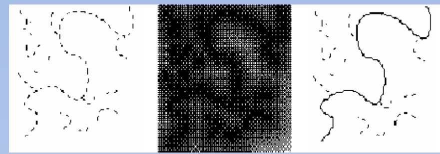
Shape context

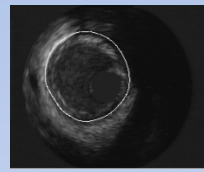
Centro de Visión por Computador,
Departament de Matemàtica Aplicada i Anàlisi,
Universitat de Barcelona

Deformable models

$$U(p) = inf_{Q_{L\_p}}\{\int_Q P(Q) + \alpha|\frac{\delta Q}{\delta s}|^2 ds\}$$



$$E_{snake} = \int_0^1 E_{int}(u(s)) + E_{ext}(u(s))\, ds.$$



$$Q_t = g(I)(c+k)\vec{n} - (\nabla g, \vec{n})\vec{n}$$

$$E(L) = \sum_{p\in\mathcal{P}} D_p(L_p) + \sum_{(p,q)\in\mathcal{N}} V_{p,q}(L_p, L_q),$$

All four models are based only on local control on the geometric shape.

Representation

- Object as set of parts
  - Generative representation
- Model:
  - Appearance of parts
  - Relative locations between parts
- Issues:
  - How to model location and mutual relations
  - How to represent appearance
  - Sparse or dense (pixels or regions)
  - How to handle occlusion/clutter

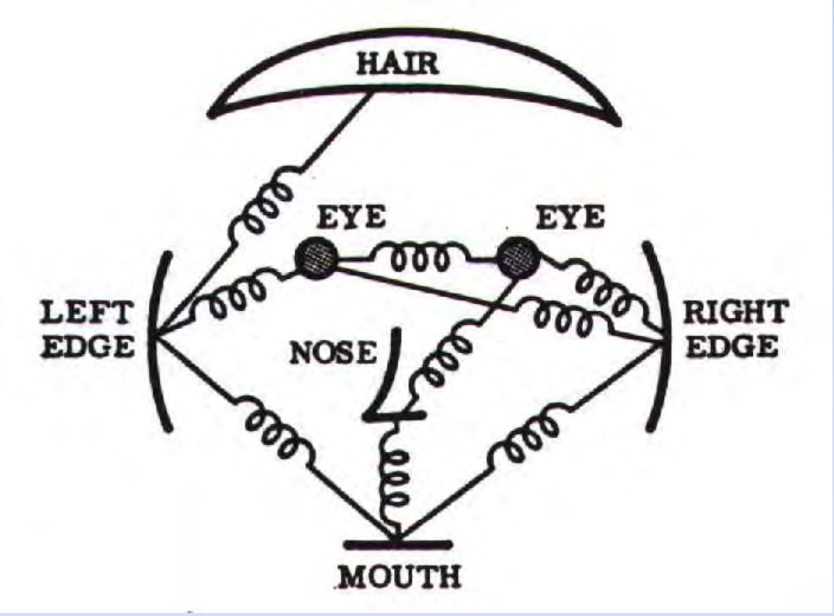


Figure from [Fischler73]

Representation

- Object as set of parts
  - Generative representation
- Model:
  - Appearance of parts
  - Relative locations between parts
- Issues:
  - How to model location and mutual relations
  - How to represent appearance
  - Sparse or dense (pixels or regions)
  - How to handle occlusion/clutter

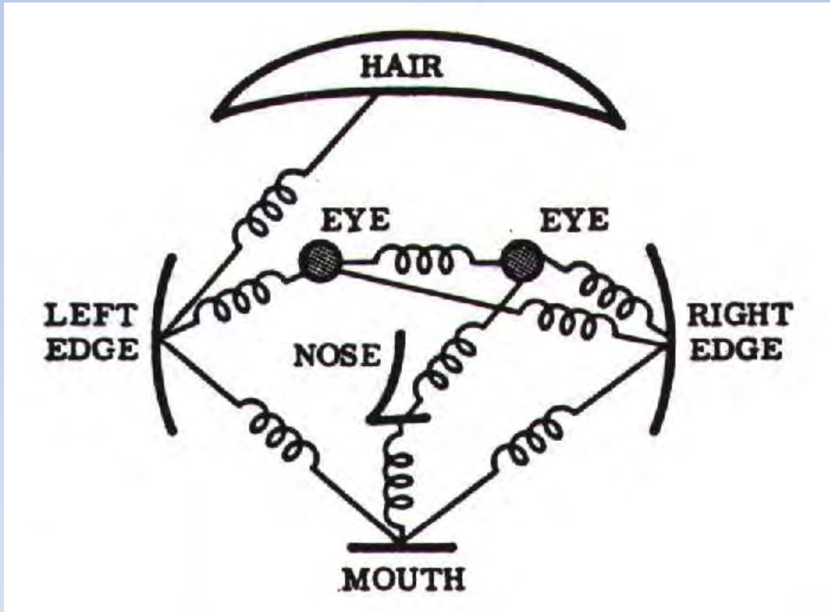


Figure from [Fischler73]

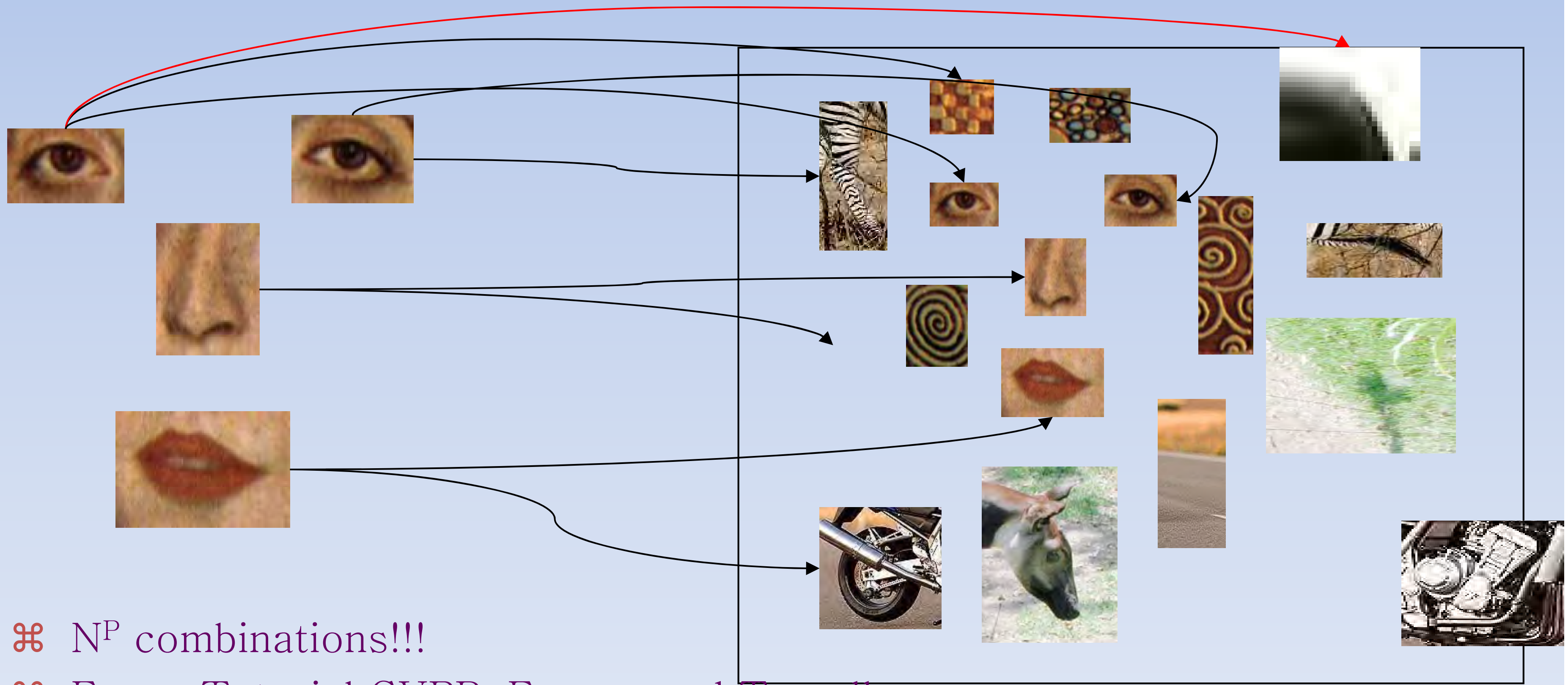
Sparse representation

+ Computationally tractable (10^5 pixels → 10^1 -- 10^2 parts)
+ Avoid modeling global variability
+ Success in specific object recognition

- Throw away most image information
- Parts need to be distinctive to separate from other classes

The correspondence problem

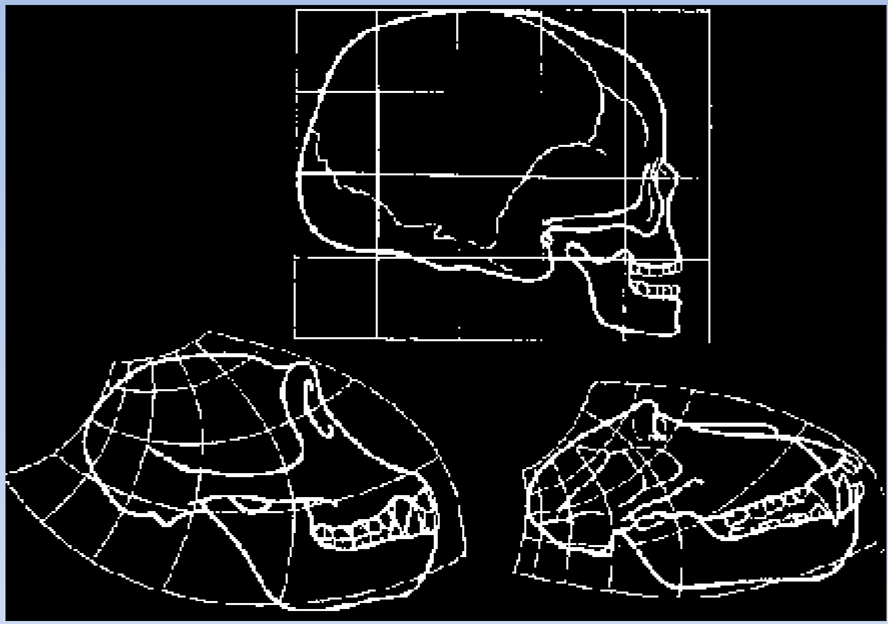
- Model with P parts
- Image with N possible locations for each part

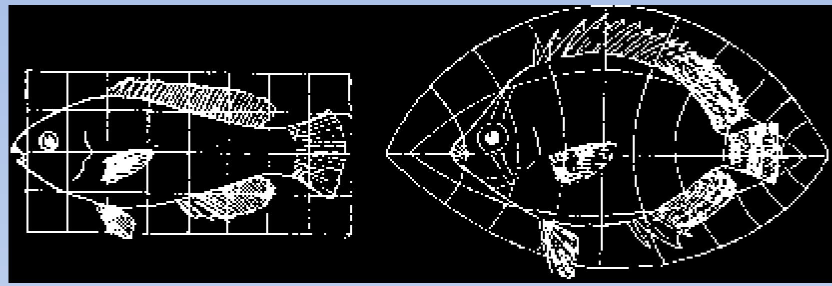


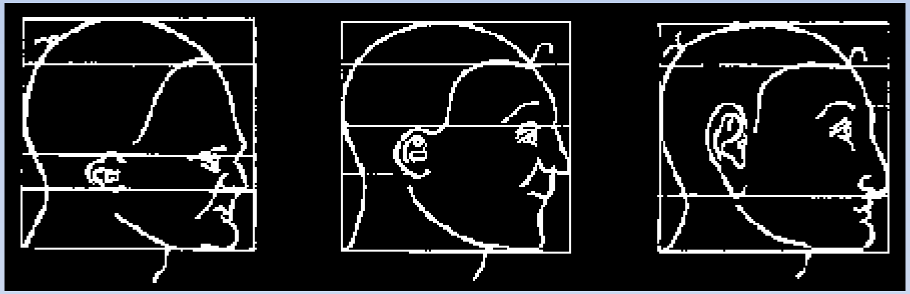
⌘ N^P combinations!!!

⌘ From: Tutorial CVPR, Fergus and Torralba.

Biological Shape

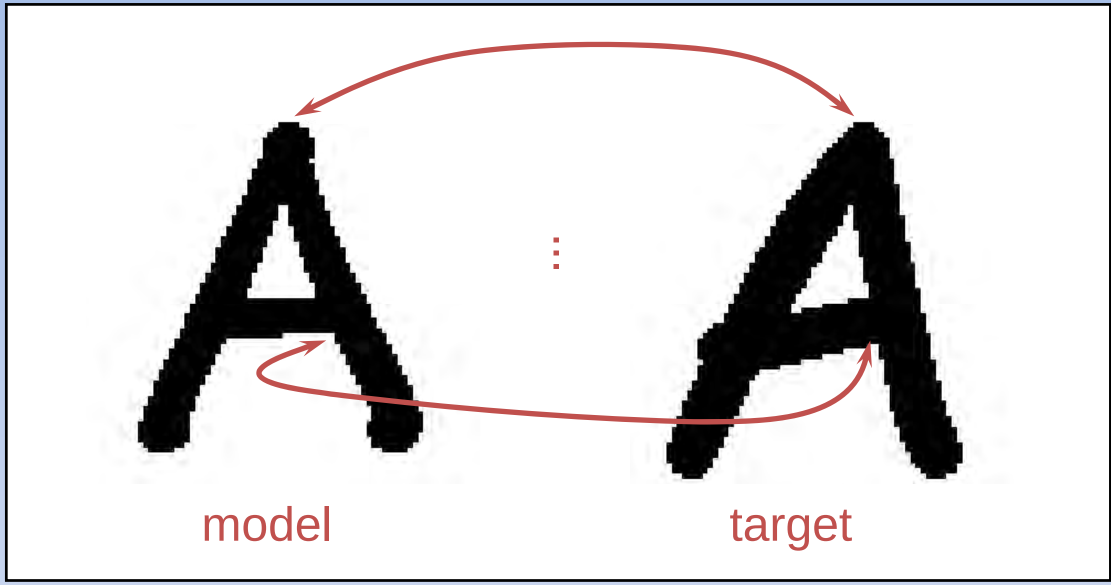






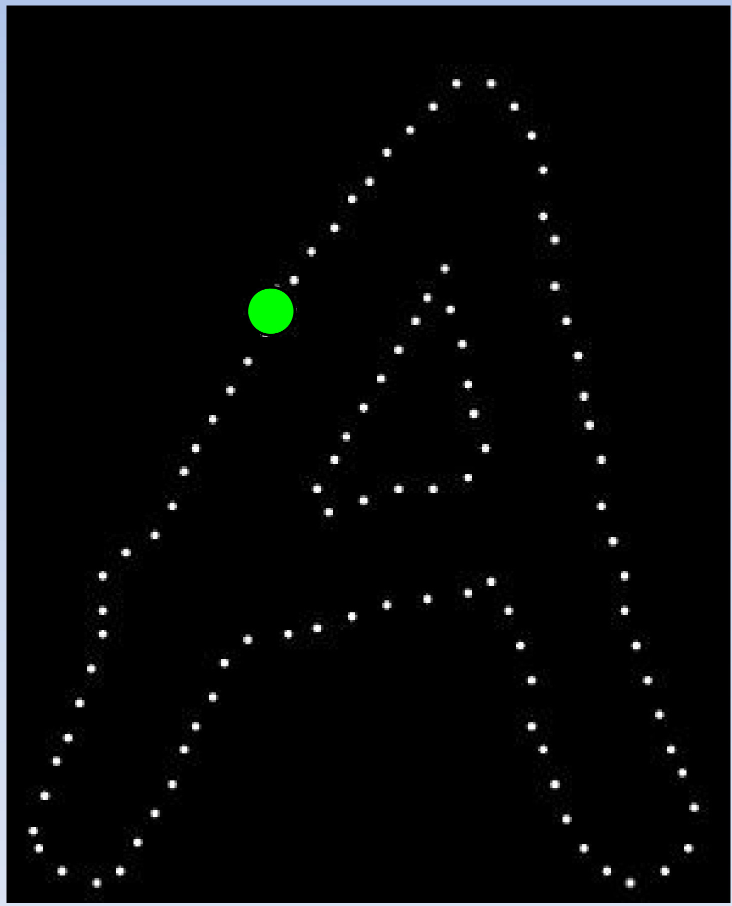
- D'Arcy Thompson: On Growth and Form, 1917
  - studied transformations between shapes of organisms

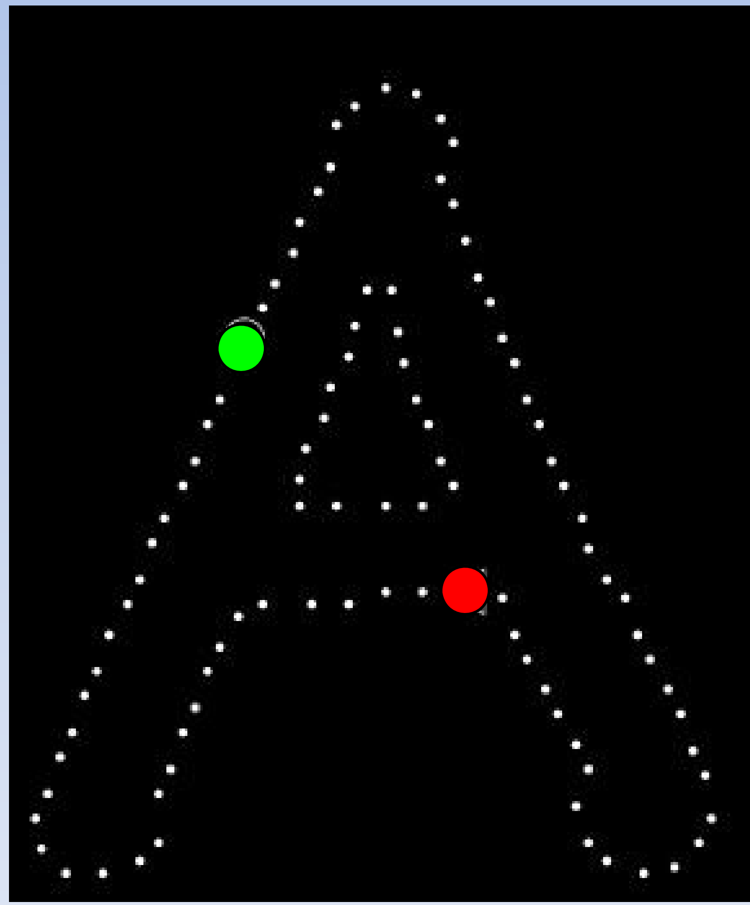
Matching Framework



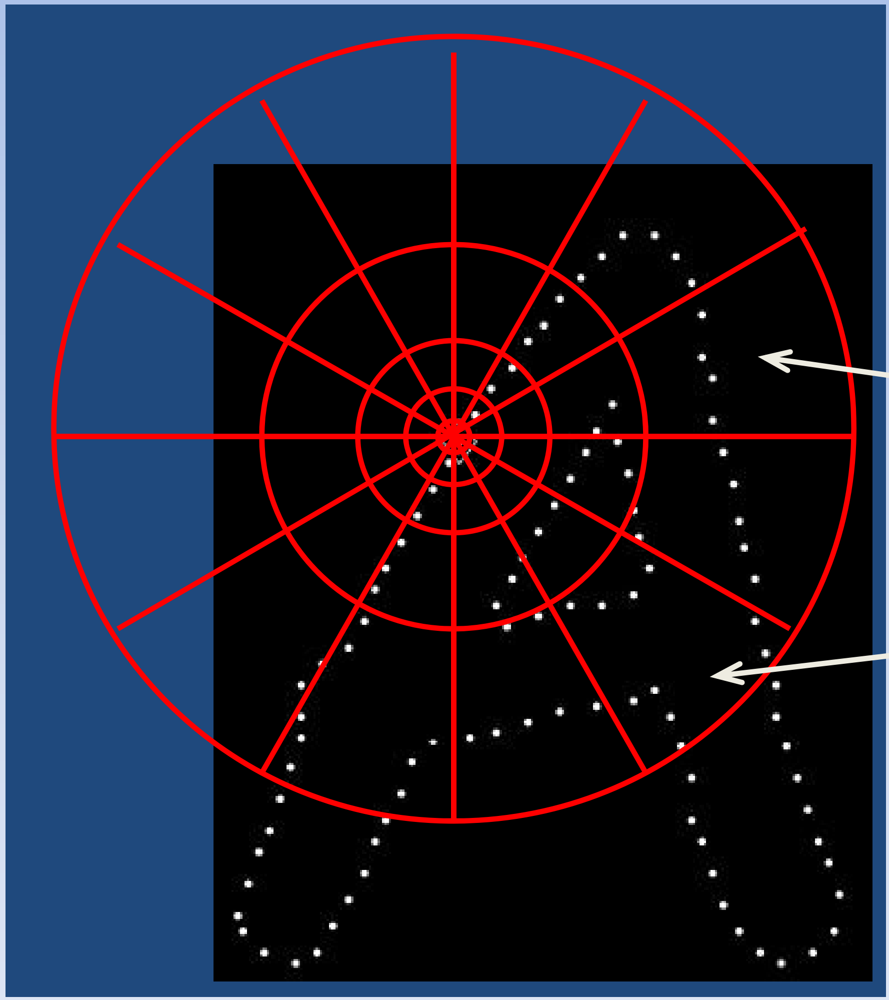
- Find correspondences between points on shape
- Estimate transformation & measure similarity
- From: Greg Mori, Jitendra Malik: Recovering 3D Human Body Configurations Using Shape Contexts. IEEE Trans. Pattern Anal. Mach. Intell. 28(7): 1052-1062 (2006)

Comparing Pointsets





Shape Context



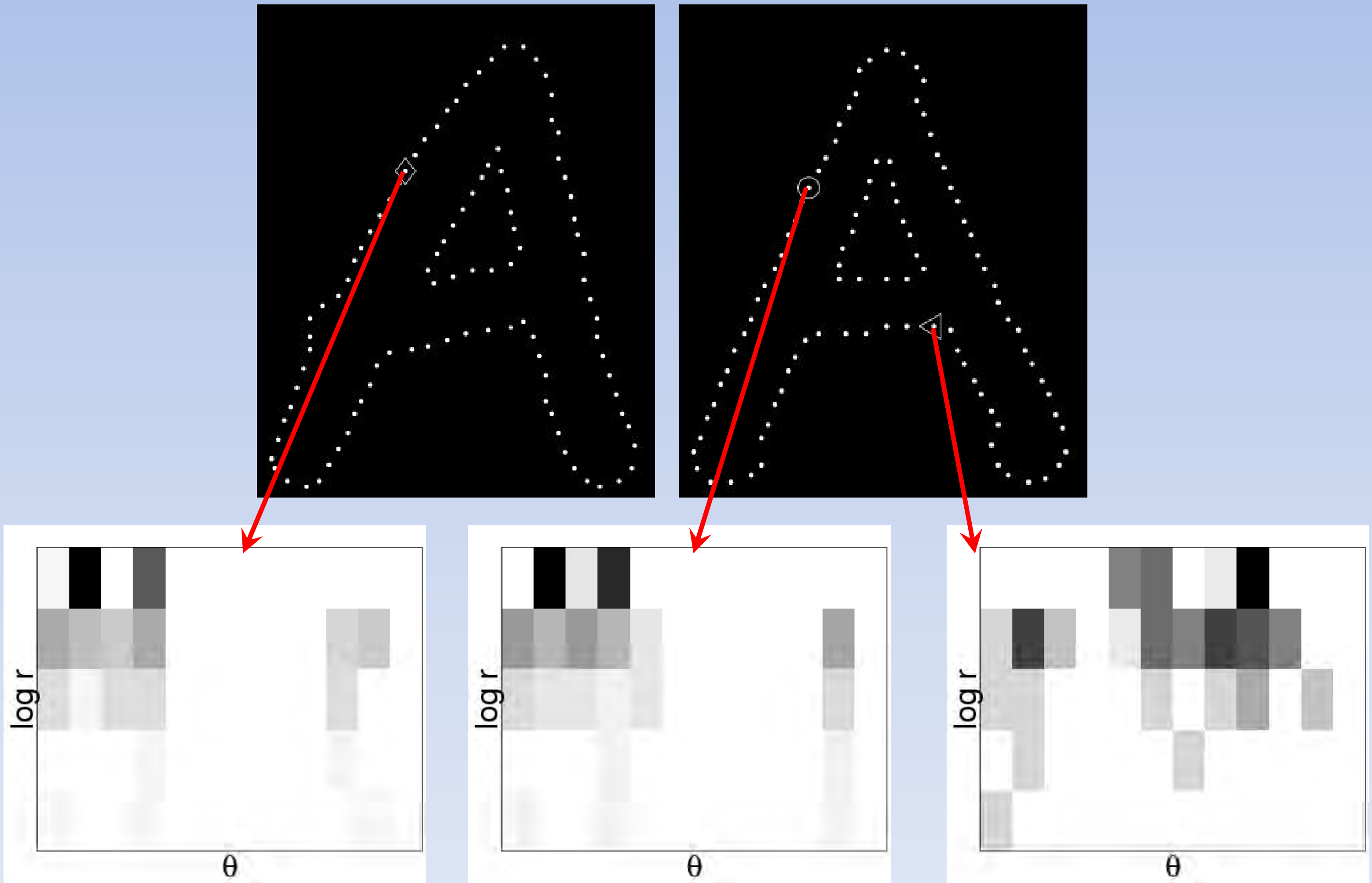
Count the number of points inside each bin, e.g.:

Count = 4

Count = 10

☞ Compact representation of distribution of points relative to each point

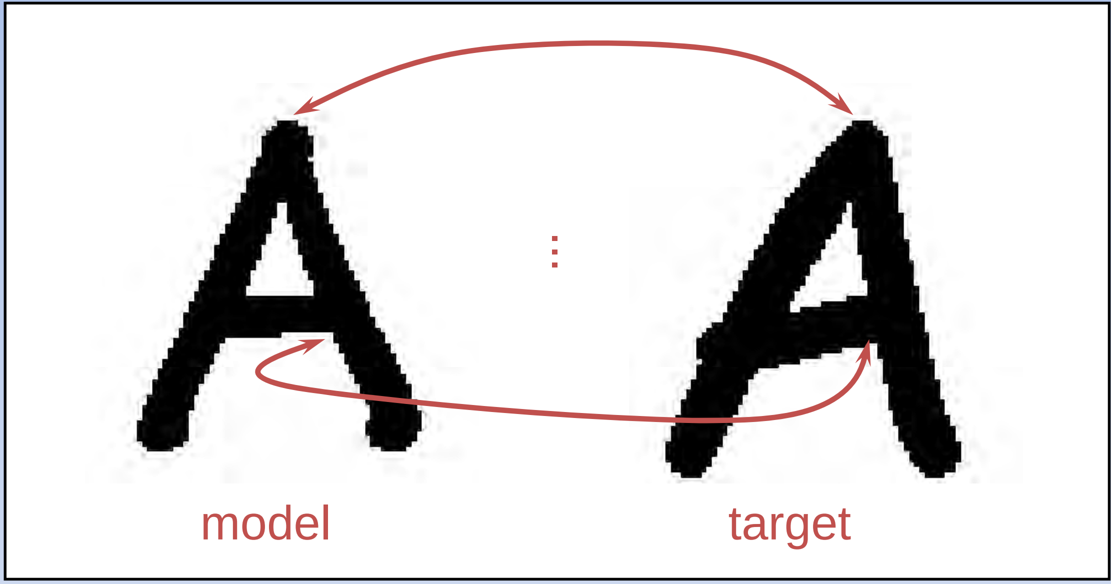
Shape Context



Shape Contexts

- Invariant under translation and scale
- Can be made invariant to rotation by using local tangent orientation frame
- Tolerant to small affine distortion
  – Log-polar bins make spatial blur proportional to *r*

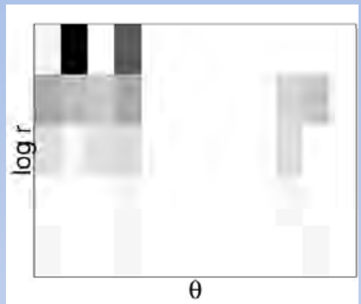
Matching Framework



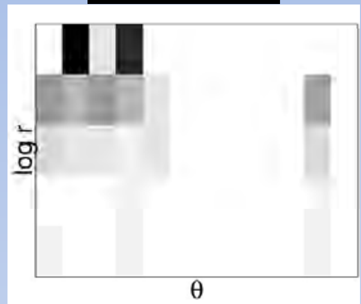
- ⌘ Find correspondences between points on shape
- ⌘ Estimate transformation & measure similarity

Comparing Shape Contexts

$h_i(k)$

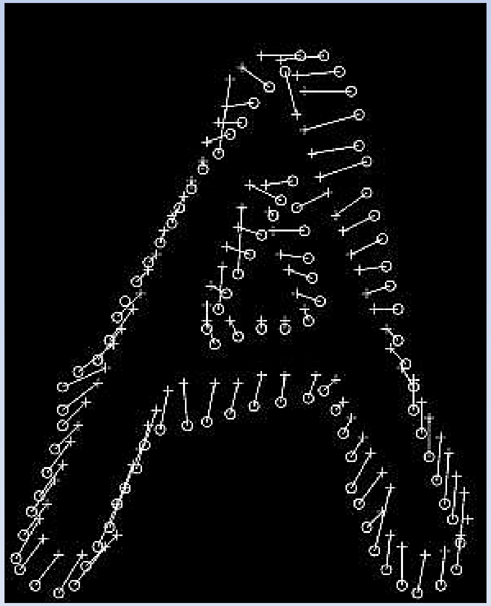


$h_j(k)$



Compute matching costs using Chi Squared distance:

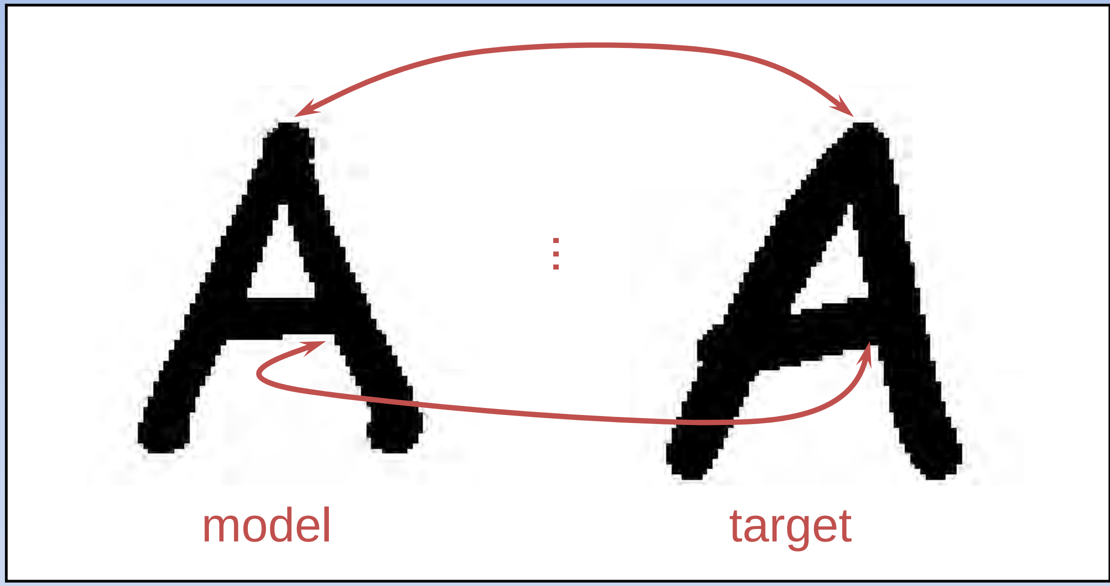
$$C_{ij} = \frac{1}{2} \sum_{k=1}^{K} \frac{[h_i(k) - h_j(k)]^2}{h_i(k) + h_j(k)}$$



Recover correspondences by solving linear assignment problem with costs C_{ij}

[Jonker & Volgenant 1987]

Matching Framework



- Find correspondences between points on shape
- Estimate transformation & measure similarity

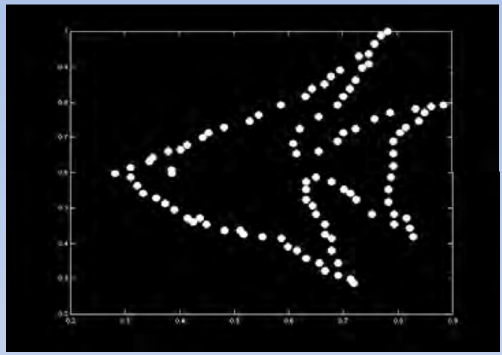
Thin Plate Spline Model

- Minimizes bending energy:

$$I_f = \int\int_{\mathbb{R}^2} \left(\frac{\partial^2 f}{\partial x^2}\right)^2 + 2\left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 + \left(\frac{\partial^2 f}{\partial y^2}\right)^2 dxdy$$

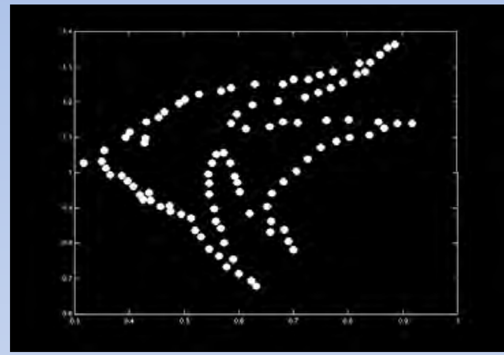
- Solve by inverting linear system
- Can be regularized when data is inexact

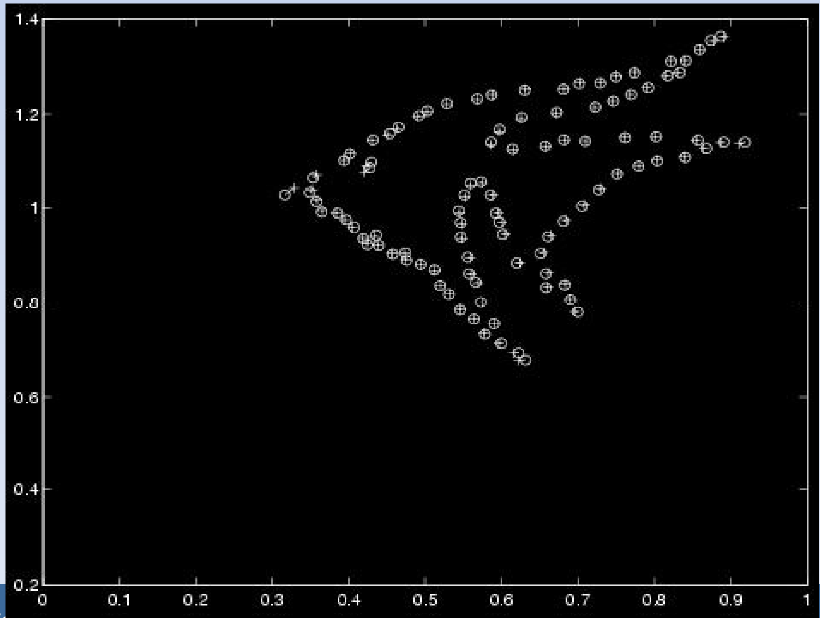
Matching Example

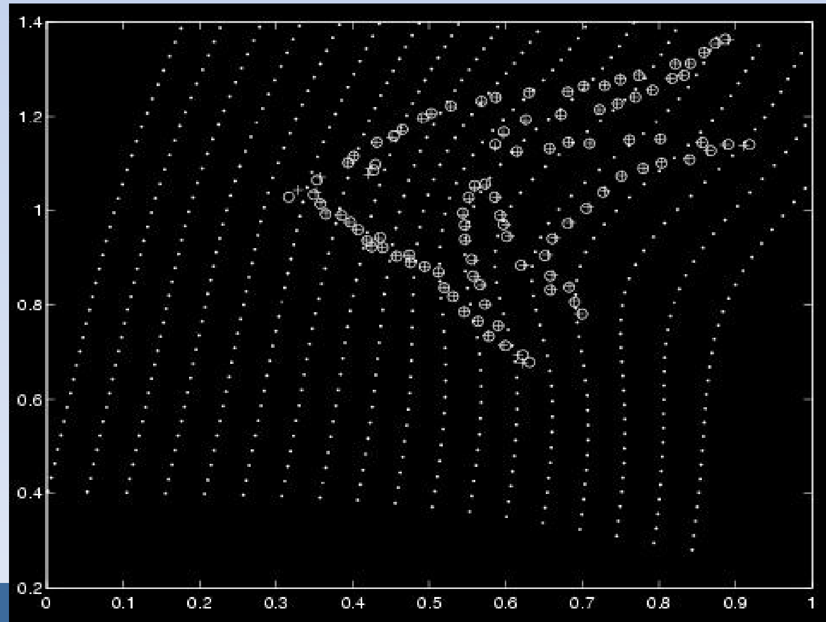


model

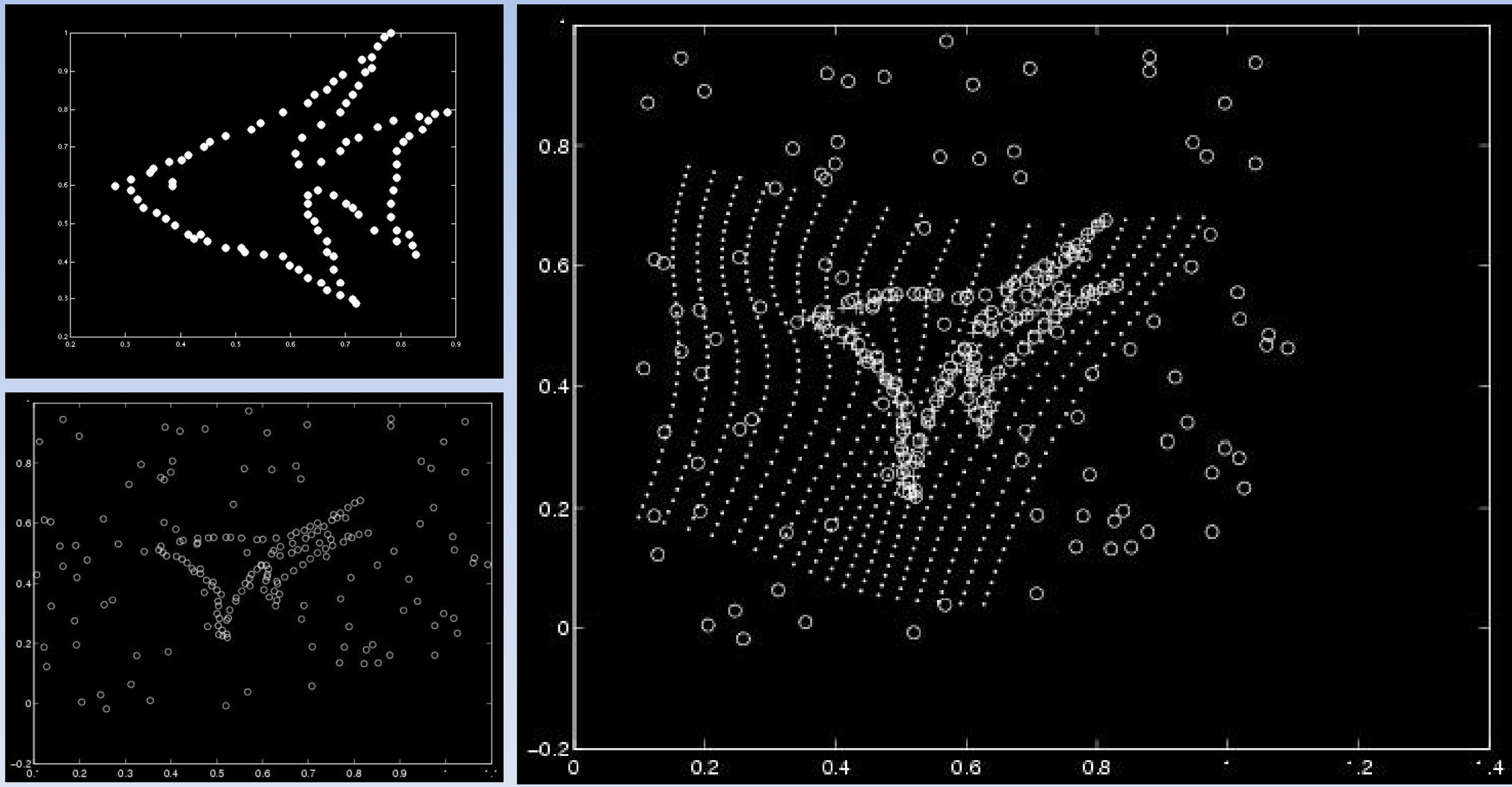
target



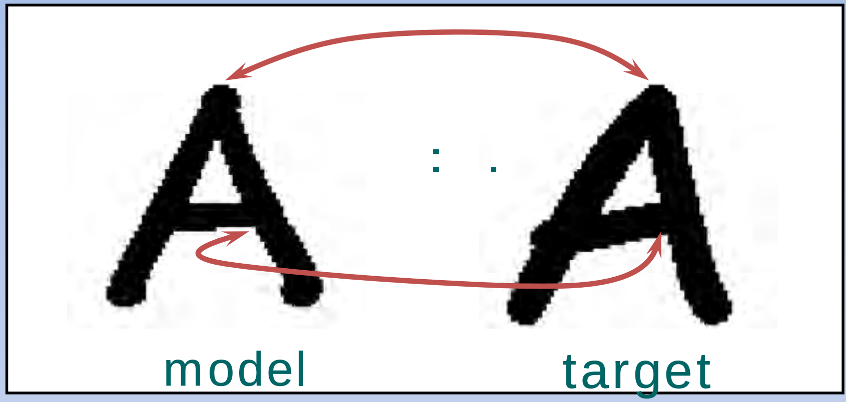


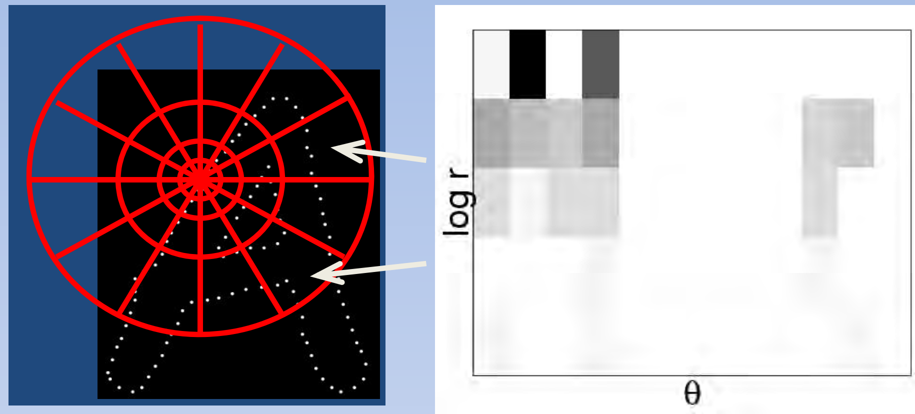


Outlier Test Example

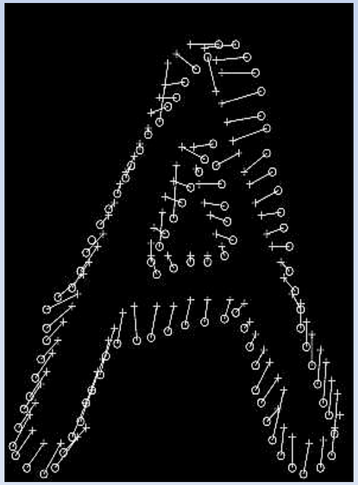


Extending features to contexts





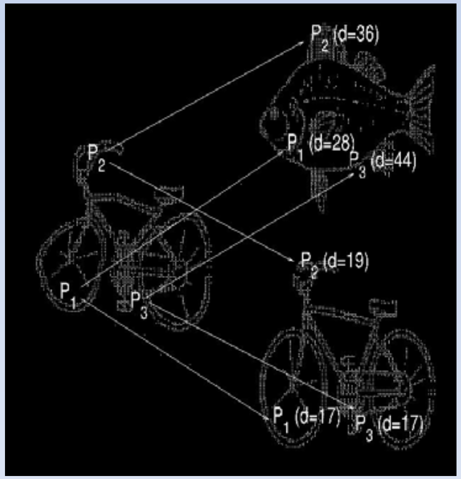
$$O_j \text{->} \{\{c_{11}, c_{12},\ldots c_{1m1}\},\{c_{21}, c_{22},\ldots\},\ldots\{c_{n1},\ldots c_{nm_n},\}\} = \{C_{1j}, C_{2j},\ldots C_{nj}\}$$



1. Landmarks correspondences based on correlograms

2. Non-rigid correspondence based on Thin-Plate Splines

3. Retrieval based on the object descriptors difference

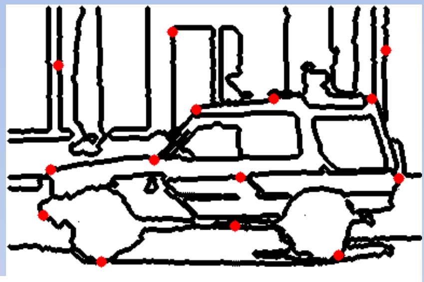


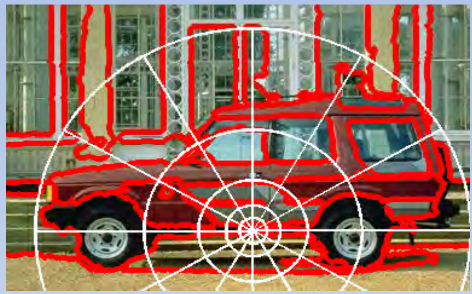
Non-rigid matching based on context

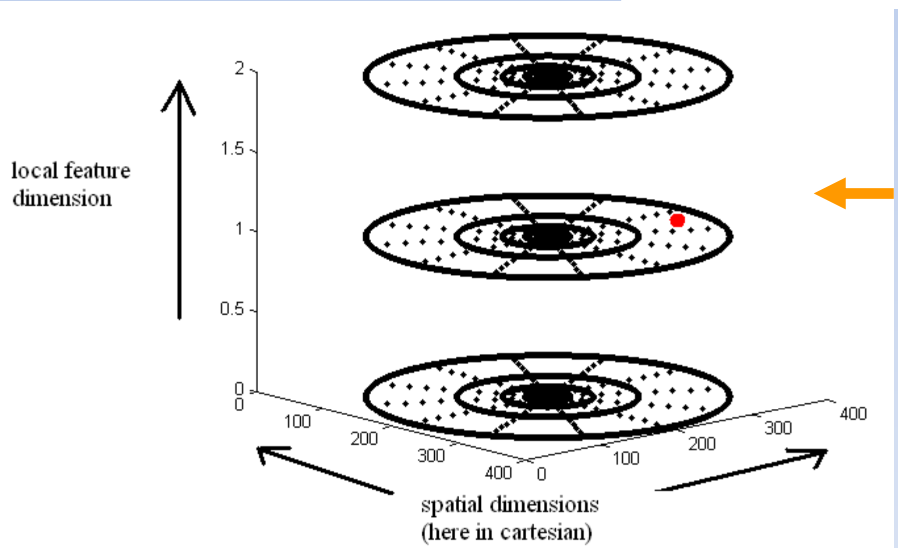
- How to choose landmarks?
- How to extend the method to real images?
- How to choose adequate local descriptors?
- How to construct compact and efficient contexts for object retrieval?
- Robust classification can help to:
    - Select the most relevant local descriptors
    - Construct compact object contexts
    - Achieve object retrieval invariant to transformations, different appearance and views of the object

The feature space

Let us consider a context consisting of local descriptors and mutual spatial relations







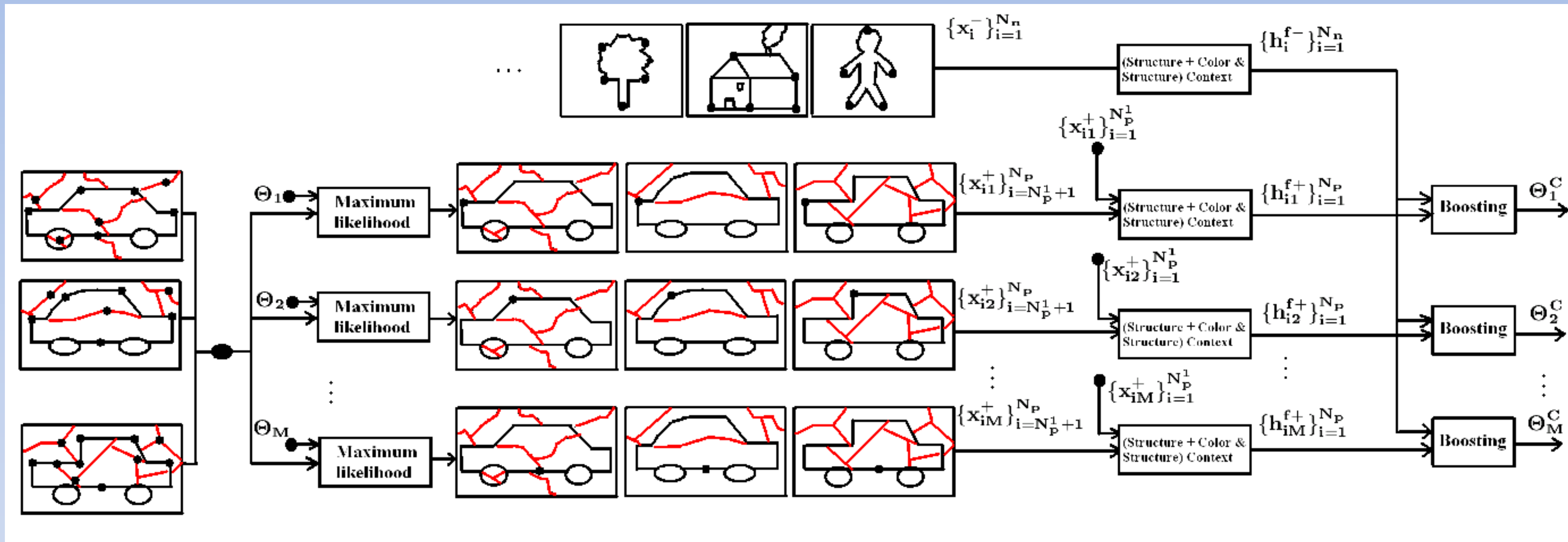
$$X = \{x_j\}_{j=1}^{M}$$

$$(p_i - x_j) \rightarrow (\alpha_{ij}, r_{ij})$$

$$v_{ij} = (\alpha_{ij}, r_{ij}, l_{i1}, l_{i2}, \ldots, l_{id})$$

From: Jaume Amores, Nicu Sebe, Petia Radeva: Fast Spatial Pattern Discovery Integrating Boosting with Constellations of Contextual Descriptors. CVPR (2) 2005: 769-774

Object descriptors



- Apply the classifiers to detect the landmarks in the non-segmented images

- Form the complete context

Terms in Similarity Score

- Shape Context difference
- Local Image appearance difference
  - orientation
  - gray-level correlation in Gaussian window
  - … (many more possible)
- Bending energy

Object Recognition Experiments

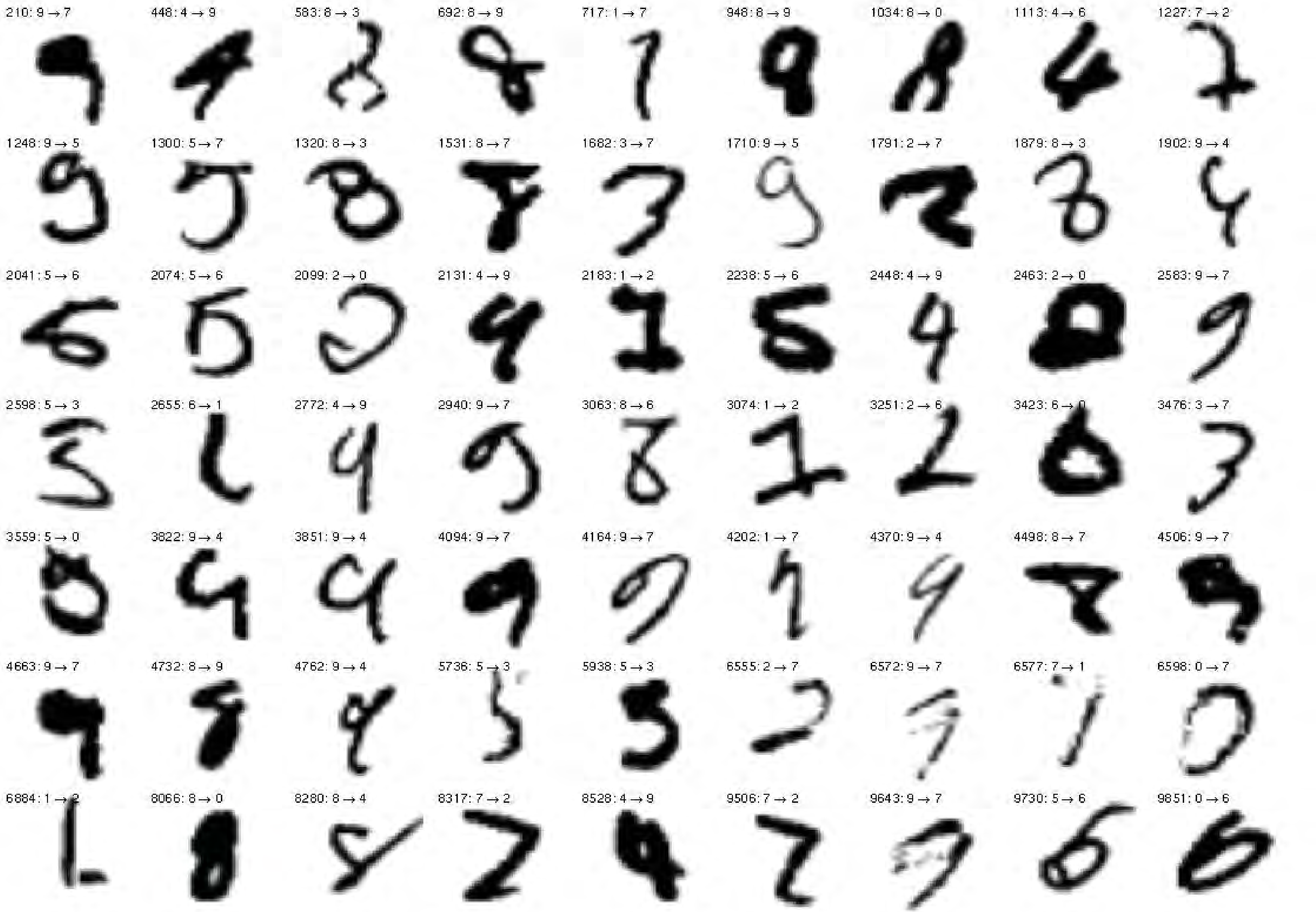
- Handwritten digits
- COIL 3D objects (Nayar-Murase)
- Human body configurations
- Trademarks

Handwritten Digit Recognition

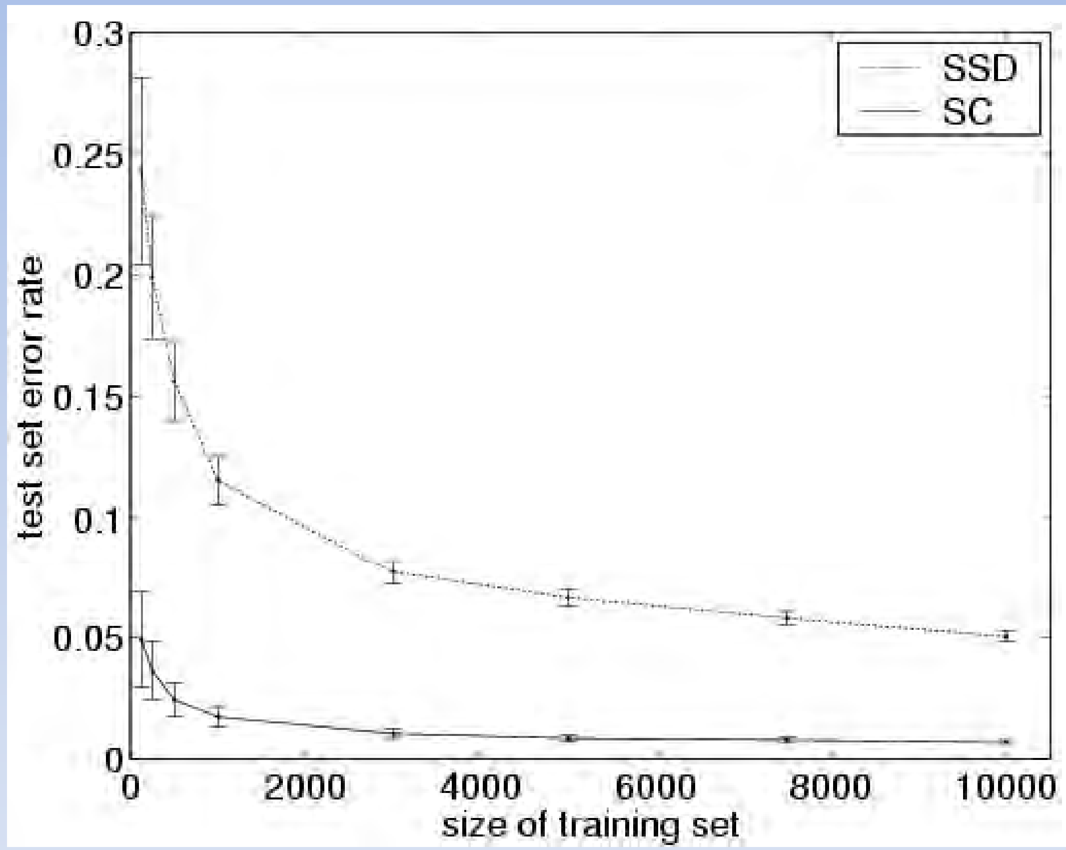
- MNIST 60 000:
  - linear: 12.0%
  - 40 PCA+ quad: 3.3%
  - 1000 RBF +linear: 3.6%
  - K-NN: 5%
  - K-NN (deskewed): 2.4%
  - K-NN (tangent dist.): 1.1%
  - SVM: 1.1%
  - LeNet 5: 0.95%

- MNIST 600 000 (distortions):
  - LeNet 5: 0.8%
  - SVM: 0.8%
  - Boosted LeNet 4: 0.7%

- MNIST 20 000:
  - K-NN, Shape Context matching: 0.63%


210: 9 → 7
448: 4 → 9
583: 8 → 3
692: 8 → 9
717: 1 → 7
948: 8 → 9
1034: 8 → 0
1113: 4 → 6
1227: 7 → 2
1248: 9 → 5
1300: 5 → 7
1320: 8 → 3
1531: 8 → 7
1682: 3 → 7
1710: 9 → 5
1791: 2 → 7
1879: 8 → 3
1902: 9 → 4
2041: 5 → 6
2074: 5 → 6
2099: 2 → 0
2131: 4 → 9
2183: 1 → 2
2238: 5 → 6
2448: 4 → 9
2463: 2 → 0
2583: 9 → 7
2598: 5 → 3
2655: 6 → 1
2772: 4 → 9
2940: 9 → 7
3063: 8 → 6
3074: 1 → 2
3251: 2 → 6
3423: 6 → 0
3476: 3 → 7
3559: 5 → 0
3822: 9 → 4
3851: 9 → 4
4094: 9 → 7
4164: 9 → 7
4202: 1 → 7
4370: 9 → 4
4498: 8 → 7
4506: 9 → 7
4663: 9 → 7
4732: 8 → 9
4762: 9 → 4
5736: 5 → 3
5938: 5 → 3
6555: 2 → 7
6572: 9 → 7
6577: 7 → 1
6598: 0 → 7
6884: 1 → 2
8066: 8 → 0
8280: 8 → 4
8317: 7 → 2
8528: 4 → 9
9506: 7 → 2
9643: 9 → 7
9730: 5 → 6
9851: 0 → 6

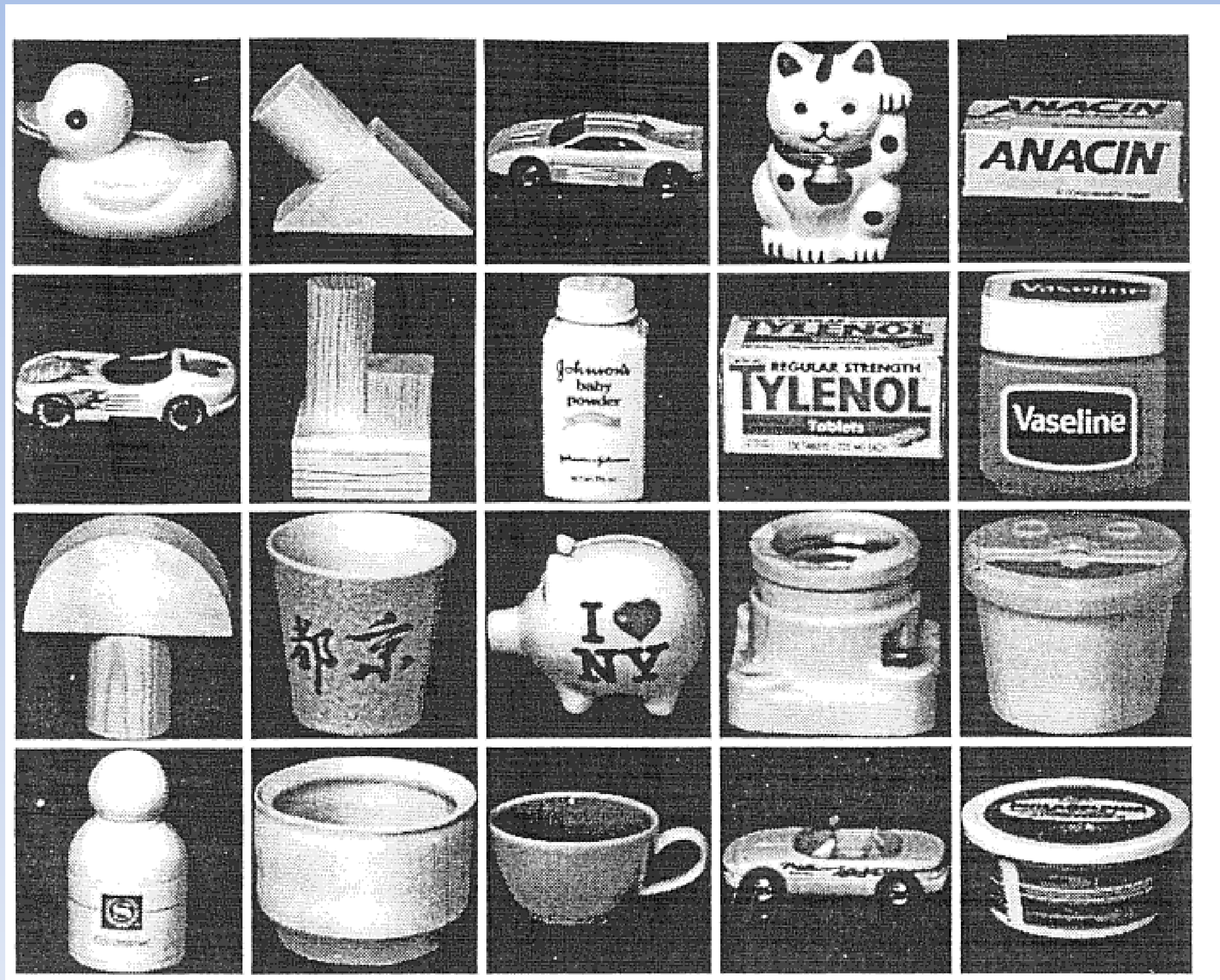
Results: Digit Recognition



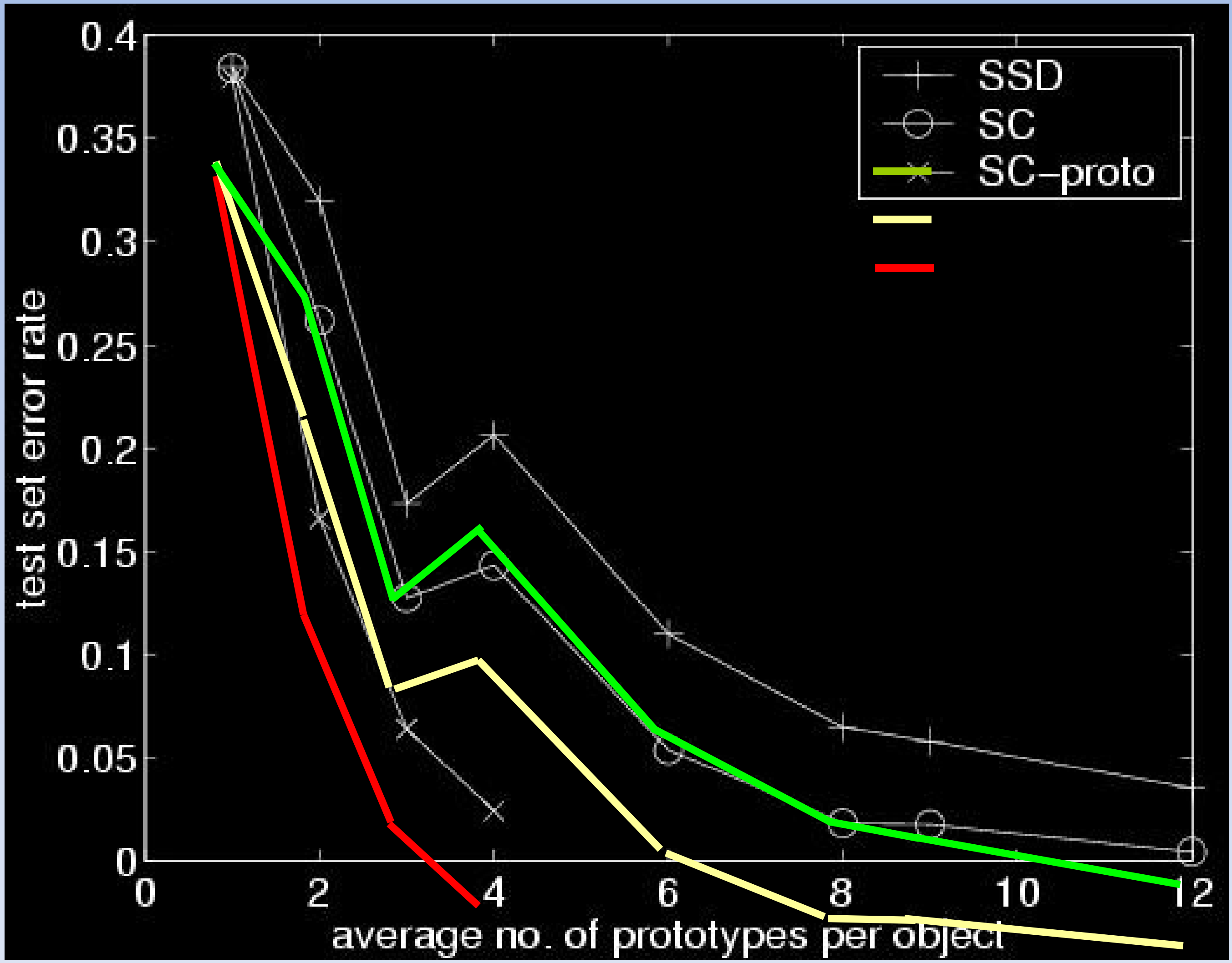
1-NN classifier using:

Shape context + 0.3 * bending + 1.6 * image appearance

COIL Object Database



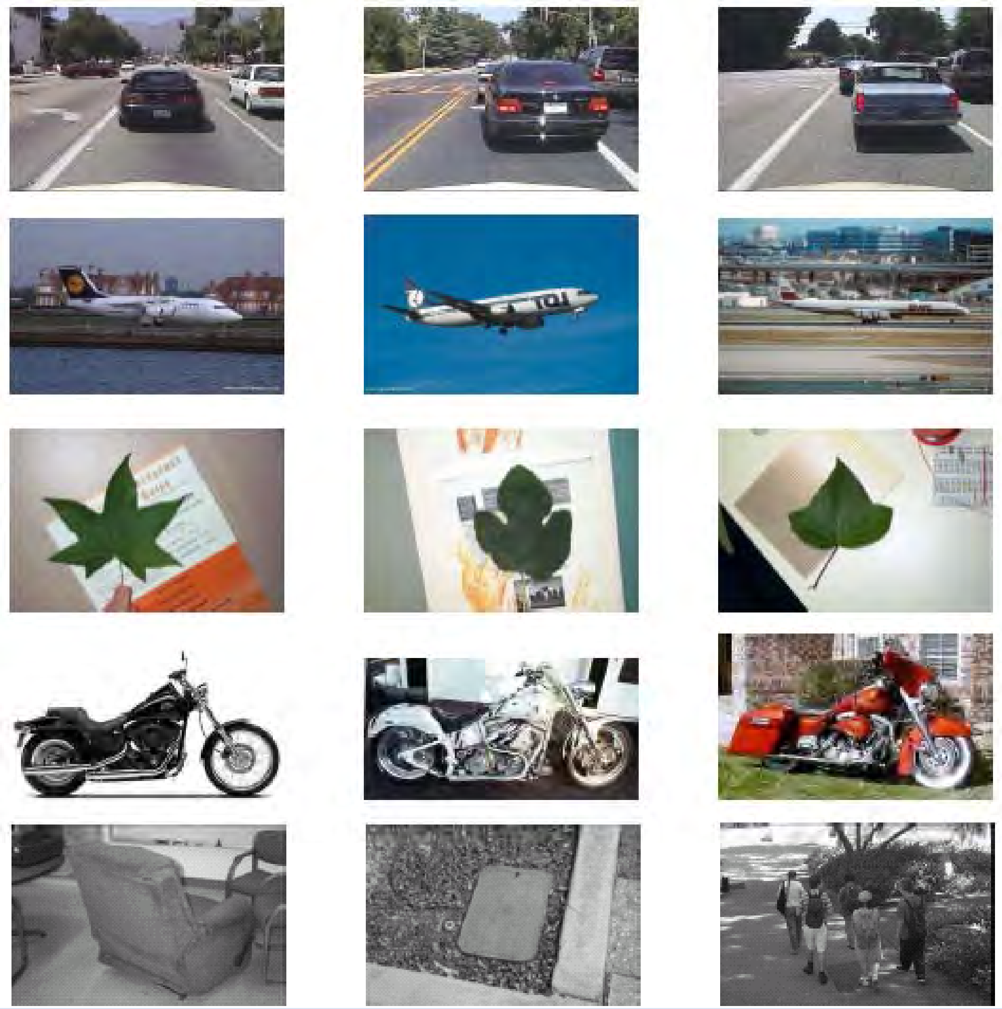
Error vs. Number of Views

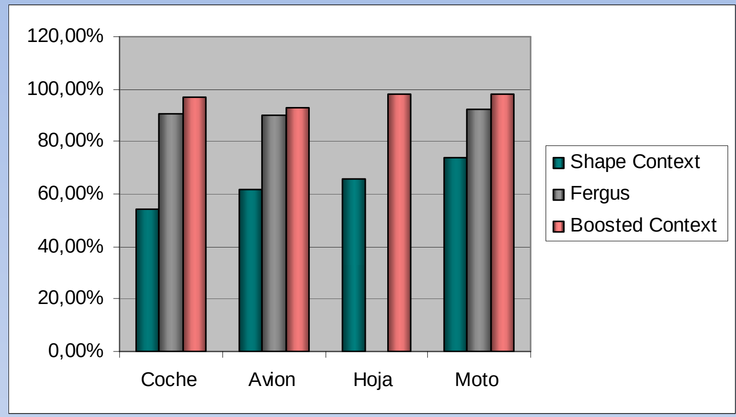


Prototypes Selected for 2 Categories

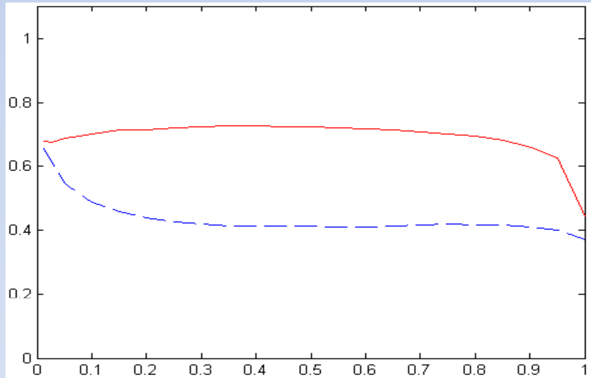
Details in Belongie, Malik & Puzicha (NIPS2000)

Retrieval validation





Error rate estimated by the ROC curve



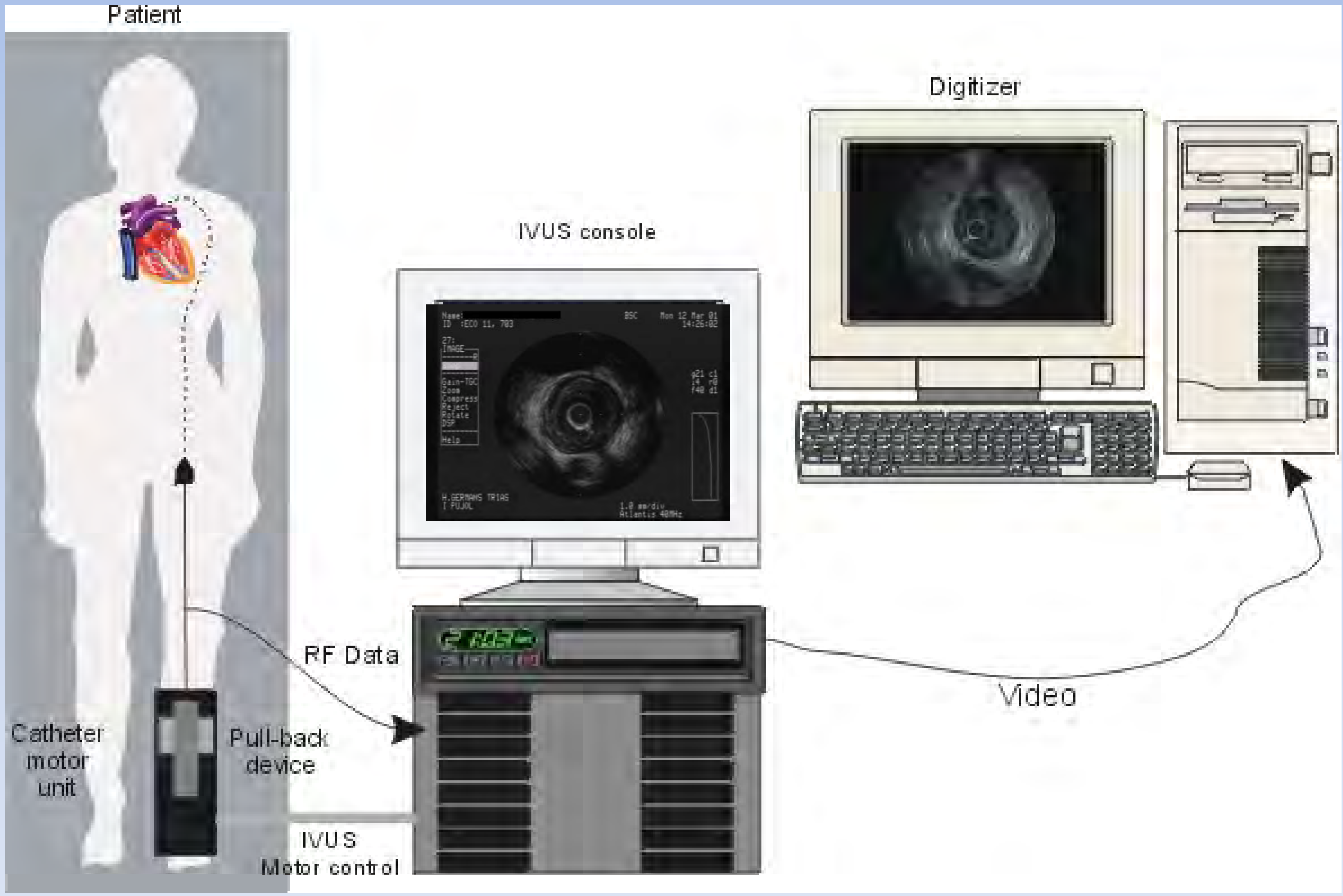
The “Precision-Recall” curve for car category

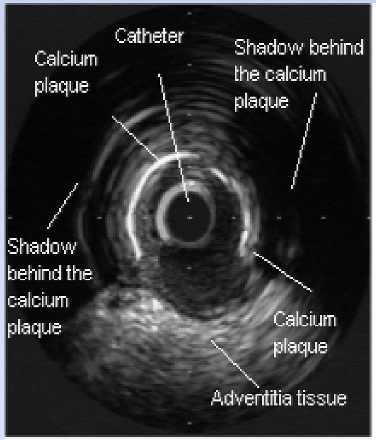
Precision=TP / (TP+FP)

Recall=TP / (TP+FN)

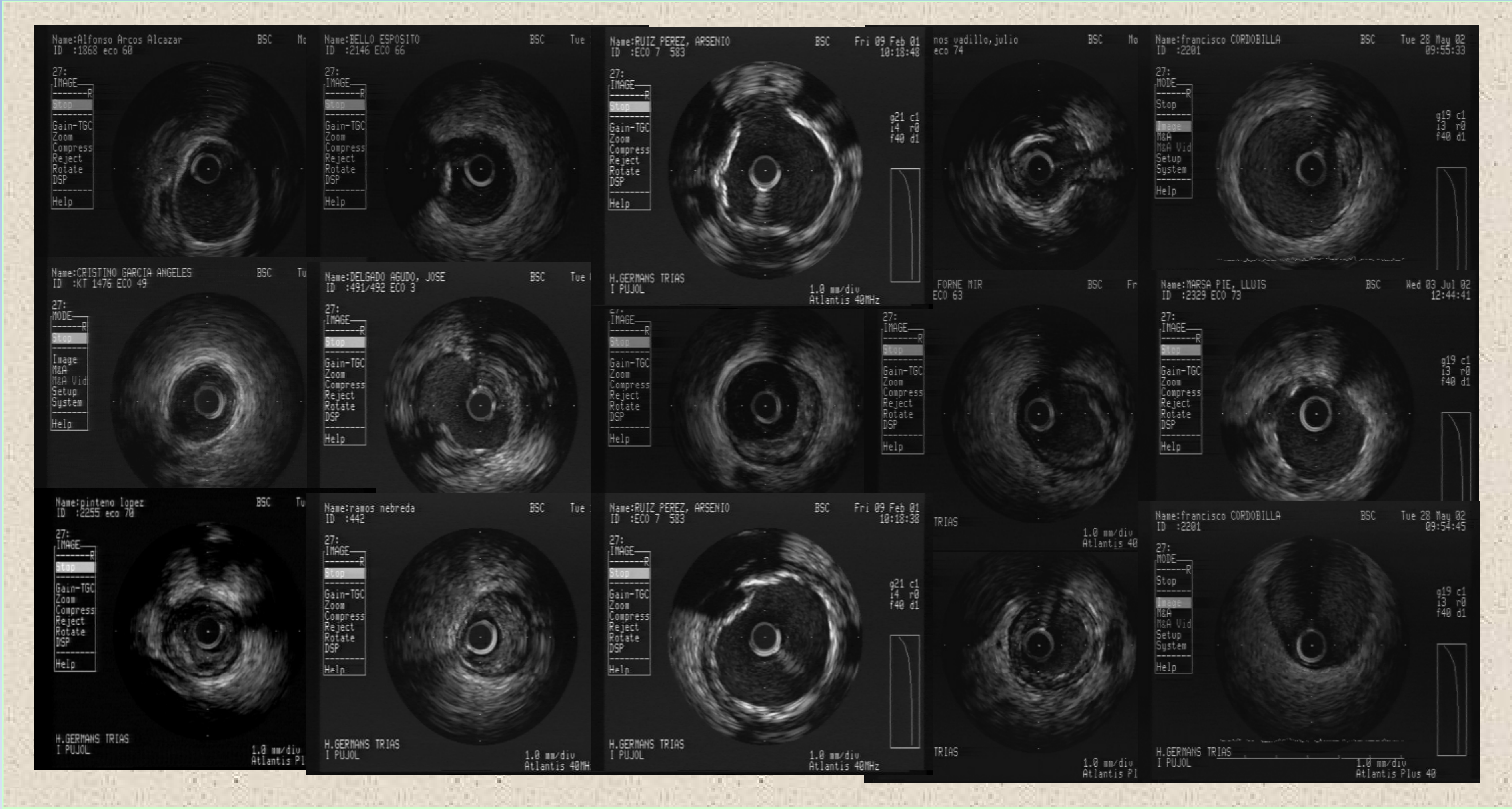
- 3512 images,
- 4 categories

IVUS retrieval by shape context





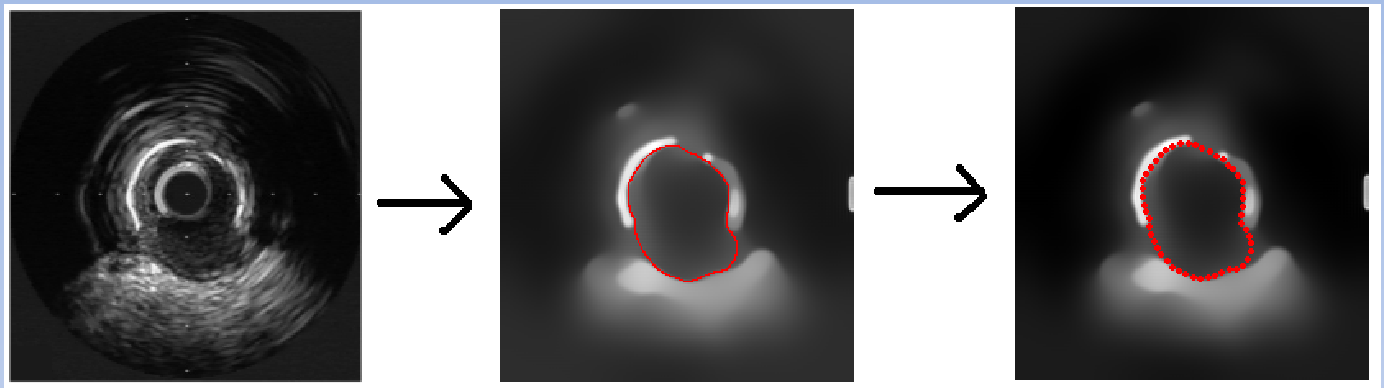
Tissue characterization and retrieval

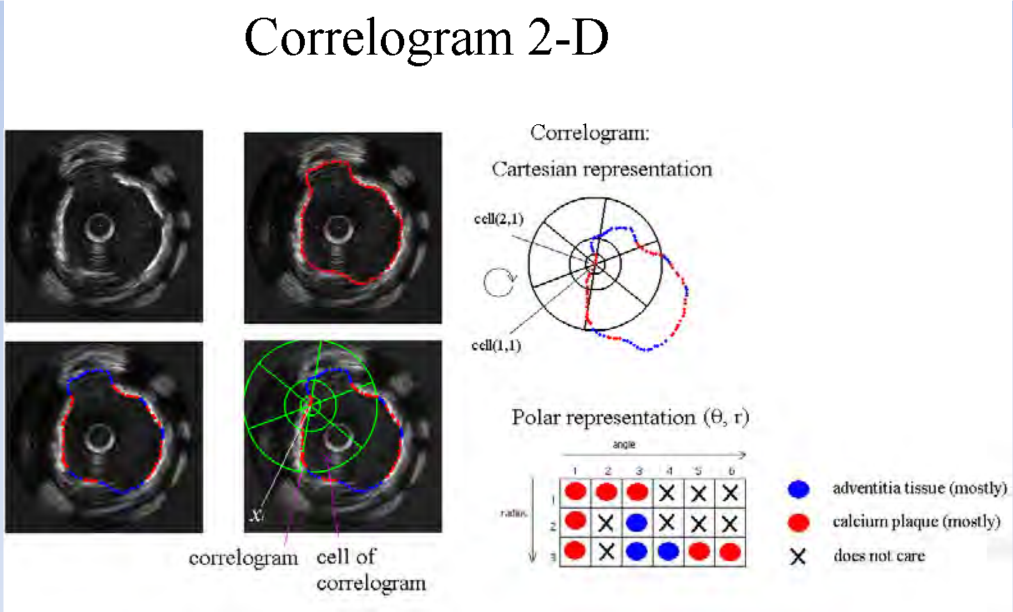


Context shape for IVUS retrieval

- What should be our features?
  - Generate context on different and customized feature set
- What should be our classifiers?
  - Generate (multiple) classifiers
  - Determine the decision rule

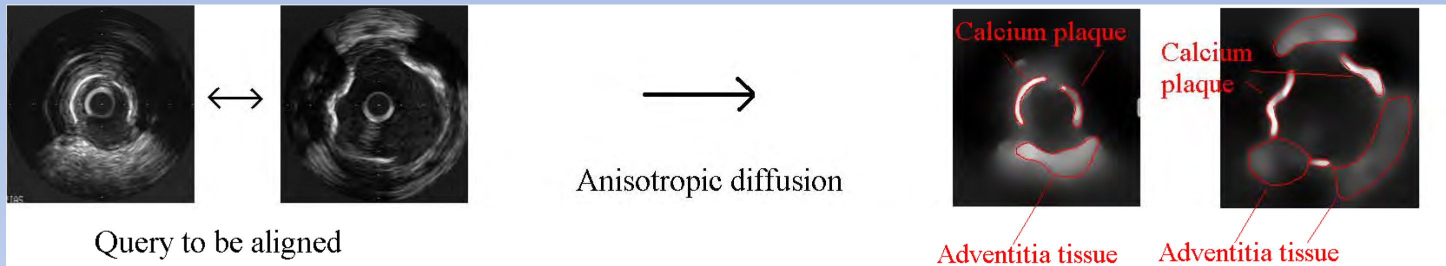
Boosted Context Retrieval of IVUS

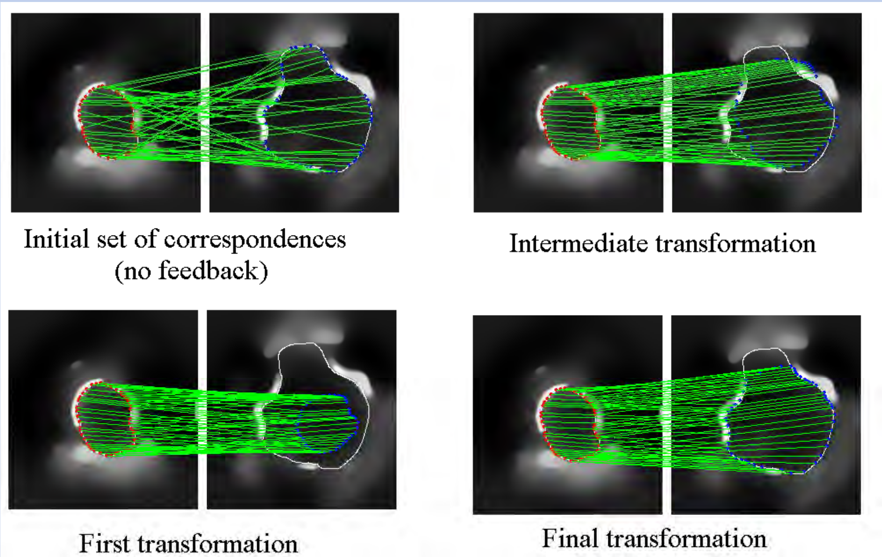


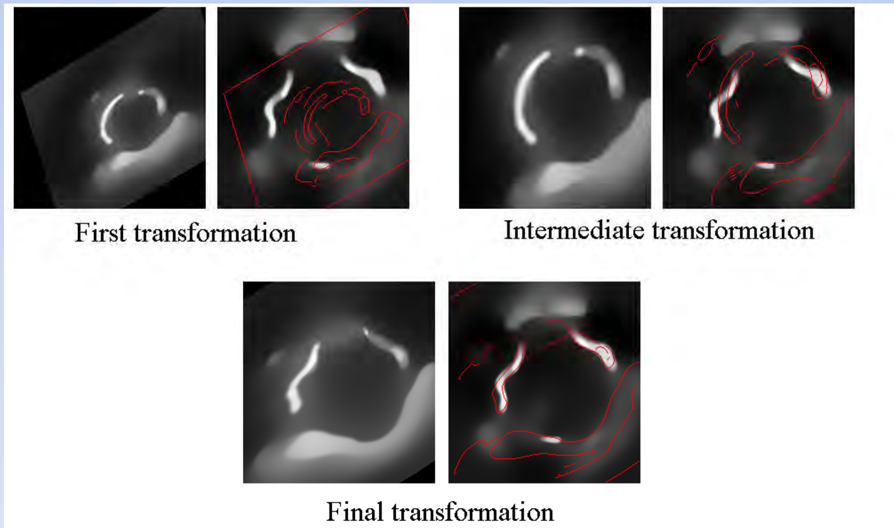


From: Jaume Amores, Petia Radeva: Retrieval of IVUS images using contextual information and elastic matching. Int. J. Intell. Syst. 20(5): 541-559 (2005)

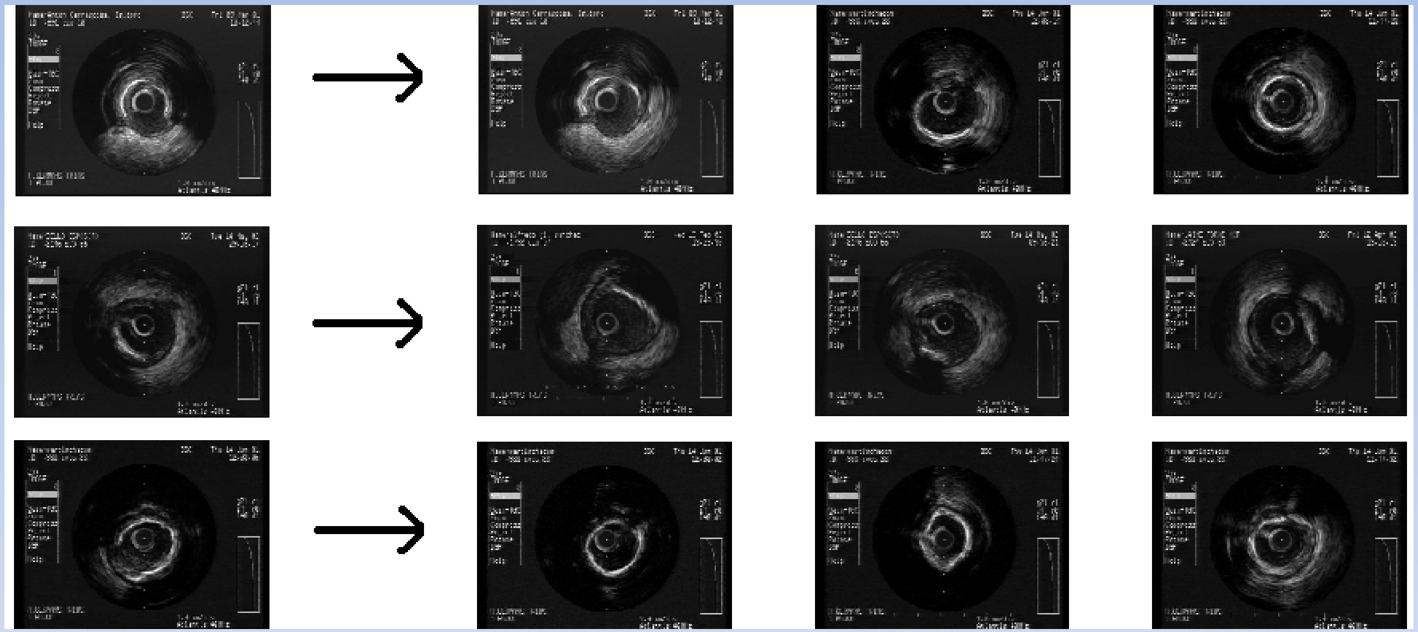
Context Matching of IVUS







Context Retrieval



Test: the probability to appear the most similar case in the first 3 automatically retrieved is 85%.

Conclusions

- Elegant way to integrate relation betweens parts of the object into the registration process

- Not necessarily connected object boundaries

- Allows to extend to 3D and 4D

- Allows to integrate different feature descriptors

- Validated in extensive public domain databases and medical problems