


Optical Handwritten Character/Numeral Recognition

9th North-East Workshop
on
Computational Information Processing

Pulak Purkait, JRF
ECSU, Indian Statistical Institute

Handwriting Categories

- Off-line character recognition


KOLKATA TRAFFIC POLICE

Book No: 04/09 Sl. No: [] Date: 15/04/09
 Beat No: 18 TP Guard: TB GPF No: CALIB39972

VEHICLE - NUMBER							VECH TYPE	TIME (RLY)	OFF CODE	DIR								
W	B	0	2	R	9	8	5	8	1	2	1	1	4	0	0	1	0	3
W	B	1	1	A	8	4	0	2	0	9	1	1	4	5	0	1	0	3
W	B	0	3	B	9	4	2	5	0	9	1	2	5	5	1	1	0	2
W	B	0	6	7	2	5	3	1	2	0	0	0	0	0	1	0	3	
W	B	2	8	5	2	0	0	5	0	3	1	3	0	5	0	1	0	3
W	B	T			5	3	5	2	0	6	1	3	0	2	0	1	0	3
W	B	1	9	C	3	5	2	4	0	6	1	3	2	0	0	1	0	3
W	B	Y			2	7	0	4	0	9	1	3	2	5	0	1	0	3

Rank & Name: of Tapsh Kr. Paul No. 2591 Signature: T. Paul

OCR Systems

- OCR systems consist of following major stages :
 - Pre-processing
 - Segmentation
 - Feature Extraction
 - Classification
 - Post-processing

Pre-processing

□ The raw data is subjected to a number of preliminary processing steps to make it usable in the descriptive stages of character analysis. Pre-processing aims to produce data that are easy for the OCR systems to operate accurately. The main objectives of pre-processing are :

- Binarization
- Noise reduction
- Skew correction
- Slant removal

Binarization

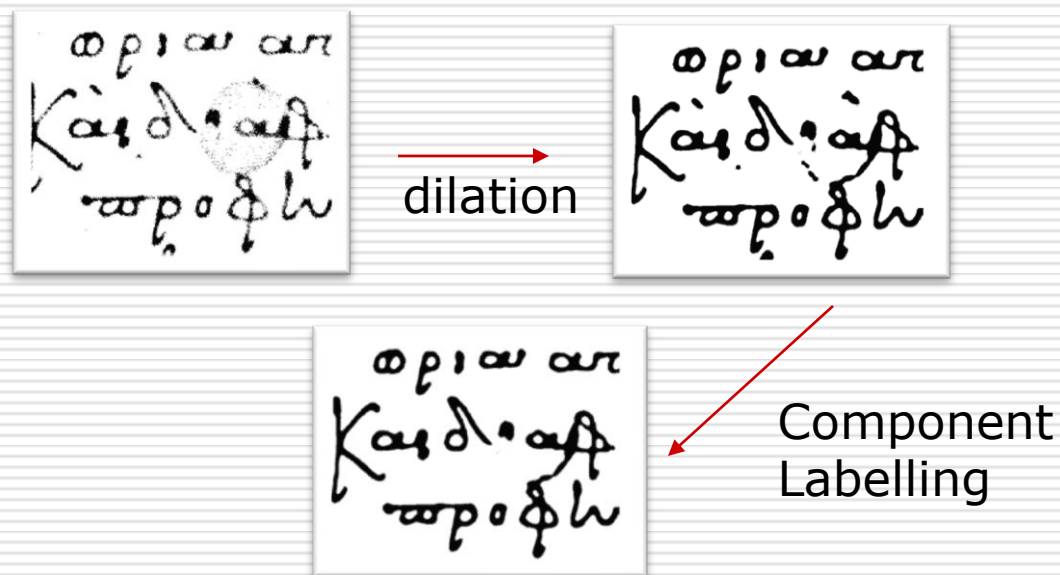


- Document image binarization (thresholding) refers to the conversion of a gray-scale image into a binary image. Two categories of thresholding:
 - Global, picks one threshold value for the entire document image which is often based on an estimation of the background level from the intensity histogram of the image. Sometimes global threshold parameter chosen using Otsu's method.
 - Adaptive (local), uses different values for each pixel according to the local area information



















































Noise Reduction

□ Noise reduction improves the quality of the document. Three main approaches:

- Filtering (masks, Gaussian filter, etc)
- Morphological Operations (erosion, dilation, etc)
- Component Labelling



Example.(Preprocessing on ISI Numeral Database)

	original	binarize	resize	skeleton	dilation
Zero					
one					
two					
three					
four					
five					
six					
seven					
eight					
nine					

(a) (b) (c) (d) (e)

Skew Correction

- Skew Correction methods are used to align the paper document with the coordinate system of the scanner. Main approaches for skew detection include correlation, projection profiles, Hough transform.

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

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

The hill and dale writing is also not
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 usual problem
making difficult the skew angle estimation

The hill and dale writing is also not
as well as the slanted and connected
characters.

Slant Removal

- ❑ The slant of handwritten texts varies from user to user. Slant removal methods are used to normalize the all characters to a standard form.
- ❑ Popular deslanting techniques are:



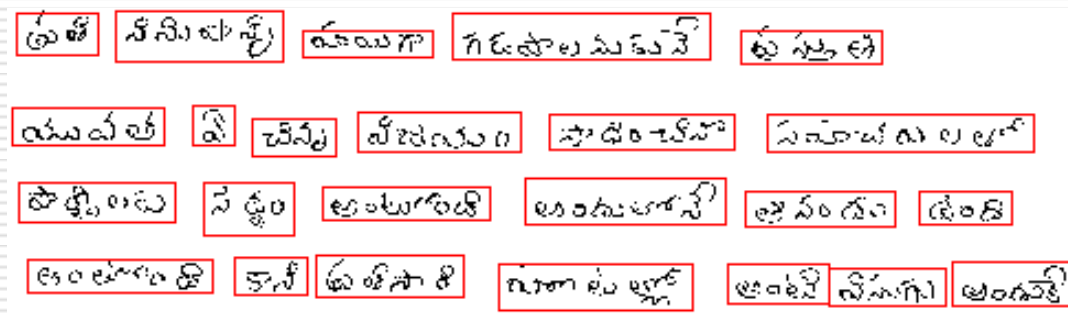
Slant Removal

- ❑ The slant of handwritten texts varies from user to user. Slant removal methods are used to normalize the all characters to a standard form.
- ❑ Popular deslanting techniques are:
 - Bozinovic – Shrihari Method (BSM)
 - Calculation of the average angle of near-vertical elements



Segmentation

- ❑ Text Line Detection (Hough Transform, projections, smearing)
- ❑ Word Extraction (vertical projections, connected component analysis)



Word Segmentation on Telugu Database

Segmentation

□ Explicit Segmentation

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. .



Segmentation

Explicit Segmentation

Implicit Segmentation

In implicit approaches the words are recognized entirely without segmenting them into letters. This is most effective and viable only when the set of possible words is small and known in advance, such as the recognition of bank checks and postal address

Feature Extraction

- 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

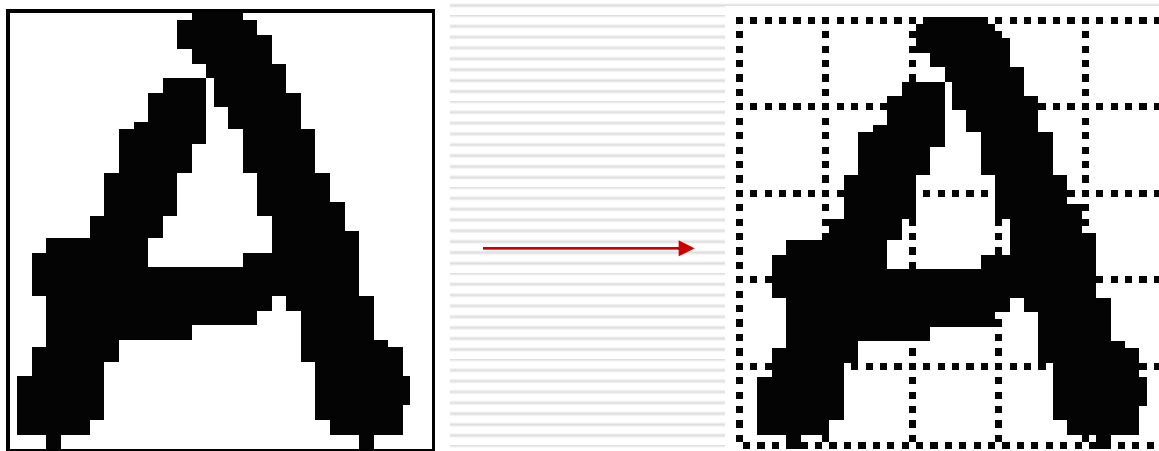
Statistical Features

- 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

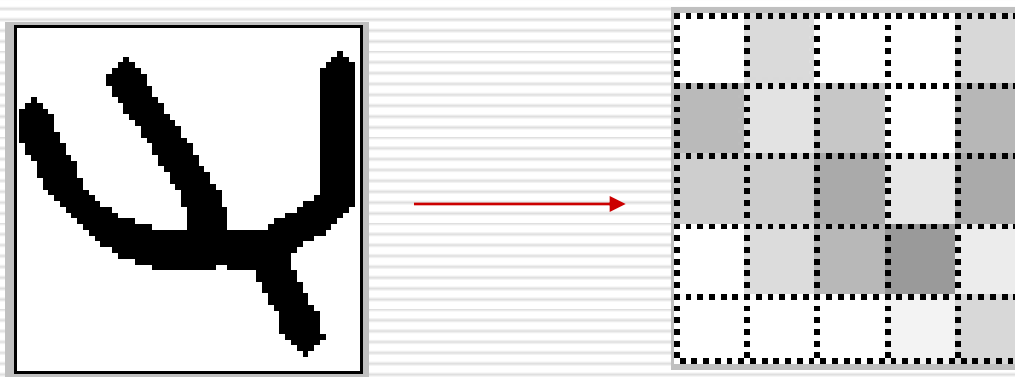
Zoning

- 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



Zoning – Density Features

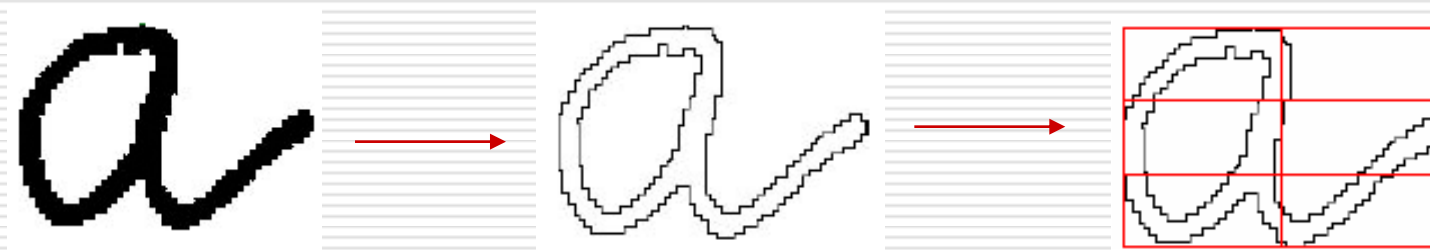
- The number of foreground pixels, or the normalized number of foreground pixels, in each cell is considered a feature.



Darker squares indicate higher density of zone pixels.

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

Zoning – Direction Features

- Based on the skeleton of the character image



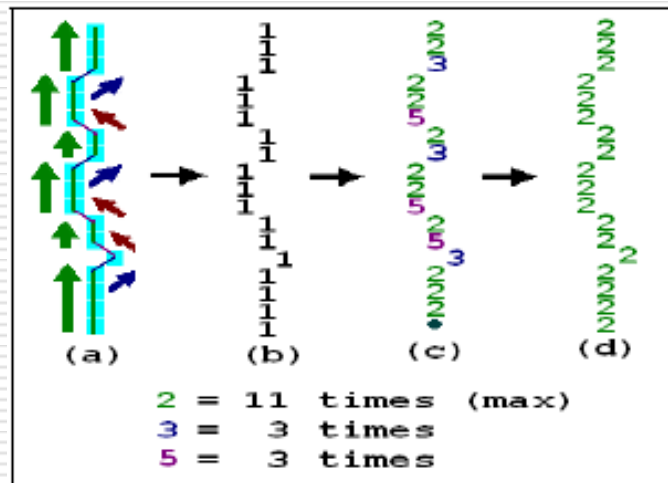
- Distinguish individual line segments
- Labeling line segment information
 - Line segments are coded with a direction number
 - 2 = vertical line segment
 - 3 = right diagonal line segment
 - 4 = horizontal line segment
 - 5 = left diagonal line segment

Zoning – Direction Features

- Based on the skeleton of the character image



- Distinguish individual line segments
- Labeling line segment information
- Line type normalization

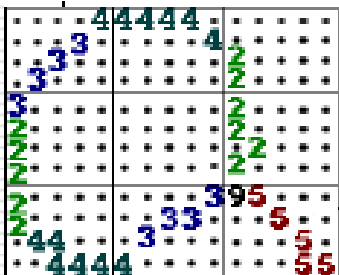


Zoning – Direction Features

- Based on the skeleton of the character image



- Distinguish individual line segments
- Labeling line segment information
- Line type normalization
- Formation of feature vector through zoning



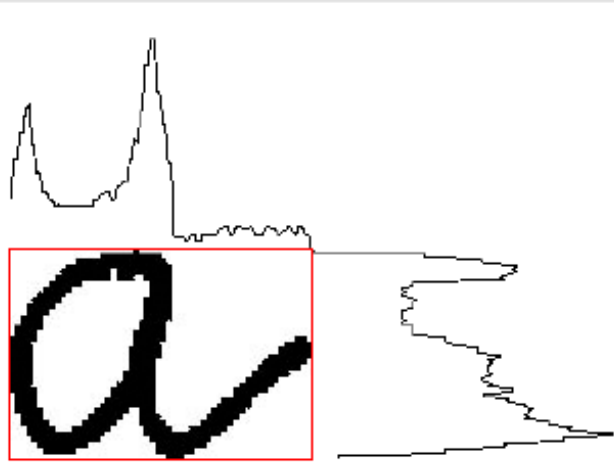
number of horizontal lines	total length of horizontal lines	number of right diagonal lines	total length of right diagonal lines	number of vertical lines	total length of vertical lines	number of left diagonal lines	total length of left diagonal lines	number of intersection points
----------------------------	----------------------------------	--------------------------------	--------------------------------------	--------------------------	--------------------------------	-------------------------------	-------------------------------------	-------------------------------

$$\text{value} = 1 - ((\text{number of lines}/10) \times 2)$$

$$\text{length} = \frac{\text{number of pixels in a particular direction}}{(\text{window height or width}) \times 2}$$

Projection Histograms

- 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" .



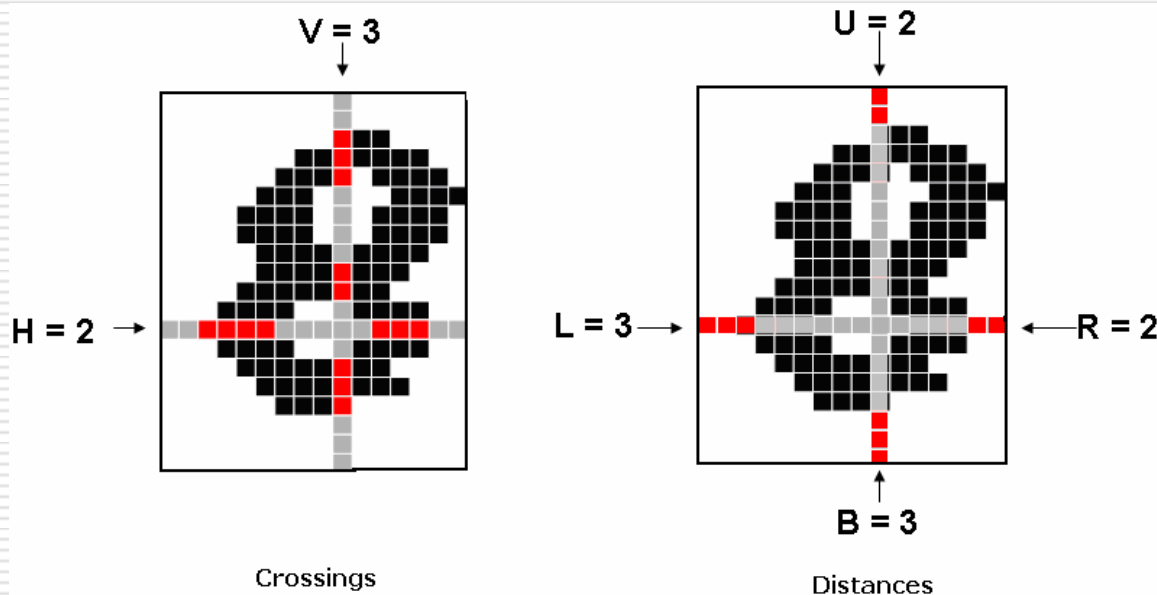
Profiles

- The profile counts the number of pixels (distance) between the bounding box of the character image and the edge of the character. The profiles describe well the external shapes of characters and allow to distinguish between a great number of letters, such as “p” and “q”.



Crossings and Distances

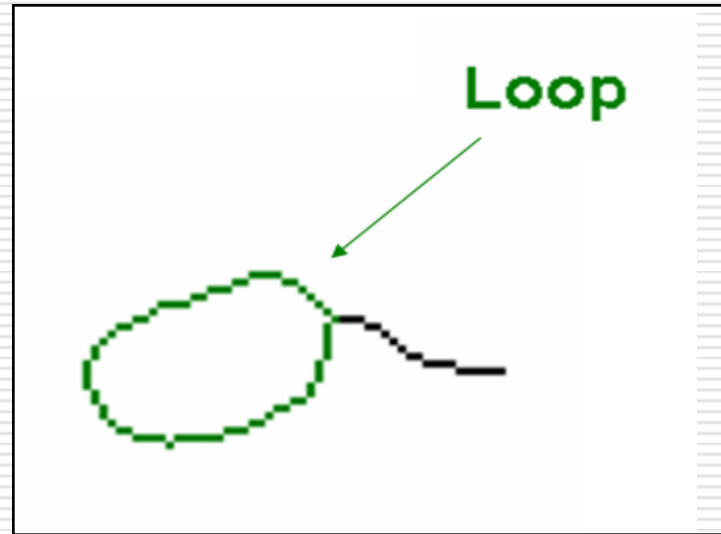
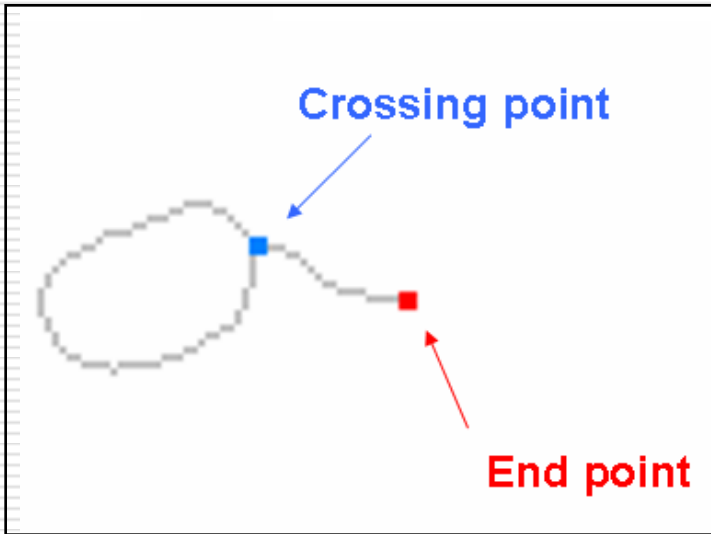
□ **Crossings** count the number of transitions from background to foreground pixels along vertical and horizontal lines through the character image and **Distances** calculate the distances of the first image pixel detected from the upper and lower boundaries, of the image, along vertical lines and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



Structural Features

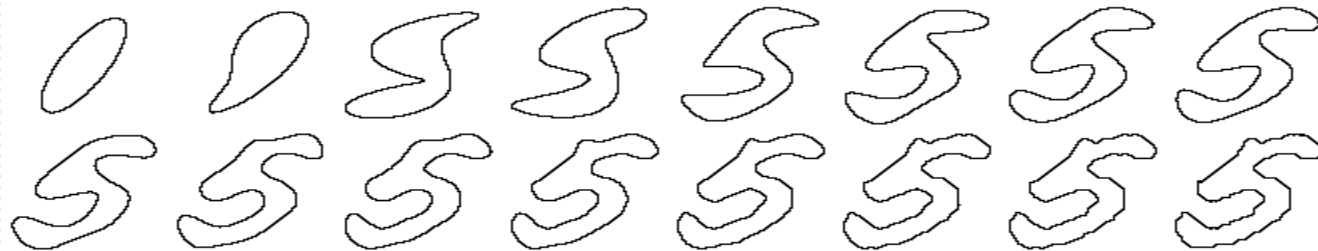
- Characters can be represented by structural features with high tolerance to distortions and style variations. This type of representation may also encode some knowledge about the structure of the object or may provide some knowledge as to what sort of components make up that object.
- Structural features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

Structural Features



Global Transformations - 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.



Global Transformations - 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, Zenrike 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.

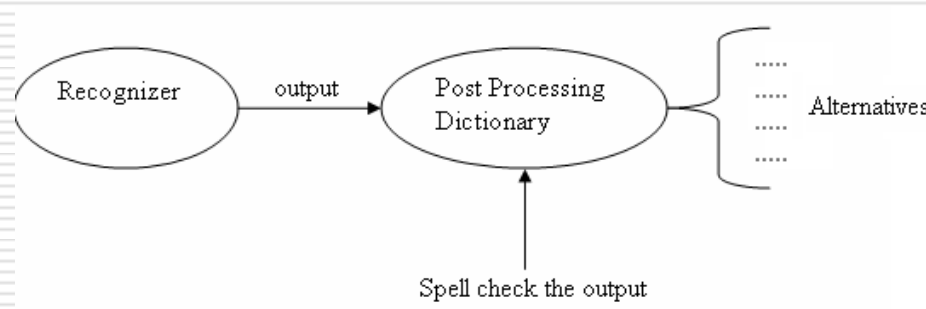
Classification

□ 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.

Post-processing

- Goal : the incorporation of context and shape information in all the stages of OCR systems is necessary for meaningful improvements in recognition rates.
- The simplest way of incorporating the context information is the utilization of a dictionary for correcting the minor mistakes.



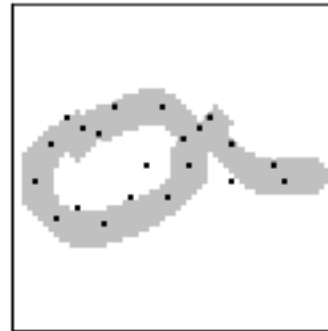
- Drawback : Unrecoverable OCR decisions.

Post-processing

- Goal : the incorporation of context and shape information in all the stages of OCR systems is necessary for meaningful improvements in recognition rates.
- In addition to the use of a dictionary, a well-developed lexicon and a set of orthographic rules (lexicon-driven matching approaches) during or after the recognition stage for verification and improvement purpose.

Future Work

- ❑ Creating new hierarchical classification schemes based on rules after examining the corresponding confusion matrix.
- ❑ Exploiting new features to improve the current performance.





THANK YOU