Grup de Processament de la Imatge
Universitat Politècnica de Catalunya (UPC)

Part-based Object Retrieval with Binary Partition Trees

Phd thesis dissertation by Xavier Giró-i-Nieto

Supervised by Professor Ferran Marqués Acosta (UPC)

and Professor Shih-Fu Chang (Columbia)

Acknowledgements

Grup de Processament de la Imatge
Universitat Politècnica de Catalunya (UPC)

Jury

GPI @ UPC

DVMM @ Columbia

Students @ UPC

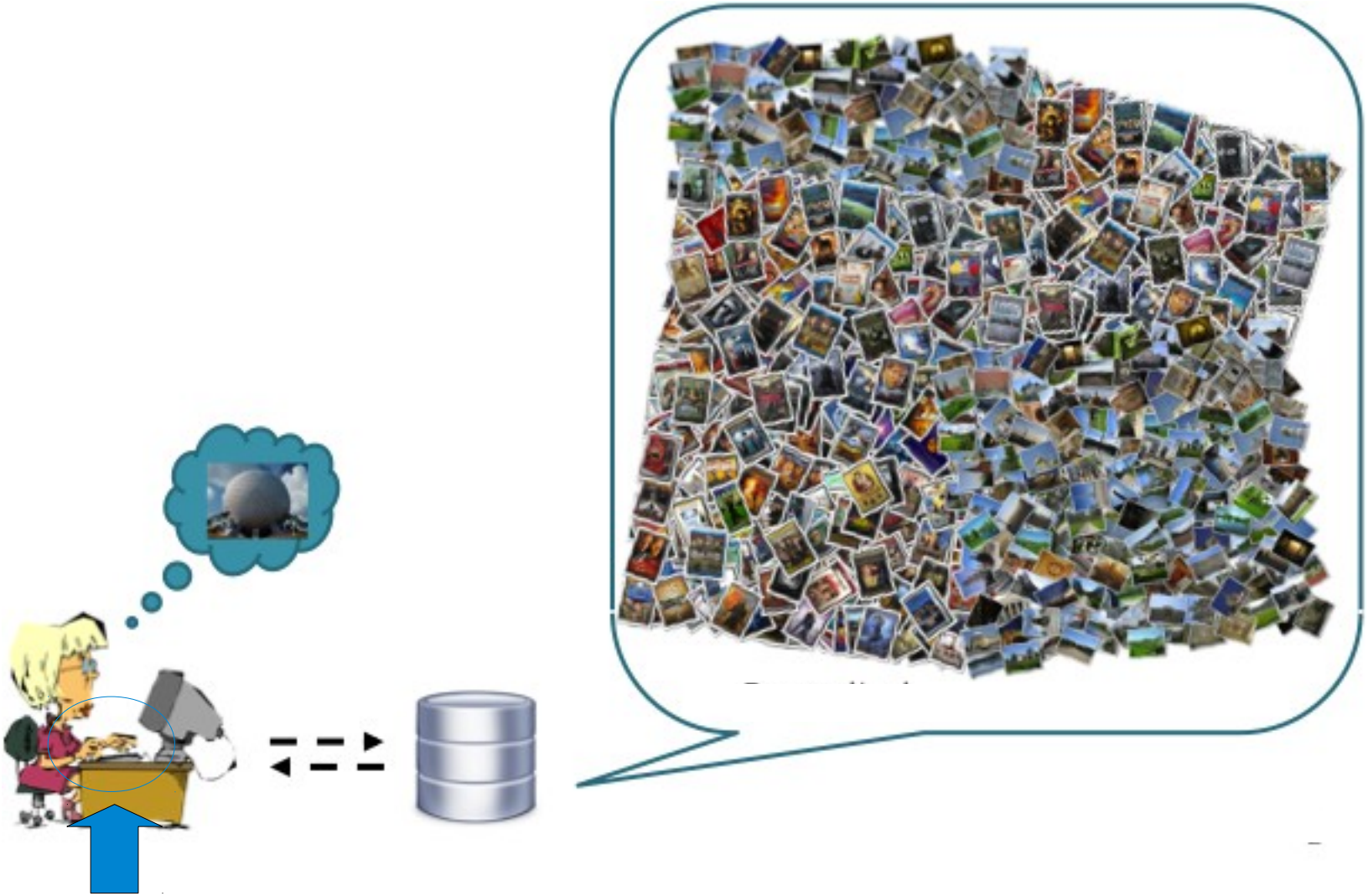
Outline

Problem statement

- Content-Based Image Retrieval

Problem statement

- Semantic gap between user query and machine data.



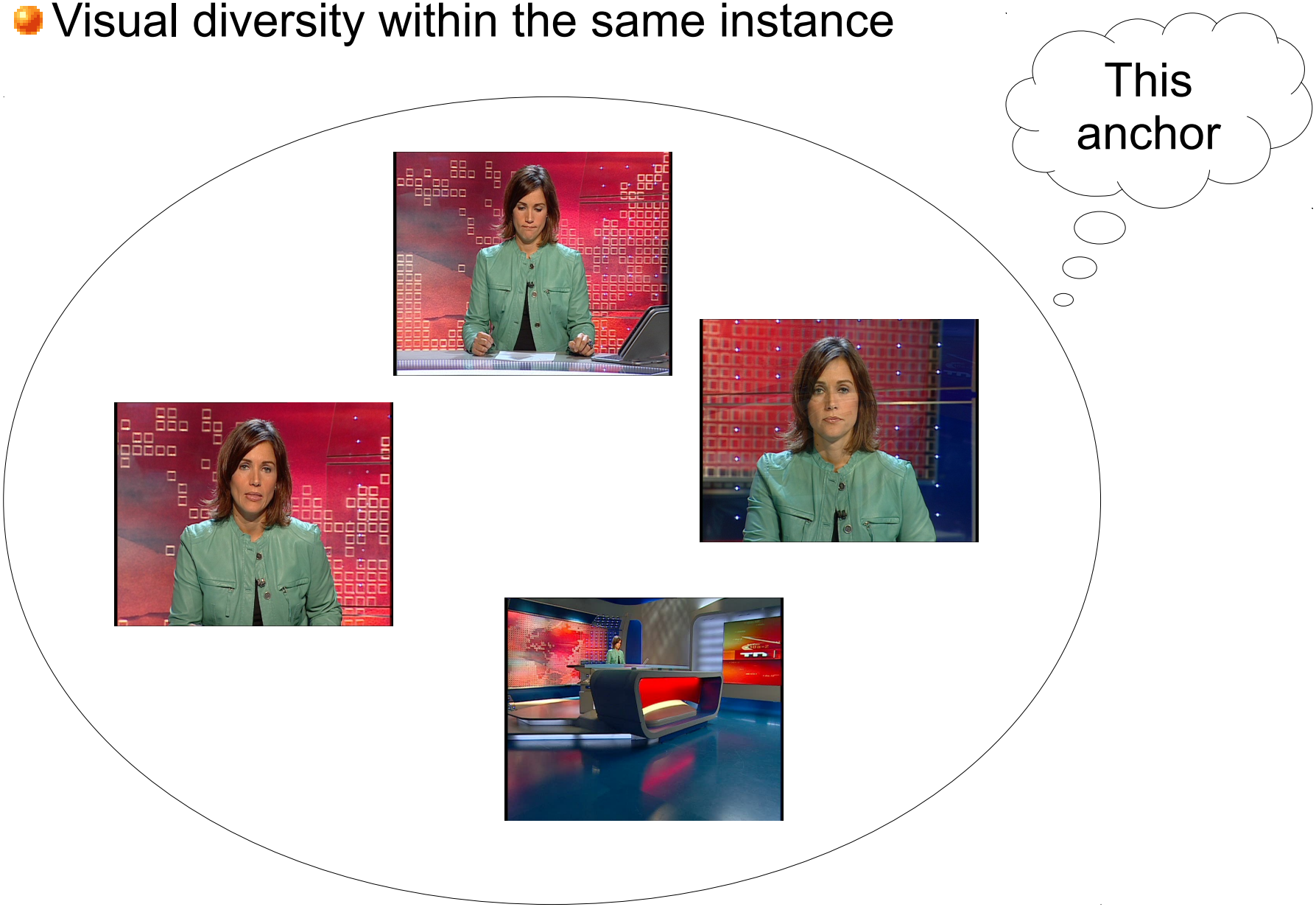
Challenges

- Visual **diversity** within the same semantic class

anchor

Challenges

- Visual diversity within the same instance



Grup de Processament de la Imatge
Universitat Politècnica de Catalunya (UPC)

Challenges

- Semantics located at multiple scales

News

Head

TV studio

Anchor

Button

Jacket

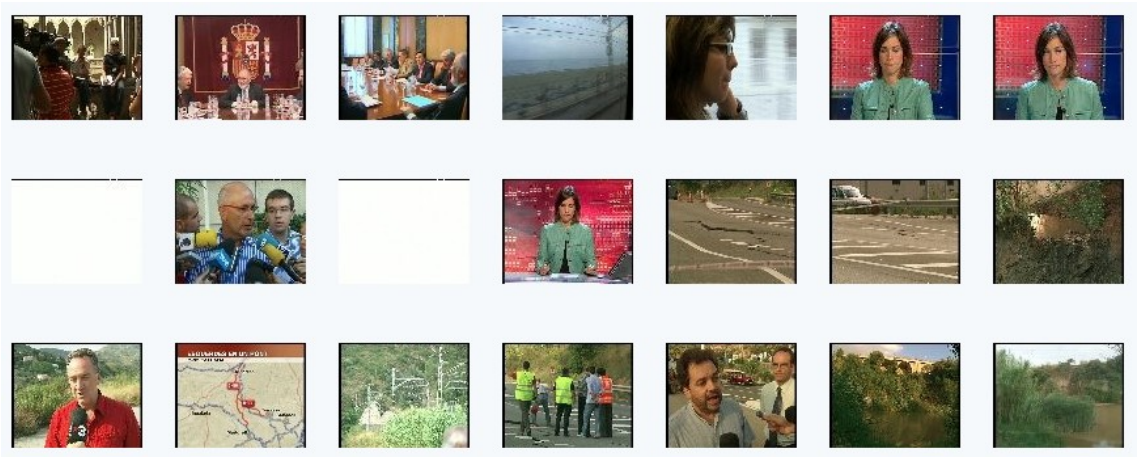
Grup de Processament de la Imatge
Universitat Politècnica de Catalunya (UPC)

Problem statement

- This thesis focuses in the **object retrieval** with one example...

Query object

Target dataset



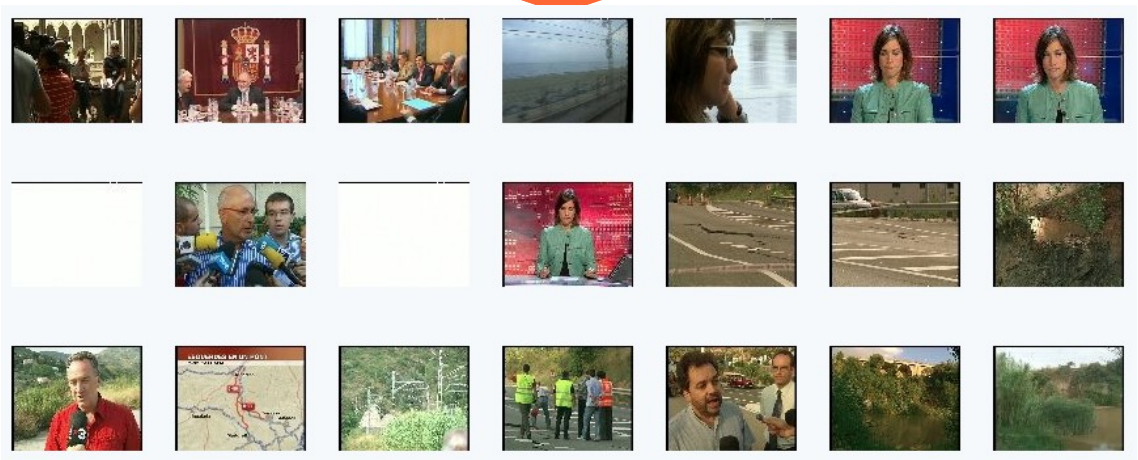
Ranked list

Problem statement

- ...and multiple examples.

Query objects

Target dataset



Ranked list

Outline

2. Hierchical Image Partition

Object analysis

Local scale

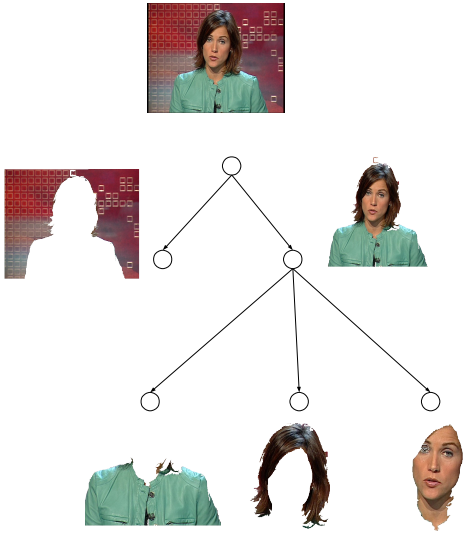
	Sliding windows	Interest points	Image partition
	[Viola-Jones'01]	SIFT [Lowe'04] SURF [Bay'06]	
Object Segmentation ?	✘	✘	✔

2. Hierchical Image Partition

Semantics at multiple scales

Multi-scale analysis

Spatial Pyramid of points	Multiple segmentations	Hierarchical partition ✔
[Lazebnik'06] [Bosch'07]	[Ravinovich'07] [Vieux'10]	[Adamek'06] [Gu'09]

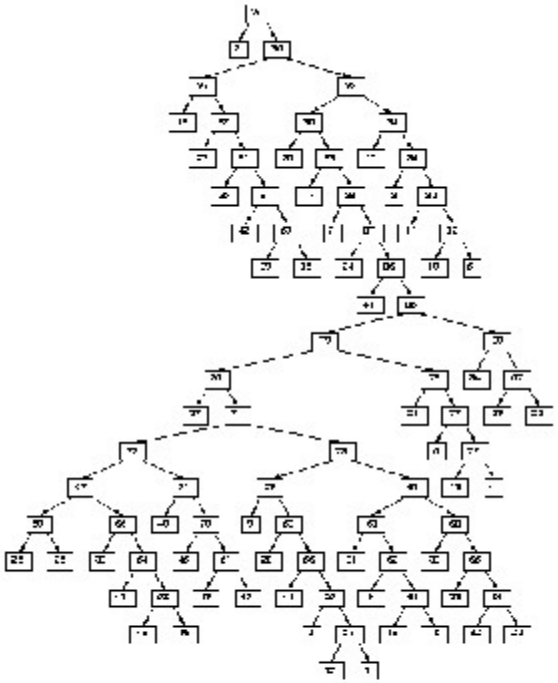


Grup de Processament de la Imatge
Universitat Politècnica de Catalunya (UPC)

2. Hierchical Image Partition

- Adopted Solution: Binary Partition Trees (BPTs) [Garrido, Salembier 2000].

(a)

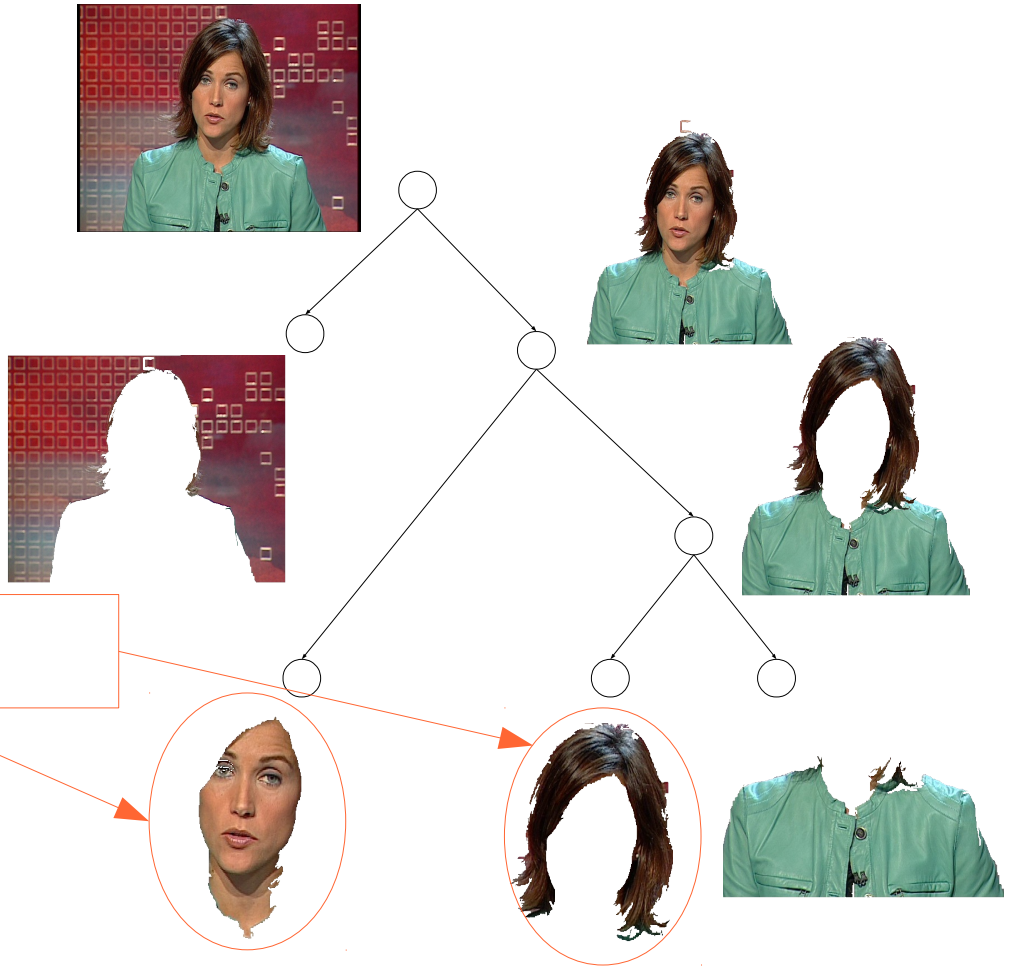


(b)

2. Hierchical Image Partition

Semantic
Object "head"

Perceptual
BPT



Thesis
topic

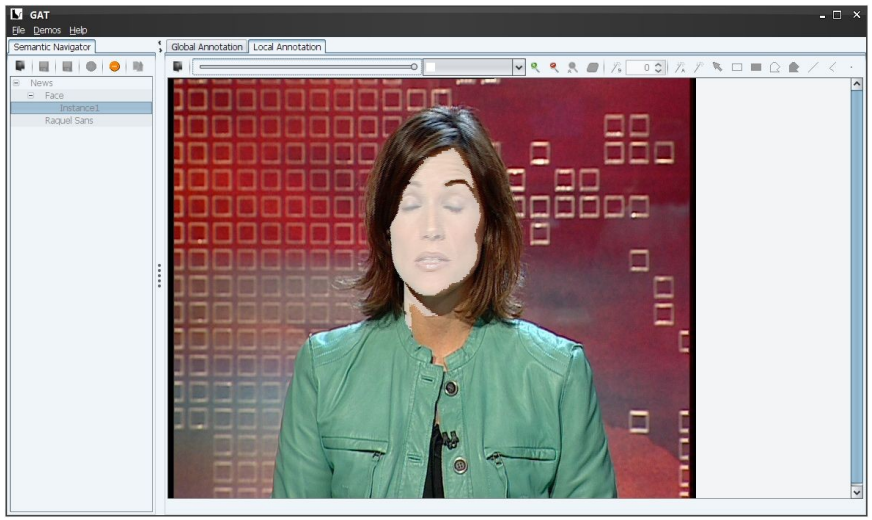
Grup de Processament de la Imatge
Universitat Politècnica de Catalunya (UPC)

Outline

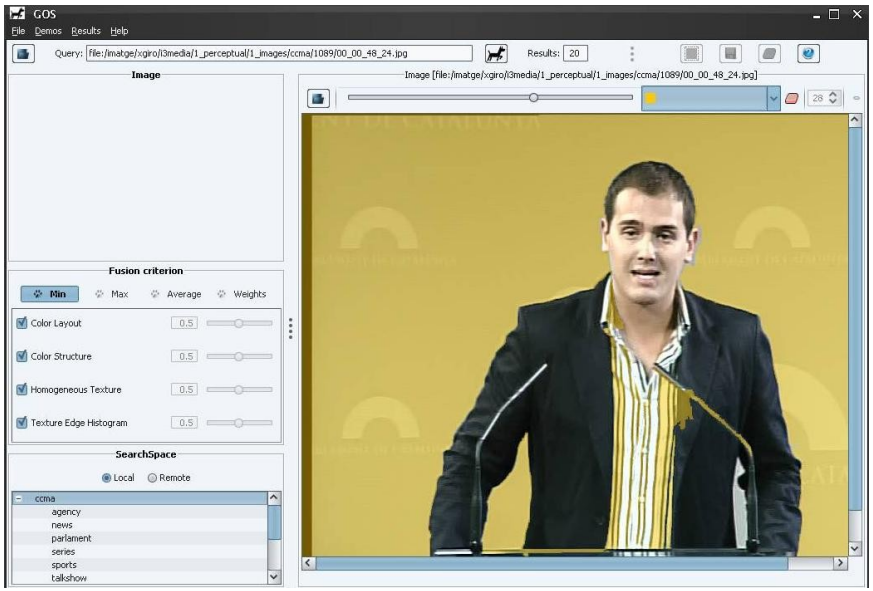
3. Interactive segmentation

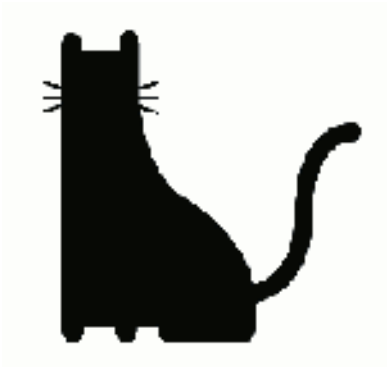
- Motivations: Ground truth generation and retrieval interface.

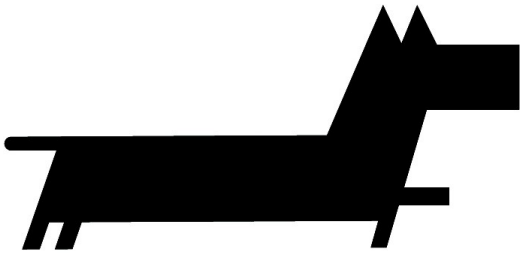
Graphical Annotation Tool (GAT)



Graphical Object Searcher (GOS)





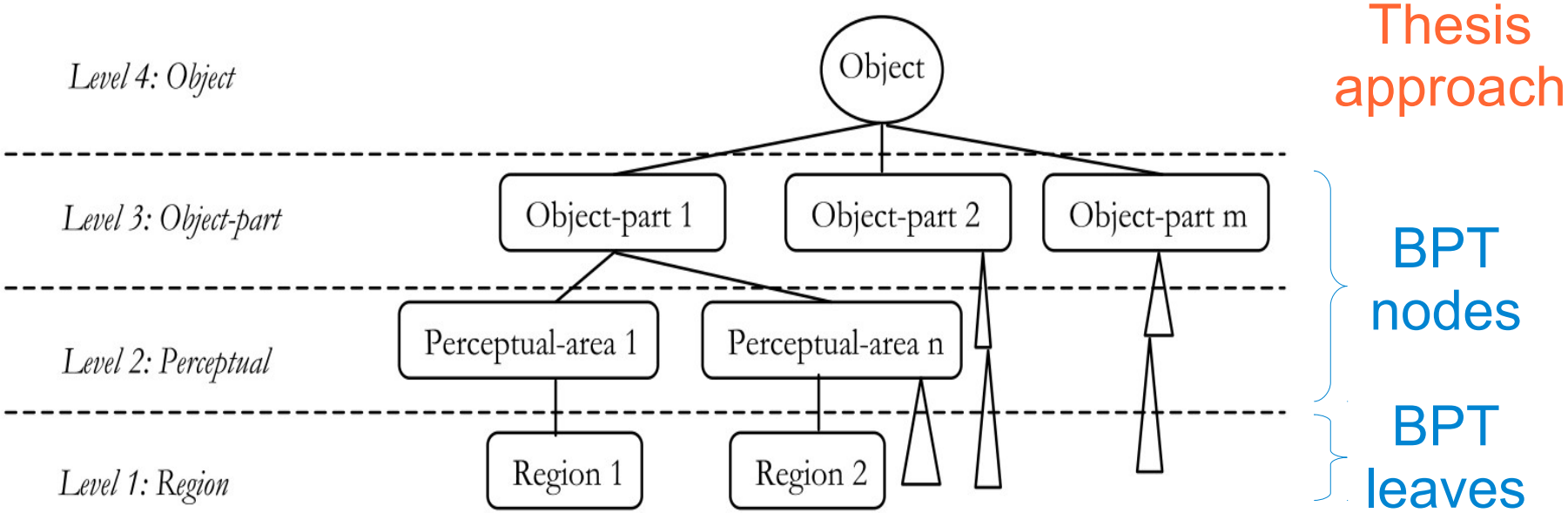


- Demo @ http://upseek.upc.edu/gat & http://upseek.upc.edu/gos

Outline

4. Object Tree

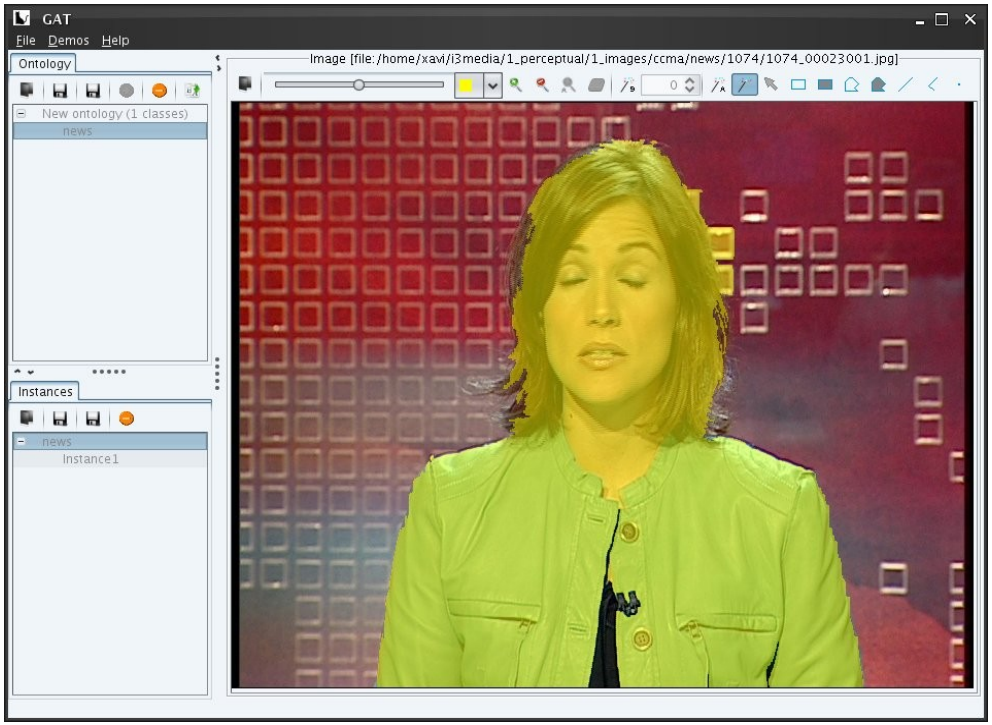
- Related work: “Definition Hierarchy”, Jaimes and Chang (2001)



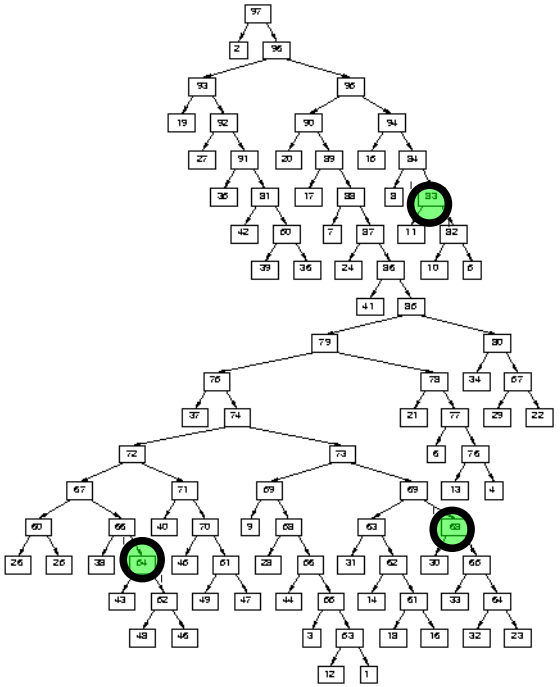
4. Object Tree

- Assumption: All objects are connected

Interactive segmentation



BPT sub-roots selection

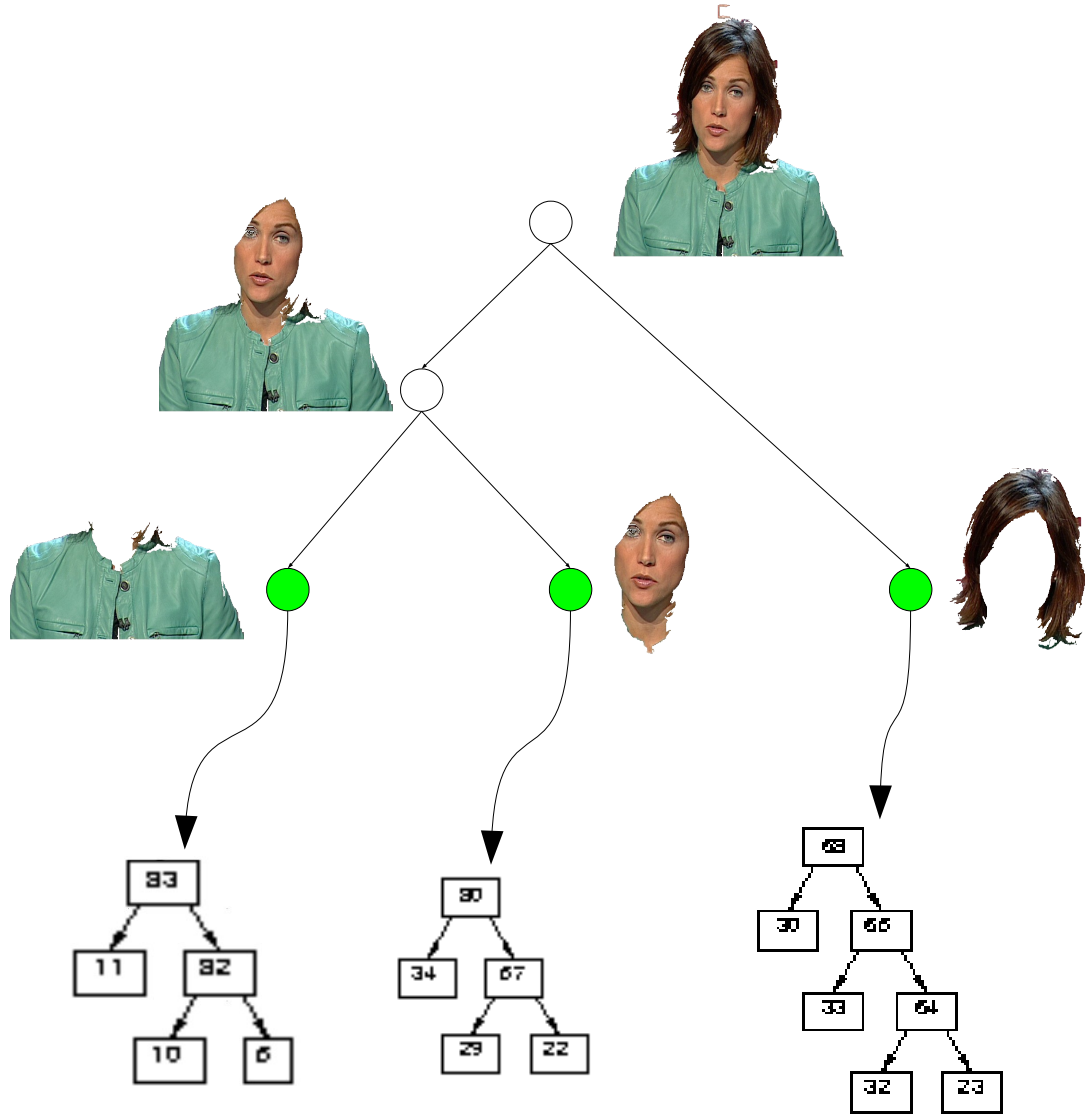


4. Object Tree

- The Object Tree (OT) represents an instance of the object as a hierarchy of regions.

User selection of sub-roots on the BPT.

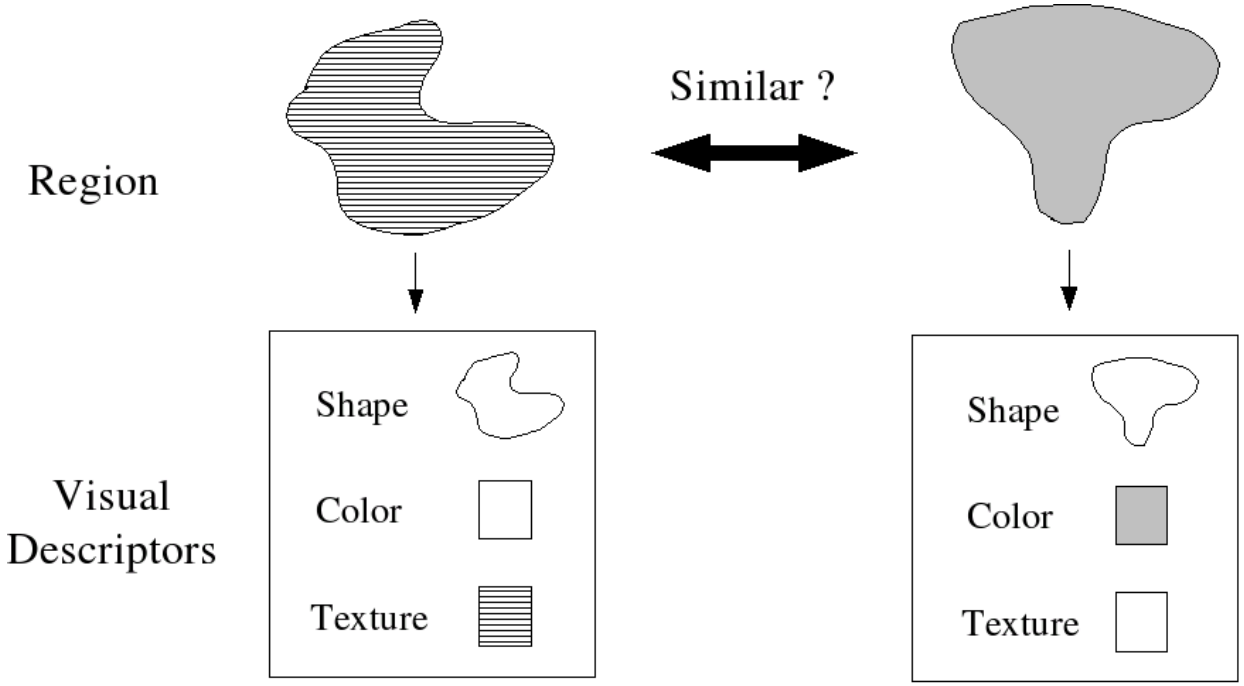
Automatic top-down expansion through the BPT



Outline

5. Region Features

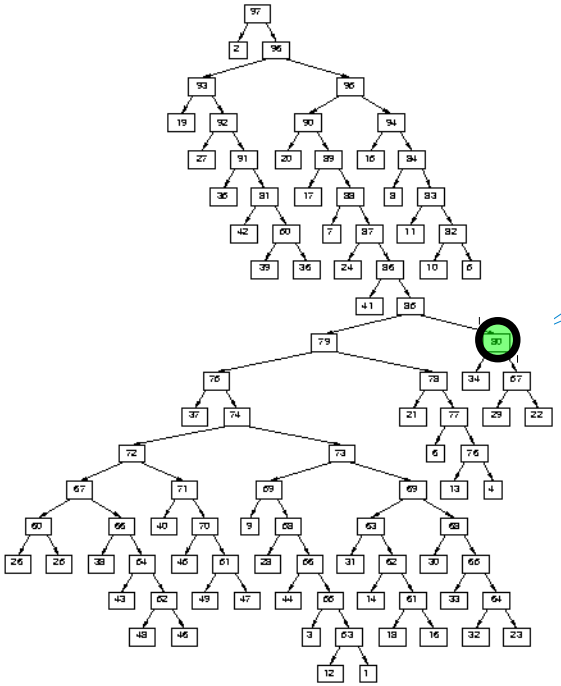
- MPEG-7 and many more, “Visual Descriptors (VD)” (2001)



5. Region Features

→ Region features in BPT

- Dominant Colors
- Contour Shape
- Edge Histogram



NOT Thesis topic

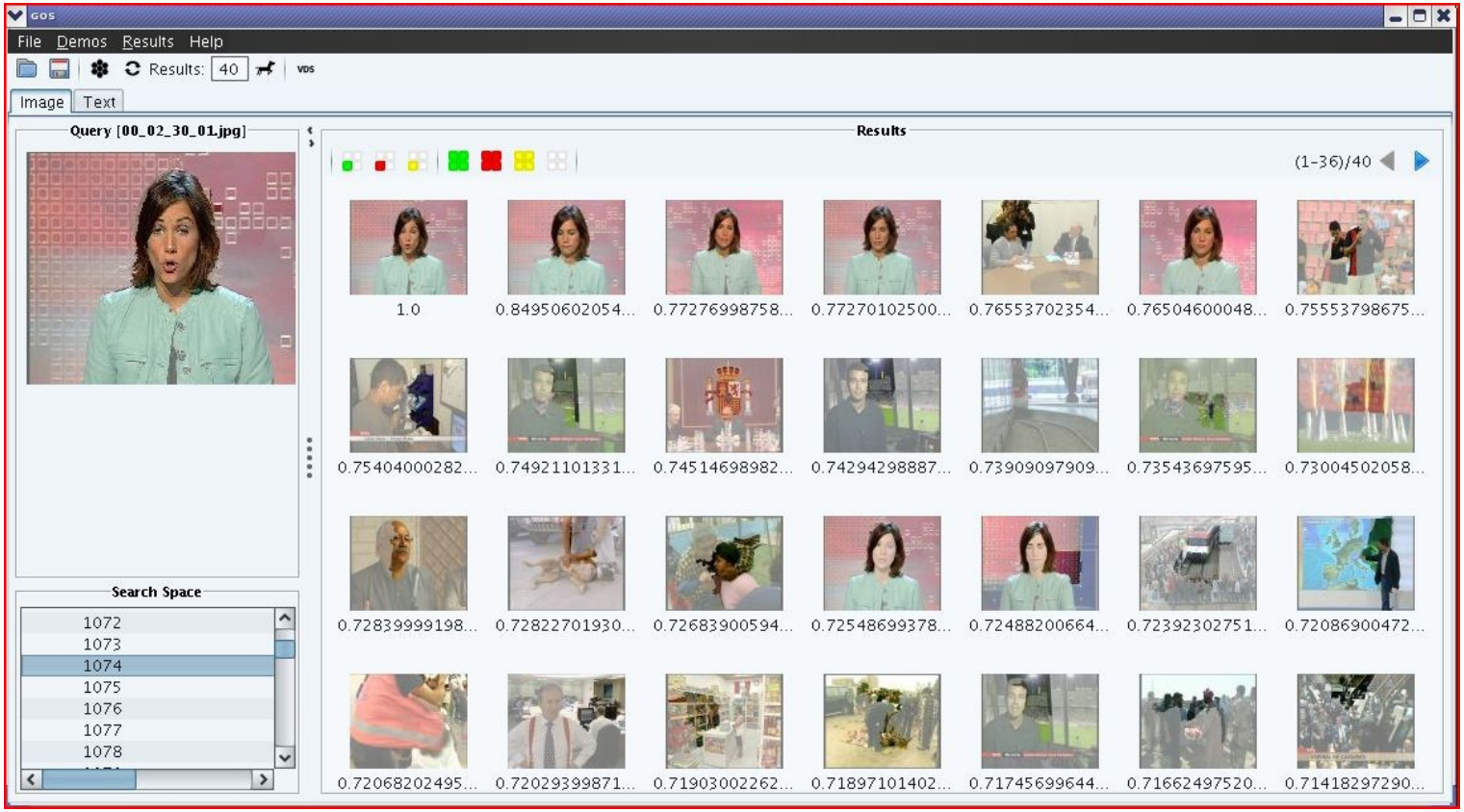
Outline

1. Introduction
2. Hierarchical Image Partitions
3. Interactive segmentation
4. Object Tree
5. **Region Features ←**
6. Codebooks
7. Fusion of Visual Modalities
8. Partition Tree Matching
9. Part-based Object Model
10. Conclusions

Part I: 2–4
Part II: 5–7
Part III: 8–9
Evaluation: 5–9

Evaluation

- How to evaluate an object retrieval system ?



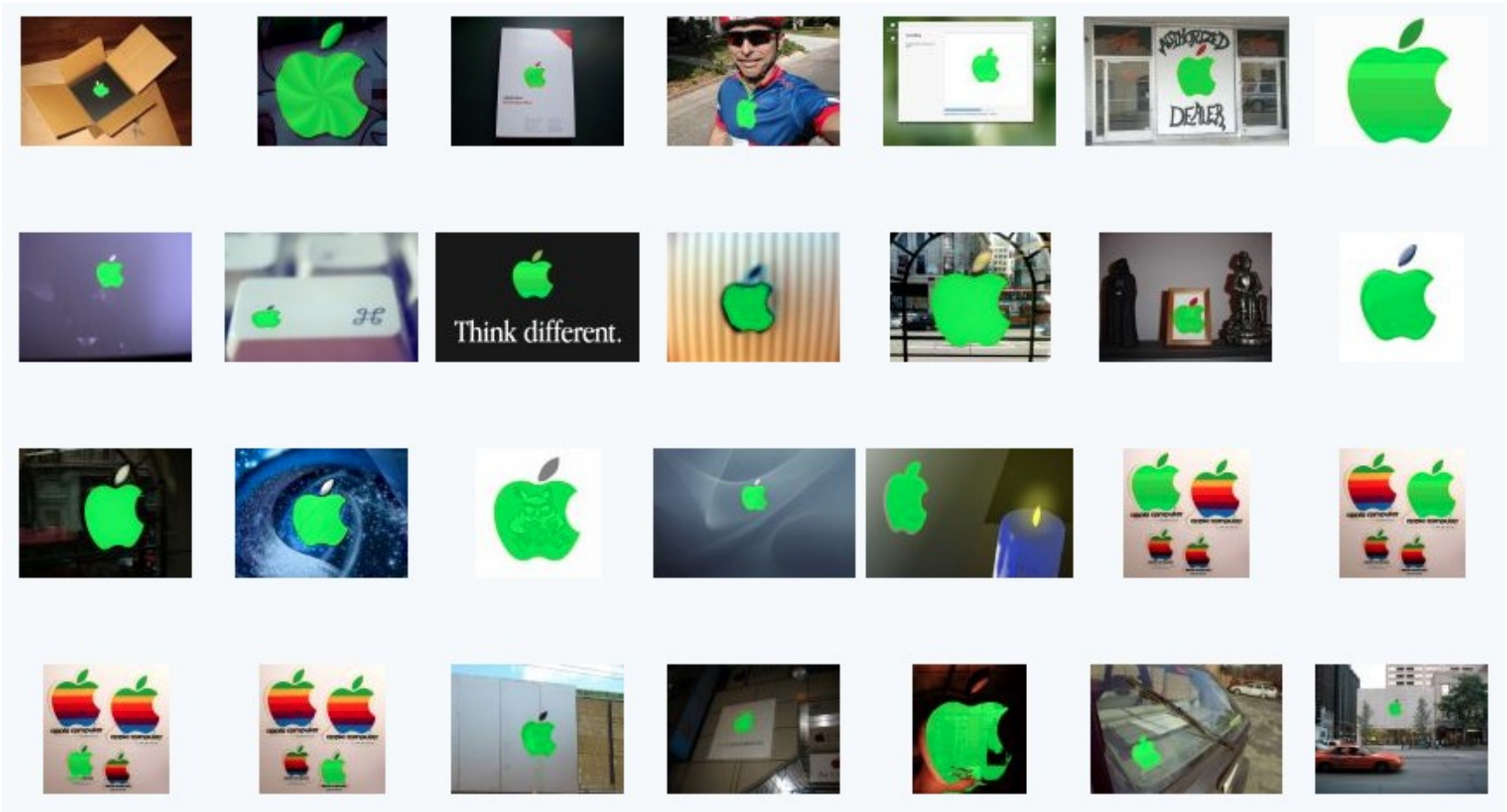
Annotated dataset + Metrics

1. Image Retrieval
2. Object Detection
3. Object Segmentation
4. Search time

Grup de Processament de la Imatge
Universitat Politècnica de Catalunya (UPC)

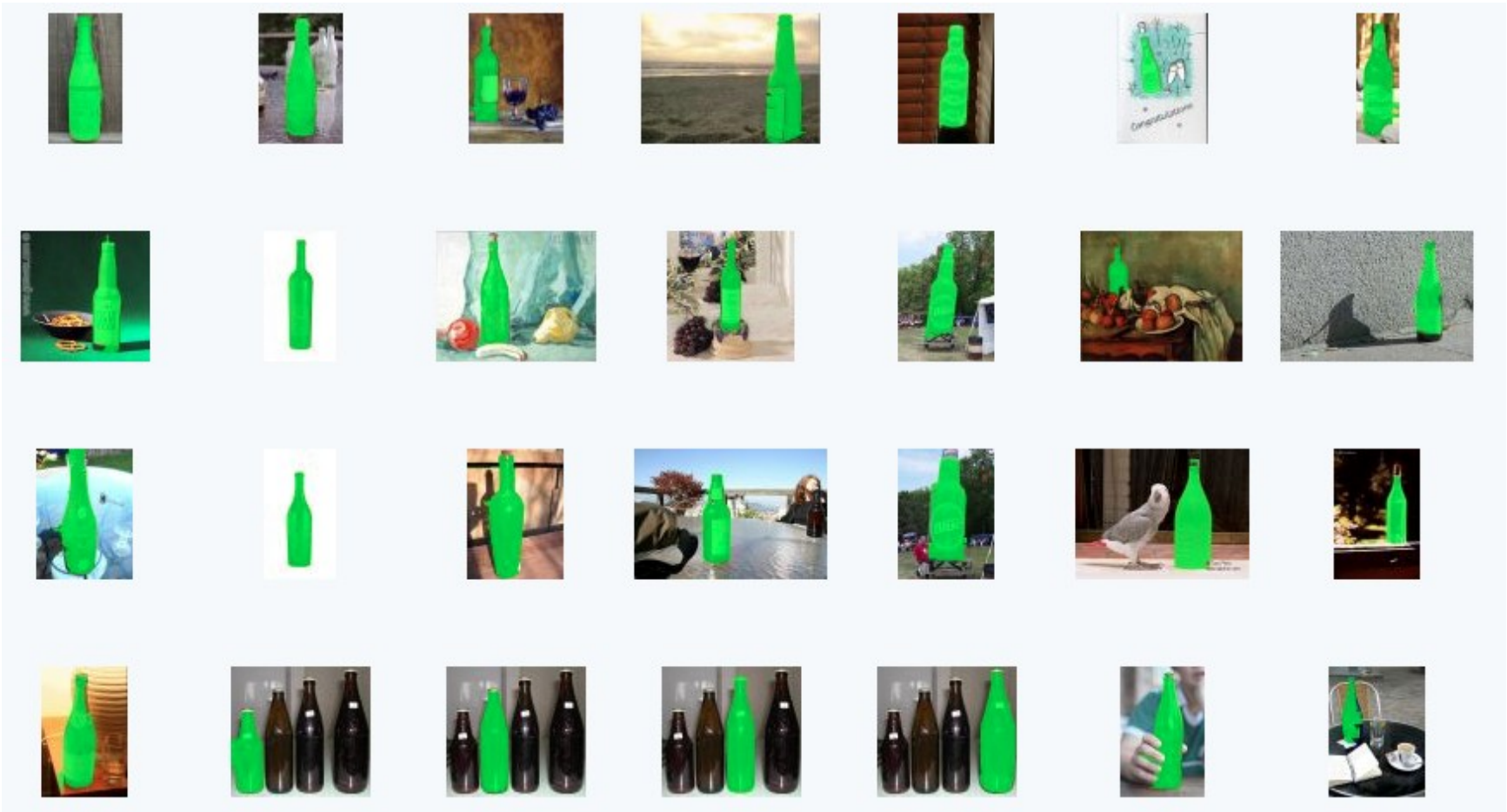
Evaluation (in Part II)

Annotated Dataset ETHZ: Category #1 “Apple logos”



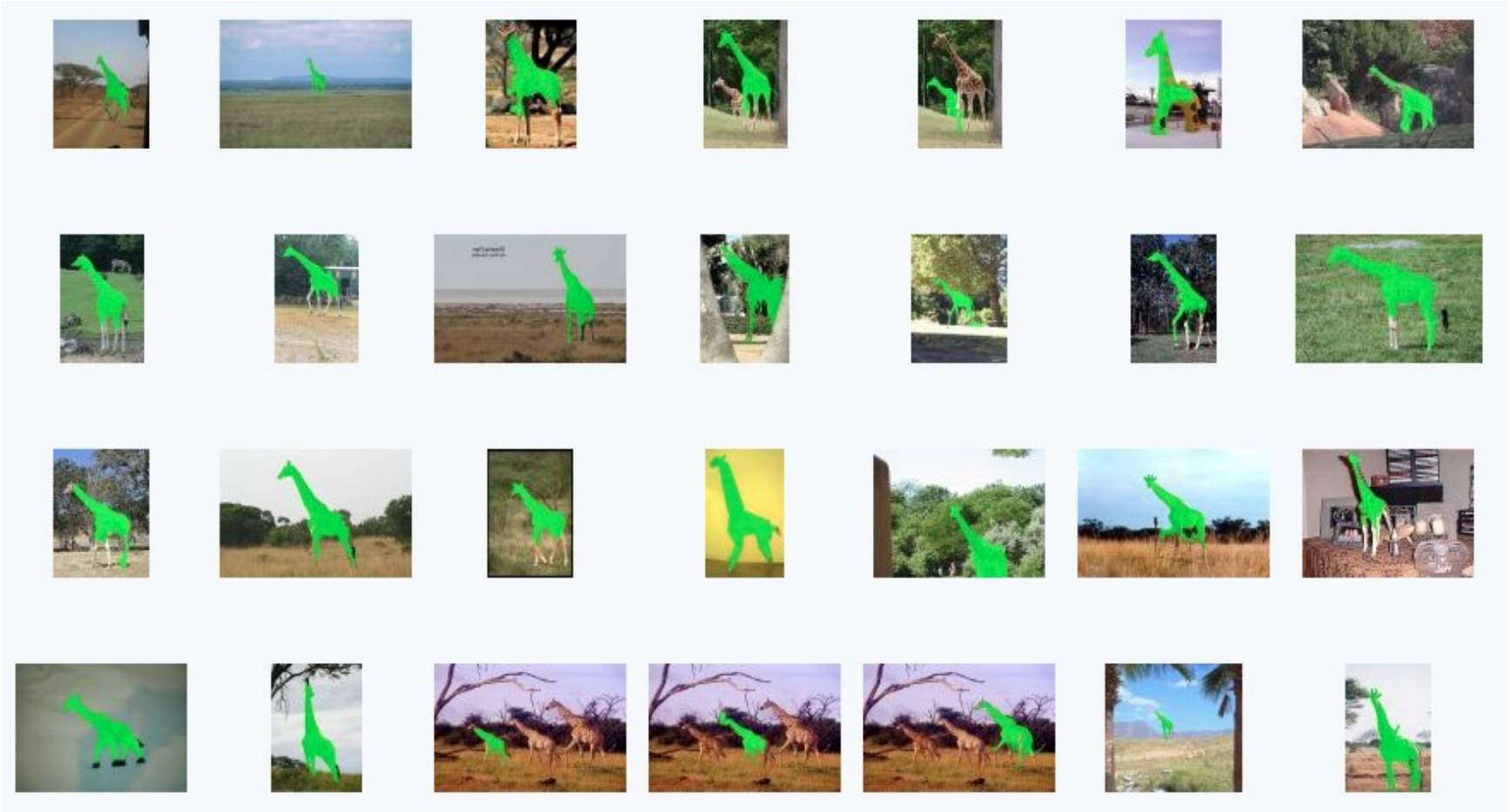
Evaluation (in Part II)

Annotated Dataset ETHZ: Category #2 “Bottles”



Evaluation (in Part II)

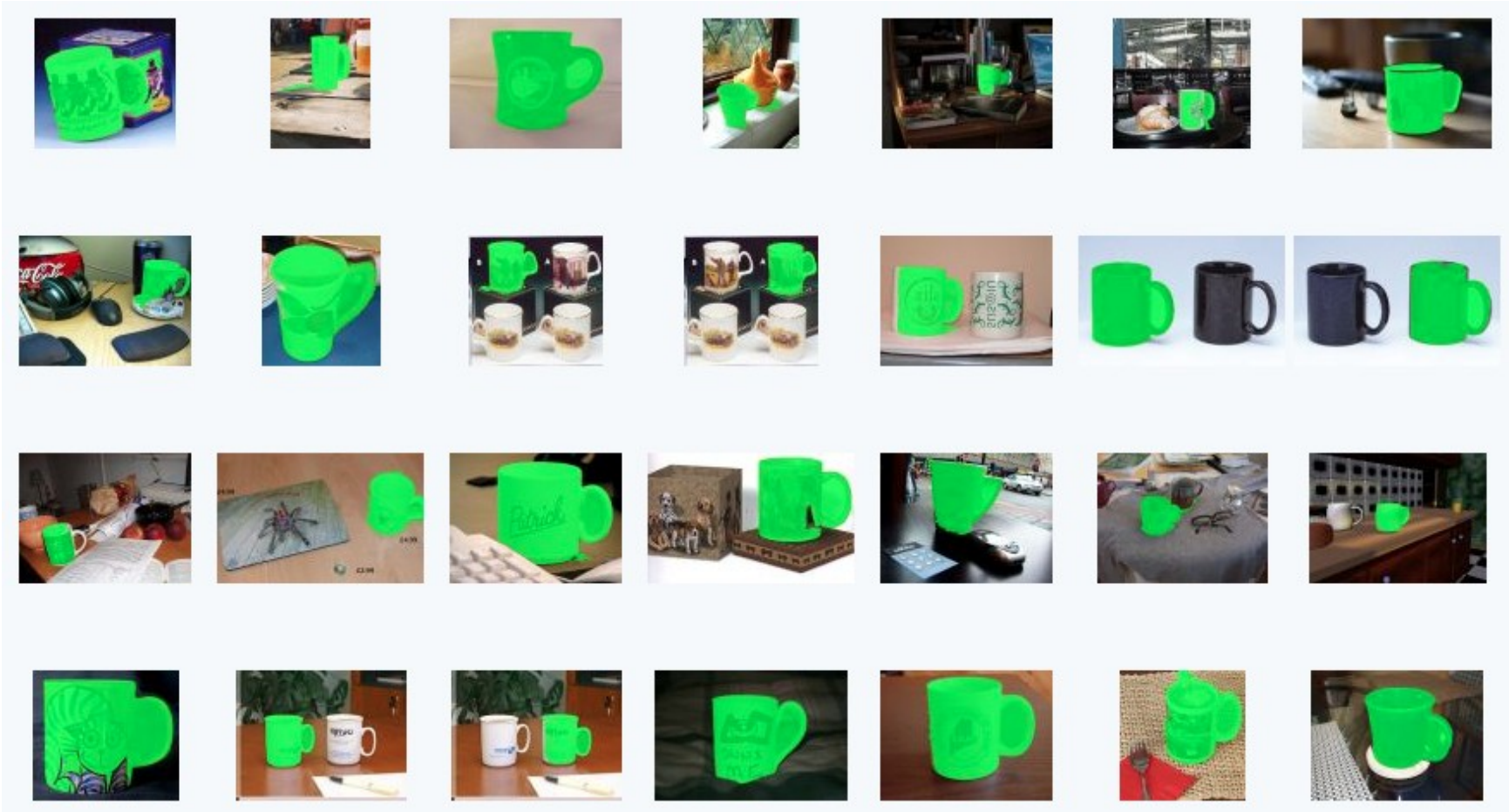
Annotated Dataset ETHZ: Category #3 “Giraffes”



Grup de Processament de la Imatge
Universitat Politècnica de Catalunya (UPC)

Evaluation (in Part II)

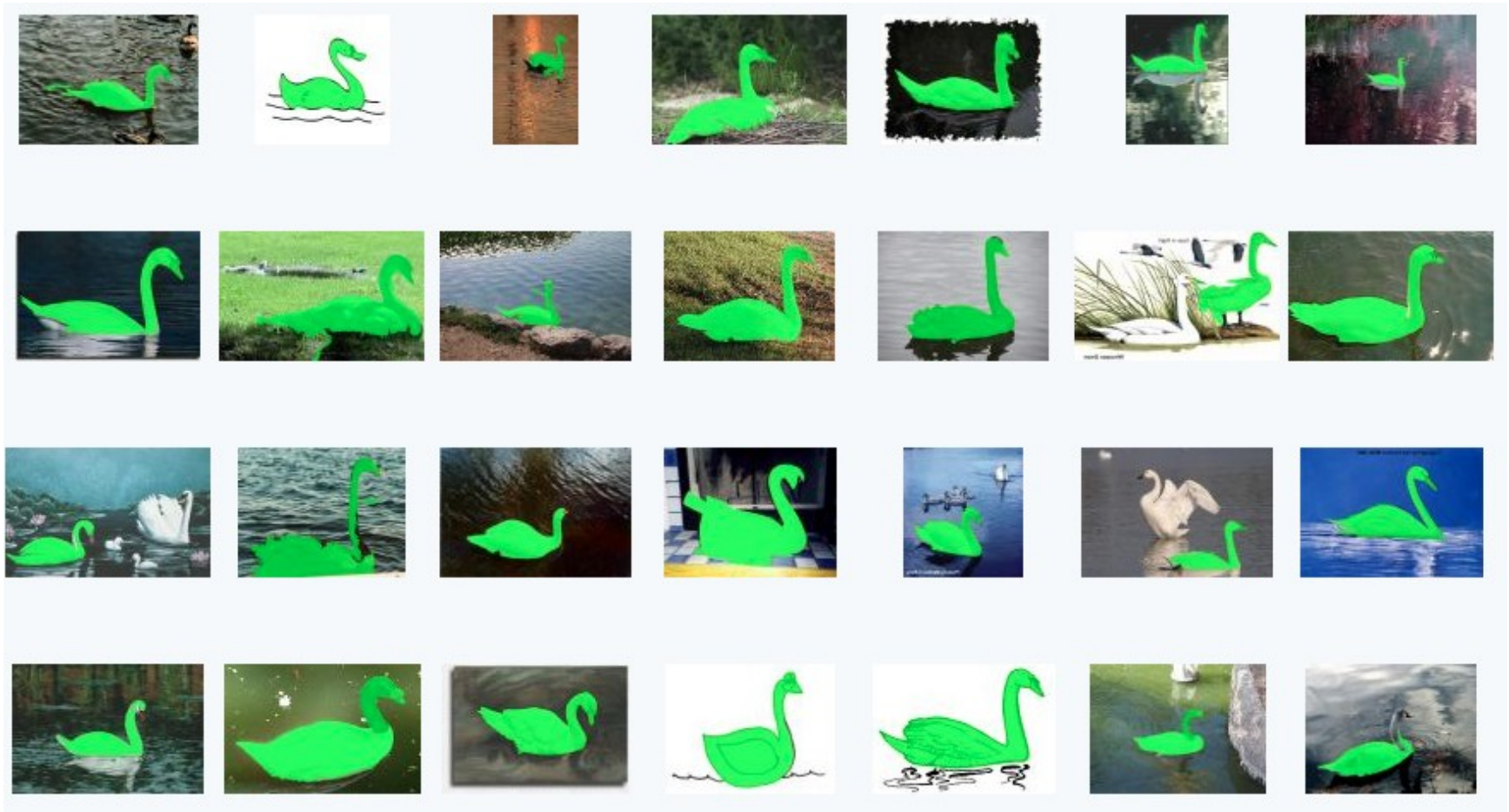
Annotated Dataset ETHZ: Category #4 “Mugs”



Grup de Processament de la Imatge
Universitat Politècnica de Catalunya (UPC)

Evaluation (in Part II)

Annotated Dataset ETHZ: Category #5 “Swans”



Evaluation

1. Image Retrieval (global scale)

- Only global scale is considered.

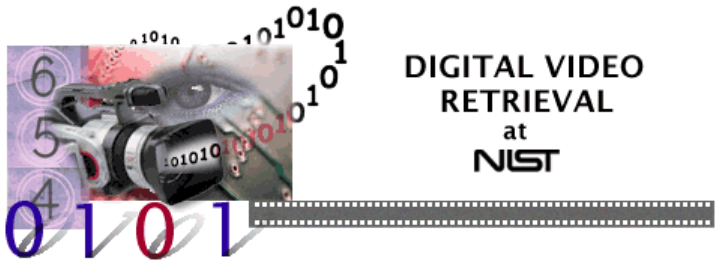
User Query (local)

Ranked list of images (global)

Evaluation

1. Image Retrieval (global scale)

- Mean Average Precision (MAP) combines Precision and Recall for a set of queries.



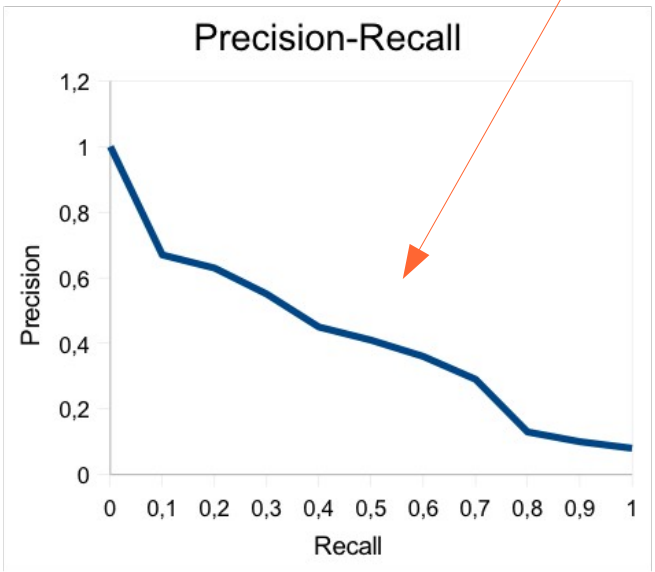
Precision

$$P(k)=\frac{|\{relevant\}\cap\{retrieved(k)\}|}{k}$$

Recall

$$R(k)=\frac{|\{relevant\}\cap\{retrieved(k)\}|}{|relevant|}$$

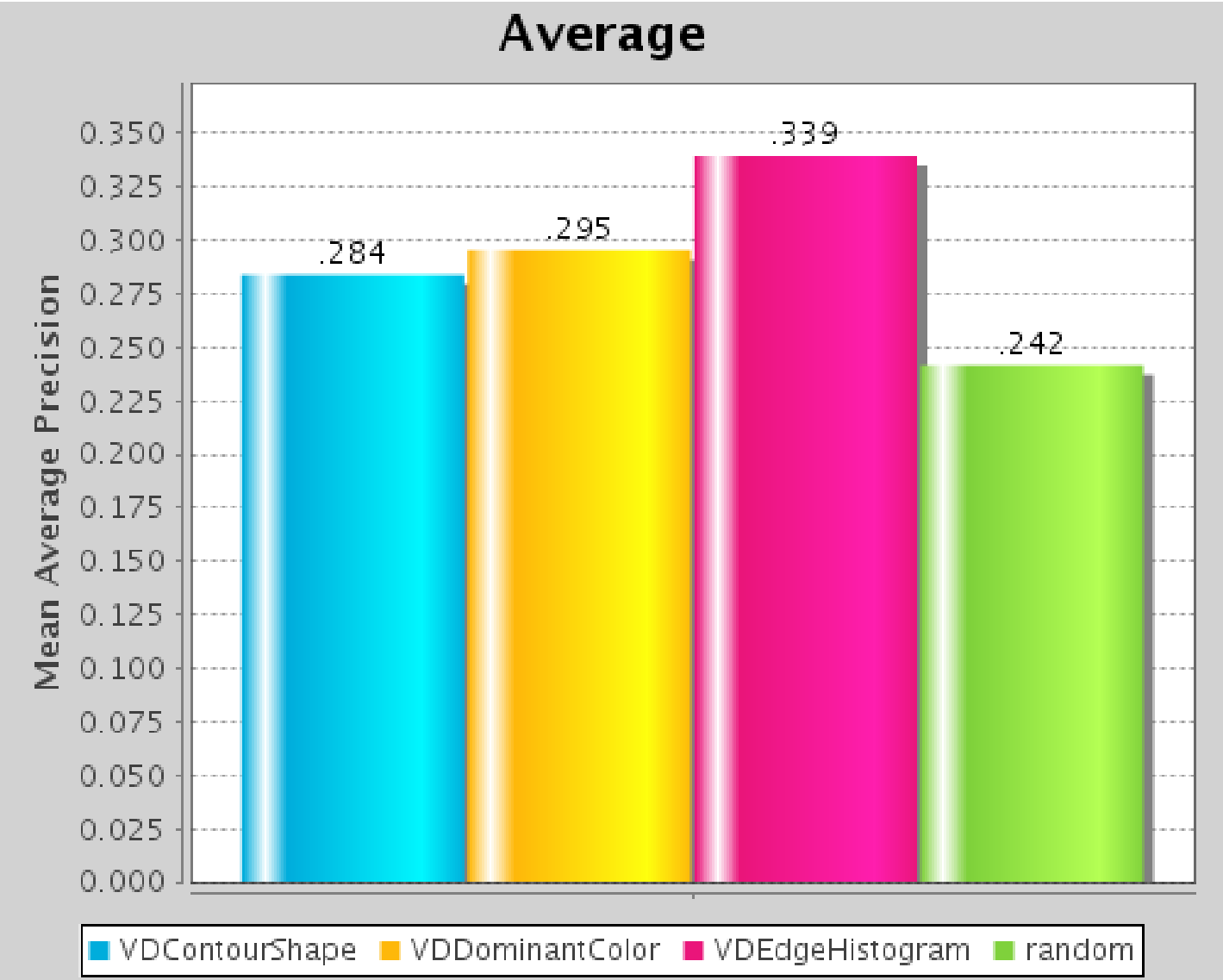
Plotted by considering different *k*'s



Grup de Processament de la Imatge
Universitat Politècnica de Catalunya (UPC)

5. Region features

Image Retrieval (global scale)

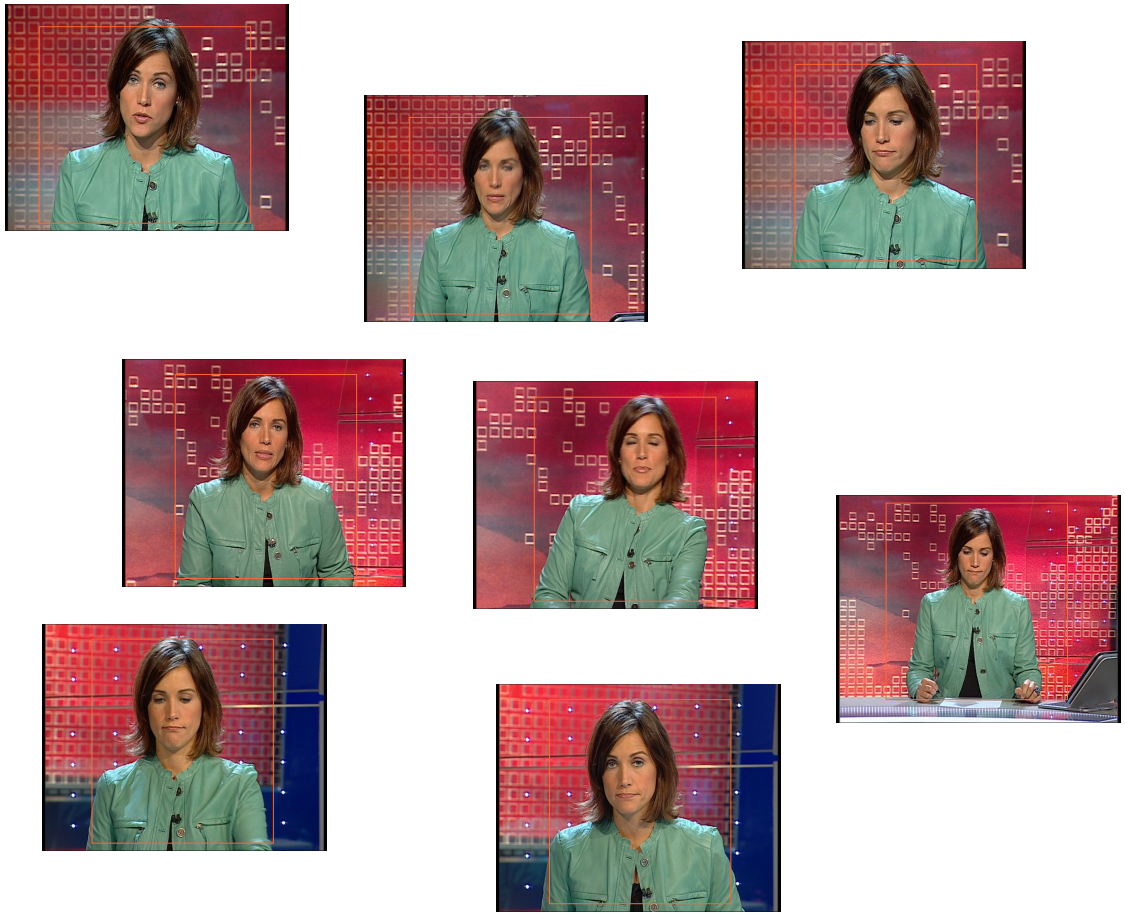


Evaluation

2. Object Detection (local scale-rough)

User Query (local)

Retrieved boxes in relevant images

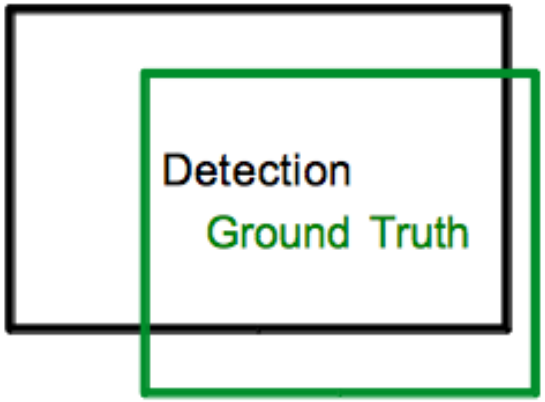


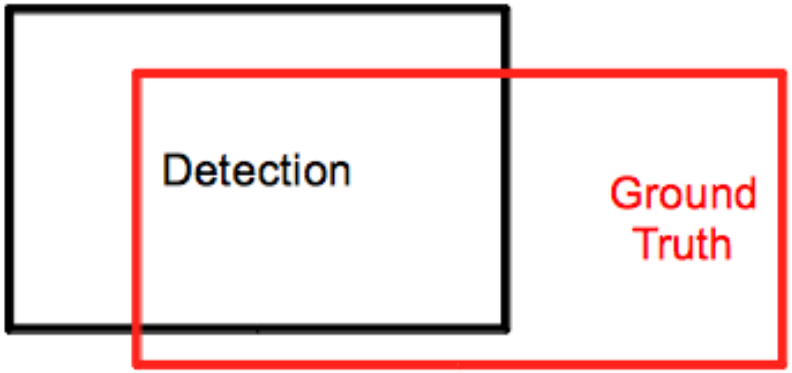
Grup de Processament de la Imatge
Universitat Politècnica de Catalunya (UPC)

Evaluation

2. Object Detection (local scale-rough)

- Pascal criterion requires a minimum of 50% overlapping in the union of detected object box and ground truth box.



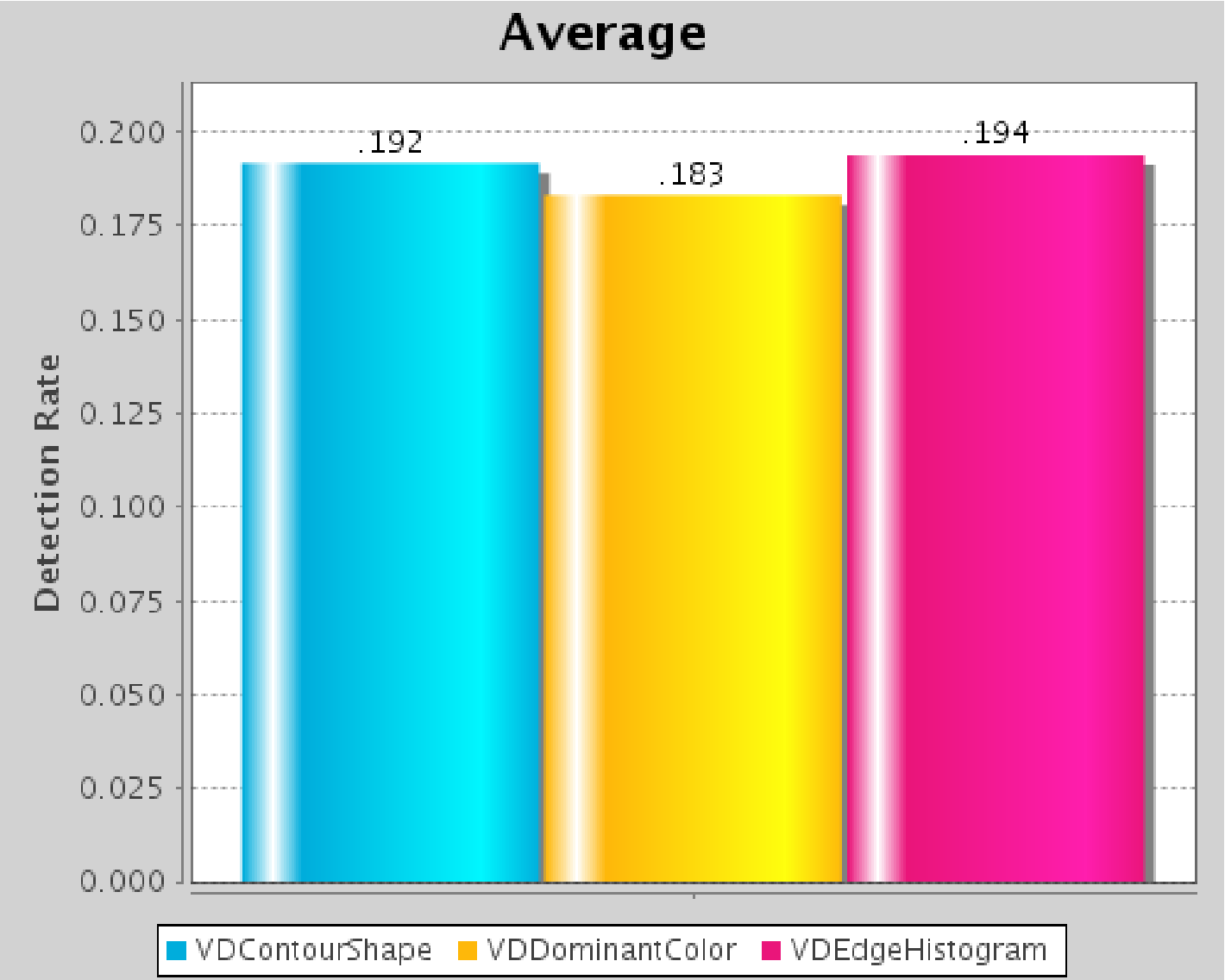


$$Detection\ rate = \frac{|\{correct\ detections\}|}{|\{relevant\}|}$$

PASCAL2
Pattern Analysis, Statistical Modelling and Computational Learning

5. Region features

Object Detection (local scale-rough)

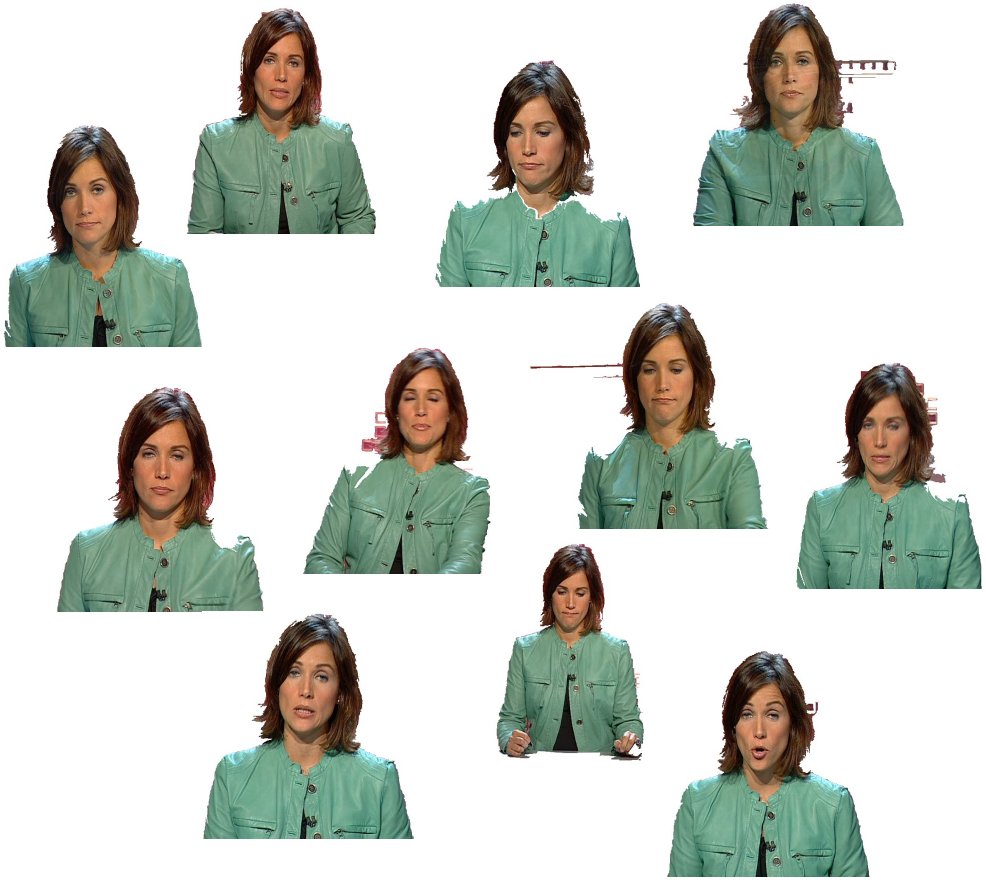


Evaluation

3. Object Segmentation (local scale-precise)

User Query (local)

Retrieved regions
from correct detections

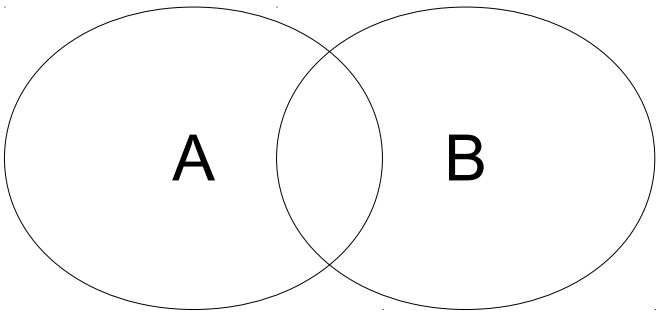


Evaluation

3. Object Segmentation (local scale-precise)

- Jaccard index-J(A,B) compares retrieved region and ground truth mask at the pixel level.

$$J(A,B)=\frac{|A\cap B|}{|A\cup B|}$$



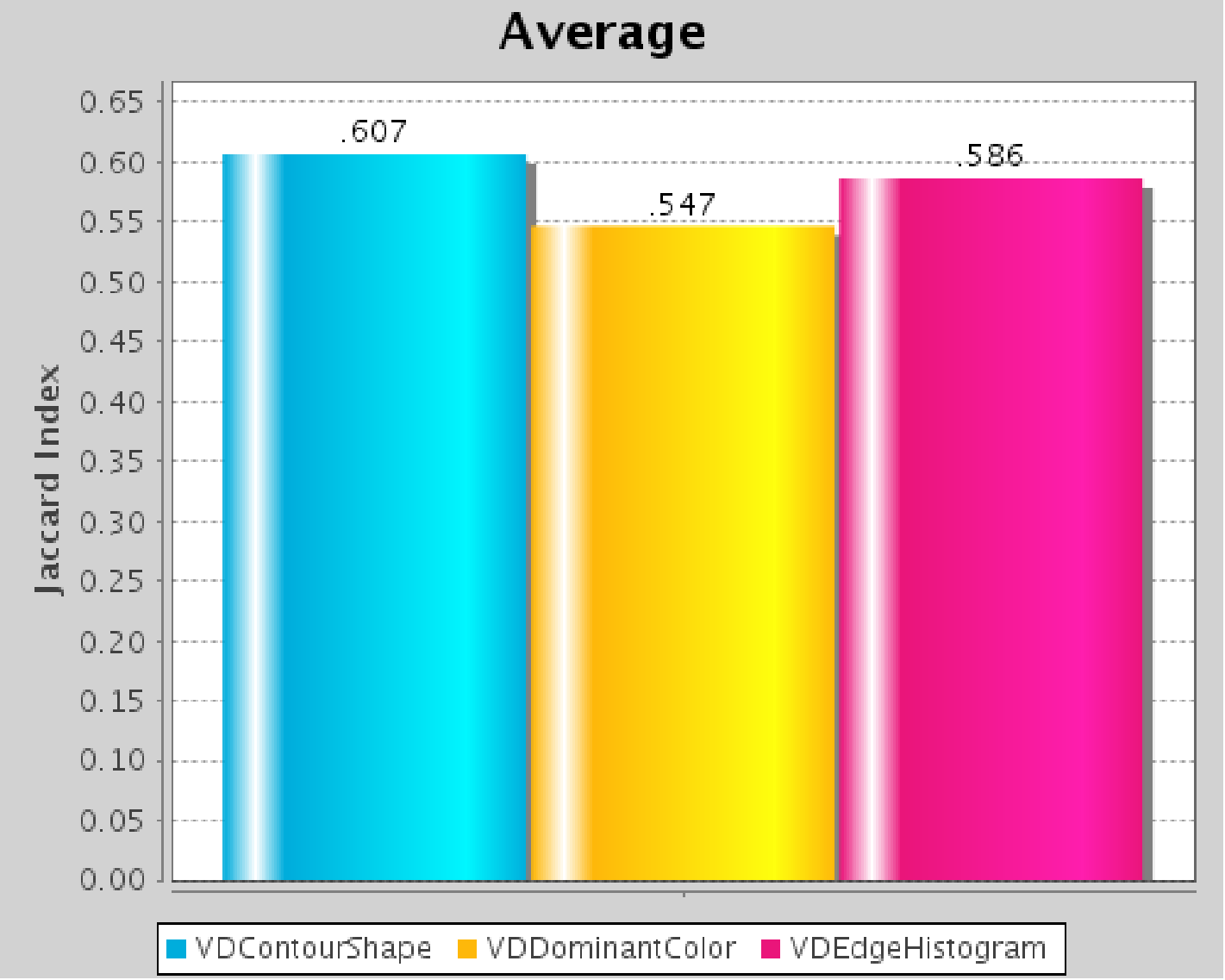
- Mean Jaccard Index combines the J's obtained by a set of correct detections D.

$$\bar{J}=\frac{1}{|D|}\sum_{i\in D} J_i$$

PASCAL2
Pattern Analysis, Statistical Modelling and Computational Learning

5. Region features

Segmentation (local scale-precise)

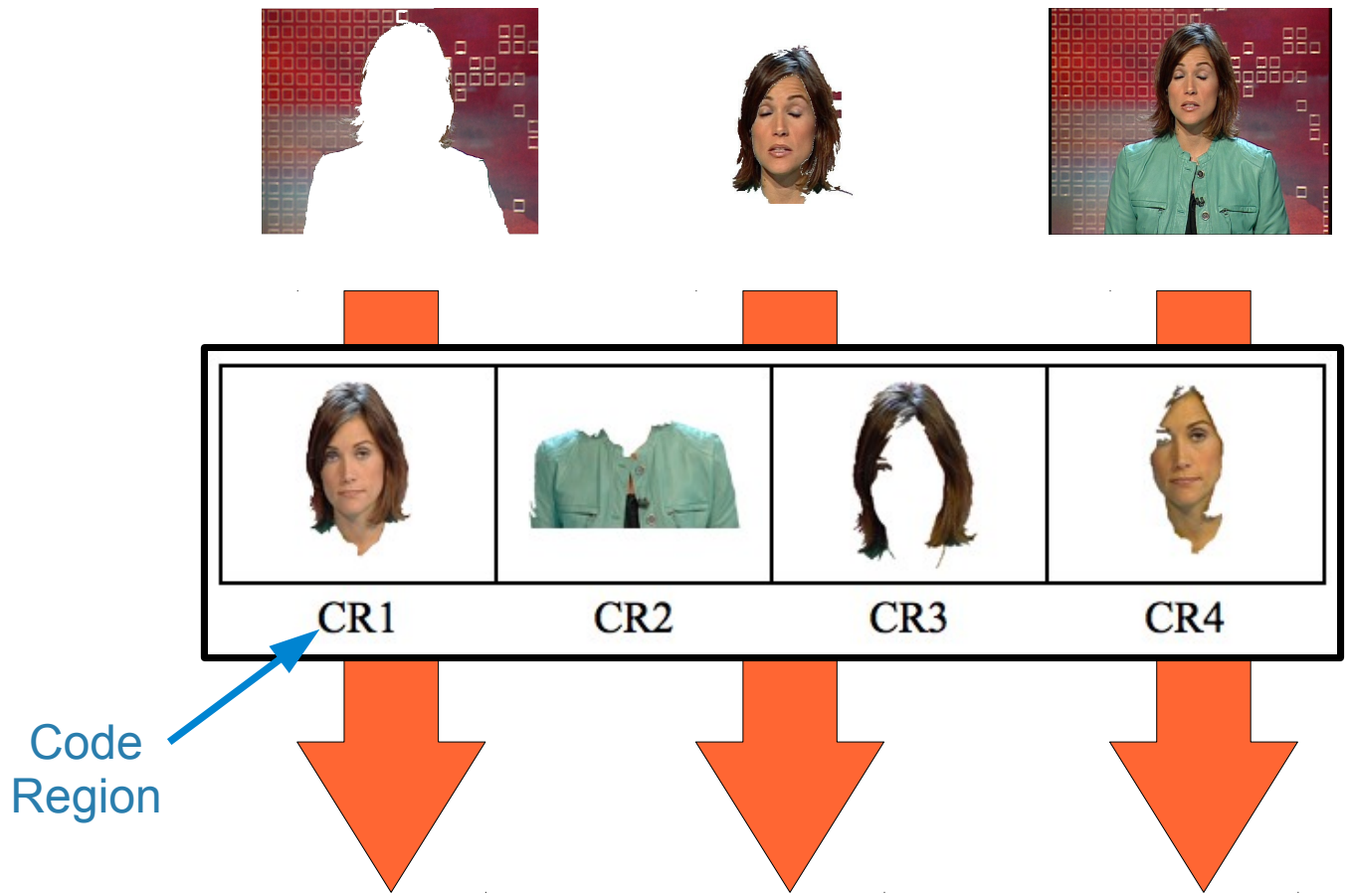


Outline

Grup de Processament de la Imatge
Universitat Politècnica de Catalunya (UPC)

6. Codebooks

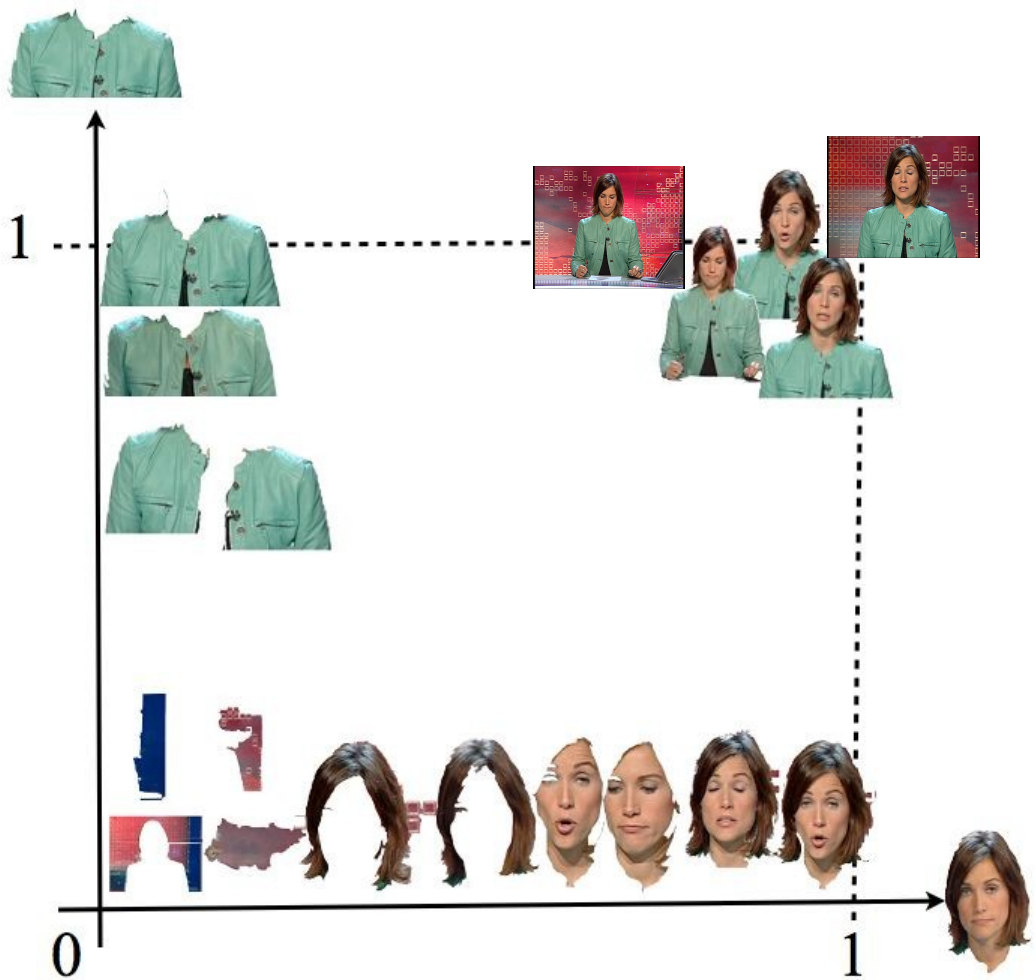
- Visual Word (VW): Vector quantization of observation (point, region...) onto a codebook [Sivic, Zissermann 2003].



6. Codebooks

- Code Regions (CR) define the basis of the **Parts Space**.

  "Regions described by their sub-parts"

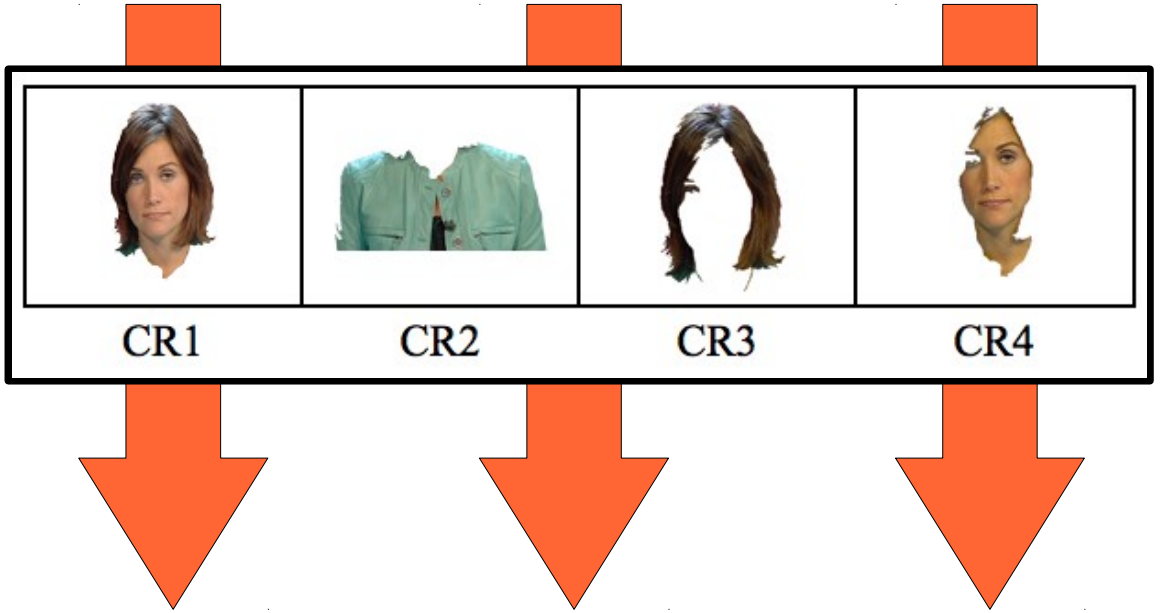


6. Codebooks

(a) Hard quantization $\xrightarrow{\text{SIFT}}$ [Sivic 2003]

Regions

Codebook



Visual Words

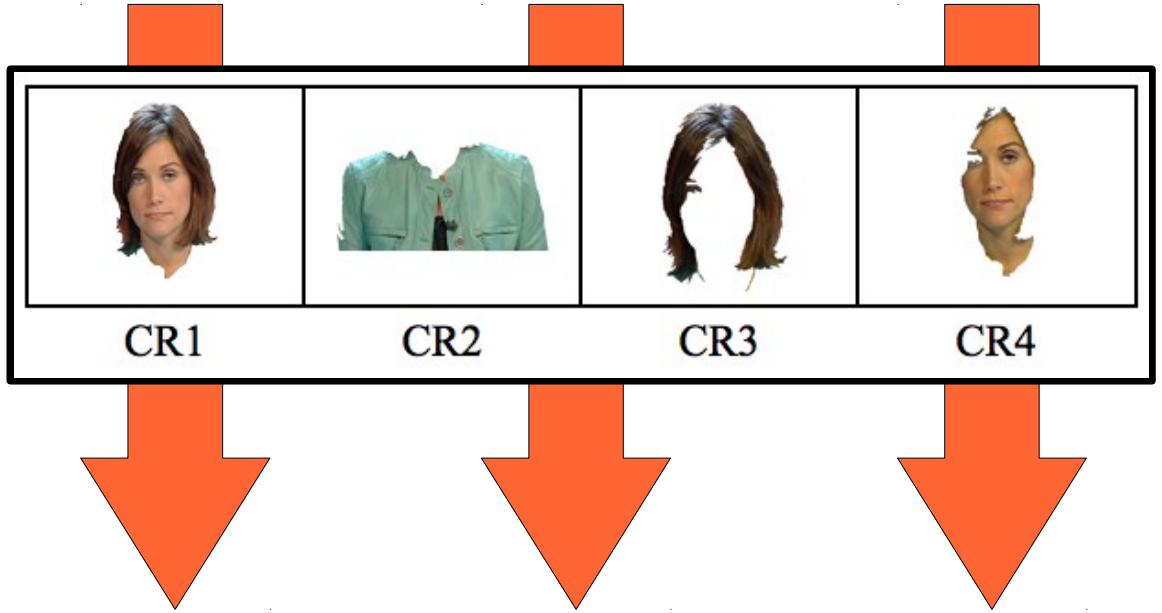
(1,0,0,0) ? ✗ (1,0,0,0) (0,1,0,0) ? ✗

6. Codebooks

(b) Probabilistic quantization [Alpaydin 1998]

Regions

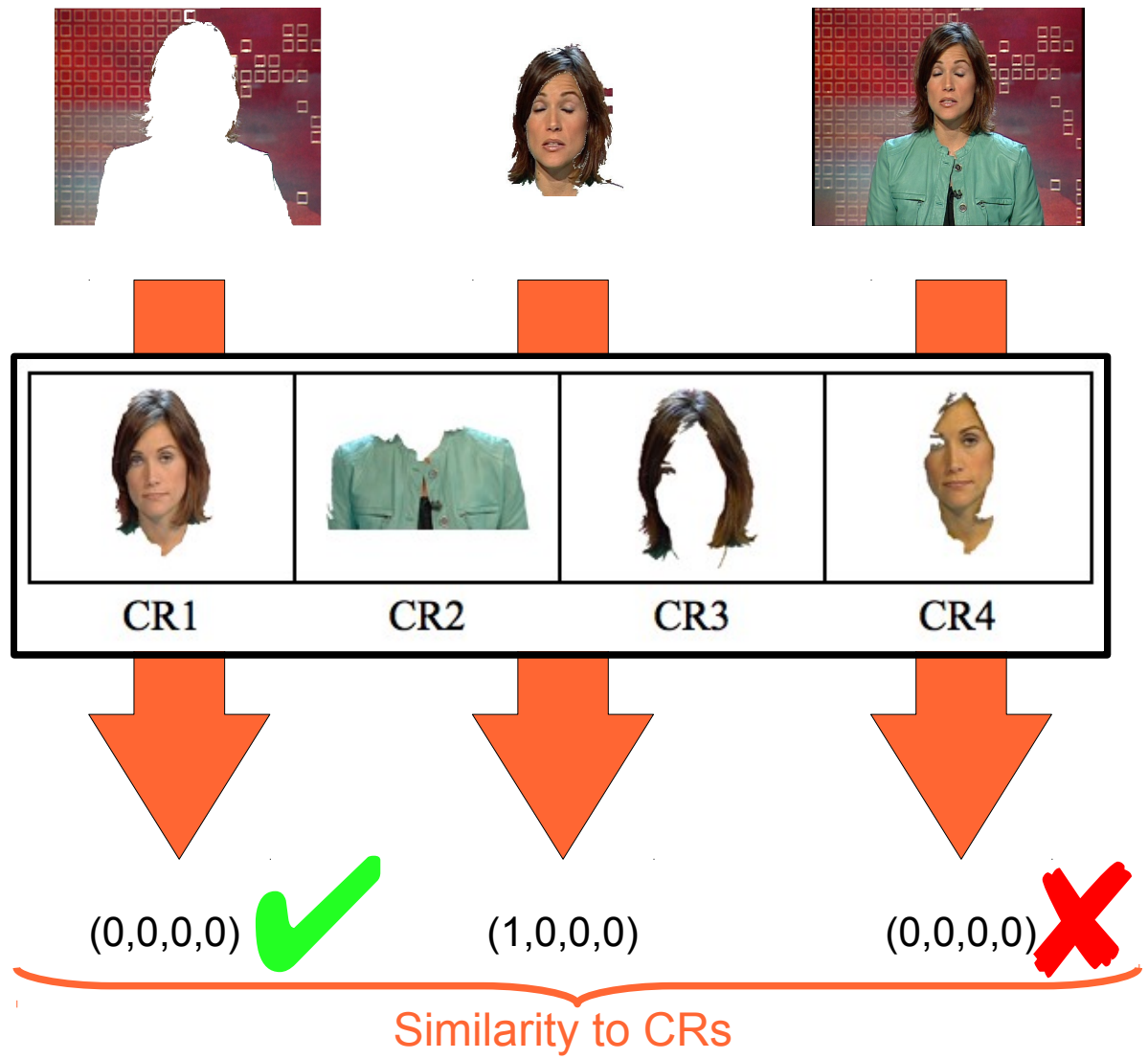
Codebook



Visual Words (0.25,0.25,0.25,0.25) ✗ (1,0,0,0) (0.25,0.25,0.25,0.25) ✗

6. Codebooks

(c) Possibilistic quantization [Bezdek 1995]

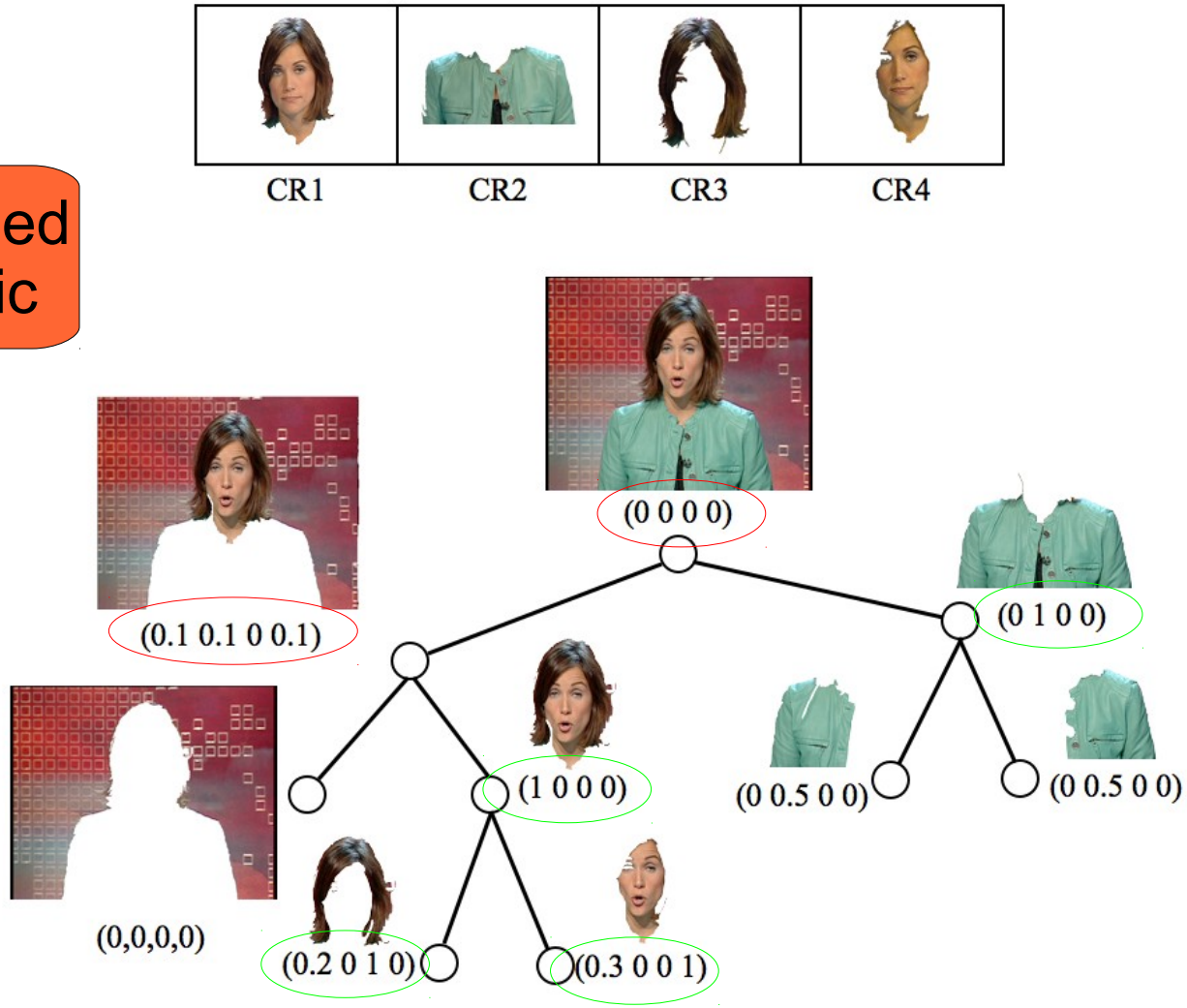


6. Codebooks

- How to define the Parts Space with a given codebook ?

Possibilistic quantization alone characterizes regions, but not their sub-parts.

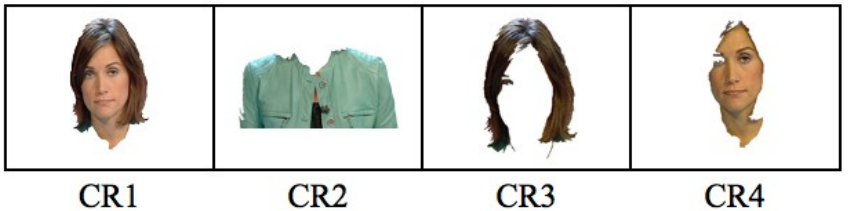
Region-based Possibilistic

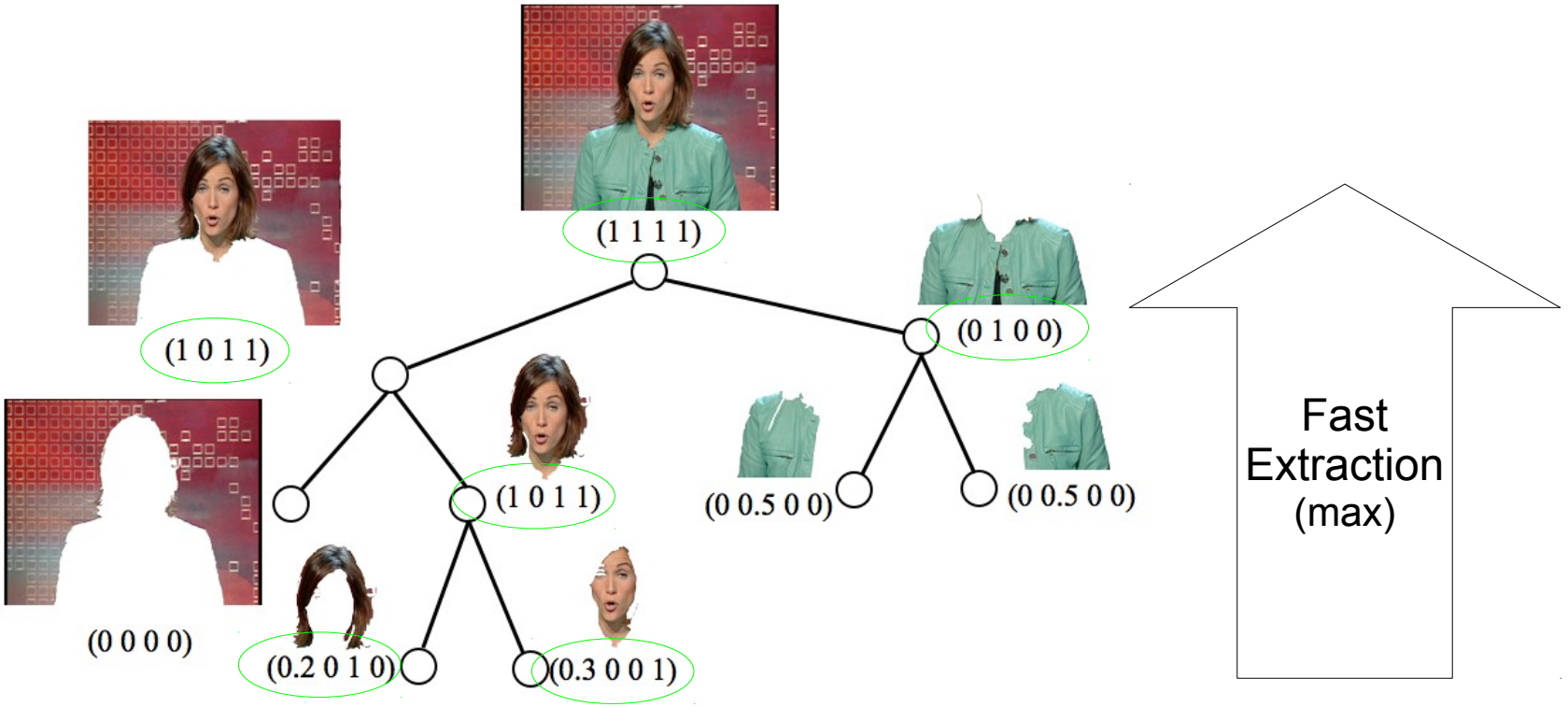


6. Codebooks

- How to define the Parts Space with a given codebook ?

Bottom-up Max Pooling through BPT ✔ [Boureau'10]





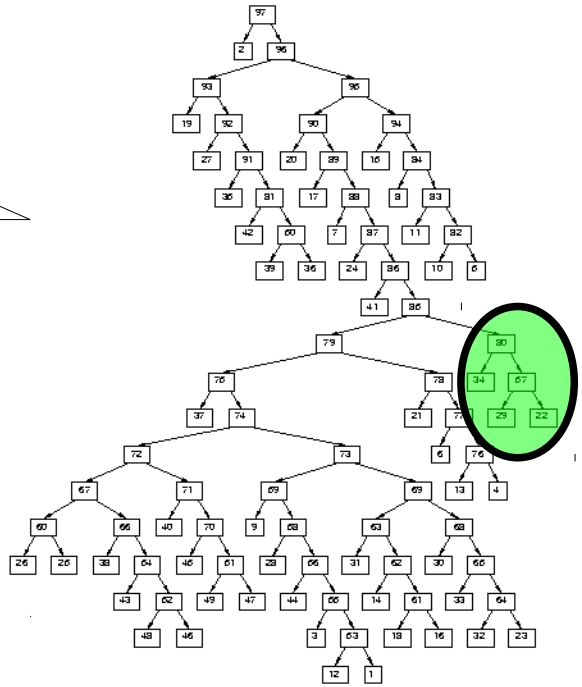
6. Codebooks

→ Hierarchy of Bags of Regions (HBoR)

HBoR

- BPT is exploited to describe sub-trees.

Fast Extraction
(max)



Efficient Analysis
(resiency to splits)

6. Codebooks

- How to generate a codebook ? (1/2)

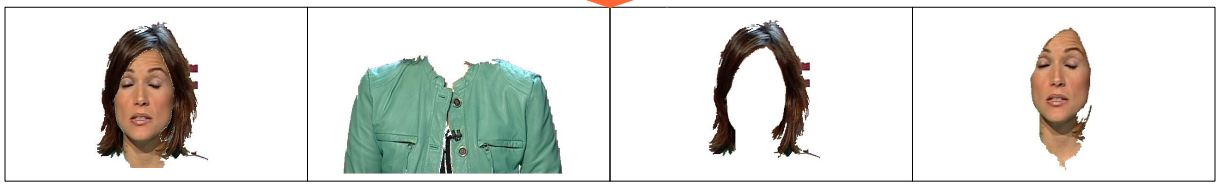
Train set

BPT-based decomposition

Object parts

K-Means Clustering

Codebook



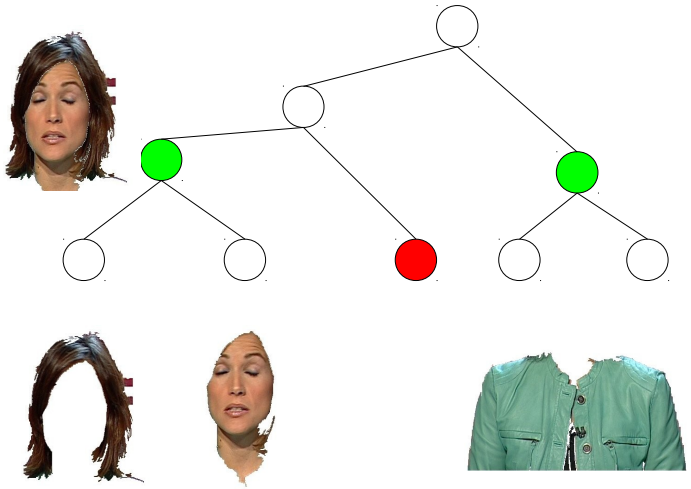
CR_1 CR_2 CR_3 CR_4

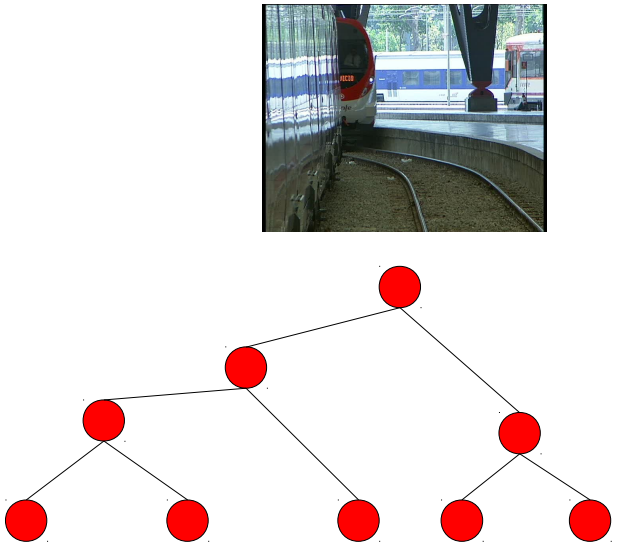
6. Codebooks

- How to generate a codebook ? (2/2)

a) Foreground and background regions - “generic”

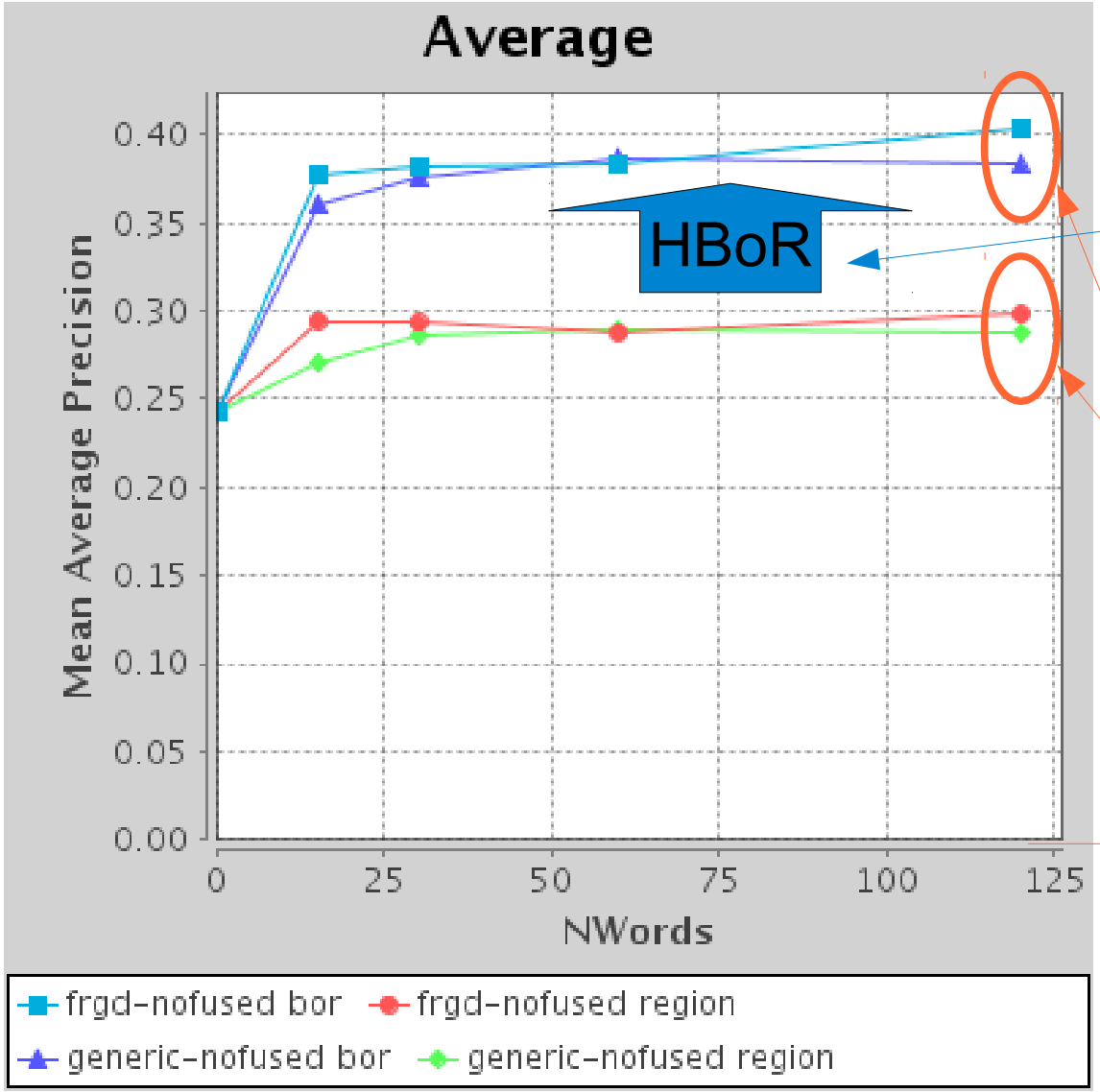
b) Only object regions - “frgd”





6. Codebooks

1. Image Retrieval (global scale) - Texture



HBoR features outstand region-based for shape & texture (not dominant colors).

Little differences between generic vs frgd codebooks (5 classes in ETHZ).

Codebook size

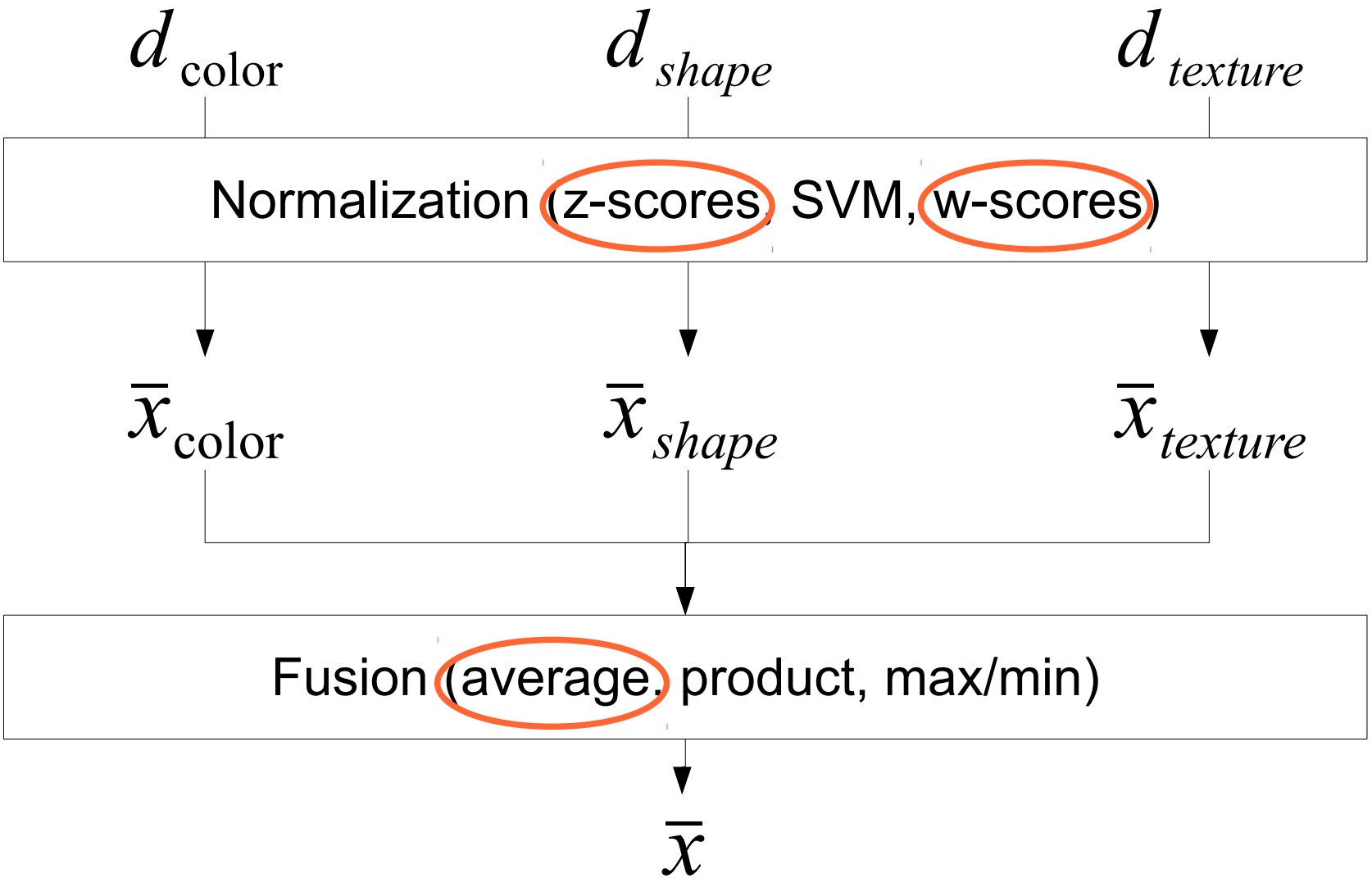
Grup de Processament de la Imatge
Universitat Politècnica de Catalunya (UPC)

Outline

Grup de Processament de la Imatge
Universitat Politècnica de Catalunya (UPC)

5. Region Features + 7. Fusion

- Problem: Similarity scores between descriptors are not comparable.
- Solution:



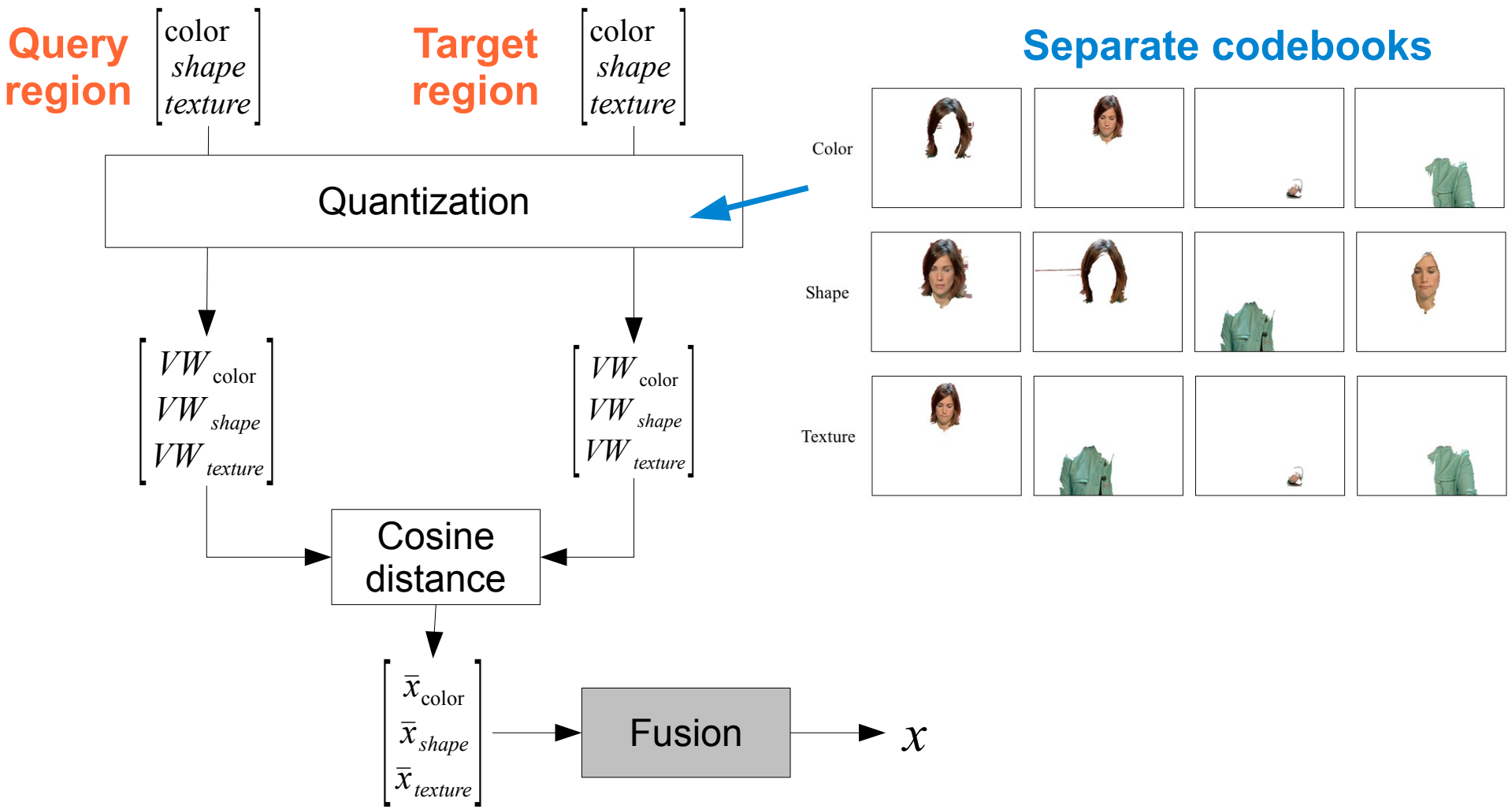
Outline

Grup de Processament de la Imatge
Universitat Politècnica de Catalunya (UPC)

6. Codebooks + 7. Fusion

- When working with codebooks, when to apply fusion ?

(a) Fusion after quantization - “nofused”

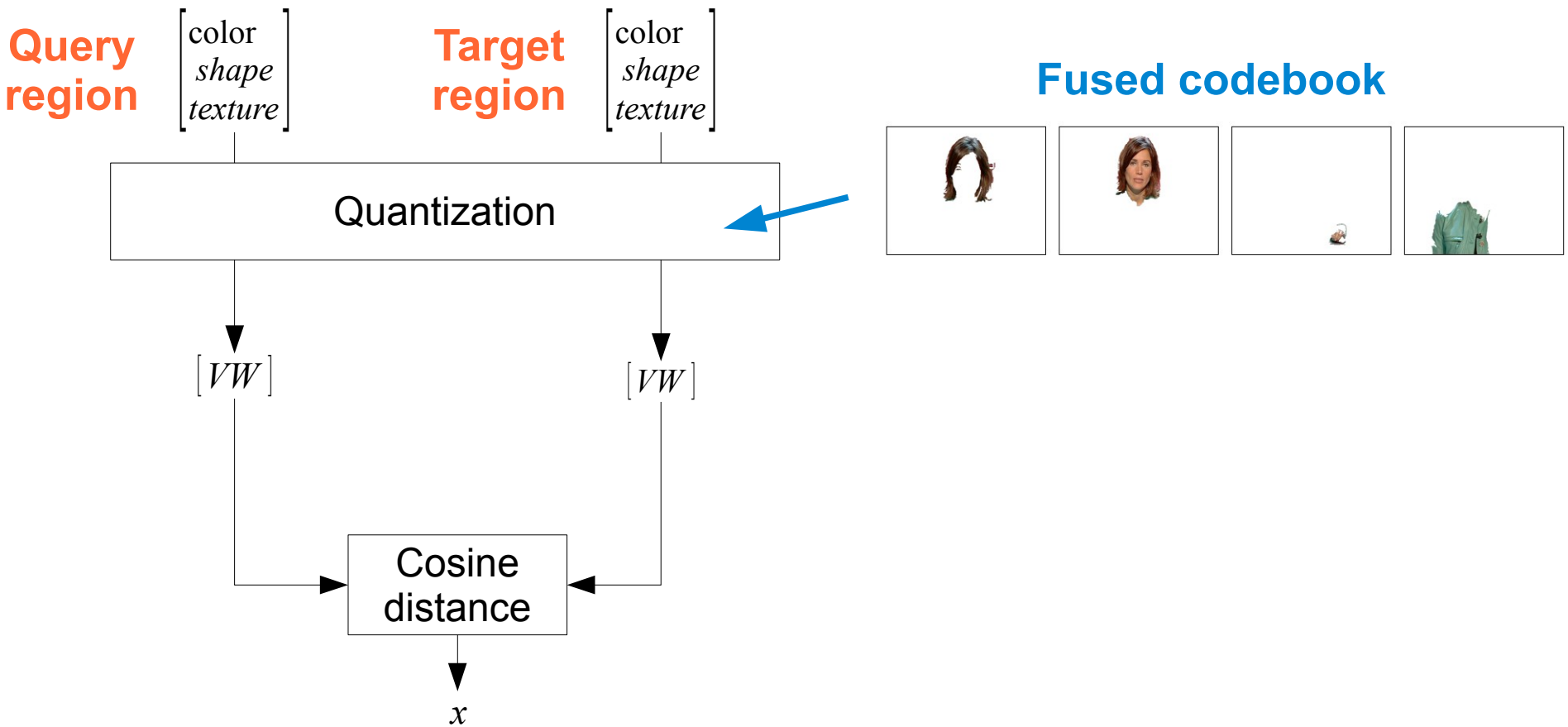


Grup de Processament de la Imatge
Universitat Politècnica de Catalunya (UPC)

6. Codebooks + 7. Fusion

- When working with codebooks, when to apply fusion ?

(b) Fusion during quantization - “fused”

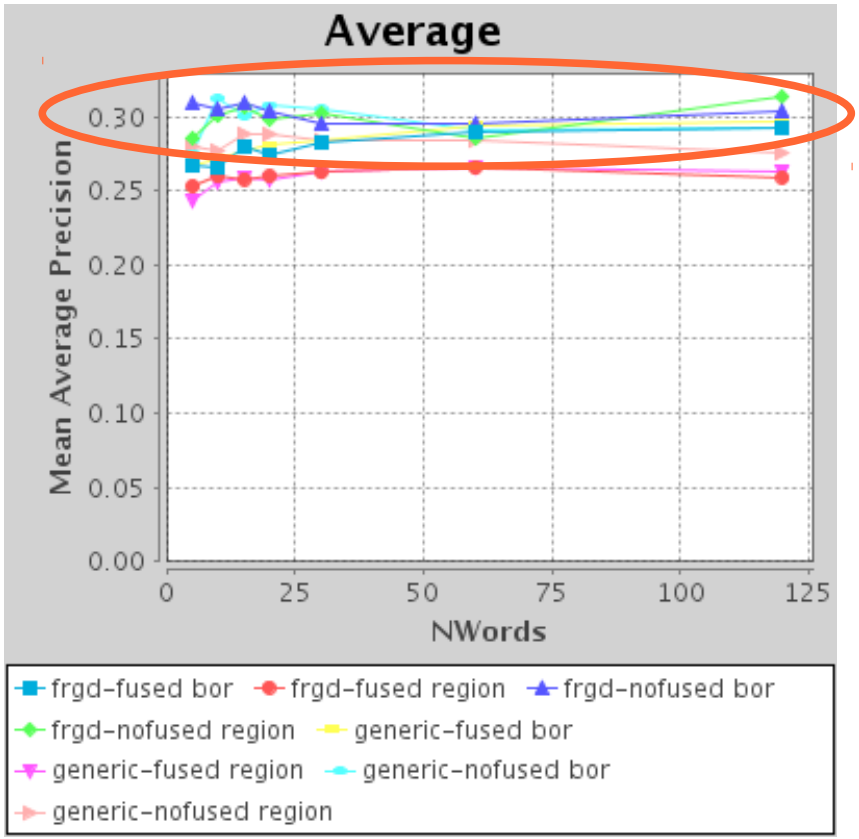


6. Codebooks + 7. Fusion

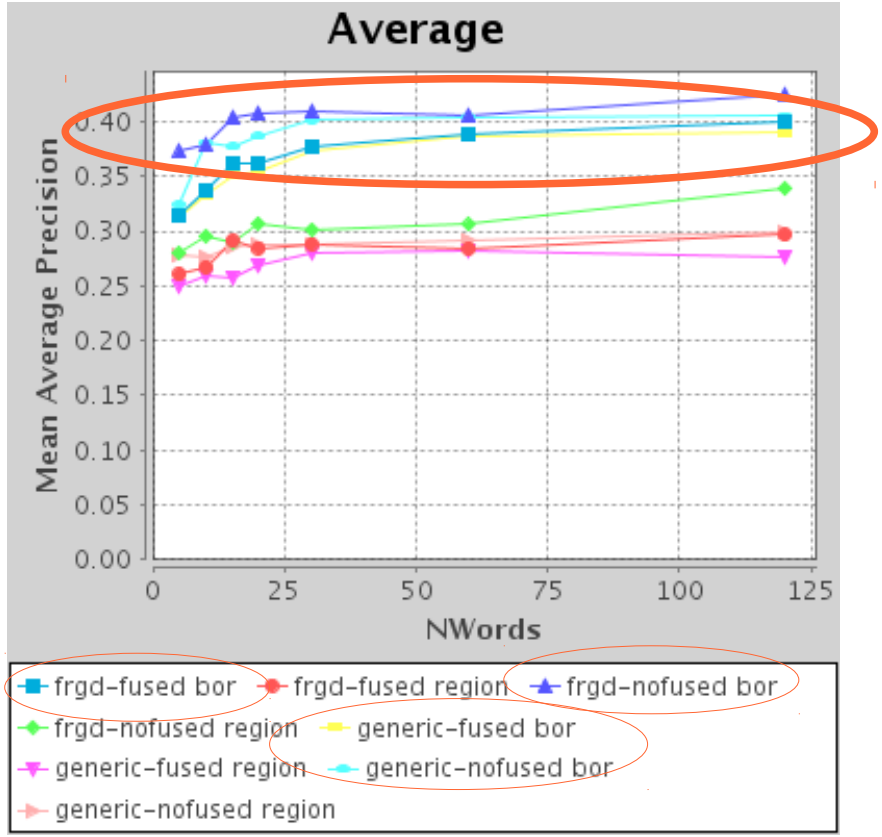
- When working with codebooks, when to apply fusion ?

1. Image Retrieval (global scale)

z-scores



w-scores



W-scores outperform z-scores when working with HBoR.

Grup de Processament de la Imatge
Universitat Politècnica de Catalunya (UPC)

Outline

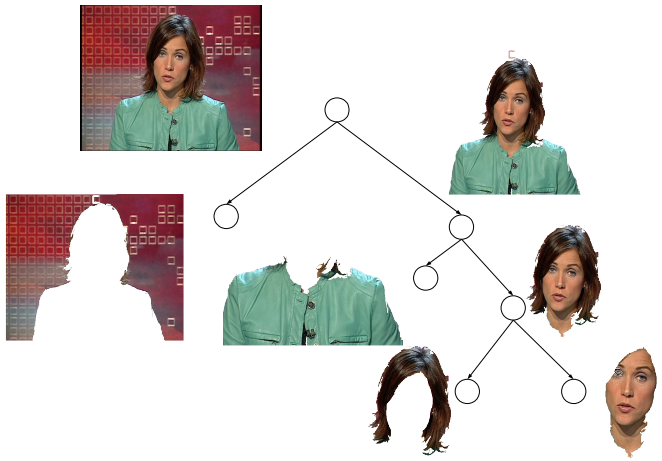
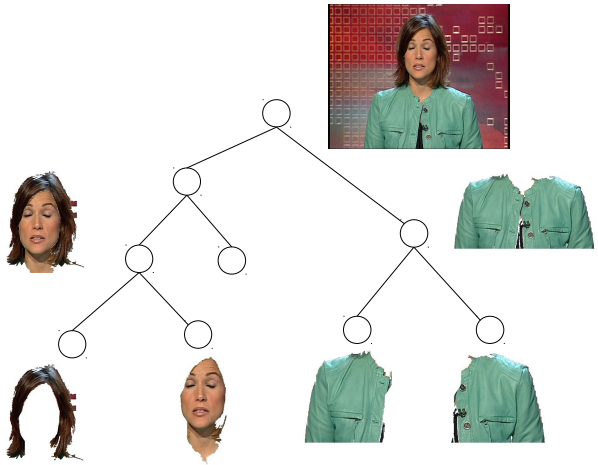
Grup de Processament de la Imatge
Universitat Politècnica de Catalunya (UPC)

Evaluation (in Part III)

- CCMA News dataset.



- Topology diversity on *11 anchors* objects.

Grup de Processament de la Imatge
Universitat Politècnica de Catalunya (UPC)

Outline

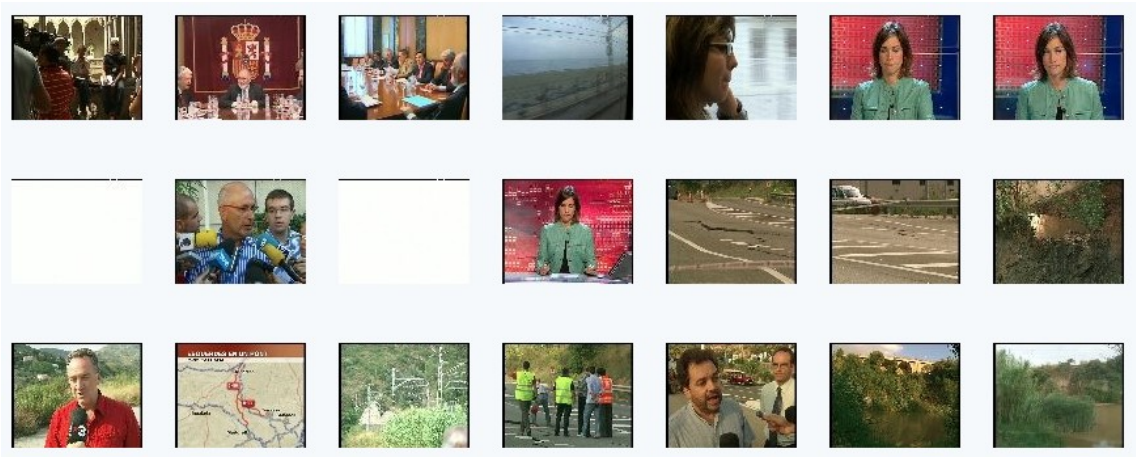
Grup de Processament de la Imatge
Universitat Politècnica de Catalunya (UPC)

8. Partition Tree Matching

- Reminder: The Object Retrieval Problem (single instance)

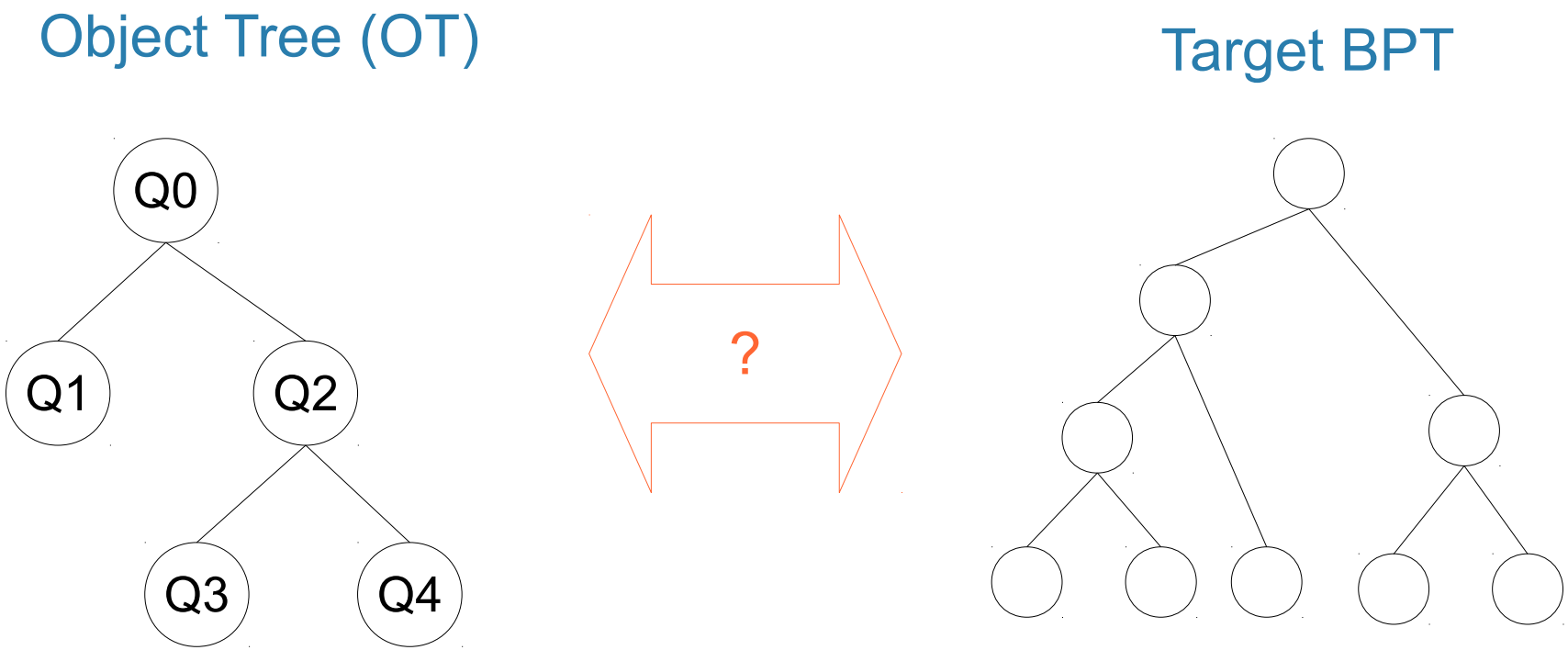
Query object

Target dataset



Ranked list

8. Partition Tree Matching



- <u>Challenges</u>:
  - (a) Distance for a given OT-BPT correspondence
  - (b) Matching algorithm

8. Partition Tree Matching: Distance

- Proposal (ad-hoc): Weighted sum of part matches

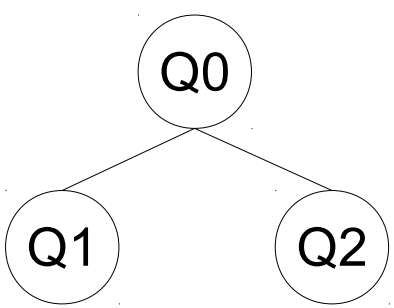
$$d=\sum_{i=0}^{N} w_i d_i \quad \text{, where} \quad w_i=\alpha_i \prod_{j\in Ancestors(Q_i)} \alpha_j=\alpha_i \cdot w_{Parent(Q_i)}$$

$$\alpha_i=\frac{1}{max(1,|C^i|)}\frac{area(i)}{area(P^i)}$$

Children of part i

Parents of part i

- Example:



$$\alpha_0=1/2=w_0$$
$$\alpha_1=area(Q_1)/area(Q_0)$$
$$\alpha_2=area(Q_2)/area(Q_0)$$

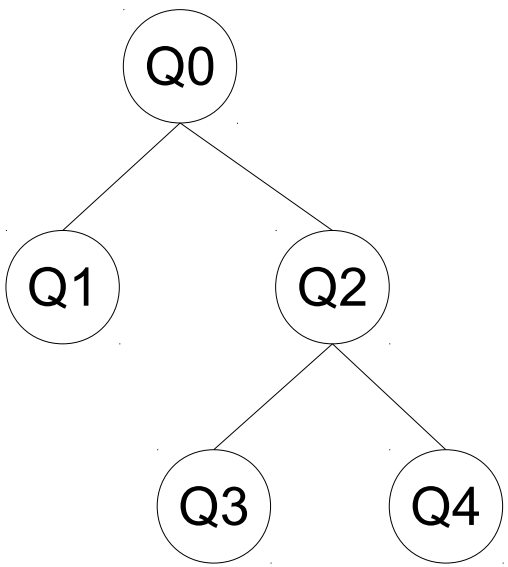
8. Partition Tree Matching: Algorithm

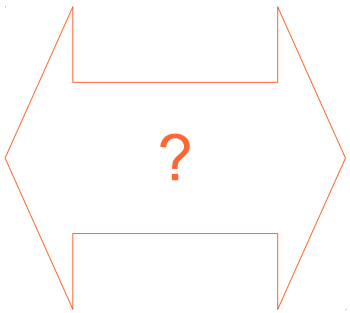
- Two proposals:

  (a) Top-down Matching

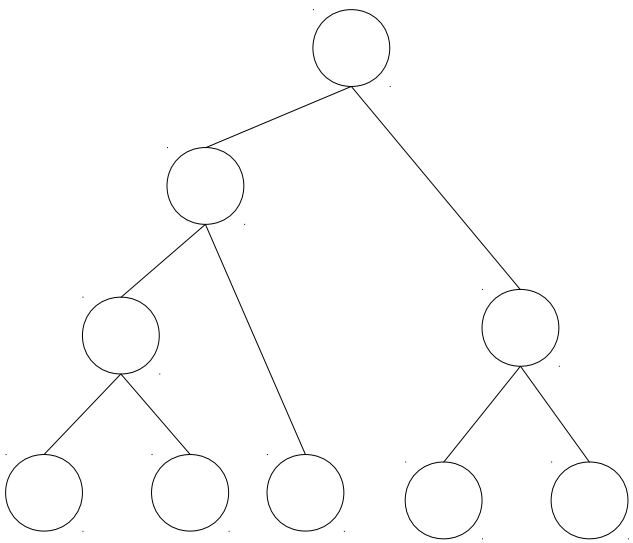
  (b) Bottom-Up Matching

Object Tree (OT)





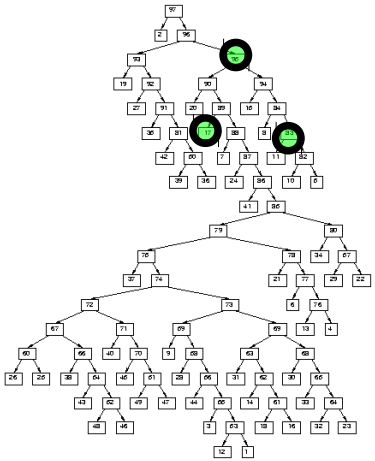
Target BPT



8. Partition Tree Matching: Algorithm

Grup de Processament de la Imatge
Universitat Politècnica de Catalunya (UPC)

- common

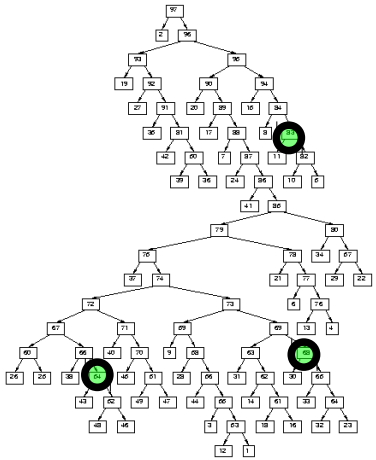


(a) Top-down Matching

- Assumption: All OT nodes are contained in a single sub-BPT.

Sets a baseline

+ common



(b) Bottom-Up Matching

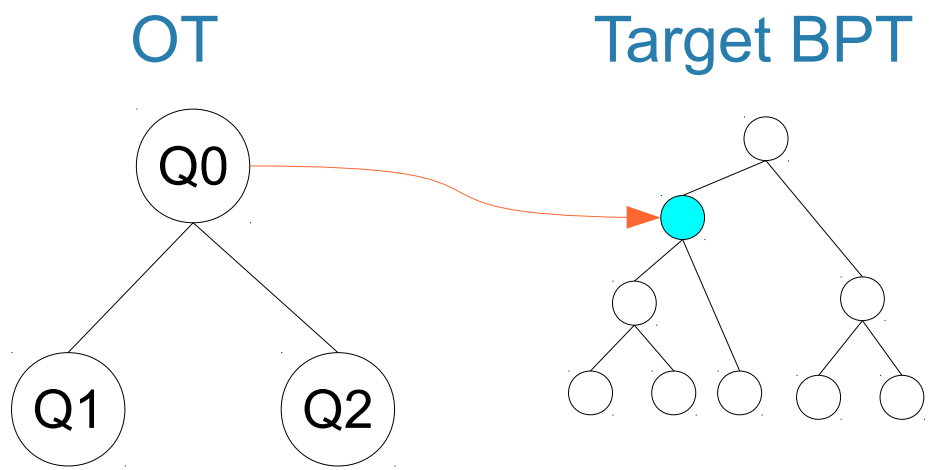
- Assumption: **OT nodes** are spread through different sub-BPTs.

8. Partition Tree Matching: Algorithm

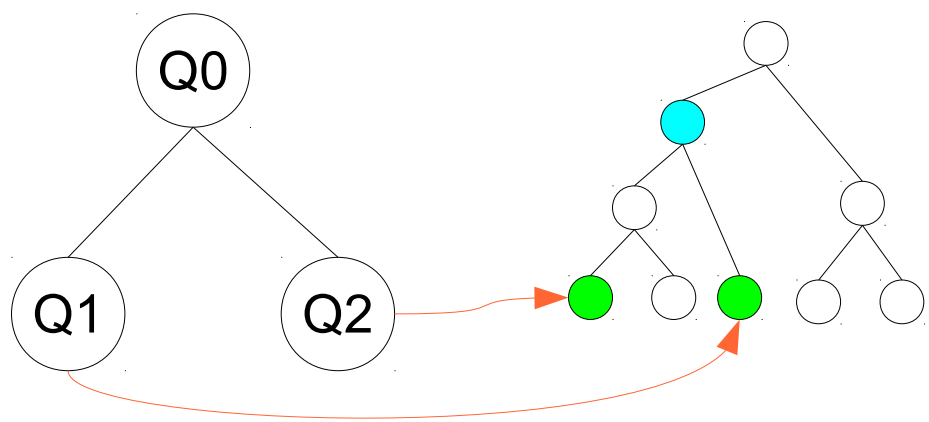
(b) Top-down QT Matching

- Assumption: The best match for the OT is contained in a single sub-tree of the target BPT.

1. Set a match for Q0



2. Find matches for Q1 and Q2

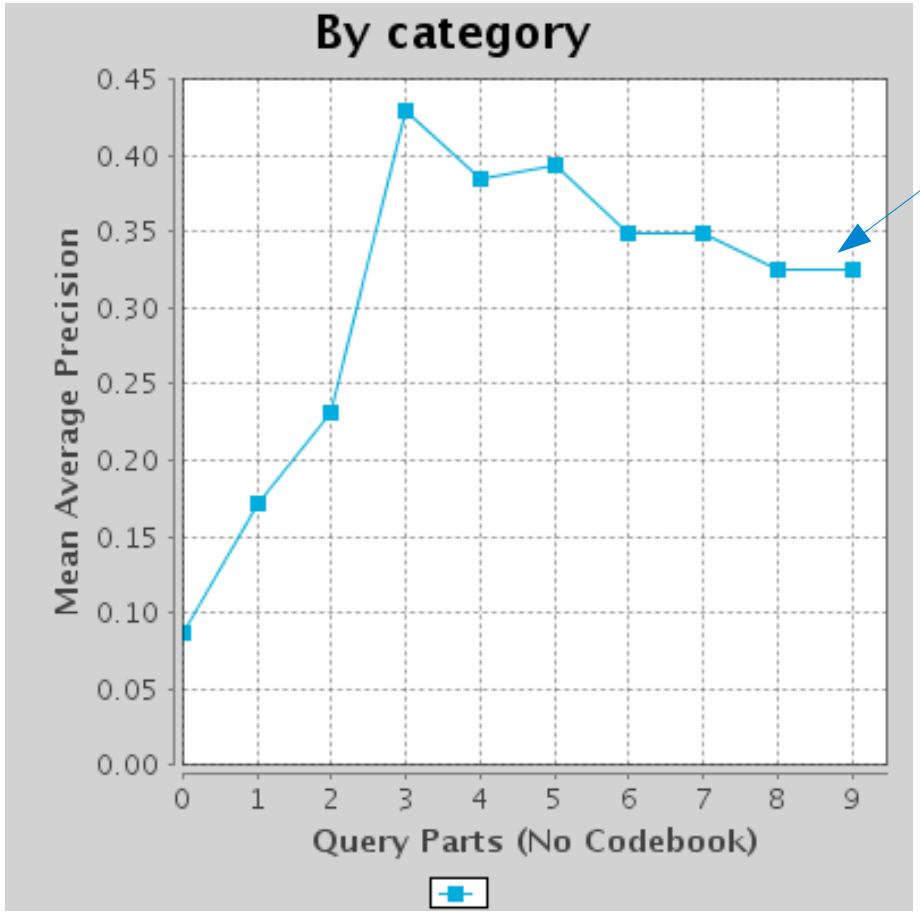


8. Partition Tree Matching

- What is the impact of decomposing the query into parts ?

Image Retrieval (global scale)

MPEG-7 Region



Top-down match

Retrieval gain until the “correct” amount of parts (3).

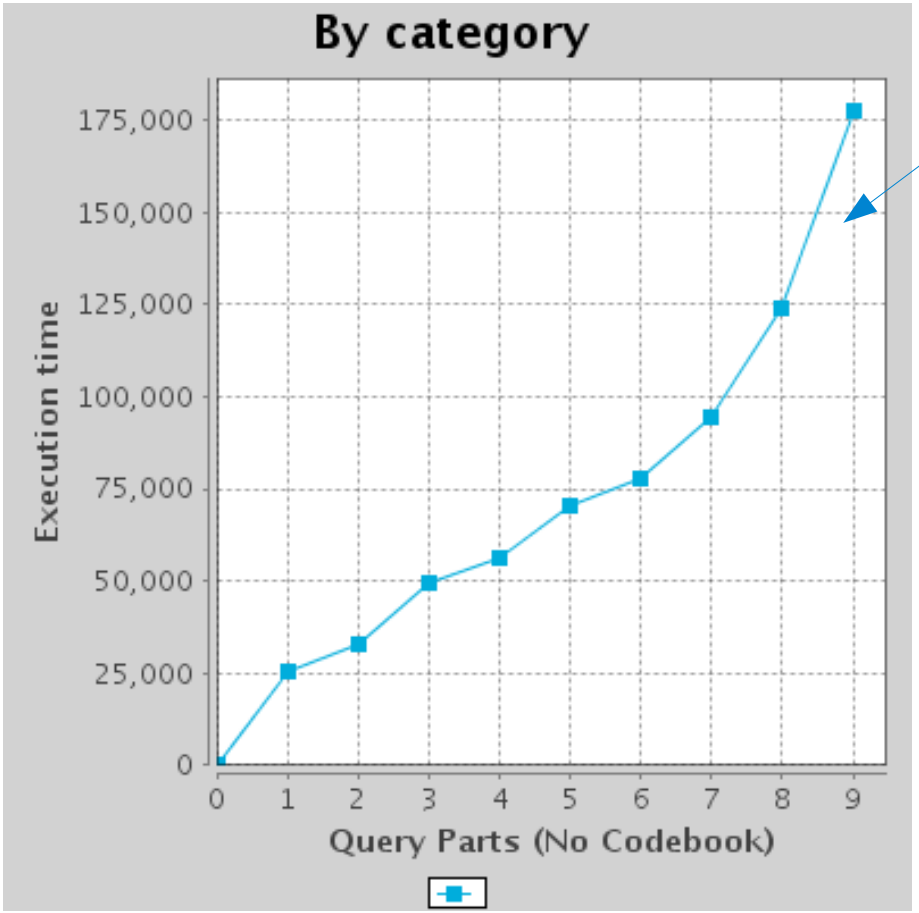
Grup de Processament de la Imatge
Universitat Politècnica de Catalunya (UPC)

8. Partition Tree Matching

- What is the impact of decomposing the query into parts ?

Search time

MPEG-7 Region



Top-down match

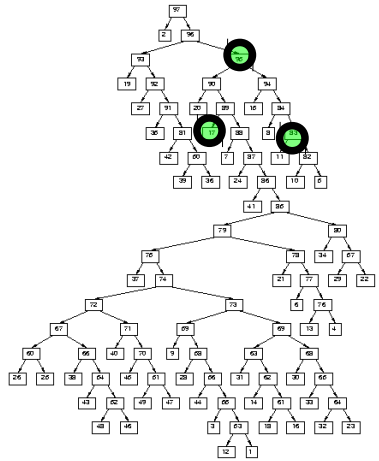
Exponential growth of search time with the amount of parts.

Grup de Processament de la Imatge
Universitat Politècnica de Catalunya (UPC)

8. Partition Tree Matching: Algorithm

Grup de Processament de la Imatge
Universitat Politècnica de Catalunya (UPC)

- common

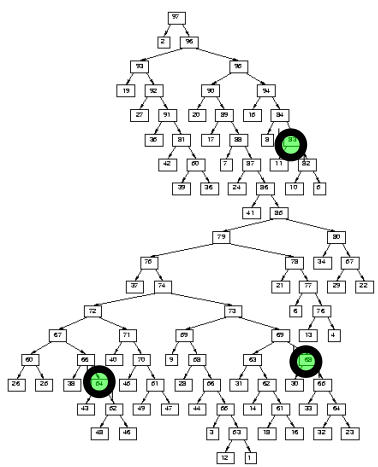


(a) Top-down Matching

- Assumption: All OT nodes are contained in a single sub-BPT.

Sets a baseline

+ common



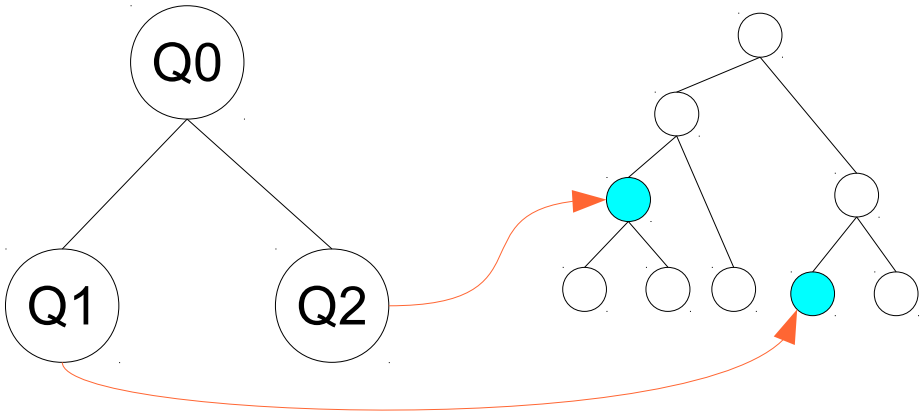
(b) Bottom-Up Matching

- Assumption: **OT nodes** are spread through different sub-BPTs.

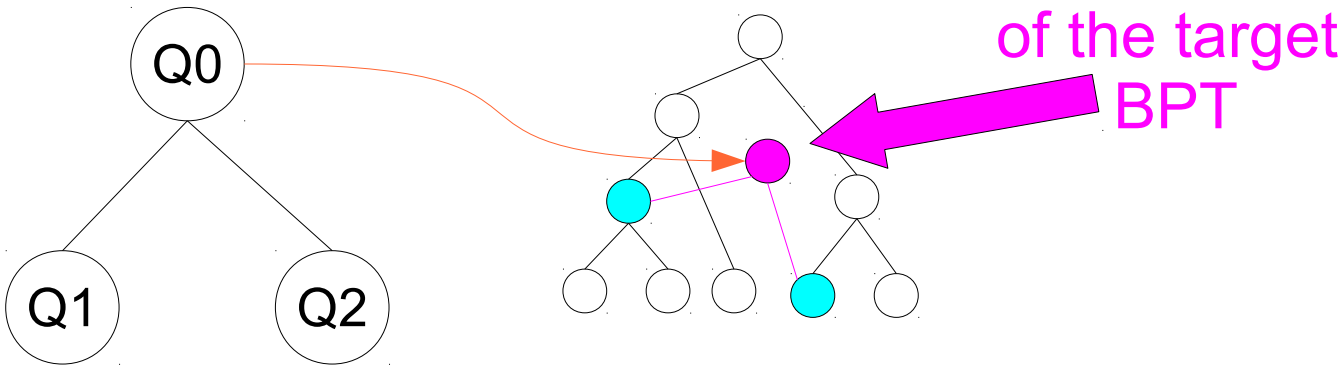
8. Partition Tree Matching: Algorithm

(c) Bottom-UP QT Matching

1. Set a match for Q1 & Q2



2. Assess the union of candidates for Q1 and Q2 to match Q0



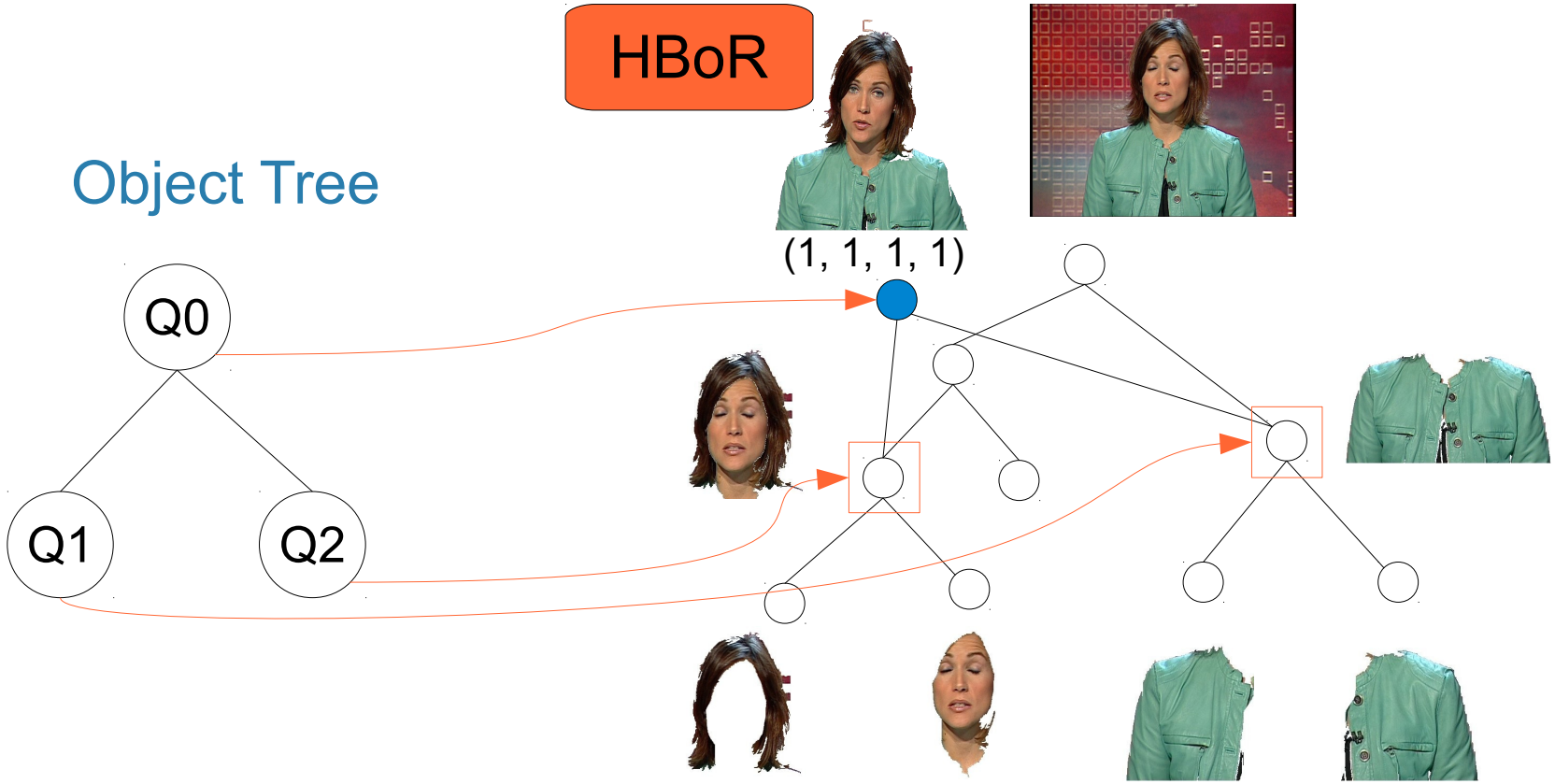
- Weaker assumption that in the top-down case:

  → OT leaves are among the nodes of the target BPT

8. Partition Tree Matching: Algorithm

(c) Bottom-UP QT Matching

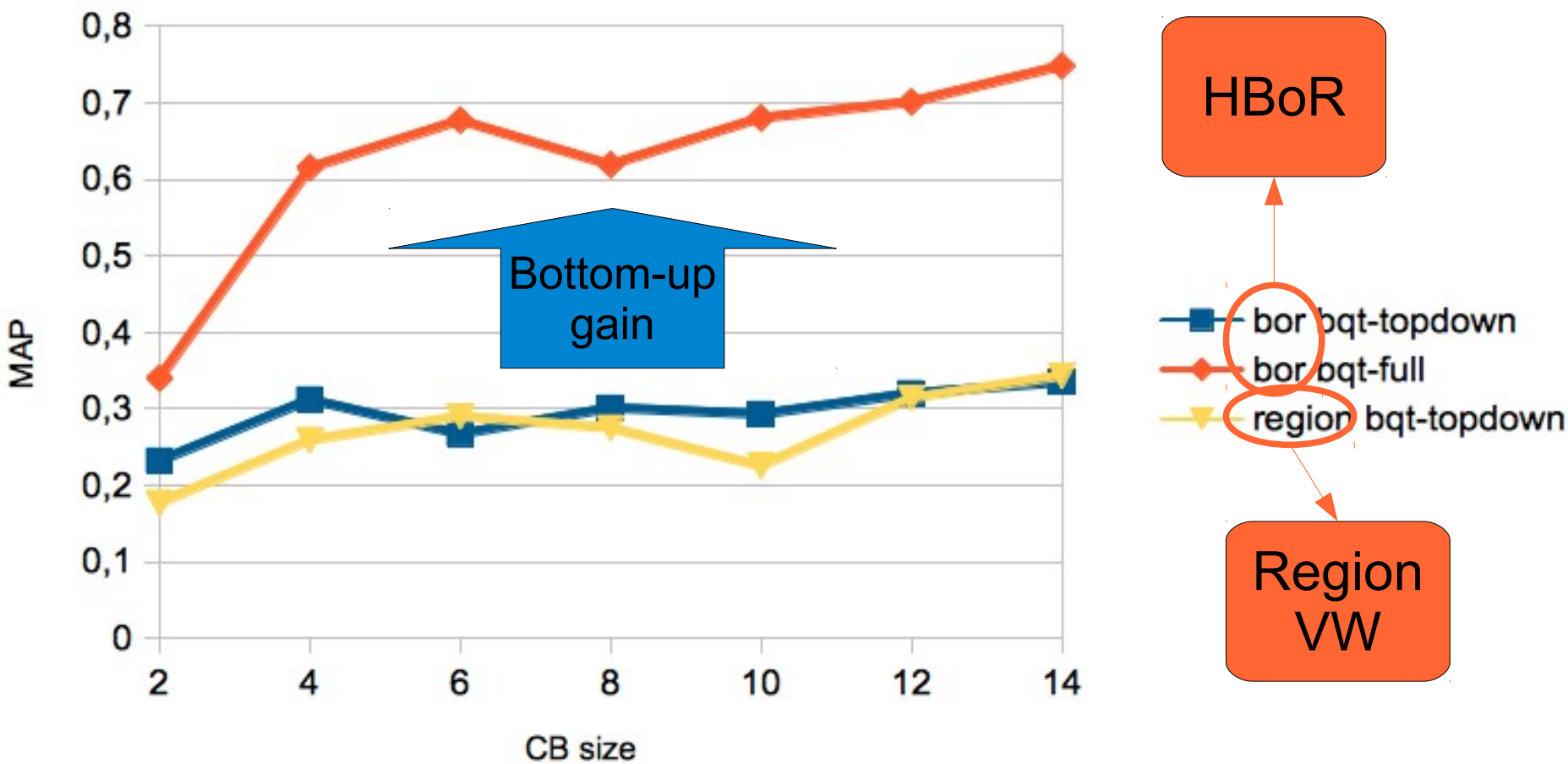
- How to rapidly describe the merged region generated during search time ?



→ Approximate HBoR satisfies this requirement.

8. Partition Tree Matching

- What is the impact of introducing the bottom-up match ?



A clear gain, the object splits in the target BPTs are solved.

Outline

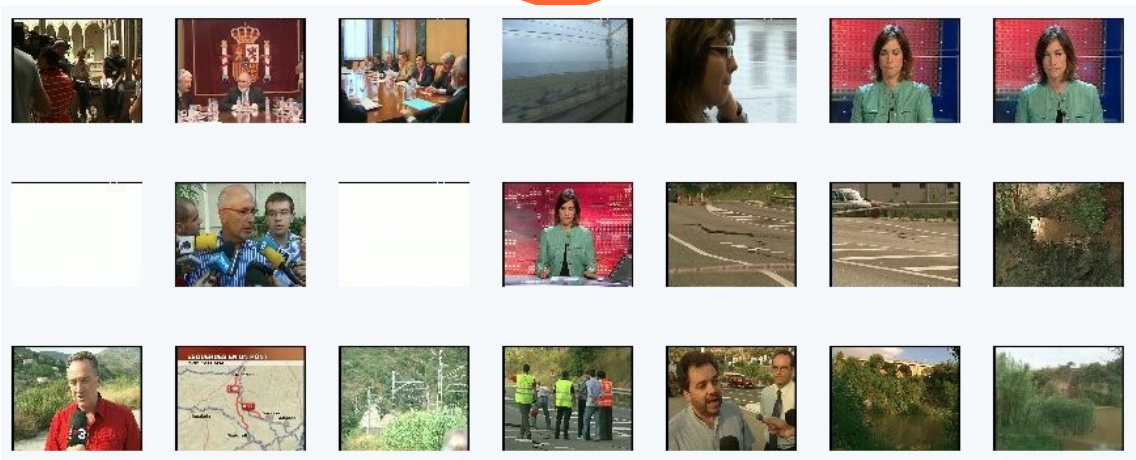
Grup de Processament de la Imatge
Universitat Politècnica de Catalunya (UPC)

9. Part-based Object Model

- Reminder: The Object Retrieval Problem (multiple instances)

Query objects

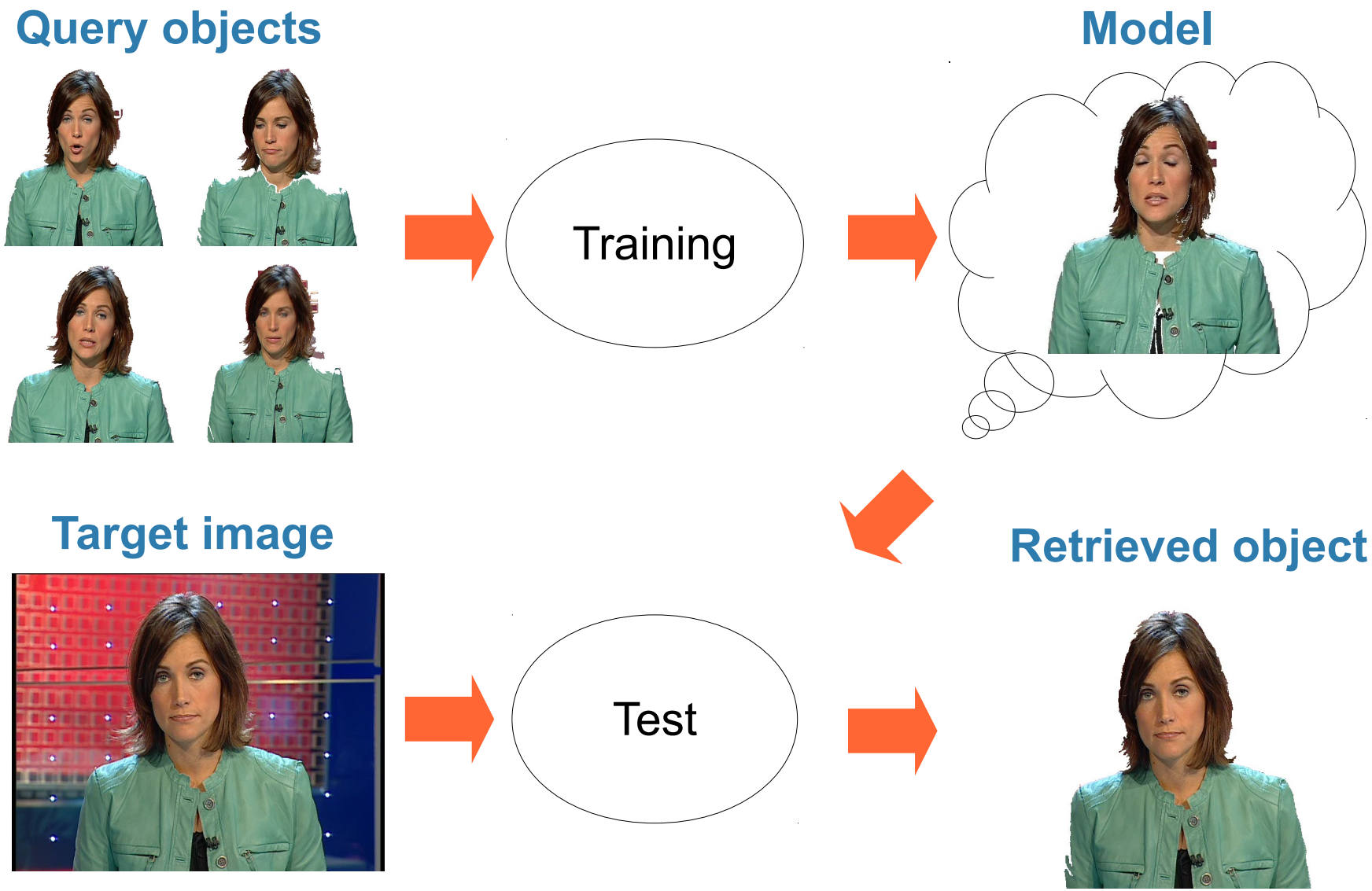
Target dataset



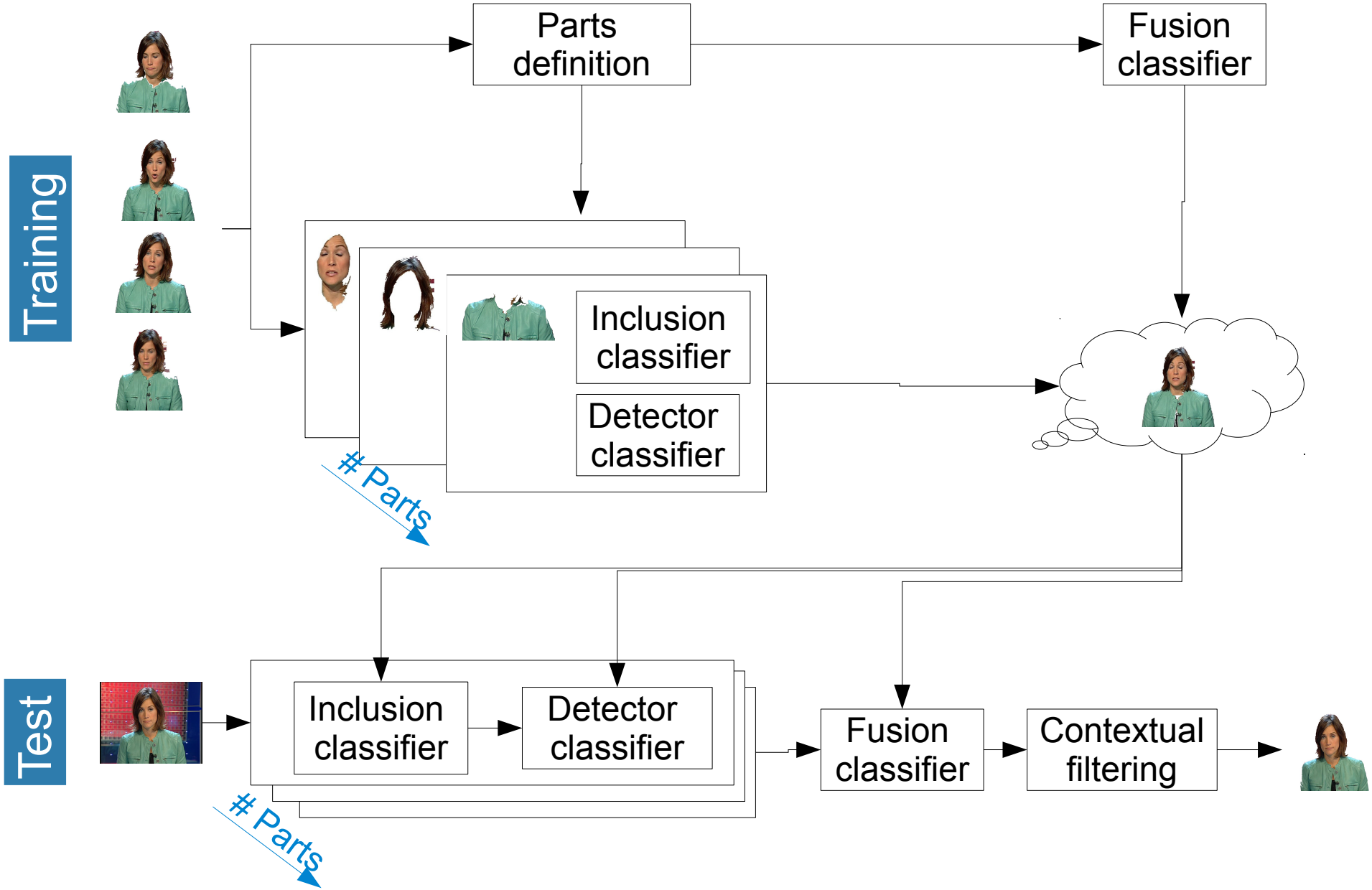
Ranked list

9. Part-based Object Model

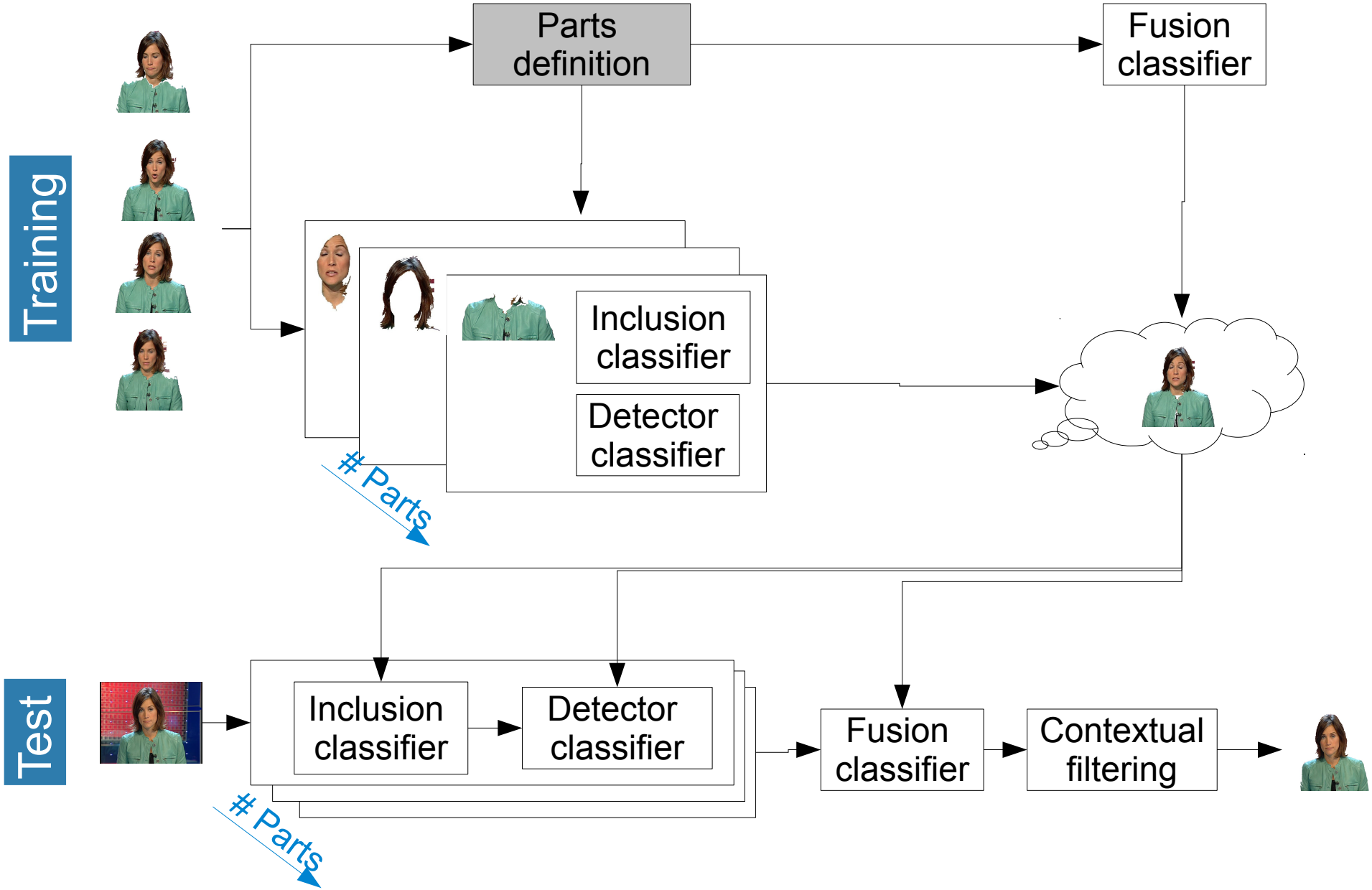
- Model-based retrieval



9. Part-based Object Model



9. Part-based Object Model

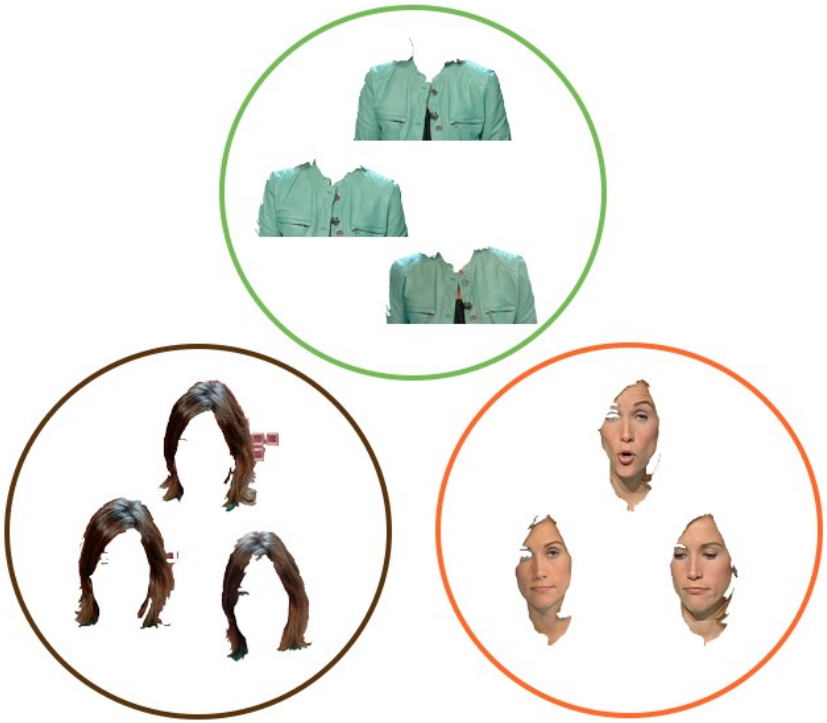


9. Part-based Object Model

Parts definition

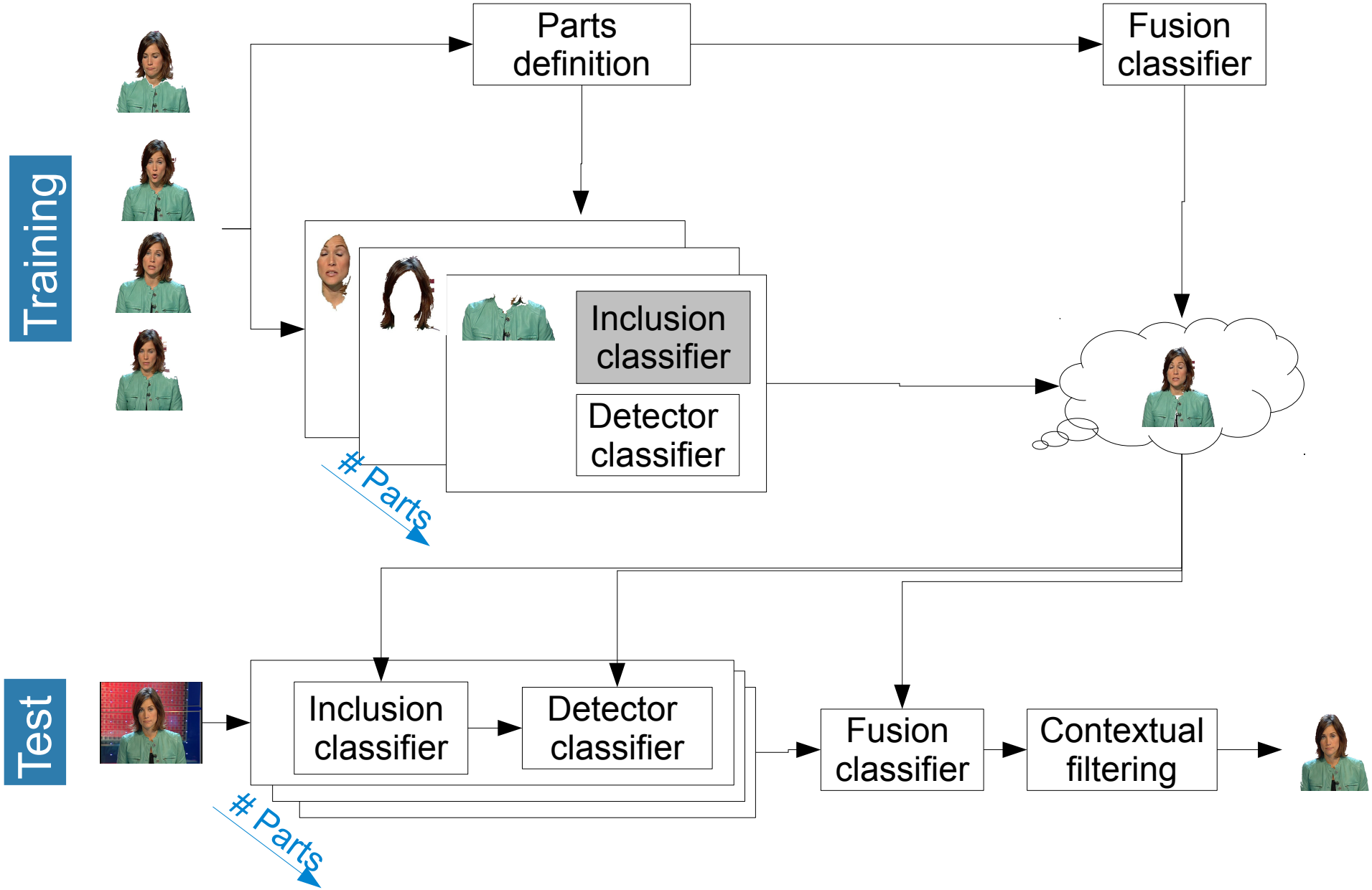
- Unsupervised clustering with Quality Threshold [Heyer 1999].

  Two parameters:
  - Mininum amount of elements in cluster
  - Maximum radius of a cluster



Grup de Processament de la Imatge
Universitat Politècnica de Catalunya (UPC)

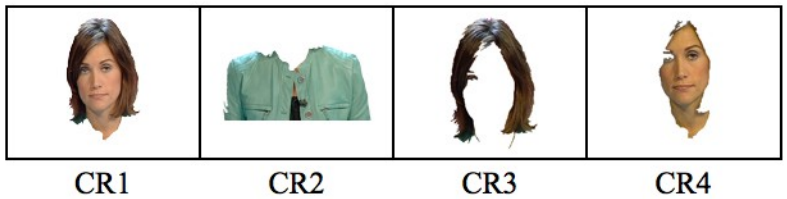
9. Part-based Object Model



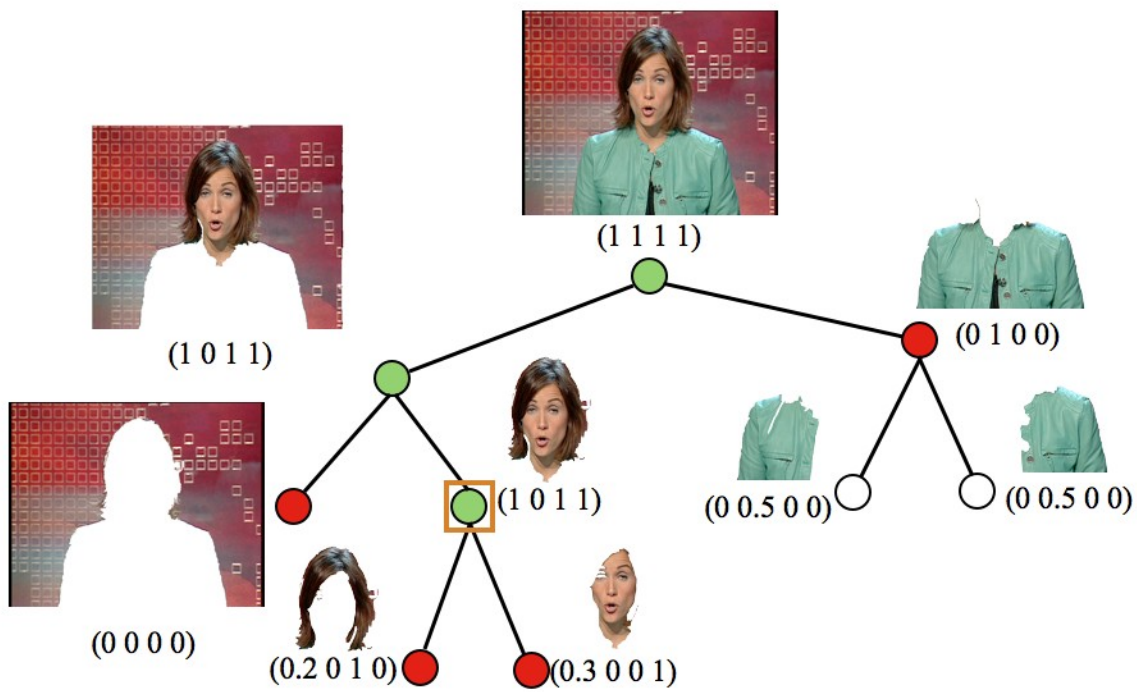
9. Part-based Object Model

Inclusion classifier (training)

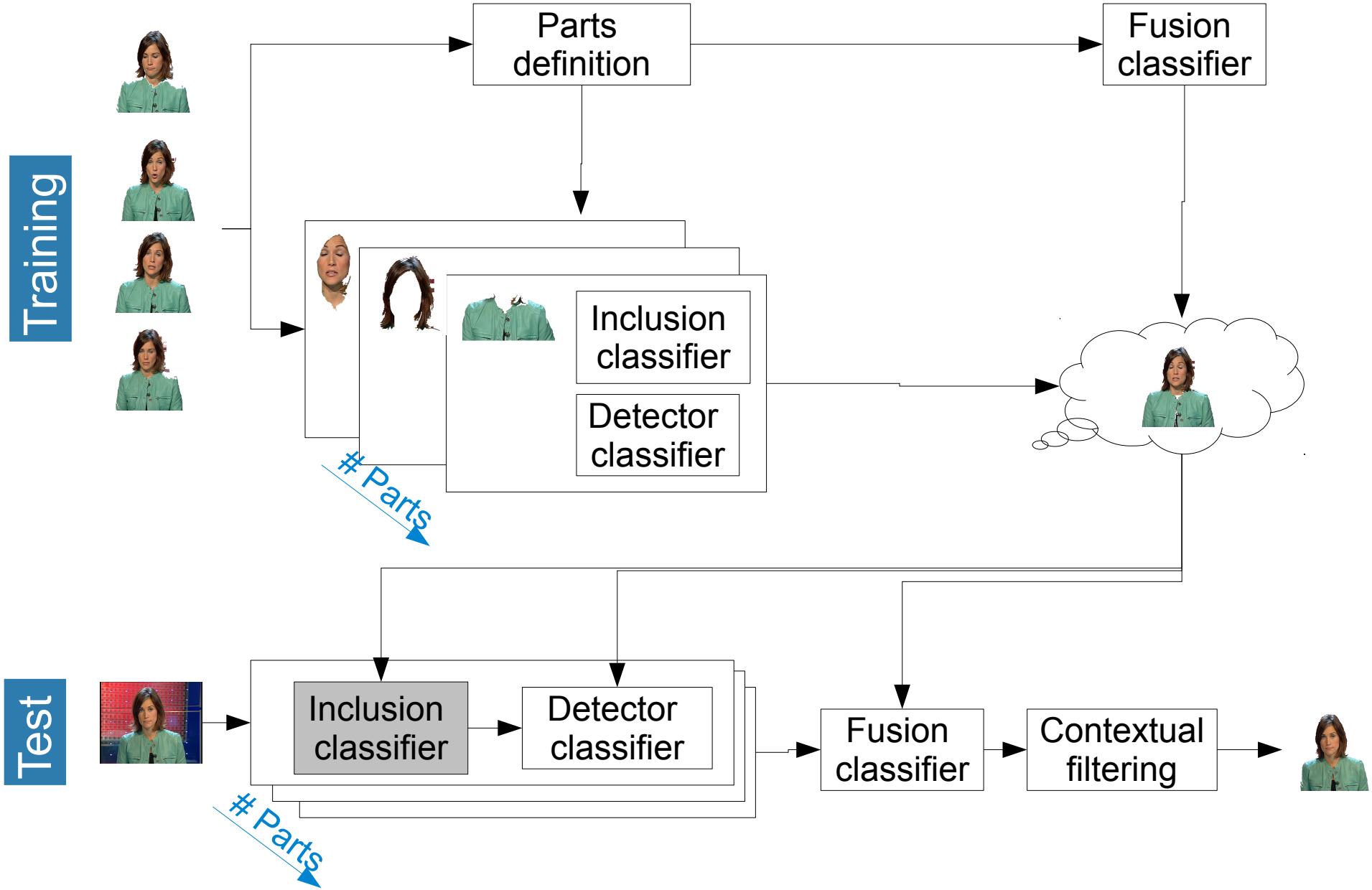
- Positive: Parts and their ancestors
- Negative: Parts' children and positive node's siblings.



HBoR



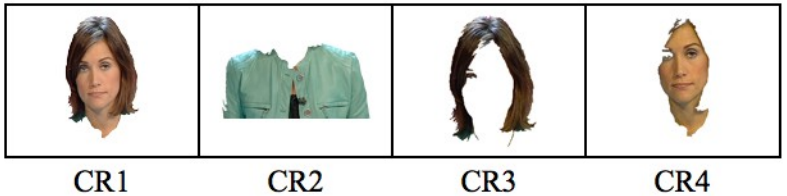
9. Part-based Object Model



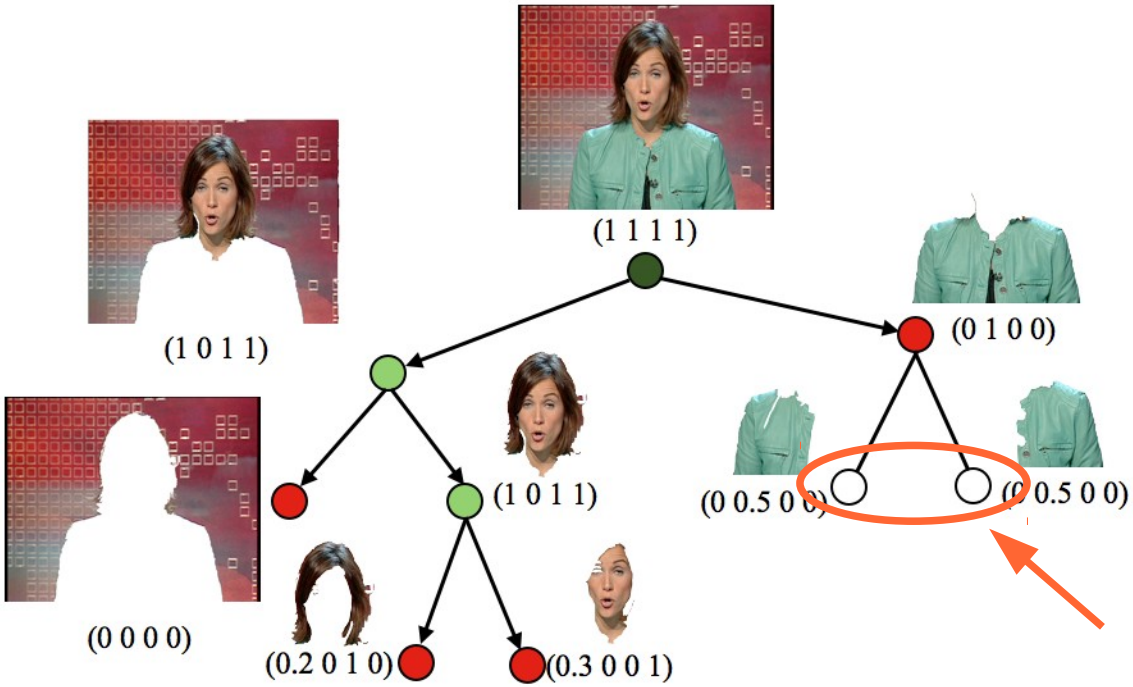
9. Part-based Object Model

Inclusion classifier (test)

- Efficient top-down exploration of the target BPT.

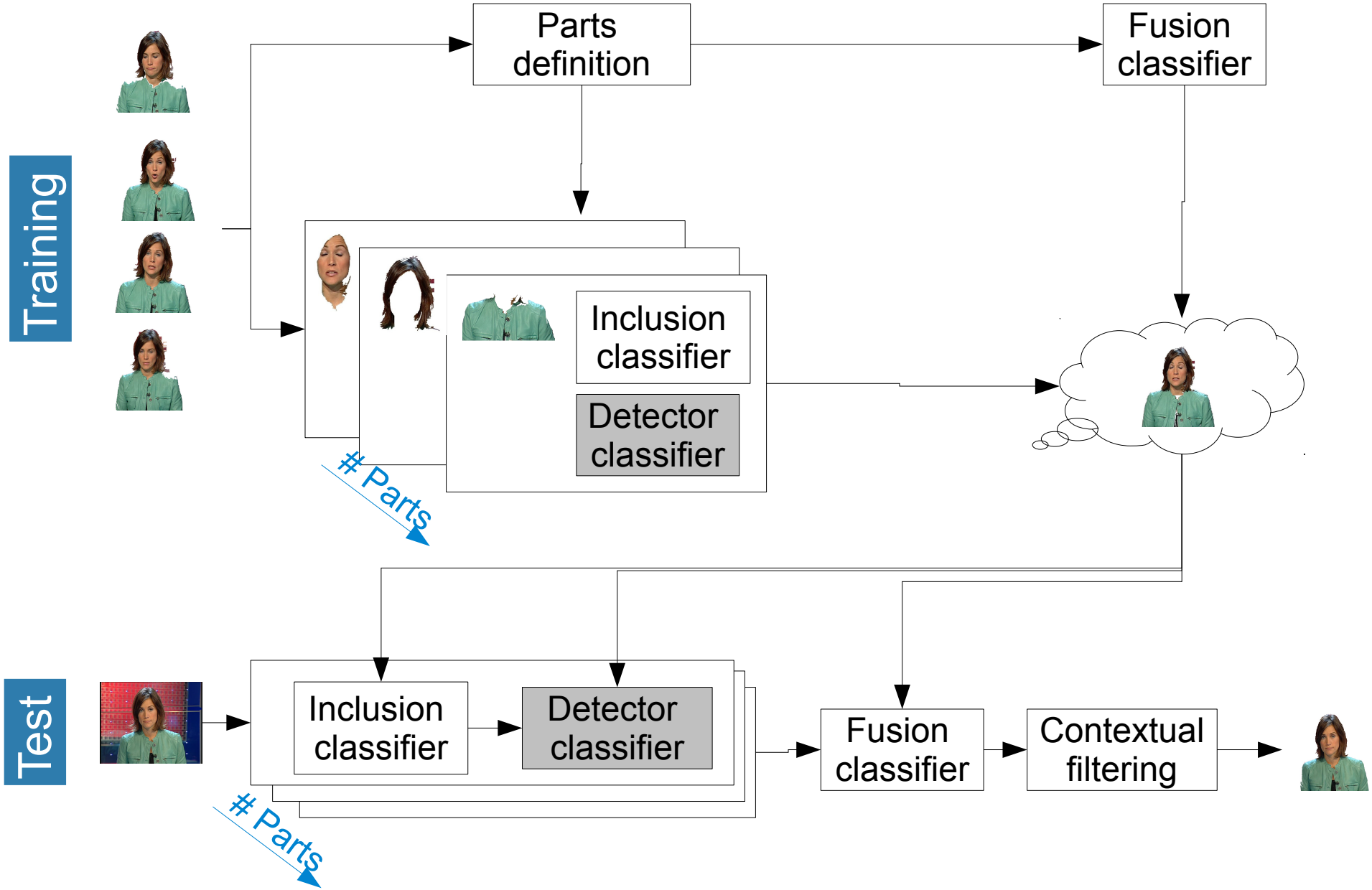


HBoR



Not explored (efficiency)

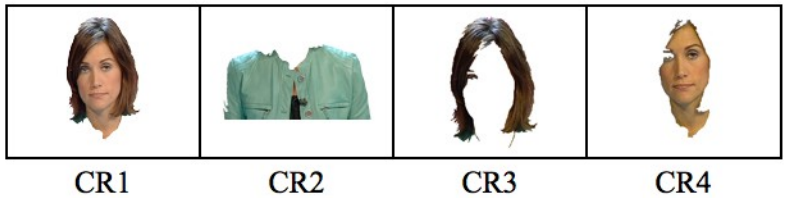
9. Part-based Object Model



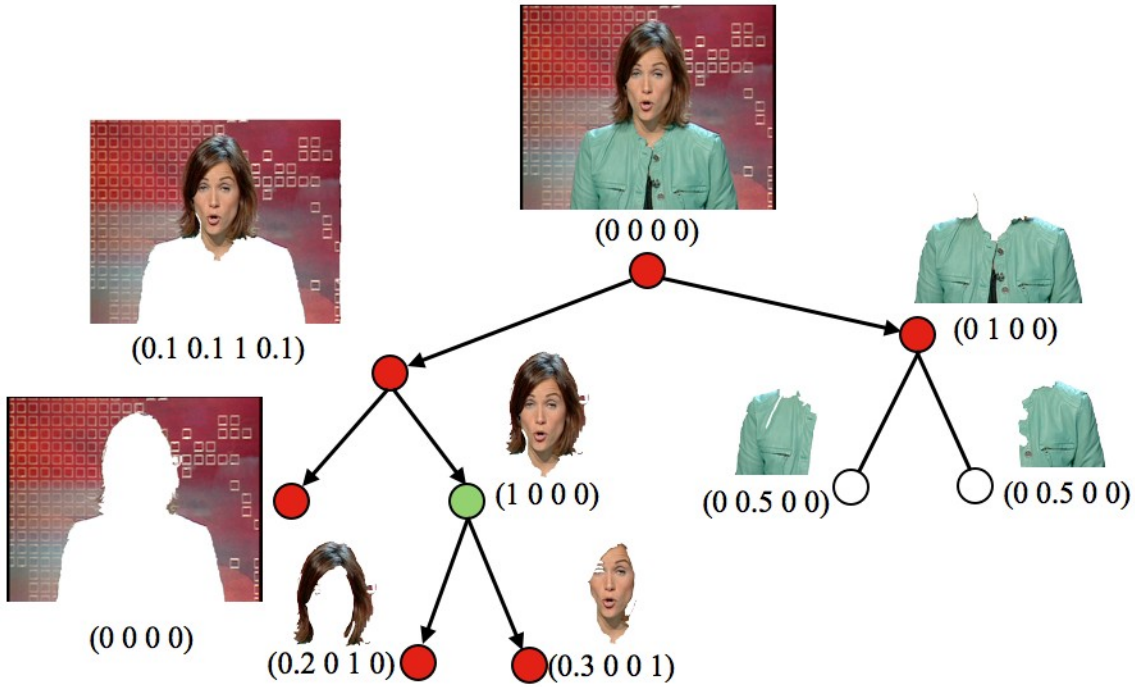
9. Part-based Object Model

Detector classifier (training)

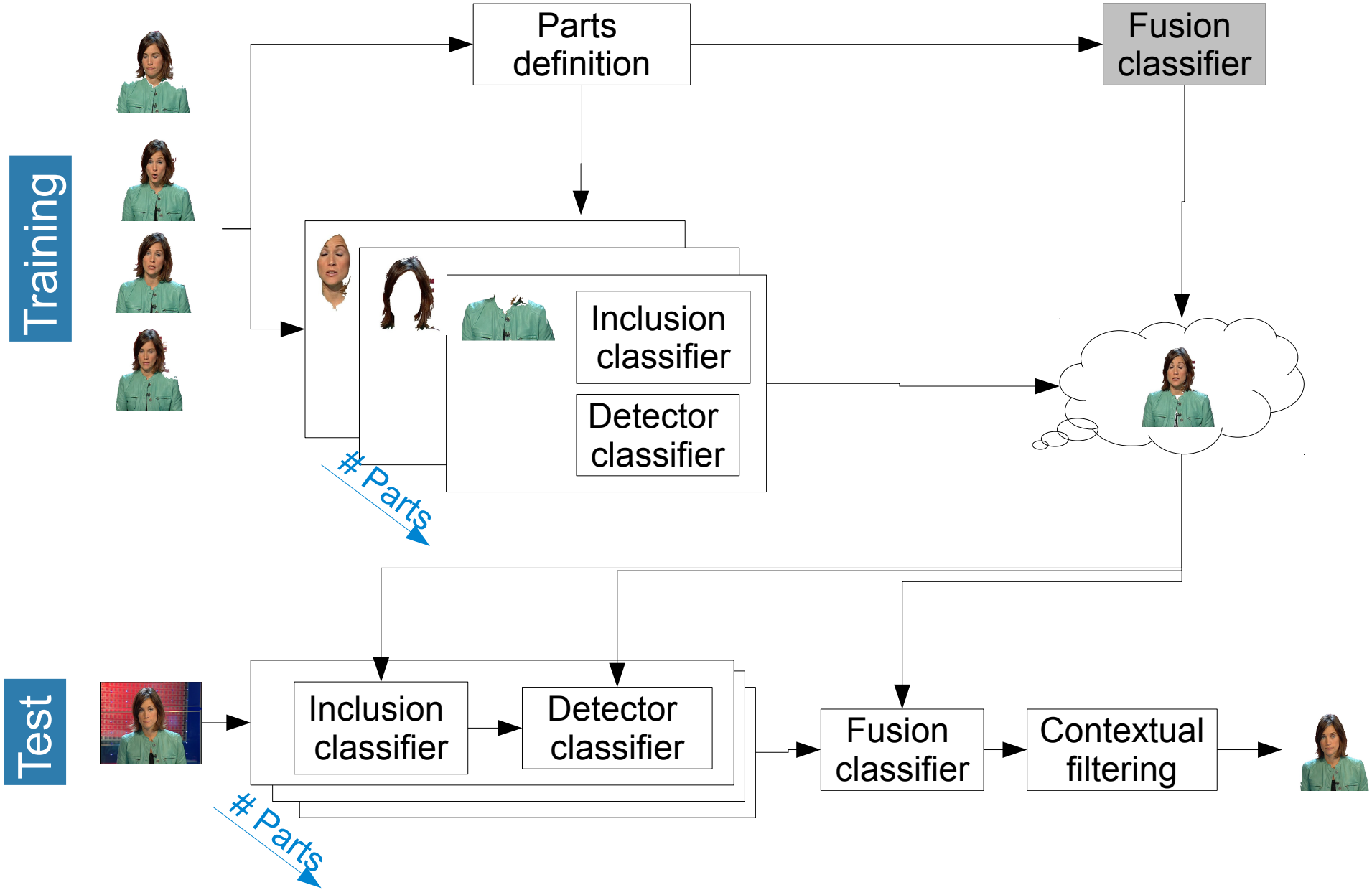
- Positive: Parts.
- Negative: Random sampling (balanced set).



Region-based Possibilistic



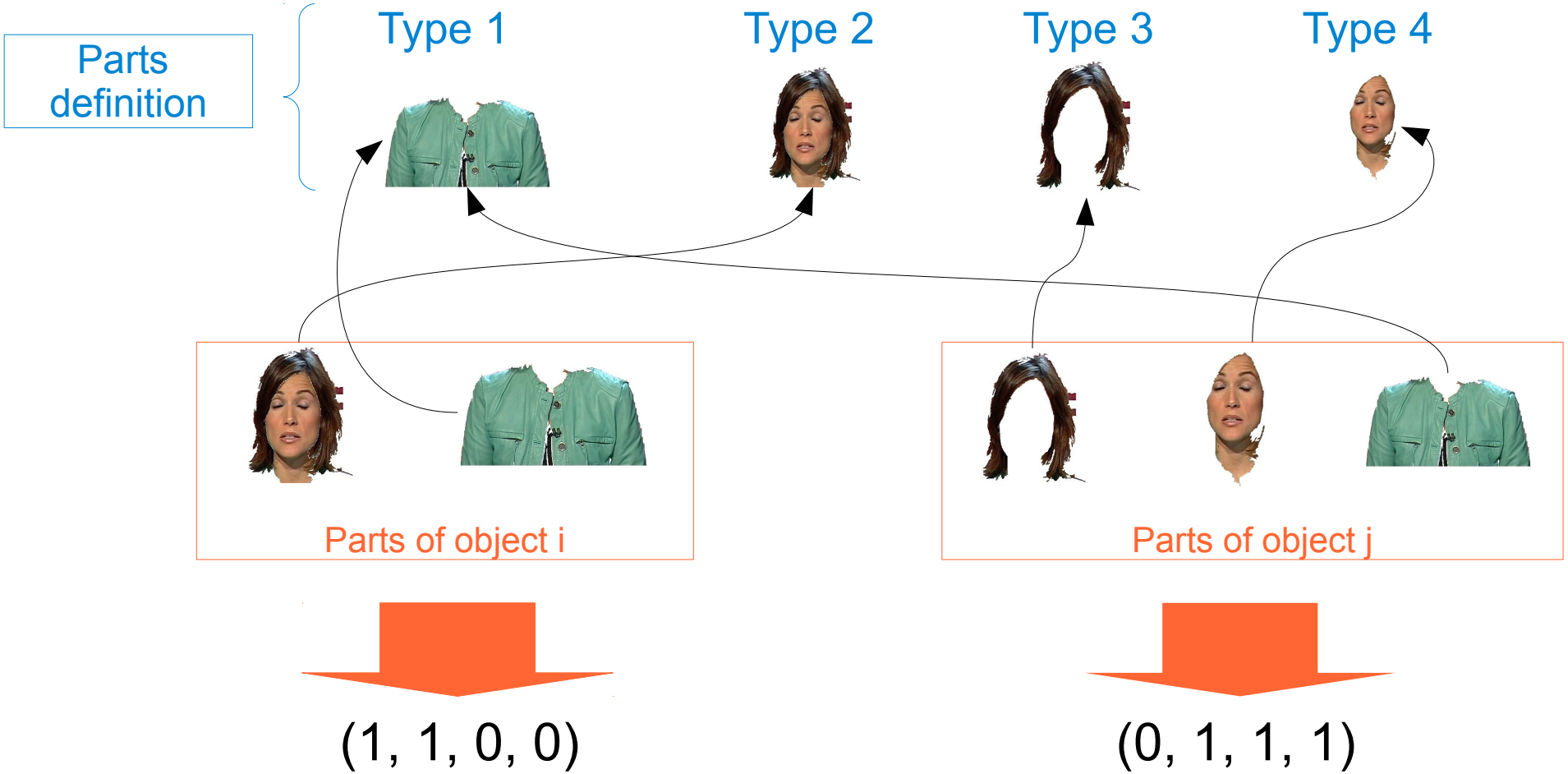
9. Part-based Object Model



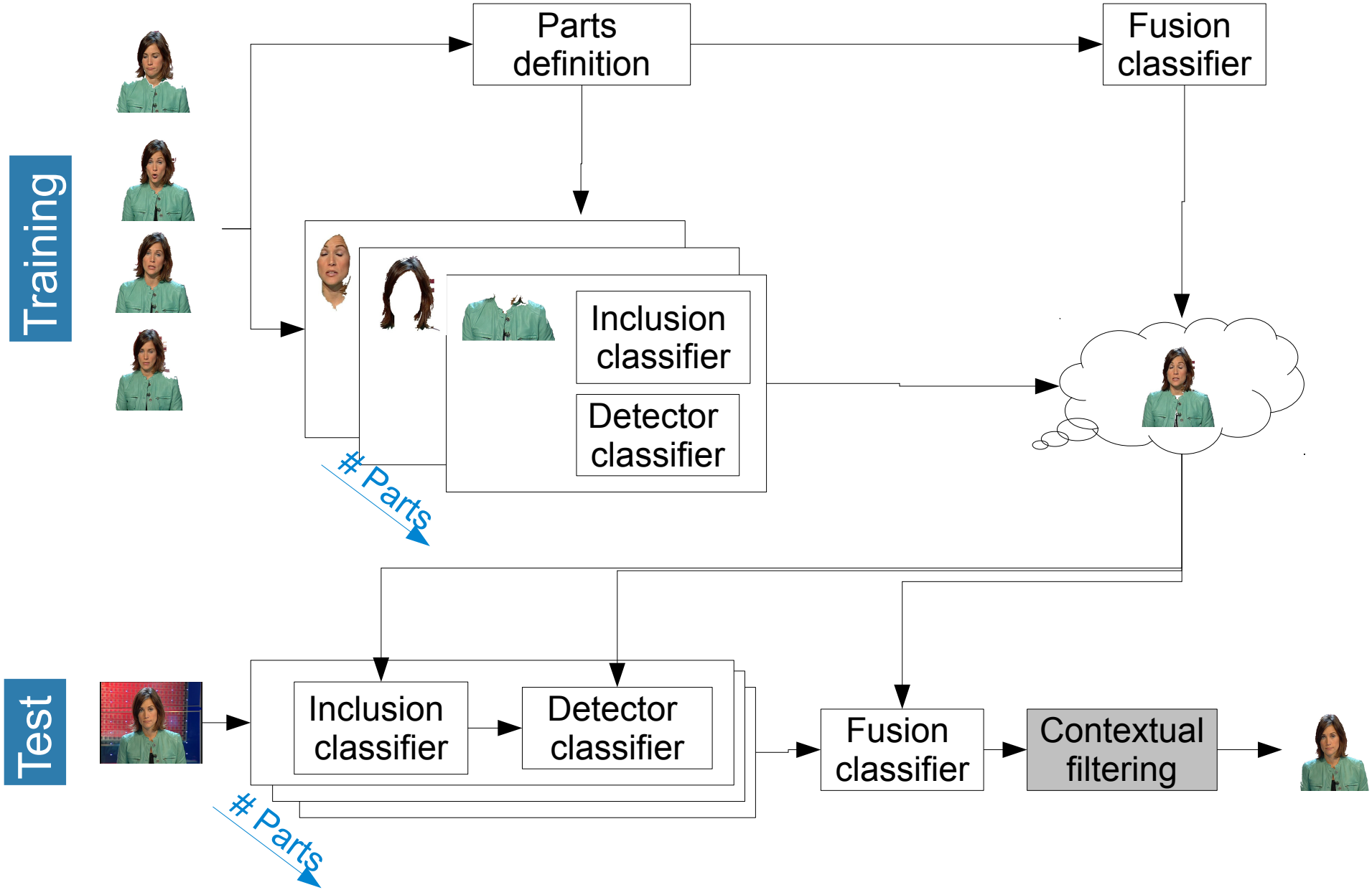
9. Part-based Object Model

Fusion classifier (training)

- Positive training samples represent valid combinations of parts.
- Negative training samples represent invalid combinations.



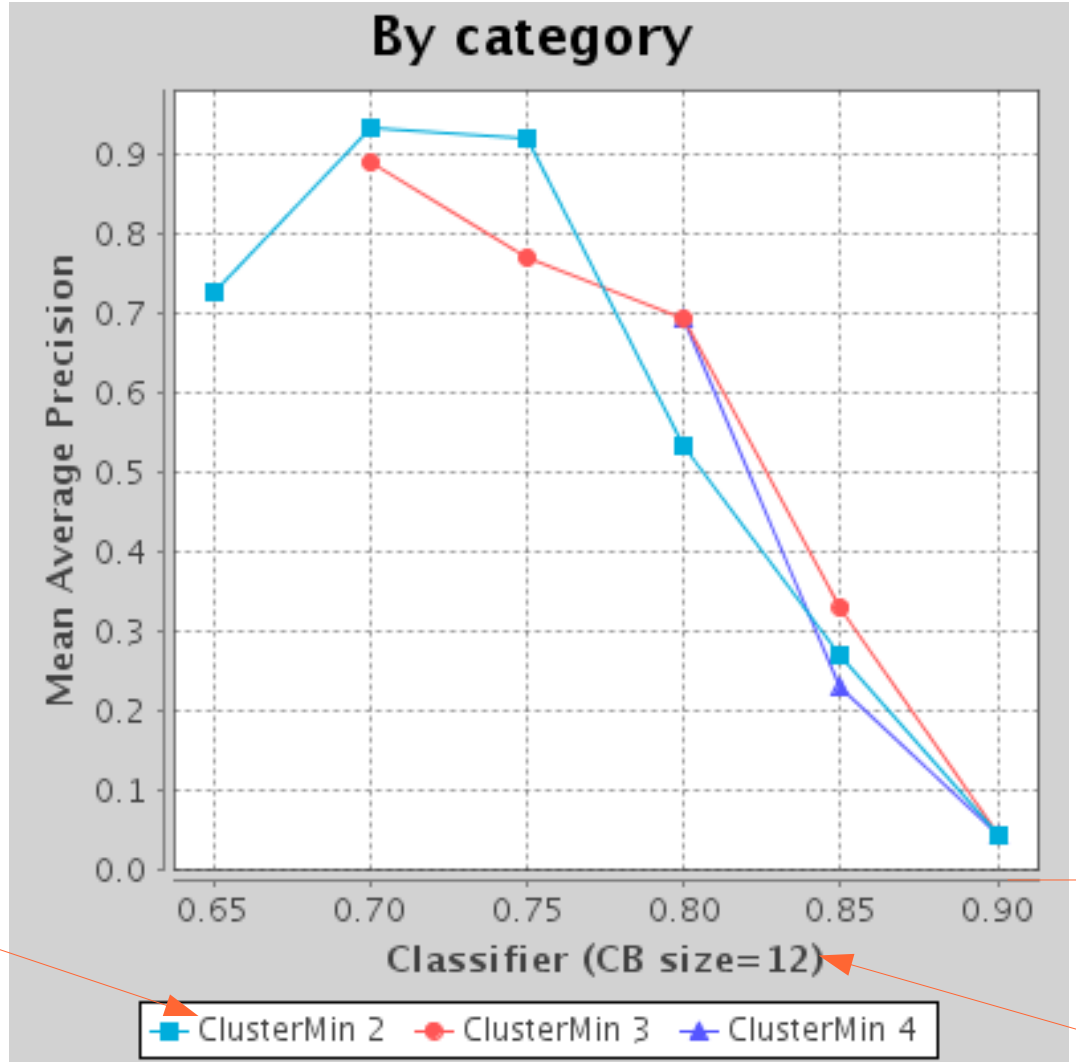
9. Part-based Object Model



9. Part-based Object Model

- How important is a correct estimation of the types of parts ?

Image Retrieval



Max radius for cluster

Min # elems in cluster

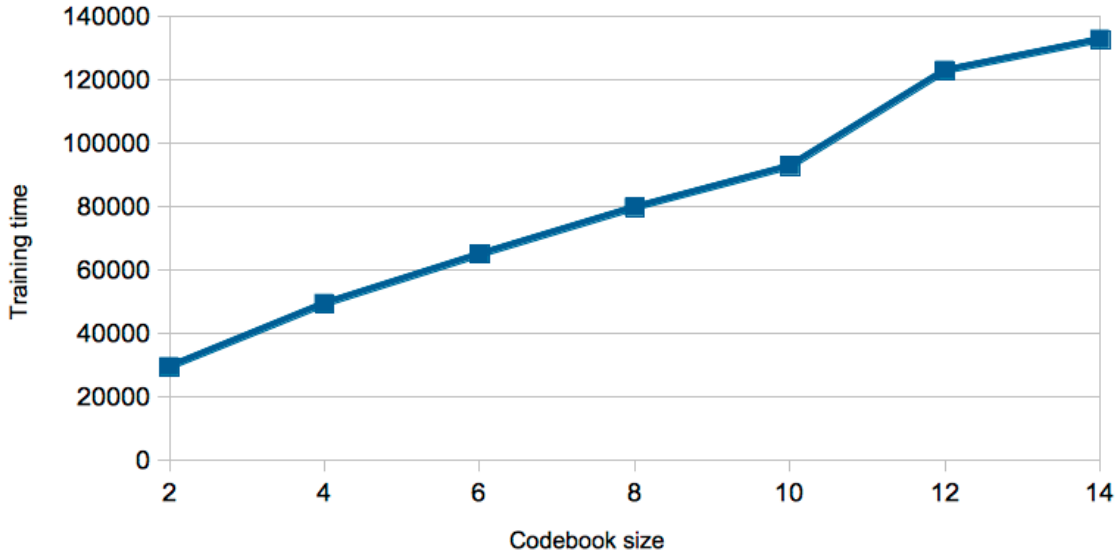
CodeBook size

The performance is dependent from the Quality Threshold parameters.

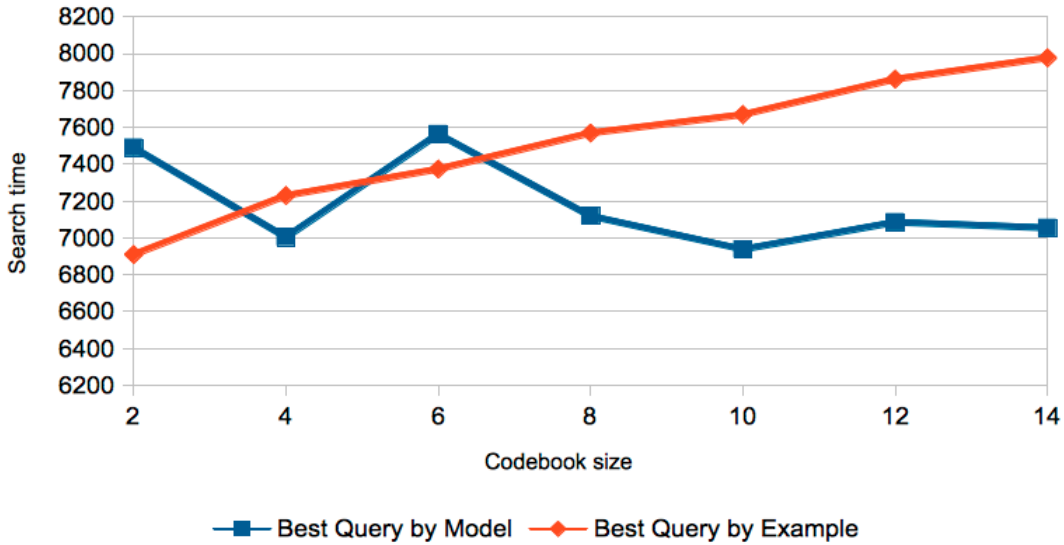
9. Part-based Object Model

- What is the impact of codebook size in terms of time ?

Training time



Search time



Search time does not grow with codebook size (Efficient exploration ??).

Grup de Processament de la Imatge
Universitat Politècnica de Catalunya (UPC)

Outline

Grup de Processament de la Imatge
Universitat Politècnica de Catalunya (UPC)

Conclusions

- BPT presents several opportunities for local analysis:

  → Interactive segmentation
  → Semantic analysis

- Three main contributions in this thesis:

Hierarchy of Bags of Regions:
- Possibilistic quantization
- Max pooling throug BPT

Object Tree - BPT matching
- Top-down
- Bottom-up

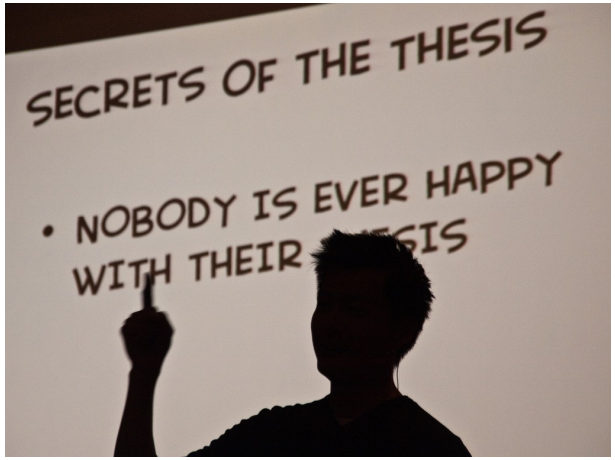
Part-based model for objects
- Automatic determination of parts
- Efficient BPT analysis.

Open work

- Test in public benchmark.
  → Instance Search task @ TRECVID 2012

- Richer image representations
  → Amount of BPT leaves adapted to content
  → Discard uninformative merges & allow multiple options

- Enrich region descriptors
  → Introduce interest points
  → Significance estimation



- Smarter codebooks
  → Unsupervised clustering
  → TF-IDF weights & Stop words

- Better model for combination of parts representing object

- Local segmentation of objects from global annotations

Outline

1. Introduction
2. Hierarchical Image Partitions
3. Interactive segmentation
4. Object Tree
5. Region Features
6. Codebooks
7. Fusion of Visual Modalities
8. Partition Tree Matching
9. Part-based Object Model
10. Conclusions

Part I: 2–4
Part II: 5–7
Part III: 8–9

Thank you, moltes gràcies !