

نظام تعرّف الحرف العربي متعدد الأنماط والأحجام

د. محمد بسام الكردي
المعهد العالي للعلوم التطبيقية والتكنولوجيا

ملخص:

تعمل أنظمة تعرّف الحرف الضوئي على تحويل النصوص المطبوعة على الورق والمسحوبة باستخدام ماسح ضوئي إلى نصوص مخزنة في ذاكرة الحاسب بغية الأرشفة السهلة والسريعة.

تتميز الطريقة التي استُخدمت في هذا المنتج بقلّة الحساسيّة للضجيج والميلانات الخفيفة. وهي سريعة جداً، حيث تحدد آلياً حجم البنت المقروء، وغيره من المتحولات اللازمة.

دعم النظام بقدرات تمكنه من التمييز الآلي واليدوي بين اللغة العربية واللغات اللاتينية وكذلك التعرف الآلي واليدوي على الجداول والأعمدة. يستخدم هذا النظام خوارزمية إعادة البناء من أجل تقسيم السطر إلى مقاطع وإن كانت متراكبة كما هو الحال في بعض الحروف العربيّة.

تصل نسبة التعرف إلى 98% لبعض أنواع الخطوط، وذلك بسرعة لا تتجاوز خمس ثوان للصفحة الواحدة، بدقة 300 ن/إ.

تساعد أنظمة تعرّف الحروف الضوئي على تخزين كميات كبيرة من المعلومات المطبوعة على الورق في ذاكرة الحاسوب باستخدام ماسح ضوئي للصور بغية الأرشفة السهلة والسريعة، كما يساعد على تحسين الواجهات التخطيبية بين الآلة والإنسان، إضافة إلى تطبيقات أخرى.

تتميز الطريقة التي استُخدمت في هذا المنتج بقلّة الحساسيّة للضجيج والميالات الخفيفة، وهي سريعة جداً، حيث تحدد آلياً حجم البنط font المقروء وغيره من المتحولات اللازمة، يجري دعم النظام بقدرات تمكنه من التمييز الآلي واليدوي بين اللغة العربيّة واللغات اللاتينية وكذلك التعرف الآلي واليدوي علي الجداول والأعمدة.

في الحقيقة، يستخدم هذا النظام خوارزمية إعادة البناء Reconstruction Algorithm، المعروفة في علم المورفولوجيا الرياضية Mathematical Morphology، من أجل تقسيم السطر إلى مقاطع وإن كانت متراكبة (كما هو الحال في بعض حروف اللغة العربية)، ثمّ تتم عملية التعرف على مرحلتين (مستويين) لزيادة السرعة باستخدام شكل الحرف العربي أو اللاتيني معياراً في ذلك، حيث نستخدم في المستوى الأول معلومات غير قياسيةّ متعلقة بالتقعرات والتحدبات والحلقات المغلقة، يتم فيها فرز الحروف ضمن صفوف، لا يتجاوز عدد حروف الصف سبعة أو ثمانية أحرف، وتعتبر هذه الصفوف دخل المستوى الثاني للتعرف، حيث يتمّ خلاله التمييز بين الأحرف المكونة لصف ما، وذلك بطريقة تجريبية بحتة.

في الختام، لقد استعملنا طريقة تمتاز بالسهولة والمتانة بآنٍ واحد، كما أنها قابلة بسهولة للتحويل إلى الواجهة الصلبة.

طرائق تعرّف الحرف العربي

تمهيد:

تعددت في الآونة الأخيرة أنظمة تعرّف الحرف العربي - وخصوصاً المطبوع - والدراسات في هذا الاتجاه، ولكن هذه الدراسات والنتائج تبقى فقيرة جداً نسبياً وهي غير مرضية ومتواضعة من جهة السرعة والأداء والسهولة.

يتألف نظام القراءة الضوئية، عادةً، من ثلاث مراحل: مرحلة ما قبل المعالجة أو ترشيح الحرف، مرحلة استخلاص المميزات ومرحلة عملية القرار (الشكل 1-1).

يمكننا فصل طرائق تعرف الحروف إلى نوعين متباينين وضوحاً من أساليب المعالجة (الشكل 1-1):

أسلوب من النوع القراري وهو من طبيعة إحصائية في أغلب الحالات.

أسلوب من النوع النحوي أو البنيوي: وهذا الأسلوب يمثل عموماً عملية وصف وتأويل ذات مستويات عدة، يقابل المستوى الأعلى منها المطابقة النهائية للحرف بينما تقابل المستويات الأدنى منه التأويلات المتوسطة للشكل.

إن الأسلوب الذي اتبعناه في هذا النظام والمفصل فيما يلي هو الأسلوب البنيوي (الأسلوب الثاني). يكمن الابتكار الأساسي لهذا النظام في انتقاء الأصول primitives وفي الطريقة التي تعين على الحصول عليها، تعتمد هذه الطريقة على أدوات المورفولوجية الرياضية (2).

استخلاص الأصول primitives أو المميزات: المحاور الأساسية في البحث:

قبل وصف الطريقة المقترحة، لنذكر، وبشكل مختصر، الطرائق الممكنة في حال استخلاص المميزات، وهي المرحلة الهامة جداً في أي عملية في القراءة الضوئية. لقد كرّس قسم هام من أعمالٍ عدة لهذه الدراسة المكتبية (راجع (1، 3)). يمكننا تصنيف هذه الطرائق بفئات حسب الفكرة التي نريد إظهارها. لا يمكن لهذا التصنيف أن يدّعي الموضوعية، لأن، وكما سنشاهد فيما يلي، هذه الفئات تتفاعل فيما بينها بشكل واسع.

آ - الطرائق العامة:

الطريقة الأسهل التي يمكننا اعتمادها، هي التي نستعمل فيها النافذة الحاوية على الحرف المطلوب تعرّفه ونفحص كل عنصر صورة (بكسل pixel) لمعرفة إذا كان أسوداً أو أبيضاً (0 أو 1). هذه الطريقة ذات حساسية عالية للتعرجات وتطبيقها لا يجدي إلا في حالة الطباعة أحادية الخط (monofont) (4).

إن أفضل طريقة من الطرائق العامة هي، باعتقادنا (راجع (5))، تلك التي تحاول التقليل ما أمكن من عدد البكسلات (pixels) المفحوصة، وذلك بفضل عملية البدء بالتلقين لإنشاء شجرة من قاموس الأشكال (الأحرف) المطلوب الفصل فيما بينها. تندرج الطرائق التي تستخدم الأقنعة أيضاً تحت هذه الفئة آ حيث كانت هي أولى المقترحات من قبل مصممي الأنظمة. هذه الطرائق يمكن استعمالها بعد إجراء على الحرف التحويلات الجديرة بتسهيل استخلاص المميزات، كتحويلات فوربيه (6) أو كارهوين-لوف (Karhuen-Loeve). وفي الختام، في علم المورفولوجية الرياضية، يمكننا ذكر، ضمن الطرائق العامة، تلك التي تستعمل عدد أولر Euler، عدد التقعرات والتحدبات أو التي، بعد عملية التنحيف، تبحث عن عدد النهايات أو التقاطعات.

ب- المميزات المحلية:

ضمن هذه الفئة ب، تسعى كل الطرائق لاستخراج المعلومات الخاصة لكل اتجاه (في مصفوفة الصورة) أو المعلومات المتعلقة بالزوايا وذلك باستعمال المحل الهندسي لهذه المعلومات (تقسيم النافذة لعدة مناطق خاصة). من أجل استخراج الخواص المحلية، عدة أنظمة مصممة قد

استعملت الأقنعة ذات الأمكنة الثابتة أو المتغيرة.

ج- طريقة إتباع المحيط:

إن معلومة المحيط تسمح بالحصول على مختلف الخواص المحلية (الميول، الزوايا، ...الخ). كذلك، إن التمثيل الدقيق للمحيط يعطي ميزة كبرى للتعرف. والتمثيل الأكثر استعمالاً هو تمثيل فريمن H. Freeman (7).

د- الميول وأنصاف أقطار الانحناءات:

إن المعلومات المتعلقة بهذه الطريقة هي من نوع المقاس (المتري) إذاً هي ذات حساسية للتعرجات ولكن، إذا أضيفت إلى مميزات أخرى كالتقعرات أو التحدبات، يمكن استخدامها بنجاح للفصل بين الحروف المطبوعة.

هـ- استعمال شبكة:

استعملت عدة عمليات (3) هذه الطريقة التي هي عبارة عن وضع شبكة على الحرف ومراقبة عدد التقاطعات الناتجة عن كل مستقيم من المستقيمت المكونة للشبكة. يمكن اختيار اتجاه المستقيمت عمودياً، أفقياً أو بطريقة عشوائية.

و- تحليل التقعرات والحلقات:

هذه المميزات هي من نوع طوبولوجية وتمتاز بجودة عالية لمناعتها ضد التعرجات.

ز- التمثيلات من نوع نحوي:

تعتبر هذه التقانة البنيوية الطريق الذي لاقى اهتماماً كبيراً خلال السنوات الماضية حيث تم التوجه نحو الذكاء الصناعي لحل مشكلة تعرف الأشكال. يبنى هذا النوع من التمثيل (8، 9، 10) على عملية ذات ثلاث مراحل (الشكل 1-1):

آ- اختيار العناصر المكونة المسماة "الأصول" primitives لتعريف مجموعة من الرموز

المكونة لمفردات مرجعية.

ب- وصف علاقات بنيوية بين هذه الأصول primitives (شكل التسلسل المنطقي للمقاطع، ... الخ) لتشكيل، كذلك، جمل أو بيانات وصفية ناتجة عن جمع أو تسلسل أسماء الأصول primitives.

ج- تعريف قواعد أو إنشاء مجموعة من القواعد لإعادة كتابة تجميع الرموز.

يتم التعرف، بعدئذ، بواسطة محلل بنوي syntactical analyzer.

3- مكونات نظام تعرّف الحرف:

2-1- التقاط وترشيح وتقسيم الصورة:

يمكننا من أجل التقاط الصور، استعمال إما الكاميرا أو الماسح أو وحدات تخزين الصور المأخوذة سابقاً (الأقراص المرنة، الشرائط، الخ ...). إن الخطوة الأولى هي الحصول على صورة ثنائية binary بعد المسح وذلك باختيار عتبة مناسبة.

إذا كان تقسيم بعض الصور باستعمال العتبة يتم دون صعوبة تذكر، فإن أغلبية الحالات (بسبب عنصر إلكتروني، التصوير الفوتوكوبي، ...)، يكون تباين التركيب بشكل لا يسمح بتقطيعه بمجرد استعمال عتبة بسيطة. من أجل حل هذه المشكلة، يمكننا اللجوء إلى تحويلة مورفولوجية تدعى "القبعة العالية" (13) TOP HAT تستخرج هذه التحويلة من الصورة قمم (على التوالي حُفَر) التباين التي تحوي على ارتفاع (على التوالي عمق) وعلى عرض معطين وتسمح بذلك، بجودة عالية، بتقسيم الصورة، حيث يندمج الشكل والأرضية على الدوام.

ومن ثمّ تجري عملية التقطيع لأسطر ثم لمقاطع وأخيراً لحروف مستقلة أو زوج من الحروف لتدخل عملية التعرف بمستويها الأول والثاني.

وسنسرّد فيما يلي الطرائق المستخدمة بالترتيب الأفضلي في كل جزء من الأجزاء الأربعة التي يتكون منها أي نظام للقراءة الضوئية:

1- تقطيع النص إلى أسطر:

1-1- الطريقة الكلاسيكية:

وهي ببساطة البحث عن سطر فارغ بين سطرين، أي البحث عن سطر من مصفوفة النص يحوي أصفاراً بكامله إذا فرضنا أن النص المكتوب ممثل بمصفوفة آحاد. ولكن المشكلة هي تداخل بعض الأسطر مع بعضها في بعض النصوص، وخصوصاً على سبيل المثال حرف "م" من السطر الأعلى وحرف "أ" من السطر الأدنى.

1-2- طريقة استخدام العتبة threshold:

وهي كسابقتها مع محاولة إهمال بعض التقاطعات وذلك باستخدام عتبة، أي يُعتبر السطر فارغاً إذا حوى عدد من الآحاد لا يتجاوز العتبة. وتكمن المشكلة هنا في التقطيع نفسه، إذ لا يمكن إلحاق قسم من حرف من السطر الأعلى بالسطر الأدنى ولا العكس.

1-3- طريقة الكثافة النقطية العظمى maximum density line:

وهي من أفضل الطرائق، حيث يتم إجراء إحصاء للكثافة النقطية histogram للنص بكامله وتكون عندئذ القمم العليا للكثافة هي الأسطر المراد استخراجها، وتُنزع extracting باستخدام خوارزمية البناء reconstruction algorithm المعروف في علم المورفولوجيا الرياضية Mathematical Morphology.

2- تقطيع السطر إلى مركبات متصلة:

1-2- الطريقة الكلاسيكية:

تتبع الطريقة الكلاسيكية نفس الأسلوب في التقطيع إلى أسطر، ولكن بالبحث عن سطر فارغ (صف من الأصفار) شاقولياً، تظهر هنا مشكلة التراكم بين الحروف overlapping مثل تراكم حرفي الراء والياء في كلمة "يرى"، حيث لا يمكنه فصل المركبة "ير" عن "ى".

2-2- طريقة البناء Reconstruction:

وهي الطريقة المتبعة، وتلخص بإيجاد أي نقطة سوداء (تساوي الواحد) من المقطع المتصل ثم بناء

المركبة المتصلة كلها بملاحقة كل النقاط السوداء المتصلة بهذه النقطة السوداء حتى ولو كانت تمتد تحت المركبة التي تليها.

3- تقطيع المركبة المتصلة إلى أحرف:

3-1- الطريقة الكلاسيكية:

نحاول تقطيع المركبة إلى أحرف منفصلة باستخدام قواعد الكتابة في اللغة العربية، وذلك بالتقطيع الشاقولي والأفقي كما في الأمثلة التالية "لم"، "نم"، "تم"، ... الخ، التي من الصعب تقطيعها أفقياً.

3-2- طريقة الأحرف الثنائية:

وهي تعتبر كل المركبات ذات الأحرف المتراكبة أحرفاً ثنائية نتعرف عليها كجزء (حرف وحيد) له رمزين عوض عن رمز واحد بعد التعرف. في هذه الطريقة نزيد في عدد الحروف في اللغة العربية ولكن تبقى دون المئة ونحل هكذا مشكلة التقطيع الأفقي كالأمثلة السابقة.

MIDDLE DOWN → C_P
MIDDLE UP DOWN →
UP MIDDLE →
UP DOWN →
UP MIDDLE C_P →

Example:

:UP MIDDLE UP MIDDLE DOWN قيل
:MIDDLE DOWN ن
:UP DOWN ل

4- تعرف الحرف:

4-1- الطرائق الكلاسيكية:

تعتمد الطرائق الكلاسيكية على مقارنة صورة الحرف الناتج مع مجموعة من الصور المرجعية المحفوظة في الذاكرة. هذه الطريقة بطيئة جداً بسبب ما تتطلبه عملية المقارنة من زمن كما أنها تحتاج إلى حجم ذاكرة كبير من أجل بنط معين من الخطوط، وهي طريقة غير متعددة البنوط

4-2- طريقة التثخيف والملاحقة **Thinning and following**:

تعتمد الطريقة على تثخيف الحرف **thinning** ليصبح ذا سماكة دنيا (تُقدر بواحد بكسل (pixel ثم ملاحقة الحرف من بدايته لنهايته مع تسجيل تغير حركته من صعود وهبوط ودوران... الخ، ليتم تعرفه. من مساوي هذه الطريقة أنها طويلة زمنياً بشكل يُخرج النظام من كونه منتج بالإضافة إلى المشاكل الخاصة باللغة العربية مثل الأحرف المطموسة كالعين في وسط الكلمة تصبح كالنبرة!.

4-3- الطريقة المورفولوجية **Morphology**:

وهي طريقة متعددة البنوط والأنماط حيث تتم عملية التعرف خلال مرحلتين (مستويين) لزيادة السرعة باستخدام شكل الحرف العربي معياراً في ذلك. حيث نستخدم في المستوى الأول معلومات غير قياسية متعلقة بالتقعرات والتحدبات والحلقات المغلقة، يتم فيها فرز الحروف ضمن صفوف، لا يتجاوز عدد حروف الصف عشرة أحرف. وتعتبر هذه الصفوف دخل المستوى الثاني للتعرف، حيث يتم خلاله التمييز بين الأحرف المكونة لصف ما، وذلك بطريقة تجريبية بحثة **heuristic**.

2-2- وصف المستوى الأول:

بما أننا اعتمدنا التوصيف ذا المستويات المتعددة (اثنان في حالتنا)، بعد اختيار الأصول **primitives** أو المعلومات التي نأخذ بها بعين الاعتبار في هذا الوصف، هناك سؤال يطرح نفسه: ما هي الأصول **primitives** المختارة في المستوى الأول؟

في الحقيقة، إن ترتيب تتابع الخواص، تماماً مثل تتابع التحولات المورفولوجية (12) في المعالجة الآلية، لا يمكن أخذه عشوائياً لأنه يؤثر بشكل غير قابل للجدل على نتيجة التعرف. كذلك، بما أن مدخل المصنّف في المستوى الثاني هو ناتج المصنّف في المستوى الأول، يجدر بنا أخذ من أجل هذا الأخير، حيث يتطلب دقة كبيرة، الأصول **primitives** الأكثر أداءً. نتيجة لذلك أخذنا، كميزات في المستوى الأول، تقعرات الحرف المتعلقة بكل من الجهات الأربعة للمستطيل الحاوي على الحرف كذلك أيضاً عدد الحلقات المغلقة. في الحقيقة، إن الخواص

المتعلقة بجهة التقعرات، مكانها وعددها تسمح بتقليل وبشكل واضح مجموعة البحث في البداية. بنهاية هذه المرحلة، تُتبع الشكل الناتج لصف من مجموعة من الأشكال المعدودة.

بالنسبة للأحرف والأرقام، فإن عدد التقعرات في الجهات الأربعة للمستطيل المحيط بالحرف تساوي 0 أو 1 أو 2 ويمكنها عندئذ أن تُرمز بمتحولين ثنائيين (على سبيل المثال bits (L1, L2): للجهة اليسارية، (R1, R2) للجهة اليمينية، (B1, B2) للجهة السفلى، (T1, T2) للجهة العليا. يُرمز عدد الحلقات أيضاً بمتحولين ثنائيين (O1, O2).

في حال انعدام التقعرات من الجهة اليسارية، على سبيل المثال، فإن المتحولين L1 و L2 تساوي 0.

وهكذا، في المستوى الأول للتصنيف، كل حرف مرمز على 10 bits (10 متحولات ثنائية) b1, b2,...,b10 كما نحصل على قيمة المتحولين O1 و O2 مباشرةً اعتباراً من حساب عدد أولر Euler للحرف المعزول.

2-3- التصنيف في المستوى الأول:

بعد الحصول على المعلومات المتعلقة بالتقعرات والحلقات على شكل 10 متحولات ثنائية b1, b2, ..., b10، يمكننا إلحاق الحرف، الخاضع للتعريف، بأحد الصفوف الحاوية على 38 حرف. تنفذ هذه العملية بحساب دلالة الحرف في جدول المقابلة (LUT)، يحوي الصف على قيمة العنصر في الجدول ذي الشكل التالي:

$$LUT = \sum_{j=1}^{10} b_j \cdot 2^{j-1}$$

يمكننا أن نلاحظ أن الصفوف الناتجة عن هذا التصنيف ليست بالضرورة منفصلة. بالحقيقة، إن سبب تغيير خط الكتابة، يمكن للحرف الواحد الانتماء إلى أكثر من صف واحد.

2-4- وصف المستوى الثاني:

تناولنا، في المرحلة السابقة، وصف لتمثيل أولي للشكل، قليل الإسهاب، الذي سمح للمصنّف إجراء فرز أولي يُلحق الحرف بصفٍ من الحروف المتشابهة (10 أحرف على الأكثر).

في هذا المستوى الثاني، نقترح، اعتماداً على مخرج التصنيف للمستوى الأدنى، بتهذيب وصف الشكل للصفوف التي تحوي أكثر من حرف واحد، بهدف إتمام التصنيف. استعملنا في المستوى الأول، التقعرات لتمييز الأشكال. وبغية الحصول على الدقة والسرعة (عدد أقل للمتحولات الشائبة)، اكتفينا بعدد هذه التقعرات والحلقات من أجل ترميز الحروف. ومن أجل إتمام وصف الأشكال، رأينا من المفيد متابعة، في المستوى الثاني، بنفس الطريق إلى أن نتقدم نحو التصنيف النهائي.

اخترنا من بين المعلومات المتممة القادرة على المساهمة بتقليل مجموعة الصفوف، مكان التقعرات (أو التحدبات) وعمقها.

إن فحص كل صف من الحروف، التي لم يسمح التصنيف في المستوى الأول ولا المعلومات المتممة المذكورة أعلاه في تنفيذها، قد قادنا لتجزئة المستطيل المحيط بالحرف إلى مناطق خاصة واعتبار، من أجل كل منها، تواترات الدخول في الحرف في كل من الجهات الأربع الرئيسية في المصفوفة الرباعية.

لقد اعتبرنا أن خيار الخمس مناطق الخاصة (الشكل 2-2) لإجراء القياسات الإسقاطية في الاتجاهات الأربعة الأساسية أو الثمانية (في حال إضافة الاتجاهات القطرية الأربعة الأخرى) (يتعلق الاتجاه بالمنطقة المعنية (الشكل 2-2)) يسمح بإزالة الغموض ويفصل الحرف بنجاح تام. تؤكد لنا النتائج الواردة هذا الأمر.

الشكل 2-2: المناطق الخمس الخاصة للحرف.

2-5- التصنيف في المستوى الثاني:

يقدم لنا وصف المستوى الثاني معلومات متممة لتلك التي سمحت بإجراء الفرز الأول، ويسمح ذلك بإنشاء تجريبية (إبداعية) heuristic تتمم التصنيف بتعرّف الحرف الخاضع لنظام التعرف. مستخدمين المتحولات القياسية (المترية) (كعمق النقر أو مكانه... الخ) والتنقيط وما شابه ذلك.

المراجع:

- GAUTHIER M., KLEIN J.C., SERRA J., "Enquête sur la lecture optique". (1)
Rapport interne du Centre de Morphologie Mathématique, Ecole des Mines,
Fontainebleau, 1981.**
- SERRA J., "Image analysis and mathematical morphology". Academic (2)
Press, London, 1982.**
- GAILLAT G., BERTHOD M., "Panorama des techniques d'extraction de (3)
traits caractéristiques en lecture optique des caractères". Congrès AFCET-IRIA,
Reconnaissance des Formes et Intelligence Artificielle, Toulouse, Sept. 1979.**
- STENTIFORD F.W.M., "Automatic feature design for optical character (4)
recognition using an evolutionary search procedure". IEEE Trans. on Pattern
Analysis and Machine Intelligence, Vol. Pami-7, N. 3, May 1985.**
- THIESSE B., MARTHON P., BRUEL A., "Recherche optimale de points (5)
caractéristiques dans un dictionnaire d'images. Applications à la reconnaissance de
caractères imprimés". Congrès AFCET-IRIA, Reconnaissance des Formes et
Intelligence Artificielle, Toulouse, Sept. 1979.**
- LAI M.T.Y., SUEN C.Y., "Automatic recognition of characters by Fourier (6)
descriptors and boundary line encodings". Pattern Recognition, Vol. 14, Nos. 1-6,
pp. 383-393, 1981.**
- FREEMAN H., "On the encoding of arbitrary geometric configurations". (7)
IRE Trans. on Electronic Computers, pp. 260-268, June 1961.**
- BENALI M., "Du choix des mesures dans des procédures de reconnaissance (8)
des formes et d'analyse de texture". Thèse de Doctorat, Ecole des Mines de Paris,
1986.**
- BOUVIER G., HERAULT J., "Méthode de description de caractères par (9)
approximation polygonale". Colloque sur le Traitement du Signal et ses
applications, Nice, Juin 1981.**
- CHEHIKIAN A., FERYAL H.H., "Une description structurelle, peu (10)
redondante, des caractères alphanumériques, au moyen de variables binaires".
Congrès AFCET-IRIA, Toulouse, Sept. 1979.**

- SIMON J.C., "Invariance en reconnaissance des formes". Cognitiva 85, (11)
Paris, Juin 1985.**
- SERRA J., "Structures syntaxiques en morphologie mathématique". (12)
Premier Colloque Image, Biarritz, 1984.**
- KURDY M.B., "Transformations Morphologiques Directionnelles et (13)
Adaptatives: Applications aux sciences des Matériaux". Thèse de Doctorat, Ecole
des Mines de Paris, 1990.**
- VINCENT L., "Algorithmes morphologiques à base de files d'attente et de (14)
lacets, Extension aux graphes". Thèse de Doctorat, Ecole des Mines de Paris, 1990.**
- KURDY M.B., JOUKHADAR A., "Multifont Recognition System for Arabic (15)
Characters". Proceedings of the 3rd International Conference and Exhibition on
Multilingual Computing, Durham U.K., Dec. 1992.**
- MRAYATI M., "Applied Arabic Linguistics for Informatics". Arab school (16)
(Computers and the Arabic Language), 1985.**
- KURDY M.B., JOUKHADAR A., Wabbi A. "Multifont Arabic/Latin Optical (17)
Characters Recognition System". Proceedings of the IIe International Congres
(Langue Arabe et Technologies Informatiques Avances), Maroc, Déc. 1993.**