

Document Image Analysis

July 2015

Alicia Fornés
aforres@cvc.uab.es

VISUM Summer School



VISion Understanding and Machine intelligence

Contents

Introduction

Main tasks:

- Layout analysis
- Recognition
- Classification / Writer Identification
- Search and Spotting

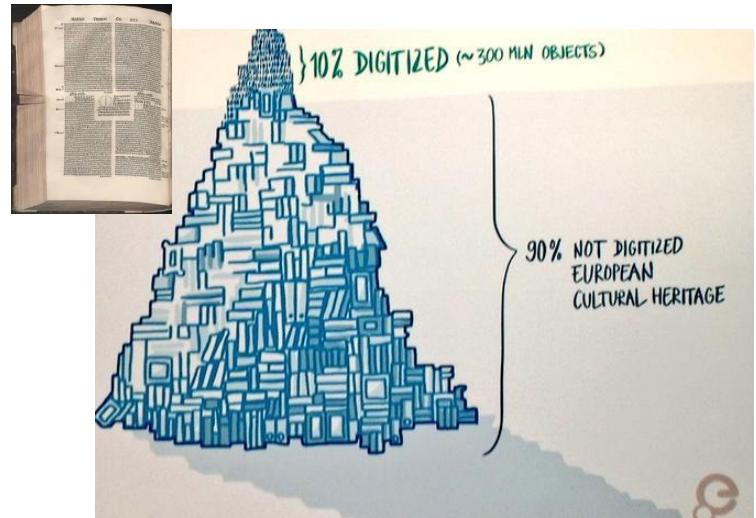
INTRODUCTION

Context and Volume of Documents

• Historical Documents

Europeana estimates that only 10% of all European cultural heritage has been digitized

- Preservation is needed
- Information Extraction?
(raw image pixels not useful)



• Administrative Documents

Datafinity (Enterprise Content Management systems provider) estimates that companies receive 3 Million documents per year with an estimated cost of manual processing of £0.15-0.25 per item.

E.g. checks, invoices, orders, administrative forms



Textual documents

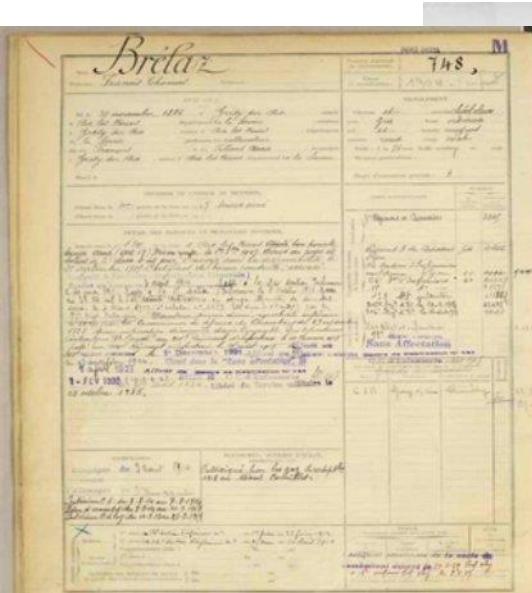
Civil registers

Proceedings of plenary sessions

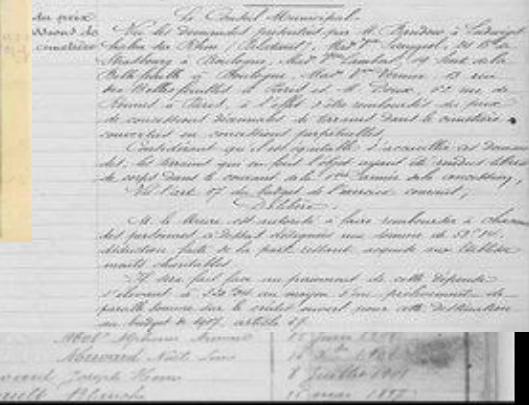
Notary records

Military records

Historic newspapers

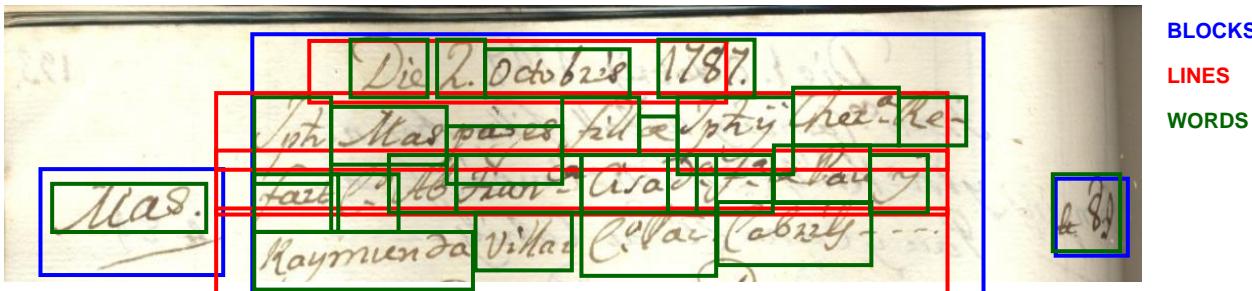


Session ord^e d'août
Séance du 29 Août 1907



Text documents: main tasks

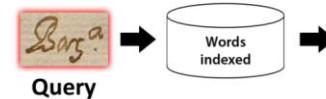
Layout analysis: to detect (crop) records, lines, words for subsequent recognition.



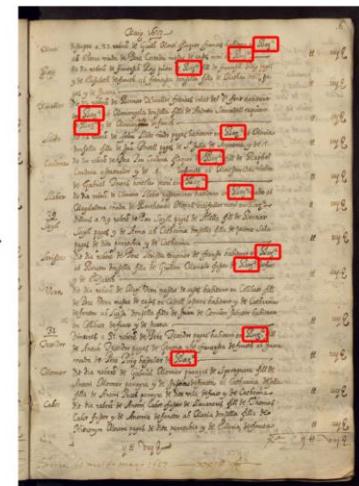
Full transcription: to convert images to editable text.

Dit dia rebere\$ de Hieronym Ponsich corder de Bar^(a) fill de Jua\$ Pon=

dit dia rebere\$ de Hieronym Ponsich corder de Bar^(a) fill de Jua\$ Pon=



Word spotting: given a query word to search, to locate at image level visually similar word snippets.



Document Classification

Invoices



Passports



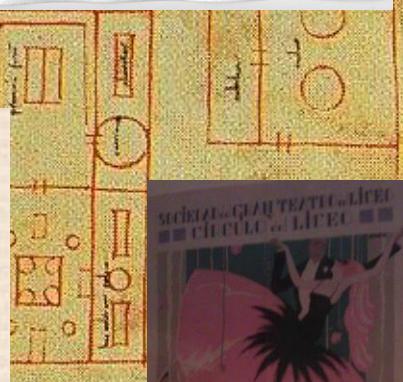
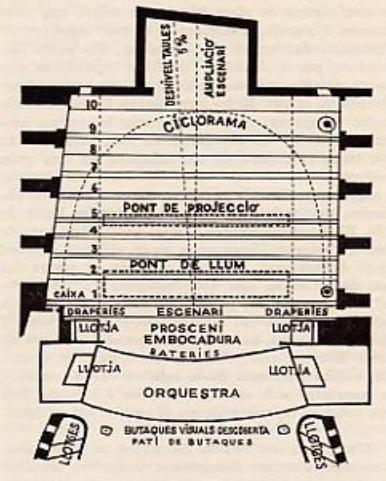
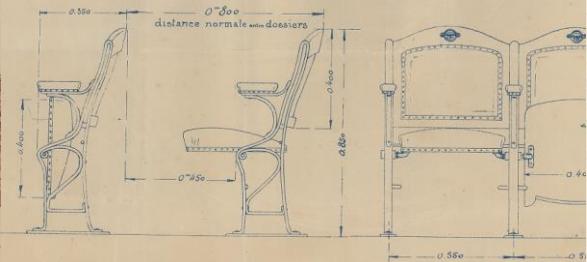
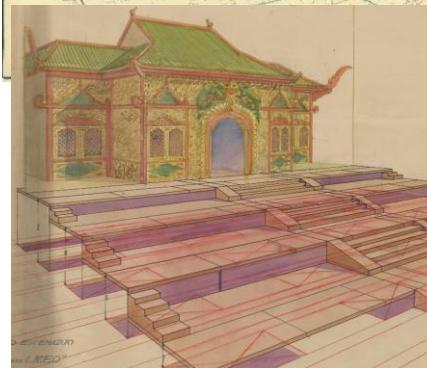
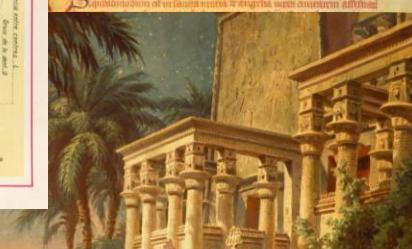
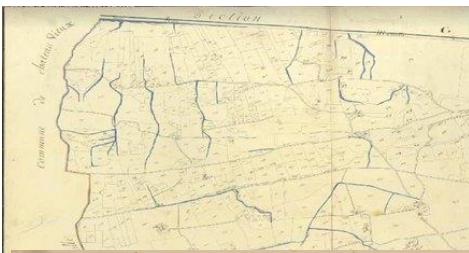
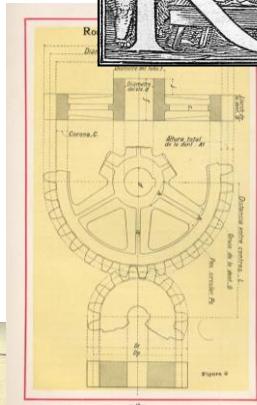
Checks

Checks

Forms

Graphic documents

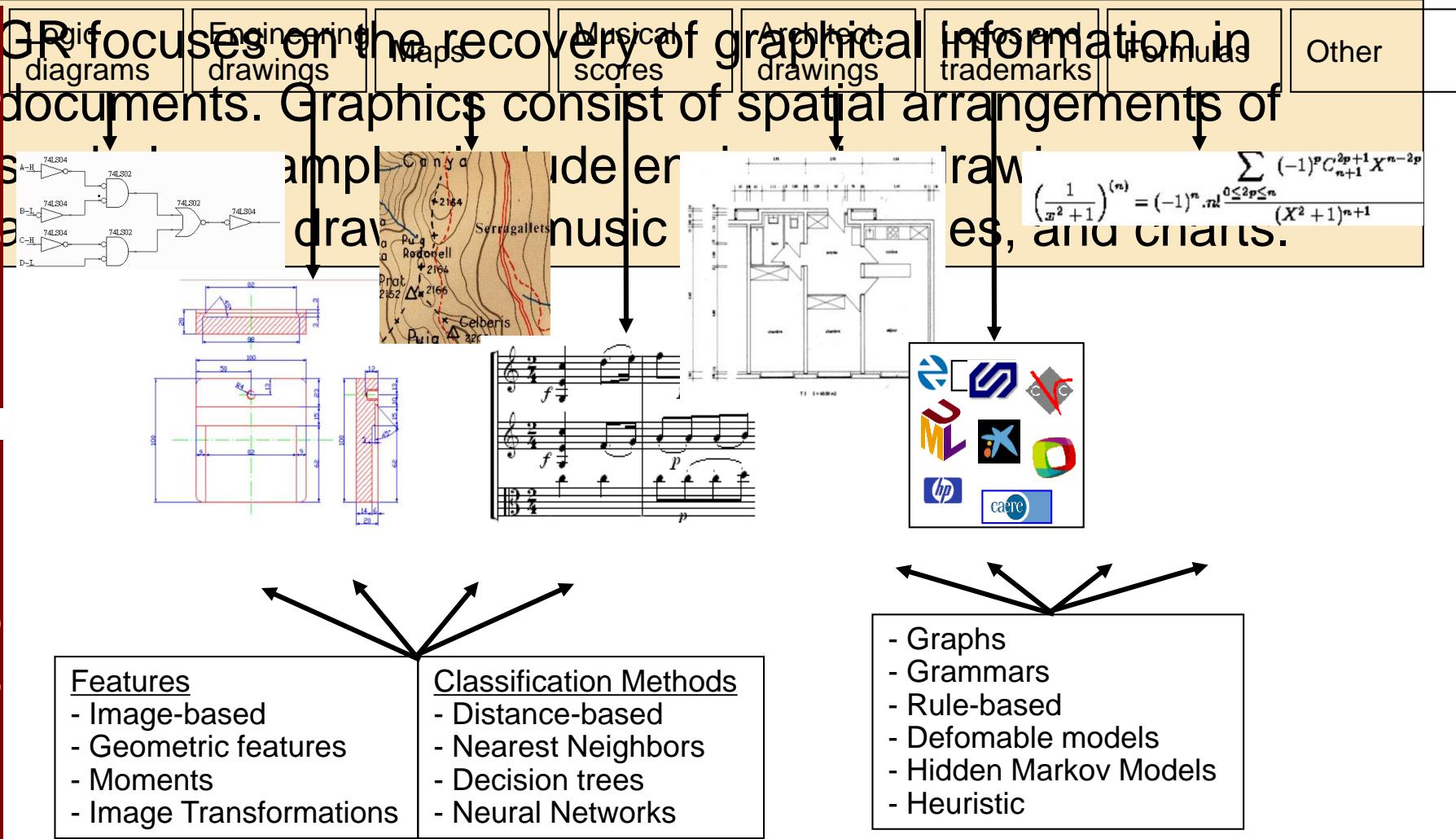
Pictures
Paintings
Maps
Music Scores
Drawings



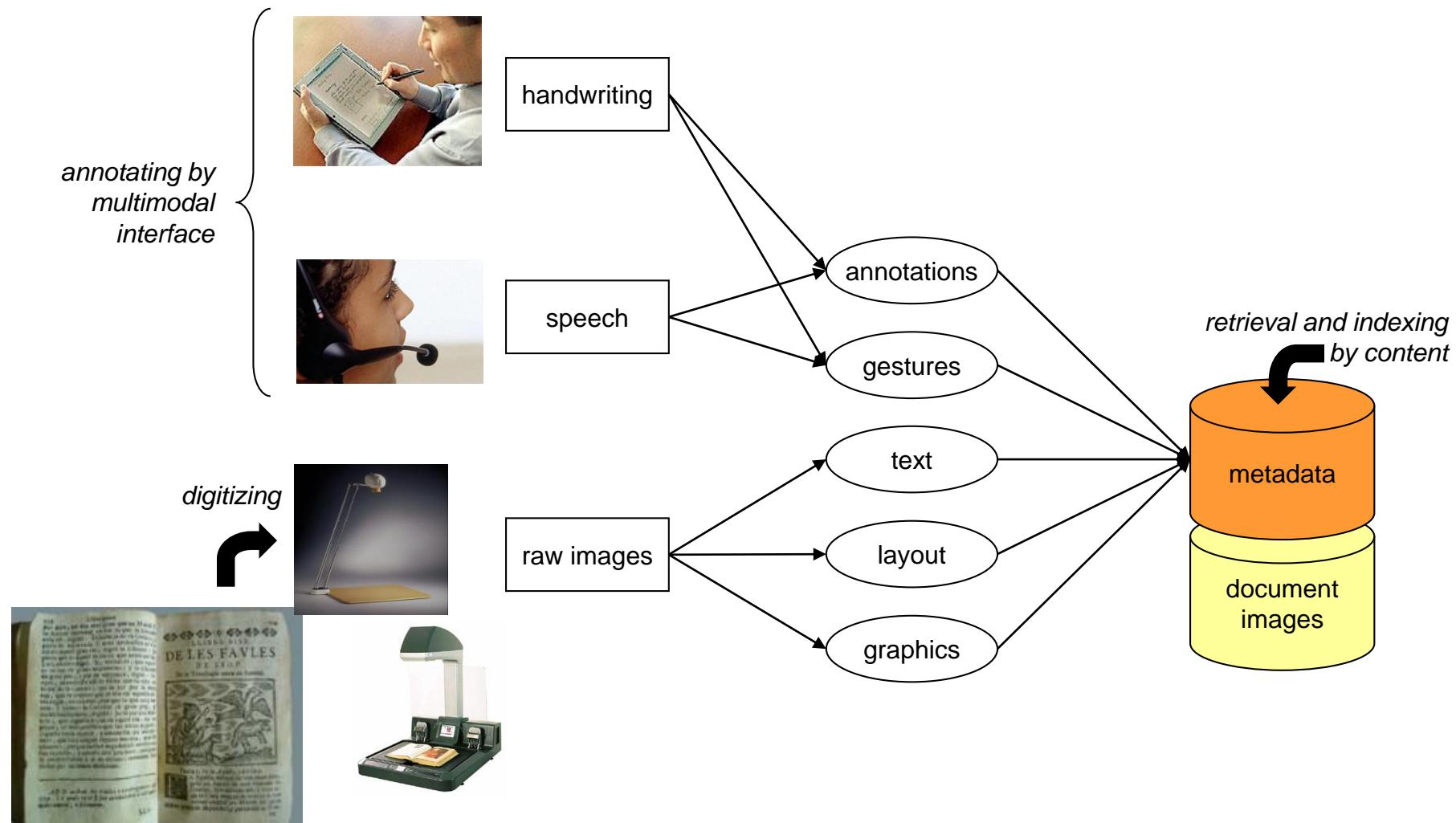
Graphics Recognition

APP. DOMAINS

METHODS



Information extraction with Pattern Recognition and Document Image Analysis techniques

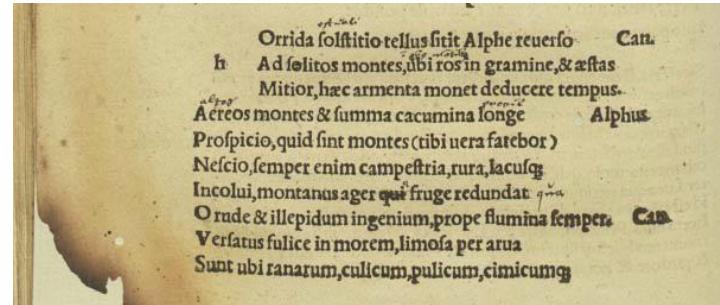


Challenges

- Legibility Enhancement
- Recognition
 - Text recognition
 - Graphics recognition
- Document Understanding
 - Layout Analysis
 - Information extraction
 - Writer identification

ENHANCEMENT

Difficulties



effectus.

Eur. 333 (37)

Kurtzer vñ warbaffter bericht
vnd vergriff / Der unwillichen gewalts / Schmiede

verständigt, so in Stuttgart. Die folgt die Taktik, in ihren Programmen W und nicht mehr zu erwähnen. Deutschen Kandidaten, so neuestens Schäffn.

für das Zollparlament sind auf den

Austria

In dem Pesther „Lloyd“ bringt die Wg des Vorgehens der Regierung in Rom eine präzise Angabe der Convenienzen Beseitigung von Oesterreich beanpruh hierauf dem österr. Botschafter in Rätes Exposé des Cultusministers zu

Intra et adverbii loci et post loci aliquas inclusum eis operosum est extra et post aliquas loci interioris exclusionem. Inde volumen neutrale inter 45 are.

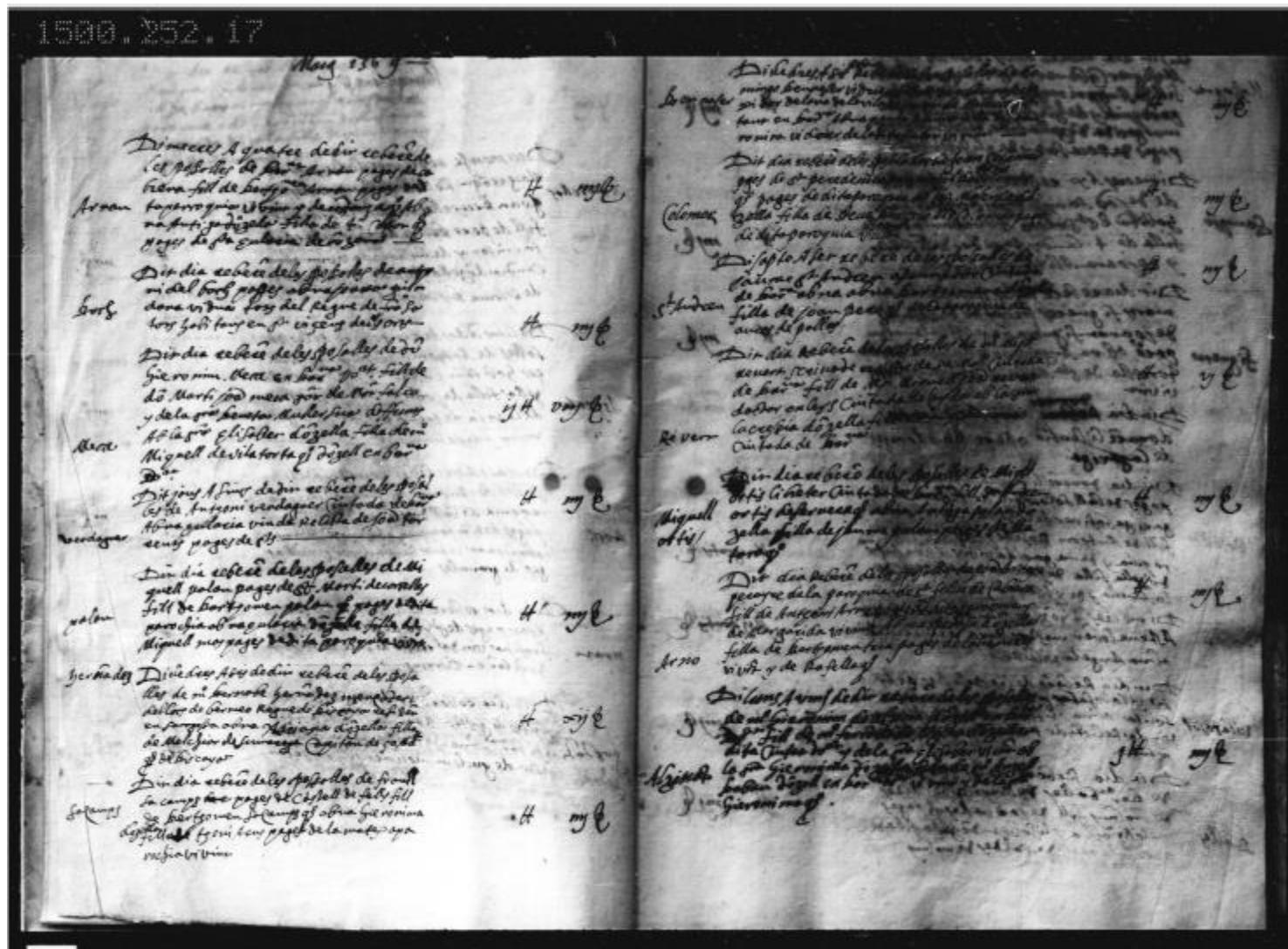
Et psalles dragma datur hinc tibi celica tracma

In medio pausa nec finis sit sine pausa

Cantibus et prosis apud hunc semper tibi prosis

Chic doceat quod et qualiter deo sit psallenda. Nisi remunerat psalmista celesti corona et quod in medio versus ultimi et fine sit fac.

Show through cancellation



Show through cancellation

1500. 252. 17

Meng 1569

Dimecres 4 quarts de diri es creu
el coster de don Joan pagos del
recau del bens d'una paga en
temporada i un altre d'abans i d'apres
en paga de la qual era filla de don Joan
paga de la qual era filla de don Joan

Dit dia es creu de la paga de la
viuda dels pags abans pagos que
dava en una foy del regne de Valo-
torn fols temps en temps del govern

Dit dia es creu de la paga de la
germana, veire en foy del regne de Valo-
torn, d'una maja que ha donat al
vila que havia de la qual era filla de
l'infant que està en la casa de
Miguel de la villa de la qual era filla de don Joan

Dit dia es creu de la paga de la
viuda dels pags abans pagos que
dava en una foy del regne de Valo-
torn fols temps en temps del govern
en paga de la qual era filla de don Joan

Dit dia es creu de la paga de la
viuda dels pags abans pagos que
dava en una foy del regne de Valo-
torn fols temps en temps del govern
en paga de la qual era filla de don Joan

Dit dia es creu de la paga de la
viuda dels pags abans pagos que
dava en una foy del regne de Valo-
torn fols temps en temps del govern
en paga de la qual era filla de don Joan

Dit dia es creu de la paga de la
viuda dels pags abans pagos que
dava en una foy del regne de Valo-
torn fols temps en temps del govern
en paga de la qual era filla de don Joan

H - mgf

Dit dia es creu de la paga de la
viuda dels pags abans pagos que
dava en una foy del regne de Valo-
torn fols temps en temps del govern
en paga de la qual era filla de don Joan

Dit dia es creu de la paga de la
viuda dels pags abans pagos que
dava en una foy del regne de Valo-
torn fols temps en temps del govern
en paga de la qual era filla de don Joan

Dit dia es creu de la paga de la
viuda dels pags abans pagos que
dava en una foy del regne de Valo-
torn fols temps en temps del govern
en paga de la qual era filla de don Joan

Dit dia es creu de la paga de la
viuda dels pags abans pagos que
dava en una foy del regne de Valo-
torn fols temps en temps del govern
en paga de la qual era filla de don Joan

Dit dia es creu de la paga de la
viuda dels pags abans pagos que
dava en una foy del regne de Valo-
torn fols temps en temps del govern
en paga de la qual era filla de don Joan

Dit dia es creu de la paga de la
viuda dels pags abans pagos que
dava en una foy del regne de Valo-
torn fols temps en temps del govern
en paga de la qual era filla de don Joan

Dit dia es creu de la paga de la
viuda dels pags abans pagos que
dava en una foy del regne de Valo-
torn fols temps en temps del govern
en paga de la qual era filla de don Joan

Dit dia es creu de la paga de la
viuda dels pags abans pagos que
dava en una foy del regne de Valo-
torn fols temps en temps del govern
en paga de la qual era filla de don Joan

H - mgf

Binarization

turn me of
to star saw
I very much
y for the fairs
as appear
had to command
the same go

(a)

turn me of
to star saw
I very much
y for the fairs
as appear
had to command
the same go

(b)

turn me of
to star saw
I very much
y for the fairs
as appear
had to command
the same go

(c)

turn me of
to star saw
I very much
y for the fairs
as appear
had to command
the same go

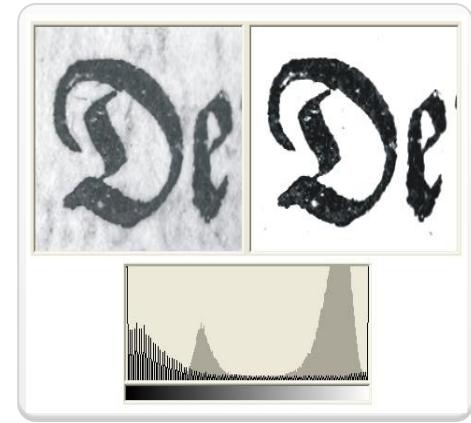
(d)

turn me of
to star saw
I very much
y for the fairs
as appear
had to command
the same go

(e)

turn me of
to star saw
I very much
y for the fairs
as appear
had to command
the same go

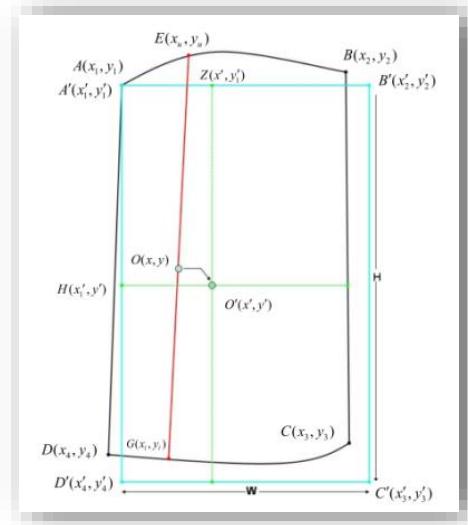
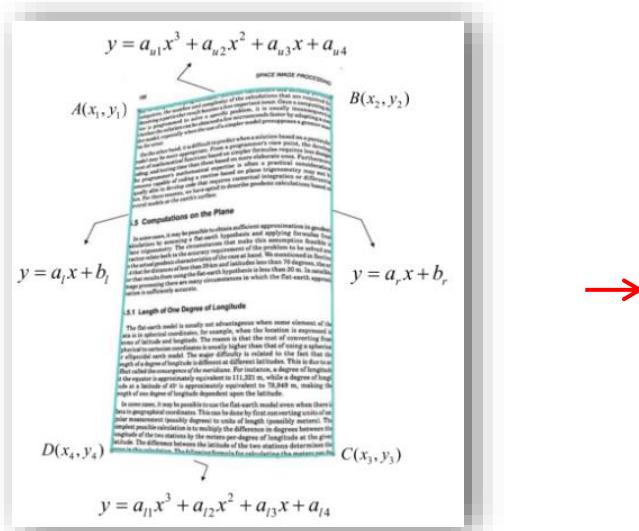
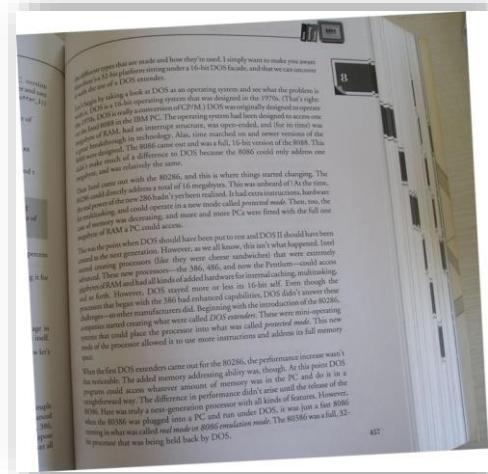
(f)



Dewarping

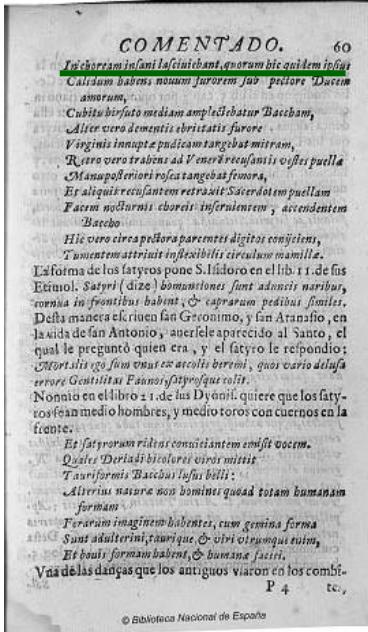
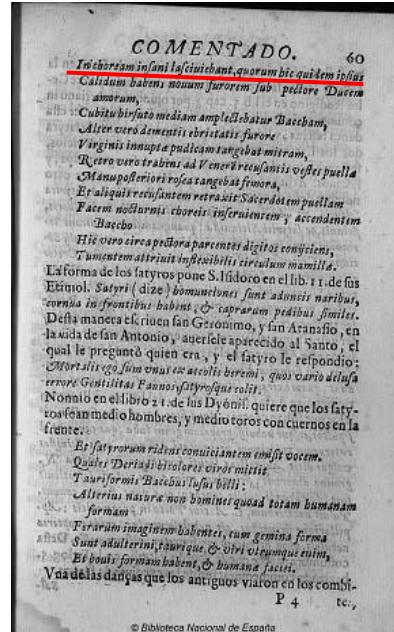
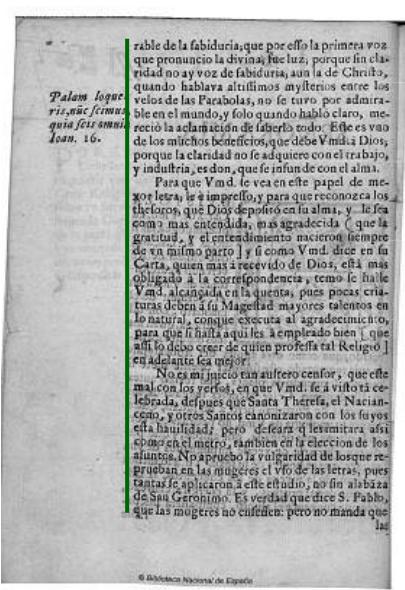
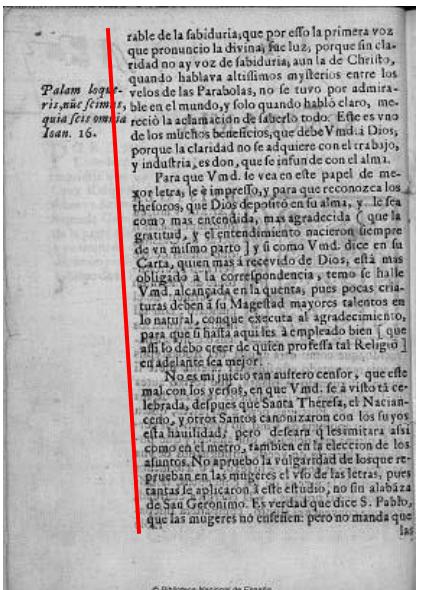
- The recognition will be easier

~~TEA WAS ready when children washed their hands and hair. Peter was sharing Brock's little room, Pam had a tiny room to herself up in the attic. She loved it because it had queer, slanting ceilings, and even boards in the floor. She looked out over her hair, humming "I was so happy in the noon,"~~



De-Skew

■ Correct the angle



LAYOUT ANALYSIS, SEGMENTATION AND NORMALIZATION

Layout analysis

Detection of the document logical structure (lines, paragraphs, columns) from the physical structure (image blocks).

NAGY, TWENTY YEARS OF DOCUMENT IMAGE ANALYSIS IN PAMI

TABLE 3
A Document Taxonomy

Type	Example	DIA Task	Ancillary data
plain text (narrative or descriptive)	Moby Dick, Gettysburg Address, NY Times, Vogue	extract correct word order; separate and reassemble articles; pointers to illustrations	English lexicon; publication-specific format
scholarly & technical text	JLBE-PAML, Dr. Dobbs Journal	index; author, title, page; pointers to refs., figs., tables, footnotes, equations	abbreviations, acronyms, units
formal text	program listing, chess, bridge, recipe	extract executable, or compilable, form	program, chess, bridge syntax
letter, envelope	information request, complaint, recommendation	extract routing info; index: sender, date, subject	directories
directory	telephone directory, street index	extract name-attribute pairs	previous edition
structured list	organization chart, table of contents, catalog	recover hierarchy; cross-references	previous edition
business form	order, invoice, subscription, survey, IRS-1040	link field content to dbms; convert to SGML or XML format	formatted data, dbms, workflow system, lexicons
engineering drawing	assembly or part drawing; isometric view	convert to CAD format	part lists, drawing standards
schematic diagram	circuits, utility maps	extract net list or convert to CAD format	P-SPLIC, manhole inventory
map	topographic grid, street map, road map	convert to GIS format	gazetteer, other maps, GIS
music score	Moonlight Sonata	recover MIDI representation	music syntax
table	airline schedules, stock quotes	construct formal model: headers <entries	airline and stock abbreviations, previous edition

compression methods simply to avoid disk access during page analysis. Run-length coding (RLC) and Freeman chain codes were used early on. Methods that came along later tried to reduce the size of the segmented regions. Niblack's method [30] on a grayscale image at the mean gray-level of a 15×15 window by a fixed fraction (0.2) of the standard deviation of the gray levels, gave the best results on their maps. (A small modification is necessary when it is evident that the entire window is covered by a large foreground blob.) They recommended postprocessing with the method of Yanowitz and Bruckstein, which iteratively creates a threshold surface that is essentially a low-pass-filtered version of the reflectance map. They also reported that character segmentation and recognition did not necessarily benefit from direct gray-scale processing as opposed to adaptive binarization [86].

Textured backgrounds are particularly difficult to handle. Liu and Sehri [53] provide a solution for postal address readers. It requires: 1) preliminary binarization sustained, thorough comparison and evaluation of published adaptive binarization methods (including their own) on hydrographic charts [83], [84], [85], [86]. Niblack's method, based on a threshold set below the mean gray-level of a 15×15 window by a fixed fraction (0.2) of the standard deviation of the gray levels, gave the best results on their maps. (A small modification is necessary when it is evident that the entire window is covered by a large foreground blob.) They recommended postprocessing with the method of Yanowitz and Bruckstein, which iteratively creates a threshold surface that is essentially a low-pass-filtered version of the reflectance map. They also reported that character segmentation and recognition did not necessarily benefit from direct gray-scale processing as opposed to adaptive binarization [86].

Most early document scanners had hardware reflectance thresholds, but current scanners typically produce 8-bit gray-scale (or color) output. Researchers from the University of Oslo and Michigan State University conducted a

2.2 Binarization

Most early document scanners had hardware reflectance thresholds, but current scanners typically produce 8-bit gray-scale (or color) output. Researchers from the University of Oslo and Michigan State University conducted a

NAGY, TWENTY YEARS OF DOCUMENT IMAGE ANALYSIS IN PAMI

Table heading

TABLE 3
A Document Taxonomy

Type	Example	DIA Task	Ancillary data
plain text (narrative or descriptive)	Moby Dick, Gettysburg Address, NY Times, Vogue	extract correct word order; separate and reassemble articles; pointers to illustrations	English lexicon; publication-specific format
scholarly & technical text	JLBE-PAML, Dr. Dobbs Journal	index; author, title, page; pointers to refs., figs., tables, footnotes, equations	abbreviations, acronyms, units
formal text	program listing, chess, bridge, recipe	extract executable, or compilable, form	program, chess, bridge syntax
letter, envelope	information request, complaint, recommendation	extract routing info; index: sender, date, subject	directories
directory	telephone directory, street index	extract name-attribute pairs	previous edition
structured list	organization chart, table of contents, catalog	recover hierarchy; cross-references	previous edition
business form	order, invoice, subscription, survey, IRS-1040	link field content to dbms; convert to SGML or XML format	formatted data, dbms, workflow system, lexicons
engineering drawing	assembly or part drawing; isometric view	convert to CAD format	part lists, drawing standards
schematic diagram	circuits, utility maps	extract net list or convert to CAD format	P-SPLIC, manhole inventory
map	topographic grid, street map, road map	convert to GIS format	gazetteer, other maps, GIS
music score	Moonlight Sonata	recover MIDI representation	music syntax
table	airline schedules, stock quotes	construct formal model: headers <entries	airline and stock abbreviations, previous edition

compression methods simply to avoid disk access during page analysis. Run-length coding (RLC) and Freeman chain codes were used early on. Methods that came along later included reduced terminal sequences of context-free grammars [43], coding on hexagonal meshes [94], production rules for subblocks [58], and filtered contours [10]. The authors of [58] used the *Proceedings of the IEEE on Selected Areas in Communications* as a testbed, reporting that their method gave better results than the others. They also reported that character segmentation and recognition did not necessarily benefit from direct gray-scale processing as opposed to adaptive binarization.

Most early document scanners had hardware reflectance thresholds, but current scanners typically produce 8-bit gray-scale (or color) output. Researchers from the University of Oslo and Michigan State University conducted a

Document heading

Table

Text columns

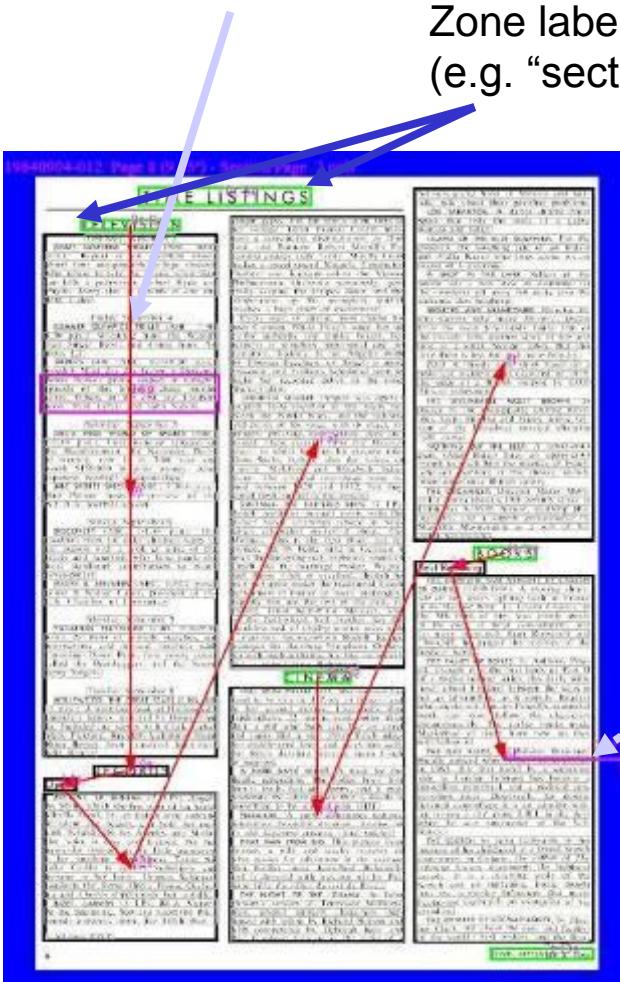
Physical Structure



Logical Structure

Logical layout analysis

Article flow links



Zone labels
(e.g. "section")

Ad. detection

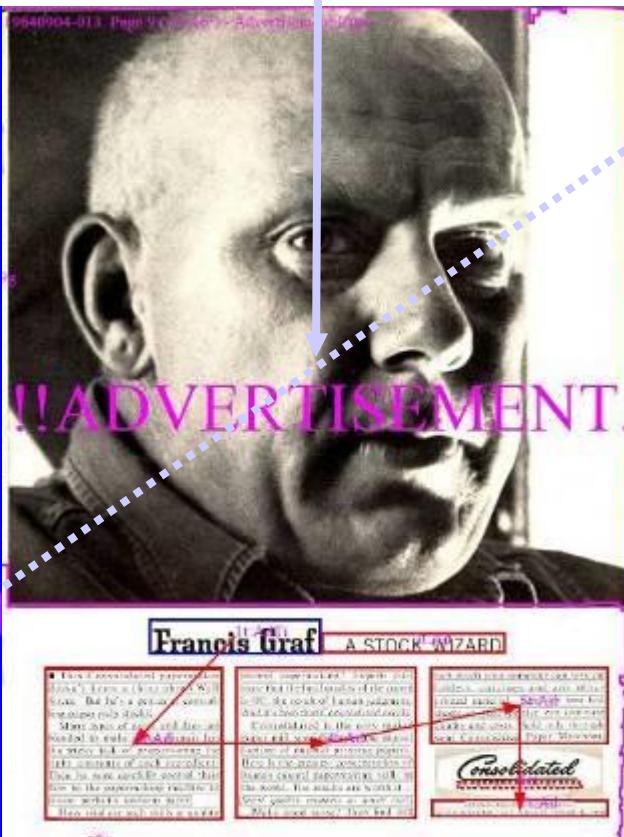
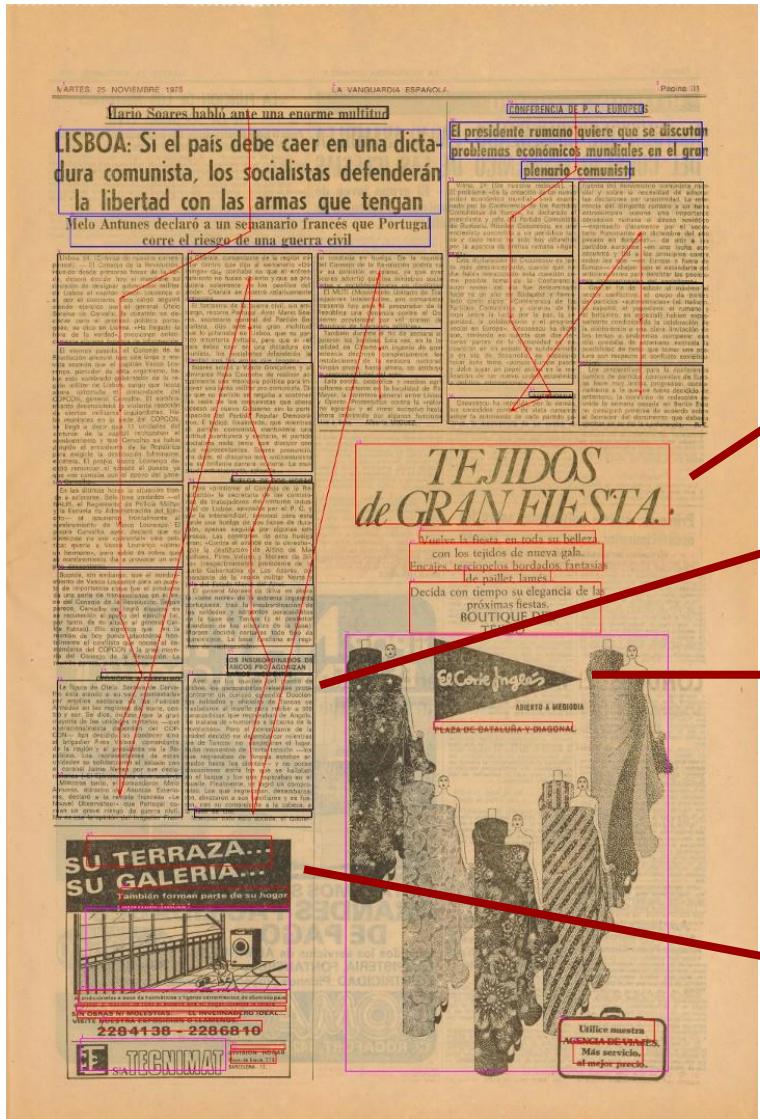


Image zone
(from *podcore*)



Cross-page link

Layout Analysis: Problems



Big fonts



Background elimination



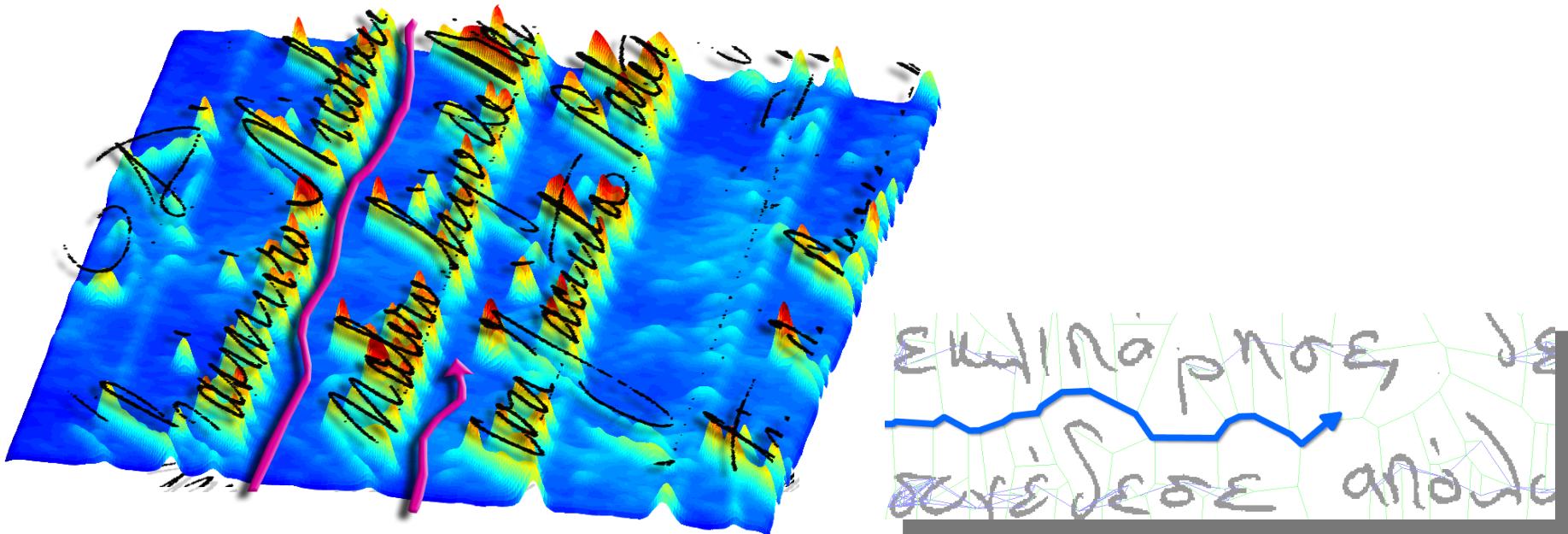
Logo recognition / advertisements



Non standard text / fonts

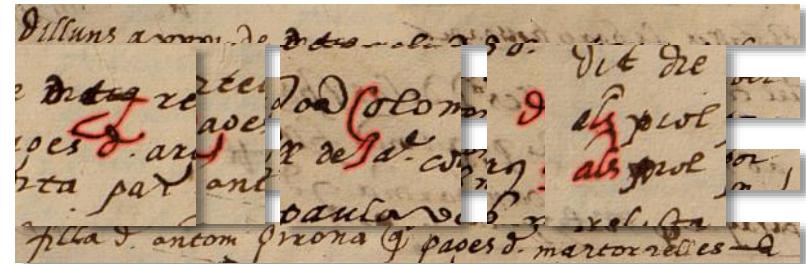
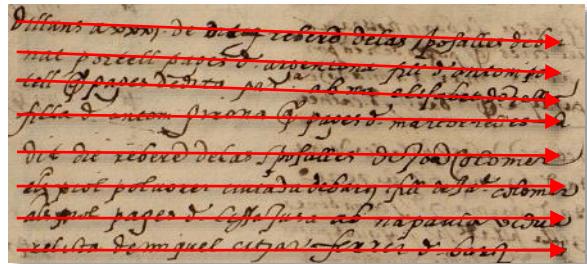
Line Segmentation

- Layout segmentation and, in particular, **line segmentation** is a key step to guarantee a good performance in the recognition.
- Humans tend to write the text in blocks with the same space between them.
- In a 3D view this characteristic can be seen as valley.
- The objective: to detect path between through the valleys of the document.

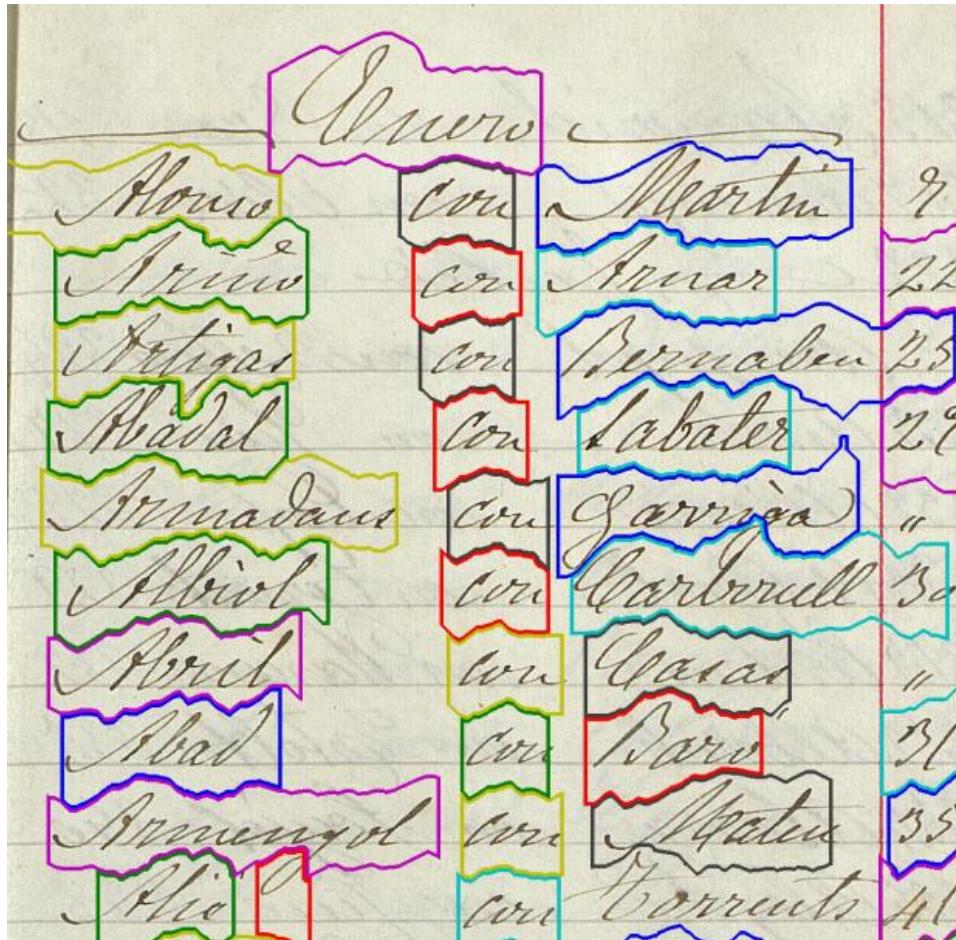


Line Segmentation: Main Difficulties

- Different skew
- Horizontal overlapping.
- Touching components.



Layout Analysis: Word segmentation

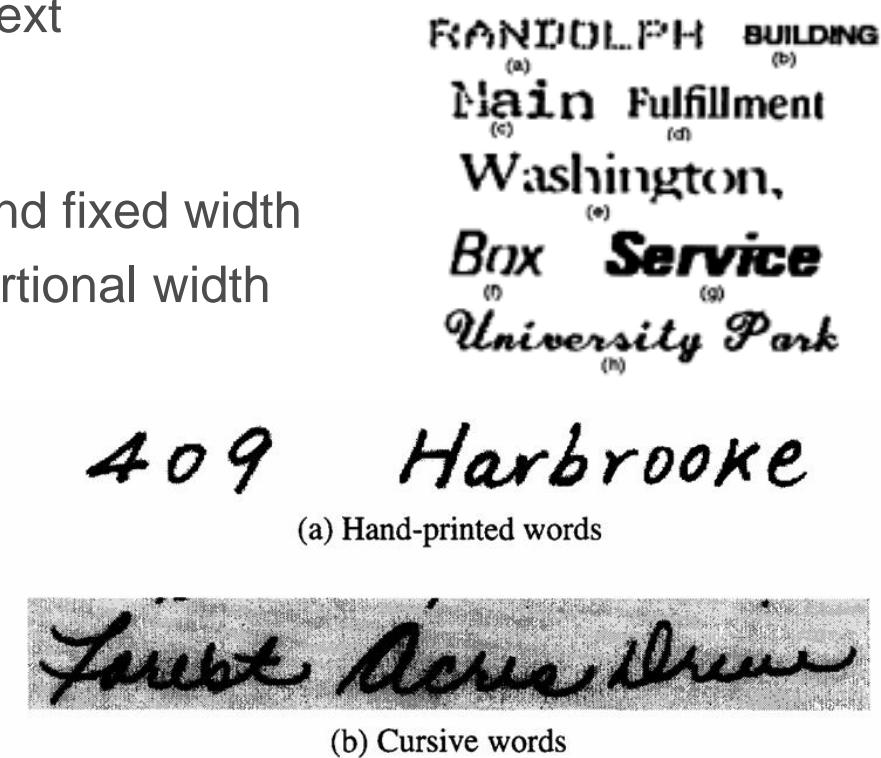


Character segmentation

Segmentation of characters in blocks of text

Levels of difficulty:

- Characters with uniform separation and fixed width
- Well separated characters with proportional width
- Broken characters
- Touching characters
- Broken and touching characters
- Cursive script
- Hand-printed words
- Handwritten cursive words.



Character segmentation: options

External segmentation:

- Segmentation before recognition. Independent processes
- The goal is to find the exact location of character separation
- Low performance with cursive script, touching characters or handwriting

Internal segmentation:

- Based on *Sayre Paradox*: a letter cannot be segmented without being recognized and cannot be recognized without being segmented
- Segmentation and recognition are done at the same time
- Recognition generates or validates segmentation hypothesis

Holistic methods

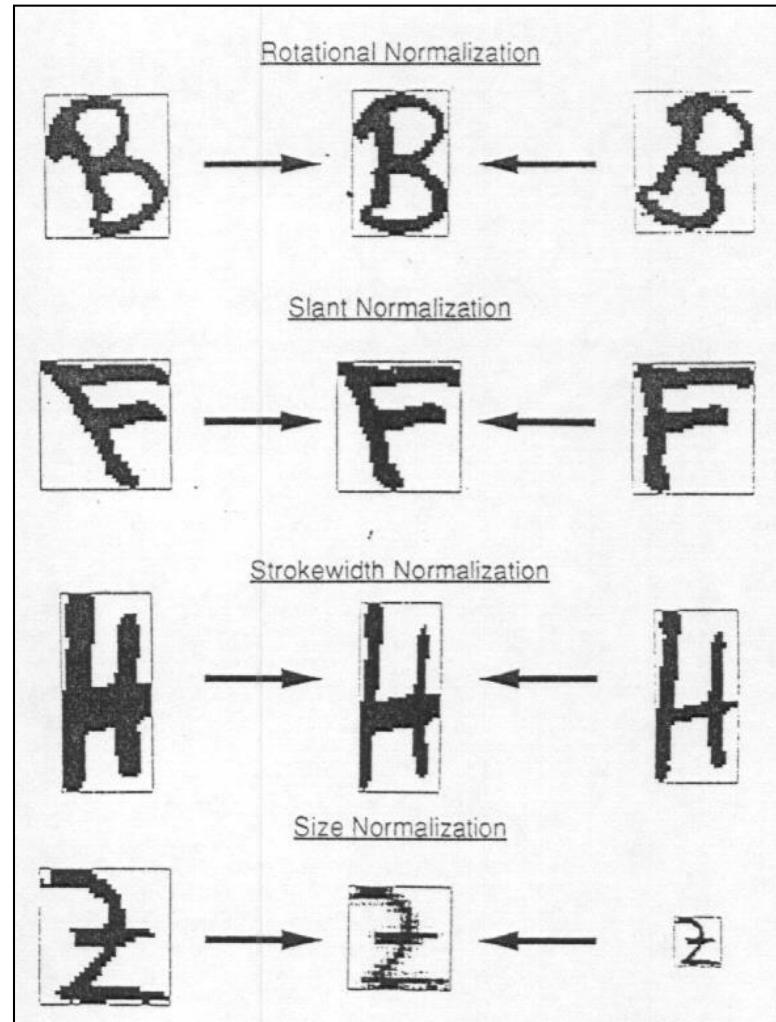
- No character segmentation
- Recognition tries to recognize words without recognizing individual characters

Normalization

Inverse transforms to reduce intra-class variance

The most usual normalization transforms are:

- **Rotation.** Rotated scan, text in graphic documents
- **Slant.** Cursive fonts or handwriting
- **Stroke thickness.** Bold fonts of very thin strokes, handwriting with different pen thickness
- **Size.** Titles, footnotes, handwriting



Links

- Software imatge enhancement Limb
www.i2s-digibook.com
- Software OCR Abby
<http://www.abbyyeu.com/es/>
- Software from IMPACT project
<http://www.digitisation.eu/>

TEXT RECOGNITION

What is an Optical Character Recognition (OCR)?

My invention relates to statistical machines of the type in which successive comparisons are made between a character and a charac-



OCR

My invention relates to statistical machines of the type in which successive comparisons are made between a character and a charac-

Some examples

Handbook of Character Recognition and Document Image Analysis, pp. 1-47
Edo. H. Barth and P. S. F. Wang
© 1997 World Scientific Publishing Company

IMAGE PROCESSING METHODS FOR DOCUMENT IMAGE ANALYSIS

THIEN M. HA
Institut für Informatik und Angewandte Mathematik
Universität Bern
Heubruckstrasse 10, CH-3012 Bern
Switzerland

H. W. WILKE
Institut für Informatik und Angewandte Mathematik
Universität Bern
Heubruckstrasse 10, CH-3012 Bern
Switzerland

This chapter describes image processing methods for document image analysis. The methods are grouped into four categories, namely, image acquisition, image transformation, feature extraction, and pattern recognition. The first section describes the process of converting a document into its numerical representation, including image coding as a means to reduce the storage requirements. Image transformation techniques are used to reduce the storage requirements. Image transformation techniques ranging from geometrical correction, filtering and figure-background separation to boundary detection and thresholding are described. Feature extraction techniques include connected component labeling, X-Y tree decomposition, run-length encoding, and rough transition graphs. Pattern extraction methods, which constitute the basis of image classification, are presented.

Keywords: Image acquisition; Image transformation; Segmentation; Feature extraction; Spatial sampling; Quantization; Image coding; Geometrical correction; Filtering; Figure-ground separation; Boundary detection; Statistical features

1. Introduction

Document image analysis, as its name indicates, is a subfield of image analysis. This implies that, on the one hand, it inherits the more general techniques of image analysis, and on the other, it can serve as a platform for testing various image processing techniques. Furthermore, with time, document image analysis has also acquired its own techniques, specifically designed for its needs. The aim of this chapter is to introduce the reader to the basic analysis techniques that have been proposed in literature or are currently used in various commercial systems. However, the main concern is to provide a tutorial rather than a description of the most up-

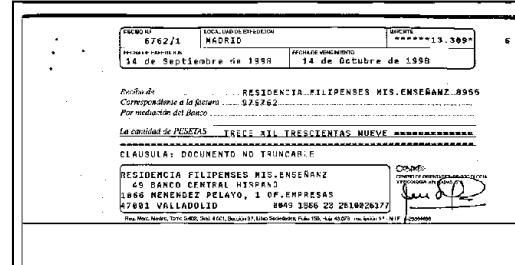
1

ILLUSTRATION

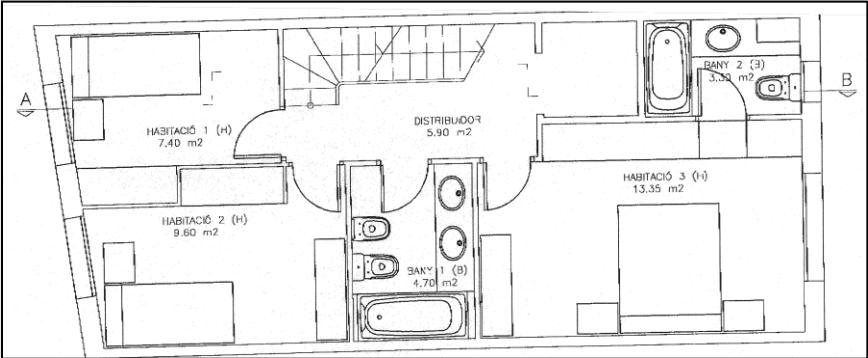
Books, journals, reports



Historical documents



Cheques, bills



Drawings, maps



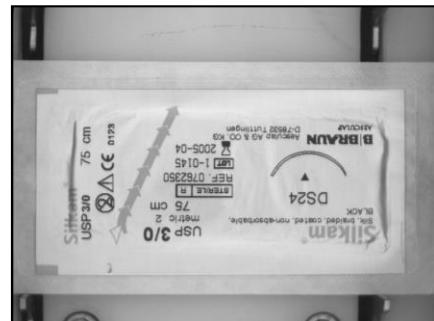
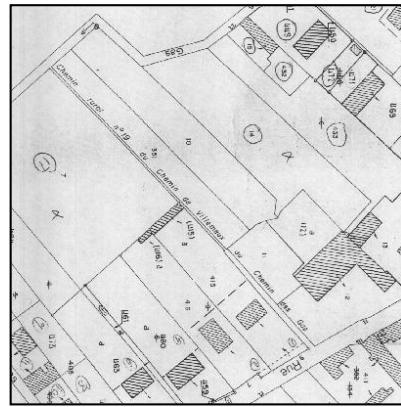
Smartphones

License plates

Identity cards

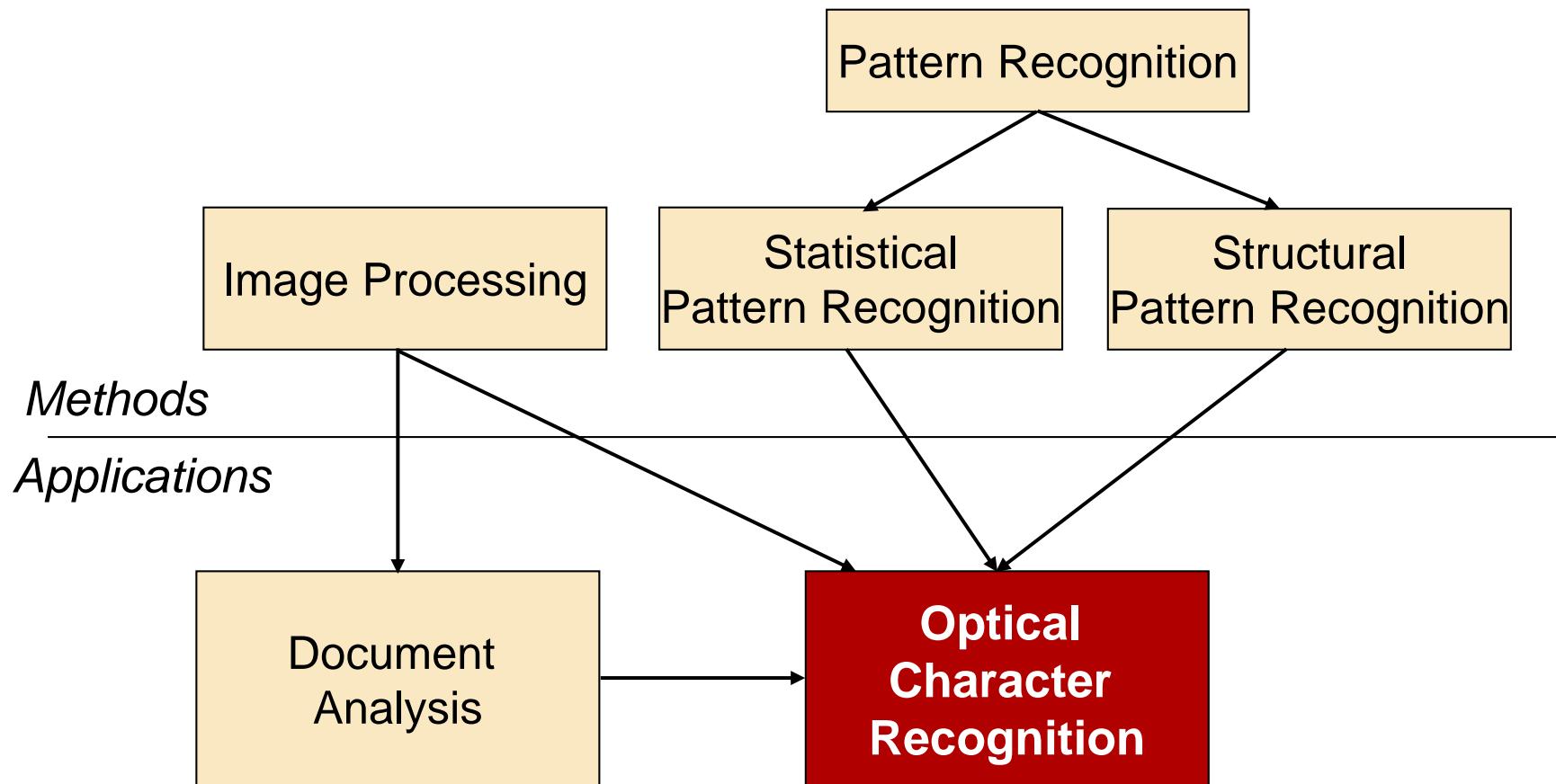


Postal addresses

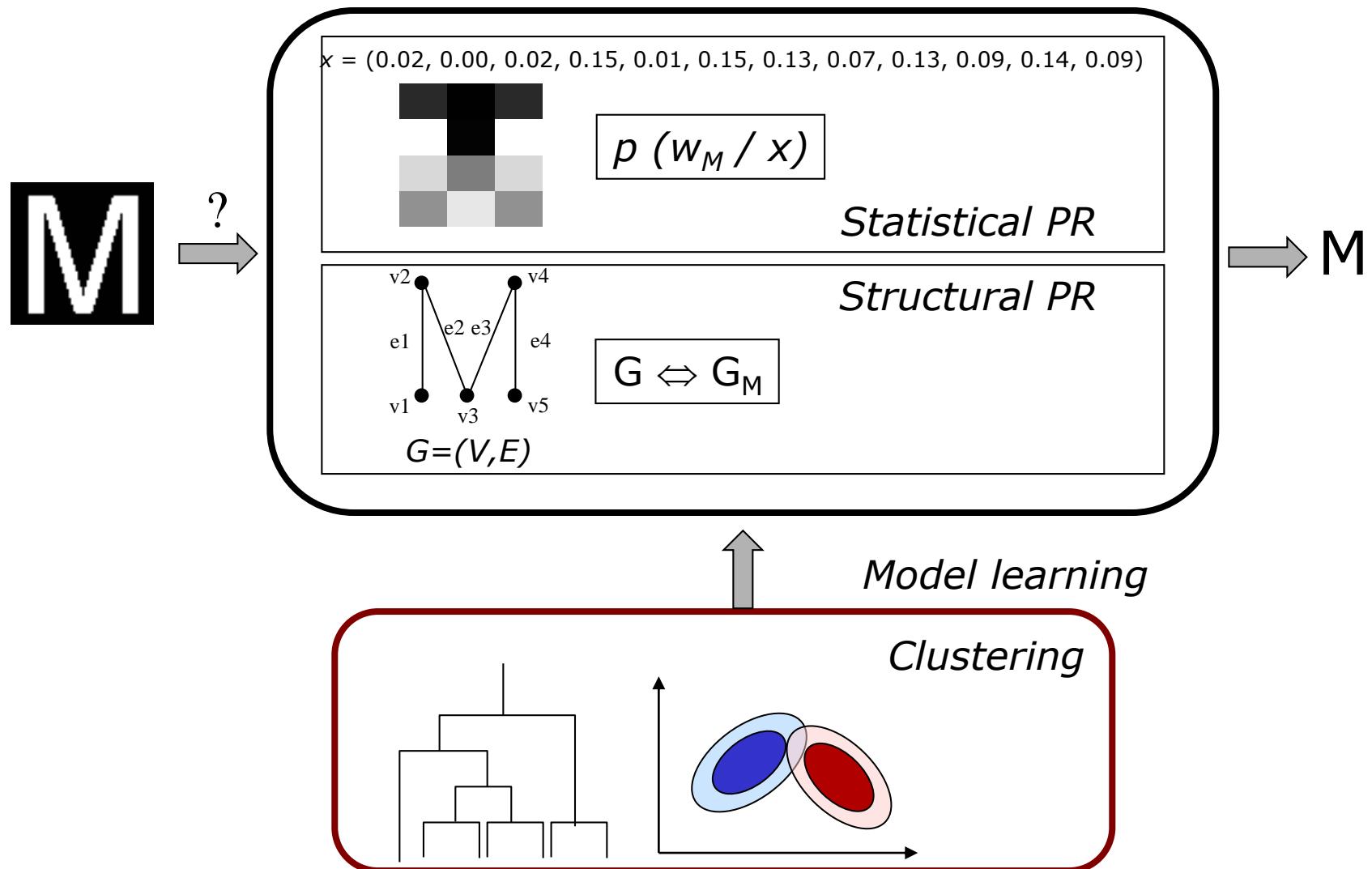


Quality control

Optical Character Recognition

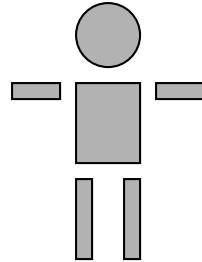


The Pattern Recognition framework

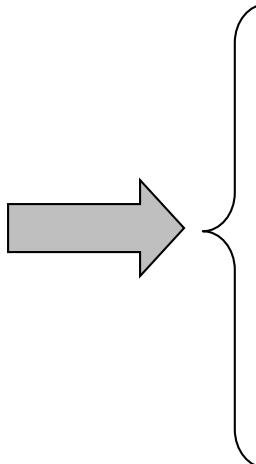


Statistical PR vs. Structural PR

- **Statistical PR** a class of patterns is described by a *vector of numerical features*.
The similarity between two shapes is formulated in terms of a distance (metric) defined on the feature space.
- **Structural PR** it is based on the explicit or implicit representation of the structure of a class, where the structure means the relational and hierarchical organization of low level features or primitives into higher level structures.



Pattern P



Statistical description: $P = (\# \text{components}, \text{height}, \text{width})$

Structural description:

Rleg, Lleg: VERTICAL RECTANGLE;

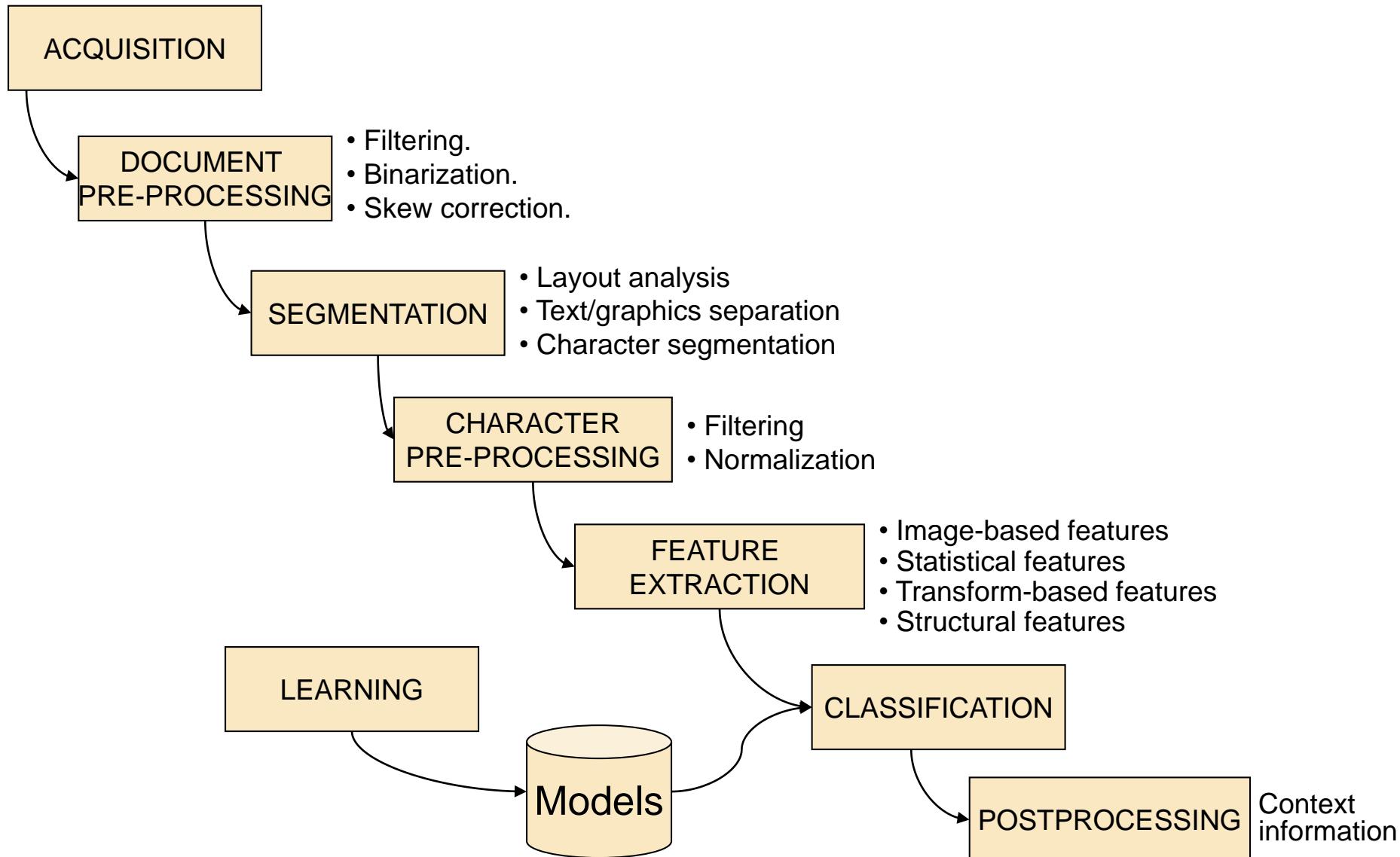
Rarm, Larm: HORIZONTAL RECTANGLE;

body: SQUARE;

head: CIRCLE;

$P = \text{head} \uparrow (\text{Larm} \Leftrightarrow (\text{body} \uparrow (\text{Lleg} \Leftrightarrow \text{Rleg})) \Leftrightarrow \text{Rarm})$

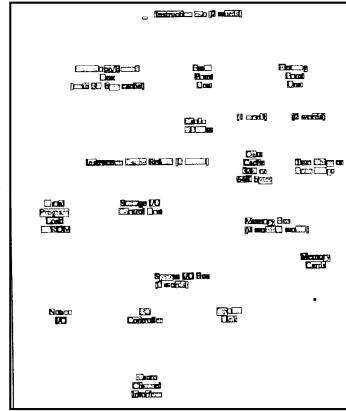
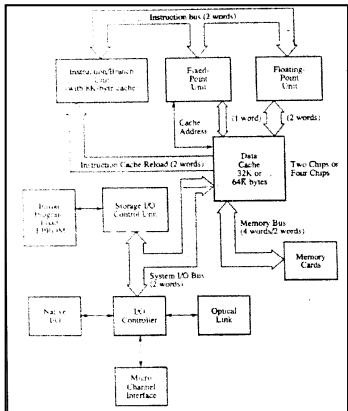
Components of an OCR system



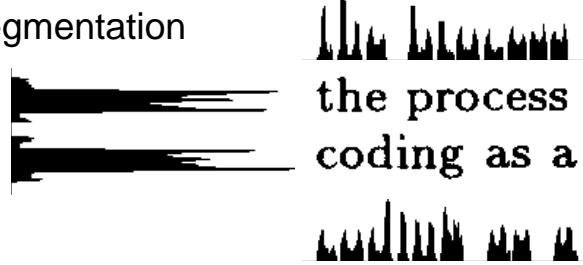
Optical Character Recognition

An OCR recognizes character by character a printed text in a digital image and transforms it to an ascii format.

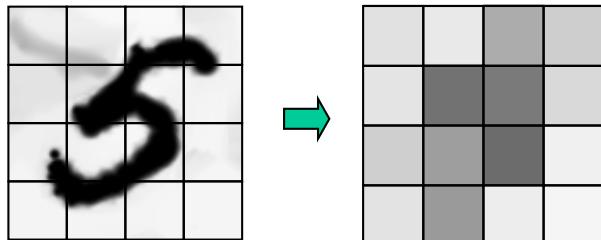
Text detection



Text segmentation



Feature extraction



S
X

Classification

Classifier

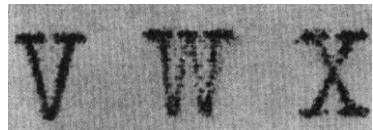
?

C17='s', c21='S', c4='g', c9='8'

Levels of difficulty in character recognition

Level 1

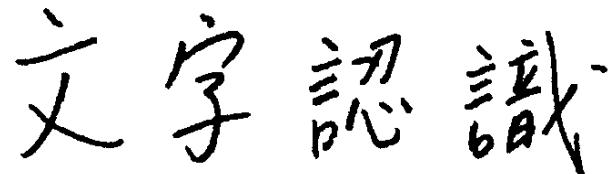
1.0. Printed characters of multiple fonts



1.1. Loosely constrained hand-printed characters

1→1 2→2 3→3
4→4 5→5 6→6
7→7 8→8 9→9

1.2. Chinese characters of few fonts.



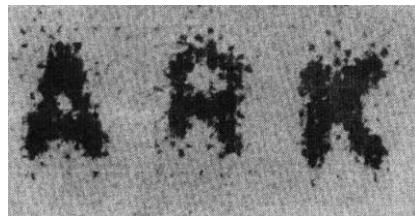
1.3. Loosely constrained hand-printed characters. N° characters ≈ 1000

→ **Solved problem**

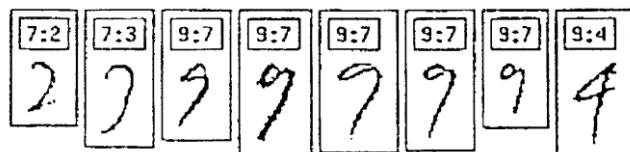
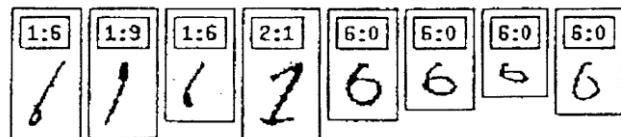
Levels of difficulty in character recognition

Level 2

2.0. Printed characters of multiple fonts



2.1. Unconstrained hand-printed characters



2.2. Affine transformed characters

Office of the Dean

Levels of difficulty in character recognition

Level 3

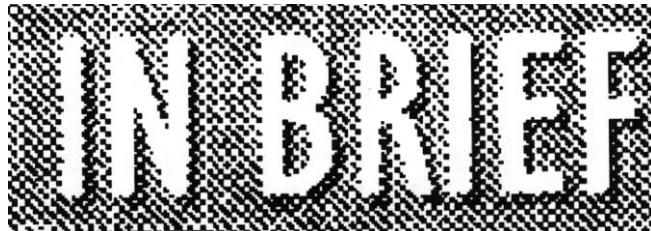
3.0. Unsegmented characters

CONTENTS A step size

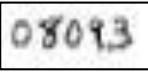
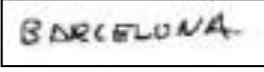
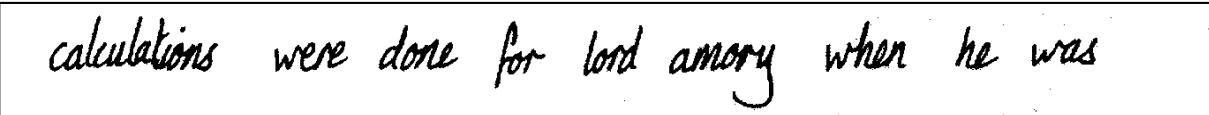
3.1. Cursive handwriting characters

calculations were done for lord amory when he was
chancellor on this basis and the conclusion was reached
that a uniform sales tax over the widest possible range of

3.2. Characters on a textured background

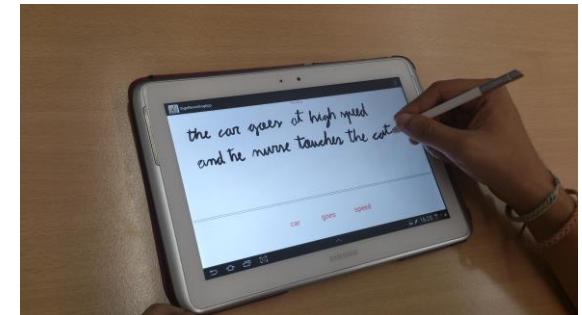


Off-line hand-written character recognition

- Hand-written
- Off-line: acquisition by a scanner or a camera
- Challenges:
 - Shape variability among images of the same character
 - Character segmentation
- Subproblems:
 - Hand-written numeral recognition: digit recognition

 - Hand-printed character recognition: well-separated characters

 - Cursive character recognition: non-separated characters


On-line hand-written character recognition

- On-line acquisition
 - Digitizer tablets
 - Digital Pen
 - Tablets
- Advantages with respect to off-line acquisition:
 - Image is acquired while the text is written
 - We can take advantage of dynamic information:
 - Temporal information: writing order, stroke segmentation, etc
 - Writing speed
 - Pen pressure
- Subproblems:
 - Cursive script recognition.
 - Signature verification/recognition.



HANDWRITTEN TEXT RECOGNITION (ISOLATED CHARACTERS/WORDS)

Difficulties

Difficulties:

- Different Handwriting styles
- Different writing instruments
- Large vocabulary
- Segmentation problem
chicken-egg

lum des

leendert

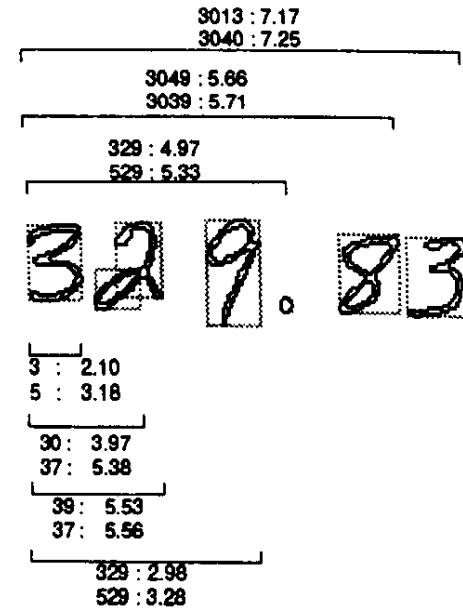
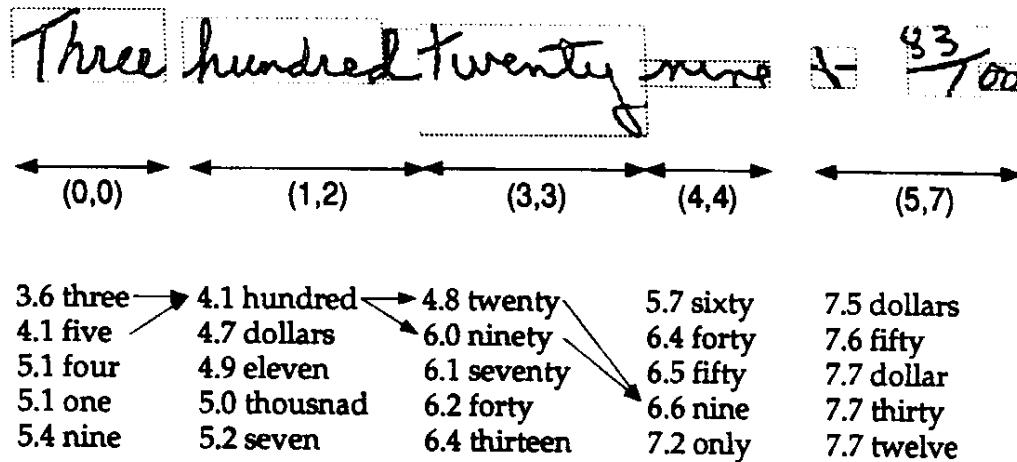
Lumdes



uuuuuuuu

Handwritten documents

- Use of contextual/grammatical models and dictionaries.
- High performance in controlled domains: c.f. bank check
- Parallel use: writer identification (biometrics).



Isolated Handwritten Character/Digits Recognition

Similar techniques than the ones for Shape Recognition

MNIST dataset and comparison methods:

<http://yann.lecun.com/exdb/mnist/>

Many different approaches:

- k-NN
- SVM
- Neural Networks (Deep)
- etc.



Current state of the art performance: over 99%

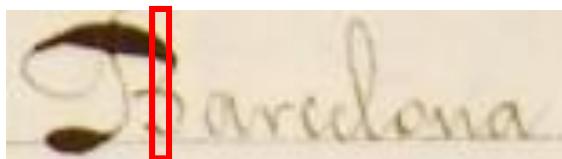
Word Recognition by sequence alignment

FEATURES:

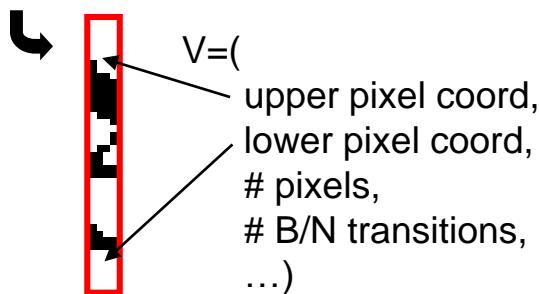
Profiles (Upper and Lower)



Pixel-based features



$$V_1, V_2, \dots, V_n$$

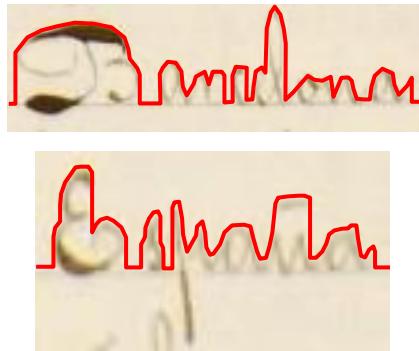


Word Recognition by sequence alignment

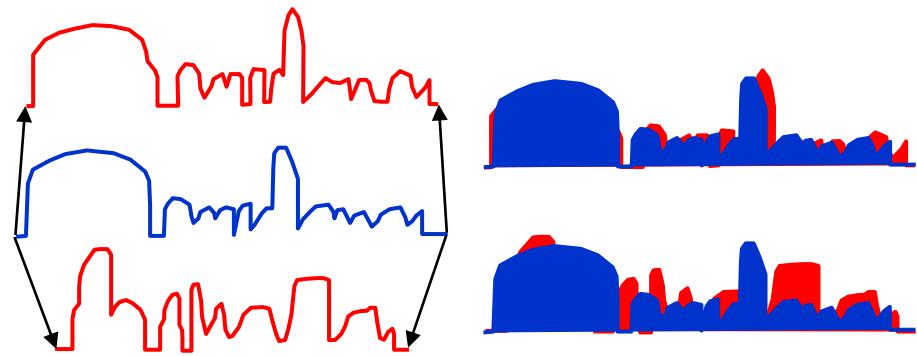
Unknown



Database (dictionary)



Dynamic Time Warping
(sequence alignment)



HANDWRITTEN TEXT RECOGNITION (TEXT LINES)

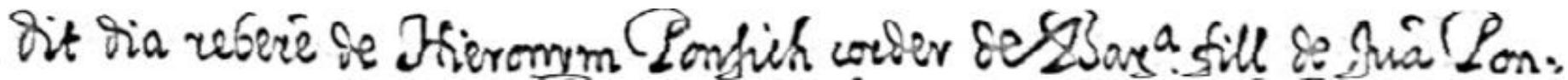
Text Line Recognition (Handwriting Recognition)

Recognition / Transcription

- Hidden Markov Models
- Neural Networks

No segmentation into words

Image:

A single line of handwritten text in black ink on a white background. The text is written in a fluid, cursive script. It appears to be in Spanish and reads "dit dia rebere\$ de Hieronym Ponsich corder de Bar^(a). fill de Jua\$ Pon=". There are some minor variations in letter height and stroke thickness throughout the line.

Text:

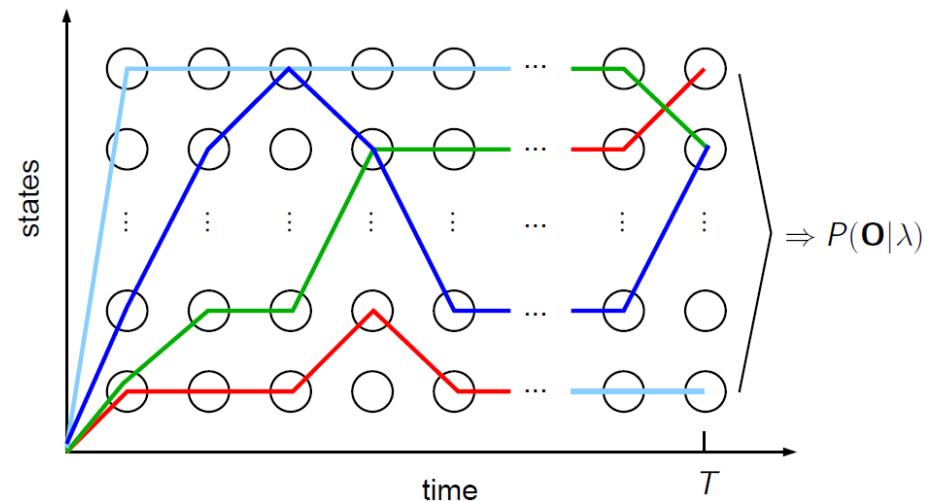
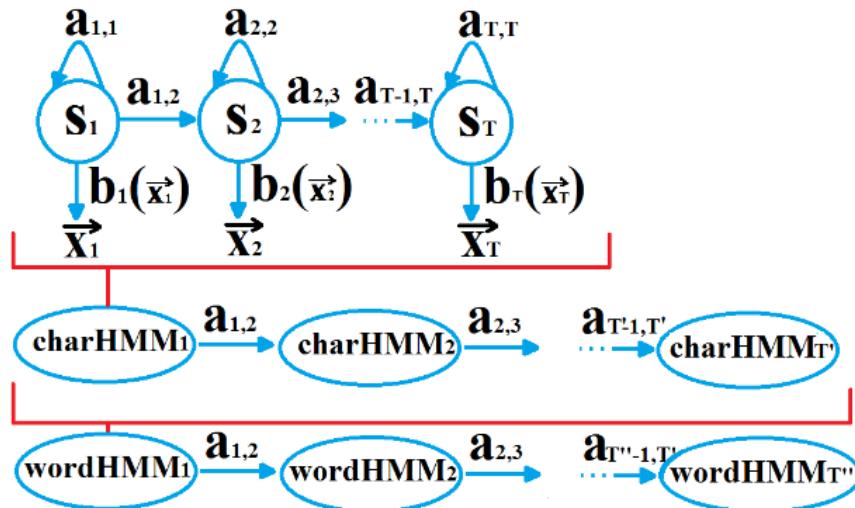
dit dia rebere\$ de Hieronym Ponsich corder de Bar^(a). fill de Jua\$ Pon=

Handwriting Recognition: HMM

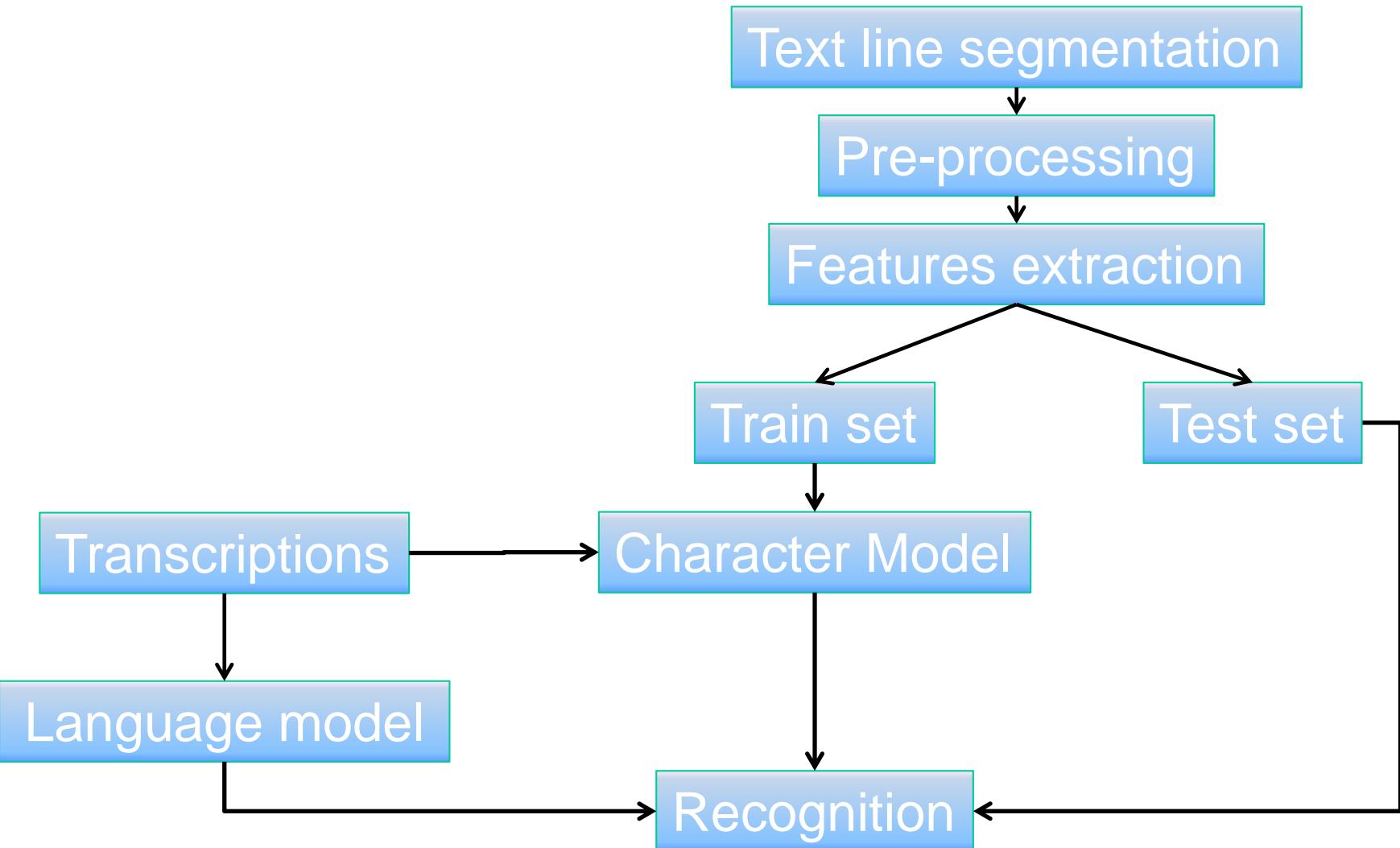
Standard HMM + n-grams

Vinciarelli et al., "Offline recognition of unconstrained handwritten texts using hmms and statistical language models", PAMI 2004

- HMM
- Word n-grams + dictionary

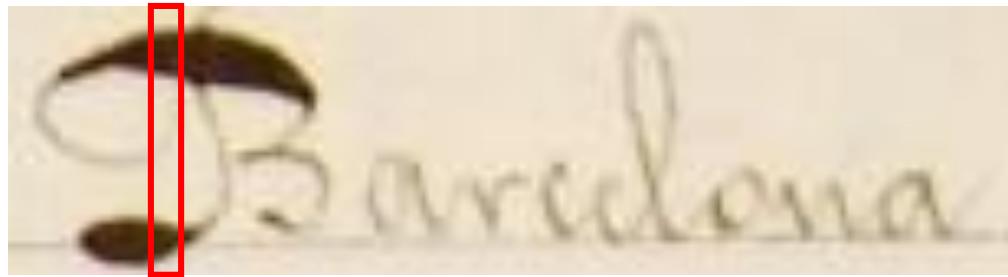


Handwriting Recognition System: Overview



Hidden Markov Models (HMM)

A HMM is an **statistical** model that represents the most probable sequence of words/characters for a given sequence of observations (features = column of pixel values).



Two aspects of modeling:

- Script (appearance) model: Features to represent words/characters.
- Language model: Restrictions for sequences of words/characters.

Text line normalization steps

- input the measurement of temperatures. This,
- skew the measurement of temperatures. This,
- slant the measurement of ~~temperatures~~. This,
- height the measurement of temperatures. This,
- width the measurement of temperatures. This,
- output the measurement of temperatures. This,

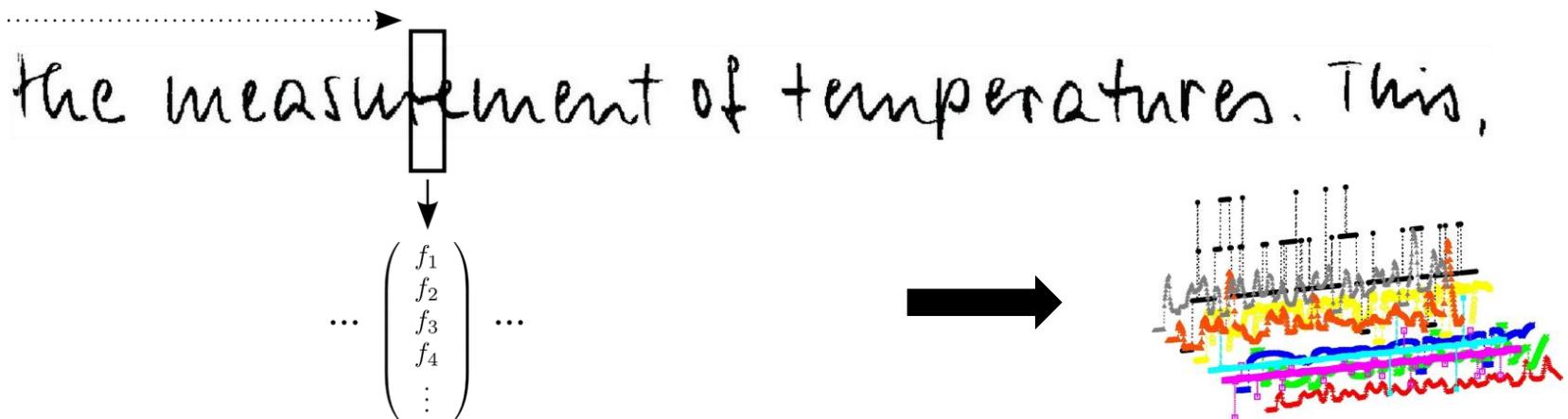
Feature Extraction via Sliding Window

Approach:

- Extract features at each position
- Concatenate vectors into a sequence

Here, we will review the features used in the presented examples:

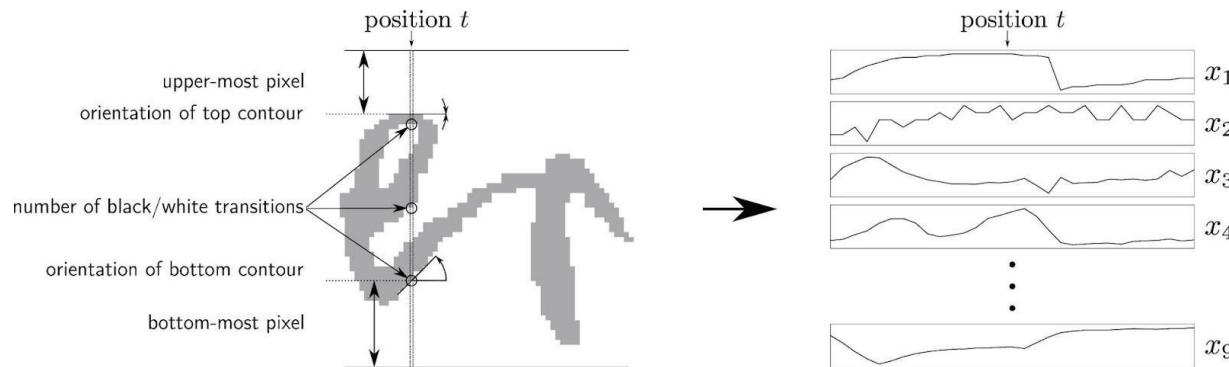
- Geometric features
- Gradient features



Geometric Features

A combination of 6 local and 9 global features.

- Window width = 1px
 - position and orientation of upper-most pixel and bottom-most pixel
 - 0th, 1st, and 2nd momentum
 - number of black/white transition
 - fraction of black pixels between highest and lowest black pixel



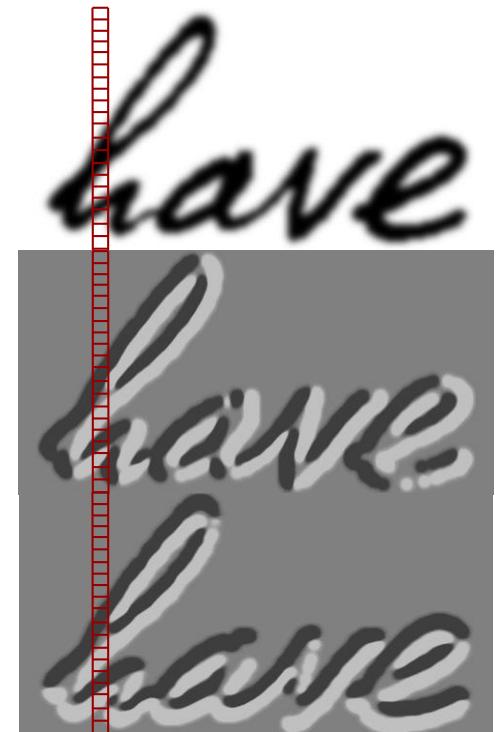
Marti, U.V. et al., Using a Statistical Language Model to Improve the Performance of an HMM-Based Cursive Handwriting Recognition System, Int'l J. of Pattern Recognition and Artificial Intelligence, 15, pp. 65–90, 2001

Grey-level and Derivative Features

Blur the image. Then split the sliding window into 1x20 cells. Use

- The grey level value
- The horizontal gradient
- The vertical gradient

60 dimensional vector



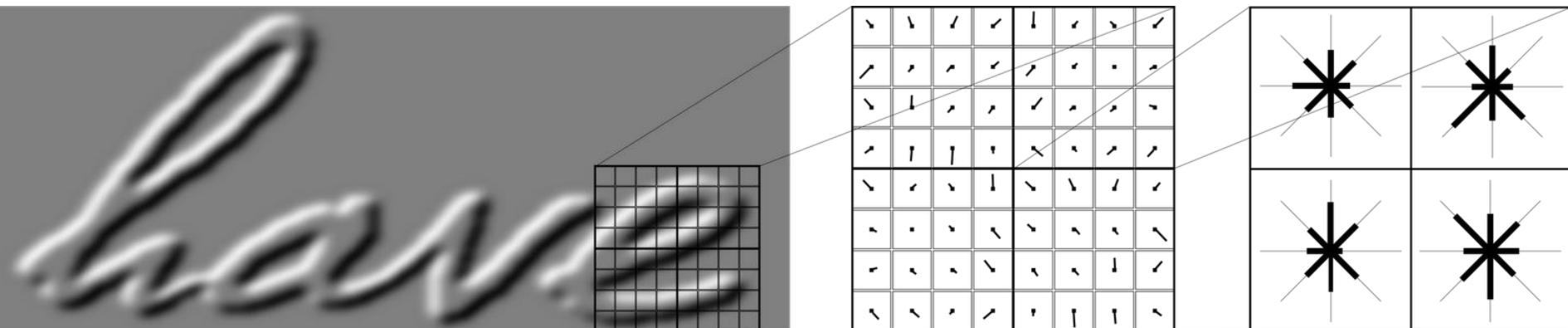
Toselli, A.H. et al, Integrated Handwriting Recognition and Interpretation using Finite-State Models, International Journal of Pattern Recognition and Artificial Intelligence, 18(4), pp. 519-539, 2004

Local Gradient Histogram Features

4x4 Gradient histograms (each 8 directions), i.e. 128 dimensional feature vector

Preprocessing:

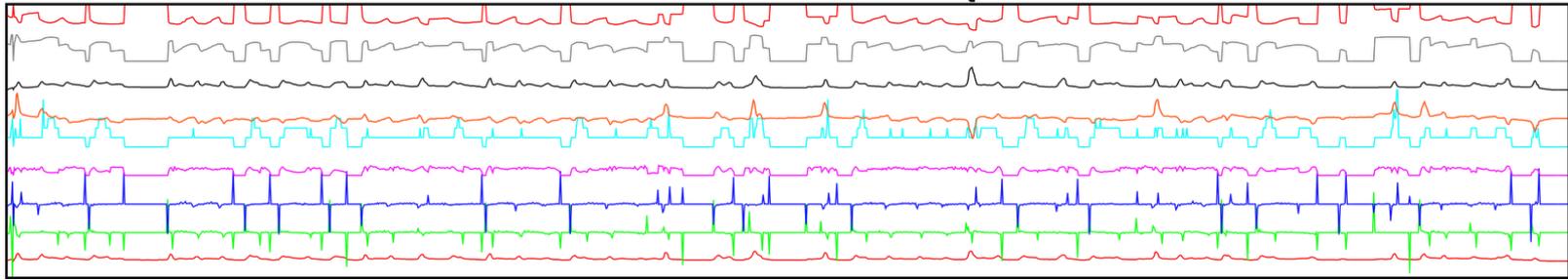
- Apply Gaussian filter to image
- Obtain horizontal and vertical gradients G_x and G_y
- Compute magnitude and direction of $(G_x, G_y)^T$



Rodríguez, J.A. et. al., Local Gradient Histogram Features for Word Spotting in Unconstrained Handwritten Documents, International Conf. on Frontiers in Handwriting Recognition, 2008

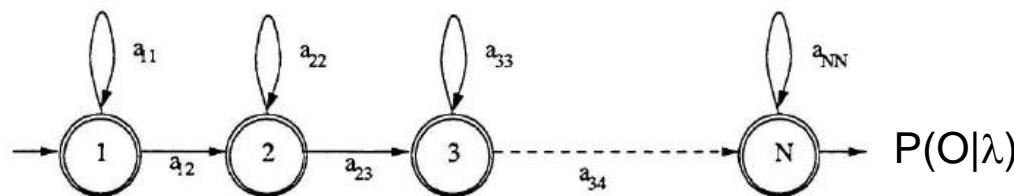
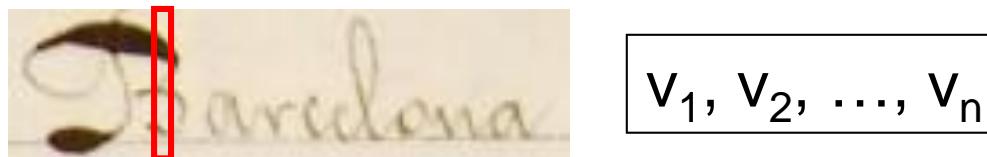
Sequential Representation of a Text Line

the measurement of temperatures. This,



Hidden Markov Model (HMM)

- Hidden Markov Models
 - A HMM is similar to a finite state machine that represents the sequence of words/characters which most probable for a given sequence of observations (features).



- Two aspects of modeling:
 - Script (appearance) model: Features to represent words/characters.
 - Language model: Restrictions for sequences of words/characters.

Hidden Markov Models Setup

HMM = (S,A,B,P,V) with

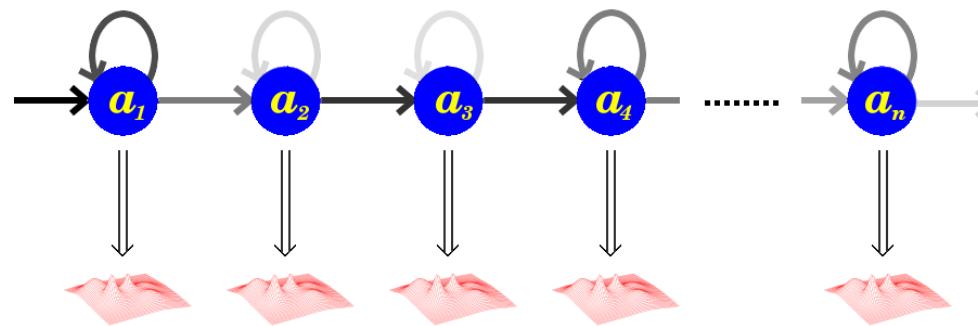
S = A set of discrete states $\{a_1, a_2, \dots\}$

A = Transition probability matrix (A_{ij})

B = Emission probability distributions $\{b_1, b_2, \dots\}$

P = Starting probabilities for each state

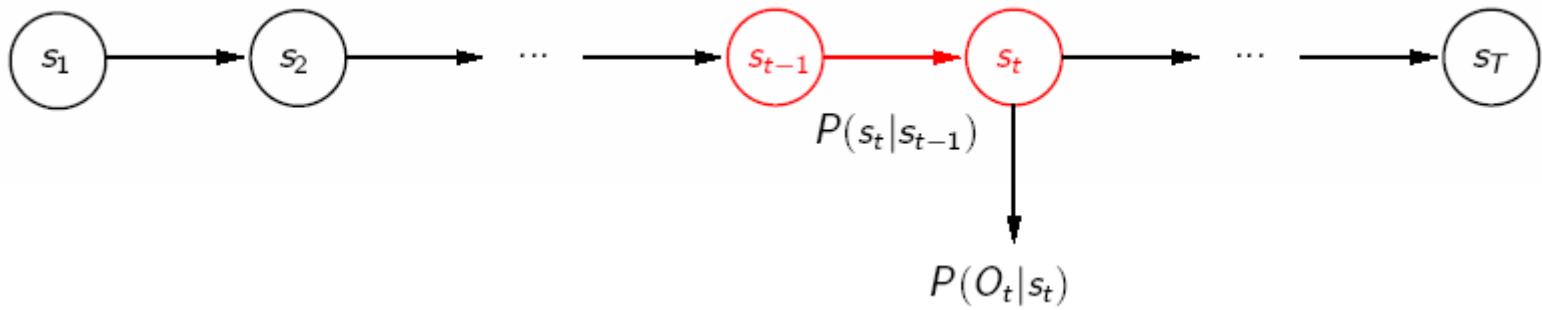
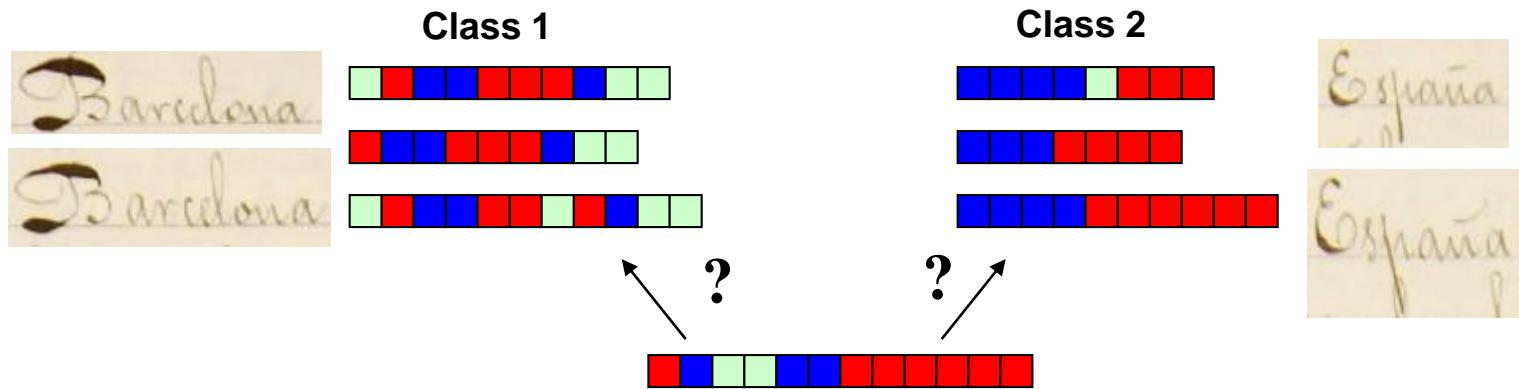
V = The feature space of the observations



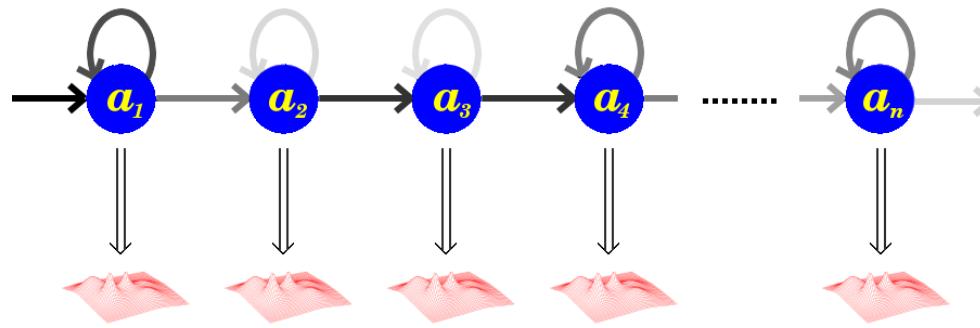
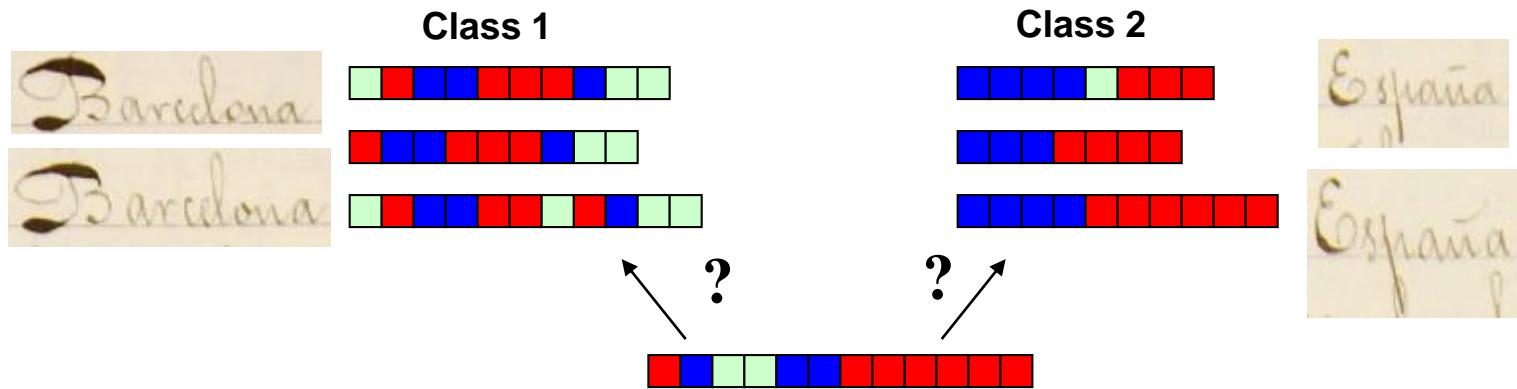
Hidden Markov Model (HMM)

- Hidden Markov Model (HMM) is a doubly stochastic process, with an underlying stochastic process that is not observable (hence the word hidden), but can be observed through another stochastic process that produces the sequence of observations (Rabiner, 1989).
- HMMs have been proven to be one of the most powerful tools for modeling speech and for real-world signals. These probabilistic models offer many desirable properties for modeling characters or words.
- One of the most important properties is the existence of efficient algorithms to automatically train the models without any need of labeling presegmented data.

Hidden Markov Models (HMM)

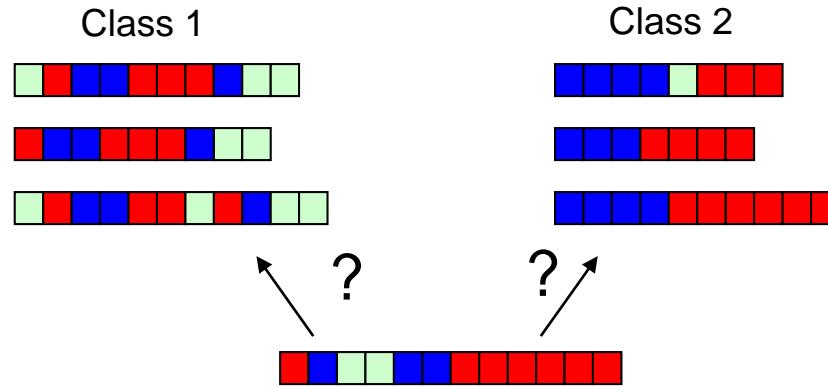


Hidden Markov Models



Hidden Markov Models

Pattern recognition: probability that a sample belongs to class i?

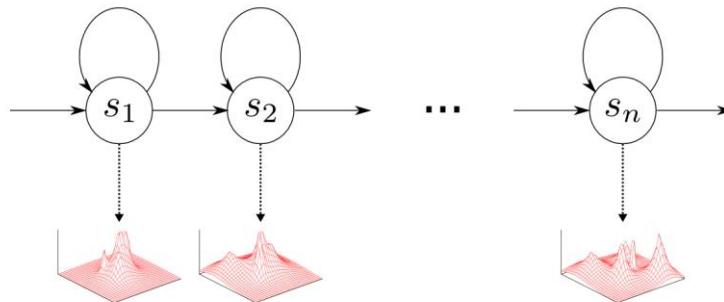


- Three algorithms:

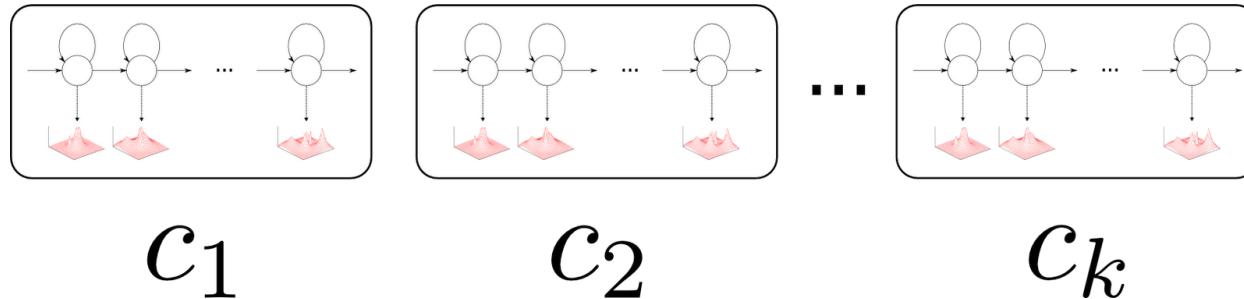
- $P(O | \lambda)$: Probability that a given observation sequence has been generated by the model ([Forward-backward algorithm](#))
- Which is the optimal state sequence? ([Viterbi algorithm](#))
- How must λ be chosen for each class? ([Baum-Welch algorithm](#))

Hidden Markov Models for words

- Character-models with linear topology and mixture of Gaussians

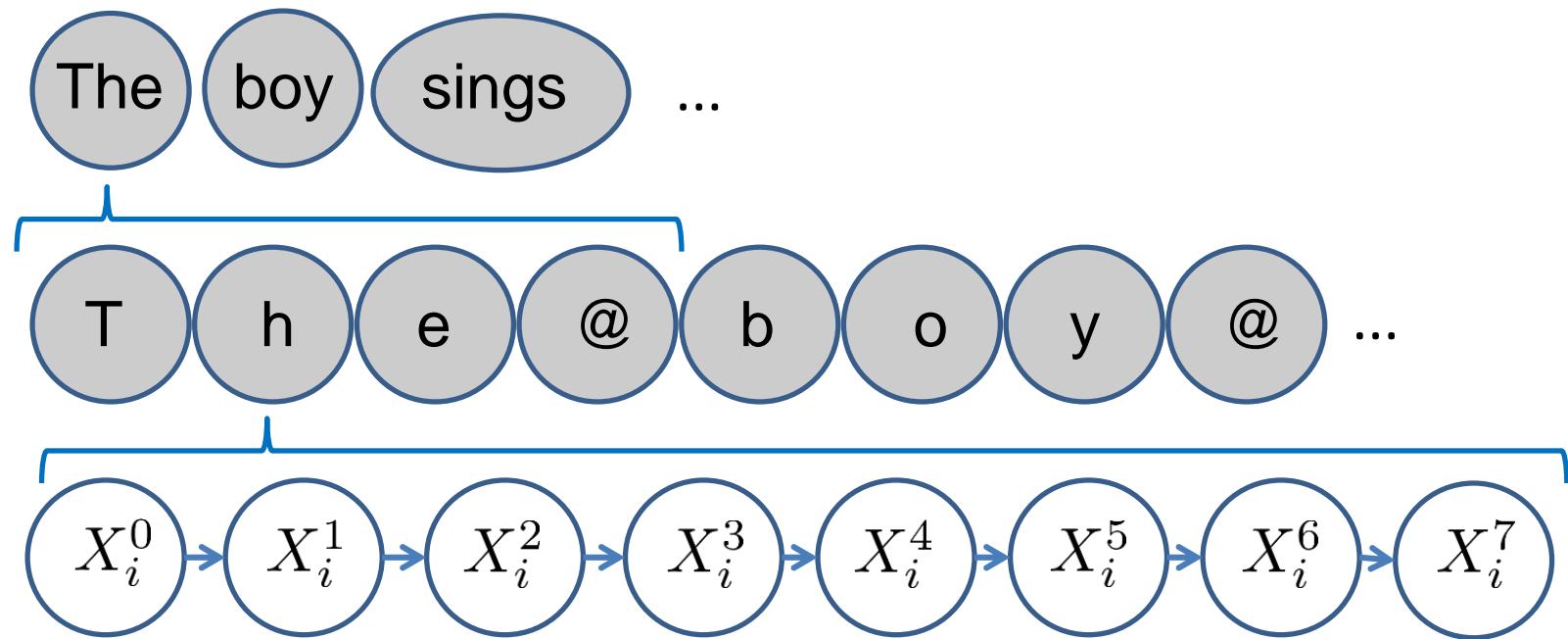


- Words are composed by concatenating according character models



Hidden Markov Models for text lines

Training of a HMM for **each** character:



X_i^0 = column of pixel values (feature descriptors)

HMM Topology?

- Number of States?
- Number of Gaussians per state?

m m m

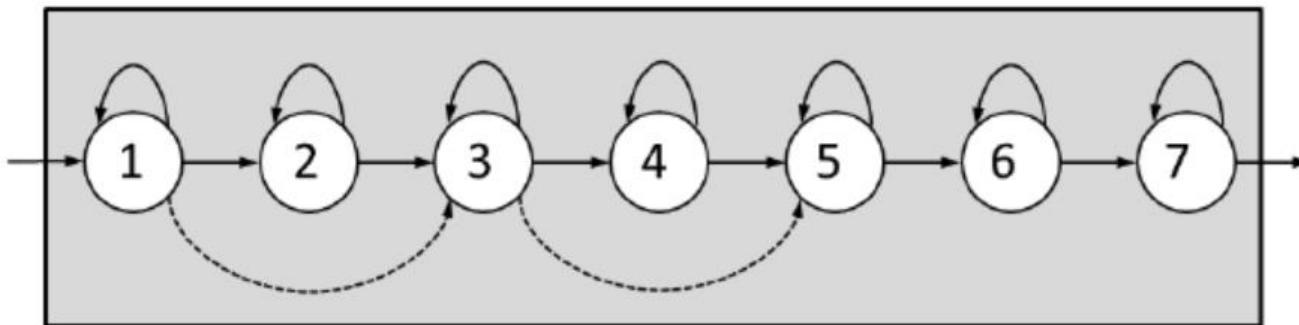
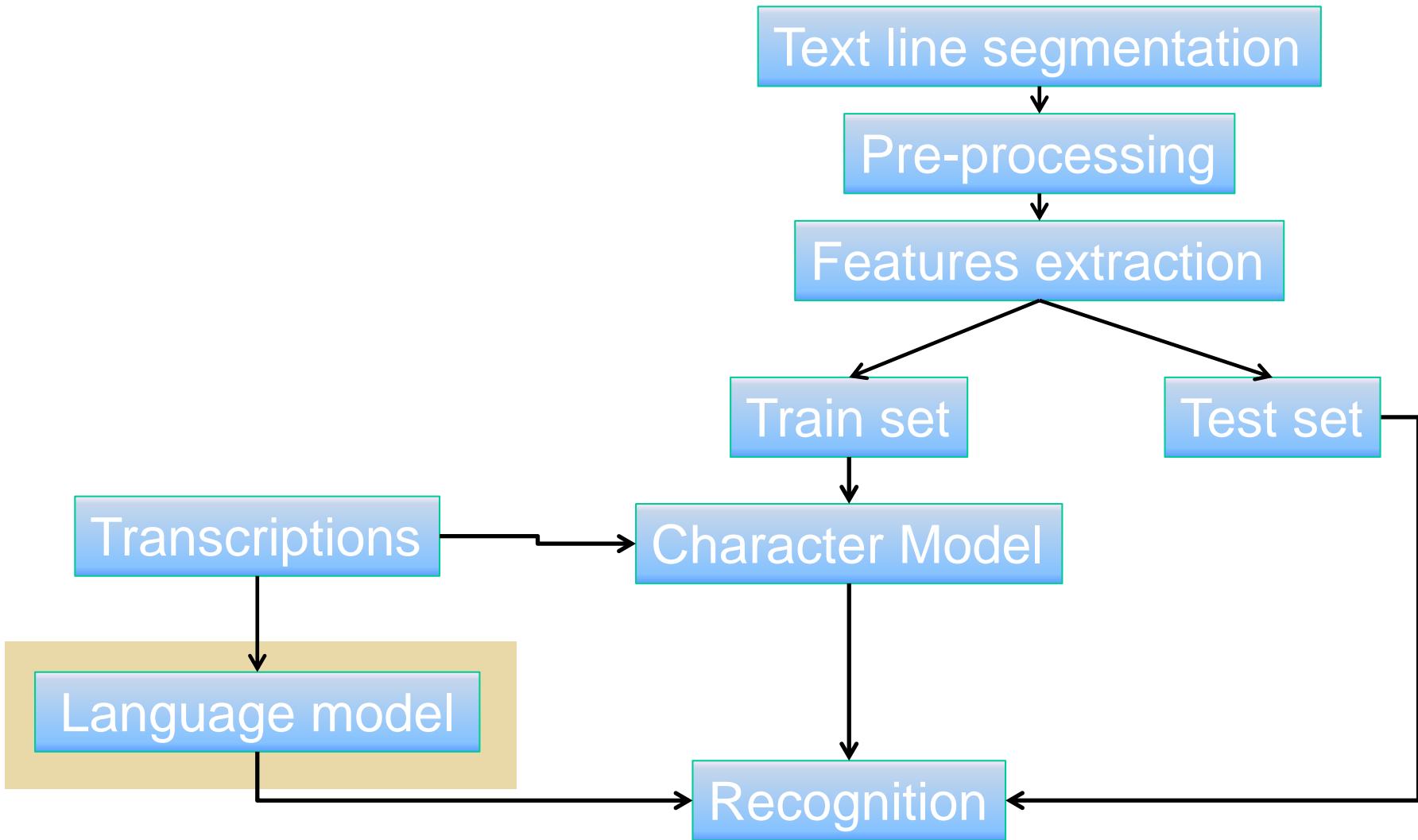


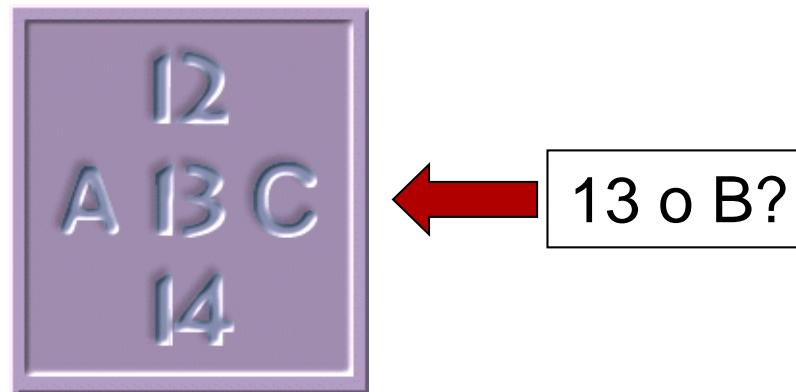
Fig. 2. Example of hidden Markov model with 7 emitting states connected from left to right with self-loops, and two skip-state transitions connecting state 1 to state 3, and state 3 to state 5. It allows a minimum number of 5 emissions. Following the notation, $NS_c=7$ and $NSkip=2$ sets $\min NE_c=5$.

N.Cirera, A.Fornés, J.Lladós. Hidden Markov Model Topology Optimization for Handwriting Recognition. International Conference on Document Analysis and Recognition (ICDAR), 2015.

Learning based: Overview



Context information



- Context analysis tries to correct errors produced by decisions taken in function of local features
- In the presence of uncertainty, hypothesis generated by the local classifier are complemented with hypothesis of neighboring characters
- Two points of view
 - Geometric context (typographic).
 - Linguistic context

The Importance of the Context

When I was in high school, my physics teacher - whose name was Mr. Bader - called me down one day after physics class and said, "You look bored; I want to tell you something interesting." Then he told me something which I found fascinating, and since then, always found fascinating... The subject ~~was~~ is this - the principle of least action.

The Importance of the Context

When I was in high school, my physics teacher - whose name was Mr. Bader - called me down one day after physics class and said, "You look bored; I want to tell you something interesting." Then he told me something which I found fascinating, and have, since then, always found fascinating....
The subject ~~#~~ is this - the principle of least action.

When I was in high school, my physics teacher - whose name was Mr. Bader - called me down one day after physics class and said, "You look bored; I want to tell you something interesting." Then he told me something which I found fascinating, and have, since then, always found fascinating....
The subject # is this - the principle of least action.

Richard P. Feynman: The Feynman Lectures, Volume II.

The Importance of the Context

Hörðpiskheiði's horomban, egg var a fríkkatavorn - Bárður úrvald
hinnið - megar hefur hinnið fríkkatavorninum í sér mæltu: „Hinnið
var fátt; næstaður mundarí nekk valam. Óndverð. " Majst
elmarðott nekkur valamit, en it ekki vildið talið um, es an-
sta is mindig ekki vildið talið um... Í leggilell hattus
elmarðið van mi.

The Importance of the Context

Középiskolás koromban, egy nap a fizikatanárom - Bader úrnak írták - magához hívott fizikaóra után és azt mondta: „Unottnak látszol; szeretnék mondani neked valami érdekeset." Majd elmondott nekem valamit, amit elbűvölőnek találtam, és az óta is mindig elbűvölőnek találom ... A legkisebb hatás elvéről van szó.

Középiskolás koromban, egy nap a fi zikatanárom - Bader úrnak ívták - magához hívott fi zikaóra után és azt mondta: "Unott-nak látszol; szeretnék mondani neked valami érdekeset." Majd elmondott valamit, amit elbűvölőnek találtam, és az óta is mindig elbűvölőnek találom ... A legkisebb hatás elvéről van szó.

Richard P. Feynman: The Feynman Lectures, Volume II.

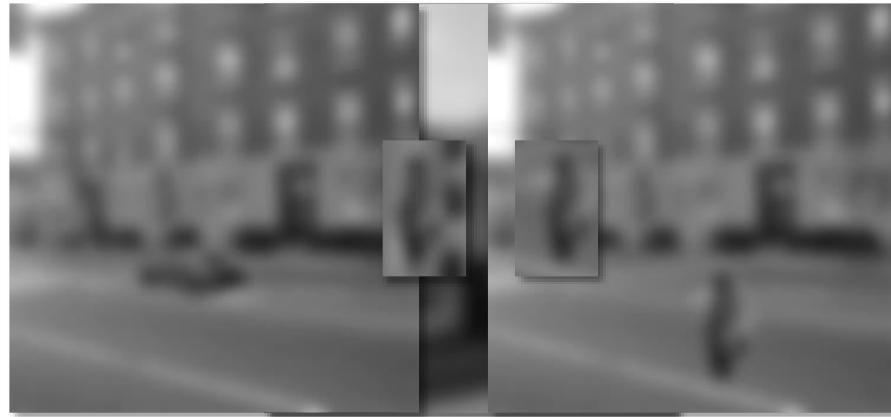
The role of the context in visual object recognition [*]

Over the last two decades much progress has been made in **object detection** using machine learning techniques.

Usually recognition is built based solely on the statistics of **local terms**.

Contextual information can provide more relevant information for the recognition of an object than intrinsic object information.

The **context** of an object in the scene can be defined in terms of other recognized objects and their mutual dependencies.



[*] Antonio Torralba. 2003. Contextual Priming for Object Detection. Int. J. Comput. Vision 53, 2 (July 2003), 169-191

Context and language models

Humans use **context information**

Vocabulary, Grammar-Syntax, Semantics, etc.

Computers use **language models**:

- regular grammars
- n-grams

Language models

Regular grammar:

$$G := (S, P, T, N)$$

where

$$T := \{w_1, \dots, w_{C_T}\}$$

set of terminal symbols

$$N := \{cat_1, \dots, cat_{C_N}\}$$

set of non-terminal symbols

$$S \in N$$

starting symbol

$$P := \{cat \rightarrow \alpha_j | cat \in N\}$$

the set of production rules

n-grams:

- frequencies of words, eg: *Look, was, he, after, etc.*
- frequencies of pairs of words, eg: *Look after, he was, was he, etc.*
- In general, frequencies of groups of consecutive words

N-grams and Grammars

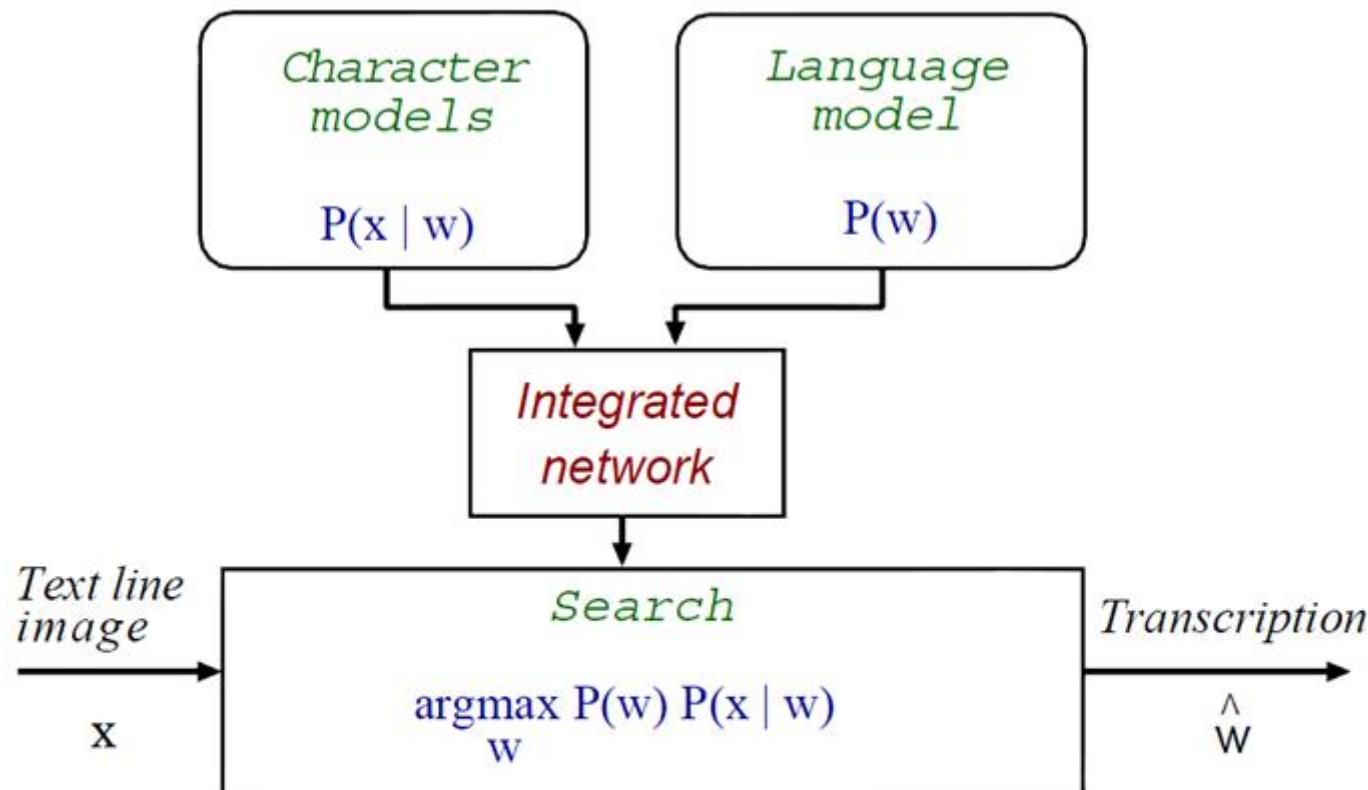
Methods based on **n-grams** (frequencies of combinations of n letters).

- Probability that an n-gram **appears** in the words of the **dictionary**
- characters with uncertainty, we take the final decision according to the n-gram with the **highest** probability

Methods based on grammars (natural language)

- used to **validate** the results of a character recognizer (OCR)
- permit to consider **variable-length** strings and recursivity

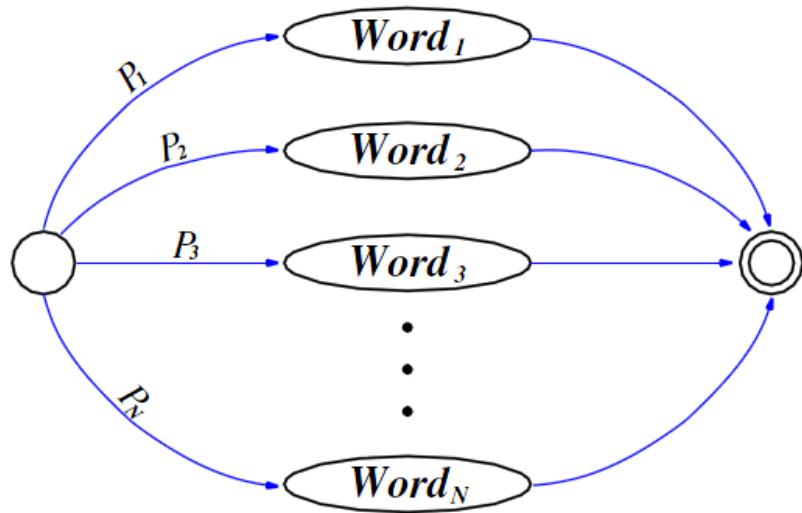
Handwriting Recognition System



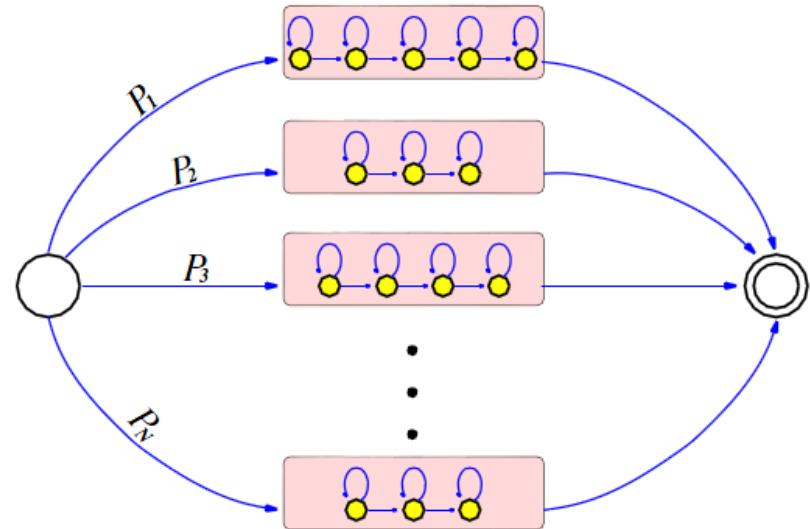
Search engine:
THE VITERBI ALGORITHM

Language model for isolated words

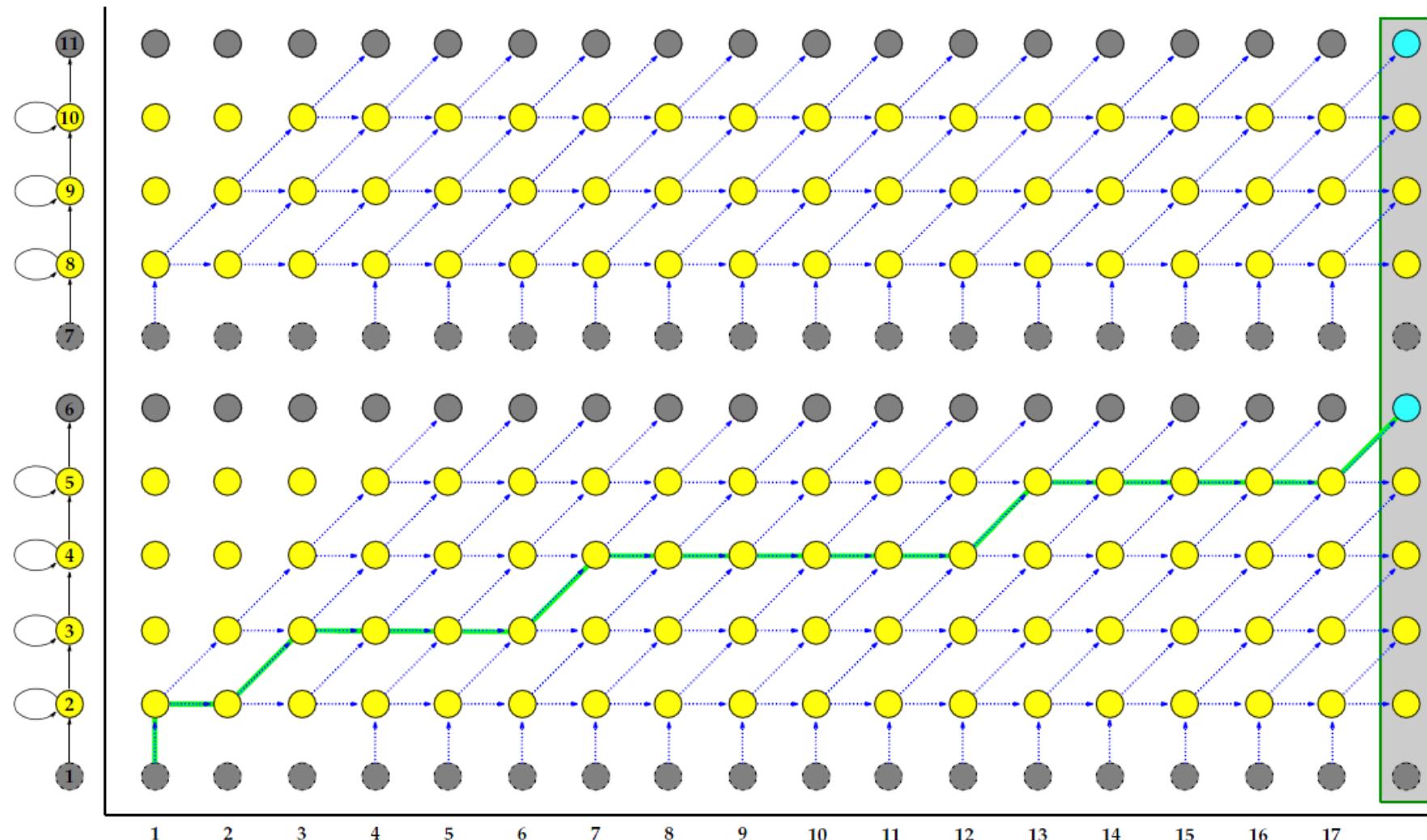
Isolated Words Language Model



Expanded Language Model

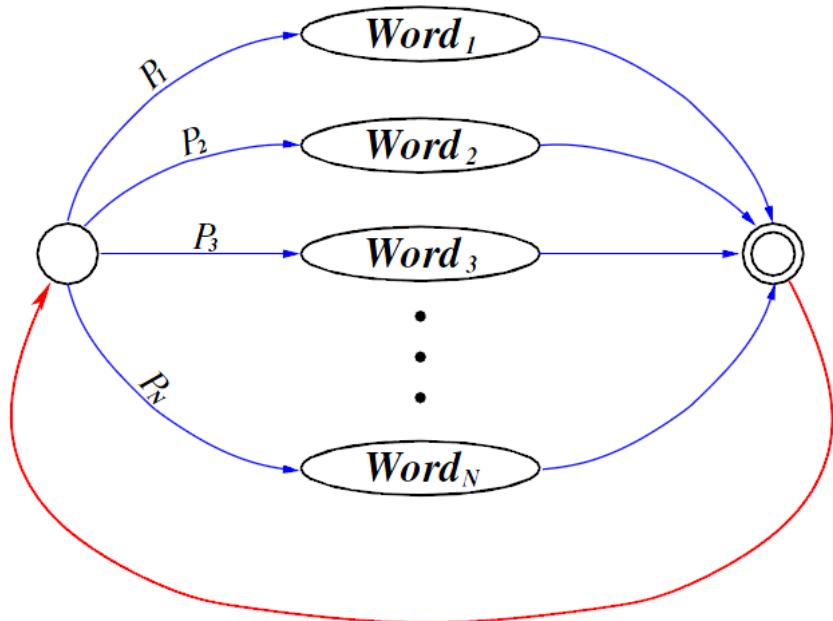


Searching Space for independent words

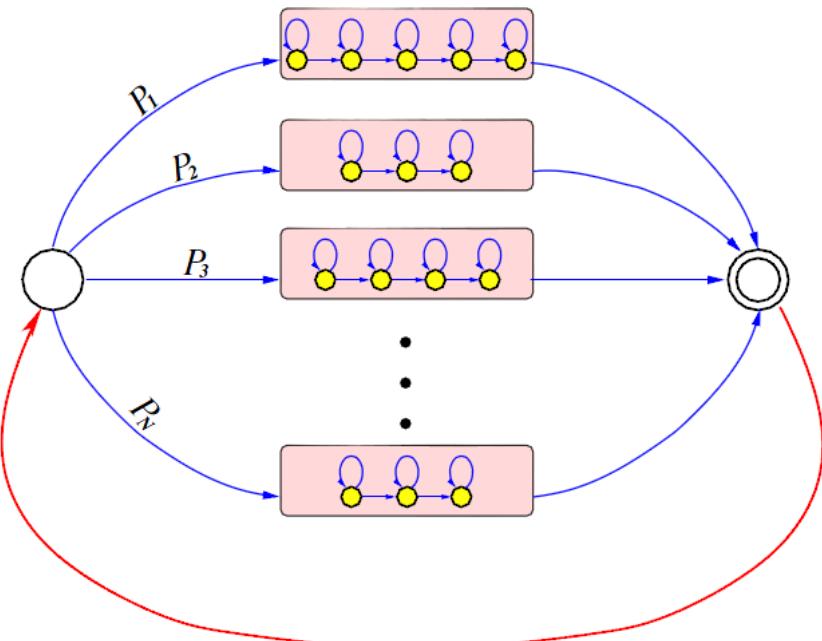


Language model for a sequence of words

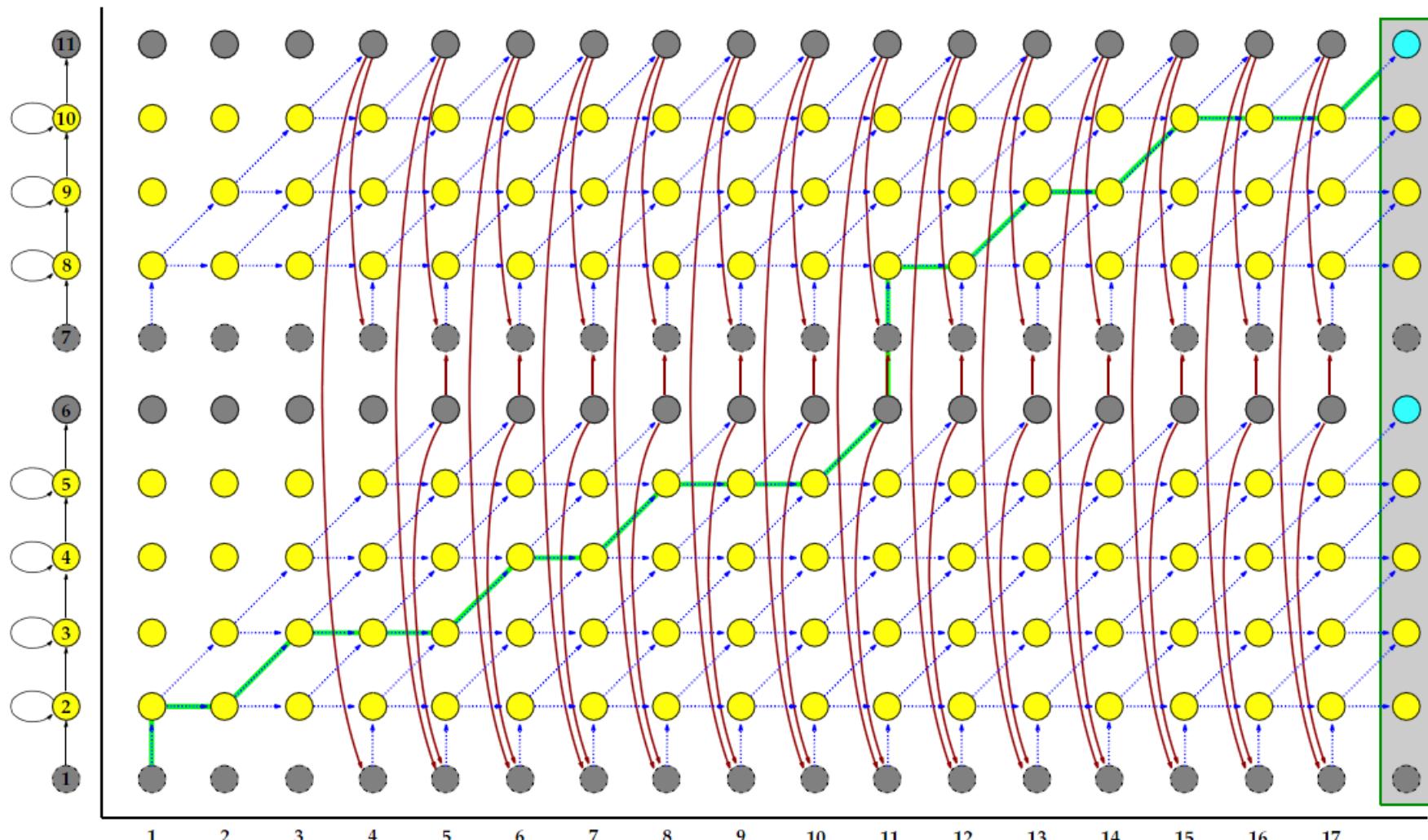
Words Language Model



Expanded Language Model

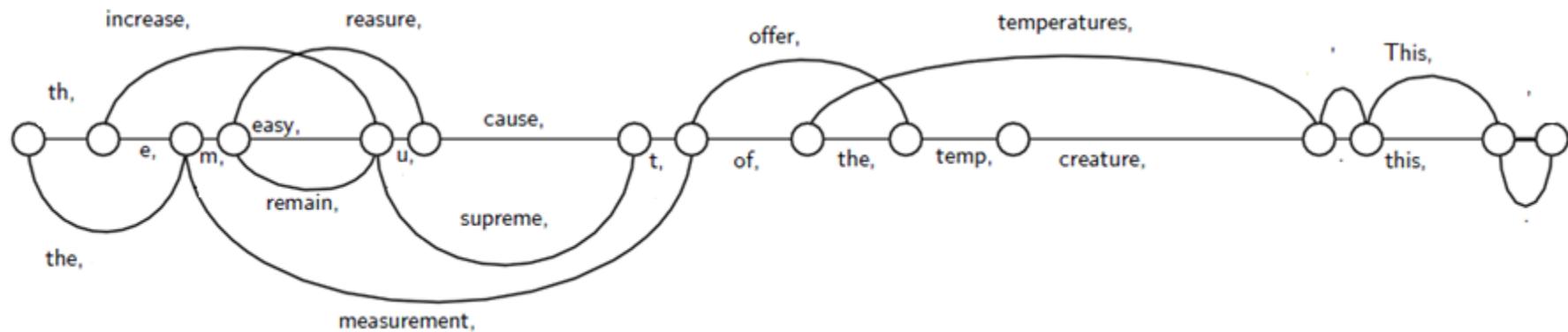


Searching Space for sequence of words



An example of a Recognition lattice

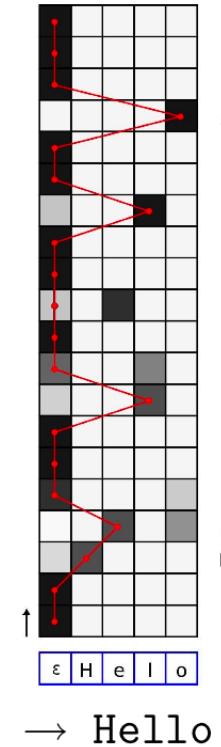
the measurement of temperatures. This,



Closed dictionaries

Contain **all** the possible words:

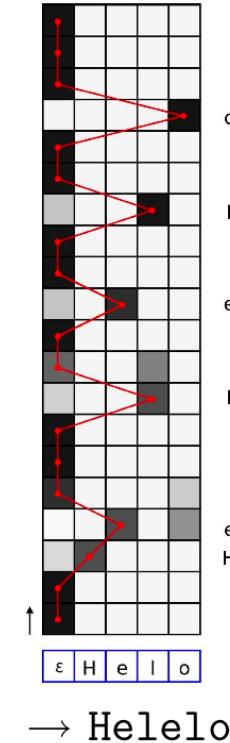
- ▶ Useful for applications where possible words are from a fixed size vocabulary
 - ▶ Numbers, city names, etc.
- ▶ error-tolerant
- ▶ slower



Open dictionaries

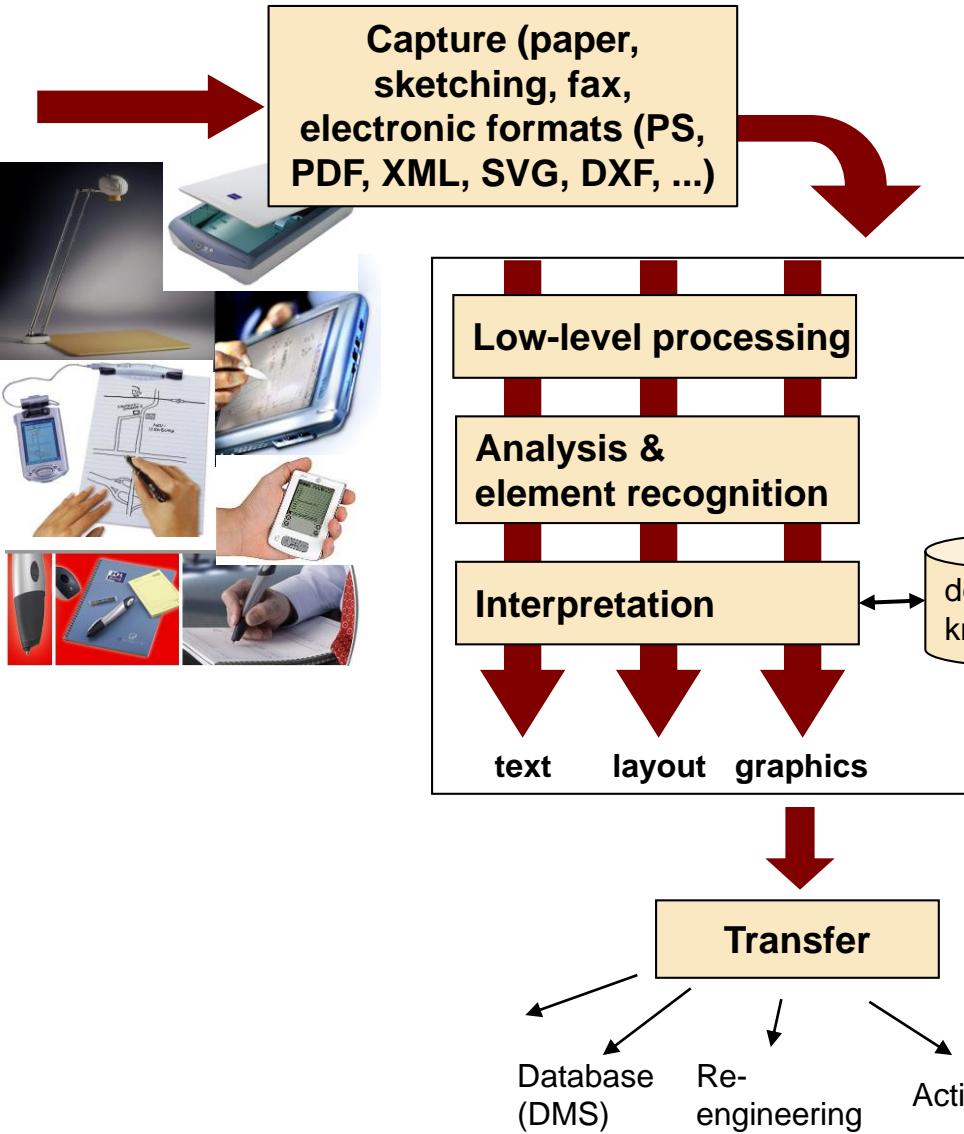
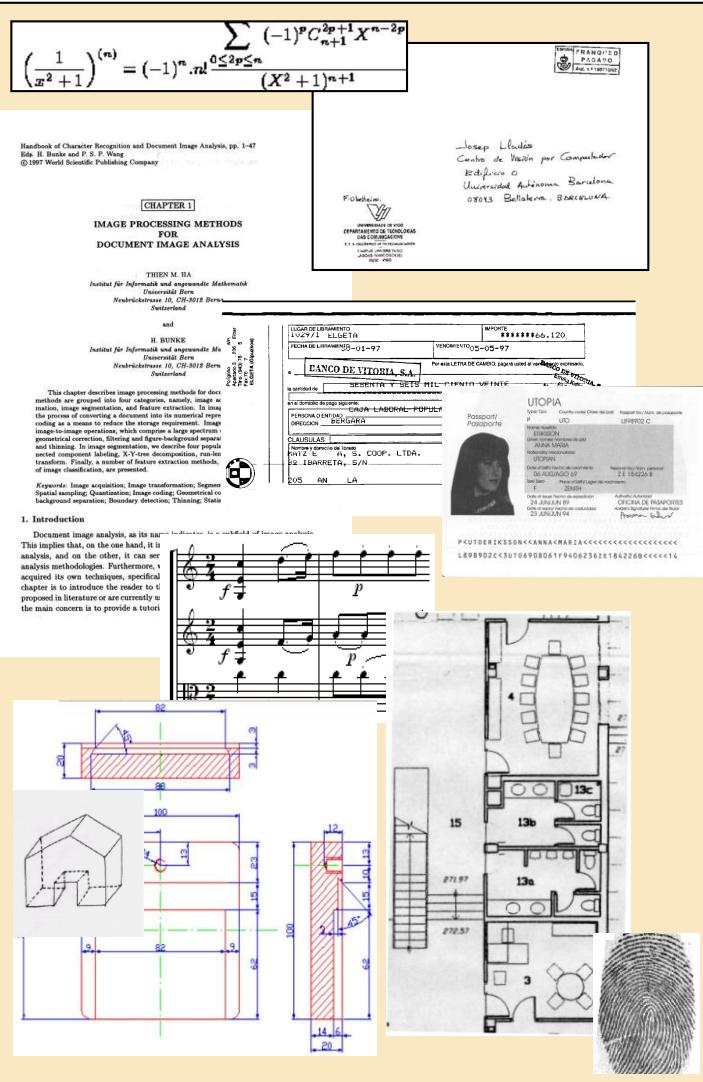
Do **no** contain **all** the possible words:

- ▶ Useful for applications where a dictionary might be too restrictive
 - ▶ Names, uncommon words, words in foreign languages, etc.
- ▶ Prone to error if text is not neatly written
- ▶ Very fast



INFORMATION EXTRACTION

Image Reading Systems in the document workflow

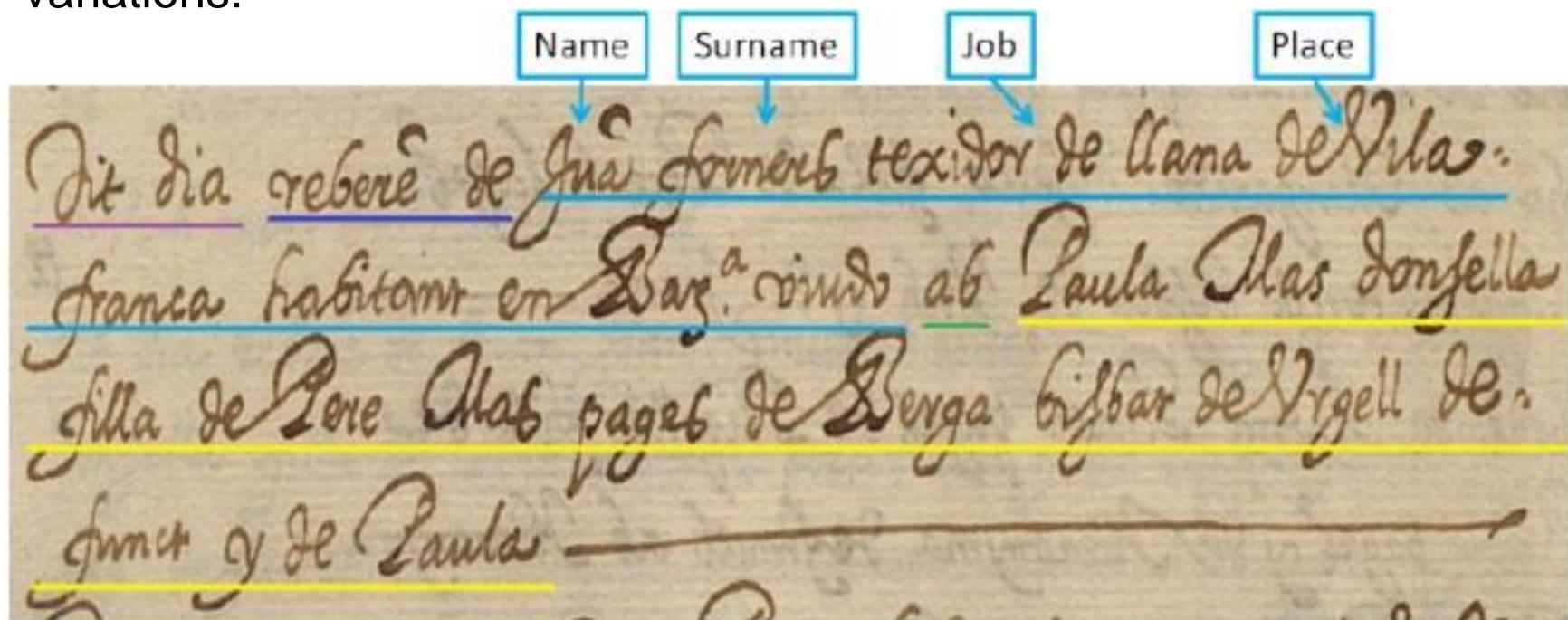


HW Recognition for Information Extraction

Objective: Information Extraction (fill a database)

Advantage:

document content share **similar syntactic structures**, but with variations:



Typical Language Models for HW

- N-grams:

$$P(c_k | c_1, \dots, c_{k-1}) = P(c_k | c_{k-n+1}, \dots, c_{k-1})$$

- Regular grammar:

$$G := (S, P, T, N)$$

where

$T := \{w_1, \dots, w_{C_T}\}$ set of terminal symbols

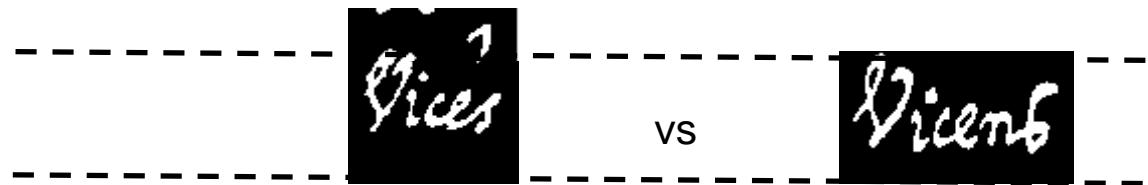
$N := \{cat_1, \dots, cat_{C_N}\}$ set of non-terminal symbols

$S \in N$ starting symbol

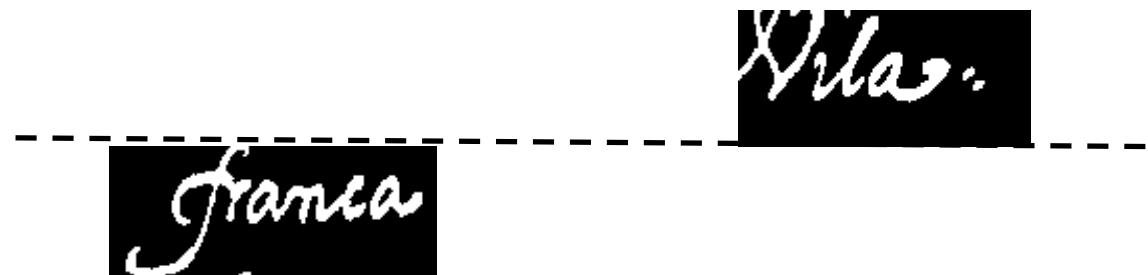
$P := \{cat \rightarrow \alpha_j | cat \in N\}$ the set of production rules

Out of Vocabulary Words??

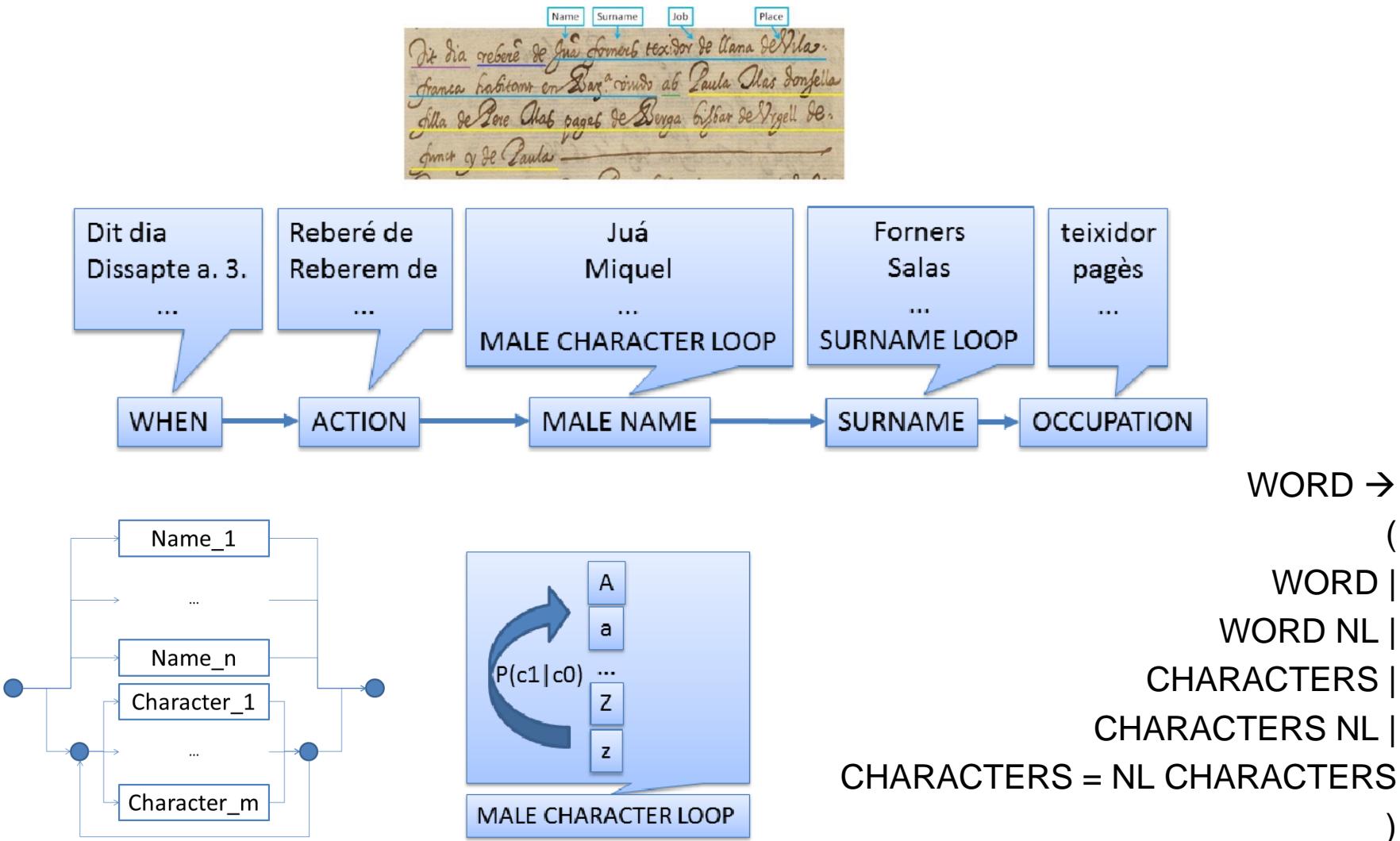
- Out of dictionary words (words not seen during training)
 - Immigration: from France, from Spain, ...
 - Misspellings:



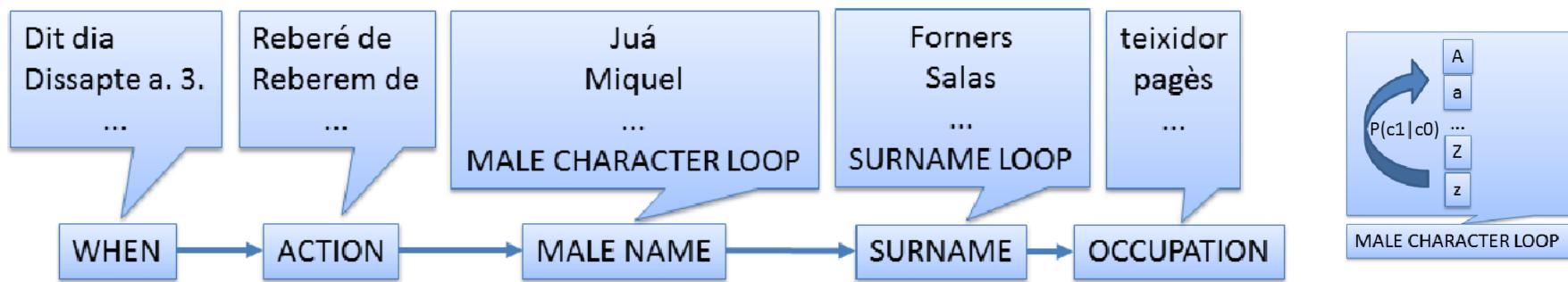
- Hyphenated words:



Hybrid method: Grammars + n-gram



Hybrid method: Grammars + n-gram



$$G' := (S, P', T', N)$$

where

$$\Gamma := \{c_i\}_1^{C_{T'}} \quad \text{set of characters}$$

$$T' := T \cup \Gamma \quad \text{set of terminal symbols}$$

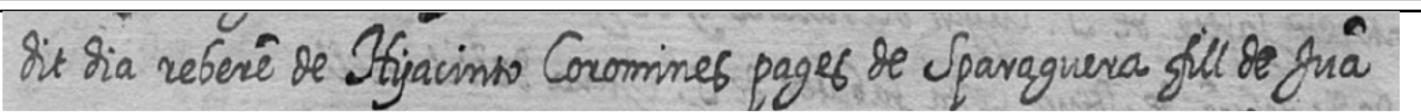
$$N := \{cat_1, \dots, cat_{C_N}\} \quad \text{set of non-terminal symbols}$$

$$S \in N \quad \text{starting symbol}$$

$$P' := P \cup \{cat_j \rightarrow c_i | c_i \in \Gamma, cat_j \in T\} \cup \text{set of production rules}$$

$$\{c_i \xrightarrow{P(c_j|c_i)} c_i c_j | c_i, c_j \in \Gamma\} \cup \{c_s \rightarrow cat_j | cat_j \in T\}$$

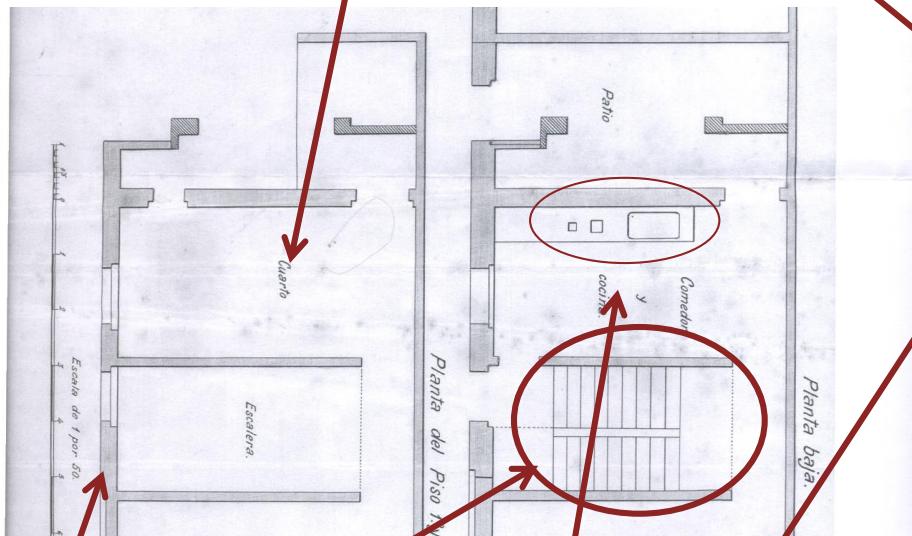
Results

Image	
Transcription	<i>dit dia rebere\$ de Hyacinto Coromines pages de Sparaguera fill de Jua\$</i>
Grammar	<i>dit dia rebere\$ de Hyacinto Coromines pages de Sparaguera fill de Jua</i>
Hybrid Gram- mar	<i>dit dia rebere\$ de Hyacinto Coromines pages de Sparaguera fill de</i> <div style="border: 1px solid green; padding: 2px; display: inline-block;"><i>_J _u _a\$ _@</i></div>

GRAPHICS RECOGNITION

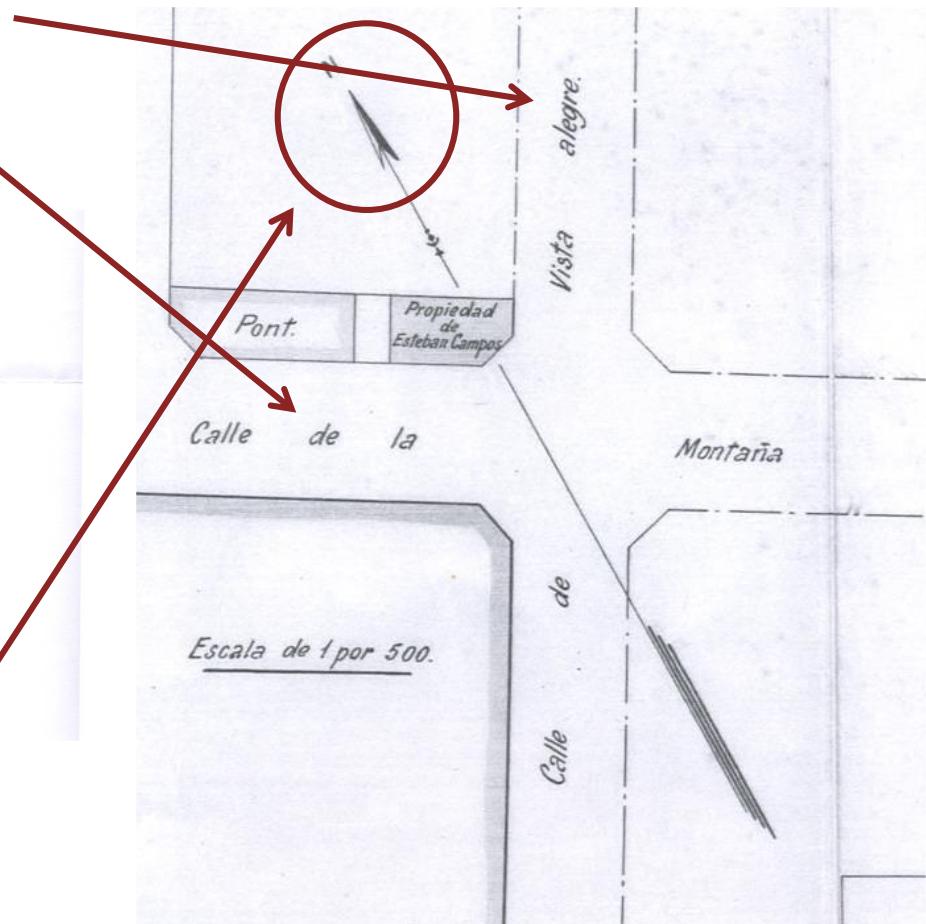
Examples of graphical documents

Text in any direction



Patterns

Symbols



Posters: Spanish Civil War

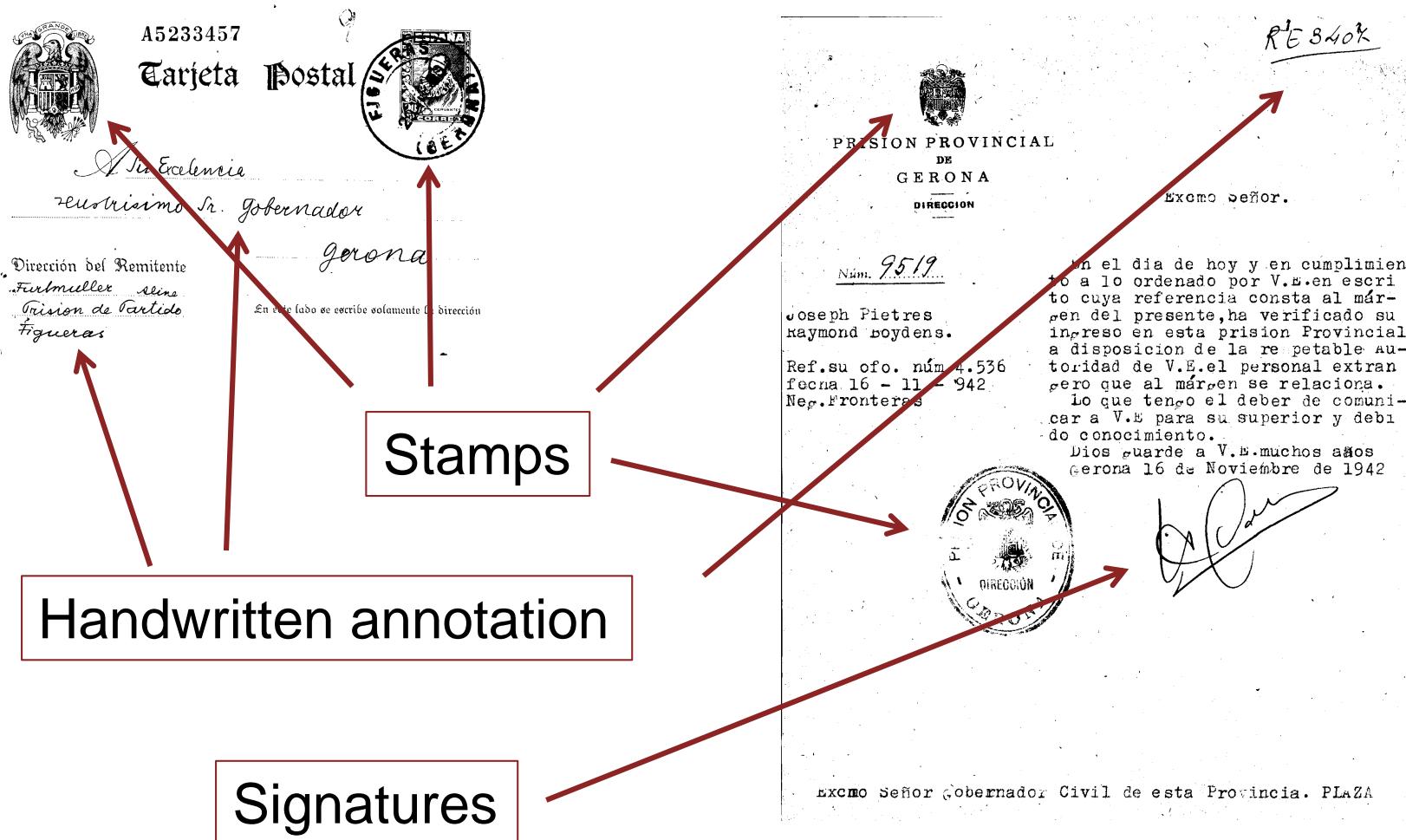
Red Cross



Artistic Text

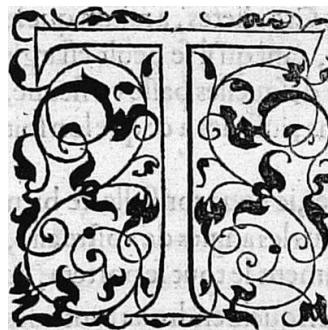
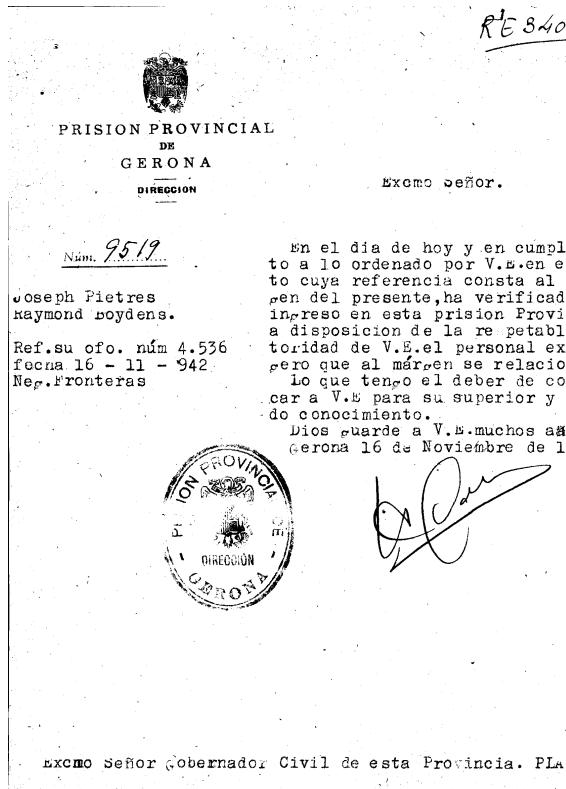


Files: The “Frontera” Archives



Main Properties

■ Complex layout



Symbol Recognition Methods

Problem in Hand-drawn

Distortions / Deformations



High variability intra-class

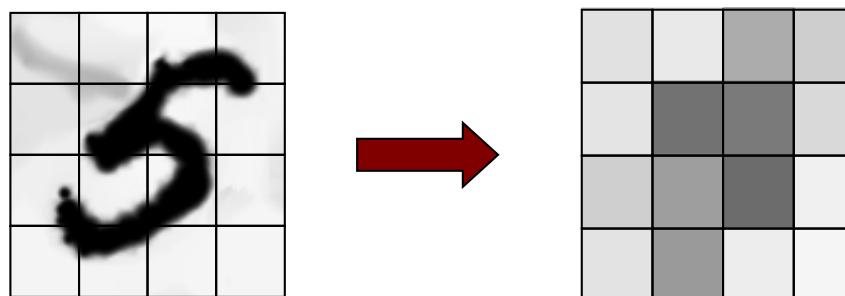


Different rotation / slant



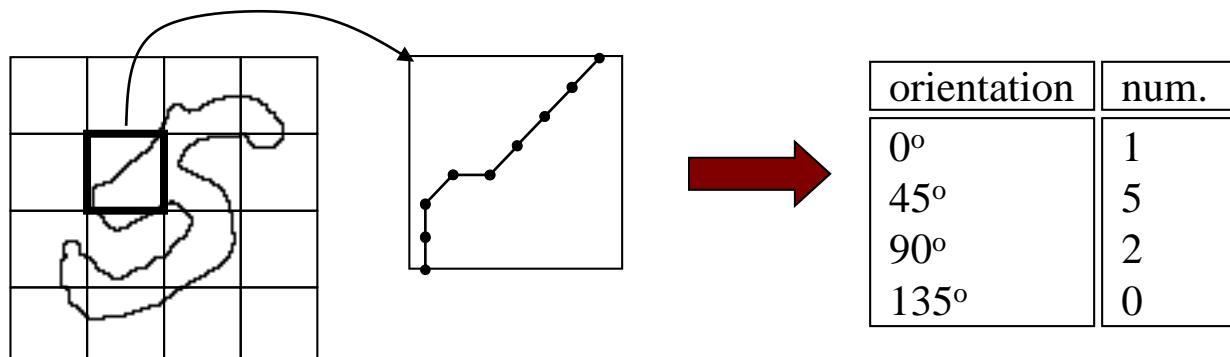
Statistical Features: Zoning and Contours

- The image is divided in $n \times m$ cells.
- For each cell the mean of gray levels is computed and all these values are joined in a feature vector of length $n \times m$.



- 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



Blurred Shape Model (BSM) descriptor

The symbol is divided in $n \times n$ cells

Each cell receives votes

The value depends on the distance to the centroid of the neighboring cells

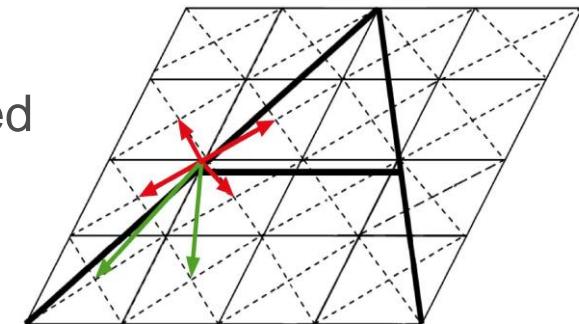
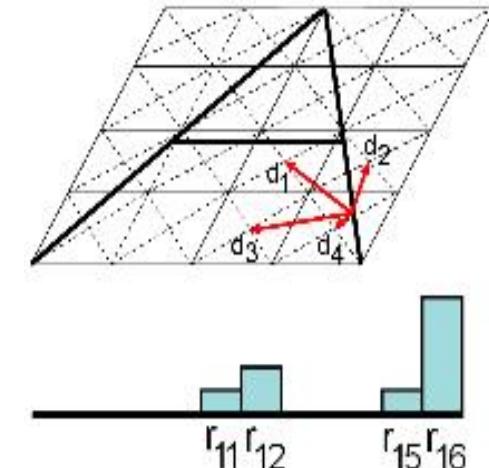
Normalize the probability density function

$$d_i = d(\mathbf{x}, r_i) = \|\mathbf{x} - c_i\|^2$$

$$v(r_i) = v(r_i) + \frac{1}{d_i D_i}, \quad D_i = \sum_{c_k \in N(r_i)} \frac{1}{\|\mathbf{x} - c_k\|^2}$$

$$v = \frac{v(i)}{\sum_{j=1}^{n^2} v(j)} \quad \forall i \in [1, \dots, n^2]$$

The number of cells determines the blurred degree allowed



S.Escalera, A.Fornés, O.Pujol, P.Radeva, G.Sánchez, J.Lladós. Blurred Shape Model for Binary and Grey-level Symbol Recognition. Pattern Recognition Letters, 2009.

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 θ of vectors joining p and the other points.
- To make the histogram more sensitive to positions of nearby points than to those of points farther away, the vectors are put into log-polar space.

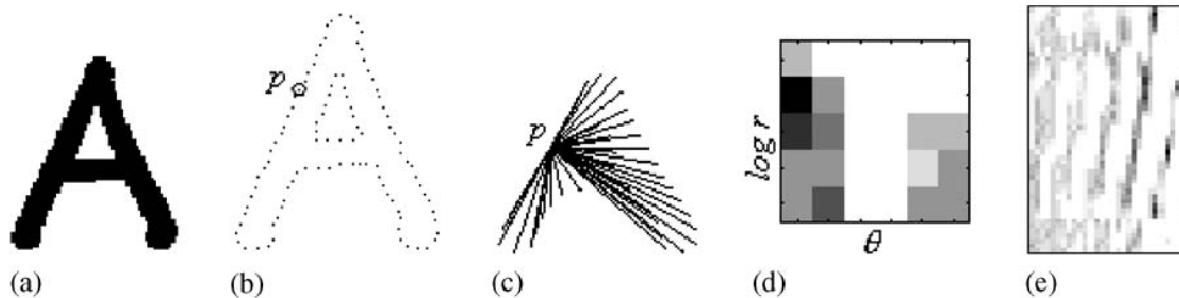
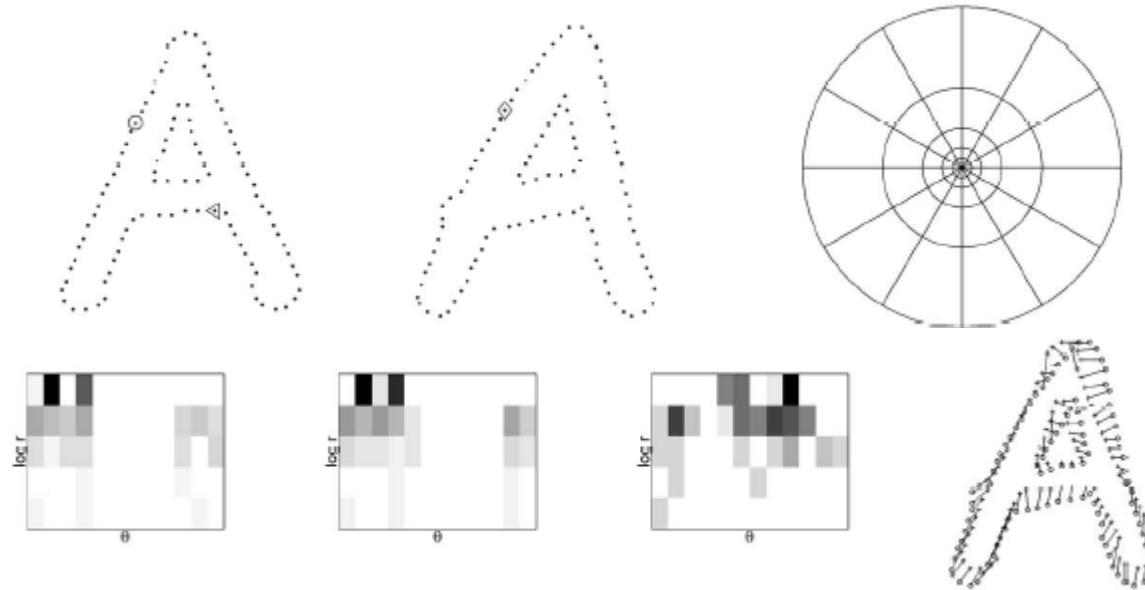


Fig. 3. Shape context. (a) a character shape; (b) edge image of (a); (c) a point p on shape (a) and all the vectors started from p ; (d) the log-polar histogram of the vectors in (c), the histogram is the context of point p ; (e) the context map of shape (a), each row of the context map is the flattened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

Histograms: Shape contexts

- The matching of two shapes is done by matching two context maps of the shapes. It minimizes the total cost of matching between one context matrix and all the permutations of another context matrix.
- In [Belongie et al, PAMI02] the matching is stated in terms of an alignment problem. It is solved by a bipartite graph matching algorithm.



CONCLUSION

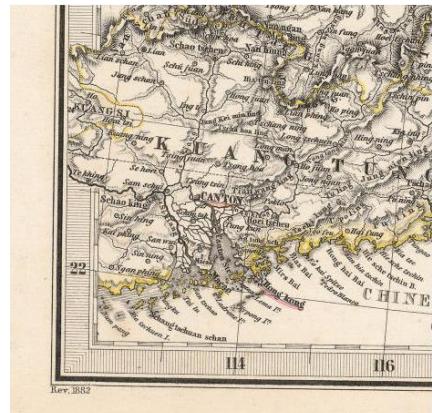
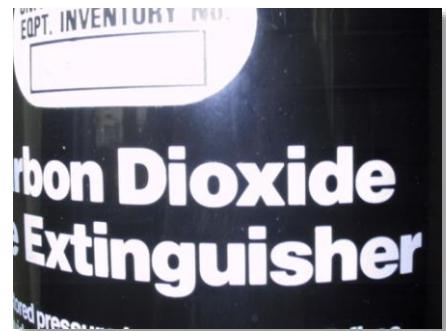
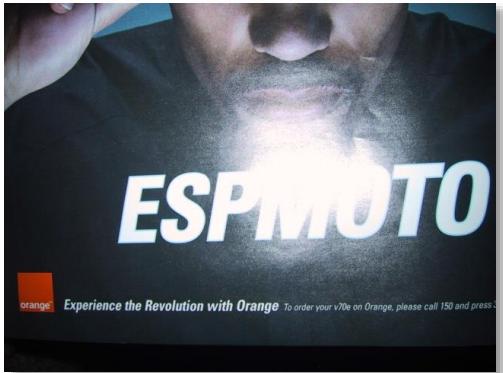
Discussion

- OCR, scientifically speaking, is a solved problem. A number of good commercial products exists.
- But, there is still room from new contributions in:
 - Highly noisy documents,
 - Systems with extremely high performance requirements (near 100%),
 - Large scale document processing (digital mail room),
 - Handwriting recognition,
 - Systems with complex layouts.
- The use of context will be very relevant for the improvement of reading systems: geometric and semantic context.

Recognition - Interesting links

- Abby Finereader OCR Software
<http://www.abbyyeu.com/es/>
- Google Tesseract OCR Software
<http://code.google.com/p/tesseract-ocr/>
- IMPACT project center of competence
<http://www.digitisation.eu/>
- Handwritten recognition demo:
<http://cat.prhlt.upv.es/iht>

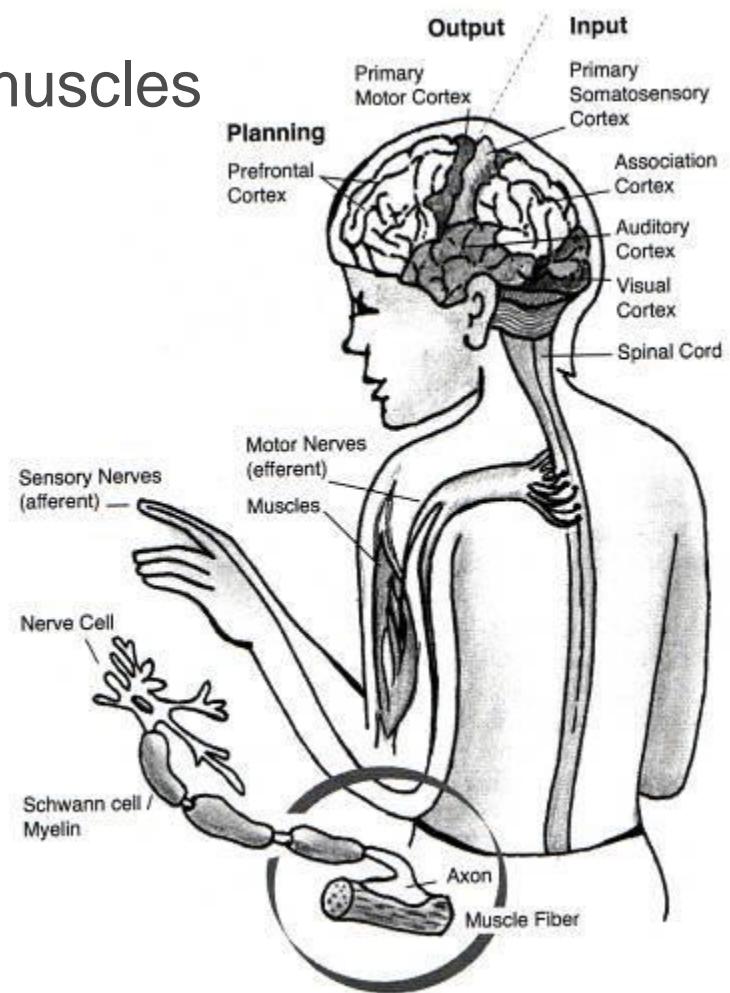
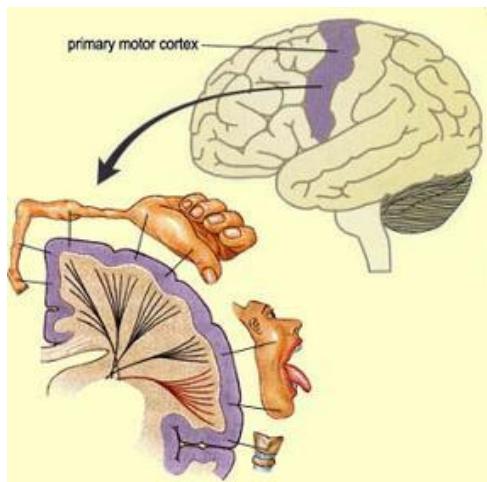
Challenges



WRITER IDENTIFICATION

What is handwriting?

- Highly automated skill
- Signals from brain to nerves to muscles
- Coordinated movements



Handwriting Fundamentals

Uniqueness



Handwriting variability

In that year Shakespeare had just turned forty and had written Hamlet two years before. Bacon was at work and Milton was just learning to read. James was followed by Charles, in whose reign came the Scottish Prayer Book in 1637. Significantly this made a deliberate return to the Book of 1549 and became the forerunner of some of the most important Prayer Books in the Anglican Communion.

The urge came from a conference of activists of Nkrumah's Convention Party "after powerful addresses by Comrades Krobo Edusei, Tewie Ademofio," and others. Strong debts followed strong words. In Takoradi a "limited state of emergency" was declared, giving the Government adequate power to maintain all essential services and ensure food supplies. Thus it becomes an offence punishable with imprisonment for anyone who "published a report likely to cause alarm or prejudicial to public safety".

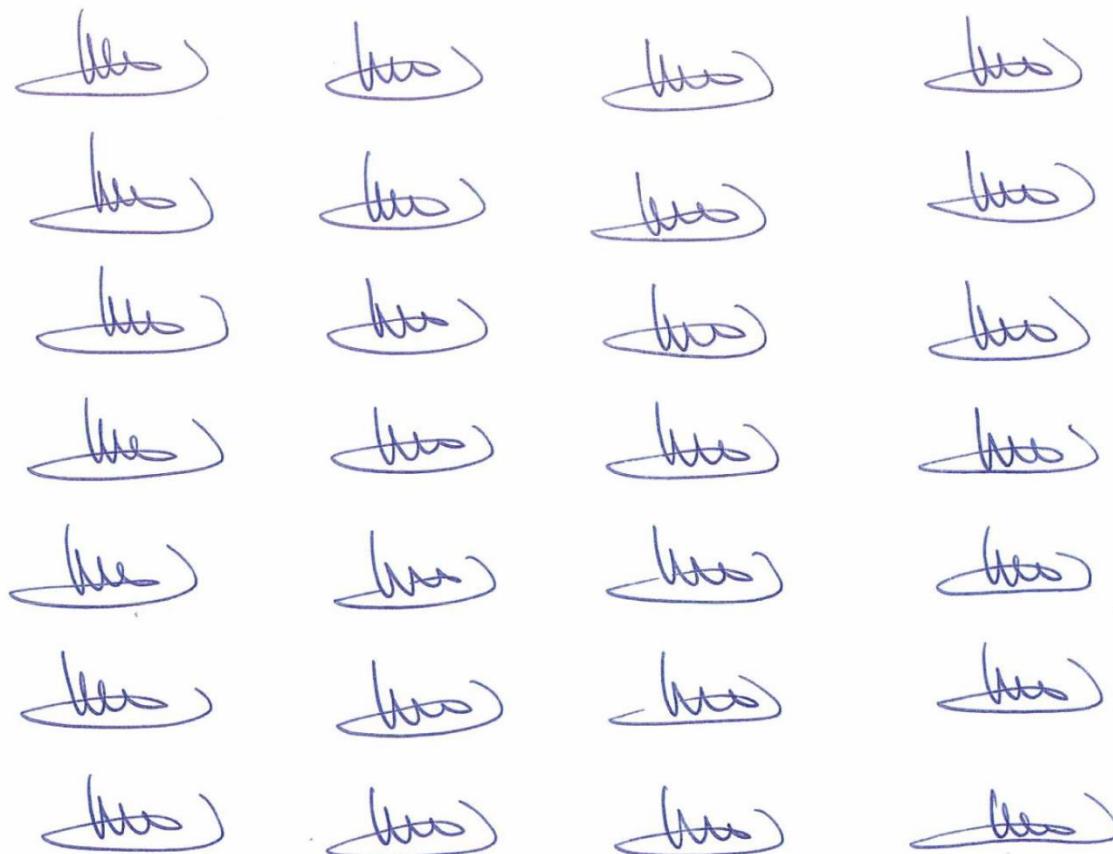
He said these concerned Mr. Weaver's alleged association with organisations black-linked by the government. Immediately Mr. Kennedy raised a like to Senator Robertson saying the Federal Bureau of Investigation had reported on Mr. Weaver. He believed he would perform "outstanding service" in his post. Senator Robertson's committee has to pass Mr. Weaver's nomination before it can be confirmed by the full Senate.

This is not a filmed play. It has been conceived throughout in terms of the cinema, and again and again it is the visual qualities of the story, and the marriage of the central characters to their background, which bring the film so vividly to life. In Fanny, which also has its premiere tomorrow, the director, Mr. Joshua Logan, attempted but failed to create the atmosphere of a city,

In fact, the Tories made it worse now for the sick and needy than Labour had to make it in 1950. And as a percentage of social service expenditure, health had fallen from 28.5 to 23.1 per cent. Then Mr. Brown swung his attack directly to the unsmiling Mr. Powell. He demanded that instead of taking it out of the patients Mr. Powell should take ruthless action against the drug making industry, whose profits had risen by up to 400 per cent. in the last eight years.

The plain, sober manner of its style all the more tellingly points up not only the horror of the case itself, which floundered on to the electrocution four years later of a German-born Bronx carpenter named Bruno Richard Hauptmann, but to the rattle-show emotionalism and sensation-hunger of that era.

But... exactly matching signature?



Handwriting: Recognition/Identification

Handwriting recognition:

Reduce characteristics of the handwriting style

Writer Identification:

Focus on these characteristics

Writer Identification (W.I.)

Writer Identification

Consists in determining the author of a piece of handwriting among a set of writers

Utilities

Forensic, such as signature verification

Typically applied to text documents

Wir beteiligen uns an einem Werk, das in mehr als 200 Ländern von ehrenamtlichen Mitarbeitern durchgeführt wird. In all diesen Ländern wird Menschen die Möglichkeit geboten, biblische Antworten auf wichtige Fragen zu erhalten wie zum Beispiel: Warum werden wir alt und sterben? Was ist der Sinn des Lebens? Wie kann man wahres Glück finden?

Pokey Mcinge
Wendy Wilson

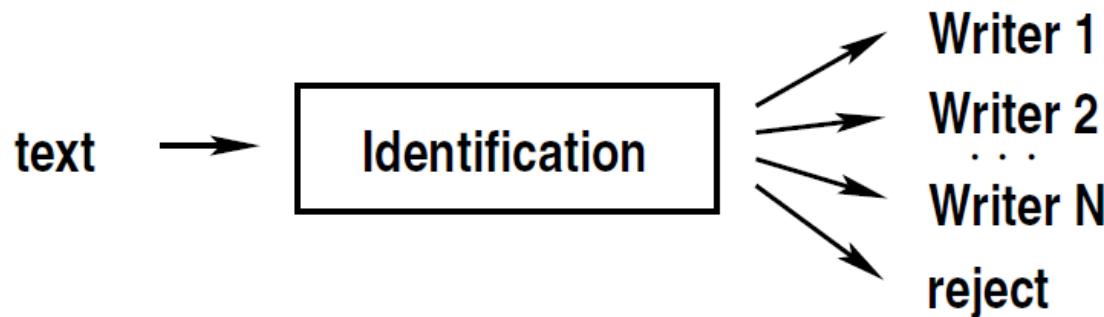
I just found out (from Lisa of Lisa writes... fame) that today is National Handwriting Day. As you can see, my handwriting is a strange mix of cursive and printing, and I don't think there's any pattern as far as when I do

Other scenarios: paintings, drawings, graffiti...



Writer Identification / Verification

Writer Identification → WHO has written a text?



Writer Verification → Is this person the author of a text?



Writer Identification: Test dependent / Text independent

In the first place it is not a great deal

In Gethsemane He prayed that the cup

Nor is she necessarily being deceitful. She really did feel tired

One of the greatest steps forward that has been

As it is, with so much of our life already

In this 200-fathom trench the herring do not touch the bottom.

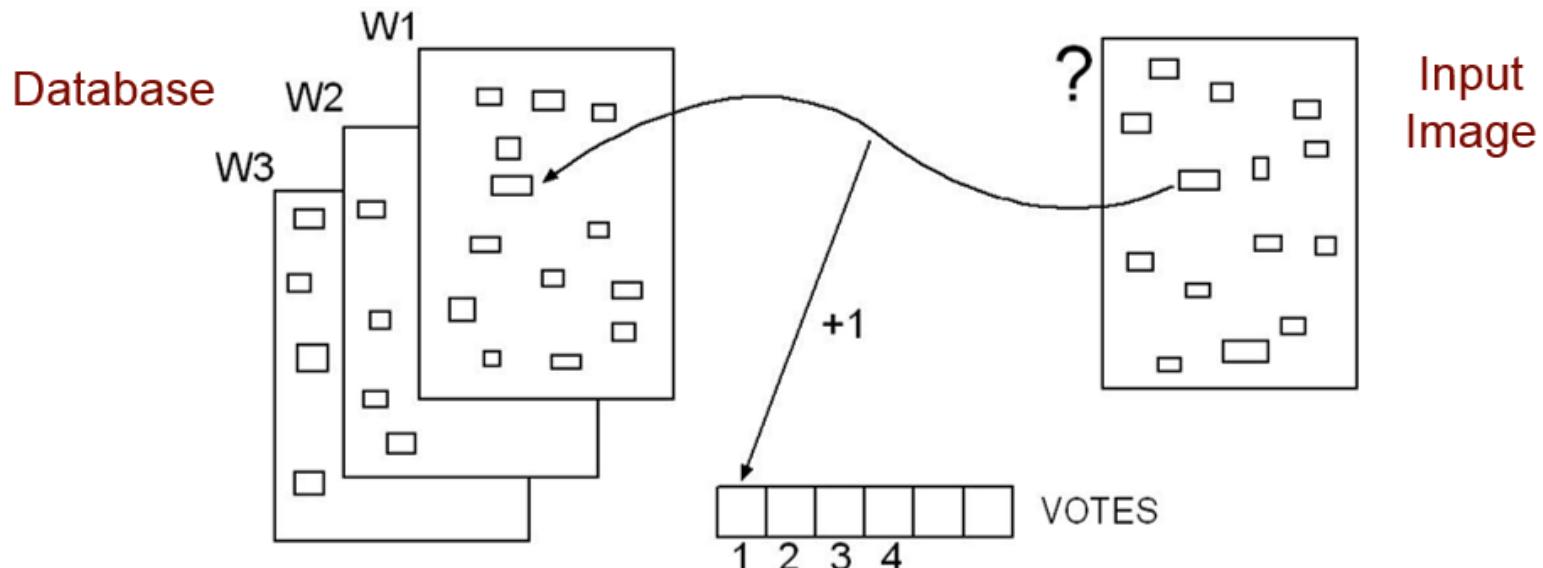
Actually the seine net has little or no cover.

Easterly winds, on the other hand,

W.I. Text-dependent

Detection of specific elements

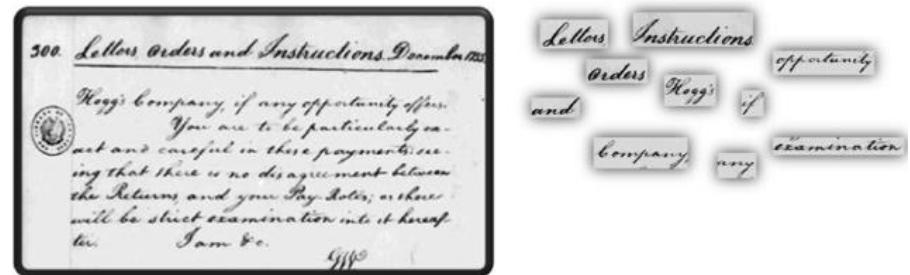
Comparison with the ones in the database



W.I. Text-dependent

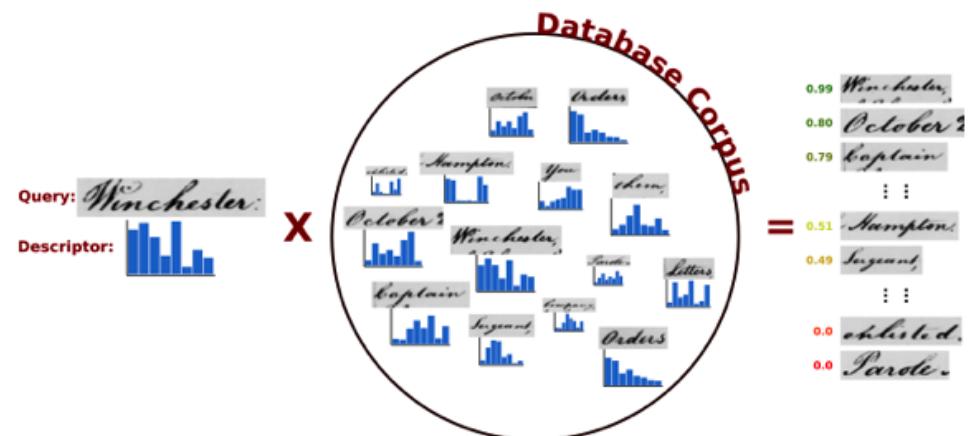
Detection of specific elements: Word Spotting

Word segmentation?



Feature extraction

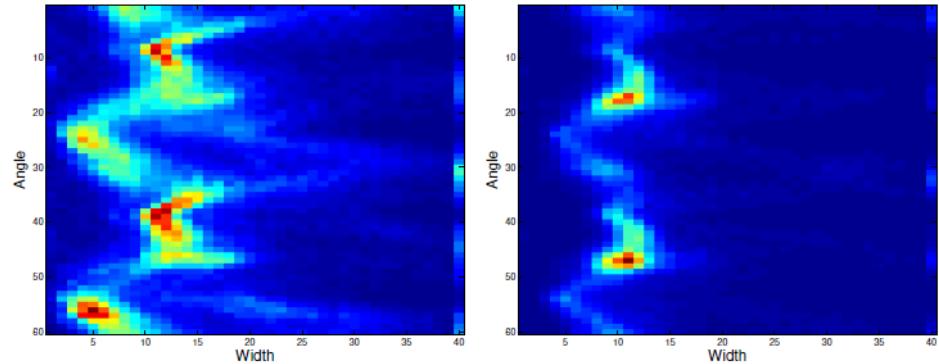
Shape comparison



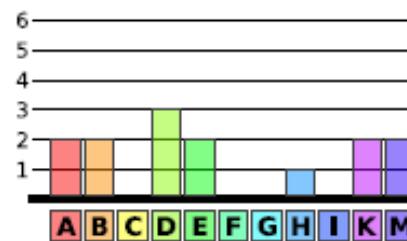
W.I. Text-independent

General aspect of the text

Quill features
(analysis of the
orientation of the contour)



Bag of Visual Words



Writer Identification System

Steps

Training

Learning from examples (genuine)

Classification

Decide the authorship of the input text

Writer Identification System: Training

Text Normalisation

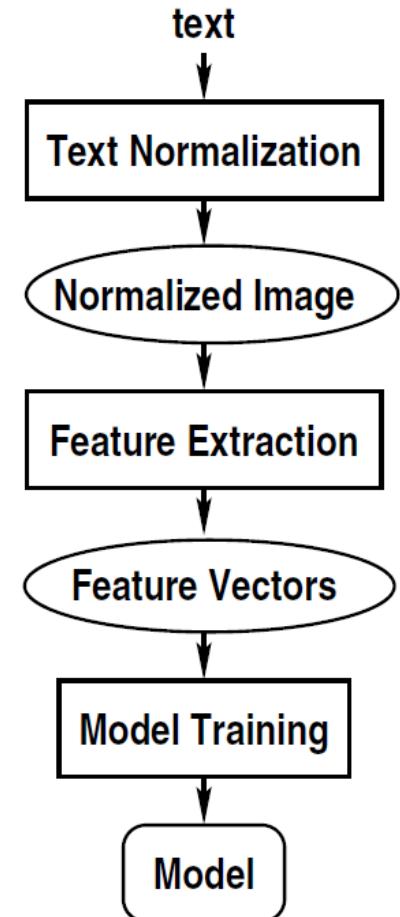
- Improve the quality of the features without removing characteristic style

Feature Extraction

- Features are designed to have a compact representation of the text that captures sufficient writer specific information

Model Training

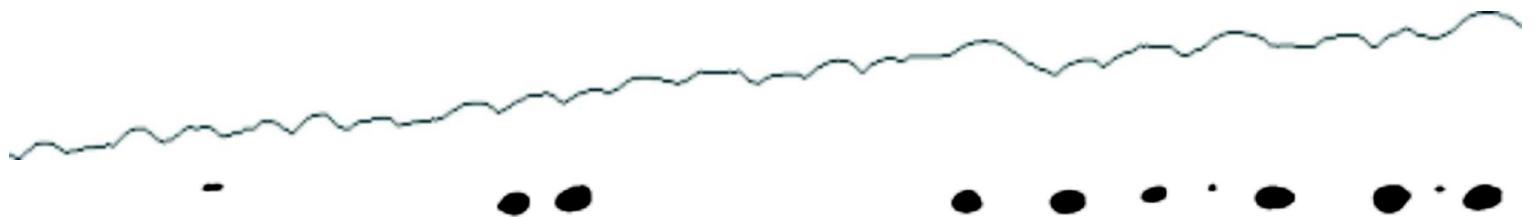
- A model represents the distribution of the sequences of features extracted from the handwritten text



Features: lines

100 features extracted for each text line:

- Basic Features
 - Slant, width of the writing, transitions, height of the 3 zones
- Connected components
 - Average of distances between cc
 - Average, median, standard deviation of the lengths of cc
- Enclosed regions (loops)
 - Average of the circularity
- Fractal features (how the text would grow with dilations)



In the first place it is not a great deal

Features: Textures

Creation of textures from the text lines

...
expectations allow
the most interests
to project that in
the next few years
will be multi-
media. Even the previous
led to the realization
of many more

uniformity of rep-
resentations behind
the excitement of
multimedia are
use I have encountered
in the next few years
be multimedia.

one of the market
answer the question
comes and what
we. Some realities
users as usual in
whether we like
ties are demanding
meeting educate

highly unfamiliar.
are informed and he
in surprise. I train
is reality in your
life. Customers a
ence of defects, but
local competitors are
more and more. A

smart, things in the nature
question, who is the author
writers that indicate that
other we like it or not, the
ing education. To quite ex-
tion, we noted in what
and change, especially a
numerous initiatives and

Compute Textural features

Textural features

Textural Features

Gabor Features



Fourier Transform in which the window is **Gaussian**

Parameters: Frequency $f = 4, 8, 16, 32$, Orientation $\alpha = 0^\circ, 45^\circ, 90^\circ, 135^\circ$

Extract mean & standard deviation of the 16 images $\rightarrow 32$ features

Grey-Scale Co-occurrence Matrices

$GSCM_{d,\alpha}(a,b)$ = number of pairs (Pixel1,Pixel2)

with a distance d and angle α , with color a and b

In binary images we have only 2 levels (black-white)

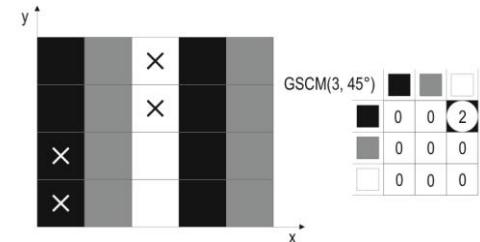
Distance $d = 1, 2, 3, 4, 5$, Angle $\alpha = 0^\circ, 45^\circ, 90^\circ, 135^\circ$

20 matrices 2×2

with 3 independent values

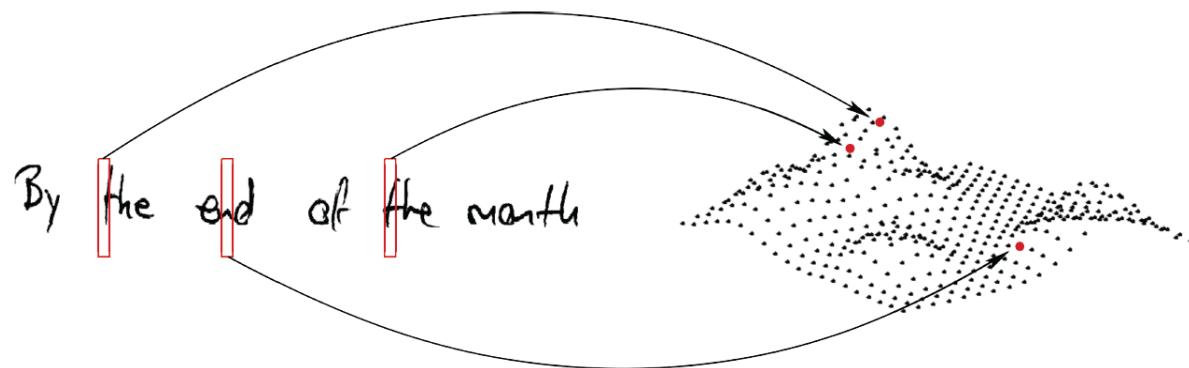
(diagonal symmetry)

Total = $20 \times 3 = 60$ features



Modelling

Once text is normalized and features are extracted

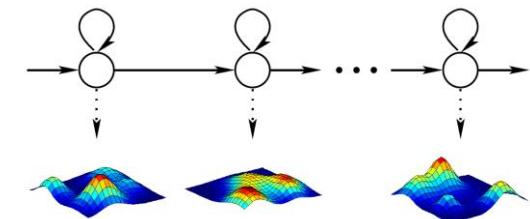
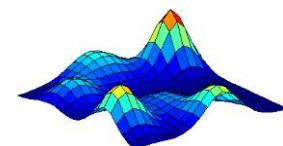


Distribution of features is modelled by a statistical model

1 model / 1 writer

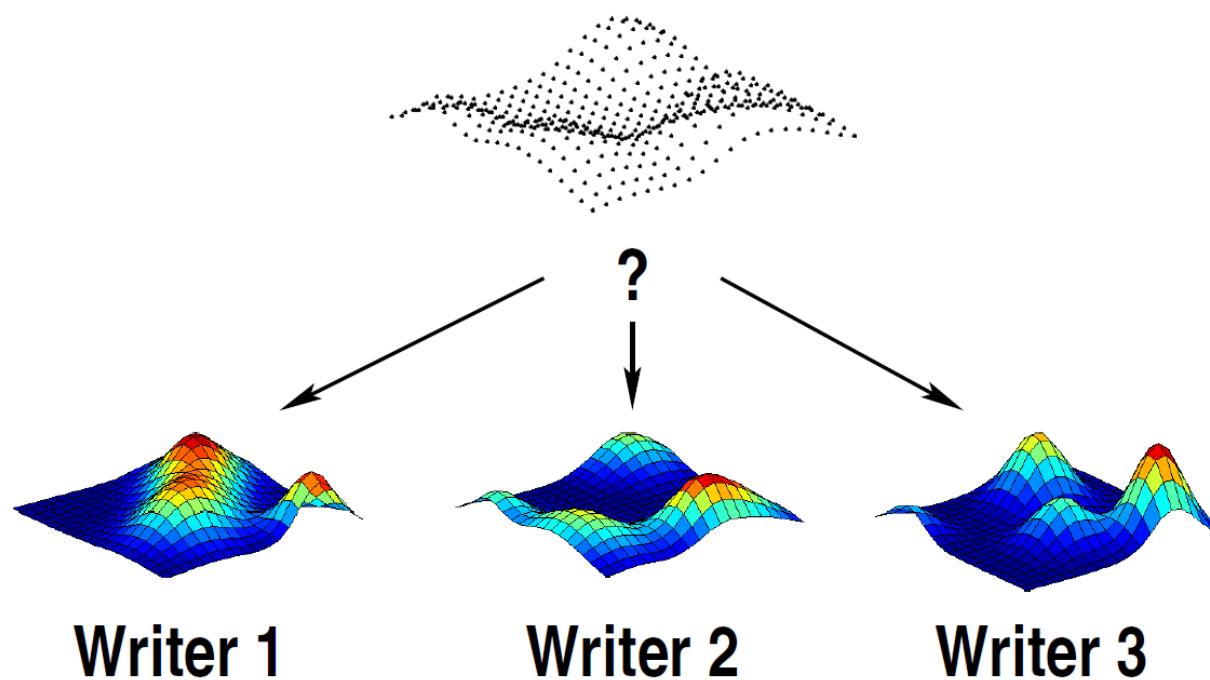
Examples of statistical models

- Gaussian Mixture Model (GMM)
- Hidden Markov Model (HMM)



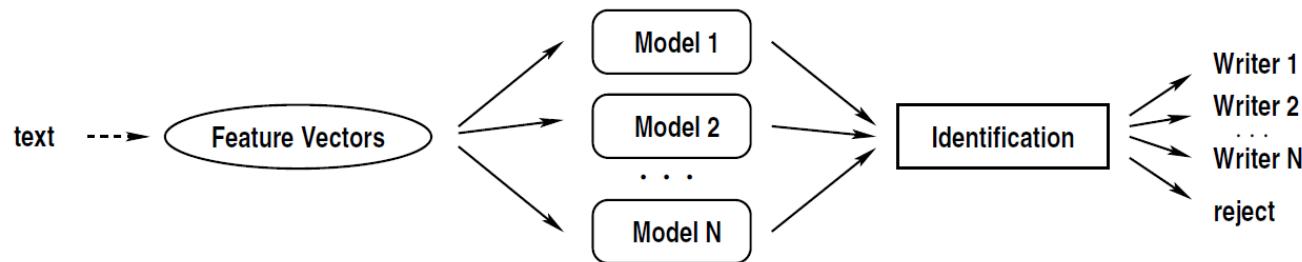
Classification

A set of features is extracted from the unknown author
The features are assigned to the model (writer) that best
represents the data

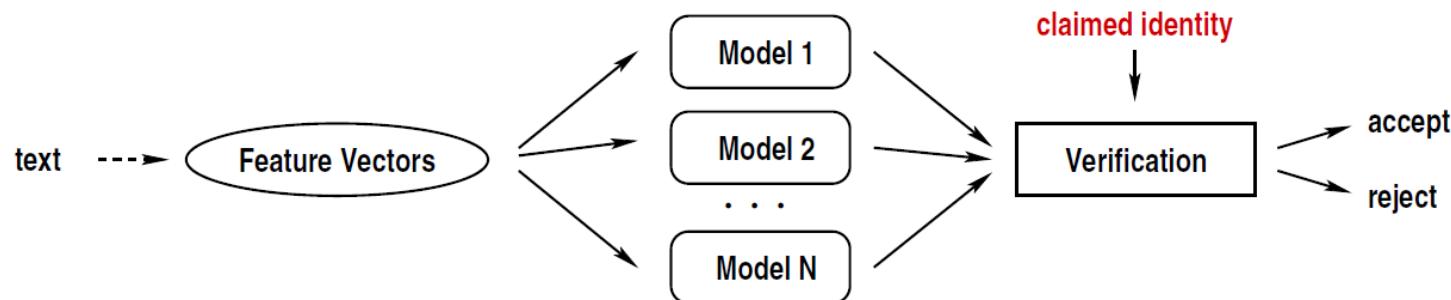


Identification / Verification

Writer identification: N+1 class problem



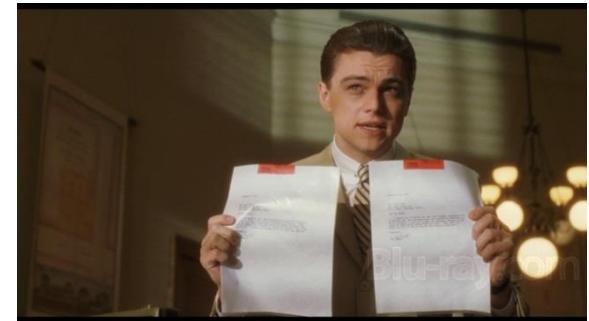
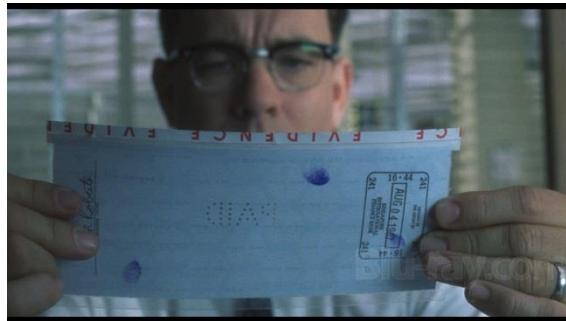
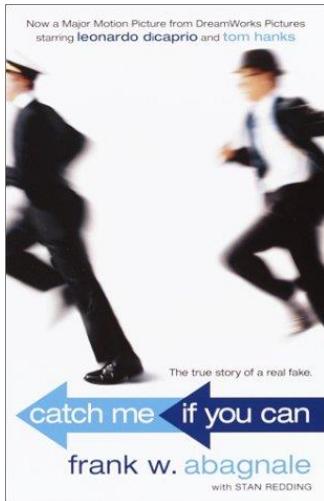
Writer verification: 2 class problem



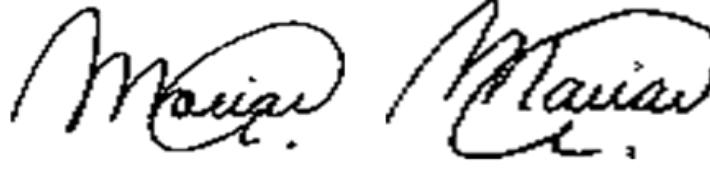
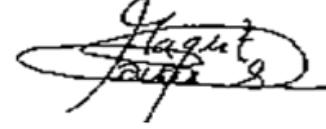
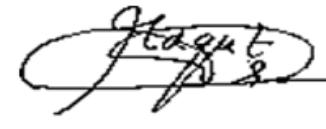
Signature Verification

Applications:

- Bank checks, contracts, receipts → fraud detection
- Personnel identification (login in your PC)

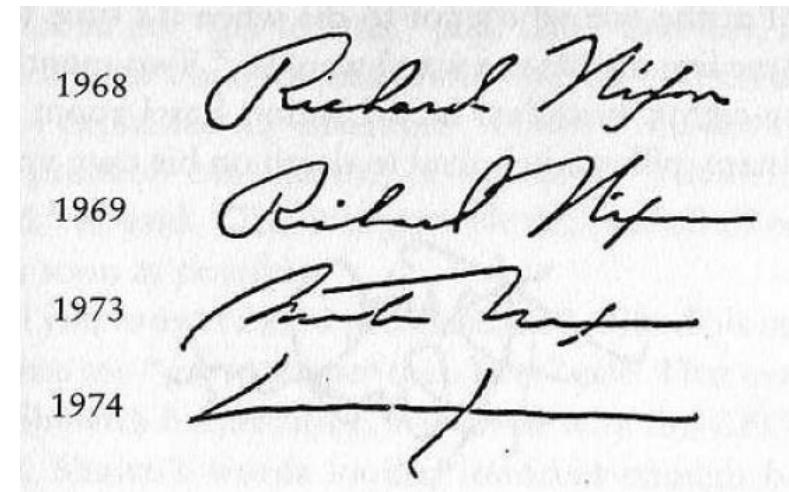
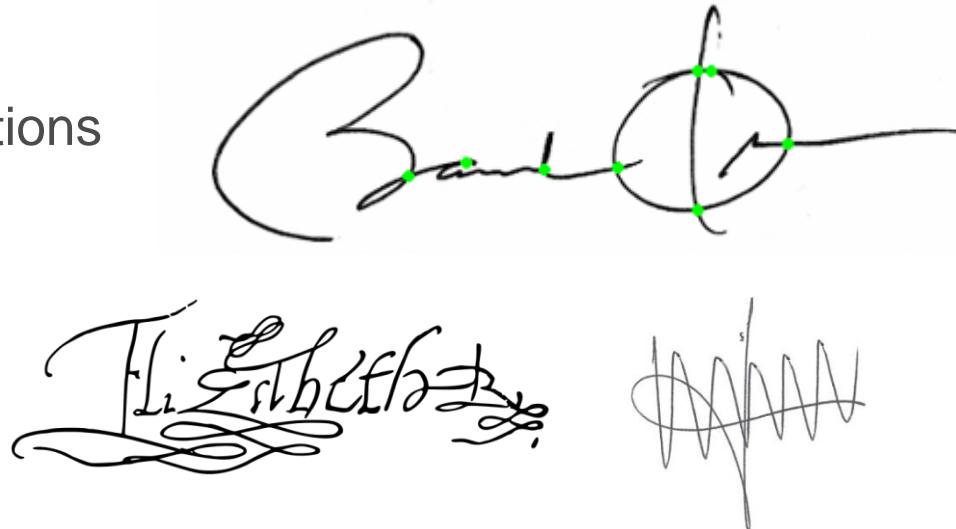


Signatures: Examples

GENUINE SAMPLES	FORGED SAMPLES		
			
			

Issues

- Complexity
 - Number of strokes, intersections
- Legibility
- Movement order (strokes)
- Pen pressure
- Slant
- Size
- Variation (over time?)



Graphical languages?

यत्र नार्यस्तु पूज्यन्ते रमन्ते तत्र देवताः ।
यत्रैतास्तु न पूज्यन्ते विनश्यत्याशु तत्कुलम् ॥
उपाध्यायान् दशाचार्यः शताचार्यस्तथा पिता ।

王 光 心

ਮਿਡੀਨਜ਼ ਆਰਾ

火 木 景
三 月 金
十 人 德
齋 翼 離

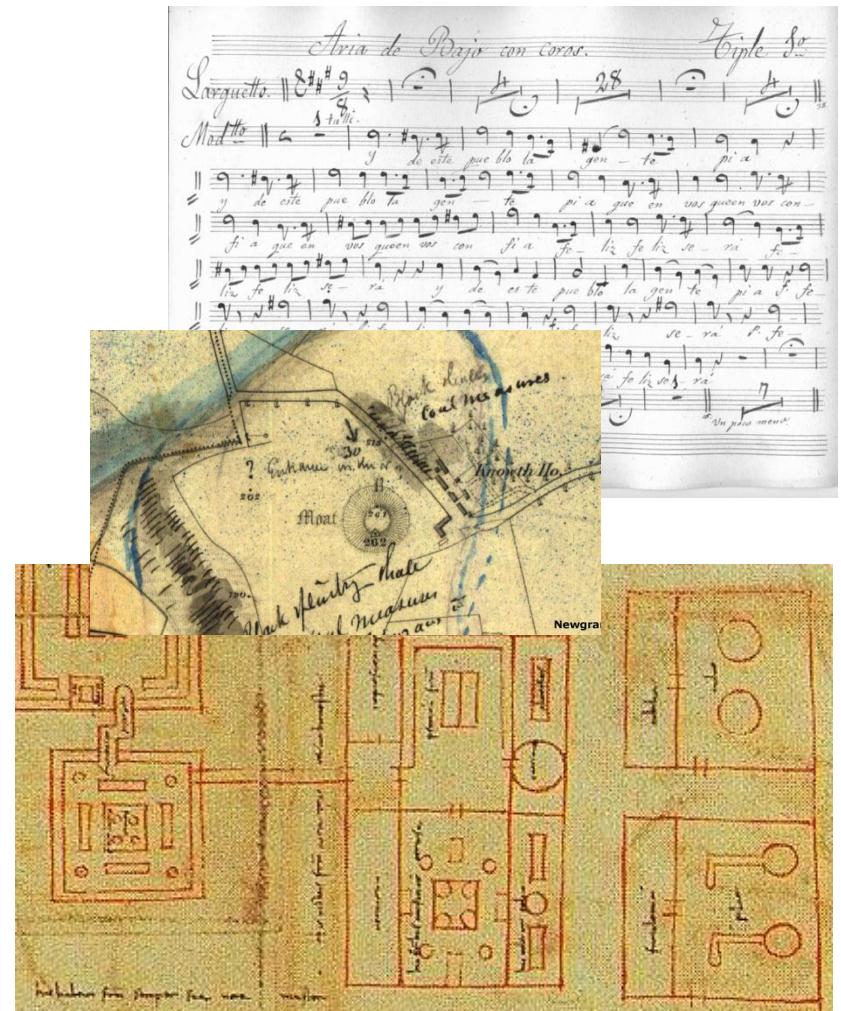
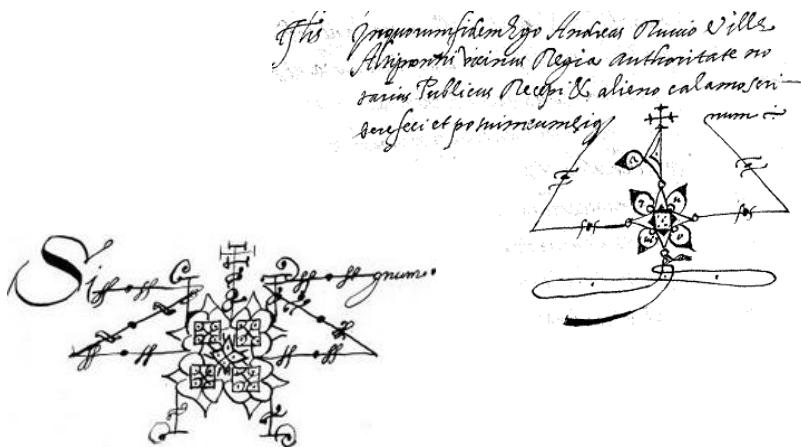
Graphic Documents

Instead of characters → symbols

Kinds of graphic documents:

maps, engineering documents, music scores, architectural plans...

Notaries' signature



Parallelism: text \leftrightarrow symbols

Text-dependent \leftrightarrow Symbol-dependent

shape of specific elements (words, symbols)

this is my handwriting

this is my handwriting.

this is my handwriting

R · q q t o q | q q q

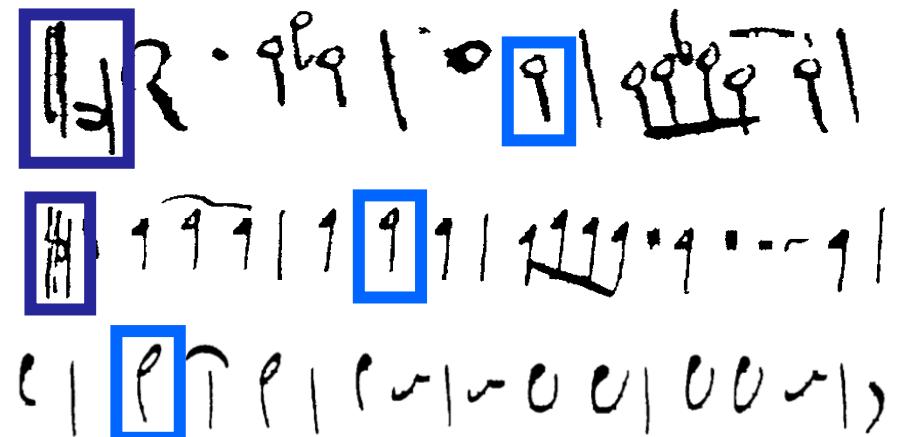
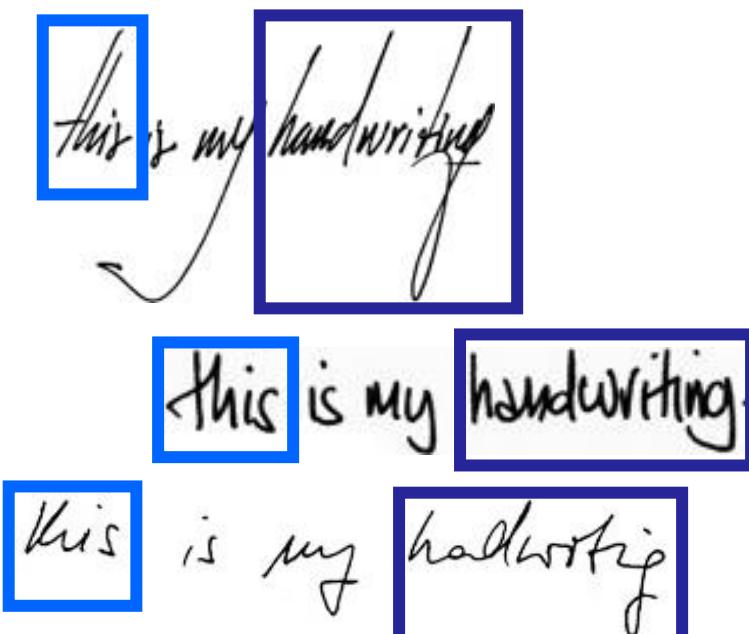
q: q q | q q | q q q - q |

l p T p | p - o o | o o - |

Parallelism: text \leftrightarrow symbols

Text-dependent \leftrightarrow Symbol-dependent

shape of specific elements (words, symbols)



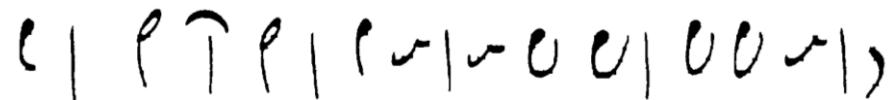
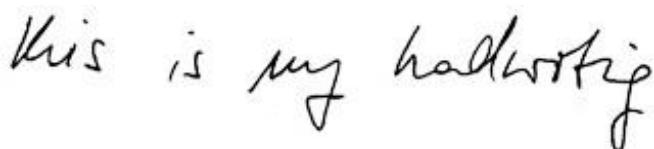
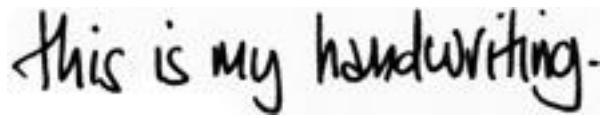
Parallelism: text \leftrightarrow symbols

Text-dependent \leftrightarrow **Symbol-dependent**

shape of specific elements (words, symbols)

Text-independent \leftrightarrow **Symbol-independent**

slant, curvature, straightness...



Parallelism: text \leftrightarrow symbols

Text-dependent \leftrightarrow **Symbol-dependent**

shape of specific elements (words, symbols)

Text-independent \leftrightarrow **Symbol-independent**

slant, curvature, straightness...

this is my handwriting

this is my handwriting

this is my handwriting

43.99 | 09 | 00 |

4: 199|199|111.9---9|

1 1 1 1 | 1 1 1 1 ,

Bag of Notes for Writer Identification

Bag of Features framework

Connected components

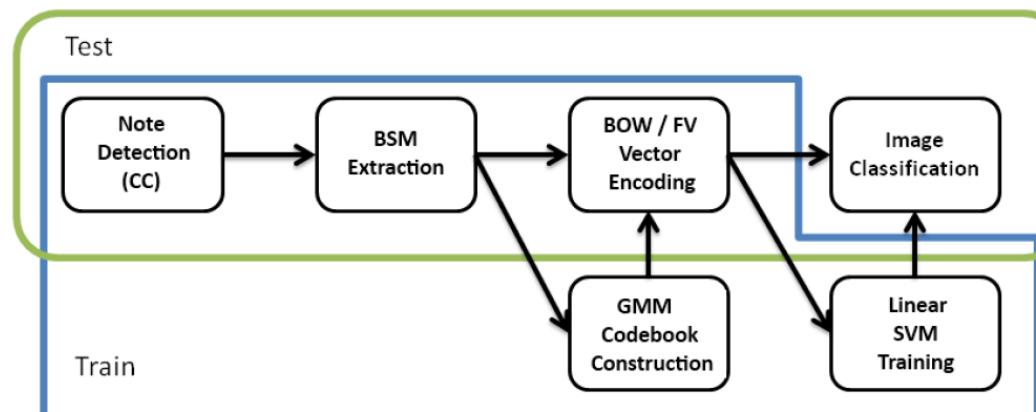
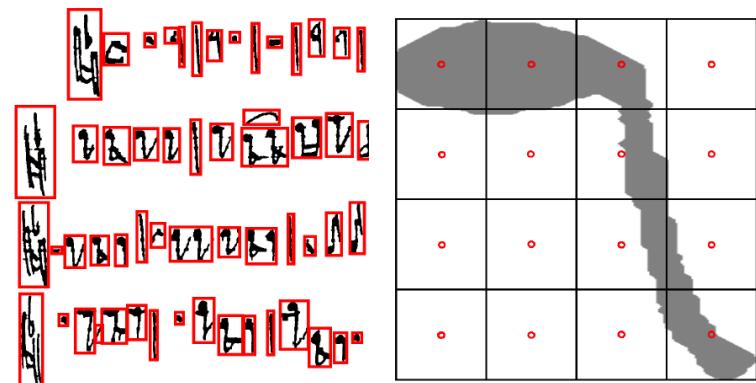
BSM descriptor

Codebook: GMM

BoW + Fisher Vectors

Frequencies of words

Position and sparseness with respect to the means and standard deviations of the codebook words (gradient vector)



A.Gordo, A.Fornés, E.Valveny. Writer identification in handwritten musical scores with bags of notes. Pattern Recognition, 2013

Dating

Evolution of handwriting (variation over time)

