentation",
Pattern Recognition, vol. 34, pp. 215-233, 2003.
- [35] Dehghan, M, Faez, K, Ahmadi, M and Shridhar, M. "Handwritten Farsi
(Arabic) Word Recognition: A Holistic Approach using Discrete HMM",
Pattern Recognition, vol. 34, n°5, pp. 1057-1065, 2001.
- [36] Duda, R.O, Hart, P.E and Stork, D.G. *Pattern Classification*, John
Wiley&Sons, New York, 2001.
- [37] El-Baati, Kheralla, Alimi et Ennaji, "De l'Hors-ligne vers un Système de
Reconnaissance En-ligne: Application à la Modélisation de l'écriture Arabe
Manuscrite Ancienne", 2006.
- [38] El-Dabi, S.S, Ramsis, R and Kamel, A. "Arabic Character Recognition
System: a Statistical Approach for Recognition Cursive Typewritten Text",
Pattern recognition, vol.23, n°5, pp. 485-495, 1990.
- [39] El-Hajj, R, Likforman-Sulem, L and Mokbel, C. "Arabic Handwriting
Recognition using Baseline Dependant Features and Hidden Markov
Modeling", *In Proceeding. International Conference on Document Analysis
and Recognition*, pp. 893-897, Seoul, Korea, 2005.
- [40] El-Sheikh, T.S and Guindi, R.M. "Automatic Recognition of Isolated
Arabic Characters", *Signal processing*, vol. 14, pp.177-184, 1988.
- [41] Farah, N, Khadir, M.T and Sellami, M. "Artificial Neural Network Fusion:
Application to Arabic Word Recognition", *In Proceedings-European
Symposium on Artificial Neural Networks (ESANN)*, pp. 151-156, 2005.
- [42] Farah, N, Souici, L and Sellami, M. "Arabic Word Recognition by
Classifiers and Context", *JCST, Journal of Computer Science and
Technology*, vol. 20, n°3, pp. 402-410, May 2005.

Ahmed, M. "Off-line Arabic Handwriting Recognition System", *International engineering in Systems Application*, pp.1-3, Nabeul-Hammamet, Tunisie, 1998.

- [44] Friedman, M and Kandel, A. *Introduction to Pattern Recognition: statistical, Structural, Neural and Fuzzy Logic Approaches*, World Scientific, Singapore, 2000.
- [45] Goraine, H and Al-Imami, S. "Off-line Arabic Character Recognition", *Computer*, pp. 71-74, July 1992.
- [46] Graupe, D. *Principles of Artificial Neural Networks*, World scientific, second edition, Chicago (USA), 2007.
- [47] Hachour, O. "Reconnaissance Hybride des Caractères Arabes Imprimés", dans un Proceeding de *Traitement Automatique de L'arabe (JEP-TALN)*, 2004.
- [48] Hagan, M.T, Demuth, H.B and Beal, M. *Neural Network design*, PWS Publishing Company, United States of America, 1996.
- [49] Halgamuge, S.K and Wang, L. *Classification and Clustering for Knowledge Discovery*, Springer, 2005.
- [50] Hammami, L and Berkani, D. "Recognition System for Printed Multi-font and Multi-size Arabic Characters", *the Arabian Journal for Science and Engineering*, vol. 27, pp. 57-72, 2002.
- [51] Hilditch, C.J, "Linear Skeletons from Square Cupboard", *Machine Intelligence*, vol. 4, pp. 403-420, 1969.
- [52] Irwin, J.D. *Supervised and Unsupervised Pattern Recognition, Feature Extraction and Computational*, CRC Press, Aitura University (USA), 2000.
- [53] Kanoun, K Alimi, A, Ennaji, A et Lecourtier, Y. "Reconnaissance de Mots Arabes par Approche Affixale", *CIFED, Colloque International Francophone sur l'Ecrit et le Document*, pp. 21-30, Hammamet, Tunisie, Octobre 2002.
- [54] Kasabov, N.K. *Foundations of Neural Networks, Fuzzy Systems, and Knowledge Engineering*, The MIT Press, second printing, London, England, 1998.
- [55] Kecman, V. *Learning and Soft Computing Support Vector Machines, Neural Networks and Fuzzy Logic Models*, The MIT Press, England, 2001.

- I.M.M and Al-Rifai, B. "Probabilistic Artificial Neural Network for the Arabic Handwritten Character", *Journal of Computer Science*, (2), pp. 879-884, 2006.
- [57] Khorsheed, M.S. "Recognition Handwriting Arabic Manuscripts using Hidden Markov Model", *Pattern Recognition Letters*, (24), pp. 2235-2242, 2003.
- [58] Khorsheed, M.S and Clocksin, W.F. "Structural Features of Cursive Arabic Script", *In Proceeding. British Machine Vision Conference*, pp. 422-431, 1999.
- [59] Kim, G, Govindaraju, V and Srihari, S.N. "An Architecture for Handwritten Text Recognition Systems", *IJDAR*, (2), pp. 37-44, 1999.
- [60] Koppurapu, S.K and Desai, U.B. *Bayesian Approach to Image Interpretation*, Kluwer Academic Publishers, New York, 2002.
- [61] Kotoulas, L and Andreadis, I. "Image Analysis using Moments", *Real-Time Imaging*, Elsevier Science, vol.10, n°6, pp. 371-378, 2000.
- [62] Kuncheva, L.I. *Combining Pattern classifiers: Methods and Algorithms*, Wiley Interscience, 2004.
- [63] Lahou, C. Contribution à l'Analyse Topologique des Images: Etude d'Algorithmes de Squelettisation pour Images 2 D et 3 D selon une Approche Topologique Digitale ou Topologique Discrète, thèse, université de Marne-la-Vallée, le 20 décembre 2001.
- [64] Lahouar, S et Ben Amara, N. "Reconnaissance de l'Écriture Arabe Imprimée par Modèles de Markov Cachés", 17^{ème} Journées tunisiennes en Electronique et Automatique, pp. 363-370, Nabeul, Tunisie, 1997.
- [65] Lorgio, L.M. "Off-line Arabic Handwriting Recognition: A Survey", *IEEE on Pattern Analysis and Machine Intelligence*, (28), pp. 712-714, 2006.
- [66] Mahjoub, M.A and Ellouz, N. "Nonstationary Hidden Markov Model: Application to on-line Arabic Character Recognition", *Computational Engineering in Systems Application*, pp. 111-116, Nabeul-Hammamet, Tunisie, 1998.
- [67] Margner, V, Pechwitz, M and El-Abed, H. "ICDAR 2005 Arabic Handwriting Recognition Competition", *ICDAR*, 2005.

- and Olivier, C. "Multi-level Arabic Handwritten Character Recognition". *Proceeding of joint International workshop on Pattern Recognition*, Sydney, Australia, 1998.
- [69] Mozaffari, S, Faez, K and Ziaratbah, M. "Structural Decomposition and Statistical Description of Farsi/Arabic Handwritten Numeric Characters", *In Proceeding. International Conference on Document Analysis and Recognition*, pp. 237-241, Seoul, Korea, 2005.
- [70] Mukundan, R. "Radial Tchebichef Invariants for Pattern Recognition". *Tencon*, pp. 114-115. Melbourne, 2005.
- [71] Mukundan, R. "Transform Coding using Discrete Tchebichef Polynomials", *Proceeding International Conference on Image and Vision Computing*, pp. 20-25, 2000.
- [72] Nabawi, A and Mahmoud, S, "Automatic Arabic Character Recognition using Back-propagation Neural Networks", *Journal of King Saudi Arabia, Computer and Sciences*, vol. 9, pp.1-29, 1997.
- [73] Nixon, M.S and Anado, A.S. *Feature extraction and image processing*, Newness, Great Britain, 2002.
- [74] Nouh, A, Sultan, A and Tolba, R. "An approach for Arabic Character Recognition", *J.Eng.Sci*, University of Riyadh, vol. 6, pp. 185-191, 1980.
- [75] Olivera, L.S, Morita, M and Sabourni, R. "Feature Selection for Ensembles Applied to Handwriting Recognition", *IJDAR*, (4), pp. 262-279, 2006.
- [76] Omidvar, O and Dayhoff, J. *Neural Networks and Pattern Recognition*, Academic Press, USA, 1998.
- [77] Parhami, B and Taraghi, M. "Automatic Recognition of Printed Farsi Texts", *Pattern Recognition*, vol. 21, pp. 293-302, 1981.
- [78] Parizeau, M. *Réseaux de Neurones*, université de Laval, France, 2004.
- [79] Pechwitz, M and Maergner, V. "Baseline Estimation for Arabic Handwritten Words". *In Proceeding. 8th International Workshop on Frontiers in Handwriting Recognition*, pp. 479-484, 2002.
- [80] Quinion, S. *Intégration de Connaissances Linguistiques pour la Reconnaissance de Textes Manuscrits En-ligne*, thèse, institut national des sciences appliqués de Rennes, le 17 décembre 2007.



Reconnaissance de Formes par Modélisation Mixte
à Base de Système d'Inférence Floue

Hiérarchisés, thèse, université de Rennes, le 28 octobre 2003.

- [82]Rousseau, L. Reconnaissance d'Écriture Manuscrite Hors-ligne par Reconstruction de l'Ordre du Trace en vue de l'Indexation de Documents d'Archives, institut national des sciences appliqués de Rennes, le 18 juin 2007.
- [83]Sari, T and Sellami, M, "Cursive Arabic Script Segmentation and Recognition System", *International Journal of Computers and Applications*, vol. 27, n°3, 2005.
- [84]Schwartz, R, Lapre, C, Makhoul, J, Raphael, C and Zhao Y. "Language-Independent OCR using a Continuous Speech Recognition System", *IEEE Proceeding. 13th International Conference on Pattern Recognition*, pp. 99-103, Vienne, Autriche, 1996.
- [85]Seb, N Cohen, I, Garg, A and THuang, T.S. *Machine learning in Computer Vision*, Springer, Netherland, 2005.
- [86]Theodoridis, S and Kontroubas, K. *Pattern recognition*, second edition, Elsevier, USA, 2003.
- [87]Touj, S, Ben Amara, N et Amiri, H. "Reconnaissance de l'Écriture Arabe Imprimée par Transformé de Hough Généralisée", *Le journal arabe international de technologie de l'information*, (2), 2005.
- [88]Trenkle, J, Gillies, A, Erlandson, E, Schlsser, S and Cavin, S. "Advances in Arabic recognition", 2001.
- [89]Vinciarelli, A. "A Survey on Off-Line Cursive Word Recognition", *Pattern Recognition*, (35), pp. 1433-1446, 2001.
- [90]Zidouri, A and Sarfraz. M. "On Optical Character Recognition of Arabic Text", *the 6 Engeneering Saoudi Conference (KFUPM)*, (4), pp. 109-118, 2002.



PDF
Complete

Your complimentary
use period has ended.
Thank you for using
PDF Complete.

[Click Here to upgrade to
Unlimited Pages and Expanded Features](#)

ARTICLES PUBLIÉS

- [1] Chergui, L et Benmohammed, M. "Classification des Mots Arabes Manuscrits par Réseaux ART", *JCI*, pp. 14. Guelma., le 20 mai 2008.
- [2] Chergui, L et Benmohammed. "Classification des Mots Arabes Manuscrits par Réseaux Fuzzy ART", *JED*, pp. a 29. Annaba, le 9 Juin 2008.
- [3] Chergui, L et Benmohammed. "Utilisation des Moments Géométriques de Tchebichef avec une Approche Neuro-flou pour la Reconnaissance de Mots Arabes Manuscrits", *IMAGE*, pp.91-94 .Guelma, le 19 novembre 2008.
- [4] Chergui, L and Benmohammed. "Fuzzy ART Network for Arabic Handwritten Recognition System", *ACIT*, Tunisie, le 18 décembre 2008.
- [5] Chergui, L et Benmohammed. "Les Réseaux ART et les Moments de Tchebichef pour la Reconnaissance de Mots Arabes Manuscrits", *JSIO*, Oran, le 16 décembre 2008.

CONCLUSION GÉNÉRALE

La reconnaissance d'écriture arabe est spécialement celle des mots manuscrits n'a pas encore atteint la maturité nécessaire pour permettre d'établir des études comparatives et approfondies des travaux effectués dans le domaine. Ceci est dû au nombre restreint de travaux menés sur les mots arabes manuscrits comparés à ceux portant sur les caractères ou mots imprimés et au fait que la plupart des modèles proposés n'ont pas dépassé le stade de tests préliminaires à cause du manque de bases de validation communes. Il est donc nécessaire d'unifier les bases et les protocoles de validation des résultats entre les différentes équipes de recherche.

Dans ce mémoire on a présenté un système de reconnaissance d'écriture arabe manuscrite hors-ligne utilisant un réseau Fuzzy ART comme classifieur. Nous avons commencé par introduire l'écriture arabe elle-même en s'attardant sur ses caractéristiques et difficultés ainsi que les approches de reconnaissance utilisées. Nous avons également présenté un historique des OCRs conçus pour l'écriture arabe. Aussi nous avons expliqué le principe des classifieurs utilisés dans les systèmes de reconnaissance de mots arabes tels que : les réseaux de neurones (PMCs et les réseaux Fuzzy ART), les SVMs, le classifieur bayésien, les HMMs et les k-ppv. Ensuite nous avons vu en détail les différentes composantes architecturales possibles d'un OCR, abordant principalement les phases : de prétraitement, de segmentation, d'extraction de primitives, de reconnaissance et de poste-traitement. Finalement nous avons présenté notre propre système de reconnaissance REMA qui est basé sur une approche globale ; où nous avons discuté des différentes phases de son implémentation en expliquant les outils et les algorithmes utilisés, ainsi que les tests et résultats enregistrés.



PDF
Complete

*Your complimentary
use period has ended.
Thank you for using
PDF Complete.*

[Click Here to upgrade to
Unlimited Pages and Expanded Features](#)

également intéressés à la reconnaissance de l'écriture
IFN/ENIT a servi pour le développement et
l'évaluation des travaux menés. Les résultats obtenus sont encourageants. Le
taux d'erreur noté est principalement dû à la mauvaise écriture et aux
problèmes de chevauchement des lettres dans les mots classés. L'étape de post-
traitement n'a pas été abordée, et une possible combinaison entre deux ou
plusieurs classifieurs est envisageable pour nos améliorations futures du
système proposé.



PDF
Complete

*Your complimentary
use period has ended.
Thank you for using
PDF Complete.*

[Click Here to upgrade to
Unlimited Pages and Expanded Features](#)

DÉDICACES

Je dédie ce mémoire à mon mari qui m'a beaucoup aidé et encouragé.

Aux deux personnes que je chérisse le plus au monde : Mon père et Ma mère.

A la mémoire de mes deux grands-mères.

A mes trois frères : Moustafa, Omar et Fouad.

A mes quatre sœurs : Samra, Rima, Imen et Maroi.



PDF
Complete

*Your complimentary
use period has ended.
Thank you for using
PDF Complete.*

[Click Here to upgrade to
Unlimited Pages and Expanded Features](#)

ABSTRACT

This memoir provides an Arabic off-line handwritten recognition system based on new classifiers: Fuzzy ART network which is a type of neural network. The proposed system employs statistical primitives and Tchebichef geometric moments as features which are novel in the domain of Arabic recognition system. The Tchebichef moments provides better feature representation capability and improved robustness with respect to image noise, over other types of moments.

Our system which is based on a holistic method includes four steps. Preprocessing, containing thinning, normalization and slant detection and correction. Feature extraction, utilizing Tchebichef moments and some statistical primitives. Training which is done with the Fuzzy ART training algorithm and classification through the Fuzzy ART classifier. The proposed system has been tested using the last set of the IFN/ENIT database, achieving good results.



PDF
Complete

*Your complimentary
use period has ended.
Thank you for using
PDF Complete.*

[Click Here to upgrade to
Unlimited Pages and Expanded Features](#)

RÉSUMÉ

Ce mémoire présente un système de reconnaissance d'écriture arabe basé sur un nouveau classifieur qui est le réseau de neurones Fuzzy ART et utilisant des primitives statistiques, et les moments de Tchebichef qui sont employés pour la première fois dans le domaine de la reconnaissance de l'écriture arabe.

Notre système qui est fondé sur une approche globale de reconnaissance inclut les étapes de : prétraitement comprenant les opérations de normalisation, de squelettisation et de redressement ; l'extraction de primitives représentée par les 49 premiers moments de Tchebichef et des primitives statistiques tels que : la moyenne, la variance et le nombre de sous-mots ; finalement l'étape de reconnaissance est réalisée à travers l'utilisation du classifieur Fuzzy ART.

Nos tests ont été effectués sur la base IFN/ENIT où nous avons obtenu un taux de reconnaissance très intéressant.

GLOSSAIRE

Arbre de décision : Un arbre de décision est une autre façon d'exprimer des règles de décision : chaque chemin dans l'arbre correspond à une règle, et les décisions figurent sur les feuilles terminales.

Arbre de décision flou : Un arbre de décision peut être flou dans le sens où les seuils sur les noeuds sont définis par des ensembles flous.

Ascendant : Hampe dans la partie supérieure d'une lettre.

Bagging : Elle appartient aux méthodes de combinaison parallèle de classifieurs par sélection statiques dont l'objectif est l'amélioration de la performance des classifieurs ayant la même structure, où les bases d'apprentissage sont construites de manière indépendante par bootstrapping.

Binarisation : Elle permet de passer d'une image de niveaux de gris à une image binaire composée de 2 valeurs 0 et 1, plus simple à traiter.

Boosting : Elle est semblable au bagging seulement la construction des bases d'apprentissage dépend systématiquement des résultats de classification antérieure.

Bruit : Un bruit (parasite) dans une image est considéré comme un phénomène de brusque variation de l'intensité d'un pixel par rapport à ses voisins, il provient de l'éclairage des dispositifs optiques et électroniques du capteur.

Chaînes de Markov Cachées Non-Stationnaires : Les NSHMMs sont caractérisées par un ensemble de paramètres de probabilité de transitions dynamiques.

Chaînes de Markov Cachées Planaires : Les PHMMs sont des HMMs où la probabilité d'observation dans chaque état est donnée par un HMM secondaire. L'architecture générale d'un PHMM inclut un modèle principal composé de super-états auxquels sont associés des modèles secondaires. Pour une image donnée, le modèle principal fait l'analyse selon une direction, et les modèles secondaires la font selon l'autre axe.

classifieurs : l'approche parallèle donne la possibilité de traiter indépendamment les uns des autres puis

fusionnent leurs réponses respectives. Cette fusion est faite :

- D'une manière démocratique, dans le sens où elle ne favorise aucun classifieur par rapport à un autre.
- D'une manière dirigée où on attribue à la réponse de chaque classifieur un poids en fonction de ses performances, donc l'ordre d'exécution des classifieurs n'intervient pas dans cette approche.

Combinaison séquentielle de classifieurs : La combinaison séquentielle, appelée également combinaison série, est organisée en niveaux successifs de décision permettant de réduire progressivement le nombre de classes possibles. Dans chaque niveau, il existe un seul classifieur qui prend en compte la réponse fournie par le classifieur placé en amont afin de traiter les rejets ou confirmer la décision obtenue sur la forme qui lui est présentée.

Contour : Les contours représentent la frontière entre les objets de l'image, ou la limite entre deux pixels dont les niveaux de gris représentent une différence significative. L'extraction de contour consiste à identifier dans l'image les points qui séparent deux textures différentes.

Contraste : C'est l'opposition marquée entre deux régions d'une image, plus précisément entre les régions sombres et les régions claires de cette image. Le contraste est défini en fonction des luminances de deux zones d'images.

Descendant : Hampe dans la partie inférieure d'une lettre.

Dimension d'une image : C'est la taille de l'image. Cette dernière se présente sous forme de matrice dont les éléments sont des valeurs numériques représentatives des intensités lumineuses (pixels).

Ecriture cursive : L'écriture est dite cursive lorsque les lettres qui composent un mot sont constituées d'un seul trait continu.

Ecriture imprimée : L'écriture est dite imprimée lorsque chaque caractère est constitué de traits qui lui sont propres.

Filtrage : Consiste à modifier la distribution fréquentielle des composantes d'un signal selon des spécifications données.

Filtre passe-bas : Ce filtre n'affecte pas les composantes de basse fréquence dans les données d'une image, mais doit atténuer les composantes de haute fréquence.



PDF Complete

Your complimentary use period has ended.
Thank you for using PDF Complete.

[Click Here to upgrade to Unlimited Pages and Expanded Features](#)

on physique ou technologique des lettres (sous informatique) et le terme caractère lorsqu'il s'agit

de considérations esthétiques.

Histogramme : L'histogramme des niveaux de gris ou des couleurs d'une image est une fonction qui donne la fréquence d'apparition de chaque niveau de gris (couleur) dans l'image.

Image : C'est une représentation d'une personne ou d'un objet par la peinture, la sculpture, le dessin, la photographie, le film, etc. C'est aussi un ensemble structuré d'informations qui, après affichage sur l'écran, ont une signification pour l'œil humain.

Image à niveaux de gris : Le niveau de gris est la valeur de l'intensité lumineuse en un point. La couleur du pixel peut prendre des valeurs allant du noir au blanc en passant par un nombre fini de niveaux intermédiaires. Donc pour représenter les images à niveaux de gris, on peut attribuer à chaque pixel de l'image une valeur correspondant à la quantité de lumière renvoyée. Cette valeur peut être comprise par exemple entre 0 et 255. Chaque pixel n'est donc plus représenté par un bit, mais par un octet. Pour cela, il faut que le matériel utilisé pour afficher l'image soit capable de produire les différents niveaux de gris correspondants.

Image numérique : L'image numérique est l'image dont la surface est divisée en éléments de tailles fixes appelés cellules ou pixels, ayant chacun comme caractéristique un niveau de gris ou de couleurs prélevées à l'emplacement correspondant dans l'image réelle, ou calculées à partir d'une description interne de la scène à représenter.

K-plus proches voisins flou : La version floue de l'algorithme des k-plus proches voisins dont le principe est d'affecter un individu à la classe qui est majoritaire parmi ses k- plus proches voisins apportant deux modifications :

- Introduction de la notion de distance dans le vote : un individu (en fait sa classe) aura d'autant plus de poids qu'il sera proche de l'individu à classer.
- Comme pour l'algorithme des c-moyennes, introduction d'une matrice d'appartenance de chaque individu à chaque classe.

Ligature : Trait qui relie deux lettres cursives.

Lissage : C'est un traitement local qui définit une nouvelle valeur de pixel en examinant son voisinage.



PDF Complete

Your complimentary use period has ended.
Thank you for using PDF Complete.

[Click Here to upgrade to Unlimited Pages and Expanded Features](#)

uminosité des points de l'image. Elle est définie
le l'intensité lumineuse d'une surface par l'aire

apparente de cette surface.

Moments de Hu : Historiquement, Hu fut le premier à présenter des travaux sur l'analyse d'image en utilisant les moments. Il utilisa les moments géométriques pour générer une famille d'invariants qui fût utilisée pour la reconnaissance automatique de caractères.

Moments invariants : Ce sont des moments invariants à la rotation, la transformation et au changement d'échelle.

Morphologie mathématique : Elle est utilisée en traitement d'image non seulement pour partitionner une image en zones homogènes selon certains critères, mais aussi pour isoler directement les objets à étudier.

N-grammes : C'est une chaîne de n caractères à laquelle est associée une probabilité, les probabilités des n-grammes sont calculées en fonction de leur fréquence dans un lexique.

Ondelettes : Elles permettent d'obtenir une information fréquentielle localisée concernant un signal ou une fonction de base choisie. Ce type de données est particulièrement intéressant pour la classification, l'inconvénient majeur de cette technique est qu'un léger décalage du caractère (mot) donnera des coefficients d'ondelettes totalement différents.

Pixel : C'est le plus petit point de l'image, c'est une entité calculable qui peut recevoir une structure et une quantification.

Point de jonction : Il connecte trois branches ou plus. On trouve deux types : les points de branchement et les points de croisement.

Point extrême : Correspond au début/fin d'un segment de ligne.

Pseudo-moments de Zernike : Ils sont basés sur les polynômes de Zernike orthogonaux offrant plus de vecteurs caractéristiques en donnant plus de performances aux images concernant le bruit.

Quantification vectorielle : C'est le fait de correspondre chaque vecteur continu de primitives à un indice discret d'un dictionnaire de référence appelé Code Book. Une fois le dictionnaire de références obtenu, cette correspondance entre les vecteurs caractéristiques et les indices du Code Book deviennent un simple calcul de type plus proche voisin.



PDF Complete

Your complimentary use period has ended.
Thank you for using PDF Complete.

[Click Here to upgrade to Unlimited Pages and Expanded Features](#)

s la reconnaissance en-ligne, les caractères sont
ils sont écrits. L'information est dynamique et

monodimensionnelle, et consiste en une séquence de traits qui suivent une échelle temporelle.

Reconnaissance hors-ligne : Dans la reconnaissance hors-ligne, le processus d'écriture n'est pas accessible. L'information ne possède pas de nature temporelle; elle est bidimensionnelle, et consiste en l'image digitalisée des caractères.

Réseau de Kohonen : Il s'agit d'un réseau non-supervisé avec un apprentissage compétitif où l'on apprend non seulement à modéliser l'espace des entrées avec des prototypes, mais également à construire une carte à une ou deux dimensions permettant de structurer cet espace. Les réseaux de Kohonen ont pour but de faire de la classification dans un espace multidimensionnel, ils reposent sur le principe que des entrées assez semblables doivent activer les mêmes neurones.

Résolution d'une image : C'est la clarté ou la finesse de détails atteinte par un moniteur ou une imprimante dans la production d'images. Sur les moniteurs d'ordinateurs, la résolution est exprimée en nombre de pixels par unité de mesure (pouce ou centimètre). On utilise aussi le mot résolution pour désigner le nombre total de pixels affichables horizontalement ou verticalement sur un moniteur; plus grand est ce nombre, meilleure est la résolution.

Scripteur : Ecrivain qui ont contribué dans la construction des bases de données des mots.

Segmentation : On est conduit pour analyser une image à la partager en un certain nombre de domaines connexes (unis): ce processus joue un rôle de plus en plus important en traitement numérique des images.

Sous-mot : C'est une partie du mot composé de caractères.

Synapse : C'est la connexion entre deux neurones.

Système mono-fonte : C'est un système qui ne peut reconnaître qu'une seule fonte à la fois, c'est-à-dire qu'il ne connaît de graphisme que d'une fonte unique.

Système multi-fonte : C'est un système qui est capable de reconnaître divers types de fonte parmi un ensemble de fontes préalablement apprises.

Système omni-fonte : C'est un système qui est capable de reconnaître toute fonte généralement sans apprentissage préalable.



PDF Complete
Your complimentary use period has ended.
Thank you for using PDF Complete.

[Click Here to upgrade to Unlimited Pages and Expanded Features](#)

est l'une des transformées les plus utilisées dans . Les caractéristiques extraites sont en fait les

descripteurs de Fourier basés sur les coefficients complexes des séries de Fourier. Elles sont invariantes aux rotations et aux changements d'échelle.

Transformée de Hough : Elle désigne des transformations qui permettent de détecter dans des images la présence de courbes ayant une forme paramétrique (droite, canonique,..) à partir d'un ensemble sélectionné de points appelés points caractéristiques.

Transformée de Hough Généralisée : Agit sur des points caractéristiques de l'image, généralement le contour. Elle permet en phase d'apprentissage de décrire chaque forme par un tableau.

Vocabulaire illimité : Le nombre de mots dépasse les 10000 mots.

Vocabulaire ouvert : Traitement de n'importe quel mot sans se baser sur un lexique.

Zonage : C'est une technique consiste à superposer une grille $n \times m$ sur l'image du caractère et pour chacune des régions résultantes, on calcule la moyenne ou le pourcentage de points en niveaux de gris, donnant ainsi un vecteur de taille $n \times m$ de caractéristiques.

INTRODUCTION GÉNÉRALE

La reconnaissance de formes (RF) est un domaine important du monde informatique dans lequel les recherches sont particulièrement actives. Elle est historiquement un chapitre de l'intelligence artificielle qui vise à automatiser le discernement de situations typiques au niveau de la perception. Ses méthodes trouvent des applications nombreuses dans divers domaines tels que : la médecine, le contrôle de procédés de fabrication, la vision robotique, le traitement de données volumineuses d'images, la reconnaissance de la parole et la lecture optique de documents.

La reconnaissance automatique de caractères, au sens large du terme, est une discipline qui a vu le jour dès l'apparition des premiers ordinateurs.

Reconnaître de l'écriture manuscrite consiste à associer une représentation symbolique à une séquence de symboles graphiques : on parle aussi de lecture automatique. Le but est de pouvoir utiliser cette représentation dans une application informatique. On distingue deux grands types d'utilisation :

- Traiter automatiquement des documents contenant de l'écriture manuscrite dont l'analyse par des individus prend trop de temps.
- Faciliter l'utilisation des ordinateurs pour des applications où un stylo est plus pratique qu'un clavier et une souris.

Un système de reconnaissance d'écriture est un logiciel dont l'entrée est un document (numérique ou numérisé) contenant de l'écriture manuscrite et qui fournit en sortie une représentation symbolique de cette écriture.

Dans ce travail, nous nous concentrons sur deux objectifs:

- Appliquer un classifieur neuro-flou non-utilisé jusqu'à présent dans le domaine de la reconnaissance d'écriture arabe et analyser ses performances par rapport aux classifieurs déjà exploités.



es moments de Tchebichef et leur qualité en tant
ant des mots manuscrits arabes.

Notre mémoire abordera la conception d'un système de reconnaissance d'écriture arabe manuscrite tout en développant les différentes étapes composant ce système, il est organisé en quatre chapitres :

Le premier chapitre étalera l'histoire de la reconnaissance de l'écriture arabe depuis l'avènement des premiers systèmes de reconnaissance optique de caractères arabes jusqu'aux derniers développements dans la reconnaissance en ligne et hors-ligne de ce type d'écriture. On trouvera aussi une description des caractéristiques et difficultés liées à la nature de l'écriture arabe elle-même, ainsi qu'une classification des principales méthodes utilisées pour la reconnaissance de ce type d'écriture.

Le deuxième chapitre présentera les différents types de classifieur utilisés dans la reconnaissance de l'écriture arabe incluant les réseaux de neurones (dont nous détaillerons deux classes, à savoir, les PMCs et les réseaux Fuzzy ART), les SVMs, les HMMs, le classifieur bayésien et les k-ppv. Nous en expliquerons ensuite leur principe, leurs avantages et inconvénients et nous présenterons également des exemples de systèmes de reconnaissance d'écriture arabe exploitant chacun d'eux.

Le troisième chapitre introduira l'architecture du système de reconnaissance d'écriture arabe, en commençant par l'étape de prétraitement contenant les opérations de : binarisation, lissage, squelettisation, redressement et normalisation. Suivie par l'étape de segmentation dont nous détaillerons deux types, à savoir : la segmentation explicite et implicite. Nous expliquerons par la suite la phase d'extraction de primitives où on parlera des différentes méthodes utilisées. Nous entamerons ensuite l'étape de reconnaissance, elle-même composée de deux phases : l'apprentissage et la classification. Enfin on terminera par l'étape de poste-traitement.



consacré à la réalisation effective de notre système
ler REMA (Reconnaissance d'Écriture Manuscrite

Arabe), où on décrira les étapes d'implémentation des différentes phases citées dans le chapitre précédant et on détaillera pour chacune d'elles les techniques et les opérations effectuées.

Pour finir, une conclusion générale mettra le point sur les problèmes rencontrés lors de la conception du système de reconnaissance d'écriture arabe manuscrite proposé, ainsi que nos perspectives et nos travaux futurs.

LISTE DES FIGURES

Figure 1.1	Sens d'écriture de la langue arabe.....	13
Figure 1.2	Forme de lettres ne pouvant être rattachées aux lettres qui les suivent.....	14
Figure 1.3	Etendu des espaces intra et inter-mots.....	14
Figure 1.4	Exemples de caractères arabes ayant une boucle.....	14
Figure 1.5	Lettres de morphologie identique distinguables par leurs points diacritiques..	15
Figure 1.6	Ecriture voyellée.....	16
Figure 1.7	Ecriture non-voyellée.....	16
Figure 1.8	Exemples de fontes arabes.....	17
Figure 1.9	Problème de segmentation d'un mot en sous-mots.....	19
Figure 1.10	Problème de chevauchement vertical.....	20
Figure 1.11	Problème de chevauchement horizontale.....	20
Figure 1.12	Problème de ligature.....	20
Figure 2.1	Perceptron Multi Couches à une couche cachée.....	29
Figure 2.2	Effet de l'apprentissage alternatif sur l'erreur.....	31
Figure 2.3	Problème de sur-apprentissage.....	31
Figure 2.4	Architecture du réseau ART.....	32
Figure 2.5	Réseau de neurones Fuzzy ART.....	35
Figure 2.6	Représentation vectorielle d'un ensemble flou E.....	36
Figure 2.7	Un séparateur à marge.....	38
Figure 2.8	SVM avec un noyau RBF.....	40
Figure 2.9	SVM avec un noyau polynomial de degré 2.....	40
Figure 2.10	Le modèle parallèle.....	46
Figure 2.11	Le modèle séquentiel.....	46
Figure 3.1	Profil de projection d'un texte arabe.....	56
Figure 3.2	Lissage par un filtre moyen.....	57
Figure 3.3	Exemple d'un filtre moyen.....	57
Figure 3.4	Histogramme de projection horizontale.....	61
Figure 3.5	Les huit directions de chaîne de code.....	62
Figure 4.1	Normalisation du mot "شعال".....	71

	t "حي الصلاح".....	71
	t "سيدي إبراهيم الزهار".....	72
Figure 4.4	Pixels voisins de P.....	72
Figure 4.5	Exemple de pixel 8-supprimable et non 4-supprimable.....	73
Figure 4.6	Echantillons squelettisés par l'algorithme de Rutovitz.....	75
Figure 4.7	Echantillons squelettisés par l'algorithme de Deutsh.....	76
Figure 4.8	Exemple de collision si on supprime x_3 , P n'est plus simple.....	77
Figure 4.9	Echantillons squelettisés par l'algorithme de Zhang et Wang.....	78
Figure 4.10	Echantillons squelettisés par l'algorithme de Zhang et Suen.....	79
Figure 4.11	Histogrammes de rotation du mot "بوزراعة".....	82
Figure 4.12	Histogrammes de rotation du mot "القلعة الصغرى".....	83
Figure 4.13	Histogrammes de rotation du mot "تونس القباضة الأصلية".....	83
Figure 4.14	Les mots "بوزراعة", "القلعة الصغرى" et "تونس القباضة الأصلية" après Redressement.....	84
Figure 4.15	Histogrammes de projection verticale des mots "بوزراعة", "القلعة الصغرى" et "تونس القباضة الأصلية".....	85
Figure 4.16	Exemple de nombre de sous-mots mal-calculés	86
Figure 4.17	Ligne de base des mots "بوزراعة", "القلعة الصغرى" et "تونس القباضة الأصلية".....	87
Figure 4.18	Exemple de formulaire rempli pour la construction de la base IFN/ENIT.....	92
Figure 4.19	Fonctionnement de l'algorithme d'apprentissage de Fuzzy ART.....	95
Figure 4.20	Régions associées à $w_i = [xy^c]^T$ (en trait plein) et à $a^1 \cap w_i$ (en trait pointillé).....	97
Figure 4.21	Squelettes du mot "بوزراعة" après l'application des quatre algorithmes de squelettisation.....	101
Figure 4.22	Squelettes du mot "القلعة الصغرى" après l'application des quatre algorithmes de squelettisation.....	101
Figure 4.23	Squelettes du mot "تونس القباضة الأصلية" après l'application des quatre algorithmes de squelettisation.....	102
Figure 4.24	Calcul des histogrammes de projection horizontale et détection de la ligne de base pour le mot "بوزراعة".....	102
Figure 4.25	Calcul des histogrammes de projection horizontale et détection de la ligne de base pour le mot "القلعة الصغرى".....	103
Figure 4.26	Calcul des histogrammes de projection horizontale et détection de la ligne de base pour le mot "تونس القباضة الأصلية".....	103

	ives statistiques et de l'histogramme de la projection premier mot de la classe "بوزراعة".....	104
Figure 4.28	Calcul des primitives statistiques et de l'histogramme de la projection verticale pour un deuxième mot de la classe "بوزراعة".....	104
Figure 4.29	Calcul des primitives statistiques et de l'histogramme de la projection verticale pour un premier mot de la classe "القلعة الصغرى".....	105
Figure 4.30	Calcul des primitives statistiques et de l'histogramme de la projection verticale pour un deuxième mot de la classe "القلعة الصغرى".....	105
Figure 4.31	Calcul des primitives statistiques et de l'histogramme de la projection verticale pour un premier mot de la classe "تونس القباضة الأصلية".....	106
Figure 4.32	Calcul des primitives statistiques et de l'histogramme de la projection verticale pour un deuxième mot de la classe "تونس القباضة الأصلية".....	106
Figure 4.33	Calcul des 49 premiers moments de Tchebichef d'un premier mot de la classe "بوزراعة".....	107
Figure 4.34	Calcul des 49 premiers moments de Tchebichef d'un deuxième mot de la classe "بوزراعة".....	107
Figure 4.35	Calcul des 49 premiers moments de Tchebichef d'un premier mot de la classe "القلعة الصغرى".....	108
Figure 4.36	Calcul des 49 premiers moments de Tchebichef d'un deuxième mot de la classe "القلعة الصغرى".....	108
Figure 4.37	Calcul des 49 premiers moments de Tchebichef d'un premier mot de la classe "تونس القباضة الأصلية".....	109
Figure 4.38	Calcul des 49 premiers moments de Tchebichef d'un deuxième mot de la classe "تونس القباضة الأصلية".....	109

LISTE DES TABLAUX

Tableau 1.1	L'alphabet arabe.....	12
Tableau 1.2	Les différentes formes d'apparence des caractères de l'alphabet arabe.....	13
Tableau 1.3	Ensemble des lettres arabes ayant des points diacritiques.....	14
Tableau 1.4	Liste des caractères pouvant avoir un "hamza".....	15
Tableau 1.5	Utilisation des signes de voyellation avec le caractère "ر".....	16
Tableau 1.6	Les sous-mots de l'écriture arabe.....	17
Tableau 1.7	Exemples de base de données de mots arabes	19
Tableau 1.8	Taxonomie des approches de reconnaissance hors-ligne de mots arabes.....	21
Tableau 2.1	Exemples des OCRs arabes exploitant les PMCs.....	28
Tableau 2.2	Exemples des OCRs arabes utilisant les SVMs.....	37
Tableau 2.3	Exemples d'un OCR arabe utilisant le classifieur bayésien	41
Tableau 2.4	Exemples des OCRs arabes utilisant les HMMs.....	44
Tableau 2.5	Exemples des OCRs arabes exploitant les k-ppv.....	48
Tableau 4.1	Les onze ongles de rotation pour le mot "بوزراعة".....	81
Tableau 4.2	Les onze ongles de rotation pour le mot "القلعة الصغرى".....	81
Tableau 4.3	Les onze ongles de rotation pour le mot "تونس القباضة الأصلية".....	82
Tableau 4.4	Comparaison entre les performances des différents moments	88
Tableau 4.5	Les 49 premiers moments de Tchebichef du premier échantillon du mot "القلعة الصغرى".....	90
Tableau 4.6	Les 49 premiers moments de Tchebichef du deuxième échantillon du mot "القلعة الصغرى".....	90
Tableau 4.7	Les 49 premiers moments de Tchebichef du troisième échantillon du mot "القلعة الصغرى".....	91
Tableau 4.8	Nombre de mots, de sous-mots et de caractères composant IFN/ENIT	93
Tableau 4.9	Age des scripteurs participant à la construction de la base IFN/ENIT.....	93
Tableau 4.10	Nombre de mots et de scripteurs pour chaque ensemble de l'IFN/ENIT.....	93
Tableau 4.11	Echantillons de mots de la base IFN/ENIT.....	94
Tableau 4.12	Résultats de la classification.....	100



PDF Complete

Your complimentary use period has ended. Thank you for using PDF Complete.

[Click Here to upgrade to Unlimited Pages and Expanded Features](#)

ts mal-écrits.....110

ts contenant des chevauchements et des ligatures.....112

Tableau 4.15 Comparaison entre les performances du système proposé et certains OCRs Arabes.....113



PDF
Complete

*Your complimentary
use period has ended.
Thank you for using
PDF Complete.*

[Click Here to upgrade to
Unlimited Pages and Expanded Features](#)

REMERCIEMENTS

Avant toute chose, je remercie dieu tout puissant de m'avoir aidé et éclairé le chemin pour la réalisation de ce mémoire.

Je tiens aussi, à exprimer toute ma gratitude à mon mari pour son soutien, son aide et surtout ses précieux conseils.

Je voudrais également remercier le Professeur Benmohammed Mohammed pour avoir accepté de rapporter sur ce travail. Merci pour tous ses commentaires et ses questions qui ont enrichi ce mémoire.

Un grand merci à Mr Bellami qui a accepté d'être président de jury ainsi qu'à Mr Chaoui et Mr Chikhi qui en ont été examinateur. Merci pour leurs remarques et leur enthousiasme.

SOMMAIRE

Introduction générale	í í í í í ..í í í í í í í ..í í í í í í í í í í í ...	XVI
1. L'écriture arabe	í í	...1
1.1 Historique des OCRs arabes	í í ..í í í í í í í í í ..í í í í ..í í í í í	...3
1.2 Caractéristiques de l'écriture arabe	í í	11
1.3 Difficultés et problèmes	í í	18
1.4 Approches de reconnaissance de mots arabes	í í	...21
1.4.1 L'approche globale	í í	...21
1.4.2 L'approche analytique	í í	..22
1.4.3 L'approche pseudo-analytique	í í	..22
1.4.4 L'approche hybride	í í	..23
1.4.5 L'approche affixale	í í	..23
1.5 Conclusion	í í	24
2. Classifieurs pour la reconnaissance d'écriture manuscrite	í í í í í í í í í í	..25
2.1 Le Perceptron Multi Couches	í í	...28
2.1.1 Architecture du réseau	í í	..28
2.1.2 Difficultés	í í	30
2.2 Les réseaux Fuzzy ART	í í	...32
2.2.1 Le dilemme : stabilité/plasticité	í í	..33
2.2.2 Structure de Fuzzy ART	í í	..34
2.2.3 Fonctionnement	í í	..35
2.2.4 Avantages	í í	...36
2.3 Les Machines à Vecteur du Support	í í	37
2.3.1 Risque structurel	í í	..38
2.3.2 Noyaux	í í39
2.3.3 Avantages	í í	40
2.3.4 Inconvénients	í í	...41
2.4 Le classifieur bayésien	í í	..41
2.4.1 Théorème de Bayes	í í	..41



PDF Complete
 Your complimentary use period has ended.
 Thank you for using PDF Complete.

[Click Here to upgrade to Unlimited Pages and Expanded Features](#)

MAÍ	í í	.67
	í í	..69
4.1.1	Normalisationí í	69
4.1.2	Squelettisationí í	...72
4.1.2.1	Algorithme de Rutovitzí í	.74
4.1.2.2	Algorithme de Deutshí í	...75
4.1.2.3	Algorithme de Zhang et Wangí í	..77
4.1.2.4	Algorithme de Zhang et Suení í	...78
4.1.3	Alignement et redressementí í	..79
4.2	Extraction de primitivesí í	84
4.2.1	Primitives statistiquesí í	..84
4.2.2	Moments de Tchebichefí í	..87
4.2.2.1	Moments orthogonaux discretsí í	87
4.2.2.2	Plan théoriqueí í	..89
4.3	Reconnaissanceí í	.91
4.3.1	Description de la baseí í	..91
4.3.2	Apprentissageí í	...95
4.3.3	Classificationí í	..100
4.3.3.1	Résultatsí í	.100
4.3.3.2	Comparaison avec d'autres systèmesí í í í í í í í í í í í í í í í í í	.112
4.4	Conclusioní í	..113
Conclusion générale	í í	114
Annexe 1 : Mots de la base IFN/ENIT	í í	...116
Annexe 2 : Algorithme d'apprentissage des PMCs	í í	..123
Annexe 3 : Algorithme de recherche des K-ppv	í í	..124
Bibliographie	í í125
Article publiés	í í	...133
Glossaire	í í	134



Your complimentary
use period has ended.
Thank you for using
PDF Complete.

[Click Here to upgrade to
Unlimited Pages and Expanded Features](#)

depuis le VI^{ème} siècle avant l'apparition de l'écriture cursive nabatéenne, et s'est progressivement répandue avec l'existence de l'Islam et la révélation coranique. L'arabe appartient au groupe des écritures sémitiques consonantiques du fait que seules les consonnes sont représentées.

Le nombre de systèmes conçus pour la reconnaissance de l'écriture arabe est moins important comparé à celui de l'écriture latine, cela est dû aux multitudes problèmes rencontrés dans le domaine de la reconnaissance de l'écriture arabe que les développeurs essayent de surpasser en proposant des systèmes non-adaptés de ceux spécifiquement créés pour l'écriture. On va citer dans un ordre chronologique une liste non-exhortative des travaux publiés concernant la reconnaissance de l'écriture arabe:

- **En 1975**, Nazif [34] a conçu et présenté le premier OCR arabe où il a travaillé sur les caractères arabes imprimés en se basant sur l'extraction de traits et leurs positions qui sont appelés des radicales, il a utilisé 20 radicales en se servant des modèles des caractères afin d'effectuer des corrélations entre les radicales et l'image du caractère, une étape de segmentation est appliquée pour segmenter le texte en caractère.
- **En 1980** Nouh, Sultan et Tolba [74] de l'université de Ryadh ont suggéré un ensemble standard de caractères arabes pour faciliter le traitement automatique des mots arabes.
- **En 1981**, Parhami et Traghi [77] ont présenté une approche pour la reconnaissance des mots arabes où les sous-mots sont segmentés et reconnus selon certaines primitives telles que : les concavités, les boucles et la connectivité en augmentant la tolérance par rapport aux variations des fontes.
- **En 1982**, Badi et Shimura [21] ont travaillé sur les mots imprimés. Ils ont obtenu un taux de reconnaissance de 89% en testant cinq types de fontes.
- **En 1985**, Amin et Masini [17] ont proposé un système de segmentation et de reconnaissance en utilisant les projections verticales et horizontales ainsi que les primitives structurelles, ils ont testé leur application sur 100 mots multi-fontes et ils ont obtenu un taux de reconnaissance de caractères de 85% et un taux de reconnaissance de mots de 95%.
- **En 1987**, Almuallim et Yamaguchi [10] ont introduit une des premières méthodes dans le domaine de reconnaissance de mots arabes où ils ont utilisé la représentation de squelette et les primitives structurelles pour la reconnaissance des mots ; ces derniers sont segmentés en sous-mots qui seront classifiés et combinés en caractères

classifieur a été construit d'un ensemble de règles de
taux de reconnaissance de 91% sur 400 mots écrits par
deux scripteurs.

- **En 1988**, un système de reconnaissance de mots basé sur les descripteurs de Fourier comme primitives et les hauteurs minimales de contour de mot comme points de segmentation, a été conçu par les deux chercheurs égyptiens El-Sheikh et Guindi [40].
- **En 1990**, de nouvelles approches ont été proposées :
 - El-Dabi, Ramsis et Kamel [38] ont utilisé les moments invariants accumulatifs qui aident à la segmentation des caractères connectés et chevauchés, la reconnaissance a été testée sur les régions ayant une largeur augmentée en levant le taux de reconnaissance jusqu'à 94%, à travers un multiplicateur empirique.
 - Al-Imami et Usher [7] ont utilisé deux types de primitives pour une écriture en-ligne; des primitives qualitatives tels que les points diacritiques et la direction de code et des primitives quantitatives comme la longueur et l'inclinaison du segment, et les coordonnées de chaque point de segment, servant tous comme entrées au classifieur d'arbre. Le taux de reconnaissance était de 86% sur une base de 50 mots contenant 150 caractères.
- **En 1991**, Amin et El-Fedaghi [16] ont proposé une architecture de système de reconnaissance de mots arabes imprimés multi-fontes comprenant quatre phases : l'acquisition, la segmentation, la reconnaissance du mot et la reconnaissance du caractère. Ils ont obtenu un taux de reconnaissance de 95,5% pour les caractères, et 81,58% pour les mots.
- **1992**, cette année fût marquée par un grand retour aux systèmes de reconnaissance de mots arabes à travers la révélation de deux nouveaux systèmes :
 - Le premier est celui de Al-Yousefi et Udpa [15] qui ont introduit une approche statistique pour la reconnaissance des caractères arabes isolés. Elle inclut les étapes suivantes : la segmentation de chaque caractère en des parties primaires et secondaires; l'extraction des primitives où ils ont utilisé neuf mesures cartésiennes ; et finalement la phase de reconnaissance est accomplie par un classifieur bayésien quadratique.
 - Le second est celui de Goraine, Usher et Al-Imami [45] qui ont employé une approche structurelle opérant sur les mots entiers. Elle a été appliquée aux mots manuscrits après leur segmentation et squelettisation. Une extraction des

 *Your complimentary use period has ended. Thank you for using PDF Complete.*

[Click Here to upgrade to Unlimited Pages and Expanded Features](#)

elles est alors effectuée ; ces dernières seront utilisées pour des règles. L'étape de post-traitement exploite un dictionnaire afin de confirmer ou de corriger les résultats obtenus. Le taux de reconnaissance a dépassé les 90% sur 180 mots incluant 600 caractères.

- **En 1995**, Al-Badr et Haralik [4] ont travaillé sur les mots arabes imprimés en se servant d'une méthode globale de reconnaissance, qui ne nécessite donc pas de segmentation. Ils ont utilisé les primitives structurales, détectées par l'application des opérations de la morphologie mathématique. Les premières expériences faites sur un lexique de 42000 mots ont donné un taux de reconnaissance de 99,4% pour le texte sans bruit, et 70% pour le texte scanné.
- **En 1996**, deux systèmes ont vu le jour :
 - Schwartz, Lapre, Makhoul, Raphael et Zhao [84] ont appliqué tout un système de reconnaissance de parole continue (le BBN BYBLOS) pour la reconnaissance de textes imprimés arabes, seuls le lexique et les données d'apprentissage ont été modifiés. Ils ont utilisé le principe de zonage avec les HMMs.
 - Namazi et Faez [65] ont travaillé sur l'écriture Arabe/Persienne imprimée multi-fontes en utilisant les pseudo-moments de Zernik comme primitives et un réseau de neurone flou comme classifieur. Ils ont testé leur système sur une base contenant 3700 exemples de caractère, le taux de reconnaissance était de 99,85%.
- **En 1997**, plusieurs travaux ont été publiés, dont nous citons :
 - Lahouar et Ben Amara [64] ont utilisé des HMMs de type gauche-droite pour la modélisation des caractères arabes imprimés préalablement segmentés. Les points diacritiques sont éliminés au cours de la phase de prétraitement, ce qui a réduit sensiblement le nombre de modèles à concevoir. Les premiers résultats obtenus sur un ensemble de 6399 caractères imprimés en fonte Jiza dans le corps 12 ont donné des résultats de 98,78%.
 - Alherbish, Ammar et Abdallah [6] ont présenté une nouvelle idée fondée sur un algorithme parallèle de reconnaissance qui peut augmenter la vitesse de reconnaissance de 5.34.
 - Nabawi et Mahmoud [72] ont présenté un algorithme de reconnaissance de caractères arabes utilisant des coupes 1-D du spectre du caractère. Les résultats expérimentaux ont montré que l'algorithme présenté est capable de reconnaître

à un taux de reconnaissance de 99.06%, en utilisant 10 projections sur l'axe X .

- **En 1998**, quatre systèmes ont été principalement développés :
 - Mahjoub et Ellouz [66] ont investi dans la modélisation de la durée d'état dans une chaîne de Markove selon deux contextes, stationnaire et non-stationnaire, où ils ont appliqué un modèle de type ergodique à durée d'états explicite à la reconnaissance en-ligne des caractères arabes. Le taux de reconnaissance obtenu varié de 87% à 92%.
 - Fehri et Ben Ahmed [43] ont proposé une méthode hybride de reconnaissance de mots manuscrits utilisant les réseaux de neurones et les chaînes de Markov de type 1D. Les premiers résultats de ce système ont été jugés encourageants.
 - Miled, Cheriet et Olivier [68] ont décrit une nouvelle stratégie de reconnaissance de mots manuscrits hybride en se basant sur une modélisation des trois niveaux perceptifs : global, pseudo-analytique et analytique, représentés, respectivement par une modélisation markovienne des mots, sous-mots et caractères. Testé sur une base de 5900 mots pour un lexique de 232 villes tunisiennes tirées de la base IFN/ENIT, les taux de reconnaissance des classifieurs : global, pseudo-analytique et analytique ont été respectivement de 67,8%, 72,5% et 81,8%.
 - Ben Amara, Belaid et Ellouz [24] ont développé une approche globale fondée sur l'emploi des chaînes de Markov planaires qui permet de modéliser les différentes variations de l'écriture arabe. Des scores de 99,84% sont atteints pour un vocabulaire de 100 pseudo-mots, et un corpus d'apprentissage de 24000 échantillons de même style.
- **En 1999**, deux applications ont été réalisées par :
 - Khorsheed et Clocksin [58] ont exploité le squelette du mot afin d'extraire des caractéristiques sans aucune segmentation antérieure, à savoir la direction et l'angle du segment.
 - Ben Amara [29] a travaillé sur les mots arabes imprimés en se basant sur une modélisation globale, pseudo-bidimensionnelle de sous-mots par des modèles de Markov cachés planaires. L'apprentissage fut effectué séparément pour chaque modèle de sous-mot sur une base de 240 échantillons par sous-mot. Les 24000 images qui ont été issues de documents imprimés de style Jiza (taille 12) ont abouti à un taux de reconnaissance de 99,84%.

rechercheurs se sont intéressés à la reconnaissance de

- Amin et Murshed [18] ont effectué la reconnaissance globale de mots arabes imprimés multi-fontes en se basant sur une approche d'apprentissage symbolique afin de créer des arbres de décision représentant les règles de classification des mots. Ce système a été testé sur 1000 mots arabes imprimés avec différentes fontes, où il a atteint un taux de reconnaissance de 92%.
- Benouareth [65] a proposé une combinaison entre une modélisation globale du mot manuscrit basée sur des caractéristiques perceptuelles avec une modélisation analytique utilisant des chaînes de Markov cachées correspondants aux lettres. Le taux de reconnaissance était élevé.
- Cheung, Bennamonn et Bergmann [34] ont conçu un système de reconnaissance d'écriture arabe manuscrite utilisant un nouveau algorithme de segmentation afin de surmonter les problèmes classiques de cette dernière dans l'écriture arabe, tels que les chevauchements. Ce système a donné un taux de reconnaissance de 90%.
- **En 2001**, principalement, deux autres systèmes sont apparus :
 - Trenkel, Gillies, Erlandson et Schollosser [88] ont préféré l'utilisation des ensembles d'arbres de décision, dont le but était de reconnaître des images de mots arabes et Farsi de qualité et de résolution faibles. Ils ont invoqué une nouvelle méthode d'apprentissage pour les ensemble d'arbre fondée sur le boosting dans l'espace des exemples et le bagging dans l'espace des primitives.
 - Dehghan, Faez, Ahmadi et Shridhar [35] ont conçu un système de reconnaissance de mots Farsi/Arabe en se basant sur une méthode globale. Le classifieur choisi est un modèle de Markov caché (HMM) de type droite-gauche combiné à un réseau de neurone de type Kohonen. Cette combinaison a abouti à des résultats encourageants.
- **En 2002**, les recherches donnent naissance à plusieurs systèmes, dont nous citons :
 - Hamami et Berkani [50] ont développé une approche structurale comportant des règles de prévention de la sur-segmentation pour manipuler plusieurs fontes, où la reconnaissance a été effectuée en deux niveaux. Le taux de reconnaissance était de 98,56%.

slati et sellami [20] ont conçu un système multi- (MCNF) pour la reconnaissance de mots manuscrits, en

se servant de la combinaison parallèle de trois classifieurs neuronaux de type Perceptron Multi Couches (PMC), basés chacun sur un type de primitives.

- Al-Ohali [12] a proposé un système de reconnaissance de mots manuscrits en utilisant une combinaison séquentielle entre un classifieur global neuronal de Kohlen et un classifieur analytique markovien. Testé sur un lexique de 67 sous-mots tirés de la base d'images Arabic Check Database, le classifieur neuronal a effectué une réduction lexicale avec un taux de réussite de 99,04% mais sa combinaison avec le classifieur markovien a augmenté le taux de reconnaissance de 3.68%.
- Kanoun, Alimi, Ennaji et Lecourtier [53] ont suggéré une nouvelle approche de reconnaissance de mots arabes imprimés basée sur les propriétés structurelles de la langue arabe en exploitant le fait que la langue arabe utilise un système de dérivation sophistiqué pour la génération des mots. Un mot décomposable est constitué par la composition des quatre entités de base qui sont les trois affixes ; préfixe, suffixe et infixes auxquelles est rajouté la racine. Les tests sur un vocabulaire de 1000 mots issus d'un lexique de 200 racines ont donné un taux de reconnaissance de 87%.
- Pechwitz et Margner [79] ont introduit une approche basée sur une segmentation implicite en utilisant une méthode par fenêtrage et une modélisation markovienne des caractères arabes. Le taux de reconnaissance a dépassé les 89% sur la base IFN/ENIT.
- Zidouri et Sarfraz [90] ont présenté un système basé sur une méthode de segmentation libre, ils ont extrait huit caractéristiques ; la largeur et la hauteur de caractère, le type de segment (horizontal, vertical), la direction gauche et droite à partir du centre et les quatre positions relatifs à la ligne de base. Le taux de reconnaissance était de 97% pour la fonte d'Ennaskh avec une vitesse de reconnaissance de 10 caractères par second.
- **En 2003**, deux systèmes sont conçus :
 - Khorsheed [57] a proposé un système où le graphe du squelette est décomposé en une suite de segments après une étape d'approximation linéaire, chaque segment est caractérisé par sa longueur et son angle d'orientation, une quantification vectorielle transforme la séquence de segments linéaires en une

iscrets qui sont présentés à un classifieur markovien
ne séquence de lettres, après un total de 12960 tests, le

taux de reconnaissance a atteint 72% sans intégrer aucune procédure de post-traitement.

- Benouareth, Ennaji et Sellami [26] ont développé un système de reconnaissance de mots manuscrits arabes basé sur les HMMs de durée d'état explicite de différents types ; Gaussien, Poissonien et Gamma. Des expérimentations significatives ont été effectuées sur la base de données IFN/ENIT donnant un taux de reconnaissance de 91,23%.
- **En 2004**, quatre systèmes ont été principalement réalisés :
 - Al-Ma'adeed, Elimam et Higgens [9] ont conçu un système de reconnaissance de mots manuscrits en combinant deux classifieurs de manière séquentielle, le premier étant basé sur des règles et le second sur des HMMs. Près de 4200 images de mots arabes issues de la base AHDB sont utilisées, les deux tiers ont servi à l'apprentissage et le reste a été réservé à la phase de test pendant laquelle le taux de reconnaissance a avoisiné les 60%.
 - Ell-Qahtani et Khorsheed [13] ont testé les primitives structurelles sur l'écriture manuscrite ayant un vocabulaire illimité en se servant des HMMs comme classifieur. L'étape de segmentation n'a pas été abordée et le système a traité les problèmes de ; ligature, de chevauchement et de variation de style.
 - Touj, Ben Amara, et Amiri [87] ont travaillé sur la reconnaissance de l'écriture arabe imprimée en utilisant la transformée de Hough généralisée. Ils ont effectué une phase de post-traitement. Les tests ont été faits sur des extraits de textes comportant environ 6400 caractères avec des pseudo-mots de longueurs variables allant jusqu'à 10 caractères successifs où le taux de reconnaissance a été de 96%.
 - Biadisy, El-Sana et Habash [28] ont proposé une nouvelle approche pour l'écriture en-ligne en se basant sur les HMMs où ils ont utilisé un filtre afin de diminuer la variation d'écriture, pour chaque point de tracé ils ont extrait trois primitives ; l'angle local, le segment supérieur et les boucles présentées. Le taux de reconnaissance a atteint 94,40%.
- **En 2005**, publication de cinq systèmes :
 - El-Hajj, Likforman-Suelem, et Mokbel [39] ont démontré l'avantage des caractéristiques basées sur la ligne de base supérieur et inférieur en

HMMs, ce système est testé sur la base IFN/ENIT
reconnaissance de 85,45%.

- Mozaffari, Faez et Ziaratban [69] ont travaillé sur les caractères arabes numériques en se basant sur des caractéristiques structurelles et statistiques, le classifieur utilisé a été celui de k-ppv. Huit chiffres ont été testés, et 280 images de chacun ont été utilisées pour l'apprentissage et 200 pour le test, ils ont obtenu un taux de reconnaissance de 94,44%.
- Farah, Souici et Sellami [42] ont appliqué une combinaison parallèle de trois classifieurs; le premier est un PMC, le deuxième est de type k-ppv tandis que le dernier est un k-ppv flou. Une base de 4800 images de mots représentant les 48 mots du lexique a été utilisée. Ils ont obtenu un taux de reconnaissance de 94%, une étape de post-traitement syntaxique basée sur la grammaire des montants littéraux a permis d'améliorer les résultats.
- Sari et Sellami [83] ont réalisé un système effectuant la reconnaissance de mots manuscrits arabes en utilisant un algorithme de segmentation explicite de mots basé sur l'exploitation des caractéristiques topologiques de l'écriture arabe. Le classifieur utilisé était un PMC, l'apprentissage et les tests ont été effectués sur la base IFN/ENIT et sur des enveloppes postales donnant un taux de reconnaissance de 86%.
- Farah, Khadir et Sellami [41] ont donné naissance à un système de reconnaissance de mots arabes manuscrits basé sur la combinaison parallèle de trois classifieurs neuronal de type PMC, le premier reçoit des primitives statistiques, le deuxième traite des primitives statistiques et structurelles tandis que le dernier travaille sur des primitives structurelles, les résultats obtenus seront combinés afin de produire la décision finale. Les tests ont donné un taux de reconnaissance de 99% sur une base de 4800 images de mots.
- **En 2006**, des travaux ont été réalisés par d'autres groupes de chercheurs :
 - El-Baati, Kherallah, Alimi et Ennaji [37] ont présenté une nouvelle approche de modélisation de l'écriture arabe manuscrite ancienne hors-ligne. Où ils ont extrait à partir de l'image scannée un signal monodimensionnel dont l'ordre temporel d'apparition des points est restauré. Afin de se rapprocher de la représentation en-ligne de l'écriture, ce signal est échantillonné selon la position des points. Une validation de l'approche est faite sur des exemples d'une base des mots extraits du patrimoine tunisien.

Al-Rifai [56] ont conçu un système de reconnaissance écrite basé sur les réseaux de neurones probabiliste, ils se sont servis des moments géométriques comme primitives en obtenant un taux de reconnaissance de 97%.

- Ben Amor et Essoukhri [25] ont proposé un système de reconnaissance de caractères imprimés multi-fontes utilisant comme classifieurs les HMMs et trois PMCs, l'extraction de caractéristiques est faite par la transformée de Hough, la base utilisée contient 85000 exemples de caractères dans cinq fontes ; Le taux de reconnaissance obtenu est de 97,36%.
- **En 2007**, trois systèmes ont été proposés :
 - Aburas et Rehiel [3] ont conçu un système pour la reconnaissance des caractères arabes manuscrits en se servant des ondelettes. Les tests sont appliqués à un ensemble de données contenant 41 groupes de l'alphabet arabe,, ils ont donné un taux de reconnaissance de 97,9%.
 - Abd et Paschos [1] ont utilisé les SVMs pour la reconnaissance de l'écriture arabe manuscrite où ils ont utilisés les moments invariants de Hu. Le taux de reconnaissance a atteint 98,99%.
 - Sarhan et Helalat [11] ont utilisé un PMC, l'étape d'extraction de caractéristiques consiste à faire entrer un vecteur de 35 primitives au réseau composé de deux couches cachées. Les résultats sont encourageants.
- **En 2008**, l'apparition du système de Al-Muhtaseb, Mahmoud et Qahwaji [11] qui ont introduit une nouvelle technique basée sur les fenêtres glissantes pour la reconnaissance d'écriture arabe imprimée hors-ligne en utilisant les HMMs, le vecteur de caractéristiques inclut 16 primitives, ils ont testé 8 fontes. Le taux de reconnaissance était entre 98,08% et 99,89%.

1.2 Caractéristiques de l'écriture arabe

De nos jours plusieurs écritures sont plus-ou-moins répandues dans le monde, quelques une d'entre-elles se démarquent par le nombre important de personnes l'utilisant. Dans ce contexte nous pouvant citer : l'écriture latine, chinoise, arabe, indienne, cyrillique ; chacune d'elles se distingue par des caractéristiques qui lui sont propres. Dans ce qui suit, nous nous intéresseront principalement à l'écriture arabe, dont les principales caractéristiques sont les suivantes :

vingt-huit lettres fondamentales. D'autres caractères représentant une deuxième et troisième voire une quatrième formes d'écriture pour une lettre donnée selon sa position dans le mot, tels que "ا, ب, ت, ث, ج, ح, خ, د, ذ, ر, ز, س, ش, ص" et "ة, ة". Le tableau ci-dessus présente les vingt-huit principaux caractères de l'alphabet arabe.

Caractère	Forme	Caractère	Forme
Alif	أ	Tad	ط
Ba	ب	Dha	ظ
Ta	ت	Ayn	ع
Tha	ث	Ghayn	غ
Jim	ج	Fa	ف
Ha	ح	Qaf	ق
Kha	خ	Kef	ك
Del	د	Lem	ل
Dhel	ذ	Mim	م
Ra	ر	Noun	ن
Zin	ز	Ha	ه
Sin	س	Wew	و
Chin	ش	Ya	ي
Sa	ص		
Dhad	ض		

Tableau 1.1: L'alphabet arabe.

- Il n'y a pas de différence entre les lettres manuscrites et les lettres imprimées ; les notions de lettre capitale et lettre minuscule n'existent pas comme c'est le cas pour l'écriture latine.
- Le sens d'écriture va de droite à gauche comme montré dans l'exemple ci-dessous tiré de la base IFN/ENIT (Figure 1.1).



Figure 1.1: Sens d'écriture de la langue arabe.

- Que sa soit du manuscrit ou de l'imprimé, la plupart des lettres s'entre-attache entre-elles. Leur forme se diffère selon leur position dans le mot, c'est-à-dire qu'elles sont soit précédées et/ou suivies d'autres lettres ou dans d'autres cas complètement isolées. C'est pourquoi l'écriture arabe est qualifiée d'être semi-cursive.
- Comme cité plus haut, et contrairement à l'alphabet latin, chacune des lettres arabes peut être écrite sous plusieurs formes suivant sa place dans le mot : début, milieu, fin ou isolée. La plupart des lettres ont ainsi quatre représentations différentes, excepté certaines lettres pour lesquelles les positions de milieu et de fin n'existent pas, ces lettres n'ont alors que deux représentations. Le tableau 1.2 donne toutes les formes possibles pour chaque lettre de l'alphabet arabe.

D	M	Fl	Fs	D	M	Fl	Fs
أ		أ	أ	ض	ض	ض	ض
ب	ب	ب	ب	ط	ط	ط	ط
ت	ت	ت	ت	ظ	ظ	ظ	ظ
ث	ث	ث	ث	ع	ع	ع	ع
ج	ج	ج	ج	غ	غ	غ	غ
ح	ح	ح	ح	ف	ف	ف	ف
خ	خ	خ	خ	ق	ق	ق	ق
د		د	د	ك	ك	ك	ك
ذ		ذ	ذ	ل	ل	ل	ل
ر		ر	ر	م	م	م	م
ز		ز	ز	ن	ن	ن	ن
س	س	س	س	ه	ه	ه	ه
ش	ش	ش	ش			و	و
ص	ص	ص	ص	ي	ي	ي	ي

D : En début du mot.

M : Au milieu du mot.

Fl : A la fin du mot mais lié à une lettre.

Fs : A la fin du mot sans être lié à la lettre.

Tableau 1.2: Les différentes formes possibles d'apparence des caractères de l'alphabet arabe.

- Certaines formes de lettres que nous avons regroupées dans la figure 1.2, ne peuvent dans aucun cas être rattachées à la lettre suivante, ce qui fait qu'un mot unique peut être entrecoupé d'un ou plusieurs espaces, lesquels sont aussi utilisés pour séparer les

ce inter-mot est généralement supérieure à l'espace attachés comme montré à la figure 1.3.

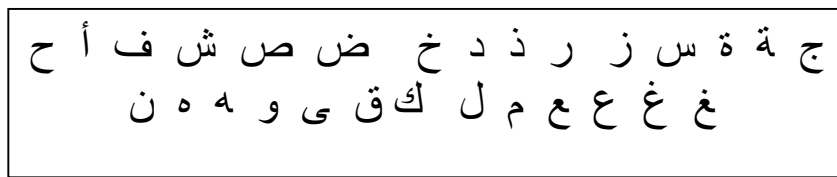


Figure 1.2: Forme de lettres ne pouvant être rattachées aux lettres qui les suivent.

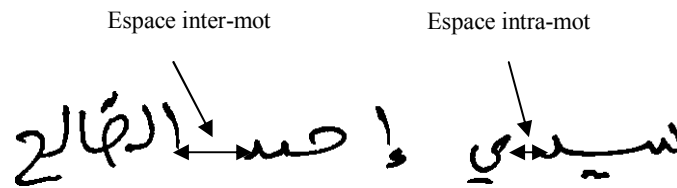


Figure 1.3: Etendu des espaces intra et inter-mots.

- Certains caractères arabes incluent une boucle qui peut avoir différentes formes. Nous pouvant trouver une, deux ou trois boucles dans le même caractère. La figure 1.4 affiche quelques caractères arabes incluant des boucles dans leurs formes.

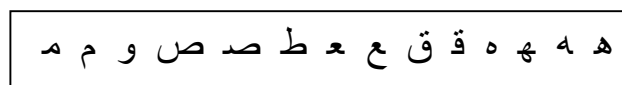


Figure 1.4: Exemple de caractères arabes ayant une boucle.

- Quinze caractères arabes incluent dans leur forme un, deux ou trois points diacritiques. Ces points peuvent se situer au-dessus ou au-dessous du caractère mais jamais en haut et en bas simultanément [7]. Le tableau 1.3 présente ces lettres en indiquant la position des points diacritiques dans leurs formes.

Position	Au-dessus	Au-dessous
Nombre de points		
Un point diacritique	ذ خ ز ض ظ غ ف ن	ب ج
Deux points diacritiques	ت ق	ي
Trois points diacritiques	ث ش	

Tableau 1.3: Ensemble des Lettres arabes ayant des points diacritiques.

(zay), qui se comporte, soit comme une lettre à part
 unique. Le tableau suivant indique les caractères qui
 peuvent avoir le "hamza" comme point diacritique en haut ou en bas.

Caractère	Caractère et hamza
ا	أ
ا	إ
و	ؤ
ى	ئ
ب	بْ

Tableau 1.4: Liste des caractères pouvant avoir un "hamza".

- L'existence de lettres différentes qui ont la même forme mais qui se distinguent par la position et le nombre des points qui leur appartiennent [34].

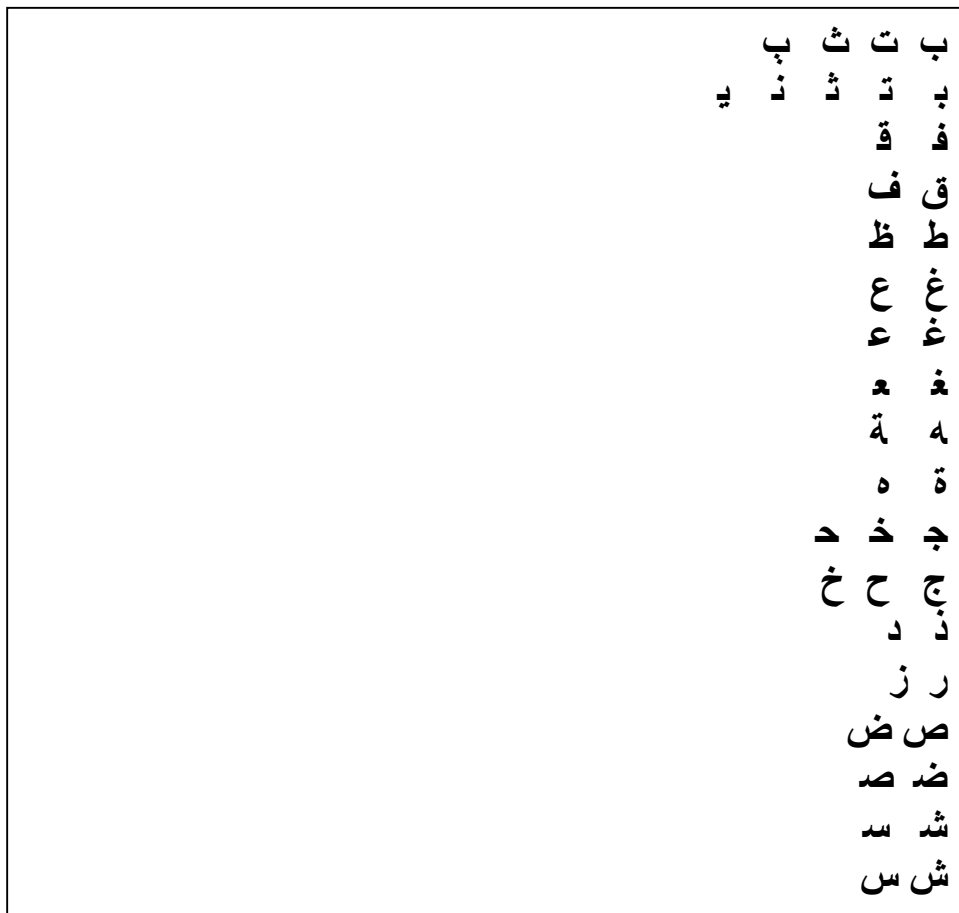


Figure 1.5: Lettres de morphologie identique distinguables par leurs points diacritiques.

ne sont pas utilisées systématiquement dans l'écriture. Les signes de voyellation qui dépendent de voyelles sont employés pour éviter des erreurs de prononciation. On trouve huit signes de voyellation qui peuvent se placer au-dessous d'une lettre tel que Kasrah " ِ " et des signes de voyellation qui peuvent se situer au-dessus tels que Fatha " َ ", Dhammah " ُ ", Soukoun " ْ " et Chaddah " ّ ", cette dernière doit être accompagnée de l'une des signes de voyellation, Fatha, Dhammah ou Kasrah. De plus trois Tanwin peuvent être formés à partir d'un double Fatha " ً ", d'un double Dhammah " ٍ " ou d'un double Kasrah " ٍ ". Nous pouvons donc distinguer deux types de textes : les textes sur lesquels figurent les signes de voyelles, tels que, le Coran et les livres d'apprentissage de la lecture et de l'écriture pour les enfants ; et d'autre ne nécessitant pas de signes de voyelles, comme par exemple les journaux, et les publications [8].

	Fatha	Dhammah	Kasrah	Soukoun
Signe de voyellation simple	َ	ُ	ِ	ْ
Signe de voyellation double	ً	ٍ	ٍ	impossible
Shaddah	ّ	ّ	ّ	impossible

Tableau 1.5: Utilisation des signes de voyellation avec la lettre "ر".

Les figures 1.6 et 1.7 donnent des exemples d'écriture voyellée et non-voyellée tirés du saint Coran.

الَّذِينَ اتَّخَذُوا دِينَهُمْ لَهْوًا وَلَعِبًا
 وَغَرَّتْهُمُ الْحَيَاةُ الدُّنْيَا فَالْيَوْمَ نَنسِفُهُمْ
 كَمَا نَسَوُا الْإِقَاءَ يَوْمَهُمْ هَذَا وَمَا كَانُوا
 بِثَانِينَ **بِحَدُوثٍ**

Figure 1.6: Écriture voyellée.

بِسْمِ اللّٰهِ الرَّحْمٰنِ الرَّحِیْمِ
 الْحَمْدُ لِلّٰهِ رَبِّ الْعَالَمِیْنَ
 الرَّحْمٰنِ الرَّحِیْمِ حَلَّكَ یَوْمَ الدِّیْنِ
 وَاِیَّاكَ نَعْبُدُ وَاِیَّاكَ نَسْتَعِیْنُ
 يَا صِدْقَ الصِّرَاطِ الْمُسْتَقِیْمِ
 صِرَاطِ الَّذِیْنَ اَنْعَمْتَ عَلَیْهِمْ غَیْرِ الْمَغضُوبِ
 عَلَیْهِمْ وَكَالضَّالِّیْنَ
 صَدَقَ اللّٰهُ الْعَلِیْمُ

Figure 1.7: Écriture non-voyellée.

Pour un caractère fixe, il peut comprendre un ou plusieurs pseudo-
 mots. Comme souvent différent de caractères, la bande centrale est
 généralement la plus chargée d'un point de vue densité d'informations en pixels, elle
 est appelée ligne de base [9]. Le tableau 1.6 donne quelques exemples tirés de la base
 IFN/ENIT ayant soit un, deux, trois ou bien quatre sous-mots.

1 Sous-mot	2 Sous-mots	3 Sous-mots	4 Sous-mots	5 Sous-mots
مثليين	لشغال	بوعثمان	كودة	بوزراقة
نقّة	شماع	مارث	الشرايع	الحارزة

Tableau 1.6: Les sous-mots dans l'écriture arabe.

- Il existe une grande variété de styles calligraphiques qui varient en fonction de la région, de l'époque, du type de document rédigé et des matériaux utilisés. Chaque style est régi par des lois particulières, et d'un style à l'autre les proportions d'une même lettre changent. Ces styles appelés encore fontes, sont regroupés sous différentes appellations : Coufique ancien, Coufique occidental, Coufique oriental, Diwani, Diwani Djali, Naskhi, Perse (Farsi), Thoulouth, Roqa et Magherebi. La figure suivante illustre un exemple d'une phrase écrite dans quelques-unes des fontes citées [1].



Figure 1.8: Exemples de fontes arabes.

ant des travaux de recherche entrepris, ces derniers demeurent en phase de prototypage. Ce retard peut en premier lieu, être attribué à la vingtaine d'année de décalage qui sépare le début timide des travaux sur la reconnaissance de l'écriture arabe de ceux portant sur les autres scripts, principalement le latin et l'asiatique. De plus, ce domaine présente deux types de problèmes qui ont freiné son évolution : le premier est inhérent à la morphologie même de l'Arabe, en particulier son aspect semi-cursif, et le second est dû au manque d'infrastructures et d'outils de test tels que les bases de données et les dictionnaires de validation.

- Difficultés liées au manque d'outils de test : l'une des difficultés majeures est l'absence d'outils de test et de validation tels que les bases de données d'image. Ces outils servent, entre autres, à valider les résultats obtenus et à évaluer de manière unifiée les divers travaux réalisés dans des environnements différents. Par conséquent, les résultats publiés ne peuvent être considérés comme valables que dans le contexte de l'étude effectuée et les performances annoncées n'ont pas réellement une signification objective. Récemment, un certain nombre de tentatives de construction de base d'images, pour la reconnaissance de textes arabes sont apparues dans la littérature, et que nous avons résumé dans le tableau 1.7.

Base	Type d'écriture	Constructeurs	Contenu
DARPA/S AIC	Imprimé	Le département américain de la défense.	345 images comportant une colonne de texte.
ERIM	Imprimé	L'institut de recherche environnemental de Michigan.	750 pages de texte arabe.
AHDB	Imprimé	L'université de Nottingham	30 000 mots.
Arabic cheque database	Manuscrit	Le CENPARMI, centre de RF de l'université de Concordia (Montréal) et le LIVIA, laboratoire d'imagerie de l'université du Québec et Al Rajhi, une grande société bancaire en Arabie Saoudite.	7000 images de chèques réels. 2499 montants littéraux. 2499 montants numériques.

		le laboratoire de système et de traitement de signal de l'école nationale d'ingénieurs de Tunis et le laboratoire de recherche sur la reconnaissance de l'écriture arabe de l'IFNb d'Allemagne.	26459 images de mot.
	IFN/ENIT	Manuscrit	
	Al-Isra database	Manuscrit	37 000 mots arabes. 10 000 chiffres. 2 500 signatures.

Tableau 1.7: Exemples de base de données de mots arabes.

Aussi le manque des outils tels que les lexiques, les dictionnaires de validation et les statistiques sur l'Arabe (succession et fréquence des caractères, bi-grammes, trigrammes et n-grammes aussi bien au niveau des caractères qu'au niveau des sous-mots ou des mots) afin d'améliorer les performances des systèmes de reconnaissance. Suite à leur absence, le post-traitement (lexical, syntaxique ou sémantique) est très peu utilisé dans les systèmes de reconnaissance de textes arabes. Les quelques systèmes qui l'intègrent sont basés sur des outils créés spécialement par les concepteurs, tels que des dictionnaires restreints au lexique considéré, ou des bases de règles grammaticales limitées au domaine applicatif abordé.

- Difficultés liées à la morphologie de l'écriture :
 - Un mot peut être constitué de plusieurs sous-mots comme on l'a indiqué dans la section précédente, et quand le nombre de sous-mot augmente la difficulté de reconnaissance augmente elle aussi.

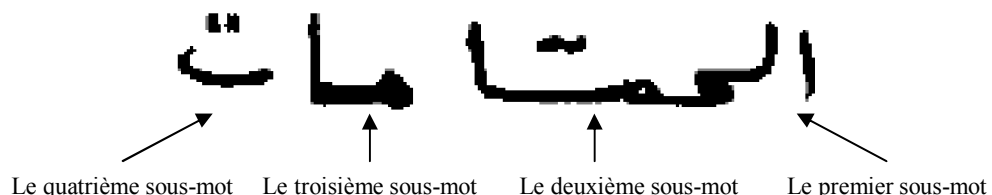


Figure 1.9: Problème de segmentation d'un mot en sous-mots.

- La présence des ligatures horizontales qui correspondent à insérer entre les caractères d'une même chaîne une ou plusieurs elongations ; ces elongations se

e du caractère courant et des ligatures (Figure 1.12) chevauchements (Figure 1.10 et Figure 1.11).



Figure 1.10: Problème de chevauchement vertical.



Figure 1.11: Problème de chevauchement horizontal.



Figure 1.12: Problème de ligature.

- Ainsi les particularités de l'écriture arabe rendent délicat le choix des opérations de prétraitement qui ne lui sont pas toutes applicables. Par exemple la squelettisation risque d'altérer la forme des points diacritiques, ce qui oblige de faire une extraction préliminaire de ces points dès le début des traitements.
- L'étape d'extraction de primitives est complexe car nous sommes confrontés à plusieurs fontes arabes et différentes formes pour chacun des caractères, ce qui rend leur modélisation plus difficile. D'autres problèmes sont à noter tels que la sensibilité des boucles et des points diacritiques aux bruits et aux dégradations.
- La segmentation de textes arabes est considérée comme étant un grand problème menant à l'augmentation du taux d'erreur lors de l'étape de reconnaissance. Les différents problèmes liés au processus de segmentations peuvent être résumés dans les points suivants:
 1. Segmentation en-lignes de texte où la plus part des méthodes sont basées sur la projection horizontale posant des problèmes d'interférences entre les ascendants et les descendants de lignes de texte voisines.
 2. Segmentation d'une ligne de texte en sous-mots qui est faite au moyen de projections verticales des différentes lignes du texte. Malheureusement cette méthode perdra son efficacité en présence de chevauchement vertical entre les sous-mots.

de mot en caractères est très difficile, surtout en lettres verticales, ces dernières peuvent être considérées comme étant une ou plusieurs lettres selon la taille prédéfinie pour un seul caractère lors de la segmentation.

1.4 Approches de reconnaissance de mots arabes

Selon les recherches entreprises jusqu'à présent, cinq approches de reconnaissance de mots arabes peuvent être considérées, en adoptant une classification fondée sur le type de l'entité de base modélisé par le système observé.

Le tableau suivant résume les principales caractéristiques des approches de reconnaissance hors-ligne de mots arabes, où l'approche globale est basée sur la reconnaissance du mot tandis que l'approche analytique considère la lettre comme entité de base ; l'approche pseudo-analytique travaille sur le sous-mot et l'approche hybride est une combinaison entre deux ou trois types des approches précédentes ; enfin l'approche affixale repose quand-à-elle sur les règles de grammaire de l'arabe.

Approches	Globale	Analytique	Pseudo-analytique	Hybride	Affixale
Entité de base	Mots	Lettres	Sous-mots	Combinaison de deux ou des trois précédentes	Morphèmes
Vocabulaire	Limité	Ouvert	Limité	Limité	Ouvert décomposable

Tableau 1.8: Taxonomie des approches de reconnaissance hors- ligne de mots arabes.

1.4.1 L'approche globale

Le mécanisme de reconnaissance d'une approche globale consiste à extraire à partir de l'image du mot, un ensemble de traits globaux permettant la discrimination du mot sans le segmenter. Cette approche est parfaitement envisageable pour une reconnaissance moyennant un lexique limité, et cela même pour des mots dégradés. Néanmoins, l'augmentation de la taille du lexique risque d'augmenter exponentiellement le taux de confiance. Notre objectif étant la reconnaissance de mots arabes manuscrits dont le vocabulaire est limité (taille<1000), nous privilégions donc l'utilisation d'une approche globale pour le problème de reconnaissance.

Généralement, les systèmes utilisant l'approche globale construisent une description du mot correspondant à la séquence de caractéristiques qui sont détectées. Malgré le début tardif de



Your complimentary
use period has ended.
Thank you for using
PDF Complete.

[Click Here to upgrade to
Unlimited Pages and Expanded Features](#)

relativement à l'analytique), plusieurs systèmes de conçus selon cette approche, tels que ceux de Azizi [12], Almaà'deed [5], El-Hajj [39], Farah [42], Khorsheed [57], Pechwitz [79], Dehghan [35], Al-Badr [4], Schwartz [84] et Benouareth [16].

1.4.2 L'approche analytique

Cette approche permet de traiter un lexique ouvert en se basant sur l'entité caractère. Le processus de reconnaissance d'un mot se résume aux points suivants :

- Partitionnement du mot en une suite de segments (lettres, ou partie de lettres appelés graphèmes).
- Extraction et identification de chaque segment.
- Combinaison de segments identifiés pour reconnaître le mot.

La principale difficulté rencontrée dans cette approche est de choisir une manière de segmenter le tracé du mot et d'obtenir de bonnes hypothèses de segmentation, c'est-à-dire les plus proches possibles du découpage en lettres souhaitées. Cette approche présente l'énorme avantage de proposer une démarche générale pour la modélisation d'un vocabulaire quelconque car elle s'appuie sur la structure alphabétique des mots pour les modéliser. Ainsi il est possible de reconnaître une infinité de mots en modélisant uniquement les 28 lettres de l'alphabet arabe sous leurs différentes formes selon leur position dans un mot. Plusieurs chercheurs ont conçu des algorithmes spéciaux de segmentation de mots en caractères comme ceux proposés par Motawa [16], Goraine [45], Fehri [43], Amin [16] et Almuallim [10].

1.4.3 L'approche pseudo-analytique

C'est un niveau intermédiaire entre les deux approches précédentes ; il est relativement peu exploité. Le niveau sous-mot ou pseudo-mot donnant naissance à une approche nouvelle et spécifique à la langue arabe, dans cette approche le sous-mot est considéré comme étant l'entité de base et les mots résultent d'une concaténation de sous-mots. Un des systèmes les plus connus de cette approche est celui de Alohalı [7] qui a travaillé sur les montants de chèques issus de la base d'images Arabic Chec Database, en utilisant une combinaison séquentielle entre un classifieur global neuronal et un classifieur analytique markovien, les sous-mots reconnus sont concaténés pour reconstruire les mots, et, par la suite les montants littéraux. Testé sur un lexique de 67 sous-mots, le classifieur neuronal donne un taux de reconnaissance de 99,04% tandis que le classifieur markovien donne 73,53%.

bit deux ou trois approches de celles présentées précédemment. On va principalement s'intéresser à deux systèmes basés sur l'hybridation ; le premier système est celui de Miled [68] qui a décrit une modélisation englobant trois niveaux perspectifs : global, pseudo-analytique et analytique, représentés respectivement par une modélisation markovienne des mots, des sous-mots et des caractères. Le deuxième est celui de Benouareth [26] qui a proposé une méthode de reconnaissance de mots manuscrits combinant une modélisation globale du mot, basée sur des caractéristiques perceptuelles avec une modélisation analytique utilisant des HMMs correspondant aux lettres.

1.4.5 L'approche affixale

Konoun [28] a proposé une nouvelle approche de reconnaissance de mots arabes, basée sur les propriétés structurelles de la langue, exploitant le fait que la langue arabe utilise un système de dérivation sophistiqué pour la génération des mots. Les mots dérivés, ou décomposables, représentent une part importante du vocabulaire, le reste du vocabulaire est constitué d'articles, adverbes, conjonctions, etc., ainsi que d'un vocabulaire non-décomposable.

Un mot décomposable est constitué par la composition des quatre entités de base ou morphèmes de base qui sont les trois affixes (préfixe, suffixe, infixe) et la racine. Partant d'un lexique de racines, le système de dérivation permet de générer tout un vocabulaire par adjonction et intercalation d'affixes qui forment un lexique fini et bien identifié, il est composé d'un ensemble de 24 préfixes, 29 suffixes et 19 infixes. Ces ensembles peuvent donc générer des combinaisons affixales, ou triades affixales (préfixe, infixe, suffixe) qui, combinées avec un lexique de racines permet de générer un vocabulaire de mots décomposables. Toutes les triades affixales ne sont pas forcément compatibles entre elles et acceptées par la langue, de même toutes les triades affixales cohérentes ne sont pas autorisées pour une racine donnée. L'ensemble de ces règles de cohérence, bien connues et énumérées, représente la notion de Restrictions Affixales. Cette notion est complétée par deux autres notions. La première est la notion de Restrictions Lexicales, qui fait référence à l'appartenance d'une racine à un dictionnaire de racines. La deuxième est la notion de Restrictions Sémantiques qui elle, complète les règles de cohérence précédentes par l'introduction de contraintes supplémentaires sur les quatre morphèmes d'un mot ; elle est engendrée par le fait que le sens d'une racine n'est pas toujours compatible avec celui que peut véhiculer une combinaison affixale donnée.



PDF Complete
Your complimentary use period has ended.
Thank you for using PDF Complete.

[Click Here to upgrade to Unlimited Pages and Expanded Features](#)

Après un bref historique des principaux OCRs réalisés pour l'écriture arabe. Quelle soit imprimée ou manuscrite, plusieurs caractéristiques spécifiques la distinguent des autres langues, engendrant de nouvelles problématiques et difficultés à surpasser.

Les différentes approches mentionnées dans ce chapitre ne peuvent être considérées qu'à travers le contexte selon lequel un système donné les exploite. Nous avons vu que certaines approches ne pouvaient être utilisées de manière efficace, que si elles agissaient sur des vocabulaires limités (telles que les méthodes globales, pseudo-analytiques et hybrides), et que d'autres étaient par contre plus destinées à des vocabulaires ouverts (comme les méthodes analytiques et affixales).

En ce qui concerne la variété intra-scripteurs, les différences sont également importantes. Ainsi l'utilisation d'un même système sur des bases différentes peut accuser une baisse de performances. Néanmoins, cet aspect est hors du cadre de ce mémoire, où nous nous limiterons à l'étude de l'écriture arabe manuscrite sur des données en provenance de scripteurs tunisiens, et sauf mention contraire, les données utilisées proviennent de la base IFN/ENIT, une base de noms de villes tunisiennes.

qui constituent un type de réseaux de neurones sont considérés comme étant un outil privilégié pour la reconnaissance de caractères manuscrits ou imprimés dans des polices d'écriture variées. Un réseau de neurones artificiels est basé sur un modèle simplifié de neurones. Ce modèle permet de simuler certaines fonctions du cerveau, comme la mémorisation associative, l'apprentissage par exemple, le travail en parallèle. On considère généralement qu'un réseau de neurones est constitué d'un grand ensemble d'unités (ou neurones), ayant chacune une petite mémoire locale. Ces unités sont reliées par des canaux de communication (ou de connexions, aussi appelées synapses d'après le terme biologique correspondant), qui transportent des données numériques. Les unités peuvent uniquement agir sur leurs données locales et sur les entrées qu'elles reçoivent par leurs connexions.

Ce type de réseau de neurones a été très utilisé dans le domaine de la reconnaissance de mots arabes manuscrits, le tableau suivant cite quelques systèmes en donnant le taux de reconnaissance de chacun ainsi que la base utilisée.

Systeme	Caractéristiques	Taux de reconnaissance	Base utilisée
Systeme de Sari et Sellami [83]	Les projections horizontales et verticales. Les densités des pixels par zone.	86%	IFN/ENIT.
Farah, Khadir et Sellami [41]	Les primitives statistiques et structurelles.	99%	4800 images de mots écrits par 100 scripteurs.

Tableau 2.1: Exemples des OCRs arabes exploitant les PMCs.

2.1.1 Architecture du réseau

Les PMCs sont des réseaux orientés de neurones artificiels organisés en couches et où l'information voyage dans un seul sens, de la couche d'entrée vers la couche de sortie sans aucune rétroaction [46]. Son apprentissage est de type supervisé, par correction des erreurs. L'architecture des PMCs se résume en au moins trois couches de neurones. Chaque neurone d'une couche donnée est totalement relié aux neurones de la couche qui le précède par des

connexions permettent de gouverner le fonctionnement

La première couche est celle des entrées, la dernière est la couche de sortie tandis que les couches intermédiaires sont des couches cachées. Ce type d'architecture permet la résolution de problème non-linéaire.

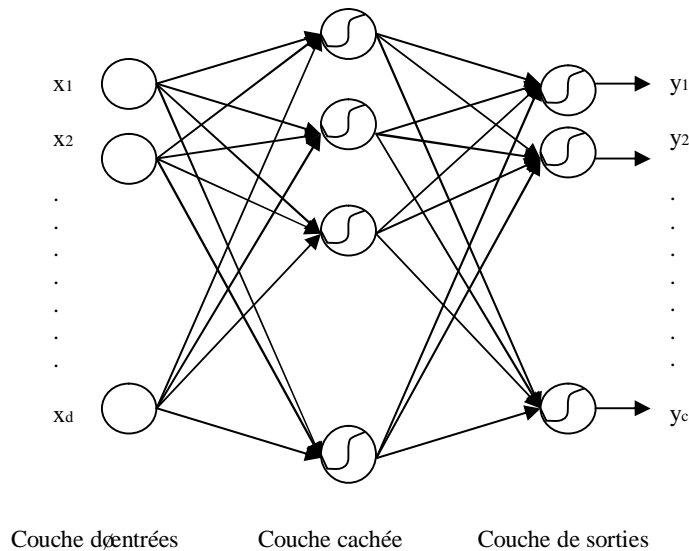


Figure 2.1: Perceptron Multi Couches à une couche cachée.

Le fonctionnement du Perceptron Multi Couche est simple : chaque neurone de la couche d'entrée génère un signal de sortie qui sera envoyé vers les neurones de la (ou les) couche(s) cachée(s). Ces derniers neurones produisent eux aussi un signal de sortie qui sera reçu par les neurones de la couche de sortie qui finalement génèrent le résultat de la prédiction. Numériquement chaque neurone est le résultat d'un calcul impliquant différentes fonctions et sommations. La valeur de sortie est une fonction pondérée des valeurs d'entrée qui est généralement modulée par la fonction sigmoïde [48].

Les PMC utilisent pour modifier leurs poids, un algorithme de rétropropagation du gradient, nécessitant des fonctions d'activations continues et dérivables. Il s'agit toujours de minimiser l'erreur quadratique. On propage la modification des poids de la couche de sortie jusqu'à la couche d'entrée. Dans le cas où une seule couche cachée est présentée, les fonctions discriminantes réalisées par un tel réseau de neurones sont de la forme:

$$\varphi_{l,i} = \left(\varphi_{2,j} - w_{2,j_0} + \sum_{j=1}^{h_1} w_{2,i_j} \varphi_{2,j} \left(-w_{1,i_0} + \sum_{q=1}^d w_{i,j_0} x_q \right) \right) \quad (2.1)$$

du neurone i de la couche l .

$W_{l,i} = [w_{l,i_0} w_{l,i_1} \dots w_{l,i_{h_{j-1}}}]^T$ est le vecteur de poids augmenté de ce même neurone.

h_l est le nombre de neurones de la couche l .

$X = [x_1 x_2 \dots x_d]^T$ est le vecteur de caractéristiques présenté à l'entrée du réseau.

2.1.2 Difficultés

Les PMCs sont probablement les réseaux de neurones qui ont été les plus utilisés dans le domaine de la reconnaissance d'écriture, néanmoins ils souffrent de plusieurs lacunes. Ci-dessous une liste non-exhaustive des différents problèmes que l'on peut rencontrer avec ce type de réseaux et les possibles solutions pour remédier.

- Le choix du nombre de neurones de la (ou les) couche(s) cachée(s) ainsi que la manière de concevoir la structure du PMC. Si l'ensemble des données d'apprentissage est regroupé sous forme de groupes ayant des caractéristiques similaires, le nombre de groupes peut être utilisé pour déterminer le nombre des neurones cachés [54]. Dans l'autre cas extrême, quand les données d'apprentissage sont séparées, donc n'ayant aucune caractéristique commune, le nombre des connexions approxime le nombre des exemples d'apprentissage pour que le réseau atteigne la convergence. Il existe plusieurs suggestions concernant le choix du nombre de neurones cachés dans les PMCs. Par exemple, le nombre minimum doit être :

$$h \geq \frac{(p-1)}{(n+2)} \quad (2.2)$$

où p est le nombre des exemples d'apprentissage et n est le nombre des entrées du réseau. Plus il y a de neurones cachés dans le réseau plus le nombre de caractéristiques détectées dans l'ensemble d'apprentissage sera important, mais plus la procédure d'apprentissage consommera de temps de calcul. Afin d'avoir la meilleure structure pour des PMCs donnés, on peut proposer un apprentissage structuré avec un oubli, c'est-à-dire que durant l'apprentissage on réduit les connexions ayant des petits poids.

- La possibilité d'oublier ce qui a été appris à partir des entrées précédentes en les remplaçant par celles des nouvelles entrées [44]. Afin d'éviter ce problème un apprentissage alternatif peut être exécuté comprenant les anciennes et les nouvelles données. Cela est illustré dans la figure 2.2 où on peut voir deux ensembles d'entrées A et B ayant déjà été présentés respectivement au réseau lors de la phase

Il est entraîné pour la deuxième fois avec le premier ensemble de données. L'erreur lors du deuxième cycle d'apprentissage est moins importante que celle du premier.

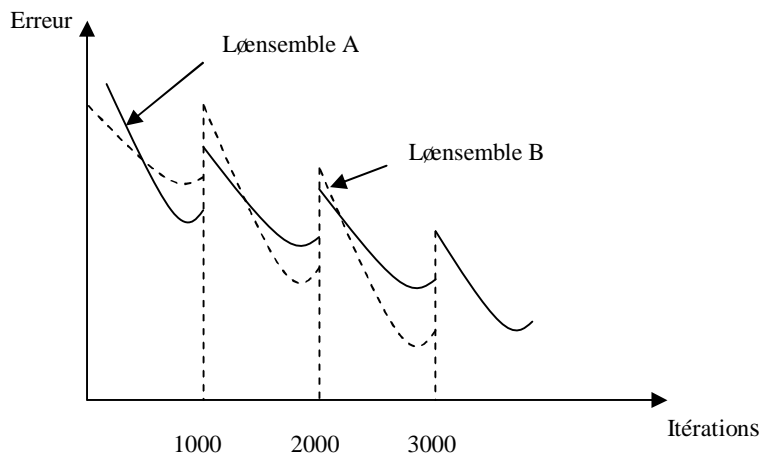


Figure 2.2: Effet de l'apprentissage alternatif sur l'erreur.

- Le sur-apprentissage qui indique que le réseau a trop approximé, c'est-à-dire qu'il a appris des exemples de données bruitées donnant ainsi une mauvaise généralisation sur les nouveaux exemples qui lui sont présentés, ce qui est indiqué par le figure 2.3. L'arrêt tôt de processus d'apprentissage est l'une des solutions qui peuvent être proposées ; pour cela, on calcule l'erreur d'apprentissage et l'erreur de test, si cette dernière est minimale alors on arrête l'apprentissage.

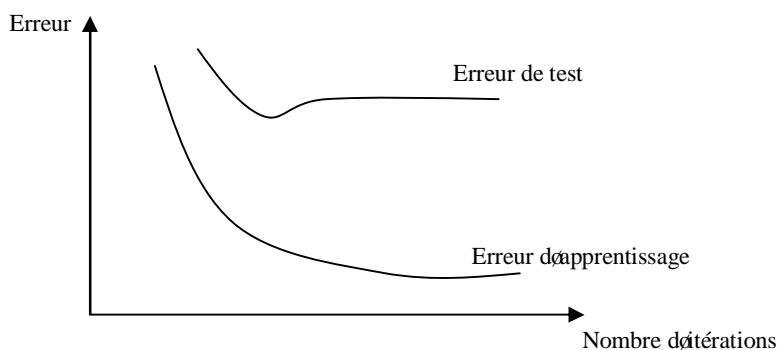


Figure 2.3: Problème de sur-apprentissage.

- La présence d'un grand nombre de paramètres à optimiser (poids des connexions).
- L'apparence de plusieurs maxima/minima locaux des fonctions coût.
- L'initialisation différente des poids peut mener à des solutions différentes. La fonction d'erreur converge souvent lentement et la complexité de calcul est de l'ordre

[76]. Ces deux facteurs font que cette méthode prend en temps de calcul et peut devenir critique pour

des problèmes réels.

2.2 Les réseaux Fuzzy ART

Les réseaux Fuzzy ART qui représentent une classe de la famille des réseaux ART (Adaptive Resonance Theory) est un modèle de réseau de neurones à architecture évolutive développé en 1987 par Carpenter et Grossberg [78]. Dans la plupart des réseaux de neurones, deux étapes sont considérées. La première est la phase d'apprentissage où les poids des connexions sont modifiés selon une règle d'apprentissage. La deuxième est la phase d'exécution où les poids ne sont plus modifiés. Dans les réseaux ART, ces deux étapes sont réalisées simultanément. Le réseau en phase de test, s'adapte à des entrées inconnues en construisant de nouvelles classes (ajout de neurones) tout en dégradant au minimum les informations déjà mémorisées.

Le réseau ART se compose de deux couches F_1 et F_2 que montre la figure 2.4 qui sont interconnectées à travers une mémoire à long terme (LTM), les entrées sont reçues à partir de F_1 tandis que la classification est réalisée au niveau de F_2 .

La classification est comparée à l'espérance du réseau qui réside dans les poids de LTM de F_2 . Si il y a une compétition les espérances seront renforcées, autrement la classification est rejetée, la mémoire courte (STM) est utilisée pour provoquer des changements progressifs dans le LTM. Il existe plusieurs versions de réseau : ART1, ART2, ART3, Fuzzy ART, ARTMAP, Fuzzy ARTMAP, etc.

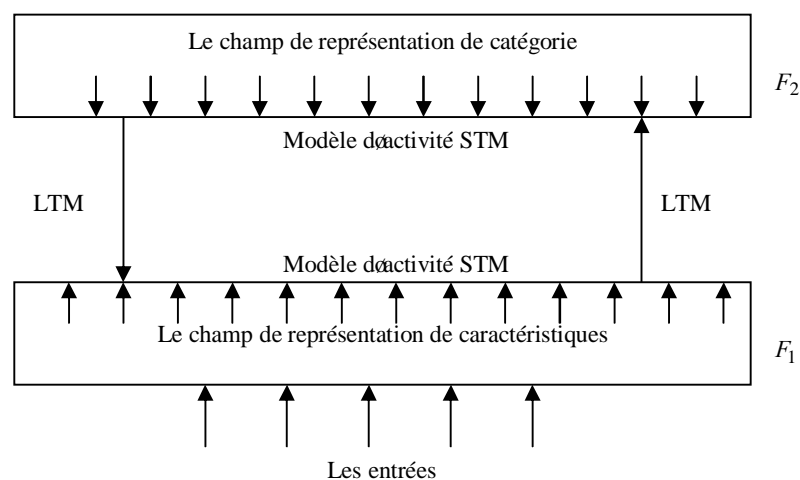


Figure 2.4: Architecture du réseau ART.



Your complimentary
use period has ended.
Thank you for using
PDF Complete.

[Click Here to upgrade to
Unlimited Pages and Expanded Features](#)

de classification, plusieurs chercheurs ont favorisé l'utilisation de réseaux neuronaux à rétropropagation. Toutefois, les recherches dans ce domaine ont rapidement démontré qu'une telle méthode et même en effectuant correctement certaines tâches de classification, ne permet pas d'intégrer de nouveaux vecteurs au processus de classification déjà entrepris. Lorsque la classification est effectuée sur un ensemble dynamique d'entrées, il est nécessaire, de reprendre l'ensemble des calculs déjà effectués pour intégrer les modifications affectant les vecteurs d'entrées. Lorsqu'elle est confrontée à d'importants ensembles de données, cette opération devient computationnellement très coûteuse. Bref, les algorithmes de rétropropagation sont très efficaces sur des corpus fermés, non-évolutifs, mais se heurtent rapidement à un important problème de plasticité.

Afin de contrer cette limite des réseaux à rétropropagation, il est possible de concevoir un réseau de neurones dont la particularité serait de réeffectuer la phase d'entraînement sur les nouveaux vecteurs d'entrées. Cette caractéristique permet au réseau de s'adapter aux différents changements de son environnement. Par opposition au réseau précédent, ce second réseau possède la capacité de s'adapter à son environnement en traitant de manière dynamique des entrées différentes, mais se ressemblant. Nous dirons alors de ce réseau qu'il possède la propriété de plasticité.

Toutefois, la majorité des réseaux possédant la propriété de plasticité ne peuvent conserver dans le temps leur apprentissage. Ainsi, la qualité des résultats obtenus par ces réseaux sensibles à l'ajout de nouveaux vecteurs d'entrées diminuera rapidement dans le temps, au fur et à mesure de l'ajout de nouvelles entrées. En ce sens, cet algorithme ne saurait satisfaire le critère de stabilité, selon lequel le système doit conserver dans le temps les structures reconnues (connaissances acquises) malgré la différence des entrées. C'est ce conflit entre la plasticité et la stabilité d'un réseau de neurones qui a été le centre des préoccupations des travaux de Grossberg [78] sur la théorie de la résonance adaptative.

Ce dilemme entre plasticité et stabilité est posé de manière plus formelle, par les points suivants :

- Comment un système d'apprentissage, tout en étant conçu pour demeurer adaptatif à l'égard d'informations pertinentes, peut-il demeurer stable à l'égard d'informations non-pertinentes?
- Comment un système peut-il varier adéquatement entre des modes de plasticité et de stabilité afin d'atteindre la stabilité sans être rigide et la plasticité sans être chaotique?



PDF Complete
Your complimentary use period has ended.
Thank you for using PDF Complete.

[Click Here to upgrade to Unlimited Pages and Expanded Features](#)

un système de conserver son apprentissage tout en quelles connaissances ?

- Qu'est-ce qui empêche le nouvel apprentissage d'éliminer ou de supprimer les apprentissages antérieurs?

Ce sont ces quatre questions qui sont à l'origine des travaux de Grossberg [46], car, jusqu'au milieu des années quatre-vingt, la majorité des algorithmes de classification sont soit stables, mais incapables de générer de nouveaux groupes de vecteurs; soit adaptatifs (donc sensibles aux nouvelles entrées), mais instables dans leur apprentissage.

Un bon système neuronal doit constamment passer d'un mode adaptatif à un mode stable et vice versa. Il doit être en mesure de conserver l'information antérieure, mais en même temps de tenir compte de la nouveauté. Ceci signifie concrètement que le système cherche constamment à s'adapter aux nouvelles entrées tout en conservant les classes antérieures. Il doit donc stabiliser les classes qu'il découvre, mais aussi les changer, si cela est nécessaire, en regard de la nouvelle réalité qui se présente à lui.

De manière générale, tous les modèles de la famille ART normalisent les entrées, réduisent le bruit et stabilisent les prototypes dans le temps. La normalisation de la fonction de transfert consiste à établir un même seuil qui est imposé à tous les neurones. Cette normalisation permet une certaine stabilisation des entrées. La réduction du bruit consiste en la définition de paramètres d'intensité avec lesquels une entrée peut agir. Finalement, la stabilisation dans le temps est la possibilité pour le système de considérer une entrée en regard des entrées antérieures; c'est en ce sens qu'un tel système ne perd pas l'information acquise antérieurement.

2.2.2 Structure de Fuzzy ART

Fuzzy ART est un réseau compétitif à deux couches de neurones complètement inter-reliées. Une couche de comparaison F_1 sert à coder les entrées avec un encodage dit complémentaire et une couche de compétition F_2 semblable à celle du réseau de Kohonen [76], ces deux couches sont activées par une entrée X. Le Fuzzy ART propose une catégorisation originale avec des classes représentées par des prototypes. Un prototype, poids du neurone est défini par un vecteur à N dimensions :

$W_j = (w_{j1}, w_{j2}, \dots, w_{jN})$ où j est l'indice du neurone, $1 \leq j \leq M$, et N est la dimension de l'entrée X [78].

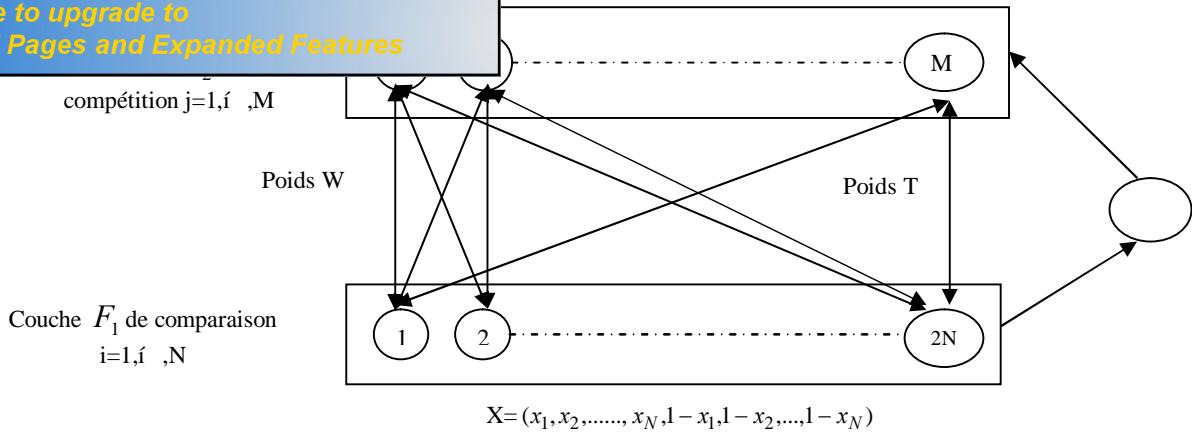


Figure 2.5: Réseau de neurones Fuzzy ART.

Cependant, tout comme les autres architectures ART, le Fuzzy ART incorpore un mécanisme de rétroaction permettant de stabiliser les prototypes appris dans les vecteurs de poids qui relient les deux couches. Ce mécanisme dit de résonance est contrôlé par un paramètre qui permet de réinitialiser au besoin la couche compétitive. Pour chaque entrée, les sorties du réseau spécifient une catégorie parmi les classes de sortie.

Ce type de classifieur n'a pas été exploité dans le domaine de reconnaissance de l'écriture arabe manuscrite ; notre système sera donc le premier à utiliser un réseau de neurones de type Fuzzy ART dans ce domaine.

2.2.3 Fonctionnement

Le Fuzzy ART est un réseau constructif où de nouveaux neurones sont alloués lors de la phase d'apprentissage. Généralement, on fixe au départ un nombre maximum de neurones S , limitant ainsi le nombre maximum de catégories possibles. Initialement, aucun neurone n'est actif. L'allocation subséquente de nouvelles catégories dépendra à la fois des entrées et des paramètres de l'algorithme.

Les entrées d'un réseau Fuzzy ART doivent être des ensembles flous définis sur un référentiel discret. Soit p la représentation vectorielle d'un ensemble flou E défini sur un référentiel discret $R = \{r_1, r_2, \dots, r_R\}$.

$$E = \{(r_j, p_j) \mid r_j \in R \text{ et } 0 \leq p_j \leq 1\}, \quad j = 1, \dots, R \quad (2.3)$$

Où $0 \leq p_j \leq 1$ représente le degré d'appartenance à E de l'élément r_j du référentiel.

Le vecteur $p = [p_1 p_2 \dots p_R]^T$ peut ainsi être interprété comme un point dans un hyper-cube unitaire de R dimensions. L'exemple décrit par la figure 2.6 montre un hyper-cube à deux

Exploiter un réseau Fuzzy ART, nous devons donc l'apprentissage de manière à les faire tous entrer à l'intérieur d'un hyper-cube unitaire, pour en faire des ensembles flous.

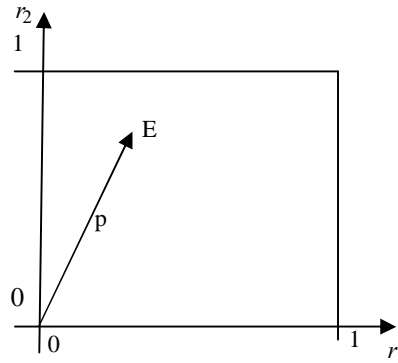


Figure 2.6: Représentation vectorielle d'un ensemble flou E.

Le réseau Fuzzy ART est contrôlé par trois paramètres. Le paramètre de sélection (ou de choix) appelé α est un nombre positif très petit de l'ordre de 10^{-5} qui est utilisé lors de la compétition entre les neurones. Le second, ρ est le paramètre de vigilance son rôle est de définir la taille de l'hyper-rectangle. Le troisième, η est le paramètre d'apprentissage qui influence sur la vitesse d'apprentissage des neurones, il est habituellement fixé à 1 pour un apprentissage rapide des neurones.

$$\alpha > 0, \quad 0 < \rho \leq 1, \quad 0 < \eta < 1 \quad (2.4)$$

2.2.4 Avantages

- L'algorithme de Fuzzy ART propose des calculs simplifiés pour la formation des classes sous forme d'hyper-boîtes, contrairement à des classes circulaires telles que retrouvées dans la plupart des algorithmes de réseaux de neurones.
- Il s'exécute en deux étapes basées sur deux critères de distance, soit l'activation et le choix.
- Il offre de bons résultats de catégorisation avec une précision modérée sur les poids des neurones.
- L'apprentissage se stabilise dans un nombre fini d'itérations.
- Le modèle Fuzzy ART exploite à fond un des avantages inhérents de l'approche neuronale ; le parallélisme.
- La forme prototype d'une classe sera immédiatement reconnue, même si elle n'a jamais été présentée, grâce aux caractéristiques pertinentes.

Support

Le SVM de l'anglais (Support Vector Machine) est un classifieur binaire appartenant aux méthodes statistiques non-paramétriques dont l'objectif est de déterminer la frontière optimale séparant les données d'apprentissage de deux classes.

L'idée principale des SVMs est de projeter des données sur un espace de plus grande dimension afin que les données non-linéairement séparables deviennent linéairement séparables donc c'est le fait de projeter les données de l'espace d'entrée (appartenant à deux classes différentes non-linéairement séparables) dans un espace de plus grande dimension appelé espace de caractéristiques où on construit un hyperplan optimal séparant les deux classes tels que :

- Les vecteurs appartenant aux différentes classes se trouvent de différents côtés de l'hyperplan.
- La plus petite distance entre les points et l'hyperplan (la marge) est maximale [55].

Les SVMs sont très utilisés dans le domaine de la reconnaissance d'écriture latine [33], mais pour l'écriture arabe on ne peut citer que deux travaux qui seront mentionnés dans le tableau 2.2 en détaillant les caractéristiques choisies, le taux de reconnaissance ainsi que la base utilisée pour l'apprentissage et pour le test.

Système	Caractéristiques	Taux de reconnaissance	Base utilisée
Système de Khorsheed et Cloksin [58]	Transformation polaire de l'image. Les moments géométriques.	Non-cité	DARPA
Système de Abd et Paschos [1]	Les moments de Hu. Le nombre de points diacritiques. La position de points diacritiques. Le nombre de trous dans le caractère. Le nombre de pixels dans la zone des points diacritiques.	98,99%	58 formes de caractères écrits par différents scripteurs.

Tableau 2.2: Exemples des OCRs arabes utilisant les SVMs.

On appelle vecteur de support les points les plus proches, qui seuls sont utilisés pour la détermination de l'hyperplan. La marge est la distance entre l'hyperplan et les exemples

figure 2.7. Le SVM suggère d'utiliser l'hyperplan au milieu des points des deux classes d'exemples.

L'hyperplan séparateur optimal est celui qui maximise la marge.

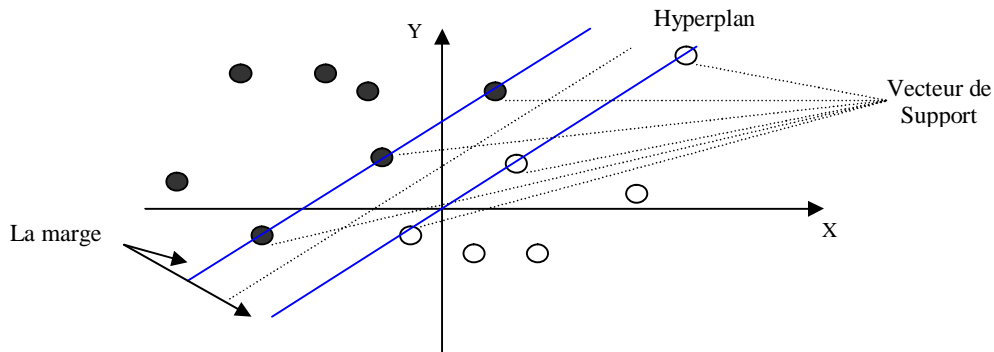


Figure 2.7: Un séparateur à marge.

2.3.1 Risque structurel

Nous introduisons ici le principe du risque structurel que minimise le classifieur SVM. Considérons un problème de classification à deux classes et soit les paires de données étiquetée : $(x_1, y_1), \dots, (x_l, y_l) \in \mathbb{R}^N \times \{\pm 1\}$ (2.5)

Où x_i représentent la $i^{\text{ème}}$ observation de l'ensemble d'apprentissage et y_i son étiquette. Et

soit l'ensemble de fonctions f_α défini : $\{f_\alpha : \alpha \in A\}, f_\alpha : \mathbb{R}^N \rightarrow \{\pm 1\}$ (2.6)

Supposons aussi qu'il existe une distribution $P(x, y)$ des données dont on ignore le modèle. La tâche d'apprentissage de f_α qui approxime au mieux cette distribution consiste à minimiser

$$\text{le risque réel donné par : } R(\alpha) = \int \frac{1}{2} |f_\alpha(x) - y| dP(x, y) \quad (2.7)$$

Ne connaissant pas $p(x, y)$, il est difficile d'estimer le risque $R(\cdot)$. Il est possible toutefois de considérer une fonction de risque empirique de la forme :

$$R_{emp}(\alpha) = \frac{1}{l} \sum_{i=1}^l \frac{1}{2} |f_\alpha(x_i) - y_i| \quad (2.8)$$

Si l'ensemble d'apprentissage est fini, la minimisation du risque empirique donné par cette équation ne garantit pas un minimum pour le risque réel. La minimisation du risque structurel évite le sur-apprentissage des données à la convergence de la procédure d'apprentissage [55].

Le risque structurel constitue une borne supérieure de l'erreur de généralisation qui s'écrit :

$$R(\alpha) \leq R_{emp}(\alpha) + \phi\left(\frac{h}{l}, \frac{\log(\eta)}{l}\right) \quad (2.9)$$

est d'au moins $1 - \eta$, et où est un terme de confiance

$$\text{donné par : } \phi\left(\frac{h}{l}, \frac{\log(\eta)}{l}\right) = \sqrt{\frac{h(\log\frac{1}{h} + 1) - \log(\frac{\eta}{4})}{l}} \quad (2.10)$$

Le paramètre h est appelé VC ou dimension de Vapnik-Chervonenkis. Il décrit la capacité de l'ensemble de fonctions solutions. Pour un problème de classification binaire, h est le nombre maximum de points séparables selon 2^h configurations par l'ensemble de fonctions solutions. La minimisation de cette borne consiste à opérer une régularisation de l'erreur empirique $R_{emp}(\alpha)$ via le terme de pénalité $\phi\left(\frac{h}{l}, \frac{\log(\eta)}{l}\right)$. Donc un minimum est garanti en minimisant à la fois l'erreur empirique et le terme de régularisation. Le terme machine d'apprentissage est utilisé pour désigner l'ensemble de fonctions de décision f_α dont la machine dispose.

Le processus de choisir le bon sous-ensemble de fonctions solutions revient à contrôler la complexité du classifieur en cherchant le meilleur compromis entre une faible erreur empirique et une complexité moindre.

2.3.2 Noyaux

Le produit scalaire des données vectorielles scalaires dans l'espace de caractéristiques est une fonction noyau $K(x, y)$ qui est un algorithme prenant en entrée une matrice symétrique positive. En pratique, quelques familles de fonctions noyau paramétrables sont connues et il revient à l'utilisateur des SVMs d'effectuer des tests pour déterminer celle qui convient le mieux pour son application. Les fonctions noyaux acceptables doivent respecter la condition de Mercer. Si pour toute fonction $g(x)$ telle que :

$$\int g(x)^2 dx \text{ est finie, on a } \iint K(x, y) g(x) g(y) dx dy \geq 0 \quad (2.11)$$

Le respect de la condition de Mercer [36] garantit que le problème quadratique possède une solution noyau polynomial :

$$K(x, y) = (\langle x, y \rangle + c)^m, m \in \mathbb{N} \text{ et } c > 0 \quad (2.12)$$

La forme générale de la fonction noyau est :

$$K(x_i, x_j) = \phi(x_i)^T \phi(x_j) \quad (2.13)$$

Les fonctions noyau encodent des connaissances essentielles sur le problème étudié, en particulier:

- Une mesure de similarité sur les données.

itions de décisions possibles.

ur les hypothèses : par exemple les fonctions noyau gaussiennes pénalisent les dérivées de tous les ordres et favorisent donc les solutions régulières.

- Le type de covariance dans l'espace des entrées stipulant comment des points en divers endroits de l'espace sont liés les uns aux autres. On peut ainsi chercher des fonctions noyau invariantes pour certaines transformations de l'espace d'entrée, comme par exemple la rotation [19].

Parmi les noyaux les plus utilisés on trouve :

- Le noyau linéaire : $K(x_i^T, x_j) = x_i^T \cdot x_j$ (2.14)

- Le noyau polynomial : $K(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0$ (2.15)

- Le noyau de fonction à base radiale RBF : $K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0$ (2.16)

- Le noyau sigmoïde : $K(x_i, x_j) = \tanh(\gamma x_i^T x_j + r)$ (2.17)

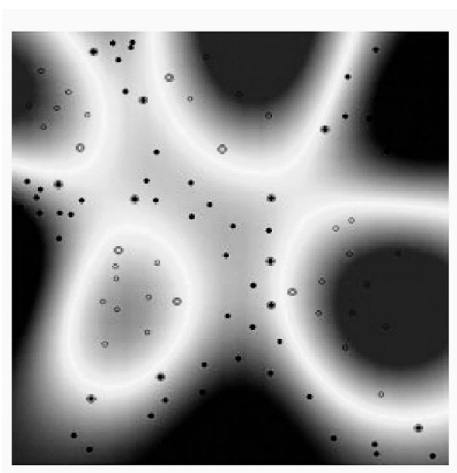


Figure 2.8: SVM avec un noyau RBF.

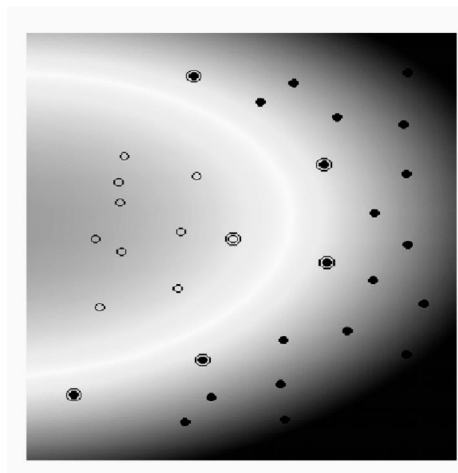


Figure 2.9: SVM avec un noyau polynomial de degré 2.

2.3.3 Avantages

- La maximisation de capacité de généralisation ; car dans la phase d'apprentissage les SVMs seront utilisés pour maximiser la marge, donc il n'y aura pas beaucoup de détérioration concernant la capacité de généralisation contrairement aux PMCs.
- L'absence des minimums locaux ; les SVMs sont formulés comme étant un problème de programmation quadratique ce qui conduit vers une solution optimale globale unique.

Le paramètre de la marge c contrôle l'erreur des entrées mal-classées. Les entrées rejetées peuvent être supprimées.

- Il y a peu de paramètres à fixer comparé à certaines autres méthodes de classification [49].

2.3.4 Inconvénients

- L'extension vers des problèmes multi-classes à cause de l'utilisation des fonctions de décision directes.
- Un temps d'apprentissage long ; car l'apprentissage des SVMs est fait par la résolution du problème dual associé, ce qui donne un nombre de variables équivalent au nombre de données d'apprentissage, donc pour un grand nombre de données d'apprentissage, la résolution du problème dual deviendra plus difficile concernant l'espace mémoire et le temps d'apprentissage.
- La tige du noyau à choisir et la sélection des paramètres. Pour la sélection du noyau approprié au problème il faut bien déterminer le paramètre de la marge c .

2.4 Le classifieur bayésien

C'est un classifieur qui a été peu utilisé dans la reconnaissance d'écriture arabe manuscrite. Nous n'avons rencontré qu'un seul système dans la littérature travaillant avec ce type de classifieur. Le tableau suivant donne les caractéristiques utilisées dans ce système, le taux de reconnaissance obtenu et la base exploitée.

Système	Caractéristiques	Taux de reconnaissance	Base utilisée
Système de Al Yousefi [9]	Les moments statistiques obtenus à partir des projections verticales et horizontales.	N'est pas mentionné.	Des caractères écrits.

Tableau 2.3: Exemple d'un OCR arabe utilisant le classifieur bayésien.

2.4.1 Théorème de Bayes

Ce classifieur est basé sur la décision bayésienne dont on cherche un modèle de production $p(x|c_i)$, qui donne pour chaque classe de formes c_i , la distribution des données qui

ne par discrimination, on cherche à approximer la probabilité à posteriori des données étant donné la

classe c_i . En pratique, cette information n'est pas toujours fournie [60].

La règle de décision bayésienne est une théorie clé en classification qui permet d'estimer la probabilité à posteriori à partir de la probabilité conditionnelle et d'émettre un vote d'appartenance de la forme traitée. Elle s'écrit sous la forme:

$$p(c_i|x) = \frac{p(x|c_i)p(c_i)}{p(x)} \quad (2.18)$$

Pour K classes, la décision bayésienne cherche la classe c_i qui maximise la probabilité à posteriori. Elle s'écrit : $c(x) = \arg \max_i p(c_i|x)$ (2.19)

Le terme $p(c_i)$ représente la probabilité a priori de la classe c_i . Il est particulièrement utile si les classes ne sont pas balancées dans l'échantillon de données considéré. La vraisemblance de l'observation $p(x)$ est une quantité constante qui peut être omise du processus de décision.

2.4.2 Types

Il existe deux types de classifieur bayésien :

- Le classifieur bayésien optimal, il permet de déterminer la classification la plus probable d'une instance selon les données d'entraînement [36]. La méthode de calcul est différente de la méthode de probabilité a posteriori maximum. La classification la plus probable d'une nouvelle instance est obtenue en combinant les prédictions de toutes les hypothèses qui sont pondérées par leurs probabilités a priori. Pour un ensemble de classes C , la valeur de classification optimale de Bayes, c , est définie comme suit: $c = \arg \max_{c_j \in C} \sum_{h_j \in H} P(c_j|h_i)P(h_i D)$ (2.20)
- Le classifieur bayésien naïf : Un classifieur bayésien naïf permet de déterminer la classification d'une instance spécifiée en terme d'attributs en présupposant que les attributs sont indépendants. Cette présupposition d'indépendance des attributs ne tient pas compte de la réalité dans beaucoup de domaines, d'où l'épithète naïve pour qualifier ce type de classifieur. Pour un ensemble de classes possibles C et une instance spécifiée par un ensemble d'attributs A , la valeur de classification bayésienne



tributs est définie comme suit:

$$(2.21)$$

2.4.3 Avantages

Le raisonnement bayésien offre plusieurs avantages :

- Le classifieur de Bayes permet d'obtenir de très bons taux de reconnaissance, à condition de connaître précisément les vraisemblances des différentes classes ainsi que leurs probabilités a priori. Dans la pratique, on ne peut obtenir que des estimations de ces probabilités.
- Il permet de supporter des données bruitées et manquantes.
- Il associe des probabilités aux prédictions, ce qui est utile dans de nombreux domaines où les connaissances sont incertaines.
- Il permet le support des connaissances a priori.
- Il fournit une approche théorique quantitative permettant l'analyse d'autres approches qui ne sont pas nécessairement basées sur des modèles probabilistes.
- Il permet le traitement incrémental des données.

2.4.4 Inconvénients

Le classifieur bayésien souffre principalement de deux inconvénients :

- Son application nécessite des probabilités dont la détermination requière typiquement de grandes quantités de données ou plusieurs connaissances a priori.
- Il nécessite un coût de calcul relativement élevé pour déterminer l'hypothèse optimale dans un cas général.

2.5 Les Modèles de Markov Cachés

Les modèles de Markov cachés (HMM) sont des modèles statistiques paramétriques de production du signal pouvant utiliser des descriptions structurelles, donc ils font partie des méthodes hybrides. L'utilisation des HMM se résume dans :

- Chaque caractère est modélisé par un HMM, un mot sera donc la combinaison des caractères et il sera modélisé par un grand HMM.
- On maintient les n meilleurs chemins.

Les HMMs sont utilisés largement dans le domaine de reconnaissance de parole et d'écriture arabe manuscrite. Le tableau suivant cite quelques systèmes utilisant les HMMs comme

ques extraites, le taux de reconnaissance obtenu et la

Système	Caractéristiques	Taux de reconnaissance	Base utilisée
Khorsheed [57]	Longueur et angle des segments, les points diacritiques, les boucles, les points de fin.	87% avec un post- traitement.	Un manuscrit ancien.
Pechwitz et Meargner [79]	Les ascendants et les descendants. L'intensité des pixels.	89%	IFN/ENIT
El-Hajj [39]	La densité des pixels, la densité des transitions, la concavité, le centre de gravité.	De 85,45% à 87,20%	IFN/ENIT
Miled, Olivier et Chriet [68]	La concavité et l'inclinaison pour les ascendants et les descendants. La densité et la forme de diacritiques.	82,5%	IFN/ENIT
Benouareth, Ennaji et Sellami [26]	Caractéristiques statistiques à partir des histogrammes. La moyenne, la variance et le mode. Les points de jonctions.	91,23%	IFN/ENIT

Tableau 2.4: Exemples d'OCRs arabes utilisant les HMMs.

2.5.1 Plan théorique

Une chaîne de Markov discrète d'ordre n est un processus stochastique discret $X = \{X = X_t \mid t=1, \dots, T\}$ avec des variables aléatoires discrètes (les réalisations de ces variables sont appelées états), vérifiant la propriété de Markov :

$$P(X_t = q_{it} \mid X_{t-1}, \dots, X_1 = q_{i1}) = P(X_t = q_{it} \mid X_{t-1} = q_{i,t-1}, \dots, X_{t-n} = q_{i,t-n}) \quad (2.22)$$

$\forall t \in [1.T]$ où $Q = \{q_1, \dots, q_n\}$ représente l'ensemble des états. C'est-à-dire l'état courant dépend de n états précédents. Dans la reconnaissance de l'écriture manuscrite, on utilise les

Un processus de Markov d'ordre 1 est stationnaire si pour tout t et

$$P(X_t = q_i | X_{t-1} = q_j) = P(X_{t+k} = q_i | X_{t+k-1} = q_j) \quad (2.23)$$

C'est-à-dire l'état courant dépend seulement de l'état précédent et il ne dépend pas du temps.

Dans ce cas on définit une matrice de probabilité de transition : $A = \{a_{ij}\}$ tel que :

$$a_{ij} = P(X_t = q_i | X_{t-1} = q_j) \quad (2.24)$$

A un instant donné d'un processus quelconque, on peut définir un HMM discret du premier ordre comme suit :

- N états représentés par $Q = \{q_1, q_2, \dots, q_N\}$.
- M symboles observés dans chaque état $V = \{o_1, o_2, \dots, o_M\}$. On désigne une observation au temps t par $O_t \in V$.
- La matrice de probabilité de transition entre les états $A = \{a_{ij}, 1 \leq i, j \leq N\}$ avec :

$$\text{➤ } a_{ij} = p(X_t = q_j | X_{t-1} = q_i) \forall i, j \in [1, N] \quad (2.25)$$

$$\text{➤ } a_{ij} \geq 0 \quad (2.26)$$

$$\text{➤ } \sum_{j=1}^N a_{ij} = 1 \quad \forall i \in [1, N] \quad (2.27)$$

- La matrice $B = \{b_j(k)\}$ de probabilité d'observation des symboles dans chacun des états du modèle, $b_j(k)$ est la probabilité d'observation de symbole quand le modèle est dans l'état j :

$$\text{➤ } b_j(k) = P(o_k | X_t = q_j) \text{ avec } k \in [1, M] \text{ et } j \in [1, N] \quad (2.28)$$

$$\text{➤ } b_j(k) \geq 0 \quad (2.29)$$

$$\text{➤ } \sum_{k=1}^M b_j(k) = 1 \quad \forall j \in [1, N] \quad (2.30)$$

- Ensemble de densité de probabilité initiale $\Pi = \{\pi_i\}, i = 1, \dots, N$

$$\text{➤ } \pi_i = P(X_0 = q_i), \forall i \in [1, N] \quad (2.31)$$

$$\text{➤ } \pi_i \geq 0 \quad (2.32)$$

$$\text{➤ } \sum_{i=1}^N \pi_i = 1 \quad (2.33)$$

Il y a deux types de HMM :

- Le modèle ergodique qui est un modèle sans contraintes où toutes les transitions d'un état vers l'autre sont possibles, c'est-à-dire $a_{ij} > 0 \quad \forall (i, j) \in [1, N]$.
- Le modèle gauche-droite qui est un modèle où il y a des contraintes sur des transitions. Seule la transition d'un état ayant un indice bas vers un état ayant un indice haut est acceptée [5]. Il existe deux sous-types de ce modèle : le modèle parallèle (figure 2.10) et le modèle séquentiel (figure 2.11). Le modèle séquentiel est souvent utilisé pour reconnaître la parole et l'écriture manuscrite car il est plus adéquat à la modélisation des signaux et il y a moins de paramètres à estimer dans la phase de l'apprentissage.

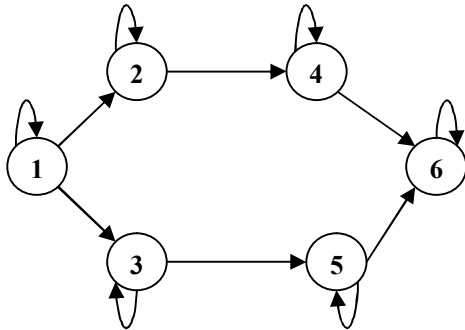


Figure 2.10: Le modèle parallèle.

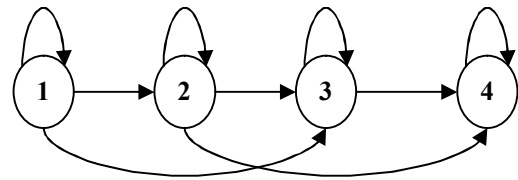


Figure 2.11: Le modèle séquentiel.

Selon le type de symboles des observations on a les HMMs discrets et les HMMs semi-continus.

- Dans les HMMs discrets les symboles des observations sont des symboles discrets dans un alphabet fini. Donc, on peut utiliser une densité discrète de probabilité à chaque état du modèle [33].
- Dans les HMMs semi-continus les symboles des observations sont des valeurs continues ou des vecteurs. Dans ce type de modèle, on met une hypothèse sur la forme de la fonction de la densité de probabilité :

$$b_j(O) = \sum_{m=1}^M c_{jm} \Pi[O, u_{jm}, U_{jm}] \quad 1 \leq j \leq N \quad (2.34)$$

Où c_{jm} : Coefficient de mixture même d'état j.

Π : Densité symétrique (par exemple Gaussian).

u_{jm} : Vecteur moyen de mixture même d'état j.

U_{jm} : Matrice de covariance de mixture même d'état j.



PDF Complete
Your complimentary use period has ended.
Thank you for using PDF Complete.

[Click Here to upgrade to Unlimited Pages and Expanded Features](#)

cités précédemment, les HMMs ont de bonnes

caractéristiques :

- Les HMMs sont robustes à la longueur des écritures. Normalement, chaque état correspond à une partie de caractère. Si le caractère est grand, il y a plusieurs échantillons dans chaque partie. Mais le HMM reste dans le même état si on ne passe pas à la partie suivante.
- La grande expérience accumulée dans l'utilisation de ces modèles pour la reconnaissance automatique de la parole.
- Les HMMs sont robustes au manque des parties d'un caractère. Si une partie d'un caractère est manquante, Le HMM va sauter l'état correspondant à la partie qui manque ce qui est appelé par la capacité d'intégration du contexte et absorption du bruit.
- Leur cadre probabilistique, qui s'accorde bien avec la nature des signaux bruités comme le cas de la parole et de l'écriture manuscrite.
- Les HMMs peuvent calculer la probabilité d'appartenance dès les premières observations. Donc, on peut éliminer les chemins qui ne sont pas prometteurs dès le début.
- En utilisant les HMMs, on peut facilement incorporer les modèles de langage comme la n-gramme et la grammaire.
- Leurs fondements théoriques qui ont permis la mise en œuvre d'algorithmes efficaces pour ajuster automatiquement et de façon itérative les paramètres du modèle.

2.5.4 Problèmes d'un HMM

- Problème 1 : étant donné le modèle $\lambda = (A, B, \pi)$, comment peut on calculer, $P(O|\lambda)$ la probabilité de l'occurrence de la séquence des observations $O = o_1, o_2, \dots, o_T$. C'est le problème de reconnaissance. Pour résoudre ce problème, on utilise la procédure aller-retour (forward-backword procedure) [36].
- Problème 2 : Etant donné le modèle $\lambda = (A, B, \pi)$, comment peut on choisir une séquence des états $Q = q_1, q_2, \dots, q_T$, afin de maximiser la probabilité $P(O, Q|\lambda)$. La solution consiste à utiliser l'algorithme de Viterbi [49].

peut ajuster les paramètres de HMM afin de résoudre le problème d'apprentissage. Pour le résoudre on utilise l'algorithme de Baum-Welch [36] utilisant des méthodes introduisant des facteurs de correction pour calculer la probabilité d'une séquence d'observation étant donné le modèle.

2.6 Le classifieur des *k*-plus proches voisins

Le classifieur des *k*-ppv est l'un des plus anciens dans le domaine de la reconnaissance de forme, c'est un classifieur non-paramétrique ne nécessitant pas une phase d'apprentissage. Le principe de cette méthode est très simple, puisqu'il consiste à attribuer au prototype que l'on cherche à identifier l'étiquette de la classe la plus représentée parmi les *k*-plus proches voisins des prototypes de la base d'apprentissage. Le voisinage est déterminé au sens d'une distance dans l'espace de représentation. La distance euclidienne, la distance de Hamming et la distance du maximum sont couramment utilisées par le classifieur de type *k*-plus proches voisins [55].

Ce classifieur a été utilisé par des chercheurs dans la reconnaissance d'écriture arabe manuscrite, ce qui est détaillé dans le tableau suivant :

Systeme	Caractéristiques	Taux de reconnaissance	Base utilisée
Mozaffari, Faez et Ziaratbah [69]	La moyenne, la variance, les points de fin, les points d'intersection.	94,44%	280 images de huit chiffres pour l'apprentissage. 200 images pour le test.
Hashizume, Yeh et Rosenfeld [65]	Les ascendants et les descendants.	N'est pas mentionné.	N'est pas mentionnée.
Burrow [29]	Le rapport entre la hauteur et la largeur, la moyenne.	81% pour une normalisation de 16×64 pixels.	IFN/ENIT.

*Tableau 2.5: Exemples d'OCRs arabes exploitant les *k*-ppv.*

classifieur le plus simple théoriquement. Donnant $V = \{v_1, v_2, \dots, v_v\}$ sera un ensemble de références étiquetées contenant v points dans R^n , indiqués comme prototypes. Les prototypes sont étiquetés dans les classes c ; c'est-à-dire pour chaque $v_i \in V$, on connaît son étiquette de classe $l(v_i) \in \Omega$. Dans la conception classique des plus proches voisins, V est la totalité de Z . Pour classifier une entrée x , les k -plus proches prototypes sont préservés à partir de V avec leurs étiquettes de classe. L'entrée x aura l'étiquette de la classe la plus représentative parmi les k -plus proches voisins [55].

Pour arriver à cette méthode de classification, on fixe k et N dans l'équation suivante :

$$p(x) \approx \frac{k}{NV_R} \quad (2.35)$$

Et on déduit la variable V_R . Supposant qu'on va travailler sur la distance euclidienne, R est l'hypersphère contenant exactement k éléments de l'ensemble de références V . la classe inconditionnelle sera approximée comme suit :

$$p(x) \approx \frac{k}{NV_R} \quad (2.36)$$

K_i dénote le nombre des éléments dans R des classes w_i , la classe conditionnelle pour

$$w_i, i = 1, \dots, c, \text{ approximée dans } R \text{ est : } p(x|w_i) \approx \frac{k_i}{N_i V_R} \quad (2.37)$$

Où N_i est le nombre des éléments des classes w_i dans Z . Les probabilités antérieures sont

$$\text{habituellement estimées par : } \hat{P}(w_i) = \frac{N_i}{N} \quad (2.38)$$

On utilise cette équation pour estimer $P(w_i)$, les probabilités postérieures sont obtenues

$$\text{comme suit : } P(w_i|x) = \frac{p(w_i|x)P(w_i)}{p(x)} \approx \frac{\frac{k_i N_i}{N_i V_R N}}{\frac{k}{NV_R}} \quad (2.39)$$

$$\text{En conséquence : } P(w_i|x) \approx \frac{k_i}{k} \quad (2.40)$$

L'erreur minimale du classifieur bayésien utilisant l'approximation en haut attribuera x à la classe ayant la plus grande probabilité postérieure, c'est-à-dire la classe la plus représentative parmi les k -plus proches voisins de x . La région R et le volume V_R , respectivement, sont



PDF Complete
Your complimentary use period has ended.
Thank you for using PDF Complete.

[Click Here to upgrade to Unlimited Pages and Expanded Features](#)

ensemble de données Z . On peut dire que le k -ppv est un classifieur bayésien optimal quand $N \rightarrow \infty$ et $V_R \rightarrow 0$. L'expression $V_R \rightarrow 0$ est équivalente à $k \rightarrow \infty$ et $k/N \rightarrow 0$. C'est-à-dire le taux d'erreur du classifieur k -ppv,

$P_{k\text{-ppv}}$ satisfait :

$$\lim_{\substack{N \rightarrow \infty \\ k \leftarrow \infty \\ k/N \rightarrow 0}} P_{k\text{-ppv}} = P_B \quad (2.41)$$

Où P_B est l'erreur de Bayes. Quand k est 1 (la règle de plus proche voisin, dénotée 1-ppv), et $N \rightarrow \infty$, le taux d'erreur $P_{1\text{-ppv}}$ est délimité par le double de celui de Bayes ; c'est-à-dire :

$$P_{1\text{-ppv}} \leq 2P_B \quad (2.42)$$

Notant que l'utilisation de plusieurs métriques aura un impacte sur les régions. Sans se soucier de la métrique utilisée, la règle de k -ppv est appliquée de la même manière. Le type de métrique ne change pas les propriétés asymptotiques des k -ppv car la forme de R n'est pas fixée.

2.6.2 Avantages

- La méthode de k -ppv présente les avantages d'être à la fois très simple à implémenter et d'application très générale car non-paramétrique.
- L'absence d'apprentissage fait du k -ppv un parfait candidat pour les classifications qui doivent être constamment ajustées ou révisées. En fait, la mise à jour de la classification ne peut pas être plus simple. Il n'y a qu'à étiqueter et indexer les nouvelles classes, et enlever les classes nuisibles.

2.6.3 Inconvénients

- Le k -ppv nécessite un nombre important d'échantillons dans la base d'apprentissage afin de remplir convenablement l'espace des caractéristiques, notons qu'un espace bien rempli engendre deux inconvénients qui ont longtemps rendu cette méthode inusitée. Ce sont la taille de la mémoire nécessaire au chargement de la base d'apprentissage et le temps nécessaire au calcul des distances entre le prototype à identifier et les différents échantillons de la base d'apprentissage.
- Il est intolérant aux bruits et à la redondance des caractéristiques.
- Il n'existe pas une mesure de similarité naturelle pour les caractéristiques nominales évaluées.



Your complimentary
use period has ended.
Thank you for using
PDF Complete.

[Click Here to upgrade to
Unlimited Pages and Expanded Features](#)

manipuler les valeurs manquantes.

Information concernant la structure des données.

2.7 Conclusion

Dans ce chapitre, nous avons présenté un aperçu des principaux classifieurs utilisés pour la reconnaissance de l'écriture arabe.

Les réseaux de neurones, et plus précisément les PMCs représentent le type de classifieur le plus employé, ceci peut être justifié par leur simplicité qui n'affecte en rien leur efficacité. L'utilisation des HMMs a permis aussi d'obtenir des résultats intéressants pour certaines applications grâce à leur capacité d'intégration du contexte et d'absorption du bruit. Les SVMs, le classifieur bayésien et les K-ppv's ont été peu exploités dans les OCRs arabes et représentent toujours un terrain fertile pour de futures recherches.

Nous avons par contre introduit un nouveau classifieur non-encore utilisé dans le domaine de l'écriture arabe qui est le réseau Fuzzy ART dont on s'est servi pour implémenter notre système. Nous y reviendrons au quatrième chapitre, où nous nous attarderons sur ses qualités de classification.

C'est la première étape de son processus de sélection dans l'espace de représentation, les informations nécessaires à l'application. Cette sélection passe souvent par l'élimination du bruit dû aux conditions d'acquisition, par la normalisation des données, ainsi que par la suppression de la redondance. Ceci ayant pour objectif de fournir des données de qualité à la phase suivante qui est l'extraction de primitives. Le prétraitement englobe plusieurs opérations dont nous citons :

3.1.1 Normalisation

La normalisation des caractères est une opération très importante, l'image de caractère est transformée en un plan standard avec une dimension prédéfinie de telle manière à ce qu'elle donne une représentation d'une dimensionnalité fixe pour la classification. Le but de la normalisation est de réduire la variation entre les formes des classes afin de faciliter le processus d'extraction de primitives. On distingue des méthodes linéaires et des méthodes non-linéaires [33].

Si la largeur et la hauteur du caractère original est respectivement W_1 et H_1 , la largeur et la hauteur du caractère normalisé seront W_2 et H_2 , et la dimension du plan standard (normalisé) est L . Le plan standard est habituellement considéré comme un carré dont les mesures sont 32×32 ou 64×64 pixels ; bien entendu d'autres mesures peuvent être utilisées.

3.1.2 Binarisation

La binarisation est sans doute le traitement de base de la plupart des systèmes d'analyse d'image. Certains, apprécient dans cette transformation de l'image, la diminution de la qualité d'informations et donc des coûts de stockage ; d'autres la préféreront pour obtenir des images plus proches de l'original. Elle permet de passer d'une image de niveaux de gris à une image binaire composée de 2 valeurs 0 et 1, plus simple à traiter.

On peut trouver dans la littérature plusieurs centaines de méthodes de binarisation. Toutefois, elles peuvent toutes être regroupées en deux grandes familles:

- Les méthodes de seuillage global où le même seuil est utilisé pour toute l'image, tels que tous les pixels de l'image (ou d'une zone d'intérêt) qui ont une intensité lumineuse supérieure à ce seuil seront définis comme blanc et les autres comme

classe ont l'avantage d'être extrêmement rapides leurs faiblesses dans le cas où l'intensité lumineuse de l'image n'est pas constante, ce qui est malheureusement très fréquent pour des documents numérisés.

- Les méthodes de seuillage local où le seuil est propre à chacun des pixels de l'image permettent d'obtenir une qualité de seuillage bien plus intéressante que les méthodes à seuillage global, puisqu'ils traitent de manière homogène les zones qui ont subi un éclairage différent, mais demandent plus d'effort de calcul [89].

3.1.3 Redressement

Le redressement est une opération fréquente en analyse de documents, souvent dû à un mauvais positionnement du document sur le scanner, conduisant à une inclinaison de l'image [49].

Il y a plusieurs méthodes pour détecter l'inclinaison dans la page, certaines comptent sur la détection des composantes connexes en trouvant la moyenne des angles connectant leurs centroids ; d'autres utilisent l'analyse de profil de projection, où la page du document est projetée sur plusieurs angles, et les variances dans le nombre des pixels noirs pour chaque ligne sont déterminées.

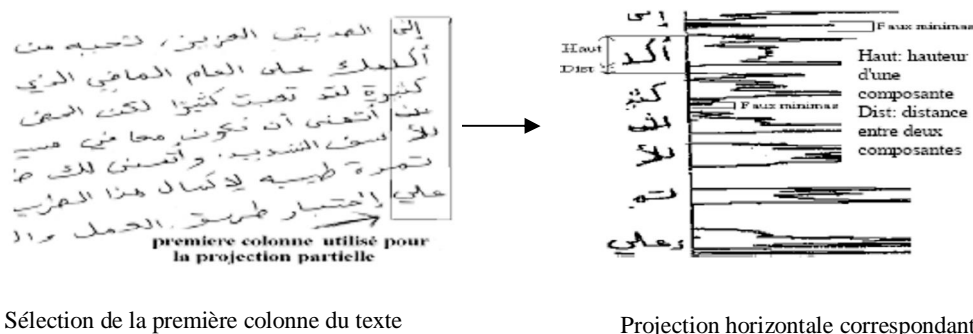


Figure 3.1: Profil de projection d'un texte arabe.

3.1.4 Lissage

C'est un traitement local qui définit une nouvelle valeur de pixel en examinant son voisinage. Il existe plusieurs opérateurs de lissage inspirés de la morphologie mathématique, ils consistent à remplacer la valeur de chaque point par une expression en fonction de ses voisins. Dans le cas de l'écriture, le lissage sert à éliminer les trous (bouchage) et les points isolés (nettoyage). Un point est mis à 0 s'il n'a pas assez de

à 1 dans le cas contraire. On distingue deux types

- Le lissage par des filtres linéaires dont la valeur du pixel de sortie est une combinaison linéaire des valeurs des pixels voisins en entrée. Les principaux filtres existants sont :
 - Le filtrage par la moyenne, qui consiste à remplacer chaque pixel par la valeur moyenne de ses voisins (y compris le pixel lui-même). Cette méthode a pour effet de modifier les niveaux de gris trop différents de leurs voisins.
 - Le filtrage gaussien quand à lui, est un opérateur de lissage utilisé pour estomper les détails et le bruit [33].

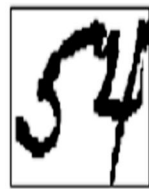


Image originale

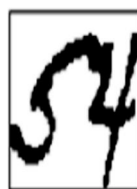


Image lissée

$$\frac{1}{9} \times \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline \end{array}$$

Figure 3.2: Lissage par un filtre moyen. **Figure 3.3:** Exemple d'un filtre moyen.

- Le lissage par des filtres non-linéaires, qui comporte le filtre médian et le filtrage par le max.
 - Le filtre médian réalisant un lissage de l'image un peu plus performant que le filtre moyen concernant les détails dans l'image. Chaque pixel est traité en considérant ses voisins sur un voisinage donné, le pixel lui-même et ses voisins forment alors un ensemble dont on calcule la médiane, le pixel sera alors remplacé par cette valeur médiane.
 - Le filtrage par le max, est un filtre de lissage qui adoucit les pixels isolés ayant un niveau de gris très différent des niveaux de gris de leur voisinage. Sa particularité est de bien préserver les contours très marqués.

3.1.5 Squelettisation

Dans le plan continu, le squelette d'une forme est un ensemble de lignes passant en son milieu. C'est la notion d'axe médian d'une forme continue qui est définie comme suit : Soit R une forme plane, B son contour et P un point de R. Un plus proche voisin

el qu'il n'existe aucun autre point sur B dont la distance PM. Si P a plus d'un plus-proche voisin, P est

appelé point squelette de R. L'union de tous les points squelettes est appelée squelette ou axe médian de R. D'après Hilditch [51] cet axe doit avoir les propriétés suivantes :

- Epaisseur : le squelette est mince. Idéalement, il n'est constitué que de lignes ayant comme épaisseur un pixel.
- Position : le squelette est idéalement positionné au centre des formes. Malheureusement le centre d'une forme discrète n'est pas unique.
- Connectivité ou homotopie : l'axe médian possède exactement le même nombre de composantes connexes, et pour chaque composante, le même nombre de trous que la forme initiale.
- Stabilité : dans le cas d'un calcul du squelette par amincissement, tout ou partie du squelette ne doit pas pouvoir être érodé à nouveau. Aussitôt qu'un pixel est considéré comme point squelette, on ne doit plus pouvoir remettre cette désignation en question ; cette propriété évite les érosions de squelette par les extrémités.
- Invariance par translation et rotation.
- Réversibilité de l'axe médian pondéré ; c'est-à-dire que la forme doit pouvoir être reconstruite à partir du squelette et des rayons des boules maximales.

Il est admis que ces conditions sont pour la plupart contradictoires. Par exemple on ne peut pas avoir un squelette homotypique qui respecte la condition d'épaisseur stricte (1 pixel de largeur). Il faut donc faire des compromis et une comparaison de différents algorithmes est nécessaire pour choisir celui qui convient le mieux à une application donnée. Il y a plusieurs algorithmes de squelettisation tels que ceux de Hilditch, Yokoi et Coll, Gonzales et Wood, Rutovitz, Deutsch, Zhang et Suen, et de Zhang et Wang [19].

3.2 Segmentation

La segmentation correspond aux différentes opérations qui doivent être effectuées pour extraire, à partir de l'image du document considéré, les différentes unités de base afin de réaliser leurs reconnaissances. Les trois premiers niveaux de segmentation, qui se terminent par la localisation des mots, ont été abordés dans le premier chapitre. Nous n'aborderons donc ici que le quatrième niveau, qui consiste à localiser les lettres



Your complimentary
use period has ended.
Thank you for using
PDF Complete.

[Click Here to upgrade to
Unlimited Pages and Expanded Features](#)

et extraits. Cette étape n'est pas toujours nécessaire la forme globale des mots. Elle s'avère cependant indispensable lorsque la taille du vocabulaire est importante et qu'une approche analytique est utilisée. Il existe deux techniques permettant la mise en œuvre de la segmentation :

3.2.1 *Segmentation implicite*

Les méthodes de segmentation implicite s'inspirent des approches utilisées dans le domaine de la reconnaissance de la parole, où le signal est divisé en intervalles de temps réguliers. Elles permettent d'assurer un taux de présence important des points de liaison entre les lettres parmi les points proposés [89]. Le système recherche dans l'image, des composantes ou des groupements de graphèmes qui correspondent à ses classes de lettres. Classiquement, il peut le faire de deux manières, soit par fenêtrage, soit par recherche de primitives.

- Dans les méthodes par fenêtrage, le principe est d'utiliser une fenêtre mobile de largeur variable pour trouver des séquences de points de segmentation potentielles qui seront confirmées ou non par la reconnaissance de caractères. En variant la taille de la fenêtre et sa position, on obtient plusieurs séquences de points de segmentation qui seront analysés par le système de reconnaissance.
- Pour les méthodes par recherche de primitives, il s'agit de détecter les combinaisons de primitives qui donneront la meilleure reconnaissance possible.

3.2.2 *Segmentation explicite*

Cette approche souvent appelée dissection, est antérieure à la reconnaissance. Les hypothèses sur les caractères sont déterminées à partir des informations de bas niveau présentées sur l'image. Ces hypothèses sont définitives, et doivent être d'une grande fiabilité, car la moindre erreur de segmentation remet en cause la totalité des traitements antérieurs.

Les méthodes de segmentation explicite s'appuient sur une analyse morphologique du mot manuscrit pour localiser des points de segmentation potentiels. Elles sont particulièrement adaptées à l'analyse de la représentation bi-dimensionnelle et donc plus souvent utilisées dans les systèmes de reconnaissance hors-ligne.

Extraction de primitives est une étape de grande importance, si elle est mal-entamée, il sera difficile, voire impossible d'effectuer une reconnaissance efficace. L'objectif de cette étape est la sélection de l'information pertinente qui se trouve noyée dans la masse d'information brute acquise [73]. La problématique de cette étape a pour origine le risque de perte d'information signifiante pour la reconnaissance. La minimisation de ce risque est conditionnée par deux dilemmes liés chacun à un paradoxe de causalité ; il s'agit du dilemme de la réduction et du dilemme de la segmentation.

Ici nous allons décrire une collection des principales méthodes d'extraction de primitives, utilisées généralement par des OCRs arabes.

3.3.1. Les moments

Ils servent à extraire des propriétés de l'image comme l'espace de la forme, le centre de masse, etc. La définition générale des fonctions de moment m_{pq} d'ordre $(p+q)$ pour la fonction $f(x,y)$ d'intensité d'une image continue $X \times Y$ est :

$$m_{pq} = \int \int \phi_{pq}(x, y) f(x, y) dx dy \quad (3.1)$$

Où p et q sont des entiers appartenant à l'intervalle $[0, \hat{O}]$, x et y sont les coordonnées d'un point de l'image et $\phi_{pq}(x, y)$ est la fonction de base. Les propriétés comme l'orthogonalité de la fonction de base sont irritées par les fonctions de moment [61].

Pareillement, la définition générale pour une image digitale $X \times Y$ peut être obtenue en remplaçant l'intégral par des sommes : $m_{pq} = \sum_y \sum_x \phi_{pq}(x, y) f(x, y)$ (3.2)

Où p et q sont des entiers appartenant à l'intervalle $[0, \hat{O}]$ représentant l'ordre et $\phi_{pq}(x, y)$ est la fonction de base. Les moments les plus utilisés dans les OCRs arabes sont :

- Les moments géométriques. Le moment géométrique d'ordre $(p+q)$ de $f(x,y)$ est

$$\text{défini par : } m_{pq} = \int \int x^p y^q f(x, y) dx dy \quad (3.3)$$

Cette méthode est coûteuse car la fonction de base du moment géométrique est non-orthogonale.

nique a défini un ensemble complet orthogonal
 lexes à travers l'espace des coordonnées polaires à

$$V_{nm}(x, y) = V_{nm}(\rho, \theta) = R_{nm}(\rho) \rho^{jm\theta} \quad (3.4)$$

Où $j = \sqrt{-1}$, $n \geq 0$, m est un entier positif ou négatif, ρ est la plus petite distance à partir de l'origine au pixel (x, y) , θ est l'angle entre le vecteur ρ et l'axe X respectant la direction inverse de l'horloge et $R_{nm}(\rho)$ est le polynôme orthogonal radial [33].

3.3.2 L'histogramme

L'histogramme est un tableau affichant la distribution d'un ensemble de pixels d'une image $f(x, y)$ d'un niveau de gris donné. L'amplitude de l'histogramme est une table $H(k)$ qui montre le nombre de pixels ayant un niveau de gris de valeur k . La définition de l'amplitude d'un histogramme pour une image digitale (x, y) est donnée par :

$$H(k) = \text{cardinal}_{k \in [0, k_{\max}], (x, y) \in [(0, 0), (x-1, y-1)]} \{f(x, y) | f(x, y) = k\} \quad (3.5)$$

Où $\text{cardinal} \{ \}$ est le nombre de pixels dans l'ensemble donné. La figure 3.4 montre l'histogramme de projection horizontale d'un mot tiré de la base IFN/ENIT.

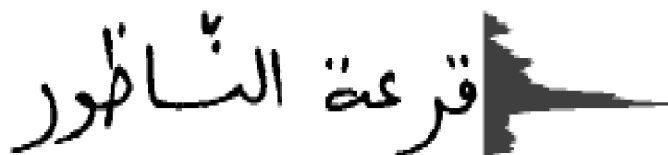


Figure 3.4: Histogramme de projection horizontale.

3.3.3 Les primitives de direction

Les caractères comprennent des traits qui sont : des lignes orientés, des courbes ou poly-lignes. L'orientation ou la direction des traits joue un rôle important dans la différenciation entre les caractères ; on cite les techniques suivantes :

- Chaîne de code, où les pixels de contour d'une image binaire sont communément codés en chaîne de code quand ils sont tracés dans un certain ordre [75]. Chaque pixel de contour pointe vers son successeur dans l'une des 8 directions, comme illustre la figure 3.5.

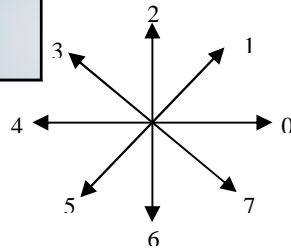


Figure 3.5: Les huit directions de chaîne de code.

- Vectorisation, qui est une procédure visant à représenter un caractère par un ensemble de segments droits, plutôt que par un ensemble de pixels. Cette représentation est vectorielle. Ici, chaque ensemble de pixels continus est remplacé par un segment ; chaque segment est codé par les 4 coordonnées de 2 extrémités (x_0, y_0) , (x_n, y_n) . Le critère est une erreur moyenne de représentation qui est définie comme la valeur des carrés des distances entre un segment et tous les pixels qu'il remplace.

3.3.4 La transformée de Hough

Elle est capable de détecter les lignes droites, courbes, et n'importe quelle forme particulière qui peut être définie par des équations paramétriques. Dans son principe, cette méthode planifie les points de l'image à partir de l'espace d'image vers un espace paramétrique, et après elle extrait les primitives. Le point dans l'espace de Hough est une ligne à certain angle et certaine distance de l'origine, la valeur de ce point est une expression de comment plusieurs coordonnées de point se trouvent sur cette ligne, donc, on utilise la TH comme étant le chemin de trouver les lignes droites dans un ensemble donné. Les points dans l'espace de Hough ayant la grande valeur traduisent le fait que plusieurs coordonnées de point se trouvent en direction de la ligne correspondante dans l'espace cartésien. Cela peut être utilisé pour détecter la ligne de base.

3.4 Reconnaissance

Cette étape consiste à déterminer la classe d'appartenance de la forme considérée en utilisant un classifieur. Ce dernier peut être vu comme une fonction de décision qui associe à une entrée e décrite dans un espace E une sortie s décrite dans un espace S . Les entrées du système représentent les formes à reconnaître et les sorties les classes



Your complimentary
use period has ended.
Thank you for using
PDF Complete.

[Click Here to upgrade to
Unlimited Pages and Expanded Features](#)

appartiennent. Le processus de reconnaissance est en deux phases. La première phase correspond à la conception du classifieur, qui est achevée par un apprentissage automatique de ce dernier. Tandis que la deuxième phase consiste à évaluer les performances du système à partir d'un ensemble de données appelé base de teste.

3.4.1 Apprentissage

Etape durant laquelle le système va acquérir et organiser la connaissance en classes ou modèles de référence ; ce qui est constitué par des groupements de formes ayant les mêmes caractéristiques. Ainsi, plus les caractéristiques sont robustes et différentes d'un ensemble de formes à l'autre, plus il est facile de séparer les formes en familles homogènes. Par contre, plus les propriétés sont redondantes, contradictoires et incomplètes, plus les répartitions en familles sont approximatives et plus il est difficile pour le système de décider à quelle famille appartient une forme donnée. On se trouve donc confronté à un compromis entre la redondance et l'altération des données, d'une part et la fiabilité de sa modélisation, de l'autre. Il existe deux types d'apprentissage :

- L'apprentissage supervisé, où les différentes familles des formes sont connues a priori ; ici, la tâche d'apprentissage est guidée par un superviseur qui indique pour chaque forme entrée, le nom de la famille lui correspondant. L'apprentissage se déroule de la manière suivante : on choisit un sous-ensemble de l'ensemble des formes à reconnaître, chacune de ces formes est analysée et décrite. Le concepteur indique le nom de classe souhaitée pour la forme entrée, c'est-à-dire la classe dans laquelle il souhaite la voir rangée. La tâche de l'apprentissage consiste alors à analyser les ressemblances entre les formes d'une famille et les dissemblances entre les formes de familles différentes pour en déduire les classes de l'apprentissage avec les meilleurs séparatrices possibles. Les paramètres décrivant cette partition sont stockés dans une base d'apprentissage. Plus cette base est grande et plus les séparatrices peuvent être déterminées avec précision et meilleurs sont les résultats.
- L'apprentissage non-supervisé ; qui consiste, à partir d'échantillons de références et de règles de regroupement, de distinguer automatiquement les classes par le système, sans intervention du concepteur. Ce mode d'apprentissage nécessite un

et des règles de construction précises et non-ambiguës, une bonne formation des classes [85].

3.4.2 Classification

A partir du vecteur descriptif d'une entrée, on recherche, parmi les classes d'apprentissage connues, celles qui sont les plus proches. La notion de proximité a un sens différent en fonction de la nature de la représentation et du type de méthode. La classification peut conduire à un succès si la réponse est unique, c'est-à-dire lorsqu'il n'existe qu'une seule classe qui répond à la description de la forme ; mais elle peut conduire à une confusion si la réponse est multiple, c'est-à-dire que plusieurs modèles correspondent à la description, ce qui dénote une certaine ambiguïté volontaire ou non au sein de la phase de l'apprentissage. Dans les deux cas la classification peut être accompagnée d'une mesure de vraisemblance, appelée aussi score ou taux de reconnaissance. La classification peut aussi conduire à un rejet de la forme si aucune des classes apprises ne correspond à la description de la forme entrée.

L'évaluation des performances d'un classifieur est une phase importante dans le processus de conception et mise en œuvre de ce dernier, dans la mesure où elle permet de savoir si le système est suffisamment performant pour l'application visée. Elle permet aussi de comparer le système réalisé avec d'autres systèmes.

Disposant d'une base de test, on procède généralement à l'évaluation du comportement d'un classifieur à partir de certaines mesures qui peuvent être globales, semi-globales ou locales.

- La mesure globale : en reconnaissance de formes, les taux de reconnaissance (TL), d'erreur (TE) et de rejet (TR) correspondent à des estimateurs classiques. Ils sont définis par:

$$TL = \frac{\text{Nombre d'entités bien reconnues}}{\text{Nombre total d'entités présentées au classifieur}} \quad (3.6)$$

$$TE = \frac{\text{Nombre d'entités mal-reconnues}}{\text{Nombre total d'entités présentées au classifieur}} \quad (3.7)$$

entités rejetées

$$TR = \frac{\sum_{i=1}^N TR_i}{N} \quad (3.8)$$

Nombre total d'entités présentées au classifieur

- Performances semi-globales : elle concerne la matrice de performances semi-globale MAT présentée par l'équation 3.9, où TL_i^k correspond au nombre d'entités de la classe C_i pour lesquelles la bonne solution est placée dans les k premières propositions rapportées au nombre total d'entités de la classe C_i ($i, k \in [1, N]$). TR_i est le nombre d'entités de chaque classe C_i rejetée. La matrice MAT contient en réalité, des taux de reconnaissance cumulés:

$$MAT = \begin{pmatrix} TL_1^1 & TL_1^2 \dots & TL_1^N & TR_1 \\ \dots & \dots & \dots & \dots \\ TL_i^1 & TL_i^2 \dots & TL_i^N & TR_i \\ \dots & \dots & \dots & \dots \\ TL_N^1 & TL_N^2 \dots & TL_N^N & TR_N \end{pmatrix} \quad (3.9)$$

- Performances locales : une analyse encore plus fine du comportement du classifieur peut être obtenue par des mesures locales comme les matrices de confusion où chaque matrice MC^k est composée de N lignes et N+1 colonnes, avec N représentant le nombre de classes. Les éléments de la diagonale sont les taux de reconnaissance TL_i^k qui correspondent au nombre d'entités de la classe C_i pour lesquelles la bonne solution apparaît dans les k premières propositions ; la moyenne de ces taux donne le taux de reconnaissance global TL_k . La dernière colonne N+1 est réservée aux taux de rejet TR_i qui correspondent au nombre d'entités rejetées pour chaque classe C_i ; la moyenne de ces taux est TR.

$$MC^k = \begin{pmatrix} TL_1^k \dots & TE_{1j}^k \dots & TE_{iN-1}^k & TR_1 \\ \dots & \dots & \dots & \dots \\ TE_{i1}^k \dots & TL_i^k \dots & TE_{iN-1}^k & TR_i \\ \dots & \dots & \dots & \dots \\ TE_{N1}^k & TE_{Nj}^k \dots & TL_N^k & TR_N \end{pmatrix} \quad (3.10)$$

Le post-traitement est effectué quand le processus de reconnaissance aboutit à la génération d'une liste de mots possibles, éventuellement classés par ordre décroissant de vraisemblance, afin de prendre compte des nombreuses relations d'ordre : logique, lexical et syntaxique entre les différents composants d'un document [59]. On peut dire que le rôle de cette étape consiste à corriger les erreurs de l'étape de reconnaissance, en tenant compte de tous les types d'informations propres au document :

- L'information lexicale : le contrôle lexical vérifie la présence d'un mot dans un lexique de référence. Dans notre cas la reconnaissance sera effectuée au niveau du mot, donc le mot présent dans le lexique sera attribué par une valeur de confiance qui dépend de celle attribuée par l'étape de classification.
- L'information syntaxique : l'utilisation d'un ensemble de règles grammaticales permet de prendre en compte le contexte syntaxique d'un document pour confirmer ou infirmer la séquence de mots proposés.
- L'information sémantique : c'est le fait de mettre en corrélation des résultats obtenus sur des champs différents de l'image, mais codant la même information ; c'est le cas des chèques où on trouve une zone pour les montants numériques et une autre zone pour les montants littéraux.

3.6 Conclusion

L'objectif de ce chapitre était de fournir une vision détaillée sur les possibles structures que peut avoir un système de reconnaissance d'écriture manuscrite hors-ligne. Nous avons aussi, expliqué les différentes phases composant ce type de système à savoir : le prétraitement, la segmentation, l'extraction des caractéristiques, la reconnaissance et le post-traitement, en détaillant les techniques utilisées dans chacune d'elles.

Les deux étapes de segmentation et de poste-traitement ne seront pas traitées lors de l'implémentation de notre système. Ceci est argumenté par l'utilisation d'une approche globale ne nécessitant pas un processus de segmentation d'une part, et la non-disponibilité d'un dictionnaire de validation d'une autre.

tées, elles contiennent des informations parasites et n'explicitent pas les informations utiles pour la reconnaissance. Le rôle du prétraitement est de préparer ces données à la phase d'extraction de primitives en éliminant le bruit, en corrigeant les erreurs, en homogénéisant, normalisant et réduisant les données. On a appliqué dans cette étape les opérations suivantes :

4.1.1 Normalisation

Elle est considérée comme étant la phase la plus importante dans l'étape de prétraitement, elle consiste à appliquer des algorithmes et des opérations afin d'avoir la même taille pour toutes les images du mot.

On a appliqué dans notre système une normalisation linéaire [33] en se basant sur un algorithme comportant trois étapes:

- Calculer la matrice M de dispersion de la forme, considérant la fonction d'image digitale $f(x,y)$ pour laquelle $f(x,y)=1$ indique que (x,y) est un pixel objet et $f(x,y)=0$ indique que (x,y) est un pixel de contour. La matrice M est l'élément clé dans le processus de normalisation. Pour la calculer on doit avoir d'abord le centroïde de la

$$\bar{x} = \sum_x \sum_y x.f(x, y) / A \quad (4.1)$$

$$\bar{y} = \sum_x \sum_y y.f(x, y) / A \quad (4.2)$$

$$\text{Où } A \text{ est le nombre total des pixels objets: } A = \sum_x \sum_y f(x, y) \quad (4.3)$$

$$\text{La matrice de dispersion est une matrice d'ordre } 2 \times 2: M = \begin{bmatrix} m_{1,1} & m_{1,2} \\ m_{2,1} & m_{2,2} \end{bmatrix} \quad (4.4)$$

Où ses éléments sont définis comme suit:

$$m_{1,1} = \left(\sum_x \sum_y x^2 . f(x, y) / A \right) - \bar{x}^2 \quad (4.5)$$

$$m_{1,2} = m_{2,1} = \left(\sum_x \sum_y x.y.f(x, y) / A \right) - \bar{x}.\bar{y} \quad (4.6)$$

$$m_{2,2} = \left(\sum_x \sum_y y^2 . f(x, y) / A \right) - \bar{y}^2 \quad (4.7)$$

- Changer l'origine des axes des coordonnées vers le centre de la forme où on doit utiliser une matrice orthogonale composée de deux vecteurs normalisés E_1 et E_2 , afin

d'abord la valeur de λ_1 et λ_2 .

$$\lambda_1 = \frac{m_{1,1} + m_{2,2} + \sqrt{(m_{1,1} - m_{2,2})^2 + 4m_{1,2}^2}}{2} \quad (4.8)$$

$$\lambda_2 = \frac{m_{1,1} + m_{2,2} - \sqrt{(m_{1,1} - m_{2,2})^2 + 4m_{1,2}^2}}{2} \quad (4.9)$$

Les vecteurs normalisés E_1 et E_2 sont donnés par :

$$E_1 = \begin{bmatrix} e_{1x} \\ e_{1y} \end{bmatrix} = \begin{bmatrix} \frac{m_{1,2}}{\sqrt{(\lambda_1 - m_{1,1})^2 + m_{1,2}^2}} \\ \frac{\lambda_1 - m_{1,1}}{\sqrt{(\lambda_1 - m_{1,1})^2 + m_{1,2}^2}} \end{bmatrix} \quad (4.10)$$

$$E_2 = \begin{bmatrix} e_{2x} \\ e_{2y} \end{bmatrix} = \begin{bmatrix} \frac{m_{1,2}}{\sqrt{(\lambda_2 - m_{1,1})^2 + m_{1,2}^2}} \\ \frac{\lambda_2 - m_{1,1}}{\sqrt{(\lambda_2 - m_{1,1})^2 + m_{1,2}^2}} \end{bmatrix} \quad (4.11)$$

Maintenant la matrice R peut être construite à partir de E_1 et E_2 par :

$$R = \begin{bmatrix} E_1^T \\ E_2^T \end{bmatrix} = \begin{bmatrix} e_{1x} & e_{1y} \\ e_{2x} & e_{2y} \end{bmatrix} \quad (4.12)$$

Tan que M est réelle et symétrique et E_1 et E_2 sont orthogonaux l'un à l'autre donc R est orthogonale ce qui garde la direction des nouveaux axes. Chaque position de pixel objet (x,y) aura une nouvelle position (x',y') donnée par :

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = R \cdot \begin{bmatrix} x - \bar{x} \\ y - \bar{y} \end{bmatrix} \quad (4.13)$$

La matrice de dispersion changée M' est donnée par :

$$M' = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \quad (4.14)$$

- Changer l'échelle de base où l'échelle des deux axes sera changé selon λ_1 et λ_2 . Pour un pixel objet (x',y').

La nouvelle position (x'',y'') est obtenue à travers une transformation définie par :

$$W = \begin{bmatrix} x'' \\ y'' \end{bmatrix} = W \cdot \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} c/\sqrt{\lambda_1} & 0 \\ 0 & c/\sqrt{\lambda_2} \end{bmatrix} \cdot \begin{bmatrix} x' \\ y' \end{bmatrix} \quad (4.15)$$

où c représente une constante.

es exemples de la base IFN/ENIT avant et après la



Image avant normalisation de taille 414×106 pixels



Image après normalisation de taille 400×100 pixels

Figure 4.1: Normalisation du mot "شغال".

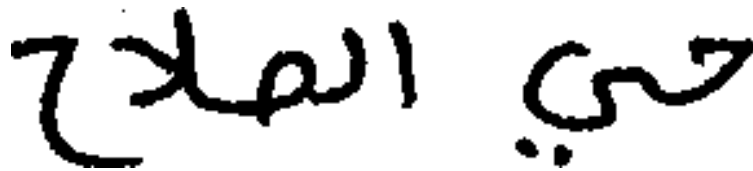


Image avant normalisation de taille 288×62 pixels

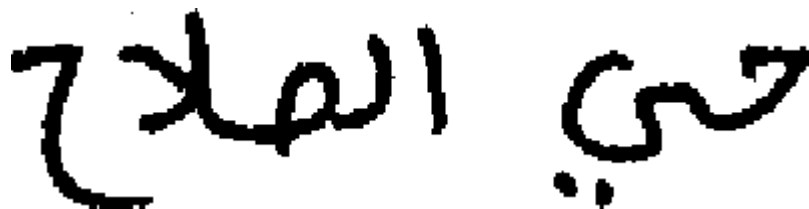
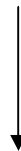


Image après normalisation de taille 400×100 pixels

Figure 4.2: Normalisation du mot "حي الصلاح".



Image avant normalisation de taille 788×91 pixels



Image avant normalisation de taille 400×100 pixels

Figure 4.3: Normalisation du mot "سیدی ابراہیم الزہار".

4.1.2 Squelettisation

L'opération de squelettisation consiste dans le cas particulier de l'écriture, à éliminer l'épaisseur du trait ou plutôt à l'amincir jusqu'à l'épaisseur minimale d'un pixel. Le choix des pixels à supprimer lors d'une itération dans un algorithme de squelettisation est délicat. Intuitivement, il semble naturel de supprimer les points de contour qui ne changent pas la connectivité. Il nous faut donc définir ce qu'est un point de contour et comment mesurer la connectivité d'un point.

X_4	X_3	X_2
X_5	P	X_1
X_6	X_7	X_8

Figure 4.4: Pixels voisins de P.

Par convention, nous noterons le voisinage $N(P)$ d'un point P de l'image comme explicité sur la figure 4.4. Par convention, nous allons considérer que les objets d'intérêt d'une image binaire sont représentés par les pixels de valeur 1 et que, par conséquent, les pixels de valeur 0 représentent le fond ou les trous. Si S désigne l'ensemble des points objets alors \bar{S} , le complément de S , désigne l'ensemble des points de fond. Nous considérerons ensuite que P

et que son complément \bar{P} est 4-connexe (afin de noterons, enfin, $b(P)$ le nombre de pixels noirs, donc

à 1, dans $N(P)$.

Un point de contour est un point qui a au moins un pixel blanc dans son 4-voisinage. Si le pixel blanc est x_3 le point est un point de contour gauche ou ouest, s'il s'agit de x_1 le point est un point de contour droit ou est, pour x_3 c'est un point de contour haut ou nord et enfin pour x_7 il s'agit d'un point de contour bas ou sud.

Un point de fin est un point tel que $b(P) = 1$. Les algorithmes de squelettisation vont donc chercher à mettre à 0 les points de contour qui ne sont pas des points de fin.

Un sous-ensemble Q d'une image binaire I , est une 8-(4-) composante de I si et seulement si pour toute paire de points (P, M) de Q , il existe un 8-(4-) chemin allant de P à M .

La 8-(4-) connectivité d'un sous-ensemble S de l'image est le nombre de 8-(4-) composantes de son complément \bar{S} .

Un point P non-nul est dit 8-(4-) supprimable si et seulement si sa mise à valeur nulle ne change pas la 8-(4-) connectivité de l'objet d'intérêt auquel il appartient. Un point 4-supprimable est 8-supprimable mais l'inverse n'est pas vrai (Figure 4.5).

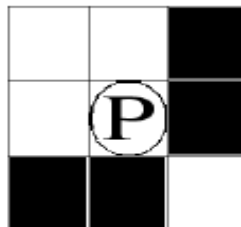


Figure 4.5: Exemple de pixel 8-supprimable et non 4-supprimable.

La squelettisation d'une forme consiste à supprimer les points de contour qui ne sont pas des points de fin et qui sont 8-(4-) supprimables (suivant le choix de la connectivité à conserver). Pour savoir si un point est 8-(4-) supprimable, il faut définir un critère calculable sur son voisinage. Ce critère est appelé, selon les auteurs :

- Nombre d'intersections,
- nombre de connectivité,
- simplicité du pixel.

C'est Rutovitz [33] qui a commencé par définir et utiliser le nombre d'intersections qui porte

son nom, qui est:
$$X_R(P) = \sum_{i=1}^8 |x_{i+1} - x_i| \quad (4.16)$$

$X_R(P)$ représente le nombre de transitions $0 \rightarrow 1$ et $1 \rightarrow 0$. Il est égal à deux fois le nombre de 4-composantes dans $N(P)$. Ainsi si $X_R(P) = 2$, il est certain que les pixels noirs de $N(P)$ sont 4-connexes (si $b(P) > 1$). Dans ce cas, et si P est un point de contour sans être un point de fin ($2 \leq b(p) \leq 6$), il est supprimable sans risque de modification pour la connectivité de la forme. En revanche, comme le critère est basé sur la 4-connectivité, les points 8-supprimables et non 4-supprimables ne seront pas supprimés. Il faut ajouter un post-traitement spécifique pour ces points, ce qui est assez lourd.

Afin de résoudre ce problème, Hilditch [51] a défini un nouveau nombre d'intersections basé, cette fois-ci, sur le nombre de 8-composantes dans $N(P)$:

$$X_H(P) = \sum_{i=1}^4 b_i \quad (4.17)$$

$$b_i = \begin{cases} 1 & \text{si } x_{2i-1} = 0 \text{ et } (x_{2i} = 1 \text{ ou } x_{2i+1} = 1) \\ 0 & \text{sin on} \end{cases} \quad (4.18)$$

$X_H(P)$ est égal au nombre de 8-composantes de $N(P)$ excepté quand tous les 4-voisins de P sont noirs mais dans ce cas P n'est pas un point de contour. Ainsi si $X_H(P) = 1$, la suppression de P ne change pas la 8-connectivité de l'objet.

Un point P est dit multiple (ou point de rupture) si $X_H(P) > 1$ et simple si $X_H(P) = 1$. Si P est un point multiple, sa suppression change la 8-connectivité de l'objet, et seuls les points simples sont supprimables.

On a appliqué quatre algorithmes ; de Rutovitz, de Zhang et Suen, de Deutch et celui de Zhang et Wang. L'utilisation de plusieurs algorithmes de squelettisation était dans le but de choisir le meilleur entre eux, c'est-à-dire celui préservant le plus la structure du mot.

4.1.2.1 Algorithme de Rutovitz

Rutovitz [33] a appliqué un balayage sur l'image suivi d'une suppression de pixels si et seulement si les quatre conditions suivantes seront satisfaites :

$$R_1 : \quad 2 \leq b(P) \leq 6 \quad (4.19)$$

$$R_2 : \quad X_R(P) = 2 \quad (4.20)$$

$$R_3 : \quad x_1 x_3 x_5 = 0 \quad \text{ou} \quad X_R(x_3) = 2 \quad (4.21)$$

$$R_4 : \quad x_1 x_3 x_7 = 0 \quad \text{ou} \quad X_R(x_1) = 2 \quad (4.22)$$

Où $b(P)$ est le nombre de pixels noirs, $X_R(P)$ est le nombre de transitions, et les x_i sont les pixels voisins du pixel du centre.

int de contour et que sa suppression ne provoquera pas
 l ne nécessite qu'une sous-itération. En revanche, il ne
 prend pas en compte la simplicité des points de voisinage pour les points supprimables, il
 entraîne donc des collisions non-gérées. Ceci implique que la stabilité du squelette n'est pas
 assurée (souvent trop érodée) et conduit à un squelette décalé par rapport à l'axe médian. Il ne
 réduit pas non plus les lignes doubles 8-connexes.

La figure suivante montre des exemples de l'application de cet algorithme sur quelques
 échantillons de la base IFN/ENIT où on remarque que cet algorithme ne préserve pas la
 structure du caractère Alif "ا" dans le cas où ce dernier est suivi par le caractère Leme "ل" où
 l'on note la disparition du Alif et du Leme pour les mots: "القلعة", "الصغرى", et "القباضة".

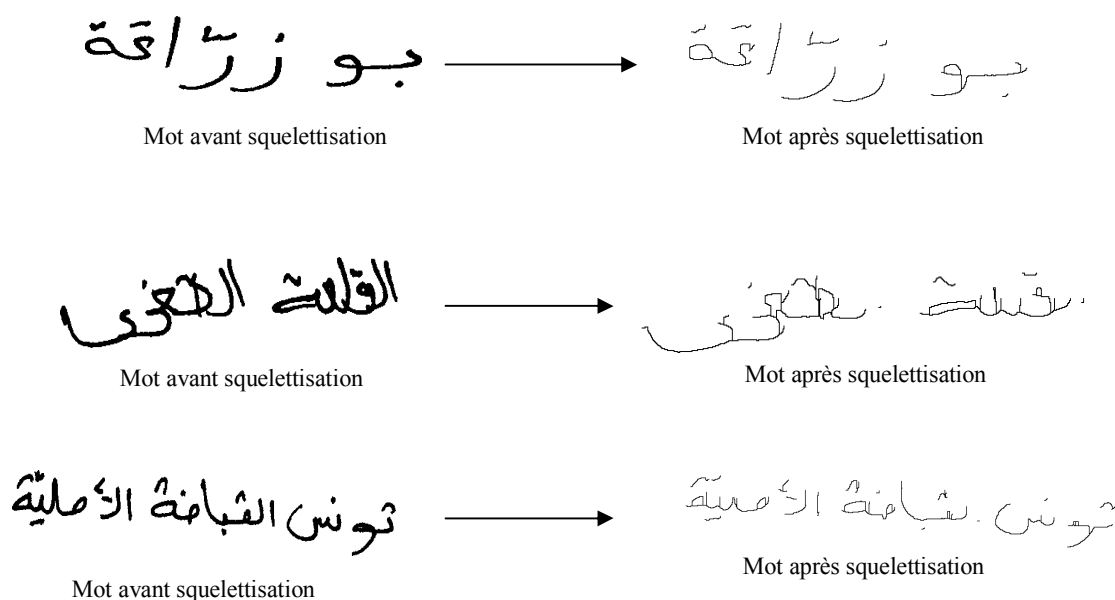


Figure 4.6: Echantillons squellettisés par l'algorithme de Rutovitz.

4.1.2.2 Algorithme de Deutsch

Deutsch [52] a ajouté une condition à celles de Rutovitz pour résoudre le problème des lignes
 doubles 8-connexes. Puis il a ajouté une sous-itération en effectuant une rotation de 180° sur
 les conditions de suppression pour tenter d'aligner le squelette avec l'axe médian. Cela donne
 les conditions suivantes :

- Première sous-itération :

$$D_1 : X_R(P) = 0,2 \quad \text{ou} \quad 4 \quad (4.23)$$

$$D_2 : b(P) \neq 1 \quad (4.24)$$

$$D_3 : x_1 x_3 x_5 = 0 \quad (4.25)$$

(4.26)

D_5 : Si $X_R(P) = 4$ alors une des conditions (a) ou (b) suivantes doit être vraie:

$$(a): \quad x_1 x_7 = 1, x_2 + x_6 \neq 0, \text{ et } x_3 + x_4 + x_8 = 0 \quad (4.27)$$

$$(b): \quad x_1 x_3 = 1, x_4 + x_8 \neq 0, \text{ et } x_2 + x_5 + x_6 + x_7 = 0 \quad (4.28)$$

- Deuxième sous-itération :

$$D_1 b: \quad X_R(P) = 0, 2 \quad \text{ou } 4 \quad (4.29)$$

$$D_2 b: \quad b(P) \neq 1 \quad (4.30)$$

$$D_3 b: \quad \text{Rotation de } 180^\circ \text{ par rapport à } D_3 \text{ c'est-à-dire } x_1 x_5 x_7 = 0 \quad (4.31)$$

$$D_4 b: \quad \text{Rotation de } 180^\circ \text{ par rapport à } D_4 \text{ c'est-à-dire } x_3 x_5 x_7 = 0 \quad (4.32)$$

$D_5 b$: Rotation de 180° par rapport à D_4 c'est-à-dire si $X_R(P) = 4$ alors une des conditions (a') ou (b') suivantes doit être vraie :

$$(a'): \quad x_3 x_5 = 1, x_2 + x_6 \neq 0, \text{ et } x_1 + x_4 + x_7 + x_8 = 0 \quad (4.33)$$

$$(b'): \quad x_5 x_7 = 1, x_4 + x_8 \neq 0, \text{ et } x_1 + x_2 + x_3 + x_6 = 0 \quad (4.34)$$

Là encore, les collisions ne sont pas gérées mais leurs conséquences sont minimisées par la double itération ; en revanche, les carrés sont souvent complètement érodés. Ce qui est montré par la figure suivante

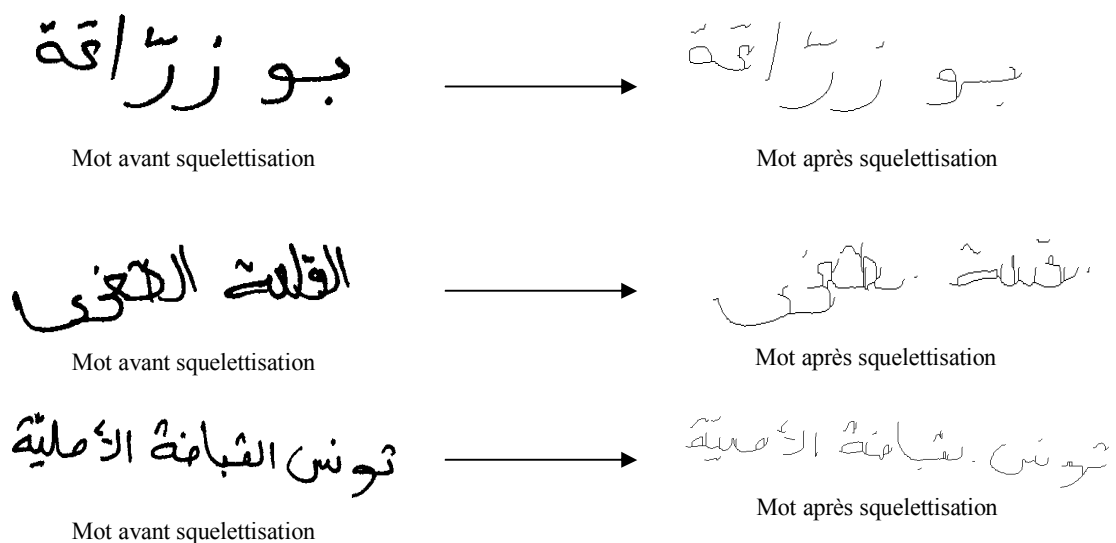


Figure 4.7: Echantillons squelettisés par l'algorithme de Deutsch.

algorithme parallèle en une seule itération qui produit des squelettes parfaitement 8-connexes et qui gère les collisions en donnant de très bons résultats et restant relativement rapide.



Figure 4.8: Exemple de collision si on supprime x_3 , P n'est plus simple.

Pour se faire, ils font l'inventaire de toutes les collisions possibles (Figure 4.8) lors de la suppression d'un point simple (caractérisé par $X_H(P) = 1$ $b(P) > 1$) suivant sa nature (Ouest, Est, Nord, Sud, Sud-ouest, Nord-ouest, Sud-est, Nord-est). Chaque point P se voit attribuer une étiquette de 0 à 8 suivant la configuration de ses voisins avec la fonction $WN : P \rightarrow \{0, \dots, 8\}$ tels que :

$WN(P) = 0$ - P est un point multiple.

$WN(P) = 1$ - P est un point simple.

$WN(P) = 2$ - P est un point simple et x_3 est un point multiple.

$WN(P) = 3$ - P est un point simple et x_5 est un point multiple.

$WN(P) = 4$ - P est un point simple et x_3 et x_5 sont des points multiples.

$WN(P) = 5$ - P est un point simple et x_7 est un point multiple.

$WN(P) = 6$ - P est un point simple et x_3 et x_7 sont des points multiples.

$WN(P) = 7$ - P est un point simple et x_3 et x_5 et x_7 sont des points multiples.

$WN(P) = 8$ - P est un point simple et x_2 est un point multiple.

La figure ci-dessous illustre quelques exemples où on remarque que cet algorithme préserve la structure du caractère Alif "ا" contrairement aux deux algorithmes cités avant. C'est pourquoi nous l'avons choisi pour être utilisé dans notre système.

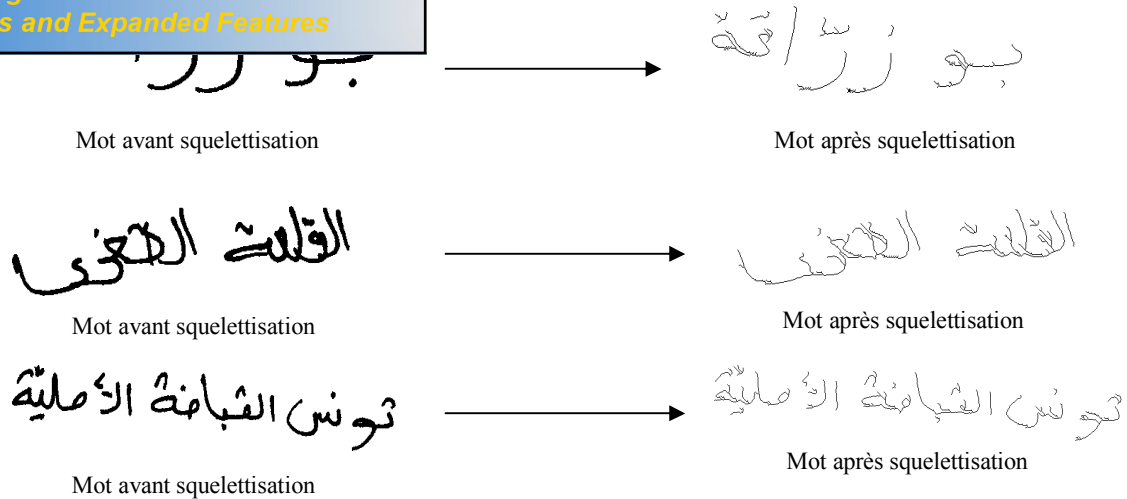


Figure 4.9: Echantillons squelettisés par l'algorithme de Zhang et Wang.

4.1.2.4 Algorithme de Zhang et Suen

Cet algorithme est le plus souvent cité et utilisé. Il consiste en deux sous-itérations pour repérer les pixels à effacer selon les critères suivants :

- Première sous-itération :

$$Z_1 : 2 \leq b(p) \leq 6 \quad (4.35)$$

$$Z_2 : X_R(P) = 2 \quad (4.36)$$

$$Z_3 : x_1 x_3 x_7 = 0 \quad (4.37)$$

$$Z_4 : x_1 x_5 x_7 = 0 \quad (4.38)$$

- Deuxième sous-itération :

$$Z_1 b : 2 \leq b(p) \leq 6 \quad (4.39)$$

$$Z_2 b : X_R(P) = 2 \quad (4.40)$$

$$Z_3 b : \text{Rotation de } 180^\circ \text{ par rapport à } Z_3 \text{ c'est-à-dire } x_3 x_5 x_7 = 0 \quad (4.41)$$

$$Z_4 b : \text{Rotation de } 180^\circ \text{ par rapport à } Z_4 \text{ c'est-à-dire } x_1 x_3 x_5 = 0 \quad (4.42)$$

Cet algorithme est relativement efficace et robuste au bruit de contour. Néanmoins certaines parties du squelette peuvent être décalées par rapport au centre.

s échantillons tirés de la base IFN/ENIT obtenus après t Suen où on remarque l'apparition du même problème rencontré avec l'algorithme de Rutovitz, à savoir la disparition des deux caractères Alif "ا" et Leme"ل".

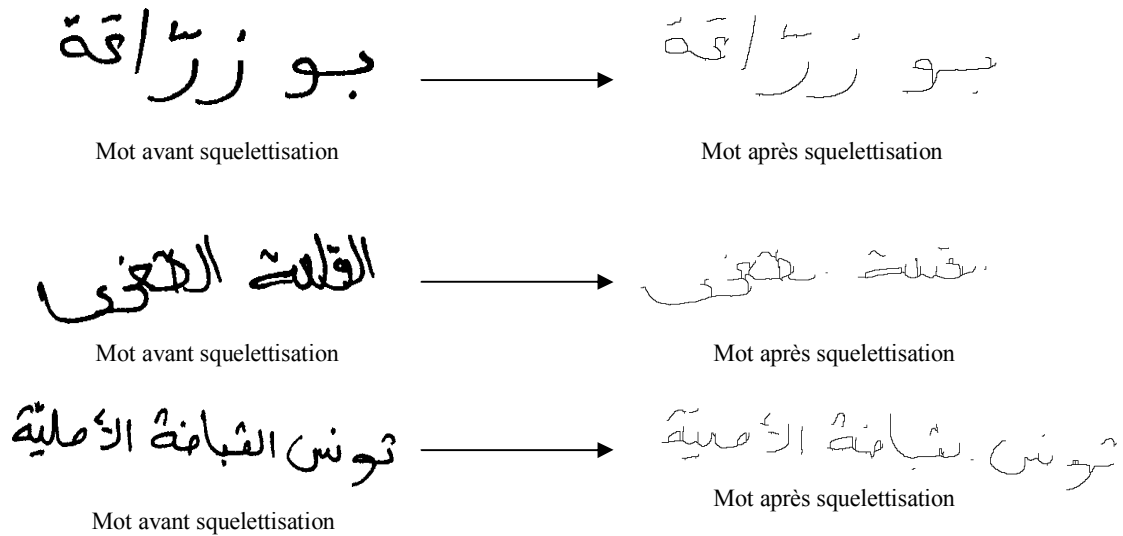


Figure 4.10: Echantillons squelettisés par l'algorithme de Zhang et Suen.

4.1.3 Alignement et Redressement

Le problème est que l'écriture cursive est généralement inclinée et les méthodes habituelles de redressement du mot entraînent souvent des distorsions causant des pertes d'information importantes. C'est d'ailleurs pour cette raison que nous proposons une technique simple de redressement en montrant que le calcul de l'entropie sur un histogramme de densité nous donne un indice supplémentaire permettant d'améliorer la détection de l'angle de rotation correcteur. Le mérite de la méthode proposée est qu'elle est plus robuste.

Dans notre méthode, nous projetons les ordonnées du contour du mot selon onze angles. Le calcul de l'entropie de l'histogramme des ordonnées pour chacun de ces angles nous permet de trouver la direction selon laquelle la projection du mot est la plus compacte. La connaissance de cette direction et l'analyse de la distribution de l'histogramme sélectionné nous permettent de choisir l'angle correcteur. L'avantage de cette technique est qu'elle ne cause aucune déformation de l'image de départ. Notre méthode est réalisée en quatre étapes:

- Extraction des points de contour du mot.
- Génération d'histogrammes des ordonnées des points suivant plusieurs inclinaisons, où chaque ordonnée des points du contour est successivement projetée dans un repère dont l'axe d'inclinaison par rapport à l'horizontale varie selon un pas de 1 degré, de -5

de 5 degré autour de l'horizontale semble raisonnable et probable qu'un mot soit écrit avec une plus grande inclinaison. Cette estimation est d'ailleurs confirmée par l'observation des exemples de notre base de données où nous avons également testé des balayages de -20 degré à +20 degré. Toutefois, si cela devait se produire, il serait possible d'augmenter le balayage. Pour chacune des 11 projections, un histogramme de densité des ordonnées est calculé. On s'attend à ce que l'inclinaison réelle du mot donne la répartition la plus compacte de l'histogramme.

- Sélection de l'histogramme le plus compact à l'aide de l'entropie. L'entropie est une mesure de l'information représentée par la formule suivante :

$$E = -\sum_i p_i \log(p_i) \quad (4.43)$$

$$\text{et } p_i = \frac{N_i}{N} \quad (4.44)$$

Où N_i est le nombre de pixels ayant l'ordonnée y_i dans le repère de projection et N est le nombre total des pixels ou de points de contour du mot. La probabilité p_i de l'histogramme désigne la fréquence d'occurrence de l'ordonnée y_i ($\sum_i p_i = 1$)

L'entropie E est maximale si toutes les probabilités p_i sont égales. Elle est minimale si toutes les probabilités p_i sauf une, sont nulles. Ainsi, plus une distribution est hétérogène, moins elle est étalée et plus petite est son entropie. Au contraire, plus une distribution est homogène, plus elle est étalée et plus grande est son entropie. Pour trouver la direction privilégiée correspondante à l'inclinaison du mot, il faudra donc minimiser l'entropie.

- Sélection de l'angle correcteur. L'angle d'inclinaison correcteur sera celui de l'histogramme le plus dense représentant la plus petite entropie.

On va montrer l'application de notre méthode sur trois exemples squelettisés de la base IFN/ENIT, la première image contient un seul mot "بوزراعة", la deuxième comprend deux mots "القلعة الصغرى" et la troisième inclut trois mots "تونس القباضة الأصلية".

Les tableaux suivants contiennent les rotations de chaque mots suivant les onze angles dans l'intervalle $[-5^\circ, +5^\circ]$.

Click Here to upgrade to Unlimited Pages and Expanded Features











 -5°	 -4°	 -3°
 -2°	 -1°	 0°
 1°	 2°	 3°
 4°	 5°	

Tableau 4.1: Les onze angles de rotation pour le mot "بوزراعة".

 -5°	 -4°	 -3°
 -2°	 -1°	 0°
 1°	 2°	 3°
 4°	 5°	

Tableau 4.2: Les onze angles de rotation pour le mot "القلعة الصغرى".

تونس القباضة الأصلية -5°	تونس القباضة الأصلية -4°	تونس القباضة الأصلية -3°
تونس القباضة الأصلية -2°	تونس القباضة الأصلية -1°	تونس القباضة الأصلية 0°
تونس القباضة الأصلية 1°	تونس القباضة الأصلية 2°	تونس القباضة الأصلية 3°
تونس القباضة الأصلية 4°	تونس القباضة الأصلية 5°	

Tableau 4.3: Les onze angles de rotation pour le mot "تونس القباضة الأصلية".

Les figures 4.11, 4.12 et 4.13 illustrent les onze histogrammes de projection horizontale associés à chacun des mots; "بوزراعة", "القلعة الصغرى" et "تونس القباضة الأصلية".

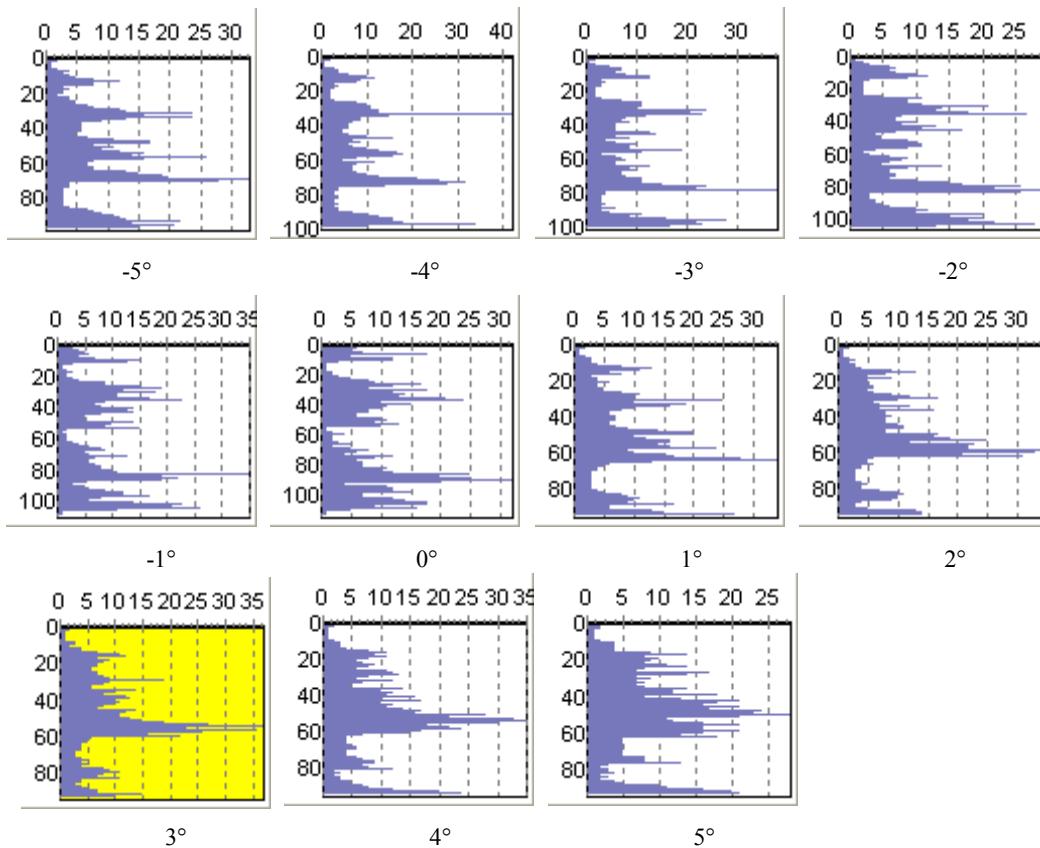


Figure 4.11: Histogrammes de rotation du mot "بوزراعة".

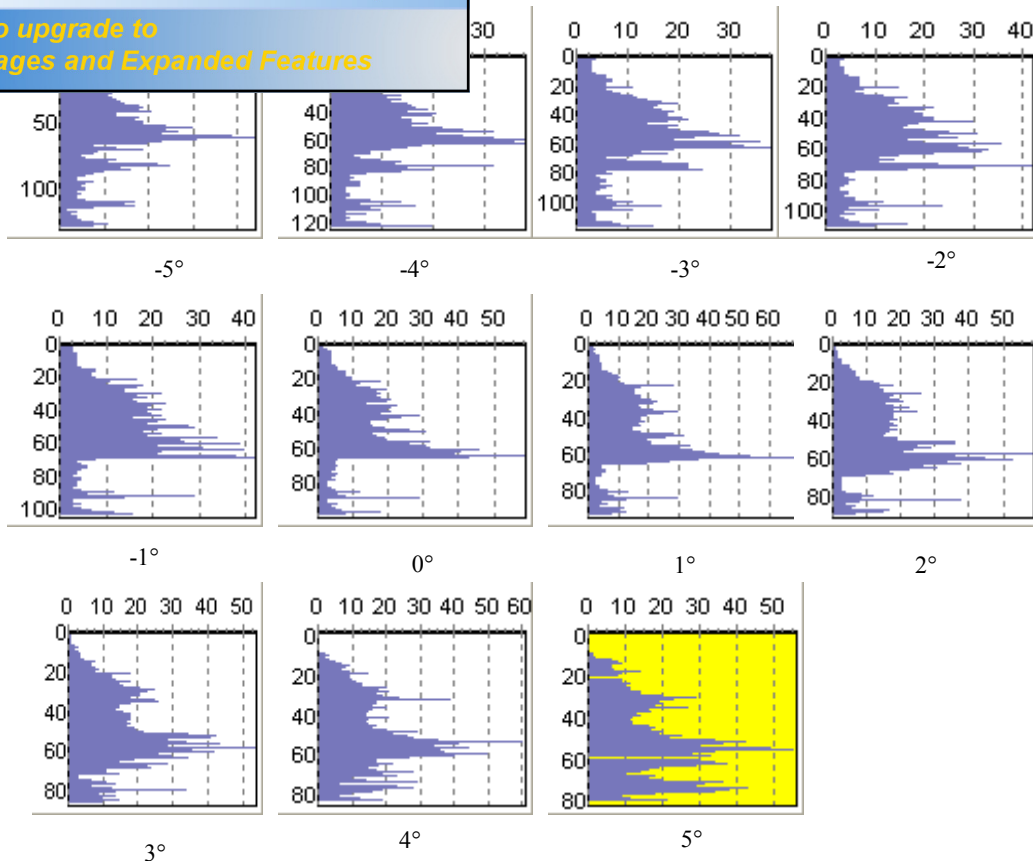


Figure 4.12: Histogrammes de rotation du mot "القلعة الصغرى".

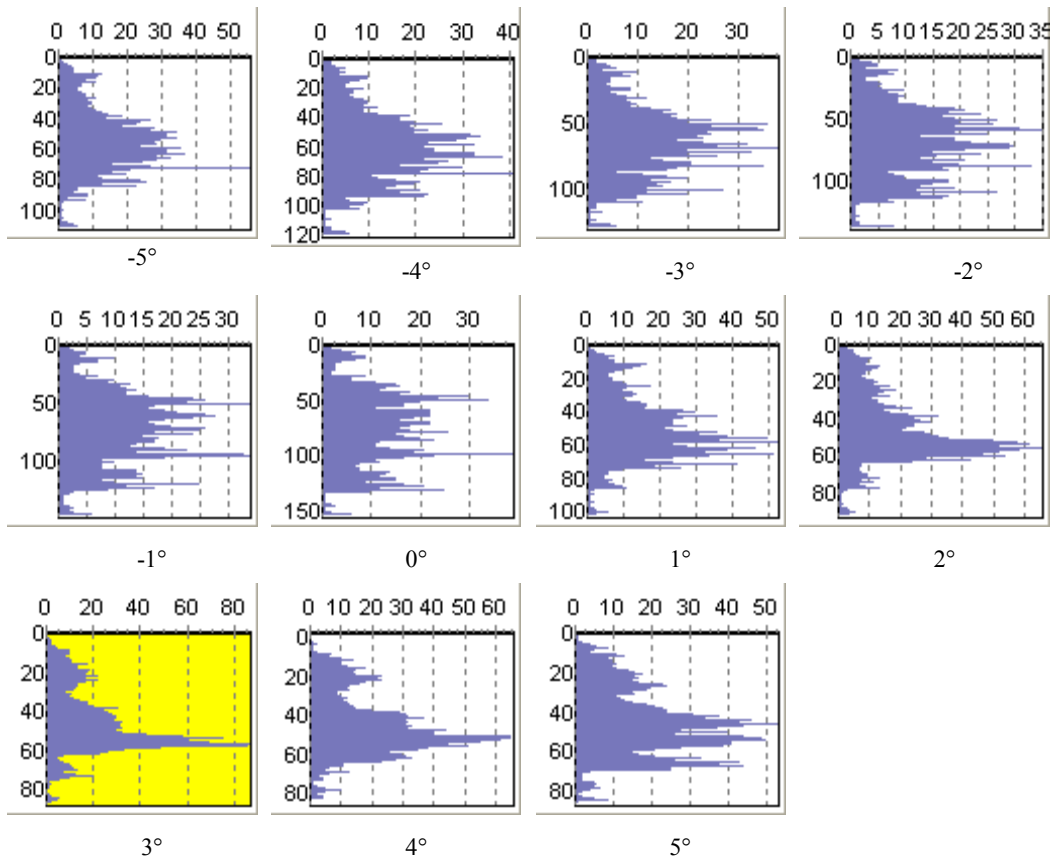


Figure 4.13: Histogrammes de rotation du mot "تونس القباضة الأصلية".

te celui le plus dense donnant ainsi l'angle correcteur.
 eur est de 3°, celui de "القلعة الصغرى" est de 5° et
 finalement le mot "تونس القباضة الأصلية" l'angle correcteur est de 3°.

La figure ci-dessous montre une correction d'inclinaison pour les mots "بوزراعة", "القلعة الصغرى" et "تونس القباضة الأصلية".

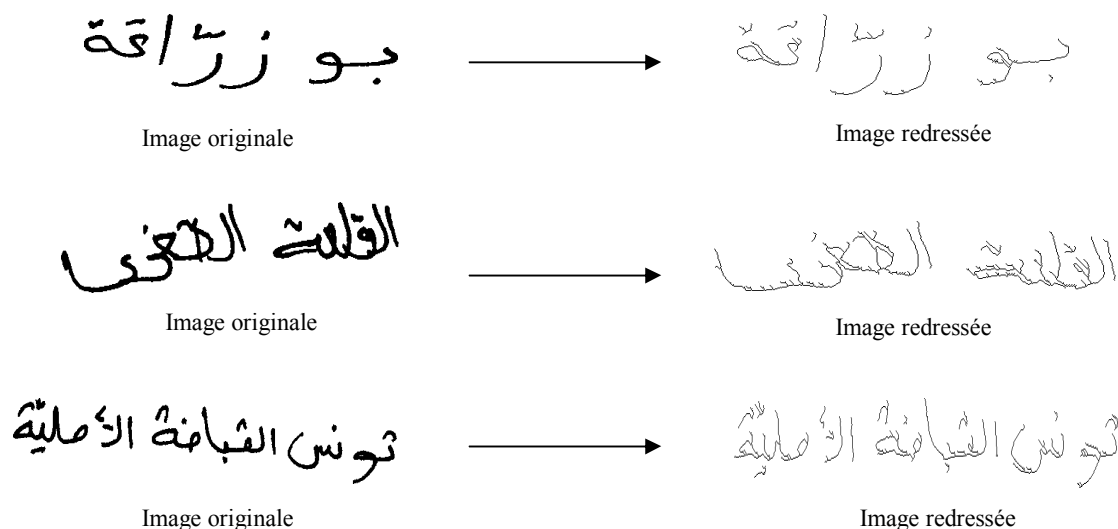


Figure 4.14: Les mots "بوزراعة", "القلعة الصغرى" et "تونس القباضة الأصلية" après redressement.

4.2 Extraction de primitives

L'étape d'extraction des primitives permet de convertir les données réelles dans un format propre à leur utilisation. Du fait de la nature numérique des systèmes actuels, cette représentation de l'information est caractérisée par une série de valeurs numériques qui traduisent, soit la présence, l'absence ou la valeur associée à la caractéristique concernée. Dans notre système nous avons utilisé les primitives suivantes :

4.2.1 Primitives statistiques

L'étude d'une image numérique débute le plus souvent par le calcul et l'analyse de son histogramme. A partir de l'histogramme on obtient les trois primitives suivantes :

La moyenne μ donnée par la relation suivante :

$$\mu = \frac{1}{256} \sum_{i=0}^{255} H(i) \quad (4.46)$$

où $H(i)$ est le nombre d'occurrence du niveau de gris i dans l'image.

$$\sigma^2 = \frac{1}{256} \sum_{i=0}^{255} H(i) \times (i - \mu)^2 \quad (4.47)$$

enfin, le nombre de sous-mots est calculé à partir de l'histogramme de projection verticale. La figure 4.15 montre que le mot "بوزراعة" contient 5 sous-mots, le mot "القلعة الصغرى" comprend également 5 sous-mots et le mot "تونس القباضة الأصلية" est composé de 8 sous-mots.

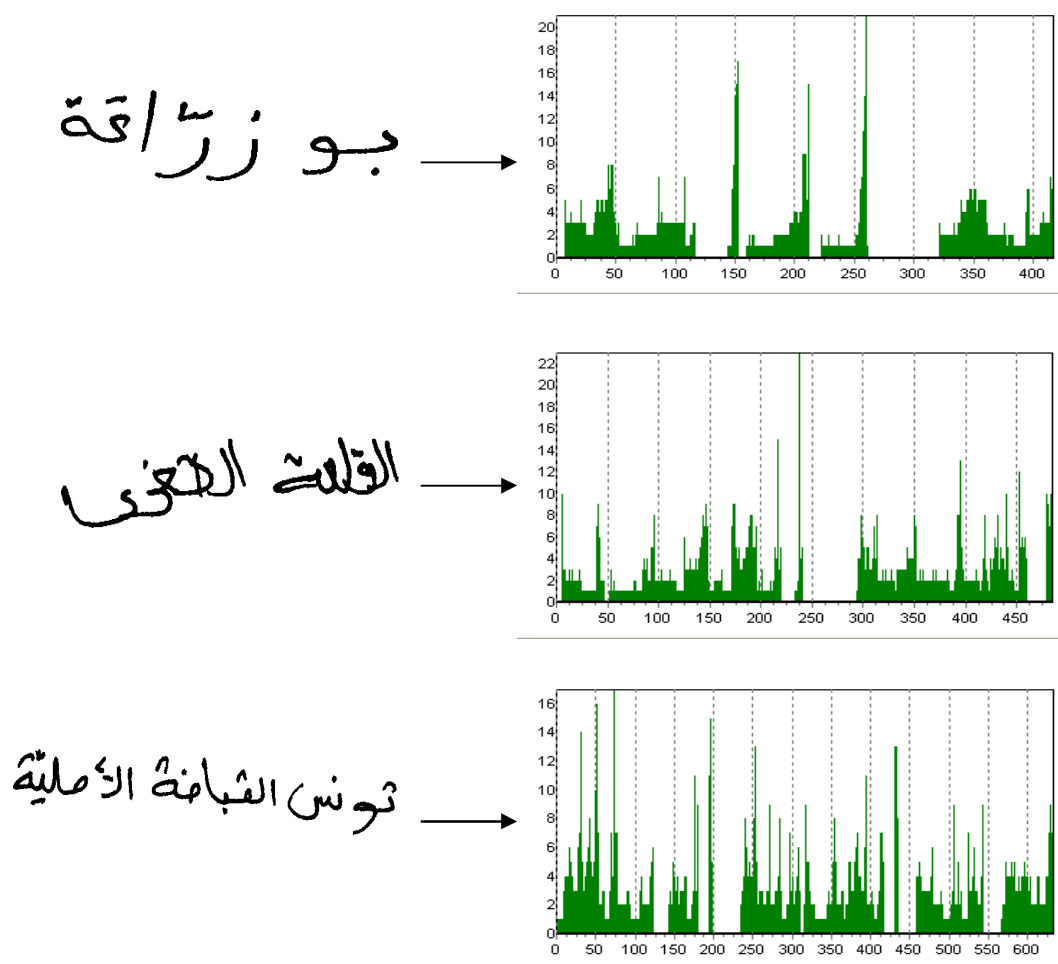
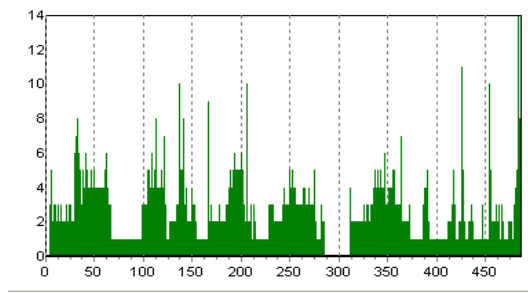


Figure 4.15: Histogrammes de projection verticale des mots "بوزراعة", "القلعة الصغرى" et "تونس القباضة الأصلية".

Les sous-mots sont séparés par des espaces dans l'histogramme, on a trouvé beaucoup de difficultés dans le calcul du nombre de sous-mots, cela est dû aux problèmes des chevauchements entre les caractères. La figure 4.16 montre deux exemples de mots contenant des chevauchements. Le premier se trouvant dans le mot "بني كلثوم" entre le caractère "و" et le caractère "م" ce qui donne un nombre de sous-mots de 2 au lieu de 3. Idem pour le mot "حفوز" le chevauchement concernant les caractères "و" et "ز" donne un seul sous-mot au lieu de 2.

بنی کلثوم →



حقوق →

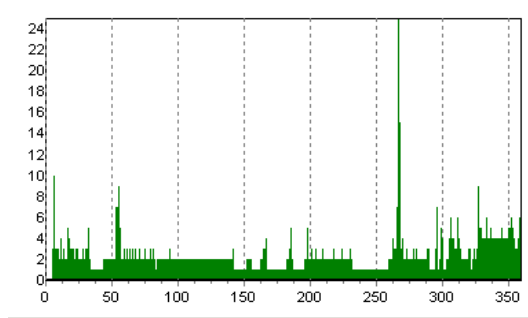


Figure 4.16: Exemple de nombre de sous-mots mal-calculé.

Pour tracer la ligne de base correspondante à une image d'un mot donné, on réutilise l'histogramme servant à la correction de l'inclinaison de l'image. Le sommet le plus haut de l'histogramme représente la ligne qui contient le plus de point objet dans l'image redressée, cette ligne représente dans le plus souvent des cas la ligne de base du mot de l'image traitée. A partir de cette ligne on utilise une méthode approximative permettant de tracer deux autres lignes, l'une au-dessus et l'autre au-dessous de la ligne de bases, sur lesquelles on se base pour déterminer le nombre d'ascendants et de descendants dans le mot sur l'image.

Pour se faire on calcule l'histogramme vertical des deux parties d'image se trouvant au-dessus et au-dessous des deux lignes tracées, à partir des quelles on calcule le nombre de sous-mots pour chacune d'eux. Le nombre de sous-mots obtenu de l'histogramme de l'image se trouvant au-dessus de la ligne haute (celle que nous avons tracé au-dessus de la ligne de base) représente le nombre d'ascendants dans le mot ; tandis que le nombre de sous-mots calculé à partir de l'histogramme de l'image se trouvant au-dessous de la ligne basse (celle que nous avons tracé au-dessous de la ligne de base) représente le nombre de descendants dans le mot de l'image.

Après plusieurs tests effectués sur différentes classes d'images, nous avons choisi de tracer nos deux lignes selon la mesure suivante : Soit T_h et T_b les tailles des deux images se trouvant respectivement au-dessus et au-dessous de la ligne de base, la ligne haute (c'est-à-

ligne de base) sera tracée à $T_h/3$, tandis que la ligne (en dessous de la ligne de base) sera tracée à $T_b/2$.

La figure 4.17 illustre la détection de la ligne de base pour les mots : "بوزراعة", "القاعة الصغرى" et "تونس القباضة الأصلية".

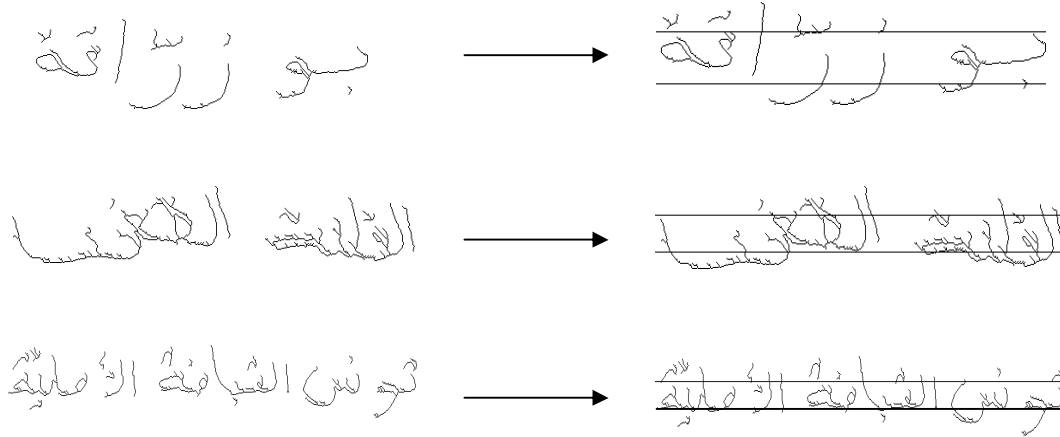


Figure 4.17: Ligne de base des mots "بوزراعة", "القاعة الصغرى" et "تونس القباضة الأصلية".

4.2.2 Moments de Tchebichef

Les moments de Tchebichef sont inclus dans la famille des moments orthogonaux discrets, on va commencer par une description concernant les moments orthogonaux ensuite on va voir la signification du terme discret. Finalement on présentera le plan théorique permettant le calcul de ces moments.

4.2.2.1 Moments orthogonaux discrets

Les moments orthogonaux ont été proposés afin d'éliminer certains problèmes causés par les moments invariants tels que; la redondance d'informations dans les vecteurs de primitives, la sensibilité aux bruits et la grande variation dans l'intervalle dynamique des valeurs menant à une instabilité numérique quand la taille de l'image est grande. Ces moments possèdent un domaine complètement différent de l'espace de coordonnées de l'image, par exemple les polynômes de Legendre et de Tchebichef sont valides seulement dans l'intervalle $[-1,+1]$ tandis que les polynômes radiaux de Zernike sont définis à l'intérieur d'un cercle.

Les moments à base orthogonale discrets qui sont proposés récemment sont faciles à implémenter et présentent une grande tolérance aux bruits et une grande précision concernant la reconstruction de l'image, leur inconvénient majeur est l'absence de l'invariance aux transformations, aux rotations et aux changements d'échelle où une étape de normalisation doit être faite avant [22]. Le problème des moments discrets est que les coefficients des polynômes sont généralement calculés en utilisant des équations récurrentes, ce qui conduit

on des erreurs en augmentant l'ordre des moments et la

Performances Types de moments		Sensibilité aux bruits	Complexité	Exactitude de reconstruction d'image	Invariance aux changements d'échelle et à la rotation
Moments géométriques		Elevé	Faible	Extrêmement difficile	Non
Moments complexes		Elevé	Faible	Extrêmement difficile	Oui
Moments à base orthogonale continus	Moments de Legendre	Faible	Elevé	Moyen	Oui
	Moments de Zernike	Faible	Très élevé	Elevé	Oui
Moments à base orthogonale discrets	Moments de Tchebichef	Faible	Moyen	Très élevé	Non
	Moments de Hahn	Faible	Moyen	Très élevé	Non

Tableau 4.4: Comparaison entre les performances des différents moments.

Le tableau 4.4 détaille les différents types de moments appliqués à l'image en faisant une comparaison sur leurs performances. Les moments géométriques et les moments complexes ont un coût de calcul moindre, mais sont très sensibles aux bruits et ont une capacité de reconstruction d'image extrêmement difficile. Les moments à base orthogonale quand à eux sont insensibles aux bruits et préservent l'avantage d'avoir une grande précision de reconstruction d'image, par contre leurs coûts de calcul reste élevé.

Les moments discrets de Tchebichef sont proposés par Mukundan [71] afin d'éliminer les problèmes des moments orthogonaux continus tels que les moments de Zernik et de Legendre liés à la reconstruction d'image. Ce sont des moments basés sur les polynômes orthogonaux

d'avantages par rapport aux autres types de moments.

- Un domaine discret de définition qui est lié à l'espace des coordonnées de l'image.
- Une absence des conditions d'approximation numérique, permettant une représentation plus précise des caractéristiques d'images en donnant comme résultat une image exactement reconstruite à partir d'un ensemble complet des moments discrets en consommant un temps de calcul moyen.

4.2.2.2 Plan théorique

Les polynômes de Tchebichef d'ordre n définis sur l'espace discret des coordonnées $x \in [0, 1, 2, \dots, N-1]$ comme suit :

$$t_n(x) = (1-N) {}_3F_2(-n, -x, 1+n; 1, 1-N, 1) \quad (4.48)$$

Où $n, x = 0, 1, 2, \dots, N-1$.

La fonction ${}_3F_2$ est une fonction hypergéométrique définie par :

$${}_3F_2(a_1, a_2, a_3; b_1, b_2; z) = \sum_{k=0}^{\infty} \frac{(a_1)_k (a_2)_k (a_3)_k z^k}{(b_1)_k (b_2)_k k!} \quad (4.49)$$

Où $(a_i)_k$ et $(b_i)_k$ sont des symboles de Pochhammer donnés par :

$$(a)_k = a(a+1)(a+2)\dots(a+k-1) \quad (4.50)$$

$$(b)_k = b(b+1)(b+2)\dots(b+k-1) \quad (4.51)$$

Les moments de Tchebichef satisfont la relation d'orthogonalité suivante :

$$\sum_{i=0}^{N-1} t_m(i) t_n(i) = \rho(n, N) \delta_{mn} \quad (4.52)$$

Pour $0 \leq m, n \leq N-1$, la fonction $\rho(n, N)$ est la norme carrée donnée par :

$$\rho(n, N) = (2n)! \binom{N+n}{2n+1} \quad (4.53)$$

La valeur du polynôme $t_n(x)$ peut être facilement calculée en utilisant la relation récurrente suivante :

$$(n+1)t_{n+1}(x) - (2n+1)(2x-N+1)t_n(x) + n(N^2-n^2)t_{n-1}(x) = 0 \quad (4.54)$$

Avec comme conditions initiales :

$$\begin{aligned} t_0(x) &= 1 \\ t_1(x) &= 2x - N + 1 \end{aligned} \quad (4.55)$$

nsion sont donnés par la formule :

$$I_p = \frac{1}{\tilde{\rho}(r, N)} \sum_{x=0}^r t_p(x) J(x) \quad (4.56)$$

Où

$$\tilde{t}_n(x) = \frac{t_n(x)}{\beta(n, N)} = \frac{1}{\beta(n, N)} n! \sum_{k=0}^n (-1)^{n-k} \binom{N-1-k}{n-k} \binom{n+k}{n} \binom{x}{k} \quad (4.57)$$

Les moments de Tchebichef à deux dimensions peuvent être exprimés par la formule :

$$T_{p,q} = \frac{1}{\tilde{\rho}(p, N)\tilde{\rho}(q, N)} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} \tilde{t}_p(x) \tilde{t}_q(y) f(x, y) \quad (4.58)$$

Les trois tableaux suivants donnent les 49 premiers moments pour trois images de la classe "القلعة الصغرى" où on remarque l'apparition des mêmes valeurs dans chaque tableau:

القلعة الصغرى

0,0308	0,0167	0,0076	0,0207	0,0052	0,9704	0,0032	0,0253	0,0140	0,0300
0,0007	0,9712	0,9884	0,0094	0,9779	0,9769	0,0173	0,9706	0,9491	0,0334
0,9785	0,9798	0,9675	0,9885	0,0087	0,9909	0,0172	0,9477	0,9939	0,9768
0,9964	0,0225	0,0223	0,0192	0,9718	0,0043	0,9806	0,9944	0,0043	0,0302
0,0187	0,9919	0,0126	0,0043	0,9952	0,9835	0,0042	0,0153	0,9976	

Tableau 4.5: Les 49 premiers moments de Tchebichef du premier échantillon du mot "القلعة الصغرى".

القلعة الصغرى

0,0329	0,0152	0,0086	0,0278	0,0066	0,9462	0,0042	0,0396	0,0120	0,0520
0,0009	0,9869	0,9769	0,0081	0,9456	0,9354	0,0143	0,9425	0,9607	0,0654
0,9621	0,9825	0,9756	0,9952	0,0063	0,9741	0,0212	0,9756	0,9705	0,9598
0,9897	0,0298	0,0245	0,0154	0,9851	0,0027	0,8756	0,9502	0,0038	0,0485
0,0201	0,9992	0,0158	0,0063	0,9259	0,9497	0,0079	0,0108	0,9698	

Tableau 4.6: Les 49 premiers moments de Tchebichef du deuxième échantillon du mot "القلعة الصغرى".

القلعة الصغرى

0,0311	0,0162	0,0068	0,0247	0,0077	0,9774	0,0022	0,0213	0,0112	0,0376
0,0009	0,9726	0,9856	0,0095	0,9833	0,9587	0,0195	0,9973	0,9416	0,0334
0,9735	0,9783	0,9689	0,9880	0,0088	0,9902	0,0254	0,9990	0,9930	0,9981
0,9979	0,0162	0,0316	0,0288	0,9605	0,0094	0,9896	0,9982	0,0408	0,0310
0,0153	0,9912	0,0171	0,0048	0,9946	0,9827	0,0000	0,0334	0,9930	

Tableau 4.7: Les 49 premiers moments de Tchebichef du troisième échantillon du mot "القلعة الصغرى".

4.3 Reconnaissance

La reconnaissance regroupe deux tâches importantes : l'apprentissage et la classification. Le résultat de l'apprentissage est, soit la réorganisation ou le renforcement des classes existantes en tenant compte de l'apport de la nouvelle classe ; soit la création d'une nouvelle classe représentant la forme entrée. Le résultat de la classification est un avis sur l'appartenance ou non de l'entrée aux classes d'apprentissage. On va commencer par décrire la base utilisée ensuite on détaillera le principe de ces deux phases.

4.3.1 Description de la base

Nous avons utilisé la base de données Tuniso-Allemande IFN/ENIT qui est une base de mots arabes manuscrits [79] dont le développement rentre dans le cadre d'un projet de coopération entre le laboratoire de Systèmes et de Traitement de Signal de l'ENIT (Ecole Nationale d'Ingénieurs de Tunis) et le laboratoire de recherche sur la reconnaissance de l'écriture arabe de l'IFN (Institute Fuer Nachrichtentechnik). Cette base a été réalisée pour avancer la recherche et le développement des systèmes de reconnaissance de mots arabes manuscrits ; plus de 30 groupes de recherche l'ont déjà utilisé, elle est disponible et gratuite sur le site : www.ifnenit.com pour les recherches non-commerciales. La base dans sa version 2.0 de niveau 1 contient 26459 images de noms composés de 1 à 4 mots, avec un total de 115585 sous-mots et 212211 caractères rédigés par 411 scripteurs représentant les 937 noms de villages tunisiens.

Un type de formulaire est illustré dans la figure 4.18 où le scripteur a écrit une liste de mots représentant les noms et codes de villes/villages tunisiens.

PLACE ↓

9046	فندان السوق	9046 فندان السوق
3024	شغال	3024 شغال
9112	الفايض	9112 الفايض
3263	تطاوين 7 نوفمبر	تطاوين 7 نوفمبر 3263
6016	عرام	6016 عرام
7141	تل الغزلان	تل الغزلان 7141
8189	جندوبة الشمالية	جندوبة الشمالية 8189
4174	حاسي الجربي	حاسي الجربي 4174
3067	مركز الدجسي	مركز الدجسي 3067
2133	قفصة حي الشباب	قفصة حي الشباب 2133
3024	شغال	3024 شغال
6020	الحامة	6020 الحامة

Age: < 20 <input type="checkbox"/>	Profession : Étudiant/élève <input checked="" type="checkbox"/>	Nom: Khalifa Nawrès
21 - 30 <input checked="" type="checkbox"/>	Enseignant <input type="checkbox"/>	Ville: Gafsa
31 - 40 <input type="checkbox"/>	Administratif <input type="checkbox"/>	
> 40 <input type="checkbox"/>	Autre <input type="checkbox"/>	

Responsable: Samia Smevli Medtoui	Numéro : A1.
--	---------------------

Figure 4.18: Exemple de formulaire rempli pour la construction de la base IFN/ENIT.

Le tableau 4.8 montre des statistiques concernant le nombre de mots, de sous-mots et de caractères composant les noms de villes/villages tunisiens tirés de la base IFN/ENIT.

composant le nom	ville	ms de mots	Nombre de sous-mots	Nombre de caractères
1	12992		40555	76827
2	10826		54722	98828
3	2599		20120	36004
4	42		188	552
Total	26459		115585	212211

Tableau 4.8: Nombre de mots, de sous-mots et de caractères composant IFN/ENIT.

Le tableau 4.9 montre des statistiques concernant l'age et la profession des scripteurs qui ont contribué dans la construction de la base.

Age	Etudiant	Professeur	Technicien	autres	Total
≤20	29%	0%	0%	0%	29%
21-30	35,6%	4,2%	3,9%	3,9%	47,6%
31-40	0,2%	3,4%	4,9%	2,0%	10,5%
>40	0%	4,1%	5,4%	3,4%	12,9%
Total	64,8%	11,7%	14,2%	9,3%	100%

Tableau 4.9: Age des scripteurs participant à la construction de la base IFN/ENIT.

La base contient dans sa version finale cinq ensembles: a, b, c, d et e, en général les quatre premiers ensembles sont utilisés pour l'apprentissage tandis que l'ensemble e est laissé pour effectuer les tests. Le tableau 4.10 donne le nombre de mots et le nombre de scripteurs participant à la création d'un ensemble donné de la base.

Ensemble	Nombre de mots	Nombre de scripteurs
A	6537	102
B	6710	102
C	6477	103
D	6735	104
E	6033	87

Tableau 4.10: Nombre de mots et de scripteurs pour chaque ensemble de l'IFN/ENIT.

chantillons d'images de mots extraites de la base

[Click Here to upgrade to
Unlimited Pages and Expanded Features](#)

لَبَّ لَهْدَان	بِنِي كَلْتُوم	الْقَالَةَ الصُّورَى	أَكُو دَه
شَم مَرِيم	شَمَحَال	بُو زْرَاعَة	بُو عَثْمَان
الْفَكَّة	الْمَا رَزَة	الشَّرَايِج	الدَّخَانِيَة
السَّعِيدَة	الرِّضَاع	الْخَلِيج	سَمَخ
كَسْرِمَا	حَمُور	هَذِيل	حَلَا الْجَمَل
مَارْت	مَثَلِين	حَيَا الصَّلَاح	كَمْبُوط
سَمْبَاط	أَوْتِيك الْجَدِيدَة	رَقَادَة	نَفْثَة
سَيِّدِي إِبْرَاهِيم الزَّمَار	سَبْعَة آبَار	تَلَّ الْغَزْلَان	تُونُو التَّبَاضَة الْأَصْلِيَّة

Tableau 4.11: Echantillons de mots de la base IFN/ENIT.

alisé sur un sous-ensemble de la base IFN/ENIT.
 L'algorithme d'apprentissage utilisé est celui du réseau Fuzzy ART qui est décrit par l'organigramme ci-dessous.

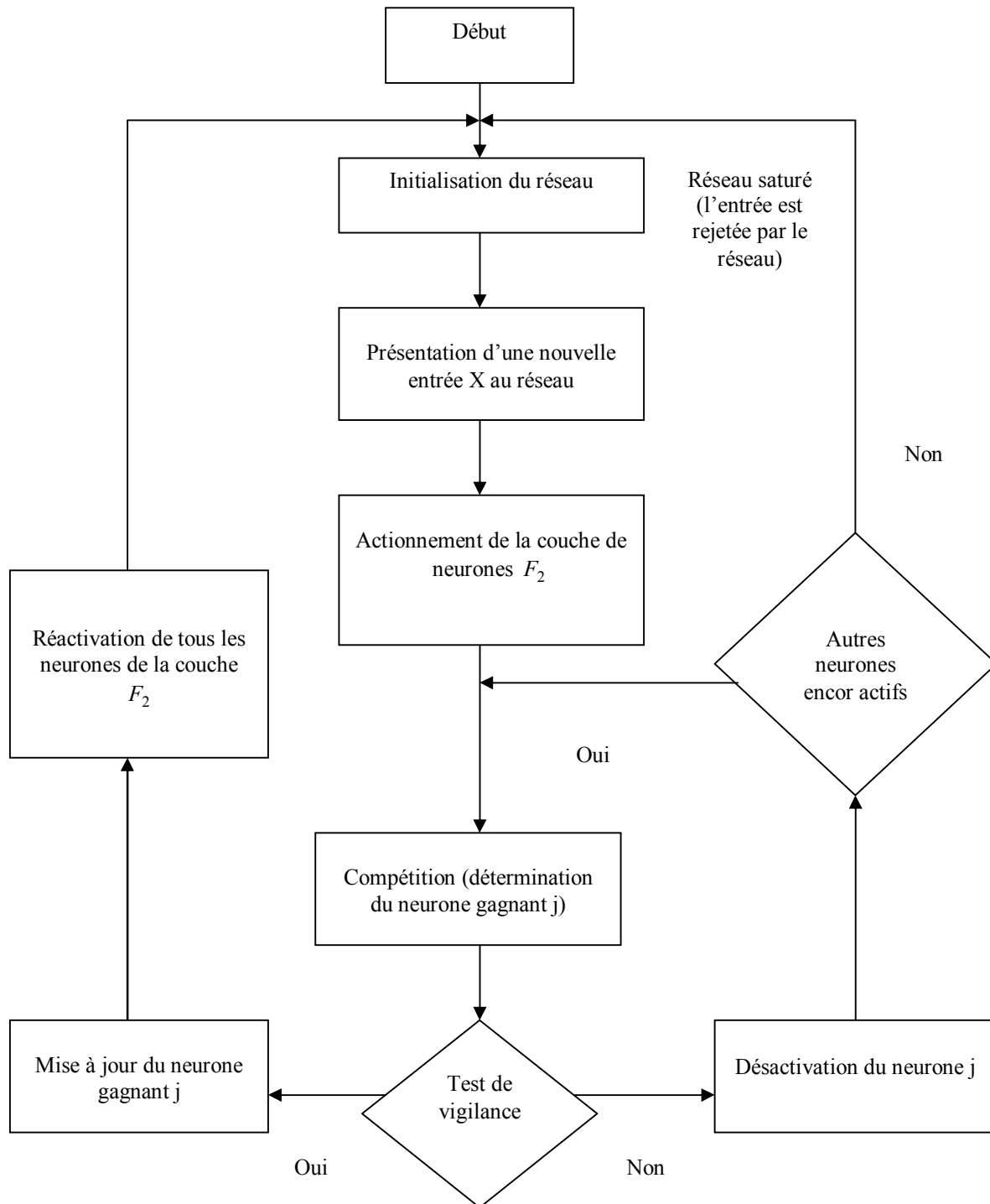


Figure 4.19: Fonctionnement de l'algorithme d'apprentissage de Fuzzy ART.

aux étapes suivantes :

..... w_s]^T avec des 1.

2. Fixer le paramètre de sélection α , le taux d'apprentissage η , et le taux de vigilance ρ .
3. $t=1$ (4.59)
4. Répéter tant que $t \leq t_{\max}$:

- (a) Choisir aléatoirement un ensemble flou $p(t)$ parmi la base d'apprentissage.
- (b) Effectuer l'encodage complémentaire

$$E^c : a^1(t) = \begin{bmatrix} p(t) \\ p^c(t) \end{bmatrix} \quad (4.60)$$

$$\text{Où } p^c = [(1 - p_1)(1 - p_2) \dots (1 - p_R)]^T \quad (4.61)$$

- (c) Calculer les niveaux d'activation n_i^1 des neurones de la première couche :

$$n_i^1(t) = \frac{|a^1(t) \cap w_i(t)|}{\alpha + |w_i(t)|} \quad (4.62)$$

$i=1, \dots, s$, où α est le taux de sélection, $|\cdot|$ désigne la norme l_1 du vecteur :

$$|x| = |[x_1 x_2 \dots x_n]^T| = \sum_{i=1}^n x_i \quad (4.63)$$

et $x \cap y$ représente l'intersection floue entre x et y :

$$\begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \end{bmatrix} \cap \begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ y_n \end{bmatrix} = \begin{bmatrix} \min(x_1, y_1) \\ \min(x_2, y_2) \\ \cdot \\ \cdot \\ \min(x_n, y_n) \end{bmatrix} \quad (4.64)$$

- (d) Calculer les sorties de la couche compétitive :

$$a_i^2(t) = \begin{cases} 1 & \text{si } n_i^1(t) = \max_{j=1, \dots, S} n_j^1(t) \\ 0 & \text{autrement} \end{cases}, i = 1, \dots, S \quad (4.65)$$

$$\text{soit } g = \arg \max_j n_j^1(t) \quad (4.66)$$

l'indice du neurone gagnant.

- (e) Si le degré de résonance du neurone g avec les sorties de la couche d'entrée est

$$\text{inférieur au seuil } \rho : \frac{|a^1(t) \cap w_g(t)|}{|a^1(t)|} < \rho \quad (4.67)$$

ement le neurone g pour qu'il ne puisse plus gagner et pour choisir le prochain gagnant.

- (f) Libérer tous les neurones bloqués.
- (g) Mettre à jour le prototype du neurone gagnant :

$$\Delta w_g(t) = \eta [a^1(t) \cap w_g(t) - w_g(t)] \quad (4.68)$$

où η est le taux d'apprentissage.

- (h) $t=t+1$ (4.69)

Fin

La première chose que l'on peut remarquer à l'équation 4.60 est que la norme des vecteurs encodés demeure toujours constante :

$$|a^1(t)| = \sum_{i=1}^R p_i + \sum_{i=1}^R (1 - p_i) = \sum_{i=1}^R p_i + R - \sum_{i=1}^R p_i = R \quad (4.70)$$

Le processus d'encodage permet donc de normaliser en amplitude les entrées. Le niveau d'activation η_i^1 du neurone i de la couche compétitive, calculé à l'équation 4.62, permet de mesurer le degré de ressemblance entre l'entrée a^1 , exprimée dans son encodage complémentaire, et le vecteur de poids w_i . En posant $w_i = [xy^c]^T$ (4.71)

Où x et y sont deux vecteurs quelconques dans l'espace des entrées, on obtient l'interprétation géométrique de la figure 4.20 pour le cas particulier d'un espace à deux dimensions. La région définie par le rectangle en trait plein correspond à la zone d'activité du neurone i . Pour tous les entrées p situées à l'intérieur de cette zone, le numérateur de l'équation 4.62 sera égal à :

$$a^1 \cap w_i = \begin{bmatrix} p \\ p^c \end{bmatrix} \cap \begin{bmatrix} x \\ y^c \end{bmatrix} = \begin{bmatrix} p \cap x \\ (p \cup y)^c \end{bmatrix} = \begin{bmatrix} x \\ y^c \end{bmatrix} = w_i \quad (4.72)$$

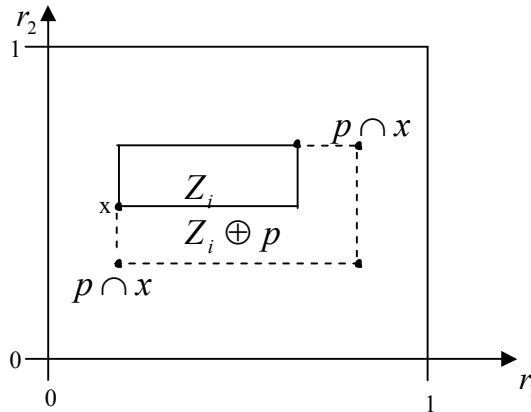


Figure 4.20: Régions associées à $w_i = [xy^c]^T$ (en trait plein) et à $a^1 \cap w_i$ (en trait pointillé).

$$a^1 \cap w_i = |w_i| \quad (4.73)$$

$$\text{Et } n_i^1 = \frac{|w_i|}{\alpha + |w_i|} \quad (4.74)$$

C'est-à-dire une valeur inférieure à 1. Pour un w_i fixé, plus α sera grand, plus le niveau d'activation sera petit et, inversement, plus w_i sera grand, plus le niveau d'activation tendra vers 1 pour un α donné.

Pour les entrées p situés à l'extérieur de la zone, comme à la figure 4.20, l'opération $a^1 \cap w_i$ engendre la nouvelle région indiquée en trait pointillé. Pour cette région, nous aurons $|a^1 \cap w_i| < |w_i|$ et le niveau d'activation calculé sera nettement inférieur à 1. Ainsi, les neurones dont la région associée englobe p gagneront la compétition avant les autres, et s'ils sont plusieurs à l'englober, c'est celui dont la région est la plus petite qui sera déclaré gagnant.

Soit Z_i la région associée à w_i . Dans le cas général d'un espace des entrées à R dimensions, il s'agit d'un hyper-rectangle défini respectivement par les coins x et y de ses composantes minimum et maximum. La dimension de cette région peut se calculer de la façon suivante :

$$|Z_i| \equiv |x - y| \quad (4.75)$$

Pour qu'un vecteur w_i puisse changer de valeur (équation 4.68), il faut que le neurone correspondant gagne la compétition de l'équation 4.65 et que celui-ci n'échoue pas le critère de résonance de l'équation 4.67. Dans ce cas seulement, la région Z_i associée à ce neurone pourra grandir et, éventuellement, englober l'entrée. Si l'on pose un taux d'apprentissage $\eta=1$ à l'équation 4.68, on parle dans ce cas d'apprentissage instantané, alors la région Z_i sera grandie juste assez pour englober l'entrée p . Dans le cas particulier d'un neurone inactif (un neurone qui n'a jamais gagné), c'est-à-dire un neurone dont le w_i ne contient que des uns (étape 1 de l'algorithme), on obtiendra $w_i(t+1) = a^1(t)$ et la région Z_i sera limitée au point $p(t)$. Par la suite, l'apprentissage ne pourra qu'augmenter la taille de la région Z_i et, par conséquent, la norme de w_i ne pourra que diminuer. En fait, toujours sous l'hypothèse de $\eta=1$, on peut montrer que la région Z_i correspondra toujours au plus petit rectangle (ou hyper-rectangle) englobant l'ensemble des entrées qui vont faire gagner et résonner le neurone i au cours de l'apprentissage.

4.67 impose une contrainte sur la taille maximum que les neurones du Fuzzy ART. En effet, le dénominateur

$|a^1(t)| = R$ est une constante et il faut que :

$$|a^1(t) \cap w_g(t)| \geq R\rho \quad (4.76)$$

Pour que le neurone gagnant puisse subir une modification de son vecteur de poids. Or :

$$\begin{aligned} |a^1(t) \cap w_g(t)| &= \left| \begin{bmatrix} p \\ p^c \end{bmatrix} \cap \begin{bmatrix} x \\ y^c \end{bmatrix} \right| \\ &= \begin{bmatrix} p \cap x \\ p^c \cap y^c \end{bmatrix} \\ &= \begin{bmatrix} p \cap x \\ (p \cup y)^c \end{bmatrix} \quad (\text{de Morgan}) \\ &= |p \cap x| + R - |p \cup y| \\ &= R - (|p \cup y| - |p \cap x|) \\ &= R - |Z_g \oplus p| \end{aligned} \quad (4.77)$$

Où $Z_g \oplus p$ désigne la plus petite région (hyper-rectangle) qui englobe à la fois Z_g et p . Par conséquent, on obtient l'expression suivante : $|Z_g \oplus p| \leq R(1 - \rho)$ (4.78)

Qui spécifie la taille maximum qu'une région peut atteindre en fonction du taux de vigilance ρ . Plus il est grand, plus les régions sont contraintes à de petites tailles. A la limite, lorsque $\rho \rightarrow 1$, le Fuzzy ART apprendra les entrées par cœur. Plus il est petit, plus les régions pourront croître. A la limite, lorsque $\rho \rightarrow 0$, un seul neurone pourrait couvrir tout l'espace d'entrée, et le réseau ne produirait plus qu'une seule catégorie. Lorsque le neurone gagnant ne respecte pas le critère de résonance, alors il est temporairement tiré de la compétition et le neurone suivant ayant le plus grand niveau d'activation est sélectionné à son tour. Ce processus est répété jusqu'à ce qu'on trouve un neurone qui résonne. Ceci sera toujours possible dans la mesure où le réseau comporte encore des neurones inactifs.

Généralement, dans le cas d'une entrée qui ne ressemble à aucune des catégories actives, c'est-à-dire lorsque le neurone gagnant était précédemment inactif, on fixera automatiquement le taux d'apprentissage $\eta=1$ pour passer en mode d'apprentissage instantané. Sinon, si l'entrée engendre la résonance d'une catégorie existante, alors on fixera un taux $0 < \eta \leq 1$. Finalement, mentionnons que le paramètre de sélection α est habituellement fixé à une petite valeur telle que, par exemple, 0.01. Ainsi, le neurone gagnant sera celui dont la région

andissement relatif pour pouvoir englober l'entrée.
ntage d'agrandissement qui compte, mais seulement la
dimension finale.

4.3.3 Classification

Elle est effectuée par le classifieur Fuzzy ART sur l'ensemble e de la base IFN/ENIT en se basant sur une méthode de classification globale, où on a obtenu les résultats suivants :

4.3.3.1 Résultats

Le taux de reconnaissance obtenu avoisine les 82%, tandis que le taux de rejet est de 10% ; les 8% restants représentent le taux de mots mal-classés. Comme on a indiqué dans le chapitre précédant ces trois taux sont calculés par les équations suivantes :

$$\text{Taux de reconnaissance} = \frac{\text{Nombre d'entités bien reconnues}}{\text{Nombre total d'entités présentées au classifieur}} \quad (4.79)$$

$$\text{Taux d'erreur} = \frac{\text{Nombre d'entités mal-reconnues}}{\text{Nombre total d'entités présentées au classifieur}} \quad (4.80)$$

$$\text{Taux de rejet} = \frac{\text{Nombre d'entités rejetées}}{\text{Nombre total d'entités présentées au classifieur}} \quad (4.81)$$

Le tableau suivant montre les résultats obtenus en variant le nombre de classes à reconnaître ainsi que le nombre d'itérations effectuées lors de la phase d'apprentissage.

Classes	Top1	Top2	Top3	Top4	Top5	Top8	Top9	Top10
150	78,66%	79,15%	80,88%	81,87%	82,52%	84,45%	86,95%	88,48%
300	74,45%	75,25%	76,89%	78,15%	80,04%	82,44%	84,58%	86,97%
450	71,52%	73,38%	75,14%	77,25%	79,79%	81,45%	83,01%	84,86%
600	69,72%	71,75%	74,63%	76,19%	78,89%	80,40%	82,05%	83,10%
750	68,15%	69,54%	72,16%	74,25%	76,75%	78,46%	81,18%	82,79%
937	66,95%	67,68%	70,85%	72,00%	76,15%	77,25%	80,50%	82,00%

Tableau 4.12: Résultats de la classification.

er l'exécution de notre système REMA sur quelques s différentes étapes. Les figures 4.21, 4.22 et 4.23 montrent l'application des quatre algorithmes de squelettisation sur trois échantillons tirés de la base IFN/ENIT, où on remarque que l'algorithme de Zhang et Wang est le meilleur concernant la préservation de la structure du mot.

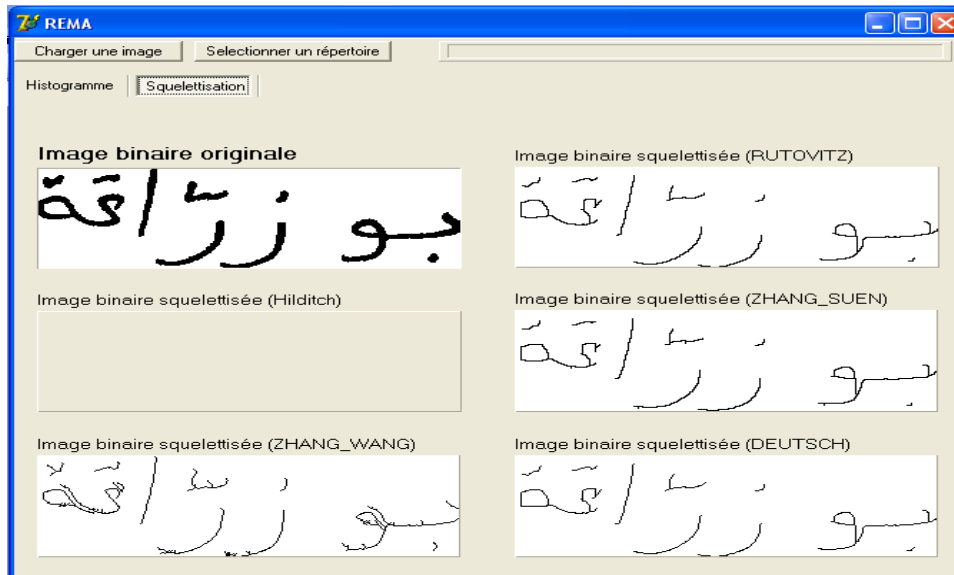


Figure 4.21: Squelettes du mot "بوزراعة" après l'application des quatre algorithmes de squelettisation.

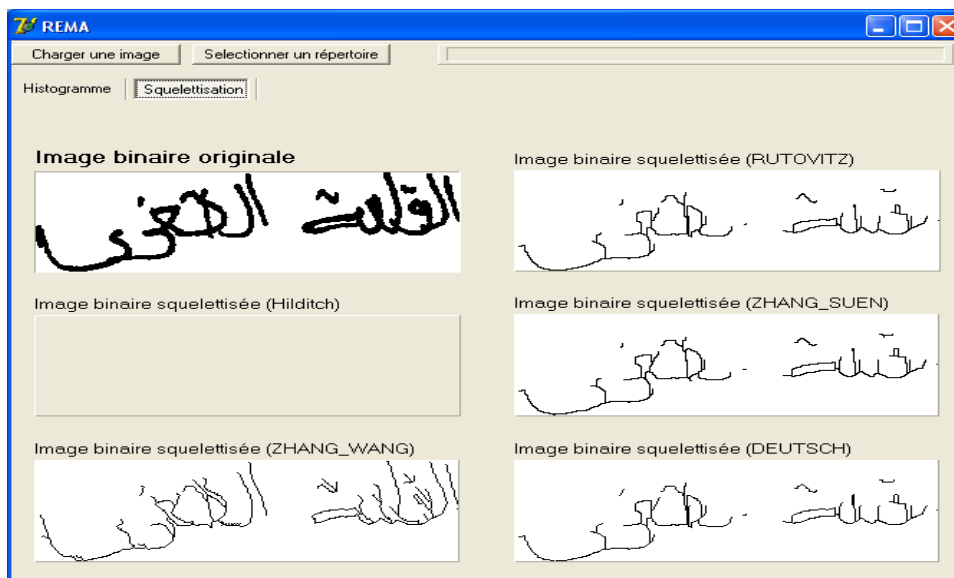


Figure 4.22: Squelettes du mot "القلعة الصغرى" après l'application des quatre algorithmes de squelettisation.



Figure 4.23 : Squelettes du mot "تونس القباضة الأصلية" après l'application des quatre algorithmes de squelettisation.

Les histogrammes de projection horizontale ainsi que la détection de la ligne de base pour les trois mots "بوزراعة", "القلعة الصغرى" et "تونس القباضة الأصلية" sont illustrés dans les figures 4.24, 4.25 et 4.26. Notons ici, que l'histogramme le plus dense est celui coloré en jaune.

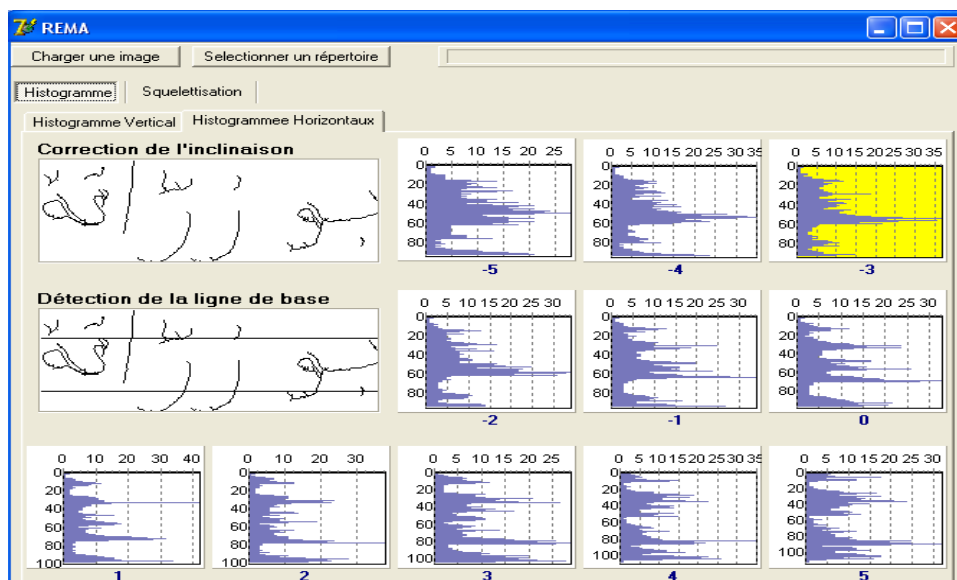


Figure 4.24: Calcul des histogrammes de projection horizontale et détection de la ligne de base pour le mot "بوزراعة".

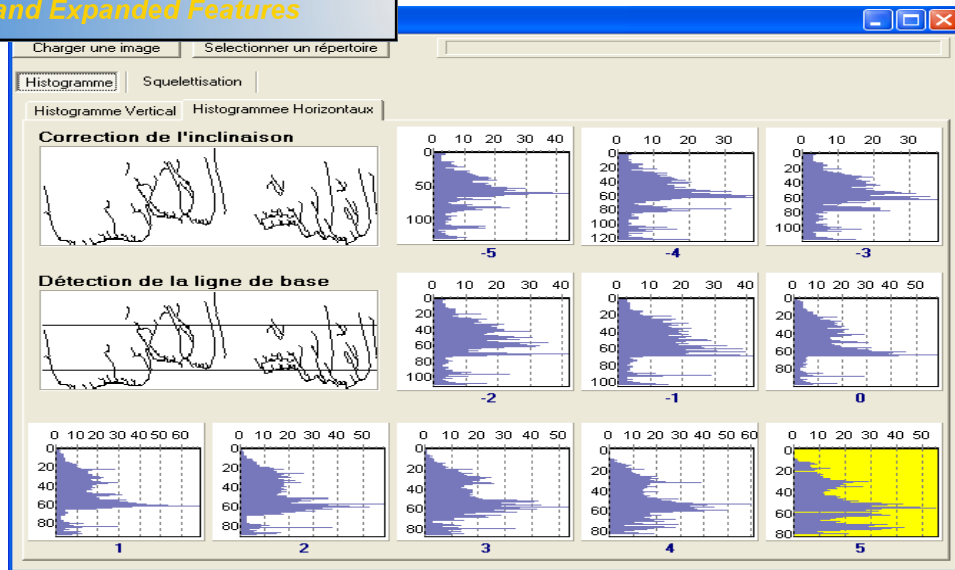


Figure 4.25: Calcul des histogrammes de projection horizontale et détection de la ligne de base pour le mot "القعدة الصغرى".

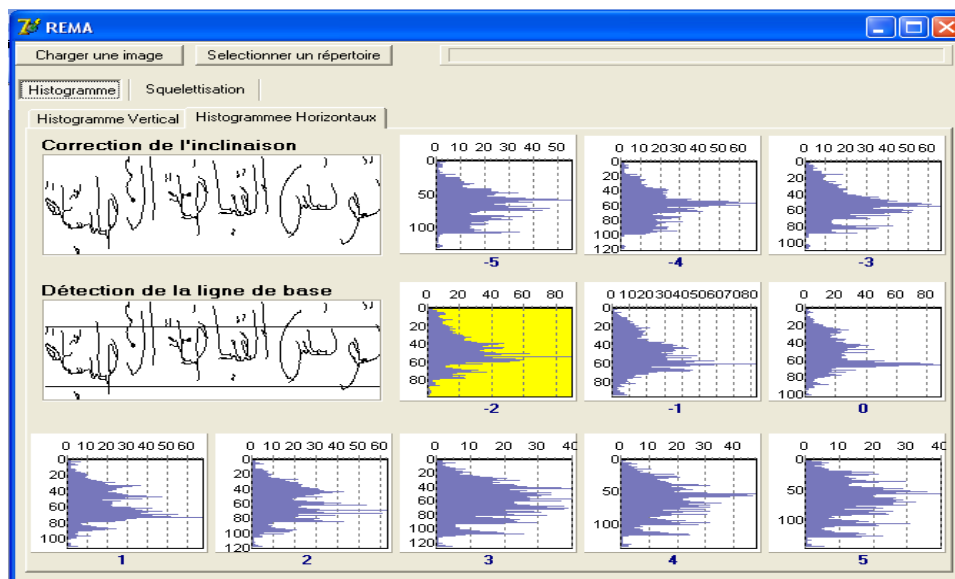


Figure 4.26: Calcul des histogrammes de projection horizontale et détection de la ligne de base pour le mot "تونس القباضة الأصلية".

Les images ci-dessous affichent les résultats obtenus lors du calcul des cinq primitives statistiques utilisées. L'algorithme de squelettisation employé est celui de Zhang et Wang. Le nombre de sous-mots quand à lui, est calculé à partir de l'histogramme de projection verticale. Deux échantillons de chacune des classes des mots "بوزراعة", "القعدة الصغرى", et "تونس القباضة الأصلية" ont été sélectionnés, afin de mieux comparer les mesures enregistrées.

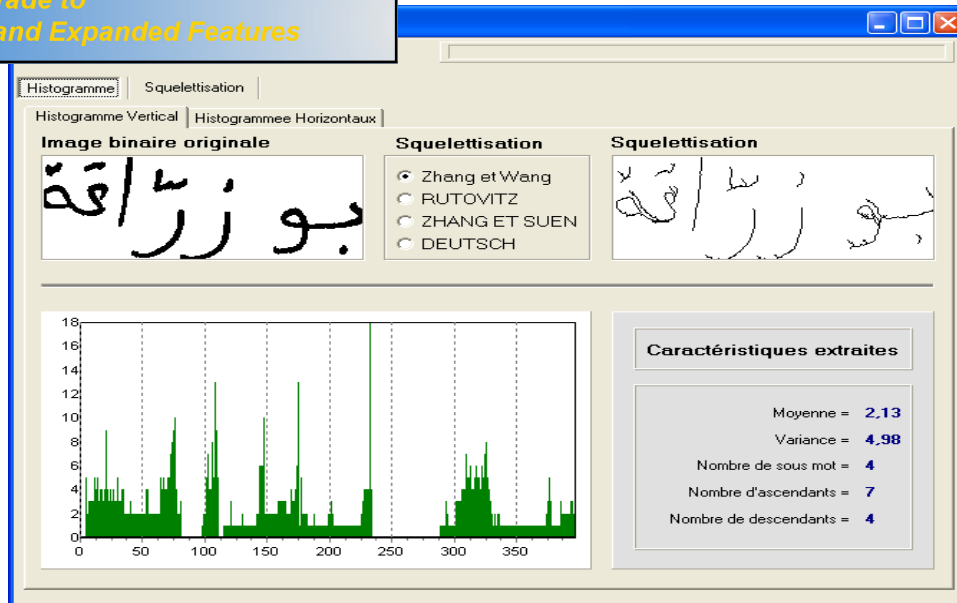


Figure 4.27: Calcul des primitives statistiques et de l'histogramme de la projection verticale pour un premier mot de la classe "بوزراعة".

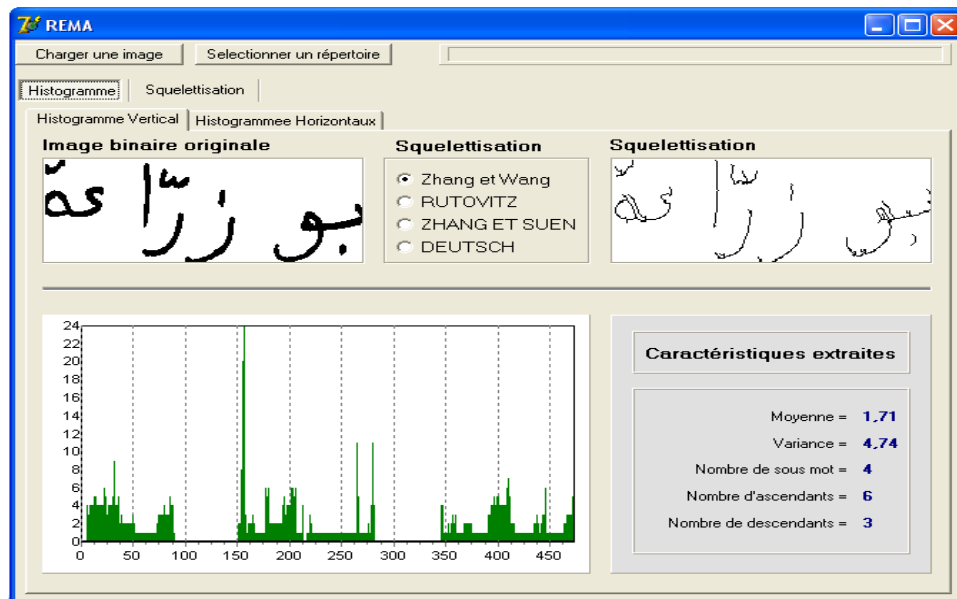


Figure 4.28: Calcul des primitives statistiques et de l'histogramme de la projection verticale pour un deuxième mot de la classe "بوزراعة".

On remarque dans les figures 4.27 et 4.28 que les mesures statistiques enregistrées pour les deux échantillons de la classe "بوزراعة" sont quasiment identiques. Néanmoins, des distinctions peuvent être remarquées concernant le nombre d'ascendants et de descendants causées par la variété intra-scripteurs et les chevauchements entre les caractères.

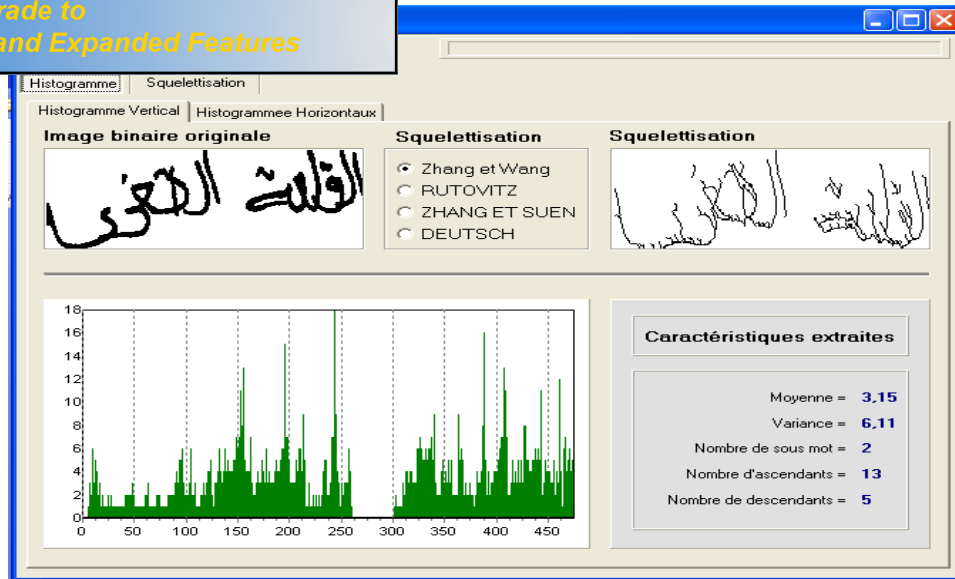


Figure 4.29: Calcul des primitives statistiques et de l'histogramme de la projection verticale pour un premier mot de la classe "القلعة الصغرى".

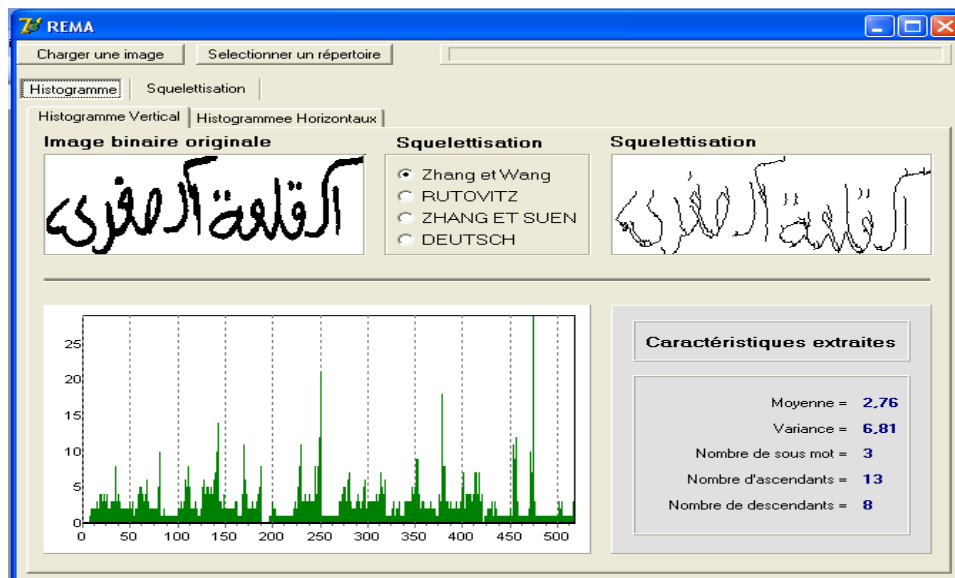


Figure 4.30: Calcul des primitives statistiques et de l'histogramme de la projection verticale pour un deuxième mot de la classe "القلعة الصغرى".

Concernant les figures 4.29 et 4.30 les problèmes cités plus haut mènent quand à eux à des différenciations entre les valeurs du nombre de sous-mots et celui des descendants des deux échantillons de la classe "القلعة الصغرى".

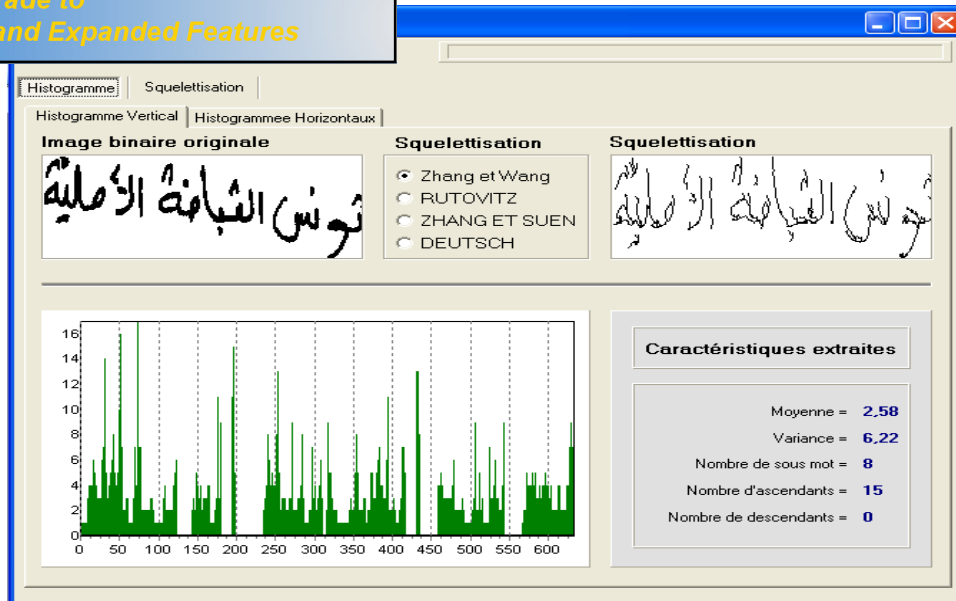


Figure 4.31: Calcul des primitives statistiques et de l'histogramme de la projection verticale pour un premier mot de la classe "تونس القباضة الأصلية".

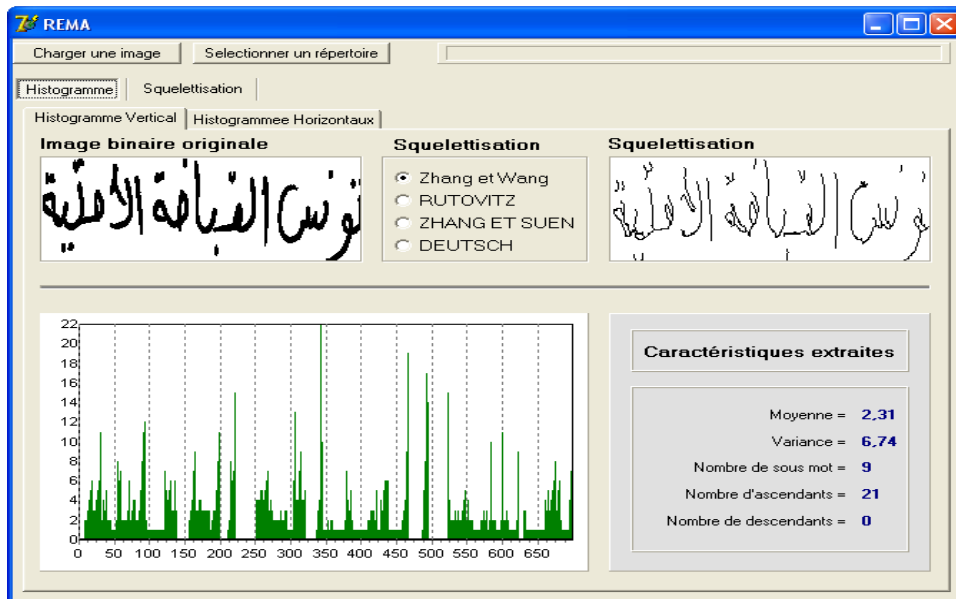


Figure 4.32: Calcul des primitives statistiques et de l'histogramme de la projection verticale pour un deuxième mot de la classe "تونس القباضة الأصلية".

Pour les mêmes problèmes, les deux échantillons du mot "تونس القباضة الأصلية" marquent une différence majeure concernant le nombre d'ascendants.

Les 49 premiers moments de Tchebichef sont indiquées. On constate que pour deux échantillons de la même classe les résultats enregistrés sont presque identiques. Ceci explique l'efficacité des moments de Tchebichef.

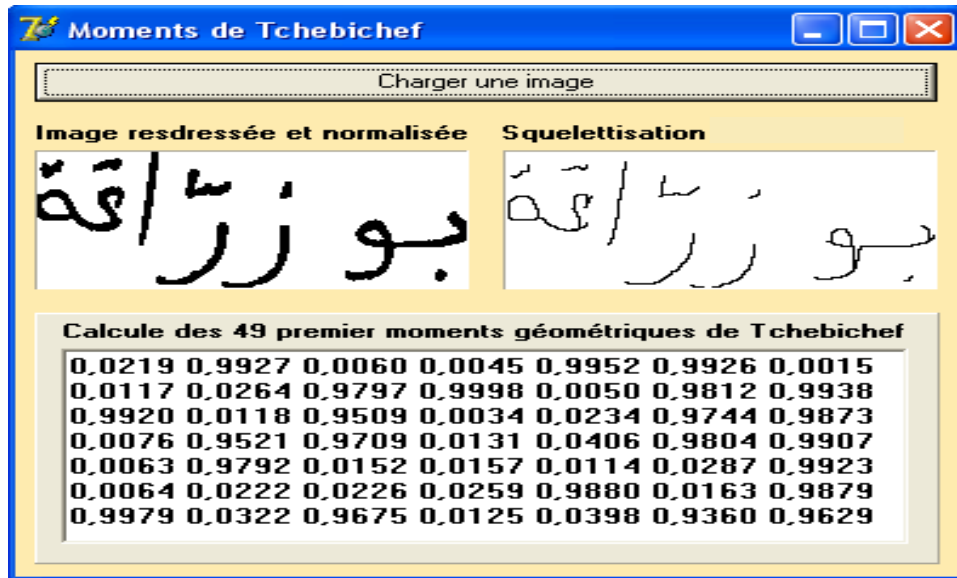


Figure 4.33: Calcul des 49 premiers moments de Tchebichef d'un premier mot de la classe "بوزراعة".

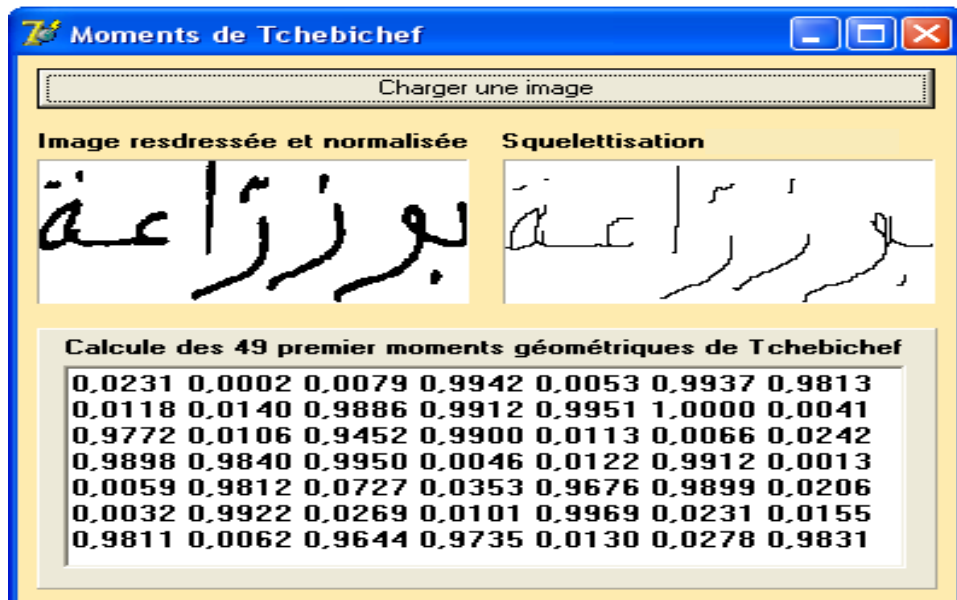


Figure 4.34: Calcul des 49 premiers moments de Tchebichef d'un deuxième mot de la classe "بوزراعة".

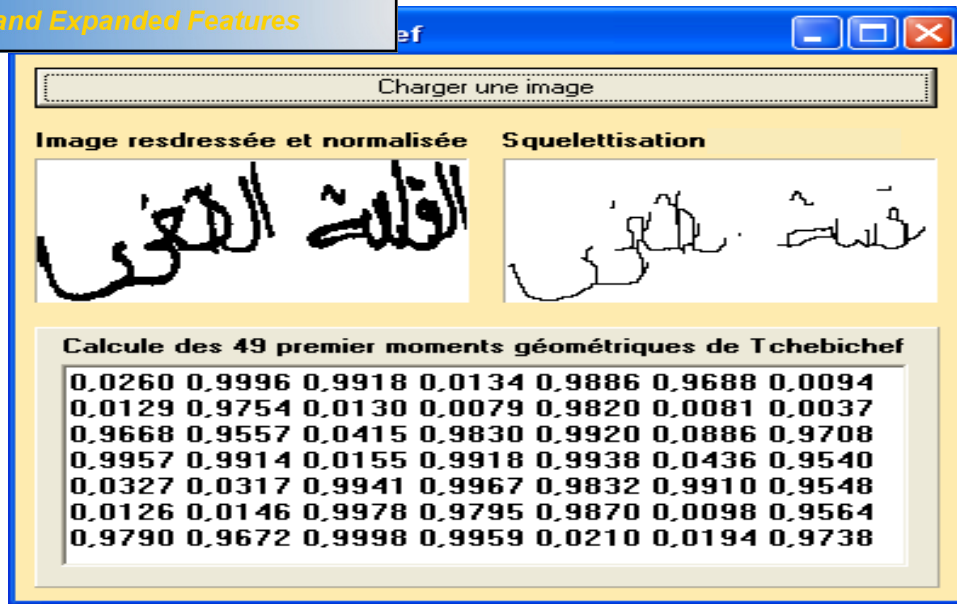


Figure 4.35: Calcul des 49 premiers moments de Tchebichef d'un premier mot de la classe "القلمة الصغرى".

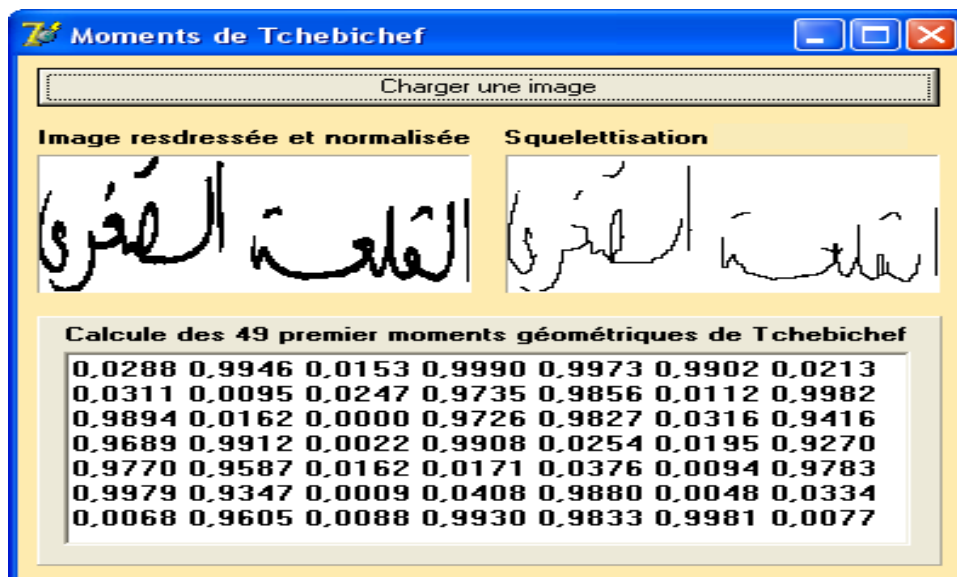


Figure 4.36: Calcul des 49 premiers moments de Tchebichef d'un deuxième mot de la classe "القلمة الصغرى".

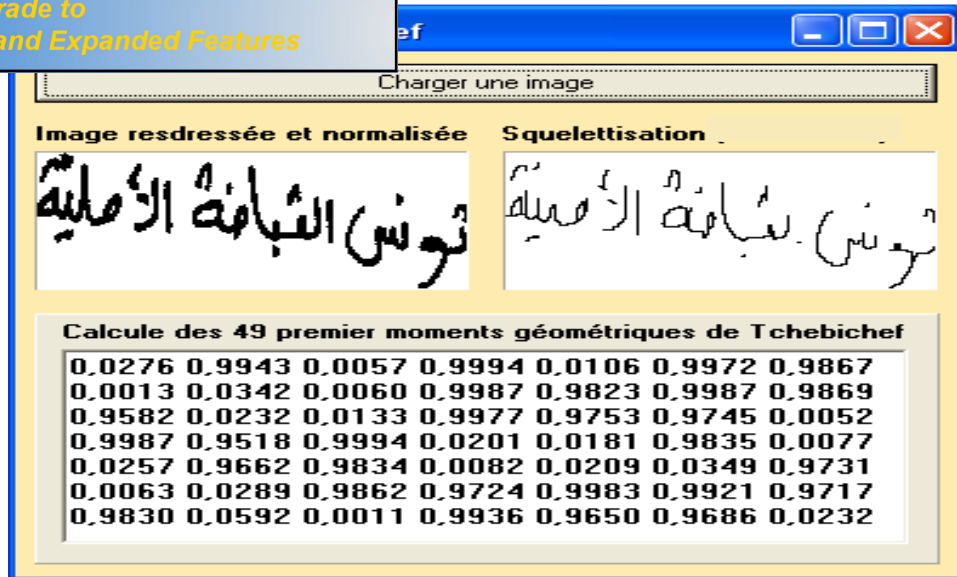


Figure 4.37: Calcul des 49 premiers moments de Tchebichef d'un premier mot de la classe "تونس القباضة الأصلية".

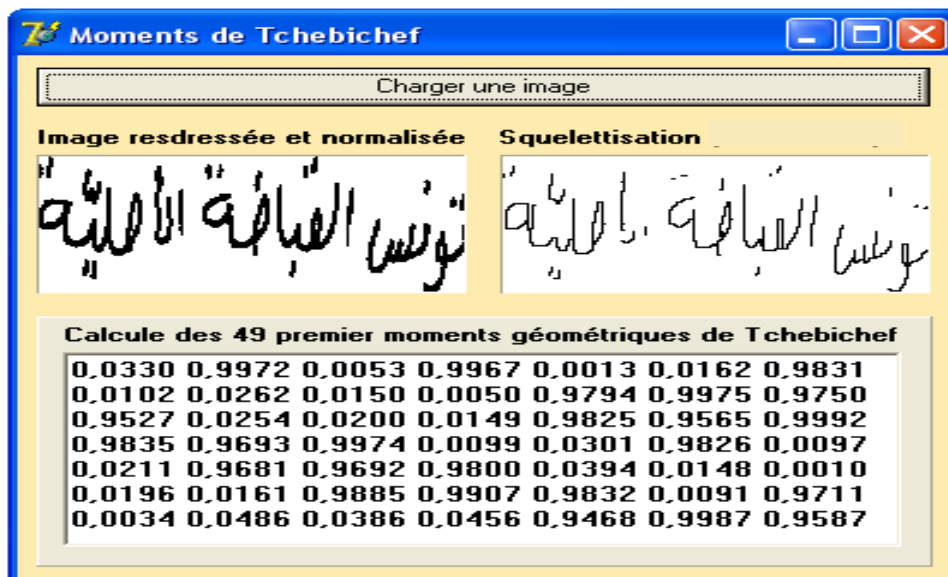


Figure 4.38: Calcul des 49 premiers moments de Tchebichef d'un deuxième mot de la classe "تونس القباضة الأصلية".

Plusieurs difficultés ont été rencontrées durant la phase de reconnaissance engendrant un classement erroné de quelques mots de la base dû : à la mauvaise écriture de certains scripteurs (ce qui est illustré par le tableau 4.13) ; et aux problèmes de chevauchements et de ligatures (indiqué par le tableau 4.14).

أم العرايين	المي	دار علفن	حمي الصلاح
ينبي زلطن	كسه العاليا	بور	بنذرت بلب ماله
حي الحجام	بلاه الدض	الدويرات	ببرء احمد
لاستفطيمية	لسيدي عويد	لمصلا	فندق الجديد
لييا قسه	شنتي	الزارات	فناقص
رادن المدينة	خنيص	مركز التثبية	زسيو
حلمة	الباسيفاتا	الدنسيف المطار	السوابي
الموارف	الكبارف	الحلحة	قصر اللمة

Tableau 4.13: Echantillons des mots mal-écrits.

مخبرة الدحول	الجديدة زراوة	الخليج العمدية
ارديف المحطة	اختمين	متلوي شاجة الحبيبة
الأس	ثالة	ز مور لالة
هبيح	نفة	كان الة
المهدية هيونا	غراف	الفيروان عيلة
عنام	بئر الدولة	مركز درويش اللوزة
بحج السدرية	العرع	تأجروين
العمران	جبل المظبية	أرسة الذراع

Tableau 4.14: Echantillons des mots contenant des chevauchements et des ligatures.



PDF Complete
Your complimentary use period has ended.
Thank you for using PDF Complete.

[Click Here to upgrade to Unlimited Pages and Expanded Features](#)

stèmes

résultats obtenus, des comparaisons avec des systèmes de reconnaissance de mots arabes manuscrits hors-lignes travaillant sur la base IFN/ENIT ont été effectuées. Les travaux concernés sont ceux de :

- **Le système ICRA (Intelligent Character Recognition for Arabic).** Conçu par Ahmed Abdulkader [67], et qui a utilisé un classifieur neuronal de type PMC. L'apprentissage s'est fait sur les ensembles a, b et c de la base IFN/ENIT, tandis que les tests ont été réalisés sur l'ensemble e de cette dernière.
- **Le système SHOCRAN (Système for Handwritten Optical Character Recognition for Arabic Names).** A été développé par un groupe de chercheurs égyptiens [67]. Le système fut entraîné sur les ensembles : a, b, c et d ; pour le test l'ensemble e a été utilisé.
- **Le système TH-OCR.** Développé par Pingping Xiu, Hua Wang, J.Jin, Yan Jiang, L.Peng et X.Din [65]. Ce système a utilisé une méthode de reconnaissance statistique basée sur les réseaux de neurones en appliquant trois niveaux de segmentation, à savoir : la segmentation de texte en lignes, la segmentation de lignes en mots et la segmentation de mots en caractères.
- **Le système UOB.** Conçu par Chafik Mokbel et El Hajj [65] de l'université du Liban, où ils ont utilisé les chaînes de Markov cachées (HMM). les quatre premiers ensembles de la base ont été utilisés pour l'apprentissage et le dernier pour le test.
- **Le système REAM (Reconnaissance de l'Écriture Arabe Manuscrite).** C'est le résultat d'un travail de groupe de chercheurs tunisiens comprenant S. Masmoudi-Touj, N. Essoukhri-Ben Amara, H. Amiri et N. Ellouz [67], et qui ont utilisé une chaîne de Markov de type planaire dans la phase de reconnaissance.
- **Le système ARAB-IFN.** Conçu par Pechwitz [33], utilisant les chaîne de Markov cachées semi-continus d'ordre 1 comme classifieur. Il a appliqué les opérations de prétraitement de normalisation et de squelettisation afin de détecter la ligne de base. Les tests ont été faits sur l'ensemble e de la base IFN/ENIT.

L'évaluation des performances de ces cinq systèmes de reconnaissance de l'écriture arabe manuscrite est indiquée dans le tableau 4.15, où on remarque que les résultats enregistrés par notre système sont nettement meilleurs que ceux obtenus par les systèmes; SHOCRAN, TH-OCR et REAM ; tandis que les systèmes ICRA, ARABE-IFN et UOB donnent des taux de reconnaissances relativement meilleurs que le notre, ceci peut être expliqué par l'expérience



on des types de classifieurs utilisés par ces derniers, à

Nom de système	Top1	Top5	Top10
ICRA	65,74%	83,95%	87,75%
SHOCRAN	35,70%	51,62%	51,62%
TH-OCR	29,62%	43,96%	50,14%
UOB	75,93%	87,99%	90,88%
REAM	15,36%	18,52%	19,86%
ARAB-IFN	74,69%	87,07%	89,77%
REMA	66,95%	76,15%	82,00%

Tableau 4.15: Comparaison entre les performances du système proposé et certains OCRs arabes.

4.4 Conclusion

Nous avons mis au point un système de reconnaissance d'écriture arabe manuscrite, utilisant un réseau Fuzzy ART comme classifieur et les moments de Tchebichef ainsi que quelques caractéristiques statistiques comme primitives. Dans ce chapitre, nous avons présenté l'architecture générale du système proposé et expliqué ses différentes étapes d'implémentation.

Les résultats obtenus sont encourageants comparés à d'autres systèmes travaillant eux aussi sur la base IFN/ENIT. Le taux d'erreur enregistré est principalement dû à la mauvaise écriture et aux problèmes de chevauchement des lettres dans les mots classés.

Notons ici que le taux de reconnaissance peut être amélioré par une combinaison avec d'autres classifieurs et l'ajout d'une étape de poste-traitement