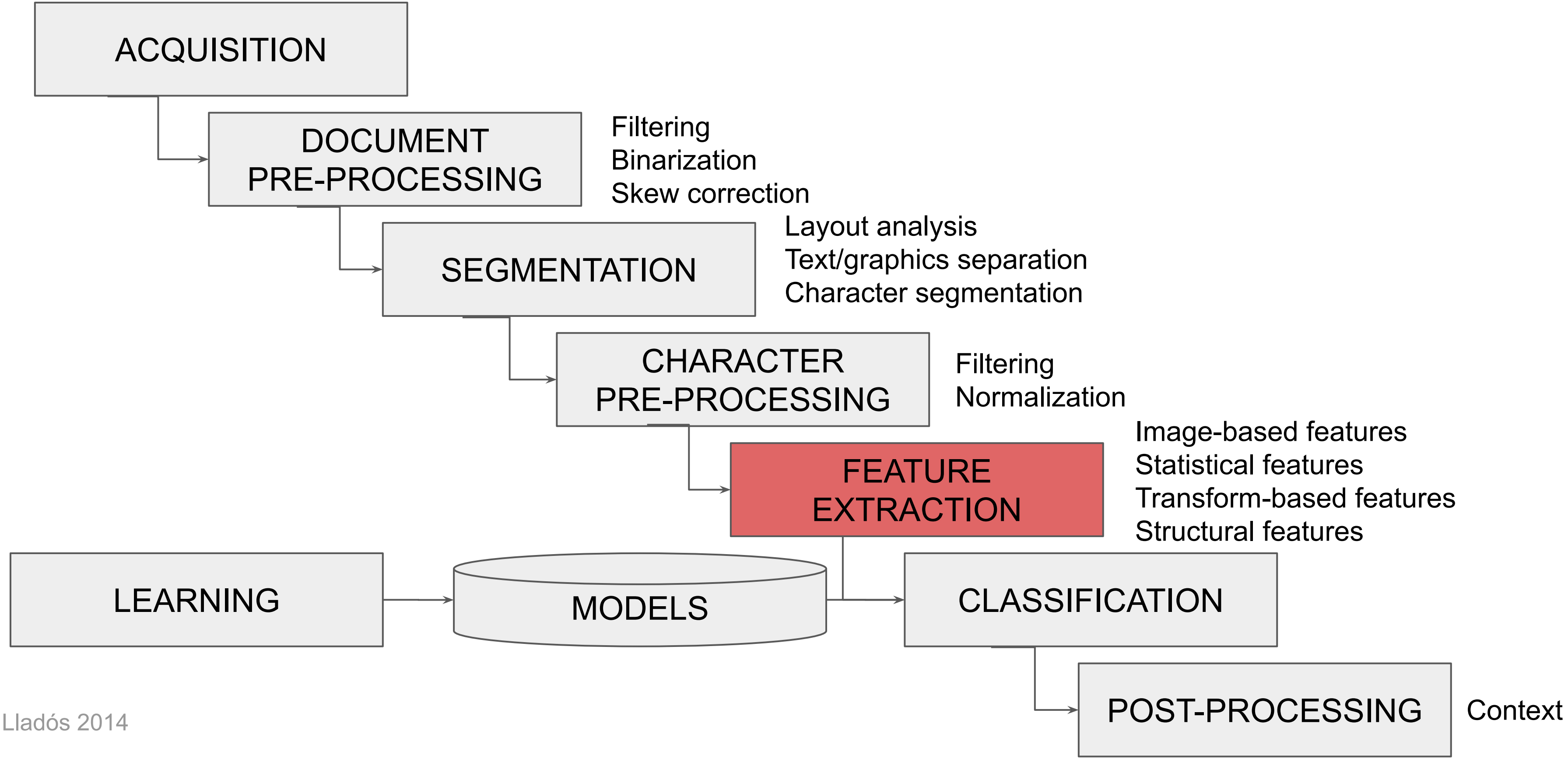
MLRF Lecture 02

J. Chazalon, LRDE/EPITA, 2019

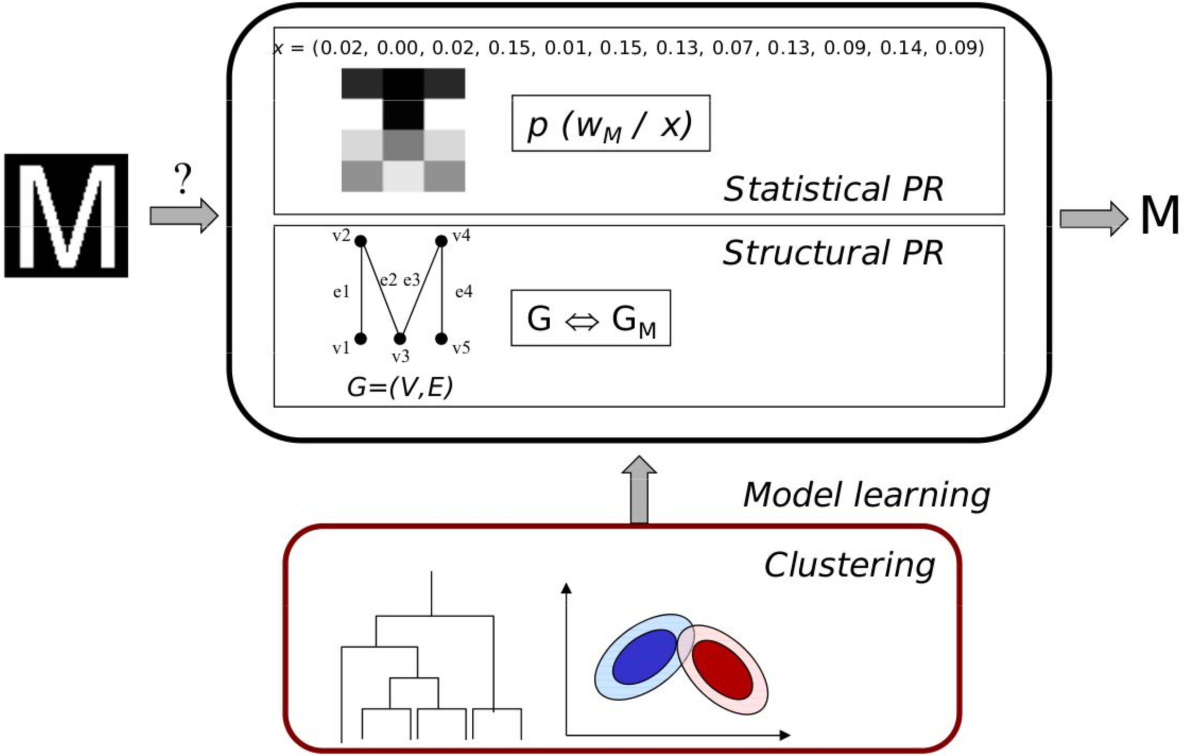
Character descriptors

Lecture 02 part 05

Components of an OCR system



Pattern Recognition: Statistical and Structural



OCR: Feature Extraction

Image-based features:

Projection, Profiles, Crossings

Statistical features:

Moments, Zoning, Histograms

Global transforms and series expansion:

Fourier descriptors...

Structural analysis:

Contour analysis, Skeleton analysis, Topological and geometric features

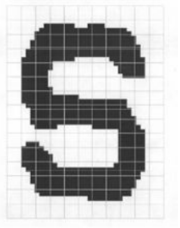
O.D. Trier, A.K. Jain, T. Taxt. Feature Extraction Methods for Character Recognition - A Survey. Pattern Recognition, Vol.29, No. 4, pp. 641-662, 1996.

Image-based features

Image-based features

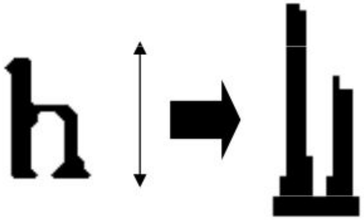
All the image as feature vector

- Classification by correlation
- Very sensitive to noise, character distortion and similarity between classes.



x and/or y projections

- We can use the accumulated projection too
- Sensitive to rotation, distortion and large number of characters



Peephole

- Coding with a binary number some pre-selected pixels of the image
- Pre-selected pixels can vary depending on the character to be recognized

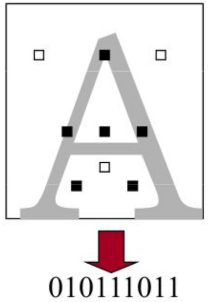


Image-based features

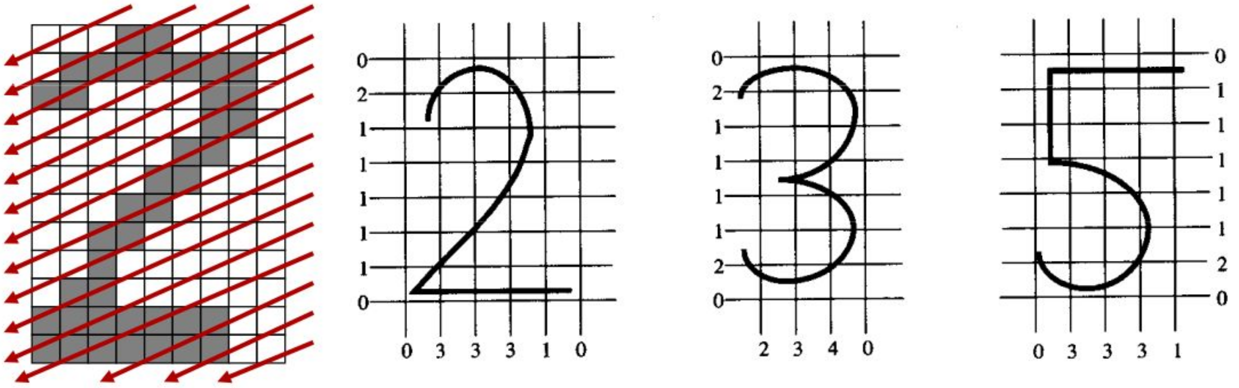
Crossing method

Computed from the number of times a character is crossed along some orientations, for example 0°, 45°, 90°, 135°

Used on legacy commercial system because of speed and low complexity

Robust to some distortions and noise

Sensitive to size variations

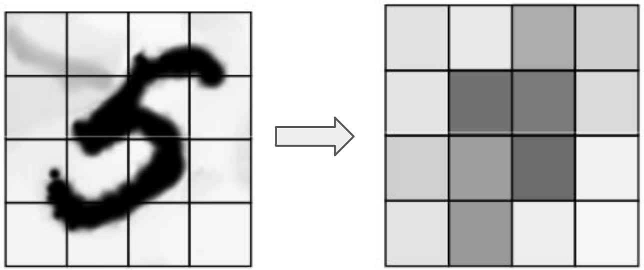


Statistical features

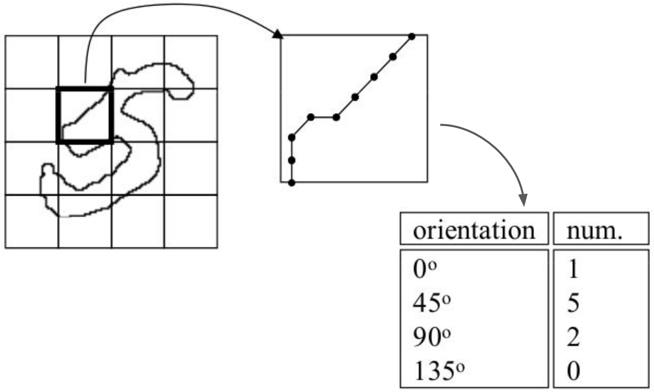
Statistical Features: Zoning

The image is divided in n x m cells

For each cell the mean of gray levels is computed and all these values are joined in a feature vector of length n x m



We can also use information from the contour or any other feature computed in every zone



F. Kimura, M. Shridhar: Handwritten numeral recognition based on multiple algorithms. Pattern Recognition, 24(10), pp. 969-983, 1991

Statistical Features: Geometric Moments

Moments of order (p+q) of image f:

$$m_{pq} = \sum_{x=1}^{N}\sum_{y=1}^{N} f(x,y)(x)^p(y)^q$$

m_{00} = character area (in binary images)

Center of gravity of the character:

$$\bar{x} = \frac{m_{10}}{m_{00}} \quad \bar{y} = \frac{m_{01}}{m_{00}}$$

Central moments (centering the character at the center of gravity):

$$\mu_{pq} = \sum_{x=1}^{N}\sum_{y=1}^{N} f(x,y)(x-\bar{x})^p(y-\bar{y})^q$$

Central moments of order 2 (20 , 02 , 11) allow to compute:

- Main inertia axes
- Character length
- Character orientation

$$\theta = \frac{1}{2} atan\left(\frac{2\mu_{11}}{\mu_{20}-\mu_{02}}\right)$$

M. Hu: Visual pattern recognition by moment invariants. IEEE Trans. Inf. Theory 8, pp. 179-187, 1962

Statistical Features: Geometric Moments

Invariant moments (based on central moments):

- Central moments μ_{pq} are translation-invariant
- Scale-invariants →

$$\nu_{pq} = \frac{\mu_{pq}}{\mu_{00}^{(1+(p+q)/2)}}, \quad p+q \geq 2$$

- Rotation-invariant (order 2) →

$$\phi_1 = \nu_{20} + \nu_{02}$$
$$\phi_2 = (\nu_{20} - \nu_{02})^2 + \nu_{11}^2$$

- Invariant to general linear transforms →

$$I_1 = \mu_{20}\mu_{02} - \mu_{11}^2$$
$$I_2 = (\mu_{30}\mu_{03} - \mu_{21}\mu_{12})^2 - 4(\mu_{30}\mu_{12} - \mu_{21}^2)(\mu_{21}\mu_{03} - \mu_{12}^2)$$

$$\psi_1 = \frac{I_1}{\mu_{00}^4} \qquad \psi_1 = \frac{I_1}{\mu_{00}^4}$$

- A set of moment invariants of different orders can be defined in a similar way

T.H. Reiss: The revised fundamental theorem of moment invariants. IEEE Trans. PAMI, vol. 13, no 8, pp. 830-834, 1991

Statistical Features: Zernike Moments

Geometric moments project the function f(x,y) over the monomial $x^p y^q$

No orthogonality => information redundancy

Zernike moments:

- Change to **polar coordinates** to achieve **orthogonality** and **rotation invariance**
- Project of the image over the Zernike polynomials V_{nm} which are orthogonal inside the unitary circle $x^2+y^2=1$

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

$$\text{where } j=\sqrt{-1}, \theta = a\tan(y/x),\ \rho=\sqrt{x^2+y^2}$$

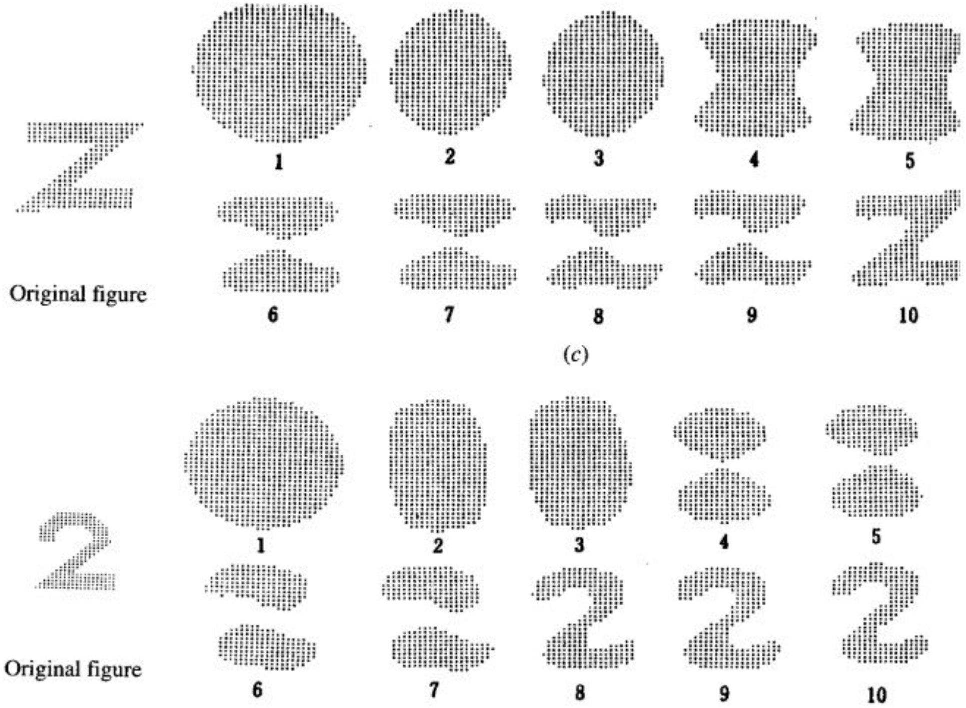
$$n \geq 0,\ |m| \leq n,\ n-|m| \text{ is even, and}$$

$$R_{nm}(\rho) = \sum_{s=0}^{(n-|m|)/2} \frac{(-1)^s \rho^{n-2s}(n-s)!}{s!\left(\frac{n+|m|}{2}-s\right)!\left(\frac{n-|m|}{2}-s\right)!}$$

A. Khotanzad, Y.H. Hong: Invariant image recognition by Zernike moments. IEEE Trans. PAMI, vol. 12, no 5, pp. 489-497, 1990.

Statistical Features: Zernike Moments

Image reconstruction using moments up to order 10 (66 moments)

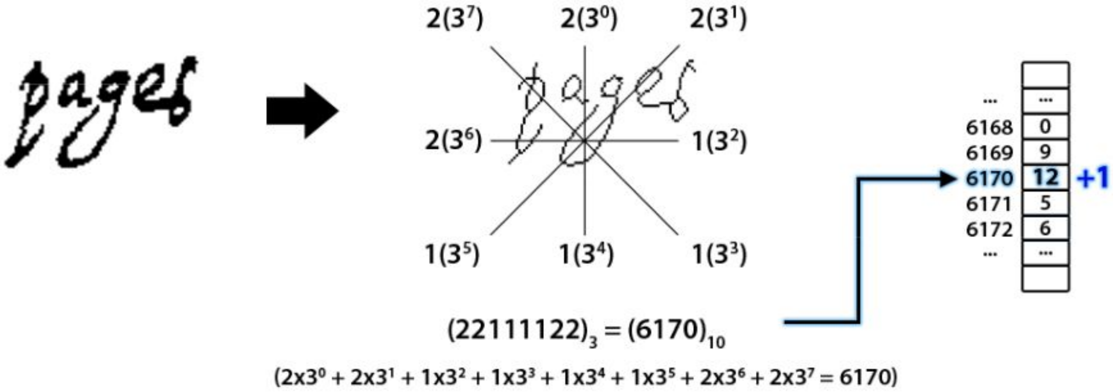


Histograms

Histograms: Characteristic LOCI features

A characteristic Loci feature in a given point p consists of the number of the intersections in four directions (up, down, right and left) or eight directions (considering the diagonals too).

Loci vectors are clustered in a codebook. The character is represented by the histogram of the most frequent codewords (Bag of Words structure).



H.A. Glucksman. Classification of mixed-font alphabets by characteristic loci. Proc. IEEE Comput. Conf., pages 138-141, September 1967.

Histograms: Shape contexts

Given a shape point p, its context is computed as an histogram of relational attributes between p and other shape points. These attributes are the length r and orientation q of vectors joining p and the other points.

To make the histogram more sensitive to positions of nearby points that to those of points farther away, the vectors are put into log-polar space.

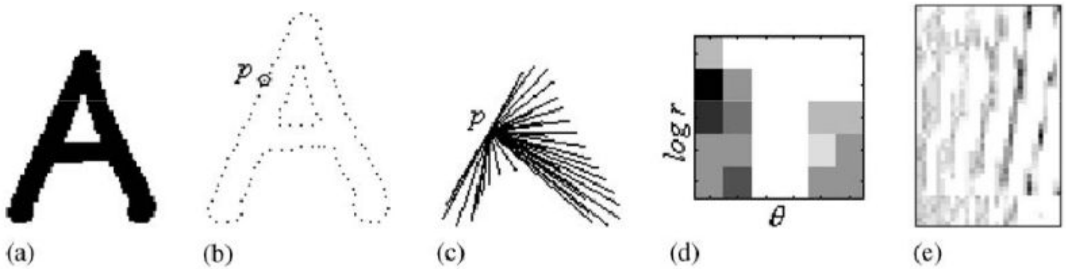


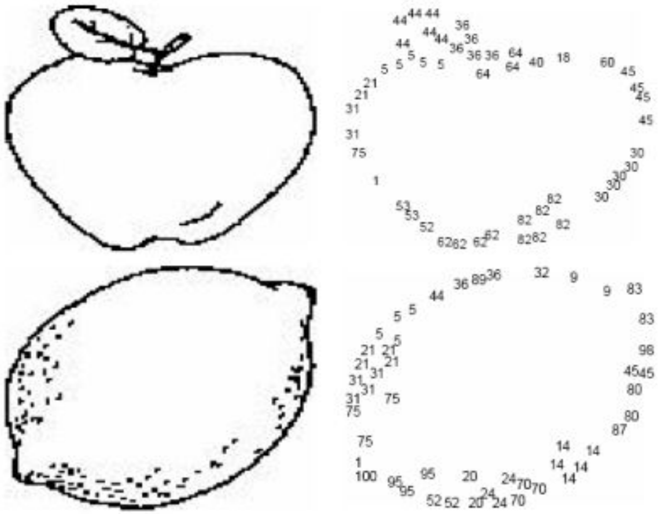
Fig. 3. Shape context. (a) a character shape; (b) edge image of (a); (c) a point p on shape (a) and all the vectors started from p; (d) the log-polar histogram of the vectors in (c), the histogram is the context of point p; (e) the context map of shape (a), each row of the context map is the flattened histogram of each point context, the number of rows is the number of sampled points. (reprinted from [10]).

Histograms: Shapemes

Given a shape context space with d bins (d-dimensional space), and s sample points of a shape

Vector quantization of the shape-context shape context space involves clustering the vectors and the representing each vector by the index of the cluster that it belongs to. Each cluster is a **shapeme**.

A shape is encoded as a **histogram** of shapeme frequencies.



(a) (b)

Figure 3: (a)Line drawing, (b)sampled points with shapeme labels. $k = 100$ shapemes were extracted from a known set of 260 shapes (26000 shape contexts). Note the similarities in shapeme labels (31,21,5 on left side, 45 on right side) between similar portions of the shapes.

Transform-based features

Transform-based features: Fourier Descriptors

Compute Fourier coefficients then extract features

High computational cost.

Can be used to describe the contour.

C. T. Zhan, R. Z. Roskies: Fourier descriptors for plane closed curves. Trans. on computers, vol. C-21, no 3, pp. 269-281,1972

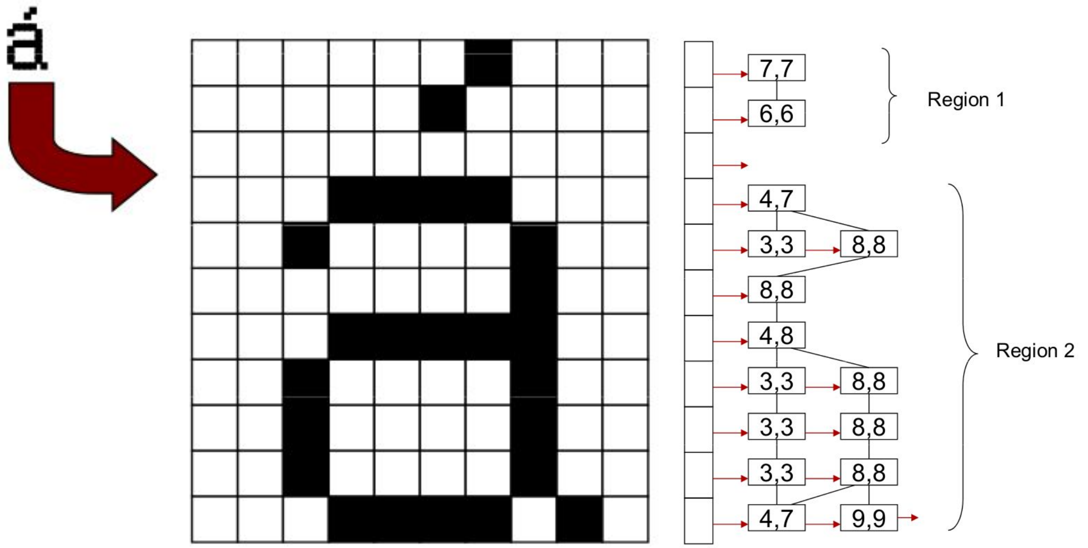
Structural Analysis

Structural Analysis: Run-length encoding

A graph is built on the run-length encoding, where:

Nodes: run-lengths.

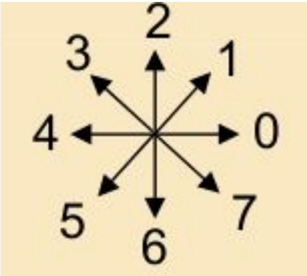
Edges: overlapping between runs in consecutive rows.



Structural Analysis: Chain-code

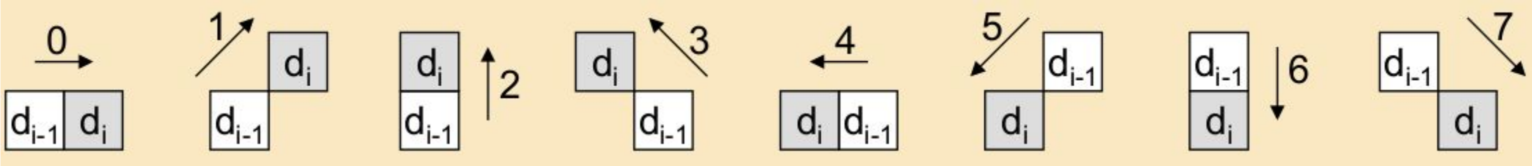
Chain-codes or *Freeman codes* are the simplest angular approximation.

They permit to code each vector d_i between two consecutive points of a contour with a code between 0 and 7.



The codification of a string S is composed of 3 fields:

- Starting coordinate
- Length
- Table of directions



Classification / comparison: string edit distance

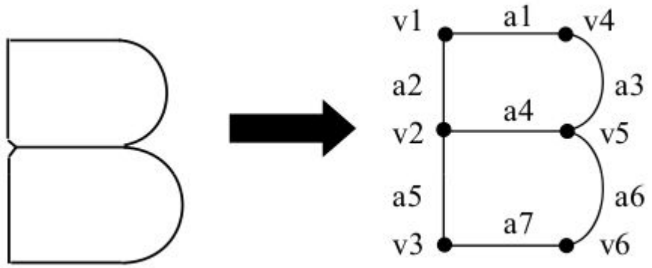
Structural Analysis: Skeleton Analysis

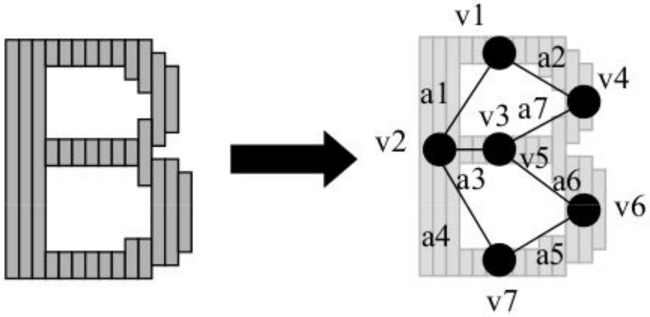
Representation with graphs or grammars.

Based on the detection of characteristic skeleton points and skeleton polygonal approximation.

Two possibilities to represent the skeleton with a graph:

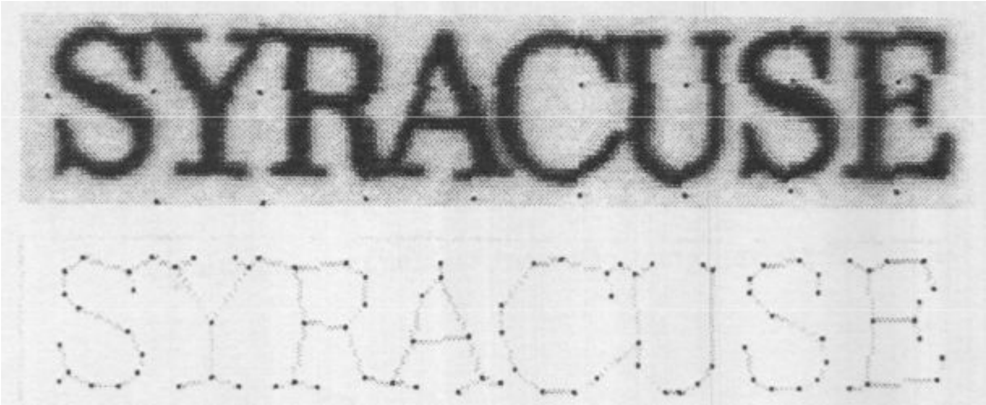
- Nodes are the characteristic points while edges are the segments joining the points
- Nodes are the segments of polygonal approximation while edges represent the adjacency relations between segments

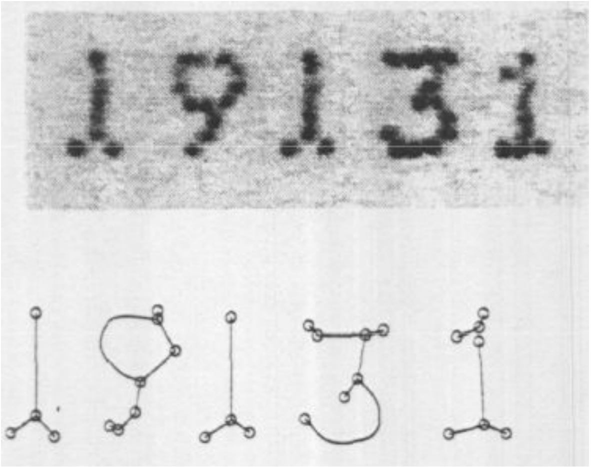




Structural Analysis: Skeleton Analysis

Representation with graphs:

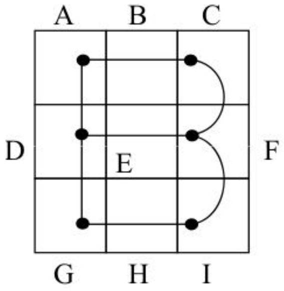




J. Rocha, T. Pavlidis: Character recognition without segmentation. IEEE Trans. on PAMI, vol. 17, no 9, pp. 903-909, 1995

Structural Analysis: Skeleton Analysis

Zoning:



Option 1: stroke length within each zone

Option 2: coding from the arcs:

ArC, ArD, CcF, DrF, DrG, FcI, GrI

where r = line, c = arc.

Discrete features:

- Number of loops
- Number of T joints and X joints
- Number of terminal points, corner points and isolated points
- Cross points with horizontal and vertical axes

Structural Analysis : Topological and geometric features

- Aspect ratio x-y
- Perimeter, area, center of gravity
- Minimal and maximal distance of the contour to the center of gravity
- Number of holes
- Euler number: (nb of connected components) - (nb of holes)
- Compacity: $(\text{perimeter})^2 / (4\pi \cdot \text{area})$
- Information about contour curvature
- *Ascenders* and *descenders*
- Concavities and holes
- Loops
- Unions, terminal points, crossings with horizontal and vertical axes
- Angular information: histogram of segment angles