Shape Descriptors, Classifier fusion and other Techniques applied to Graphics Recognition and related problems

O. Ramos Terrades

April 1, 2009

Personal information

2001: "Maîtrise" in Mathematics by the Autonomous University of Barcelona (UAB)

2003: Master in Computer Vision by the Computer Vision Centre

2006: PhD in Computer Science by the UAB and Nancy 2 University (UN2)

2007: Visiting Scientist in the CVPR Unit of the Indian Statistical Institute (ISI)

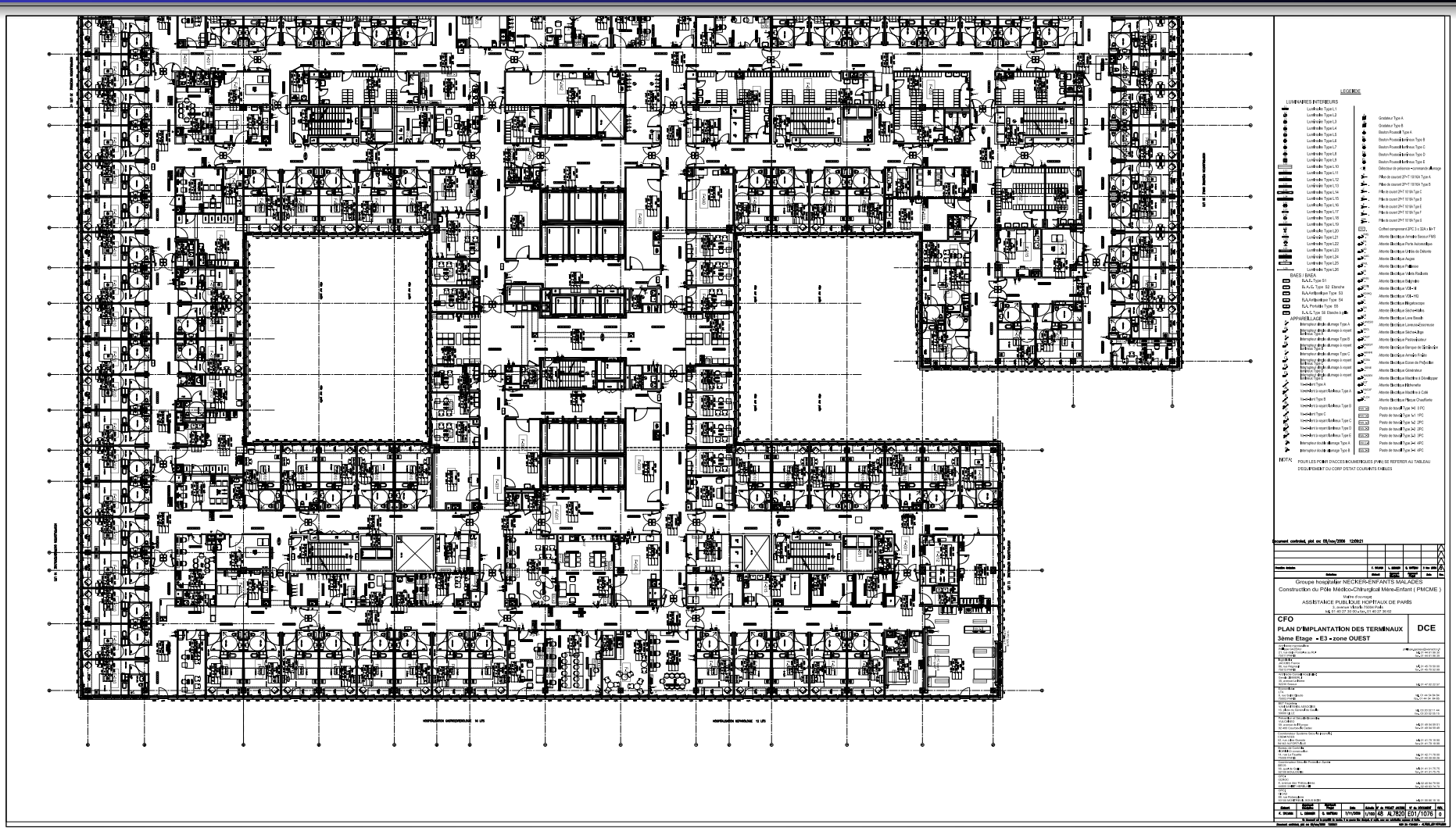
2008: ATER at UN2.

Nowadays: Research fellow at Technological Institute of Computer Science (ITI in Spanish)

Outline

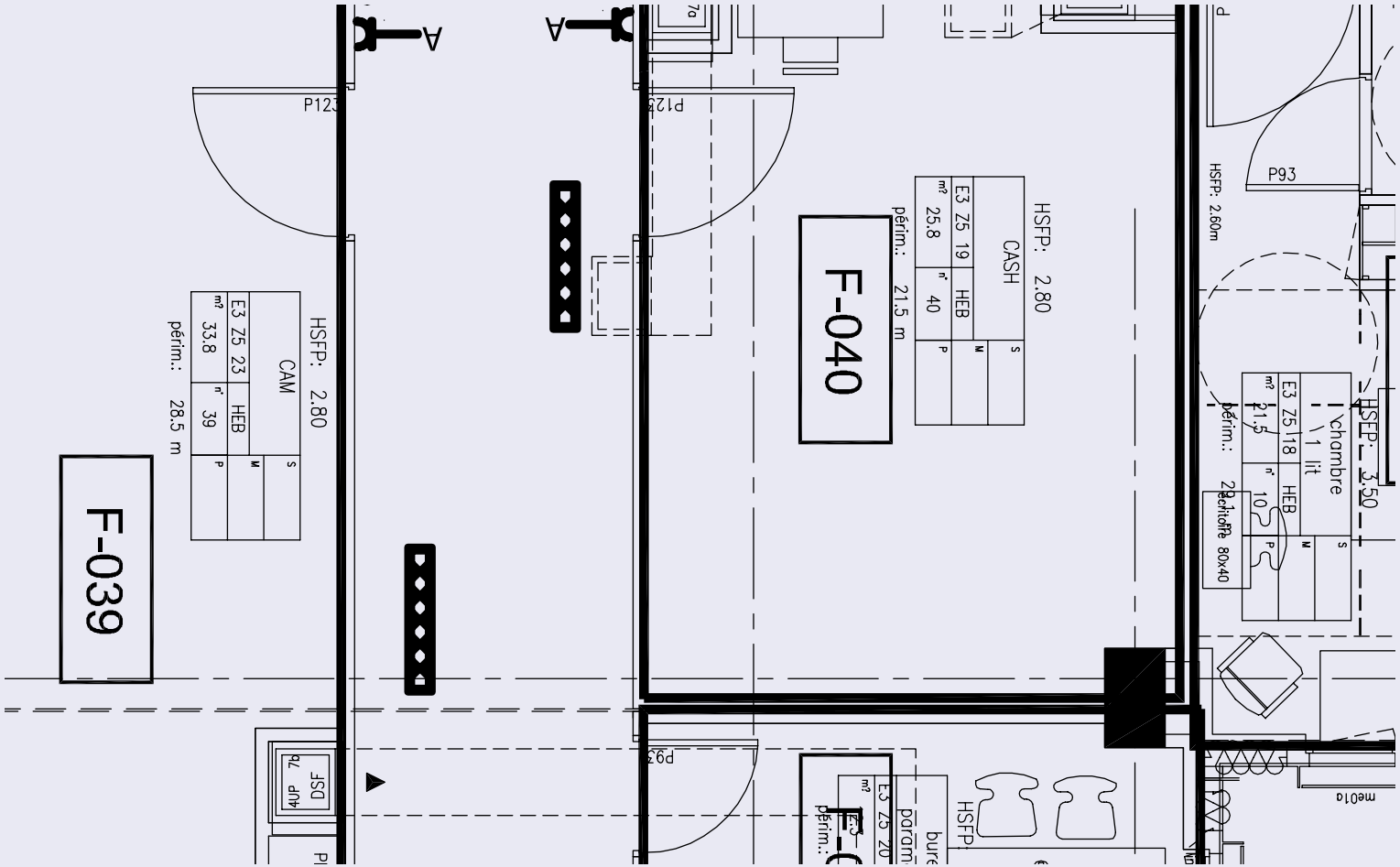
Outline

Technical Documents

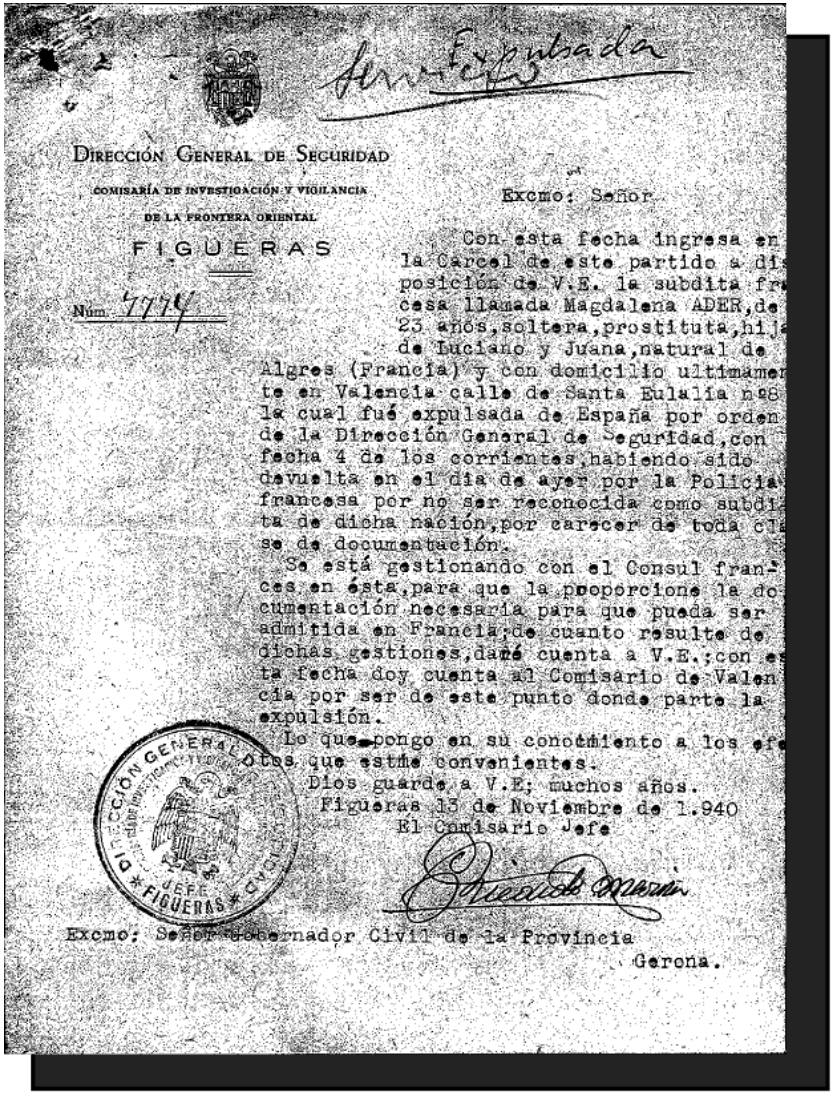


Technical Documents

Zoomed...



Historical Documents



DIRECCIÓN GENERAL DE SEGURIDAD

COMISARÍA DE INVESTIGACIÓN Y VIGILANCIA DE LA FRONTERA ORIENTAL

FIGUERAS

Núm. 7774

Expulsada

Excmo: Señor

Con esta fecha ingresa en la Carcel de este partido a di posición de V.E. la subdita fr cesa llamada Magdalena ADER,de 23 años,soltera,prostituta,hij de Luciano y Juana,natural de Algres (Francia) y con domicilio ultimame te en Valencia calle de Santa Eulalia nº8 la cual fué expulsada de España por orden de la Dirección General de Seguridad,con fecha 4 de los corrientes,habiendo sido devuelta en el dia de ayer por la Policia francesa por no ser reconocida como subdi ta de dicha nación,por carecer de toda cl se de documentación.

Se está gestionando con el Consul fran ces en ésta,para que la proporcione la do cumentación necesaria para que pueda ser admitida en Francia;de cuanto resulte de dichas gestiones,daré cuenta a V.E.;con es ta fecha doy cuenta al Comisario de Valen cia por ser de este punto donde parte la expulsión.

Lo que pongo en su conocimiento a los efe otos que estime convenientes.

Dios guarde a V.E; muchos años.

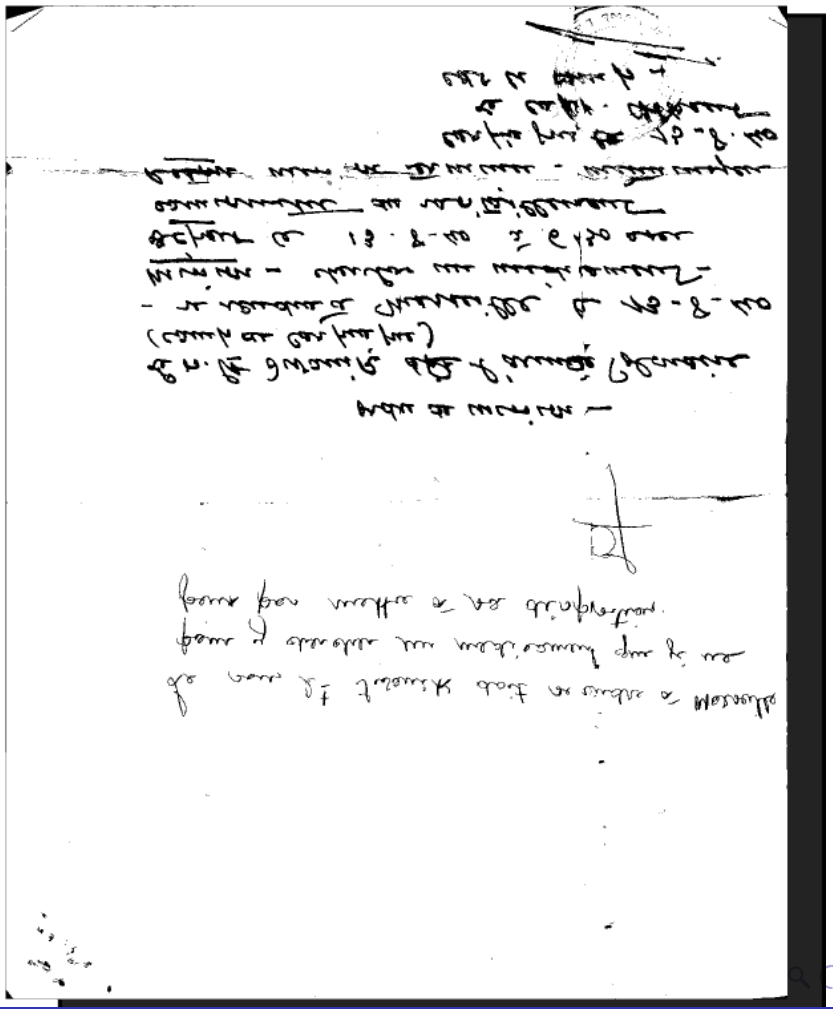
Figueras 13 de Noviembre de 1.940

El Comisario Jefe

DIRECCIÓN GENERAL DE SEGURIDAD * JEFE FIGUERAS *

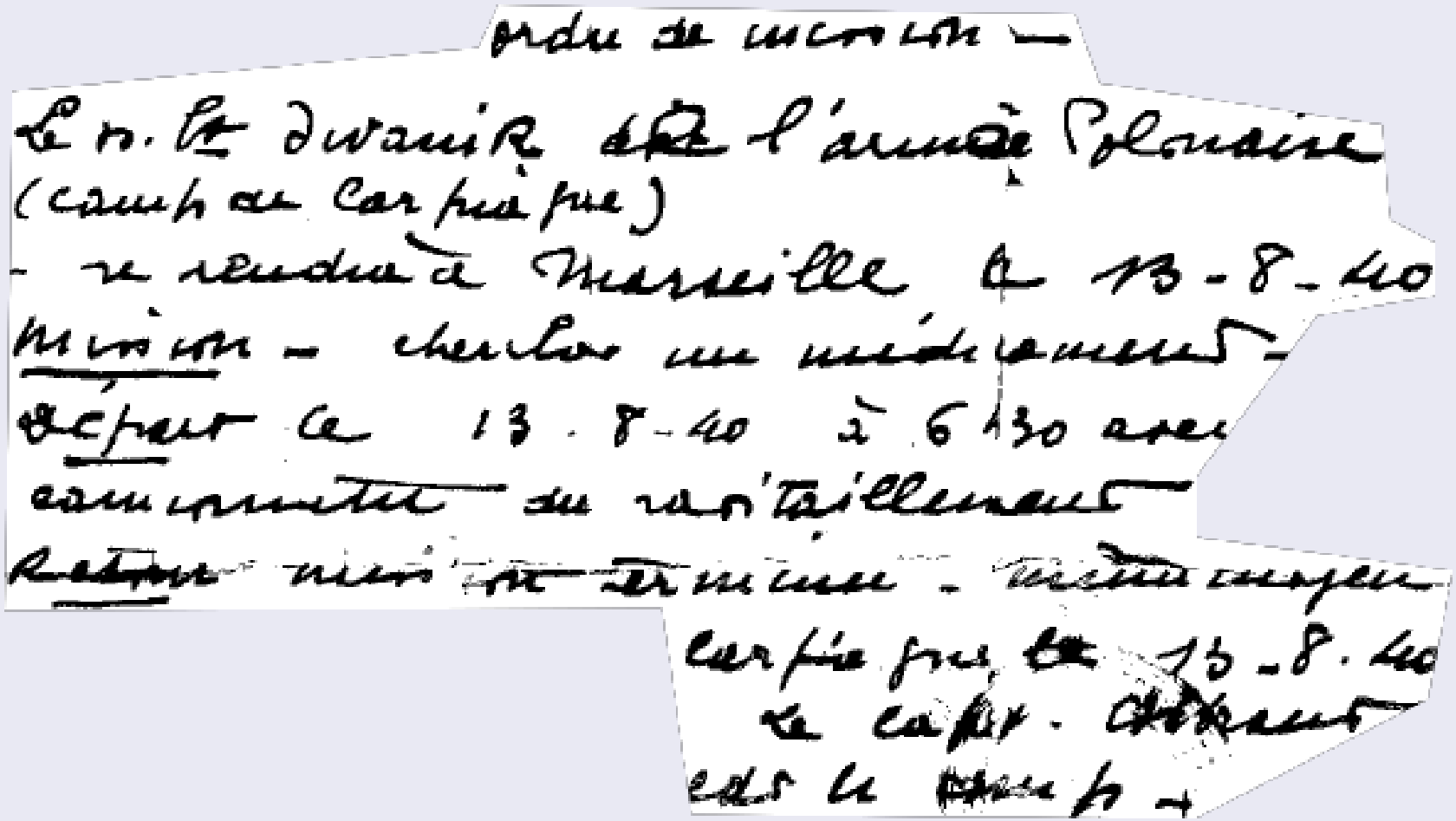
Excmo: Señor Gobernador Civil de la Provincia

Gerona.



Historical Documents

handwriten text...



Tasks of interest

- Extract information from technical and historical documents
- Find graphical information through document collections
- Script identification
- Handwriten recognition

Open questions:

- How represent graphical entities?

Open questions:

- How represent graphical entities?

By using descriptors

- How to organise information?

Open questions:

- How represent graphical entities?

By using descriptors

- How to organise information?

Applying indexing strategies

- How to determine which are the best descriptors?

Open questions:

- How represent graphical entities?

By using descriptors

- How to organise information?

Applying indexing strategies

- How to determine which are the best descriptors?

Evaluating descriptors on reference benchmarks

- How to improve performances?

Open questions:

- How represent graphical entities?

By using descriptors

- How to organise information?

Applying indexing strategies

- How to determine which are the best descriptors?

Evaluating descriptors on reference benchmarks

- How to improve performances?

Combining multiple classifiers

Methods

- Definition and computation of new descriptors

Methods

- Definition and computation of new descriptors
- Indexing and structuring data

Methods

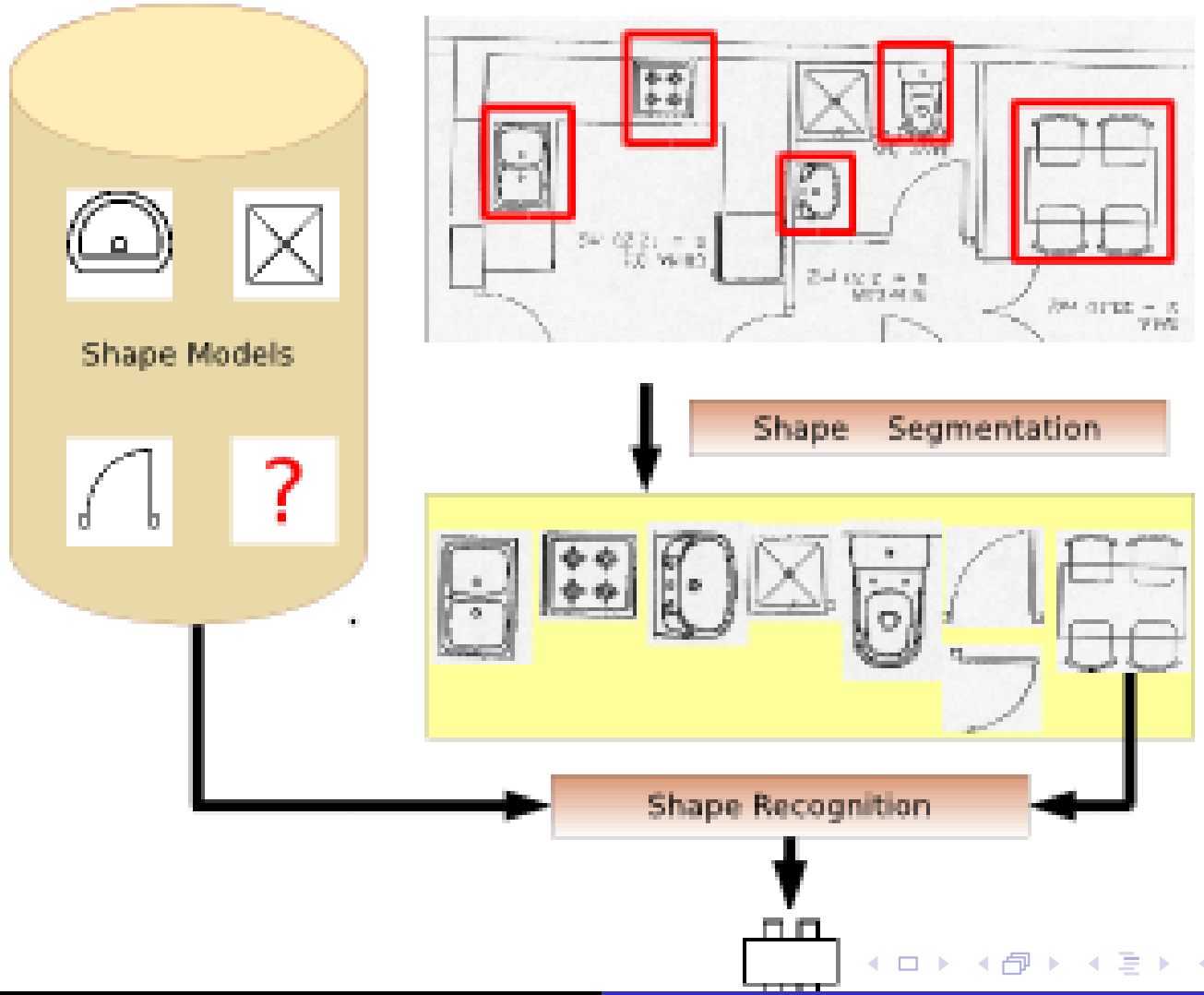
- Definition and computation of new descriptors
- Indexing and structuring data
- Performance evaluation

Methods

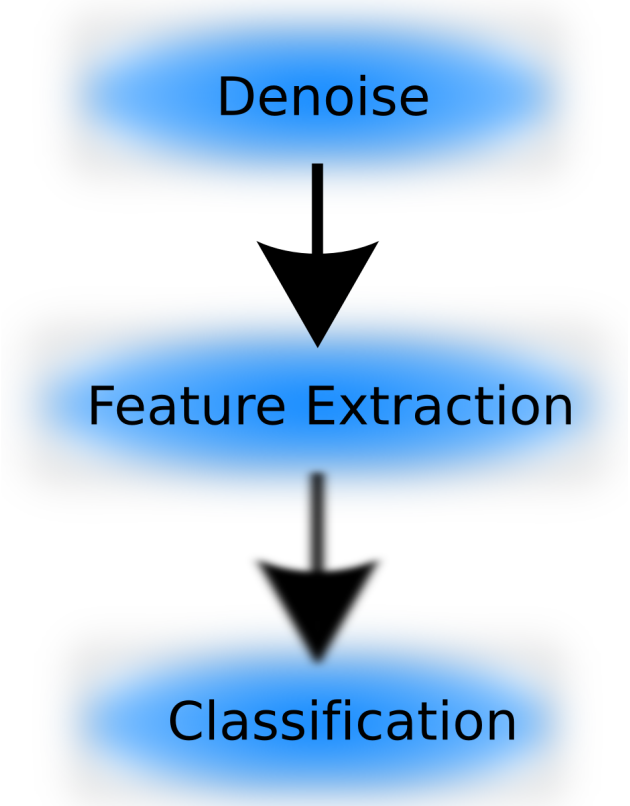
- Definition and computation of new descriptors
- Indexing and structuring data
- Performance evaluation
- Research on machine learning methods

Outline

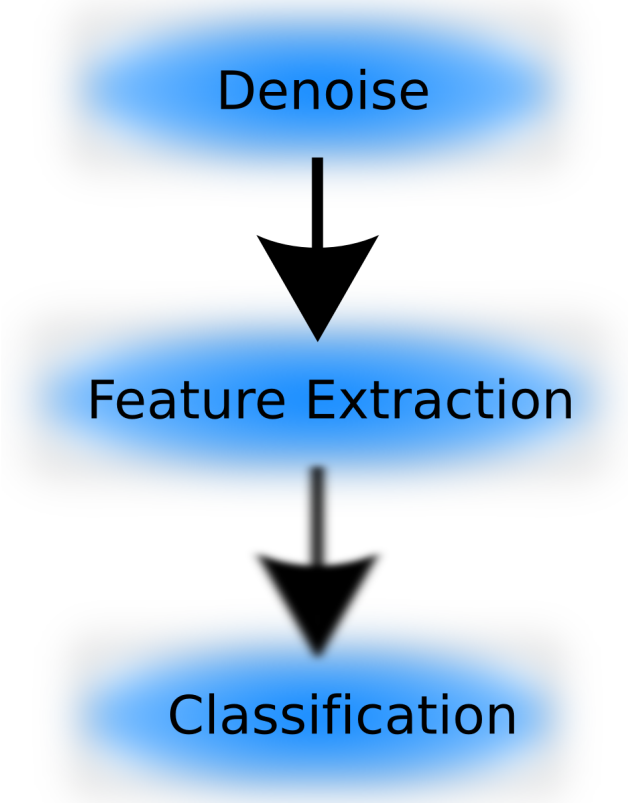
Scheme Process

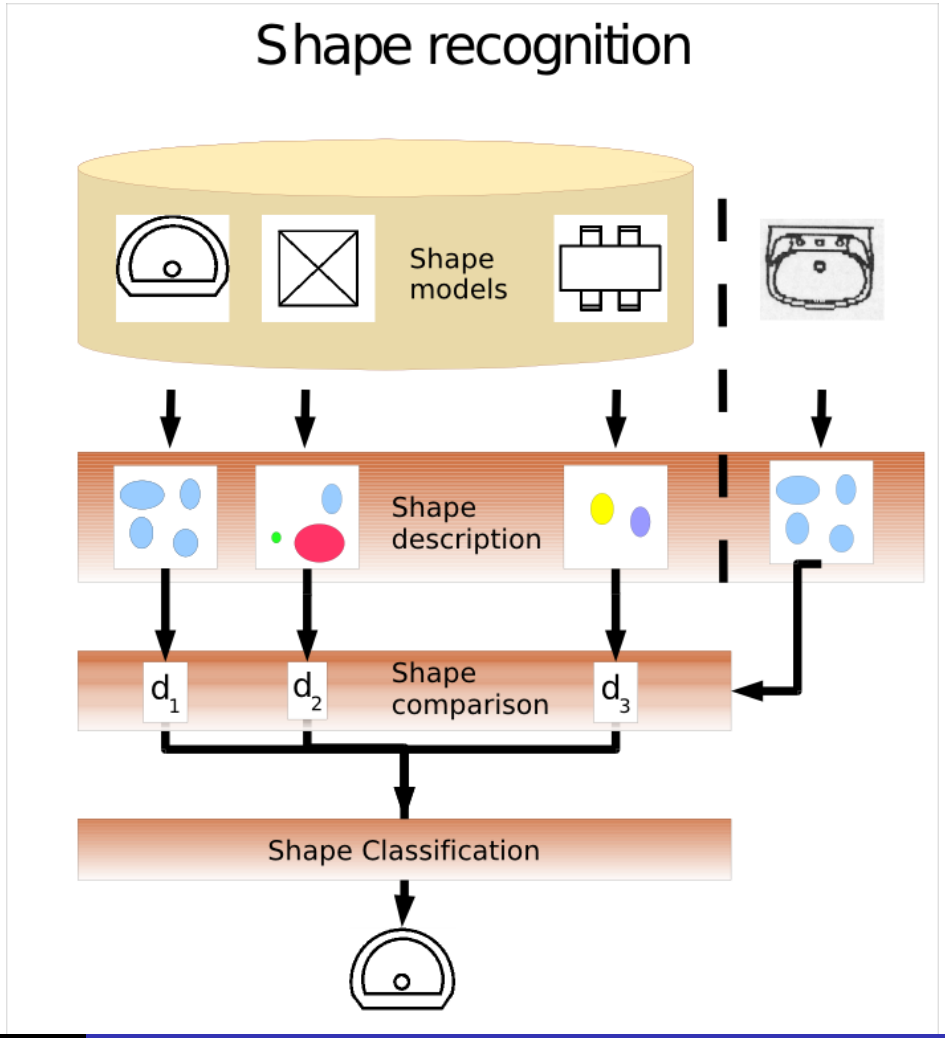


Recognition Process



Recognition Process





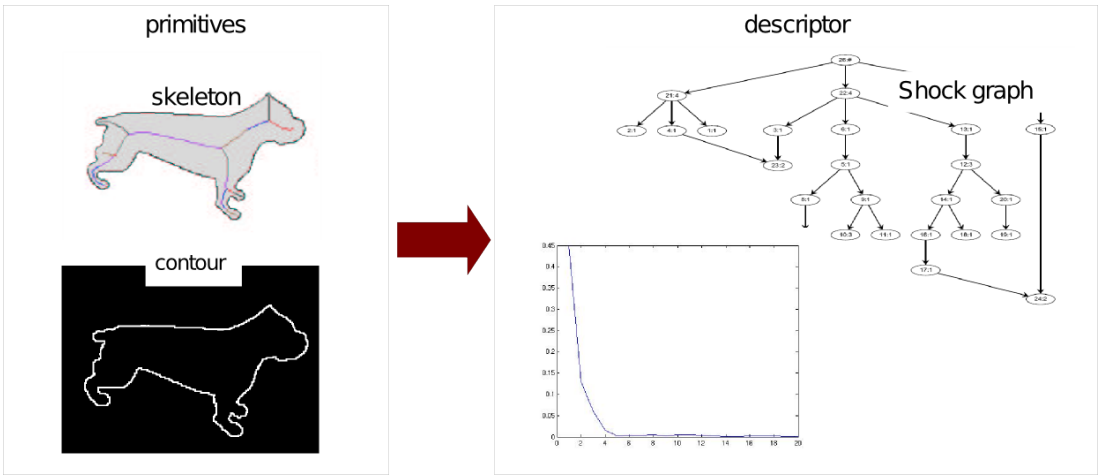
Shape Descriptors

Definitions

A Feature Extraction Method (FEM) is a map: $D : X \rightarrow Y$ such that:

The elements $x \in X$ are primitives.

For any $A \subset X$, $y_A = D(A)$ is a descriptor.



Review of Shape descriptors

Primitives:

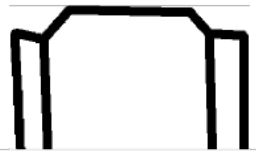
- 1D (contour):
  - Fourier.
  - Stochastic: Autoregressive methods.
  - Curvature: Curvature Scale space (CSS).
  - Geometric invariants.
- 2D (region):
  - Polar: Polar Fourier, Fourier-Mellin, Radon/Hough, Zernike moments, Angular Radial transform (ART), **ridgelets**.
  - Moment-based: geometric, Zernike, Legendre.
  - Local norm based: R-Signature, Zoning, **LNR**.

Descriptors:

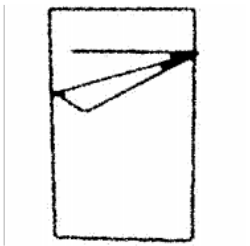
- Multiresolution:
  - Space scale: CSS
  - MRA: wavelets, **ridgelets**.
- Structural:
  - Graph-based.
  - Grammar-based.

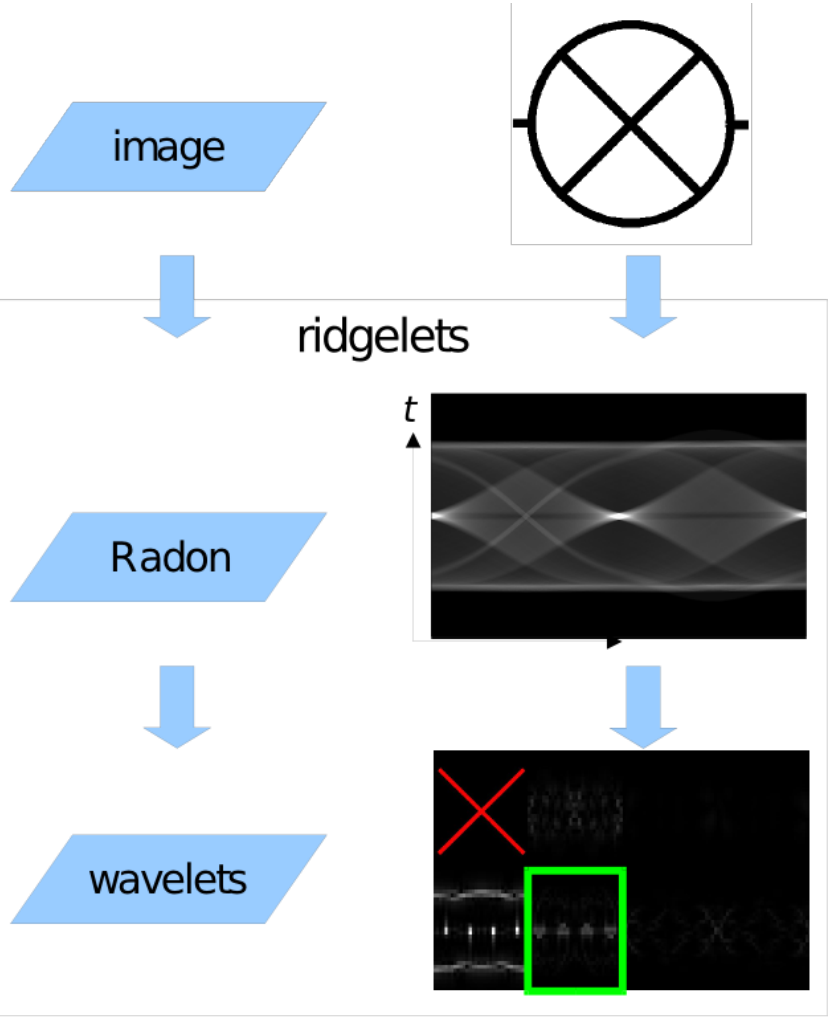
Ridgelets transform

Document degradation and vectorial distortion make difficult vectorization process

We should look for a different way to detect linear singularities

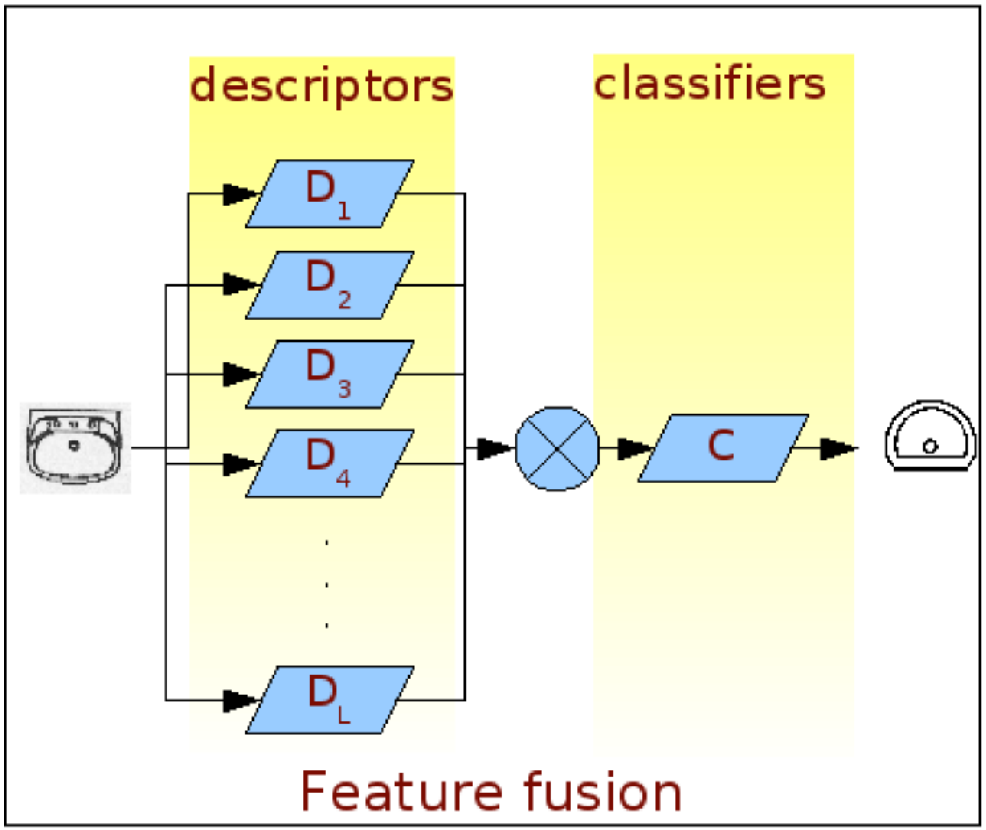


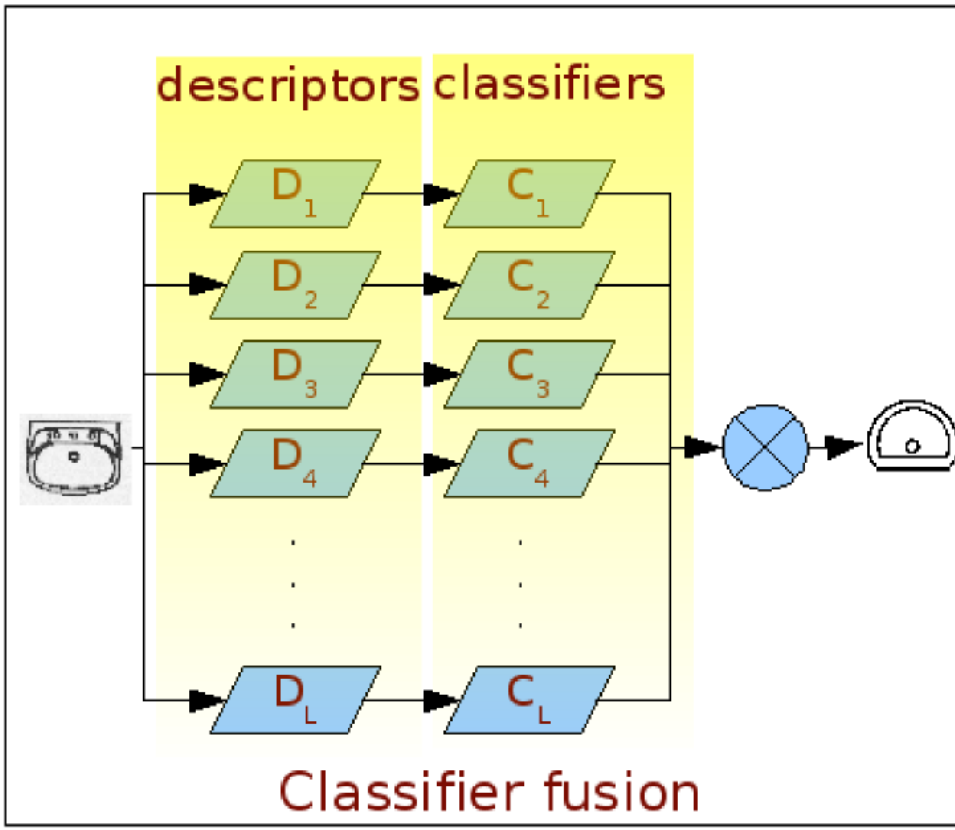



Definition of ridgelets descriptors: multiresolution, 2D, polar and

Combination Schemes

One descriptor is not usually enough.





Classifier Fusion: review

Given J Classes: $\{\omega_1, \ldots, \omega_J\}$ and L classifiers for each class, which is the best way to combine them?

- Bayesian approach: Classifiers return conditional probability: $Pr(X_l|\omega_j)$ [Kitler et al. 1998]

$$Pr(\omega_j|X_1 \ldots, X_L) = \frac{Pr(\omega_j)\prod_l Pr(X_l|\omega_j)}{Pr(X_1, \ldots, X_L)}$$

which explains some clasic combiantion rules:

$$\frac{1}{L}\prod_l Pr(X_l|\omega_j) \quad \frac{1}{L}\sum_l Pr(X_l|\omega_j) \qquad \max_l Pr(X_l|\omega_j)$$

Classifier Fusion: review (contd.)

- Logistic regression. The probability $Pr(\omega_j|X_1,\ldots,X_L)$ is fitted by an additive model ofthe **logit** function:

$$log\frac{Pr(\omega_j|X_1,\ldots,X_L)}{1-Pr(\omega_j|X_1,\ldots,X_L)} = \sum_l \alpha_l Pr(\omega_j|X_l)$$

  - Boosting algorithms: Friedman et al. [1998].
  - Generalization of the Borda Count method: Ho et al. [1994].

Notation and Simplifications

2 simplifications

- 2 class classifiers: $\{-1, 1\}$
- linear combination: $\sum_l \alpha_l C_l$

Name	Notation	Domain	Meaning
Shape	S	Ω	the shape to recognize
Label	Y	$\{-1, 1\}$	the class of shapes
Descriptor	$X = FEM(S)$	$\mathbb{R}^d$	the descriptor computed from shapes
Prediction	$Z = C(X)$	$\mathbb{R}$	the classifier output
Validation	$U = YZ$	$\mathbb{R}$	the validity of the prediction

Linear Combination of Classifiers

Problem:

With the precedent definitions of r.v. the problem of finding the optimal linear combination rule is expressed as the optimization of the following objective function:

$$\alpha_{optimal} = \arg\min_{\alpha} Pr(\sum_{I} \alpha_I U_I < 0|S)$$

with constraints:

$$\alpha_I > 0 \quad \text{for all } I$$

$$\sum_{I} \alpha_I = 1$$

IN and *DN* methods

- Denote $U = (U_1, \ldots, U_L)$.
- μ and Σ, respectively, mean and covariance matrix of U (σ_l^2 variance of U_l)
- $A = \sum_l \alpha_l^{\mathcal{N}}$ and $B = \sum_l \alpha_l^{D}$.

IN and *DN* methods

- Denote $U = (U_1, \ldots, U_L)$.
- μ and Σ, respectively, mean and covariance matrix of U (σ_l^2 variance of U_l)
- $A = \sum_l \alpha_l^{\mathcal{N}}$ and $B = \sum_l \alpha_l^{D}$.

U_l are conditional independents

- $\alpha_l^D = \mu_l$ if $\sigma_l \approx 0$.
- $\alpha_l^{\mathcal{N}} = \frac{\mu_l}{\sigma_l}$, otherwise.

$$\alpha_{op} = \lambda_{\mathcal{N}} \alpha^{\mathcal{N}} + \lambda_D \alpha^D$$

$$\lambda = \begin{cases} (\frac{A-B}{A^2}, \frac{1}{A}) & \text{if } A > B \\ (0, \frac{1}{B}) & \text{if } A \leq B \end{cases}$$

IN and *DN* methods

- Denote $U = (U_1, \ldots, U_L)$.
- μ and Σ, respectively, mean and covariance matrix of U (σ_l^2 variance of U_l)
- $A = \sum_l \alpha_l^{\mathcal{N}}$ and $B = \sum_l \alpha_l^{D}$.

U_l are conditional independents

- $\alpha_l^D = \mu_l$ if $\sigma_l \approx 0$.
- $\alpha_l^{\mathcal{N}} = \frac{\mu_l}{\sigma_l}$, otherwise.

$$\alpha_{op} = \lambda_{\mathcal{N}} \alpha^{\mathcal{N}} + \lambda_D \alpha^D$$

$$\lambda = \begin{cases} (\frac{A-B}{A^2}, \frac{1}{A}) & \text{if } A > B \\ (0, \frac{1}{B}) & \text{if } A \leq B \end{cases}$$

U_l are dependents to each other

Minimize the object function:

$$\phi(\alpha) = \left\langle \frac{\alpha}{\alpha^t \Sigma \alpha}, \mu \right\rangle$$

Subject to the constraints:

$$\begin{cases} \alpha_l > 0 & \text{for all } l \\ \sum_l \alpha_l = 1 \end{cases}$$

Correlation Filters

- Collaboration with Prof. Djemel Ziou (univ Sherbrooke, Canada) and S. Tabbone (univ. Nancy 2 - LORIA, France)
- Idea: Apply correlation filters to CBIR problems.
- Related work: Face and Object Identification, **MACE** filters

Correlation Filters

- Collaboration with Prof. Djemel Ziou (univ Sherbrooke, Canada) and S. Tabbone (univ. Nancy 2 - LORIA, France)
- Idea: Apply correlation filters to CBIR problems.
- Related work: Face and Object Identification, **MACE** filters

Variational approach: Main Idea

K semantic classes: $1, \ldots, K$; $v^{i,k}$ image collection, h_k correlation filters.

$$\max \frac{\int_{-w}^{w} v^{i,k} \circledast h_k}{\int v^{i,k} \circledast h_k}$$

Outline

Example of applications

- Use of shape descriptors and Classifier fusion methods to Graphics (segmented) recognition problems
- Use of low level features and SVM classifiers to Thai-English script identification
- Use of Correlation filters to CBIR problems: symbol spotting.

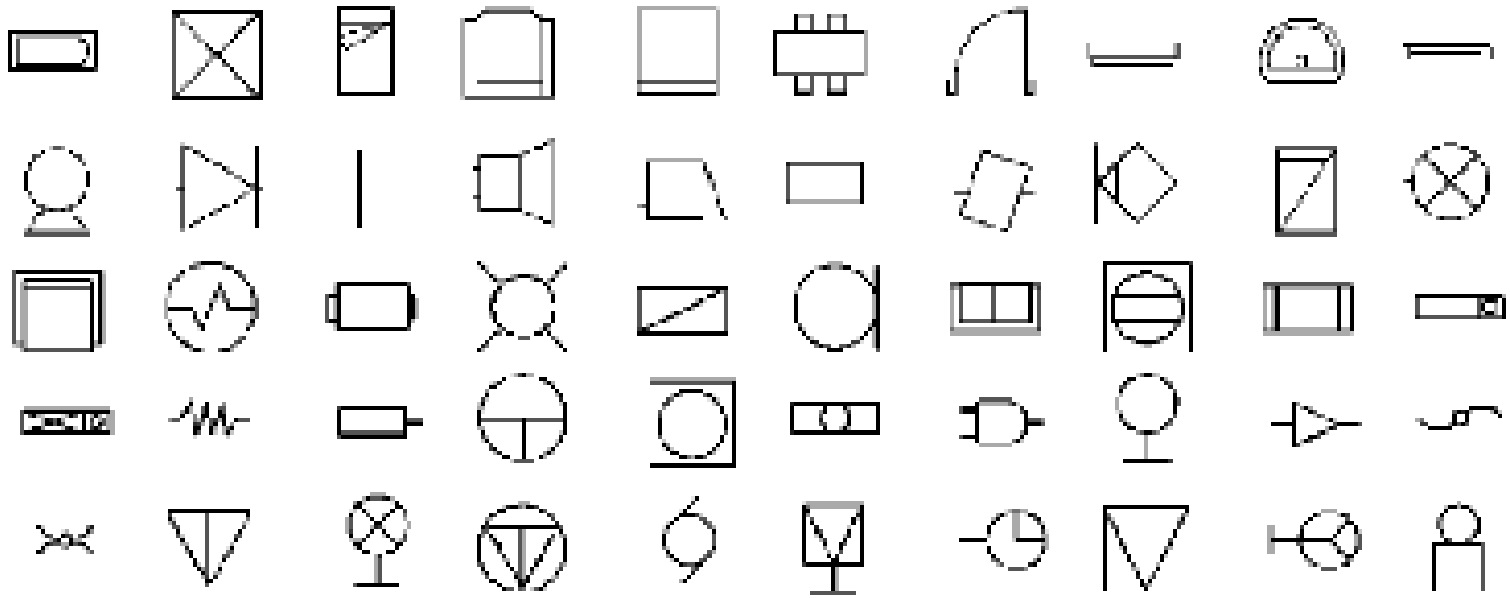
Example of applications

- Use of shape descriptors and Classifier fusion methods to Graphics (segmented) recognition problems
- Use of low level features and SVM classifiers to Thai-English script identification
- Use of Correlation filters to CBIR problems: symbol spotting.
- Use of Genetic algorithms for feature selection
- Shape context and Video Google for symbol spotting

Graphics Recognition

Validation on benchmark datasets:

- Graphics (GREC):

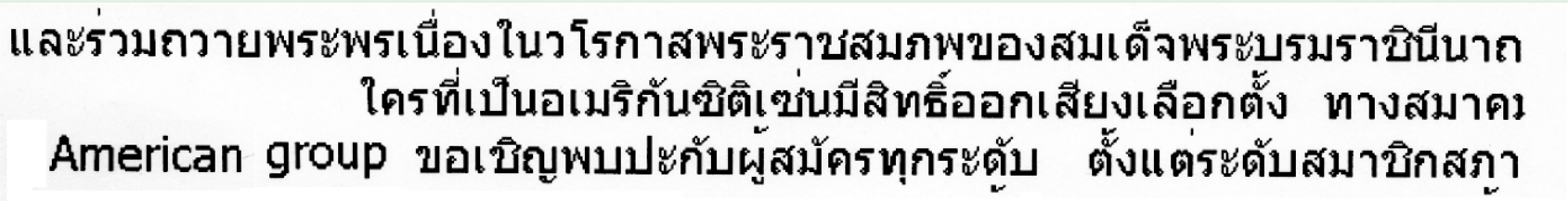


- and digits (MNIST):

Script Identification

- Collaboration with Sukalpa Chanda, PhD candidate, and Umapada Pal (ISI)
- Use of 7 features to train two-class SVM

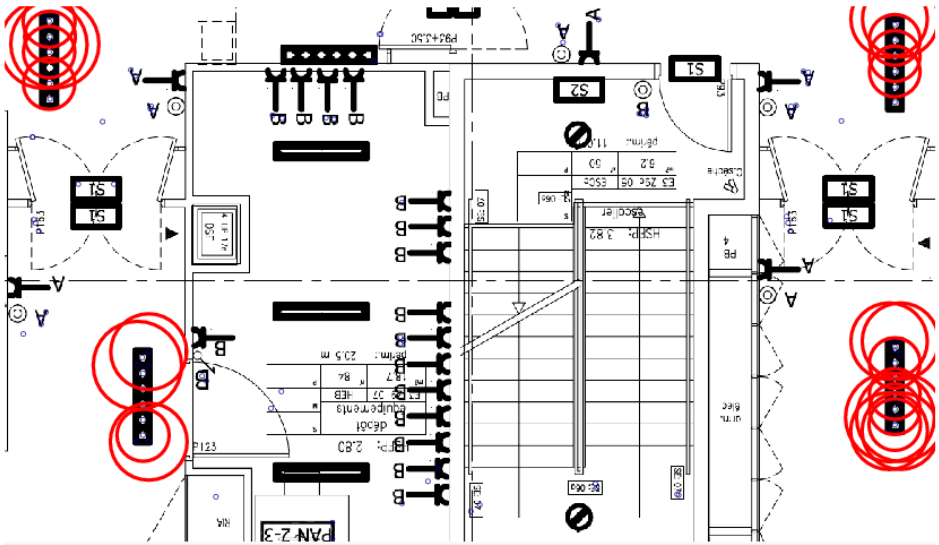
Exemple



Symbol Spotting

- Thi-Oanh Nguyen, PhD candidate, and S. Tabbone (LORIA)
- SCIP: Shape Context of Interest Points and Video Google.
- Symbol spoting: detect and localise non-segmented symbols

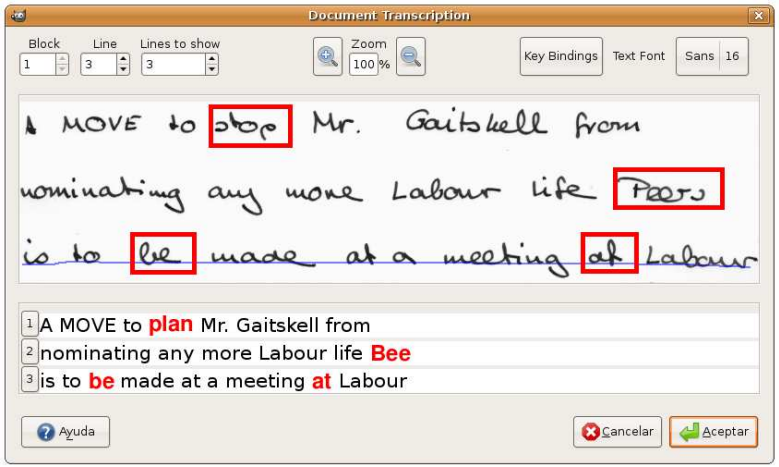
Exemple



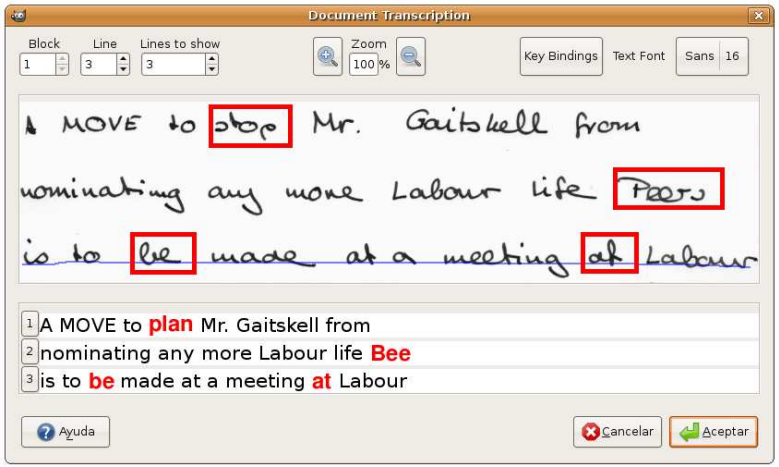
Feature Selection

- Hassan Chouaib, PhD candidate, S. Tabbone (LORIA) and Prof. N. Vincent (Paris 5)
- Combine Genetic Algorithms (GA) and boosting-based classifiers.
- Use of R-Signature, Zernike moments and pixel images as shape descriptors.

GiDoc


Document Transcription
Block
Line
Lines to show
Zoom
100 %
Key Bindings
Text Font
Sans 16
A MOVE to plan Mr. Gaitskell from
nominating any more Labour life Bee
is to be made at a meeting at Labour
1 A MOVE to plan Mr. Gaitskell from
2 nominating any more Labour life Bee
3 is to be made at a meeting at Labour
Ayuda
Cancelar
Aceptar

GiDoc



- Offers an interactive environment
- Usefull for groundtruthing
- based on gimp
- HMM are estimated by the means of HTK engine
- Language models estimated by SRILM

The End

Before concluding...

Any question?