Learning Visual Distance Function for Identification from one Example.

Same Different Same or Different?

Eric Nowak and Frederic Jurie

Bertin Technologies / CNRS

LEAR Group – INRIA - France

This is an object you've **never seen** before …
… can you recognize it in the following images?

This is an object you've **never seen** before …
… can you recognize it in the following images?

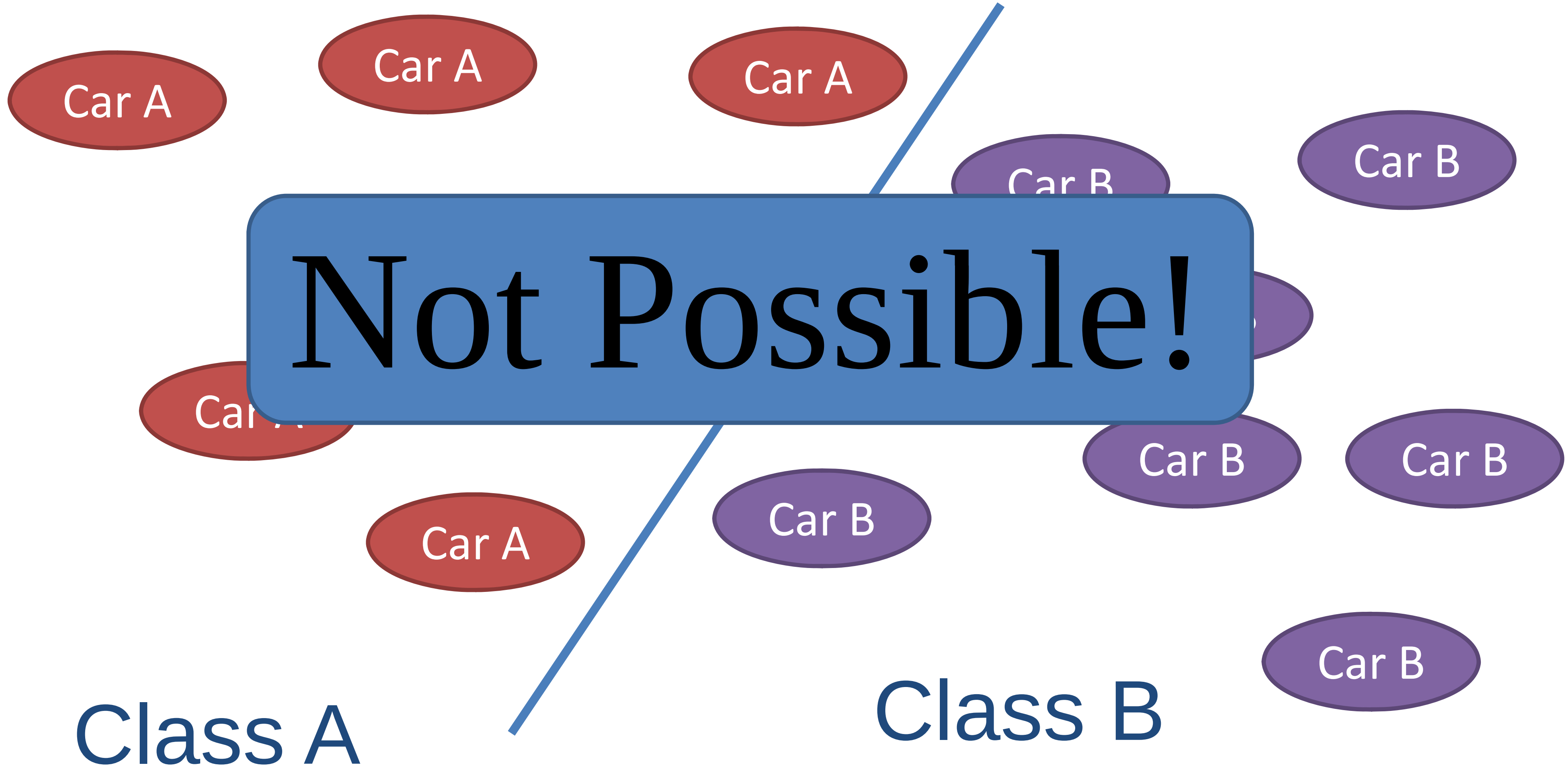
Identification from One Example.

"obviously" different

same pose and shape, but different object

different pose and light, but same object

This is an object you've **never seen** before …
… can you recognize it in the following images?



This is an object you've **never seen** before …
… can you recognize it in the following images?

S( , )

S( , 

)

S( , )

This is an object you've **never seen** before …
… can you recognize it in the following images?

S( , ) < S( , )

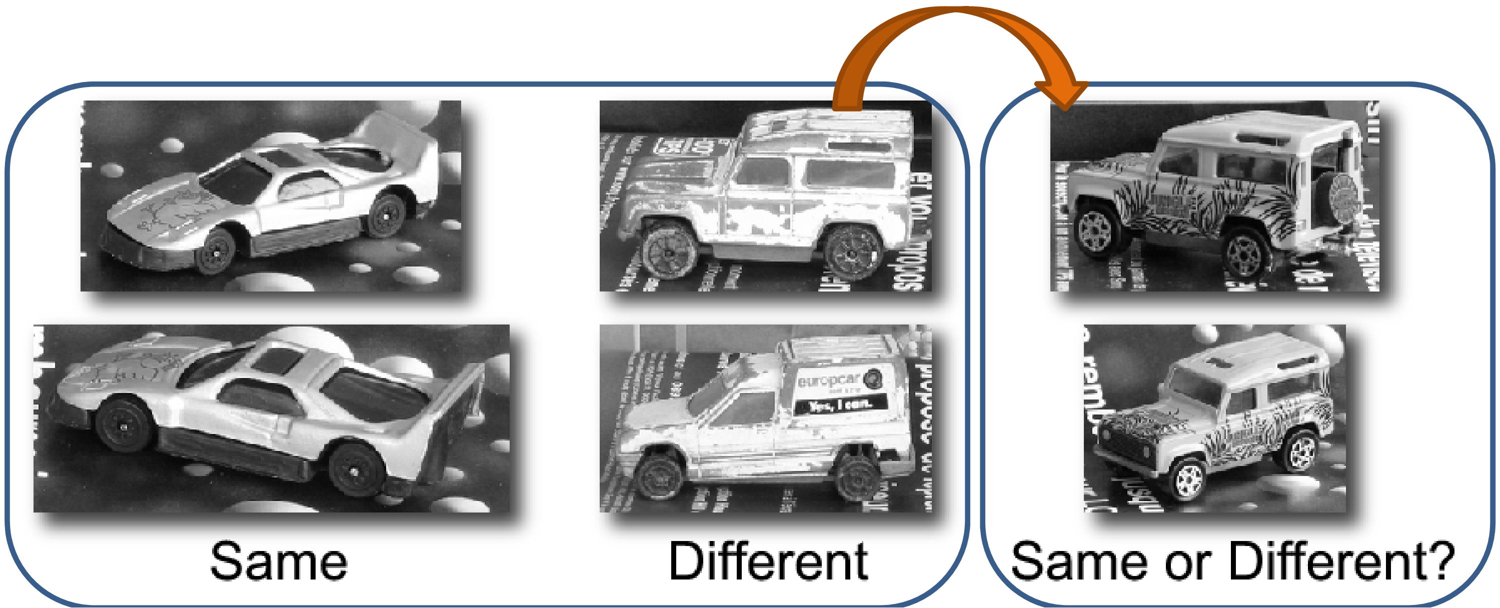
Knowledge about categories


Different

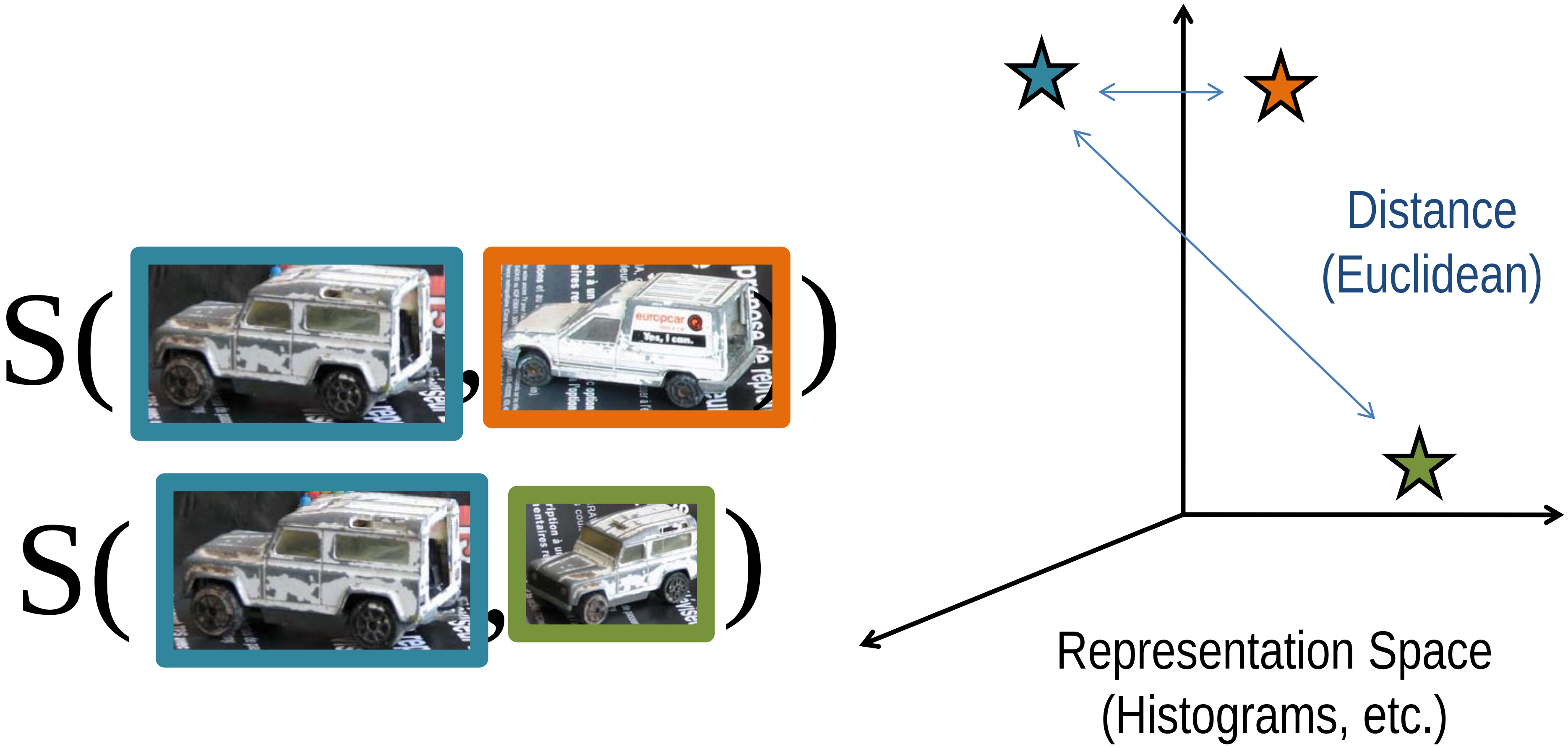

Same

Our goal: Learning from one Example with Equivalence Constraints.

- We want to **learn** a similarity measure on a generic category (e.g. cars)
- Given a training set of image pairs labelled «same» or «different»: **equivalence constraints**
- we can predict how similar two **never seen** images are
- **despite** occlusions, clutter and modifications in pose, light, ...

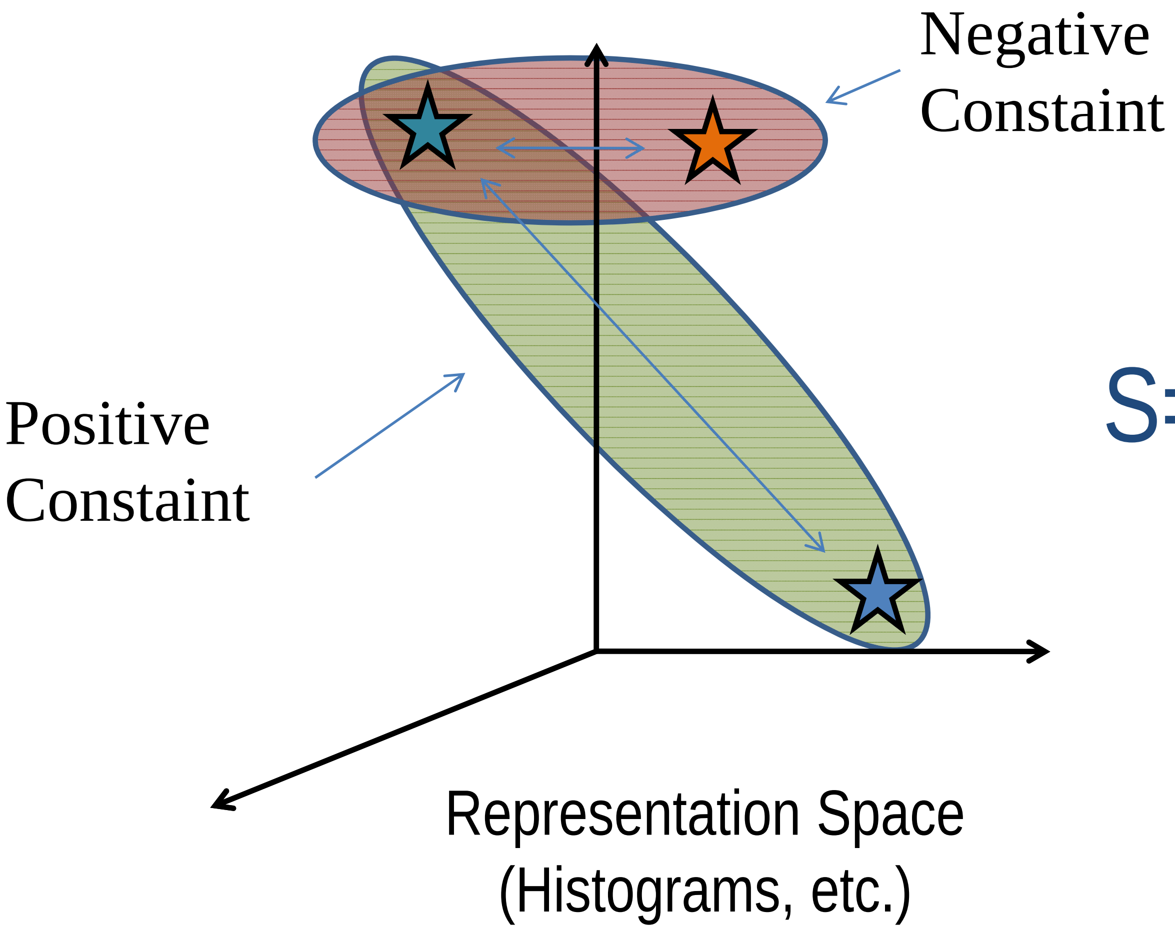


How to compare images ?



Not adapted to visual classes

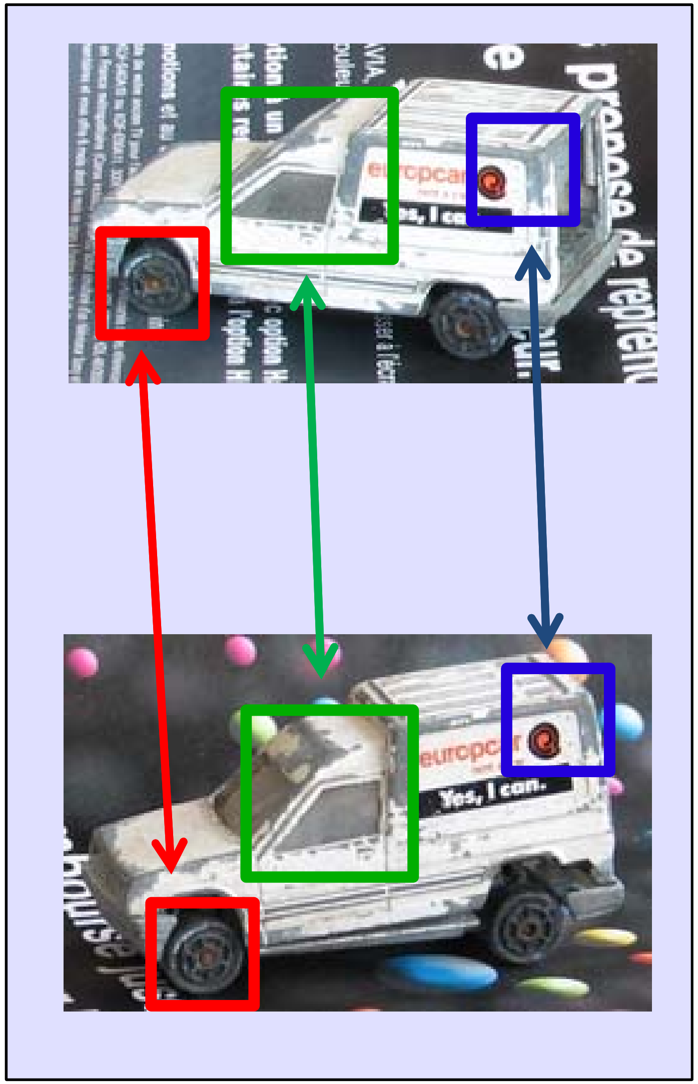
How to learn the distance ?



$S = X^t A X$

Not robust to occlusions, background

How to be robust to occlusion, view point changes ?

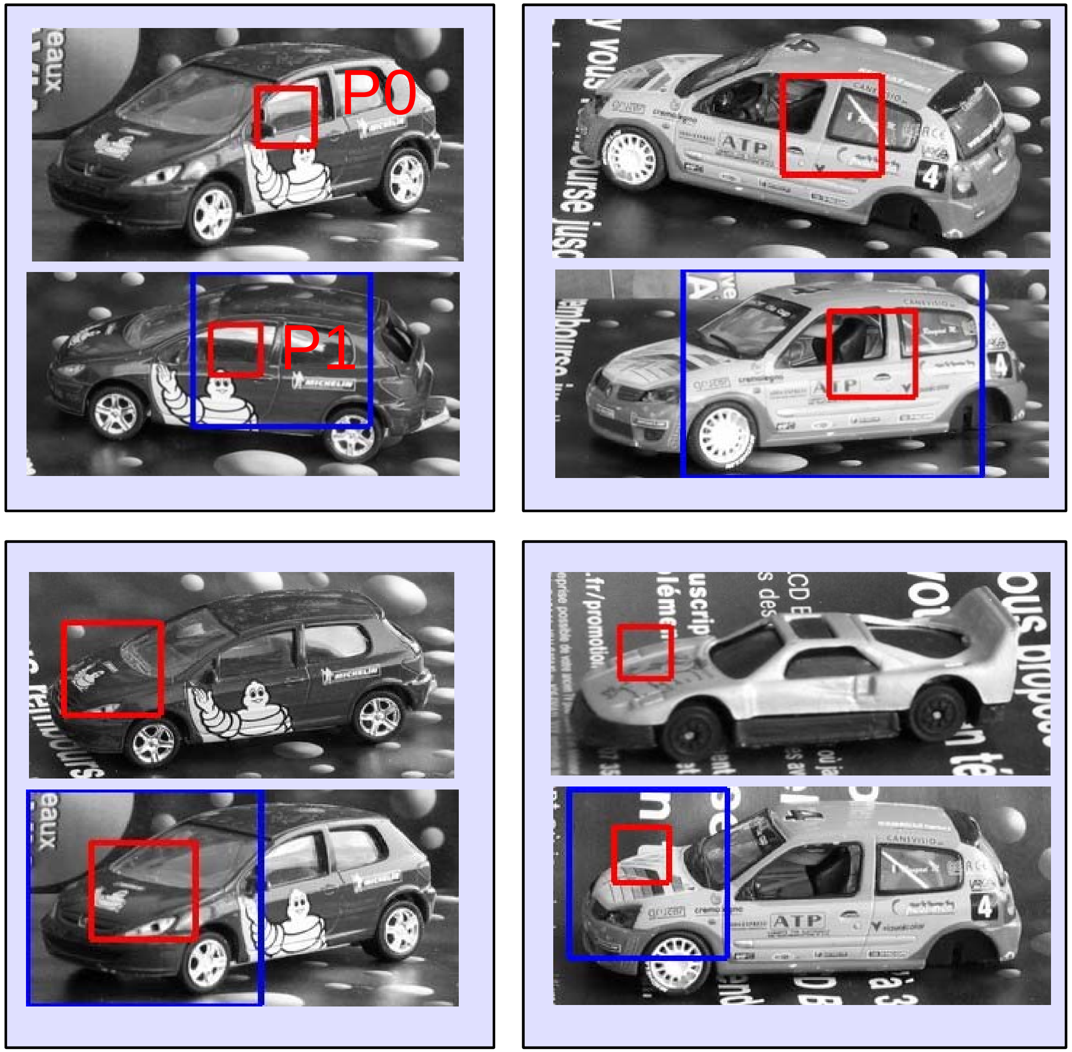


Robust combination” of local distances:

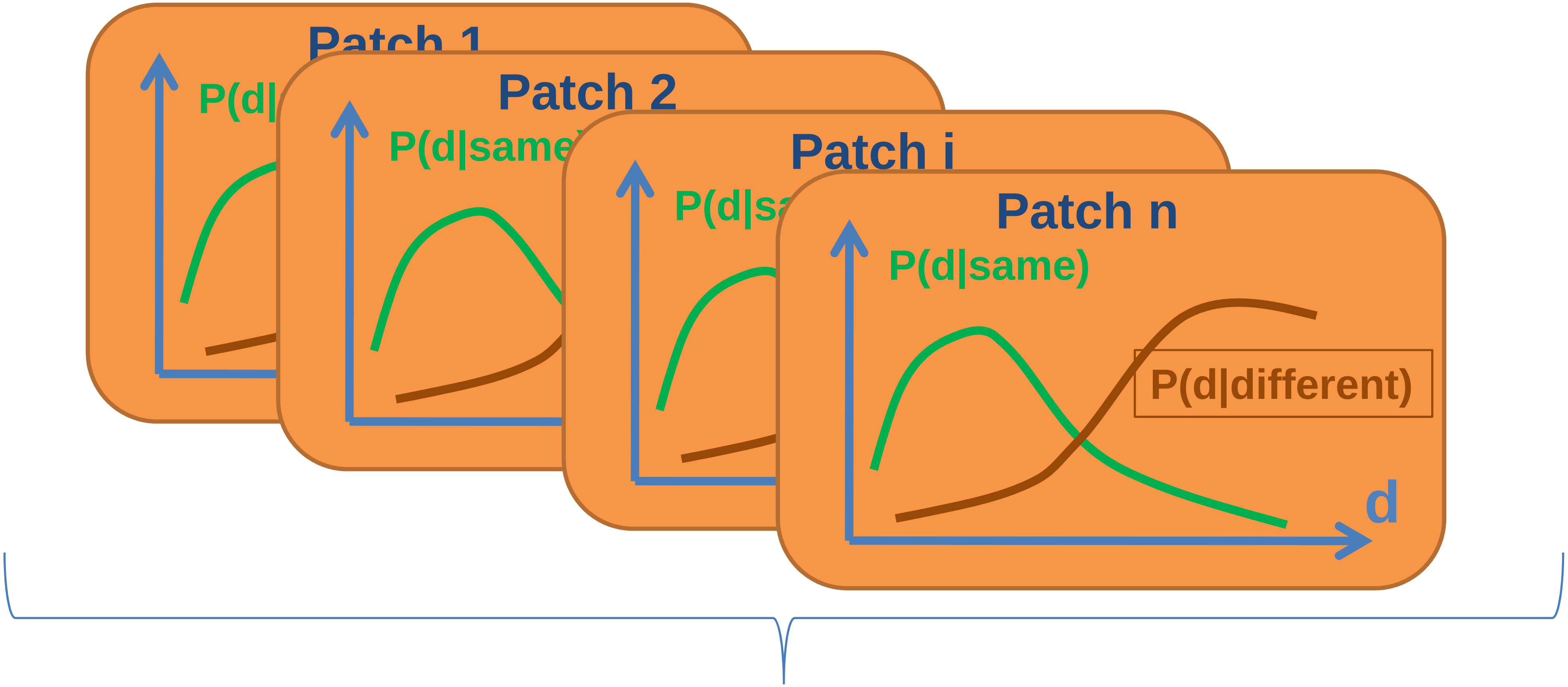
$S=f(d_1,d_2,\ldots,d_n)$

Computation of corresponding patches

- P0 in I0: sampled **randomly** (quadratic in size, uniform in position)
- P1 in I1: the **best ZNCC match** of P0 around P0. Search region: extension of P0 in all directions.
- A pair of images is **simplified** into the **np** patch pairs sampled from it.

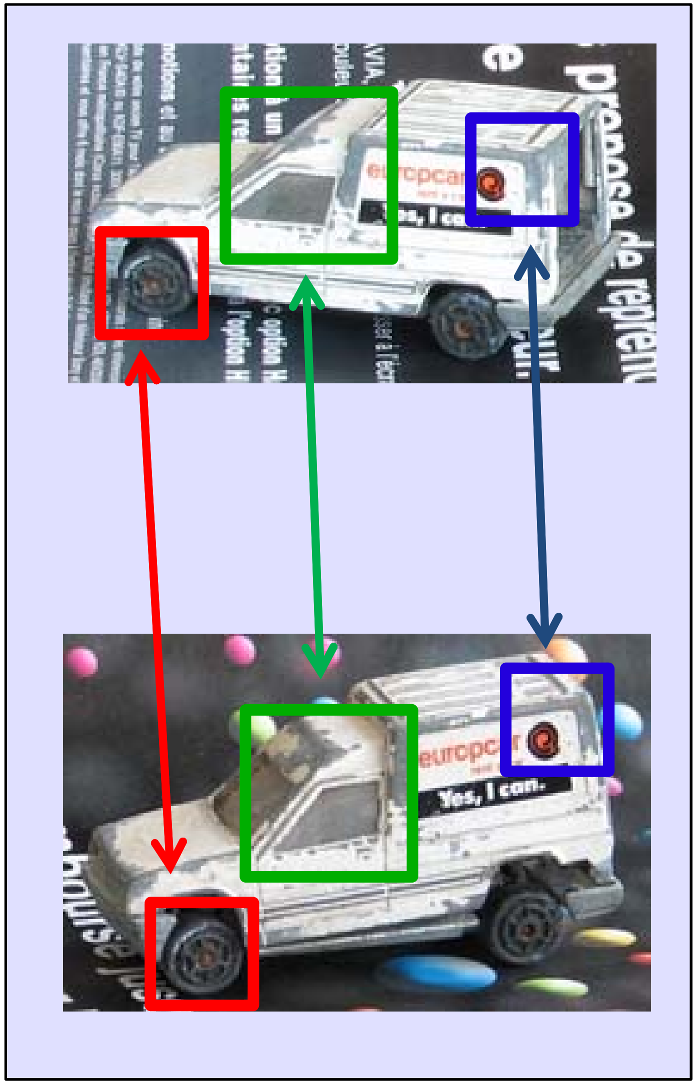


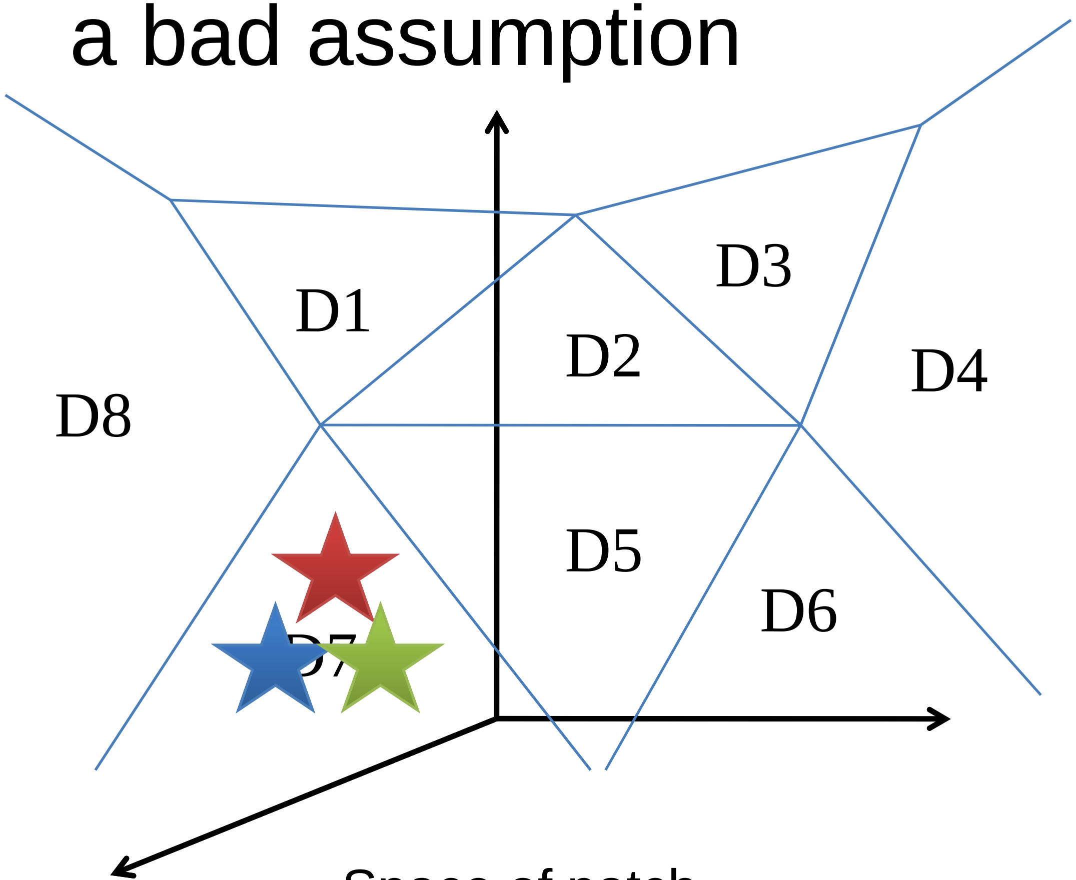
From multiple local similarities to one global similarity



Likelihood->Similarity
[Ferencz et al. Iccv 05]

Patch independence: a bad assumption

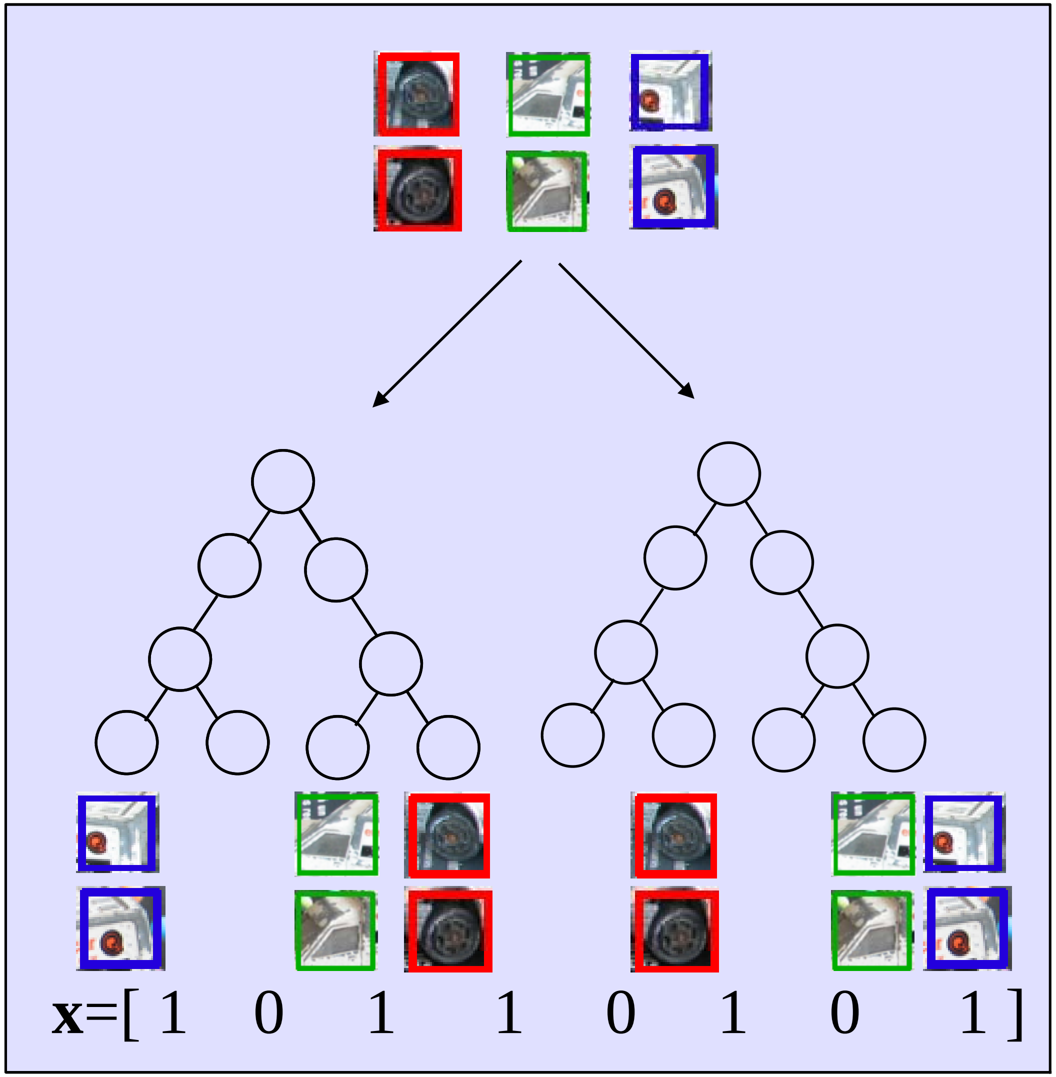




Space of patch pairs differences

=>Vector quantization

Vector quantization of pair difference

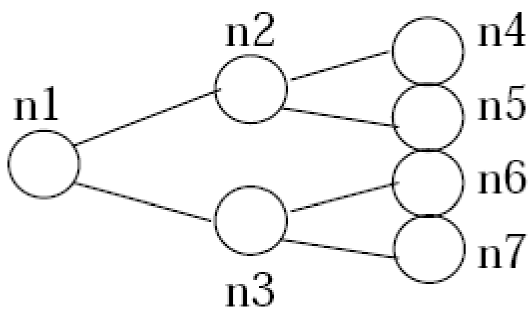


Computation of the trees

Tree creation (EXTRA-Trees [Geurts et al. ML06, Moosman et al. NIPS06]):

- create a root node with positive and negative patch pairs.
- **recursively split** the nodes until they contain only pos or neg pairs:
  - create *ncondtrial* **random split conditions**:
    simple parametric tests on pixel intensity, gradient, geometry, etc.
    random <=> parameters drawned randomly
  - select the one with the **highest information gain**
  - split the node into two sub-nodes

Node	#pos pairs	#neg pairs	Entropy	Splitcondition
n1:	10000+	10000-	H=0.693147	Abs(Delta(P(2,3)))<=0.72144
n2:	174+	317-	H=0.650115	Abs(Delta(P(9,8)))<=0.48665
n4:	19+	60-	H=0.551663	Leaf
n5:	155+	257-	H=0.66218	Leaf
n3:	9826+	9683-	H=0.69312	Abs(Delta(P(2,2)))<=0.31706
n6:	1845+	3040-	H=0.66292	Leaf
n7:	7981+	6643-	H=0.688956	Leaf

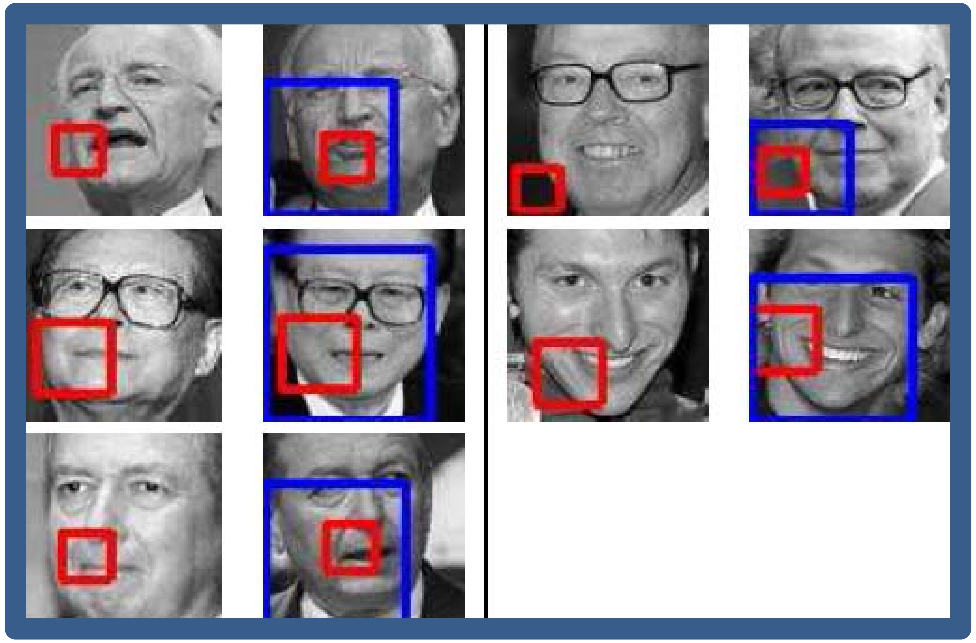


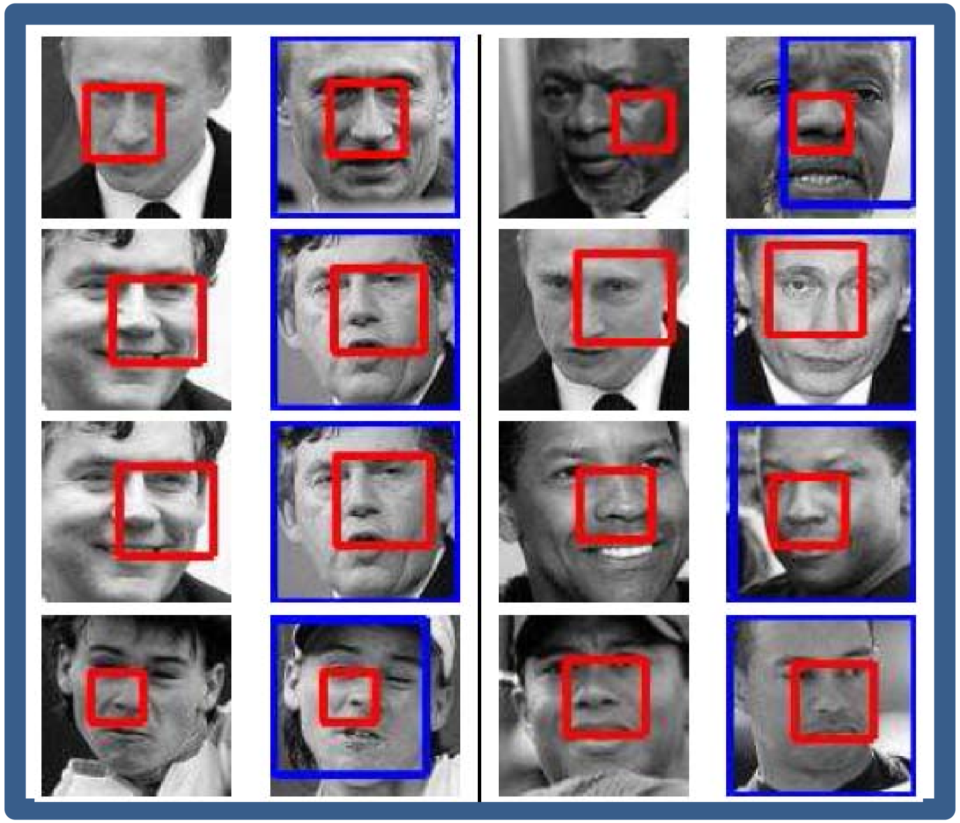
Very Fast!

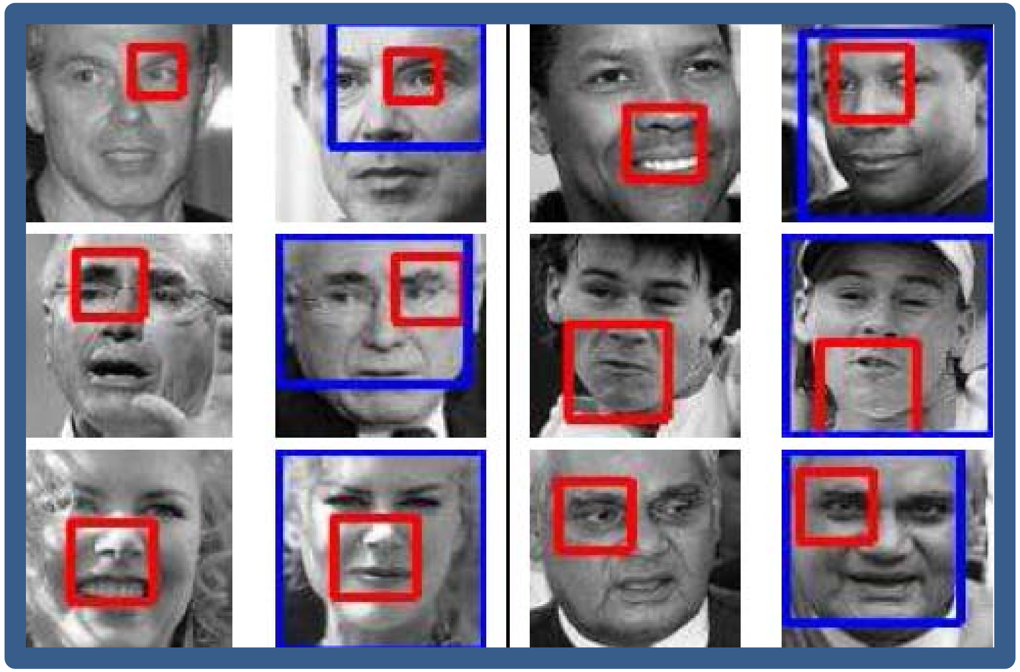
Computation of the trees

The positive patches of three different nodes during tree construction

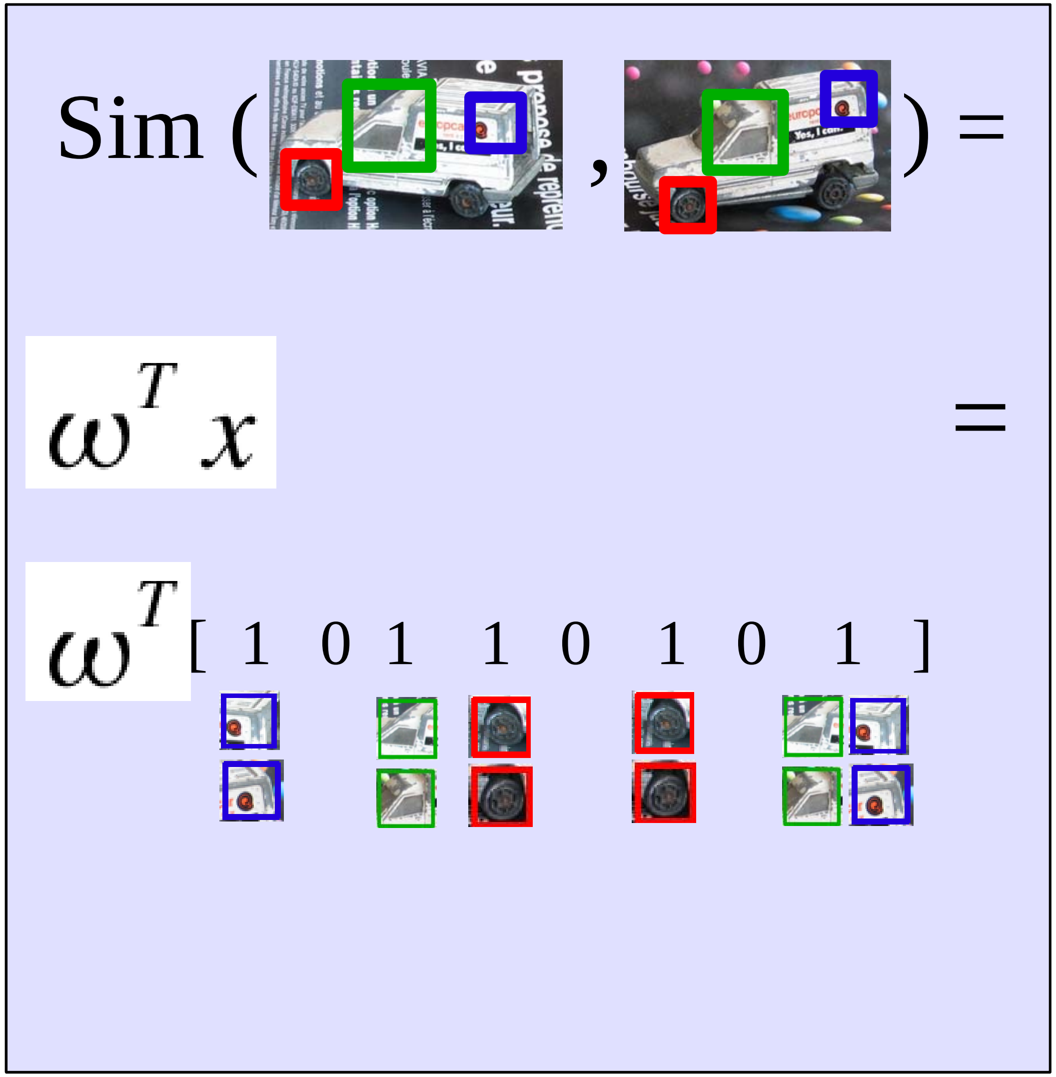
("faces in the news" dataset)







From clusters to Similarity



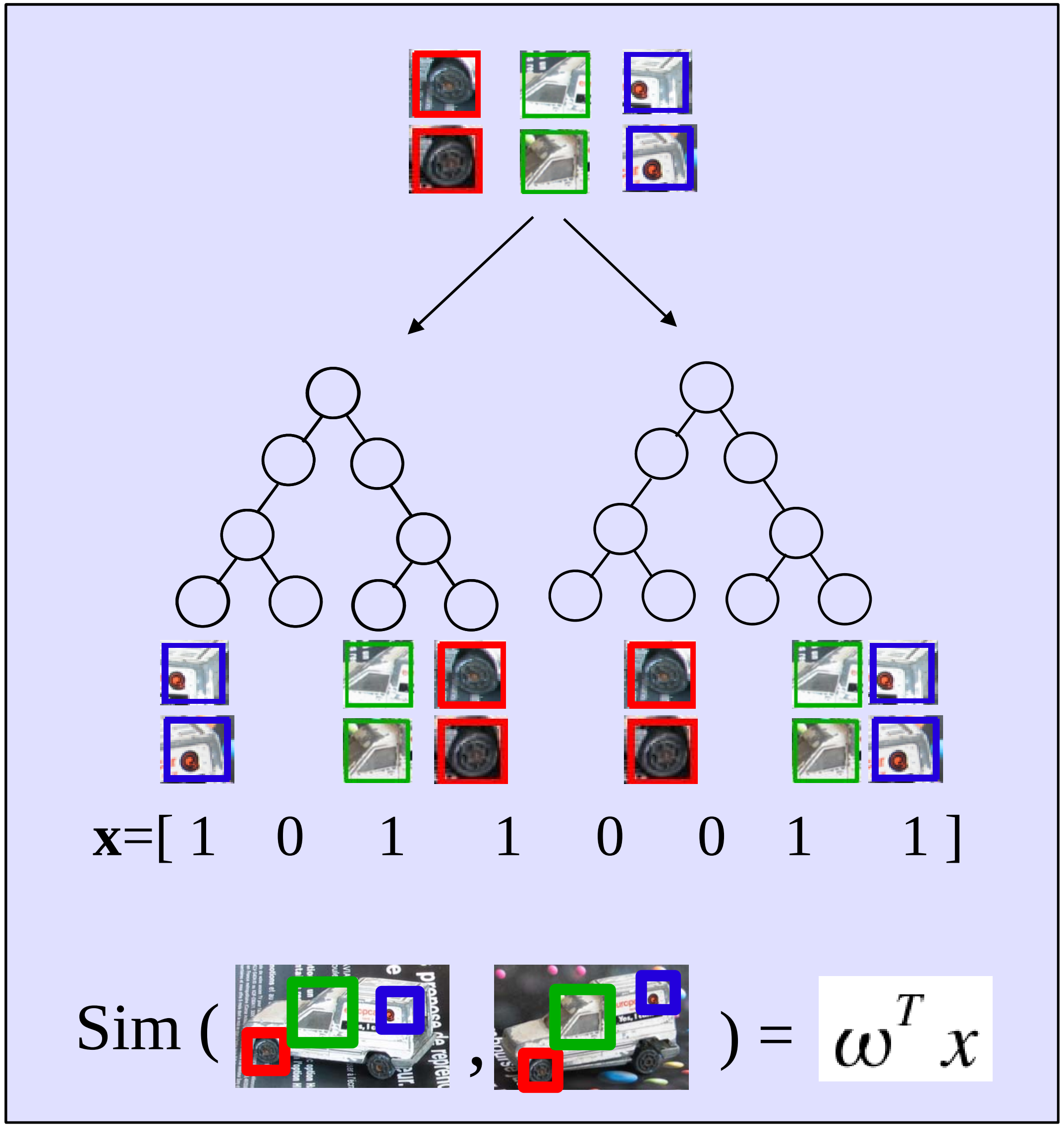
The similarity measure is a linear combination of the cluster membership

- $S(I_1, I_2) = \omega^T x$ and we want: **the larger the more similar**

- We define the weight vector as the **normal of the linear SVM hyperplane** separating the descriptors of positive and negative learn set image pairs.

Similarity measure

- **Given 2 images ...**
- Detect corresponding patch pairs.
- Affect them to clusters with extremely randomized trees.
- **The similarity measure is a linear combination of the cluster membership.**



Conclusions

- Similarity of **never seen objects**, given a set of similar and different training object pairs of the same category.
- Original method consisting in
  - (a) finding similar patches
  - (b) clustering the set of patch pair differences with an ensemble of extremely randomized trees
  - (c) combining the cluster memberships of the pairs of local regions to make a global decision about the two images.
- Can learn complex visual concepts.
- Image polysemy->of pairs of “same“ and “different” defines visual concepts
- Can automatically selects and combines most appropriate feature types
- Future works: recognize similar *object categories from a training set of equivalence* constraints.