

Reconnaissance Automatique de l'écriture

Dr. Djellali Hayet

Master 1 IATI

Département informatique

Université d'UBMA Algérie

Mars 2020

Reconnaissance Automatique de l'écriture

- * Définition
- * Applications
- * Le processus de Reconnaissance de l'écriture
- * Les méthodes de la Reconnaissance Ecriture
- * Base de données

Applications

- * Reconnaissance des textes manuscrits
- * Reconnaissance d'écriture des chèques postaux
- * Reconnaissance écriture des textes anciens
- * Reconnaissances des scripteurs

Reconnaissance d'écriture

Exemple : Pour un document

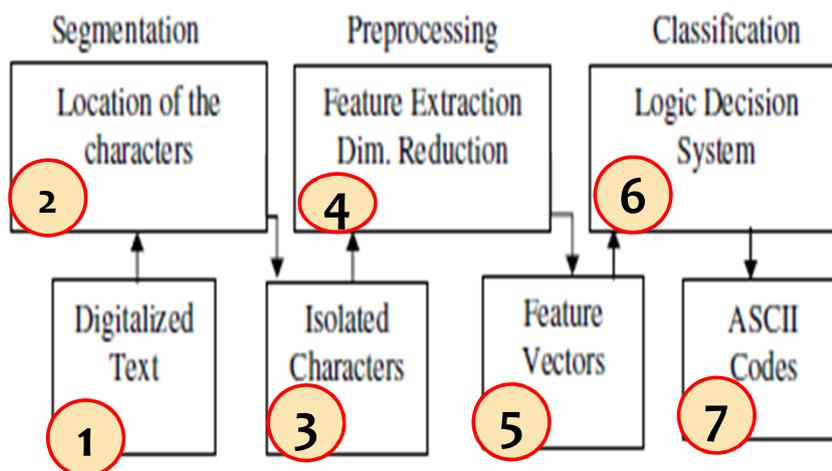
- * Phase 1:
 - * Document: identification de ligne
 - * Segmentation de ligne
 - * Localization de texte (ligne)
- * Phase 2: niveau mot
 - * Prétraitement
 - * Segmentation en caractères
 - * Reconnaissance de Caractères

Etape de la Reconnaissance de l'écriture

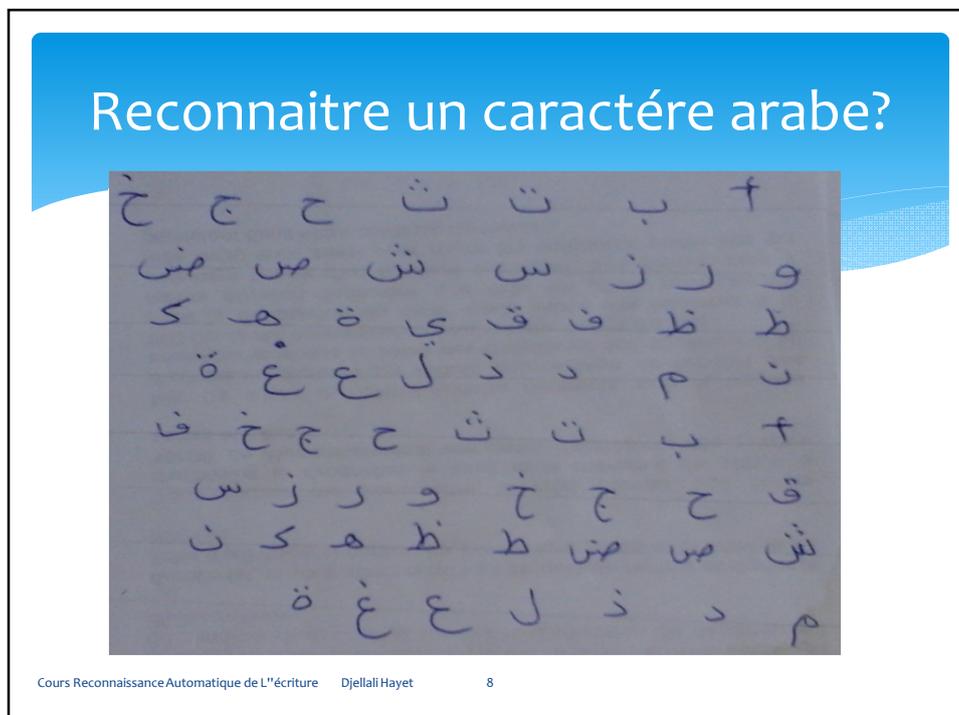
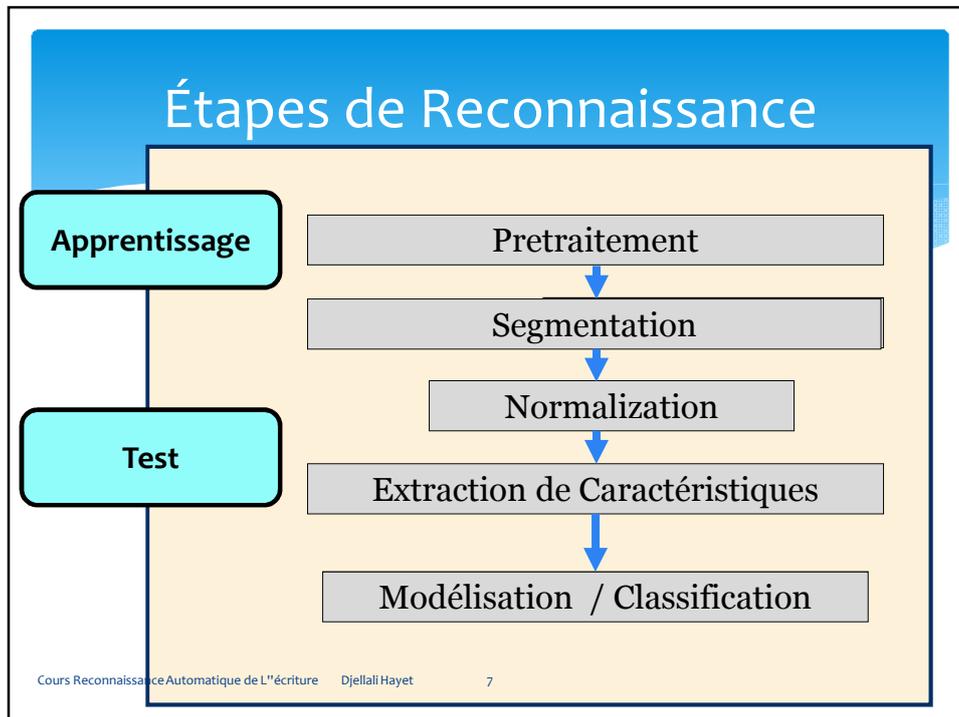
- * Les phase de la reconnaissance de l'écriture
- * Collection de la Base
- * Prétraitement
- * Extraction de caractéristiques
- * Modélisation
- * Classification

Cours Reconnaissance Automatique de L'écriture Djellali Hayet 5

Composants d'un système de reconnaissance de caractères



Cours Reconnaissance Automatique de L'écriture Djellali Hayet 6



Apprentissage Supervisé: Trouvez f ?

- * Exemple de reconnaissance de caractères:
- * La base de données base:
- * Exemple: base d'image dont chaque image contient un caractère.
- * Cette base est divisé en :
 - * Base d'apprentissage (train) : images du caractere dite X_{train} ,+sa classe dite Y_{train}
 - * Base de Test: image du caractere de test (X_{test}) + sa classe(Y_{test})
- * Etant donné: base d'apprentissage $\{(x_i, y_i) \mid i = 1 \dots n\}$
- * Trouvez: une bonne approximation de $f : X \rightarrow Y$
- * *Pas de segmentation dans ce cas*

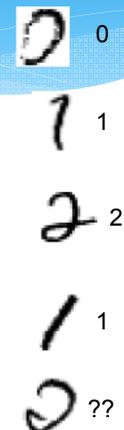
Cours Reconnaissance Automatique de L'écriture Djellali Hayet 9

Exemple 1: Reconnaître un chiffre

- * Entrée: images / pixel grids
- * Sortie: un chiffre de 0-9
- * Realisation:
 - * Il faut créer la base d'images avec la classe de chaque image.
 - * Image de zero avec la classe 0
 - * Image de un avec la classe 1 etc...
 - * On veut apprendre à prédire la classe (le chiffre) qui apparait dans une nouvelle image contenant un chiffre

- * Features: ?

0	0	0	0	0	0
1	1	1	1	1	1
2	2	2	2	2	2
3	3	3	3	3	3
4	4	4	4	4	4
5	5	5	5	5	5
6	6	6	6	6	6
7	7	7	7	7	7
8	8	8	8	8	8
9	9	9	9	9	9



Cours Reconnaissance Automatique de L'écriture Djellali Hayet

Phases de la Reconnaissance de l'écriture

- * La création un fichier des caractères écrits sous format images.
- * Etape de Prétraitement est appliqué pour trouver une image binaire normalisée du caractère écrit qui facilitera la tâche d'extraction de caractéristiques.
- * Extraction de caractéristiques est le processus d'extraire des vecteurs de caractéristiques à partir des données pretraitées assurant une bonne classification.

Exemples 2

Scripteur 1

- * Soit à reconnaître cette suite de chiffre?
- * Exemple de 2 scripteurs

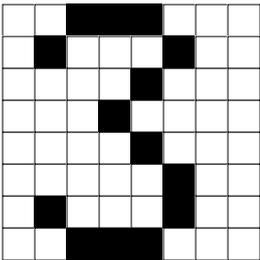
0	0	0	0	0	0
1	1	1	1	1	1
2	2	2	2	2	2
3	3	3	3	3	3
4	4	4	4	4	4
5	5	5	5	5	5
6	6	6	6	6	6
7	7	7	7	7	7
8	8	8	8	8	8
9	9	9	9	9	9

Scripteur 1

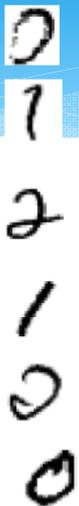
0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6
7	7	7	7	7	7	7	7
8	8	8	8	8	8	8	8
9	9	9	9	9	9	9	9

Reconnaisseur automatique de chiffres

* Entrée: pixel grids



* Sortie: a digit 0-9



Cours Reconnaissance Automatique de L'écriture Djellali Hayet 13

Etapes de la Reconnaissance de l'écriture

- * Prétraitement et Représentation (squelétisation, contours,..)
- * **Segmentation**
 - * En Mot,
 - * En lettres (unités)
- * Reconnaissance de caractères

Cours Reconnaissance Automatique de L'écriture Djellali Hayet 14

Prétraitement

- * Il vise à normaliser l'entrée et réduire le bruit sinon les performances de la reconnaissance sera faible.
- * Techniques utilisées en pretraitement sont:
 - * **Binarization**
 - * **Opérateurs Morphologiques**
 - * **Normalization de L'IMAGE**
 - * Filtrage,
 - * Skew detection(detection d'asymetrie).
 - * Suppression de bruit , slant correction, etc.
- * Si le bruit n'est pas correctement supprimé il conduira à une segmentation erronée et donc un taux de reconnaissance faible.

Binarisation

- * Le document image est enregistré en general en RGB (red, green, et blue) convertit en niveau de gris en calculant une somme pondérée de l'intensité RGB de la valeur de chaque pixel (Brink, 2011).
- * **La binarisation consiste à** distinguer la région d'intérêt (handwritten text) et le background (paper) sans modifier l'image écrite (Solihin and Leedham, 1999; Plamondon and Srihari, 2000).

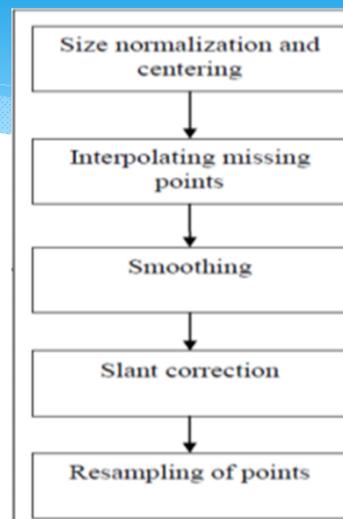
Processus de Rec. Ecriture

- * Prétraitement (Pr)
- * Segmentation (Seg)
- * Extraction de Caractéristiques (EC)
- * Modélisation
- * Décision

Prétraitement

Le prétraitement réduit le bruit dans l'image comportant le texte écrit.

- * Les bruits produits possible sont:
- * Taille irrégulière du texte.
- * Les points manquants pendant le mouvement du stylo.
- * Les inclinaisons dans l'écriture des caractères.



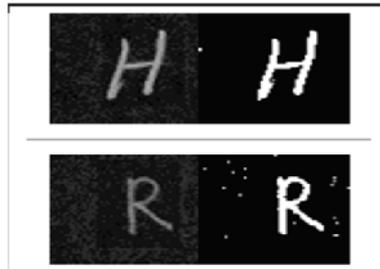
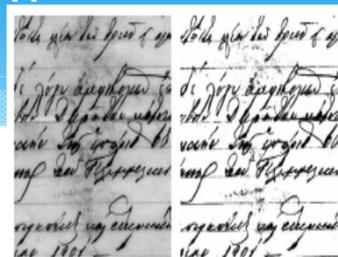
Prétraitement

- Binarization
- Réduction du bruit
- Stroke avec normalization
- correction d'inclinaison (Skew correction)
- Slant removal

Binarization

□ **Document image binarization** (thresholding)
la conversion de l'image en niveau de gris en image binaire. Deux catégories de seuillage sont utilisées.

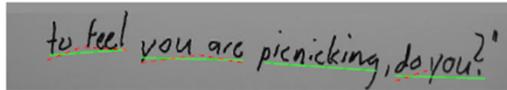
- Global, fixe une seule valeur du seuil pour tout le document image.
- Adaptive (local), plusieurs valeurs sont utilisées selon l'information de la région locale



Prétraitement

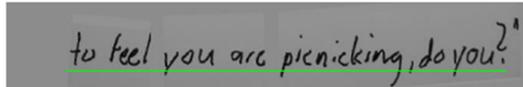
Assumption: Documents are already segmented into text lines
(Text detection and line extraction *highly* application specific!)

Baseline Estimation:



Potential method: ▶ Initial estimate based on horiz. projection histogram
▶ Iterative refinement and outlier removal (cf. [2, 10])

Skew and Displacement Correction:



Prétraitement

Détection d'inclinaison (Skew) et correction:

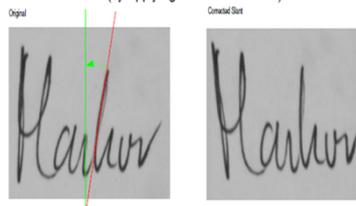
une inclinaison introduite lors du processus de numérisation ou par l'effet du sripteur.

Les lignes inclinées sont réalignés horizontalement car autrement elles peuvent produire une mauvaise segmentation ou classification.

Slant estimation: E.g. via mean orientation of edges obtained by Canny operator (cf. e.g. [12])



Slant normalization (by applying a shear transform)



Prétraitement exemple

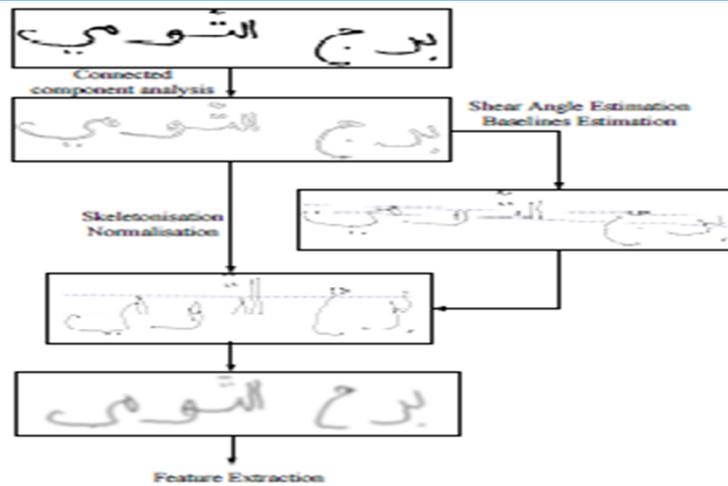


Figure 1. Example of the pre-processing and normalisation steps

Prétraitement



Fig.2. Word Image Pre-Processing (a) Input Scanned Word Images; (b) Word Images after Gray Scale Intensity Threshold; (c) Word Images in Binary Format; (d) Word Images after Thinning; (e) Word Images after Noise Removal.

Prétraitement

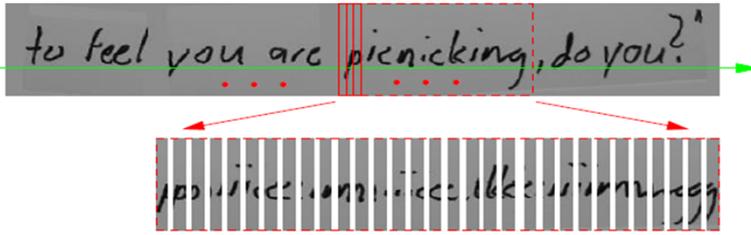
Serialization: The Sliding Window Method

Problem: Data is two-dimensional, images of writing!
 ⚡ No chronological structure inherently defined!

Exception: Logical sequence of characters within texts

Solution: Sliding-window approach
 First proposed by researchers at Daimler-Benz Research Center, Ulm [3],
 pioneered by researchers at BBN [11]

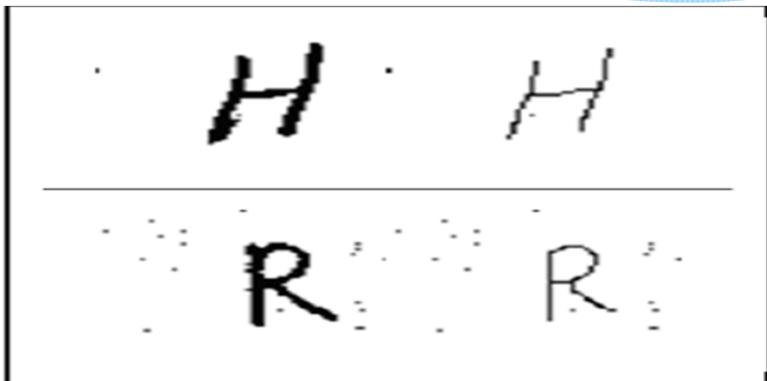
- ▶ Time axis runs in writing direction / along baseline
- ▶ Extract small overlapping analysis windows



Cours Reconnaissance Automatique de L'écriture Djellali Hayet 25

Squeléttisation (Thining) après seuillage

La squelettisation permet de réduire le traitement et ne garder que le squelette général du caractère.



Cours Reconnaissance Automatique de L'écriture Djellali Hayet 26

Exemple Ecriture manuscrite image

called	rabbit	red	black	blue
buffalo	good	age	each	easy
image	robot	word	easy	idea
design	been	bat	top	median
degraded	normalization	car	bad	vision
father	method	all	does	color

Cours Reconnaissance Automatique de L'écriture Djellali Hayet

27

Segmentation

"Traditional" Recognition Paradigm



La Segmentation de mots permet d'identifier les différents mots à reconnaître. Puis, chaque mot devra être segmenté en caractère pour pouvoir l'envoyer au module reconnaissance de caractères et ainsi produire en sortie la liste des caractères reconnus en format texte.

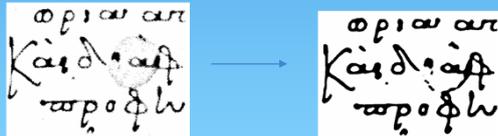
Cours Reconnaissance Automatique de L'écriture Djellali Hayet

28

Noise Reduction - Normalization

□ La reduction du bruit ameliore la qualité du document image. Deux approches principales:

- Filtres(masks)
- opérations Morphologiques (erosion, dilation, etc)



□ La Normalization produit une reduction de la taille des caractères, squeletisation (thinning) extrait l'information de forme des caracteres.



Cours Reconnaissance Automatique de L'écriture

Djellali Hayet

29

Skew Correction

□ Skew Correction methods sont utilisés pour aligner le document. Les principales approches pour la detection de skew inclut la correlation, projection, Transformé de Hough

This is a document image that will present you the common problems of OCR

The non-parallel text line is a very usually met problem making difficult the skew angle estimation

The hill and dale writing is also met as well as the slanted and connected characters.

This is a document image that will present you the common problems of OCR

The non-parallel text line is a very usually met problem making difficult the skew angle estimation

The hill and dale writing is also met as well as the slanted and connected characters.

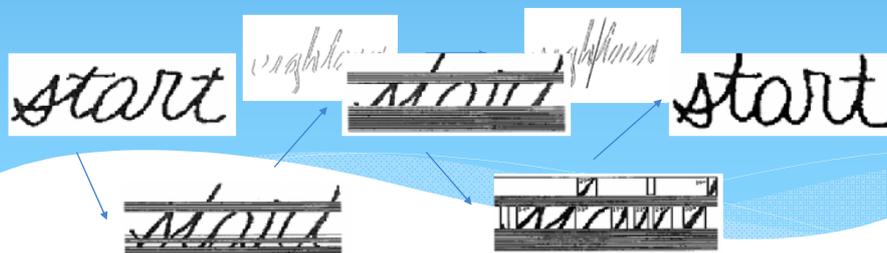
Cours Reconnaissance Automatique de L'écriture

Djellali Hayet

30

Suppression de Slant (Slant Removal)

- The slant de texte manuscrits varient d'un utilisateur à une autre . Slant removal est utilisée pour normaliser tous les caractères à une forme standard. Popular deslanting techniques are:
 - Calculation of the average angle of near-vertical elements
 - Bozinovic – Shrihari Method (BSM).



Cours Reconnaissance Automatique de L'écriture Djellali Hayet

31

Slant Removal

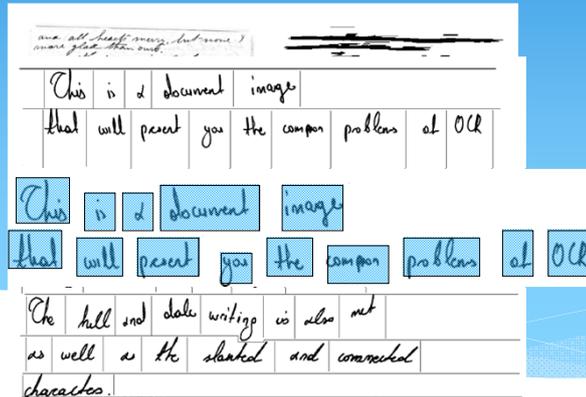
- La pente dominante du caractère est trouvée à partir de la pente du caractère qui fournit l'entropie minimale de la projection verticale d'histogramme.

Cours Reconnaissance Automatique de L'écriture Djellali Hayet

32

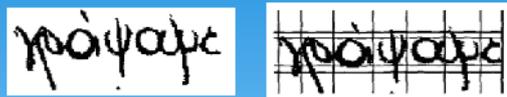
- ☐ Détection de la ligne de Texte(Text Line Detection) utilisant (Hough Transform, projections, smearing)

Extraction de Mots (projections verticales), analyse des composantes connectées (connected component analysis)
Extraction de mots 2 (RLSA)



Cours Reconnaissance Automatique de L'écriture Djellali Hayet 33

Segmentation Mots- Caractères



Segmentation Implicite

L'approche implicite : Les mots sont reconnus entièrement sans segmentation en caractères. Cette approche est effective seulement quand un ensemble de mots possibles et connus à l'avance, tels que la reconnaissance des chèques postaux et les adresses postal

Segmentation Explicite

In explicit approaches one tries to identify the smallest possible word segments (primitive segments) that may be smaller than letters, but surely cannot be segmented further.

Later in the recognition process these primitive segments are assembled into letters based on input from the character recognizer. The advantage of the first strategy is that it is robust and quite straightforward, but is not very flexible.

Cours Reconnaissance Automatique de L'écriture Djellali Hayet 34

Extraction de Caractéristiques

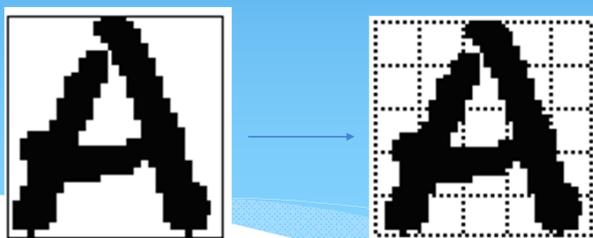
- In feature extraction stage each character is represented as a feature vector, which becomes its identity. The major goal of feature extraction is to extract a set of features, which maximizes the recognition rate with the least amount of elements.
- Due to the nature of handwriting with its high degree of variability and imprecision obtaining these features, is a difficult task. Feature extraction methods are based on 3 types of features:
 - Statistical
 - Structural
 - Global transformations and moments

Caractéristiques Statistiques (Statistical Features)

- Representation of a character image by statistical distribution of points takes care of style variations to some extent.
- The major statistical features used for character representation are:
 - Zoning
 - Projections and profiles
 - Crossings and distances

Zoning

- The character image is divided into $N \times M$ zones. From each zone features are extracted to form the feature vector. The goal of zoning is to obtain the local characteristics instead of global characteristics

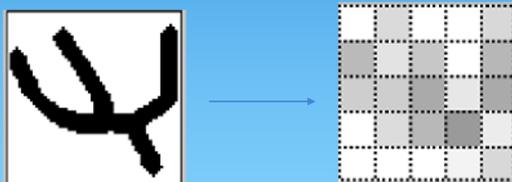


Cours Reconnaissance Automatique de L'écriture Djellali Hayet

37

Zoning – Density Features

- The number of foreground pixels, or the normalized number of foreground pixels, in each cell is considered a feature.



Darker squares indicate higher density of zone pixels.

Cours Reconnaissance Automatique de L'écriture Djellali Hayet

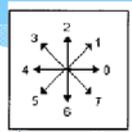
38

Zoning – Direction Features

- Based on the contour of the character image

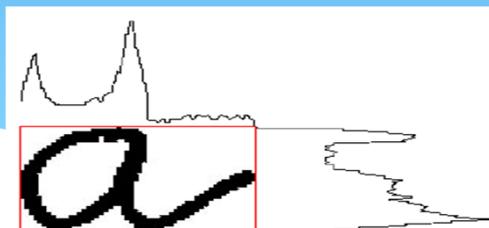


- For each zone the contour is followed and a directional histogram is obtained by analyzing the adjacent pixels in a 3x3 neighborhood



Projection d'Histogramme

- The basic idea behind using projections is that character images, which are 2-D signals, can be represented as 1-D signal. These features, although independent to noise and deformation, depend on rotation.
- Projection histograms count the number of pixels in each column and row of a character image. Projection histograms can separate characters such as “m” and “n” .



Profiles

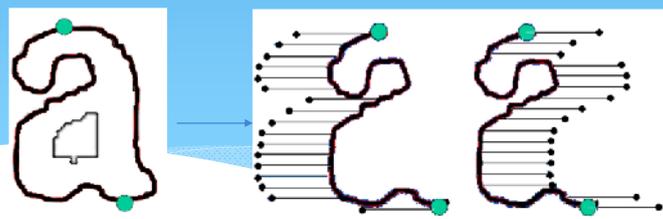
□ Le profile compte le nombre de pixels (distance) entre le rectangle entourant l'image du caractère (the bounding box) et son contour (edge) . Le profile décrit bien les formes extérieures du caractère et permet de distinguer entre un grand nombre de lettres, telsque "p" and "q".



Cours Reconnaissance Automatique de L'écriture Djellali Hayet 41

Profiles

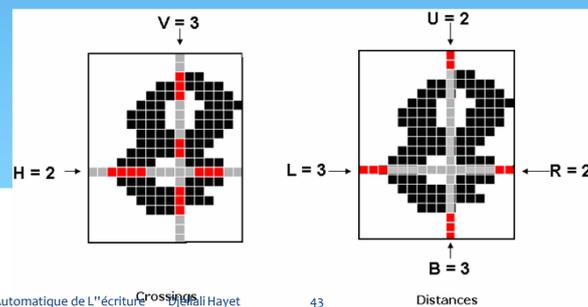
- Profiles can also be used to the contour of the character image
- Extract the contour of the character
- Locate the uppermost and the lowermost points of the contour
- Calculate the in and out profiles of the contour



Cours Reconnaissance Automatique de L'écriture Djellali Hayet 42

Crossings and Distances

□ **Crossings** count the number of transitions from background to foreground pixels along vertical and horizontal lines through the character image and **Distances** calculate the distances of the first image pixel detected from the upper and lower boundaries, of the image, along vertical lines and from the left and right boundaries along horizontal lines



Cours Reconnaissance Automatique de L'écriture Djellali Hayet

43

Distances

Caractéristiques Structurelles (Structural Features)

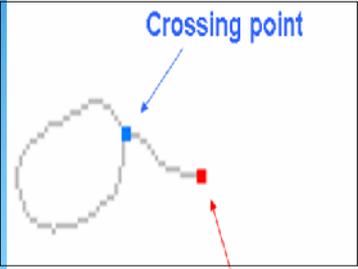
Characters can be represented by structural features with high tolerance to distortions and style variations. This type of representation may also encode some knowledge about the structure of the object or may provide some knowledge as to what sort of components make up that object.

Structural features are based on topological and geometrical properties of the character, such as aspect ratio, cross points, loops, branch points, strokes and their directions, inflection between two points, horizontal curves at top or bottom, etc.

Cours Reconnaissance Automatique de L'écriture Djellali Hayet

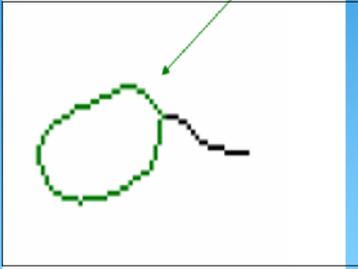
44

Structural Features



Crossing point

End point



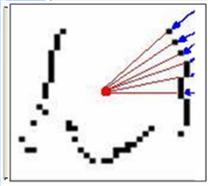
Loop

Cours Reconnaissance Automatique de L'écriture Djellali Hayet 45

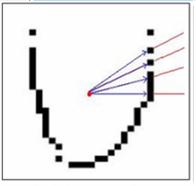
Structural Features

- Horizontal and Vertical projection histograms
- Radial histogram
- Radial out-in and radial in-out profiles


→



Radial out-in profile
radial histogram

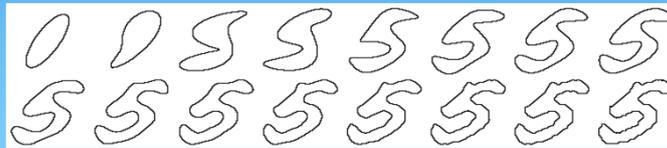


Radial in-out profile

Cours Reconnaissance Automatique de L'écriture Djellali Hayet 46

Transformations Globales- Moments

□ **The Fourier Transform (FT)** of the contour of the image is calculated. Since the first n coefficients of the FT can be used in order to reconstruct the contour, then these n coefficients are considered to be a n -dimensional feature vector that represents the character.

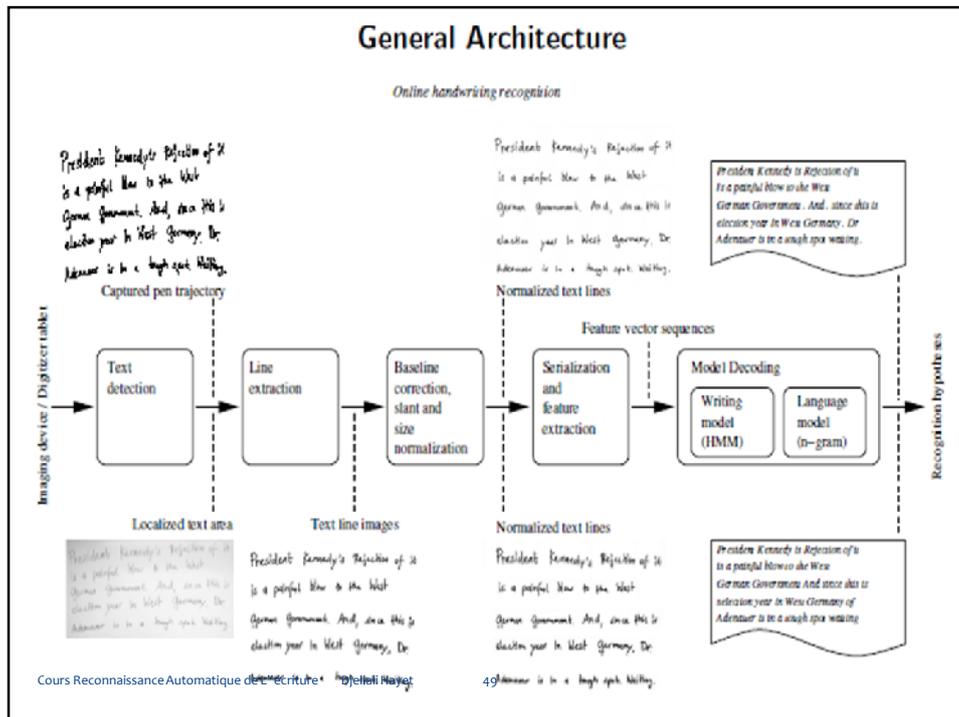


□ **Central, Zernike moments** that make the process of recognizing an object scale, translation, and rotation invariant. The original image can be completely reconstructed from the moment coefficients.

Extraction de Caractéristiques

Plusieurs paramètres à extraire tels que:

- Les paramètres géométriques
- Straight-line detection,
- Moments geometric, Moments de Zernike
- Unitary image transforms,
- Graph descriptors,
- Projection histograms,
- Contour profiles,
- Zoning,
- Stroke detection,
- Approximations de courbes spline,
- Descriptors de fourier,



Modélisation- Décision

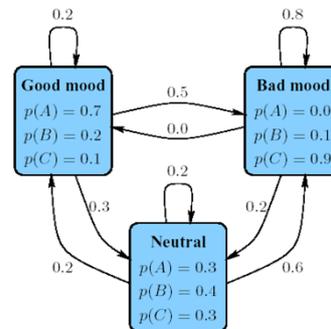
- * Les Modèles de Markov caches
- * Les Réseaux de Neurones
- * Les Méthodes hybrides
- * etc

Modèle de Markov Cachés

Parametres du Modele:

S - states {good, neutral, bad}
 a_{kl} - probability that state l is followed by k
 Σ - alphabet {A,B,C}
 $e_k(x)$ - probability that state k emit symbol $x \in \Sigma$

Comment modéliser l'écriture par les modeles de Markov cachés
 HMM
 (prochains cours)



Classification

k-Nearest Neighbour (k-NN) , Bayes Classifier, Neural Networks (NN), Hidden Markov Models (HMM), Support Vector Machines (SVM), etc

There is no such thing as the “best classifier”. The use of classifier depends on many factors, such as available training set, number of free parameters etc.

Mesures de Performances

- * Taux de bonne reconnaissance est le nombre de caractères reconnues correctement sur le total des caractères à reconnaître.
- * La précision et le rappel sont calculés(voir 1^{er} chapitre). Fmeasure

Base de Données

- * La base Mnist utilisé par lecun comporte 60000 chiffres écrits pour l'apprentissage
- * 20000 chiffres pour le test.

Pretraitement

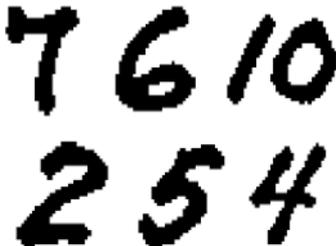



Figure 4: Digits from the hydrographic map in the binary raster representation.

Figure 5: Skeletons of the digits in Fig. 4, thinned with the method of Zhang and Suen [28]. Note that junctions are displaced and a few short false branches occur.

Cours Reconnaissance Automatique de L'écriture Djellali Hayet 55

Prétraitement



Figure 6: Contours of two of the digits in Fig. 4.

Cours Reconnaissance Automatique de L'écriture Djellali Hayet 56

Pretraitement

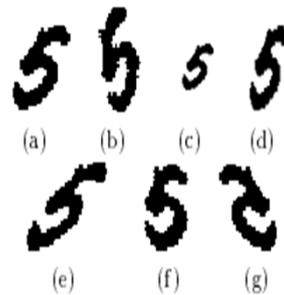


Figure 7: Transformed versions of digit '5'. (a) original, (b) rotated, (c) scaled, (d) stretched, (e) skewed, (f) de-skewed, (g) mirrored.

Feature Extraction

2.4 Zoning

The commercial OCR system by CALERA described in Bokser [40] uses zoning on solid binary characters. A straightforward generalization of this method to gray level character images is given here. An $n \times m$ grid is superimposed on the character image (Fig. 8(a)), and for each of the $n \times m$ zones, the average gray level is computed (Fig. 8(b)), giving a feature vector of length $n \times m$. However, these features are not illumination invariant.

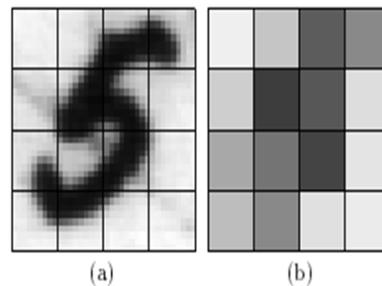
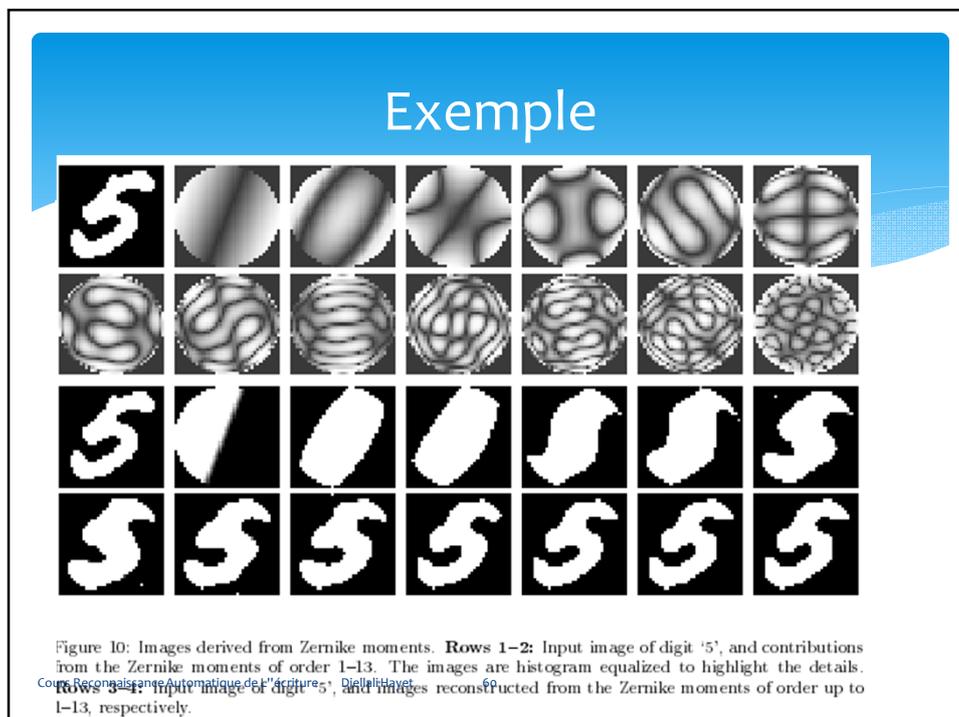
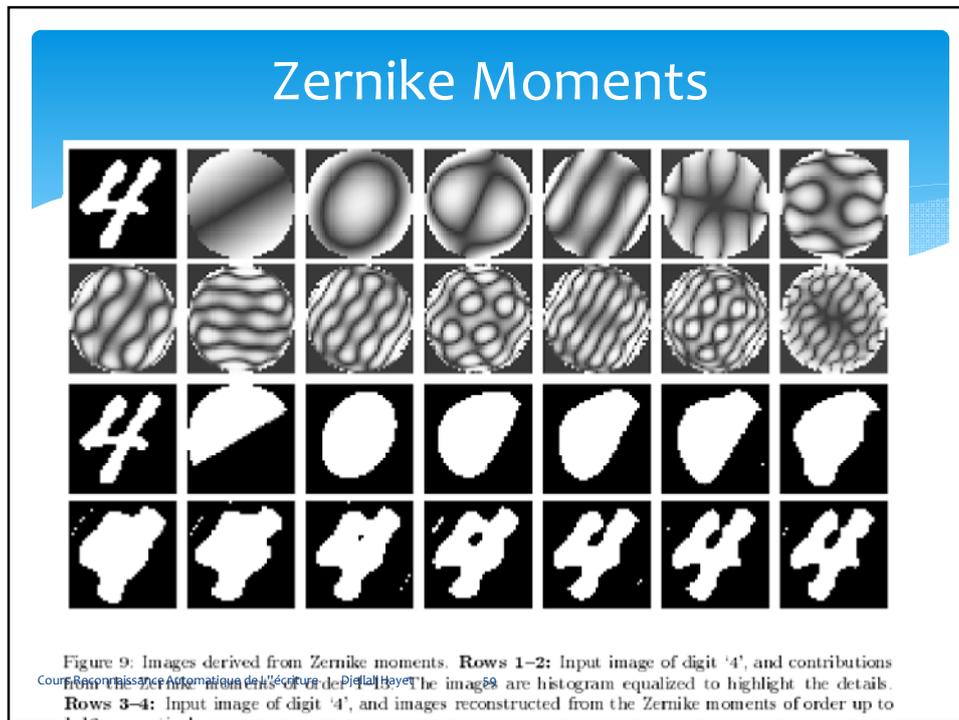


Figure 8: Zoning of gray level character images. (a) A 4×4 grid superimposed on a character image. (b) The average gray levels in each zone, which are used as features.



Projection d'histogramme

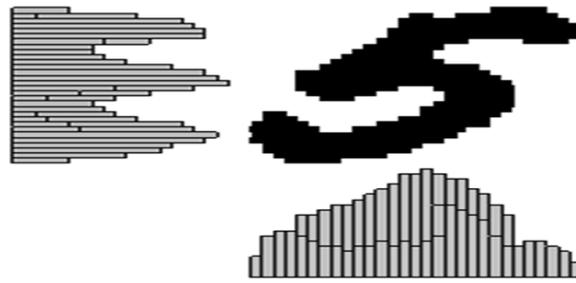


Figure 1. Horizontal and Vertical projection histogram

Exemples de Réseaux de Neurones de LeCun

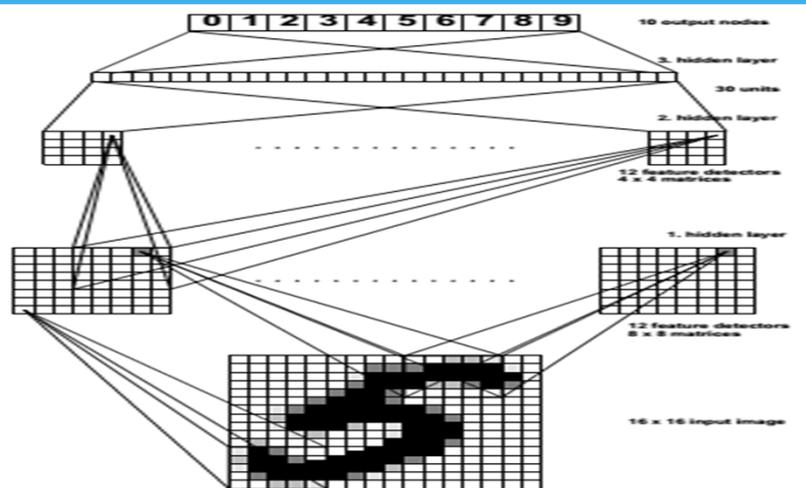


Figure 23. The neural network classifier used by Le Cun et al. [86].

détails

- * Entrées : image du chiffre
- * Les couches cachés
- * Sorties: 10 sorties (10 chiffres)

- * Les images sont extraites d'articles scientifiques
(référence rajouter ultérieurement)