BOOSTED LANDMARKS OF CONTEXTUAL DESCRIPTORS AND FOREST-ECOC

Sergio Escalera

Oriol Pujol

Petia Radeva

The 2006 International Conference of Pattern Recognition

INDEX

1. Object Recognition
2. Boosted Landmarks
3. Forest-ECOC
4. Results
5. Conclusions

Bottom-Up Approaches

Region based (full image processing) → **Keypoint detector**
- Edges, Valleys, Contours …
- Saliency (Kadir, Fergus, Fei-Fei, Perona, Zisermann)
- Harris (Lowe, Buhmann)
- Covariance Affine Regions (Torralba)

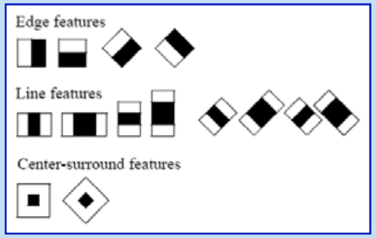
Patch based (patch) → **Description**
- SIFT (Lowe, Torralba)
- Feature histograms (Buhmann)
- Gabor, Haar-like (Torralba)
- Disassociated dipoles (Baró)
- Boosting Context (Amores), Boosted Landmarks (Escalera)

Dimensionality Reduction — **Object landmark selection**
- Mutual Information (Ullmann)
- Boosting (Torralba)
- ClusteringNDA, LDA, MDS, ND (Torralba, Buhmann)
- Fergus, Fei-Fei PCA region to probabilistic model

Parts based → **Object modelling**
- Descriptors array (Torralba, Lowe, Ullmann)
- Probabilistic models (Buhmann, Fergus, Fei-Fei, Torralba)
- Graphical Models (Bayessian networks)
- Gaussian Miture Models

Top-down Approaches (Presence of an object)
- Gist -> Steerable pyramid (Torralba)

Definition

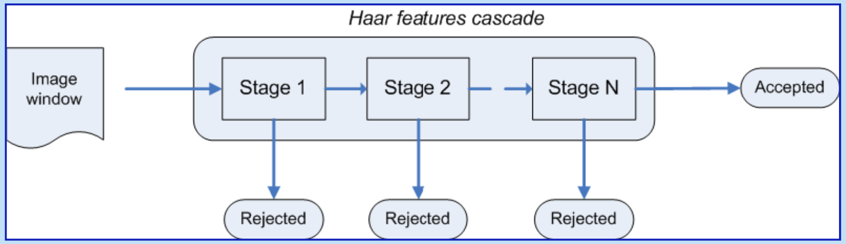


Parts

Haarlike Features

Gentle Adaboost

Cascade of weak classifiers



Contextual descriptor

$$P = \{p_i\}_{i=1}^{N}$$
$$L^1 = \{L^1_1 ... L^1_{i_1}\}, ..., L^n = \{L^n_1 ... L^n_{i_n}\}$$
$$\{L^1_{j_1}, ..., L^n_{j_n}, j_1 \in \{1, ..., i_1\}, ..., j_n \in \{1, ..., i_n\}\}$$
$$D = [D_1, ..., D_n]$$
$$D_i = \{B_i^1, ..., B_i^n\}$$
$$B_i^j = \{(o_j, h_j, x_j)\}_{i=1}^{n}$$

Points of interest

Landmarks candidates

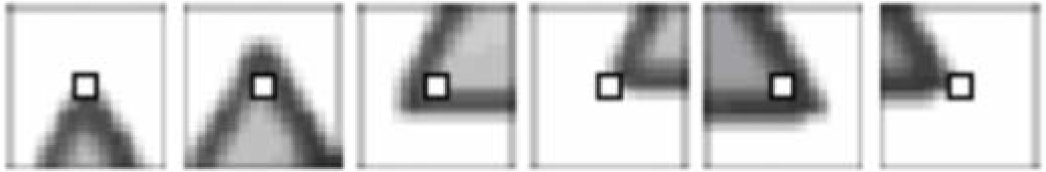
Landmark sets

Object descriptor

Landmark descriptor

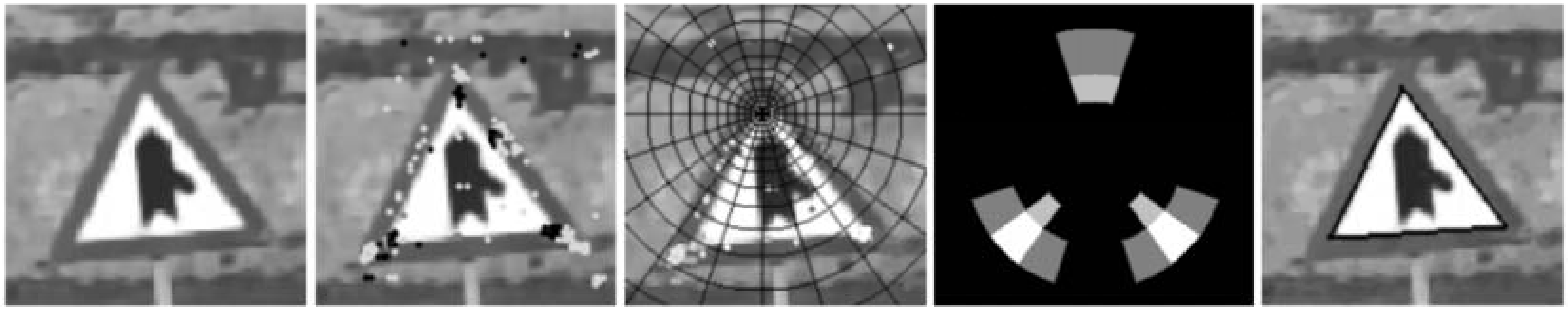
Landmark features description:

label, properties, spatial arrangement



(a) (b) (c) (d) (e) (f)

Selected landmarks for triangular signs.

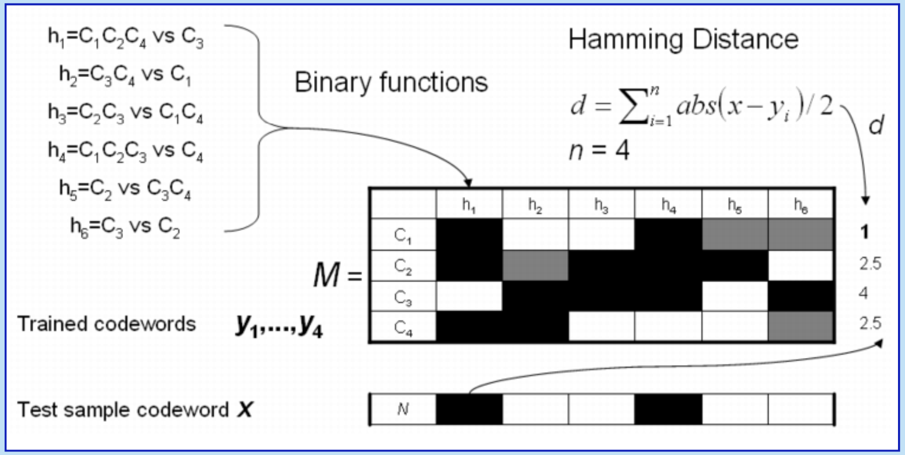


(a) (b) (c) (d) (e)

Input image. (b) Detected landmarks. (c) Contextual descriptors. (d) Resulting bins at feature selection of the correlogram of the landmark (e) Detected sign.

Error Correcting Output Codes

- Design a codeword for each class of N_c classes (up to N_c codewords).
- Arranging the codewords as rows of a matrix we define the "coding matrix" *M*, composed by –1, 0 or 1, that represented as *n* binary learning problems (dichotomies),
- Each dichotomy corresponds to a column of the ECOC matrix.
- Each dichotomy defines a partition of classes
- As a result of the outputs of the *n* binary classifiers, a code is obtained for each data point in the test set.
- This code is compared with the base codewords of each class defined in the matrix *M*, and the data point is assigned to the class with the "closest" codeword.



Coding

One-vs-one

One-vs-all

Dense Random

Sparse Random

Decoding

Hamming distance

Euclidean distance

UAB CVC RMG

Forest-ECOC

Given n classes: $C_1, \ldots, C_n$ and T the number of optimal tree structures to be embedded:

Step 1. Initialize the root node with the set $N_0 = \{C_1, \ldots C_n\}$

Step 2. Generate the tree structures:

- For each node N_j consider the T partitions $\wp_{kj} = \{\{\wp_j^1, \wp_j^2\} | N_j = \wp_j^1 \bigcup \wp_j^2\}, k = 1 \ldots T$ that attain the minimal empirical error for the subproblem defined by the partition $\wp_{kj}$.

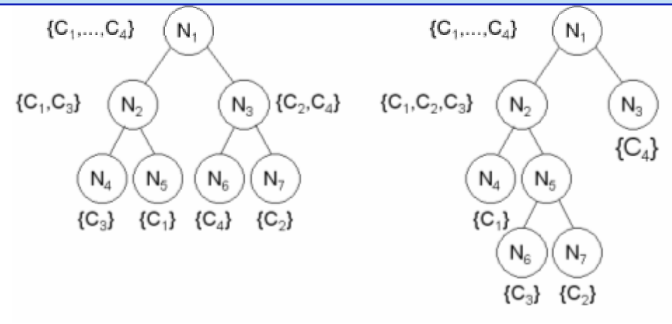
$$\wp_k = \underset{\tilde{\wp}}{argmin} \quad (e(\mathcal{H}(\tilde{\wp}, \mathrm{x}), \mathrm{l})) \qquad (1)$$

  where $e(H(.,x), l)$ stands for the empirical error between the hypothesis result $H(.,x)$ on the data set x and the respective class labels l.

- Partitions $\wp_{kj}, k = 2, .., T$ define $T - 1$ roots of new trees of the forest.

- Include each binary classifier h_j for each internal node of the trees as a column in the Forest-ECOC matrix M, using the following rule for each class C_r:

$$M(r, j) = \begin{cases} 0 & \text{if } C_r \notin N_j \\ +1 & \text{if } C_r \in \wp_j^1 \\ -1 & \text{if } C_r \in \wp_j^2 \end{cases} \qquad (2)$$

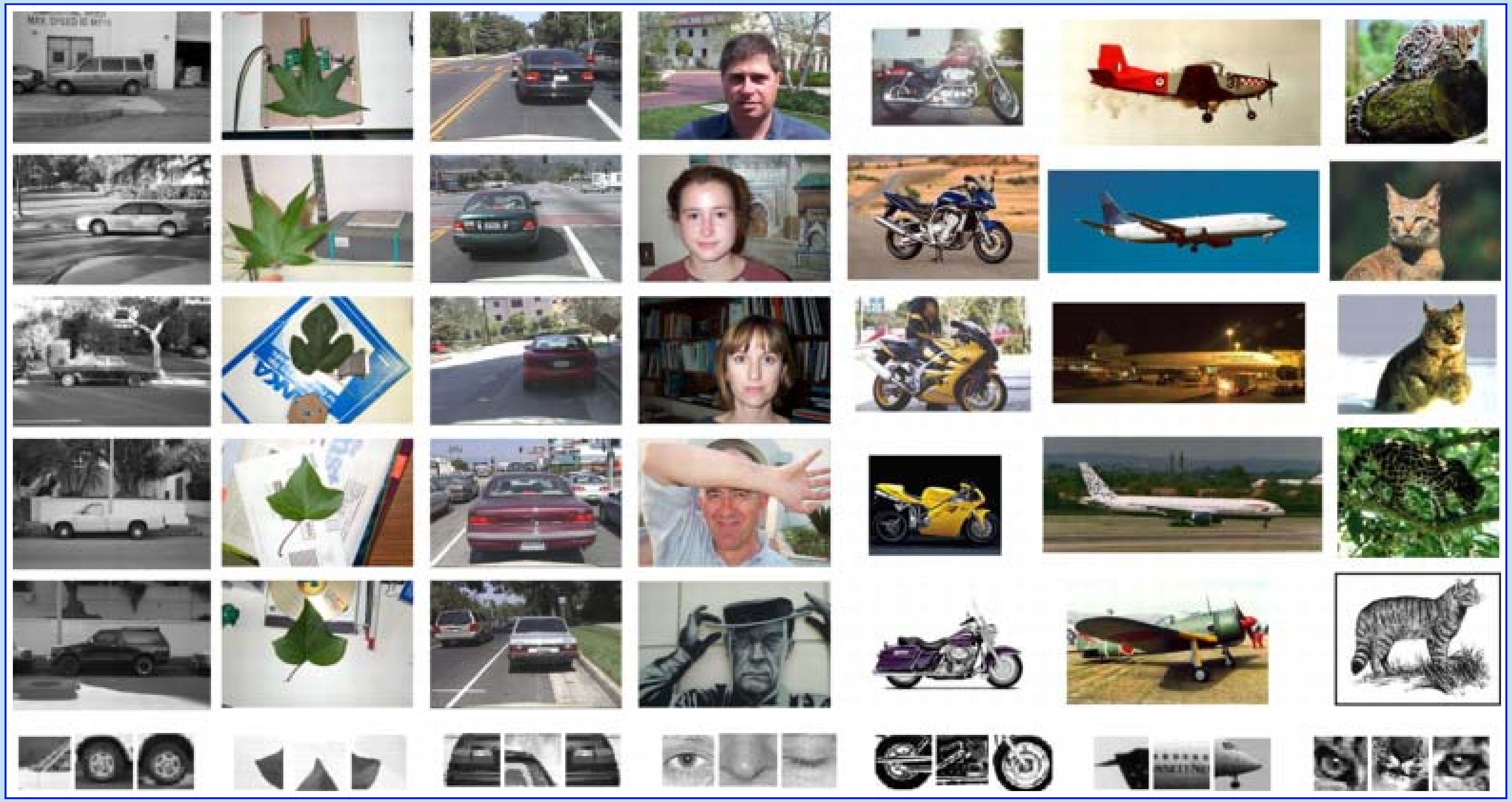


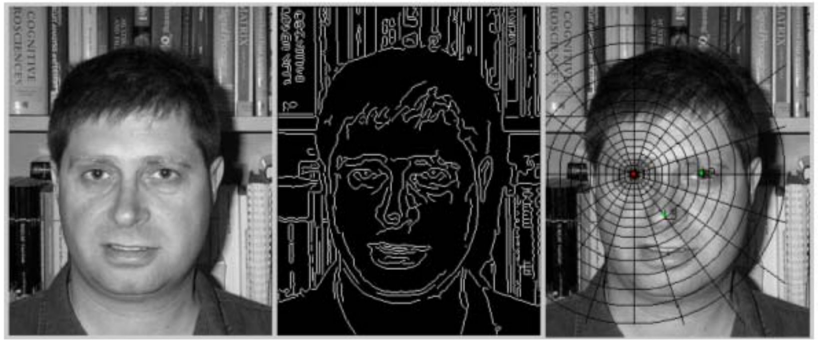
Two optimal trees for a toy problem.

	H_1	H_2	H_3	H_4	H_5	H_6
C_1	1	1	0	1	1	0
C_2	-1	0	1	1	-1	-1
C_3	1	-1	0	1	-1	1
C_4	-1	0	-1	-1	0	0

Forest-ECOC matrix for a toy problem, where H_1, H_2 and H_3 correspond to classifiers of N_1, N_2 and N_3 from the first tree of figure 5, and H_4, H_5 and H_6 to N_1', N_2' and N_5' from the second tree.

UAB CVC RMG





(a) (b) (c)

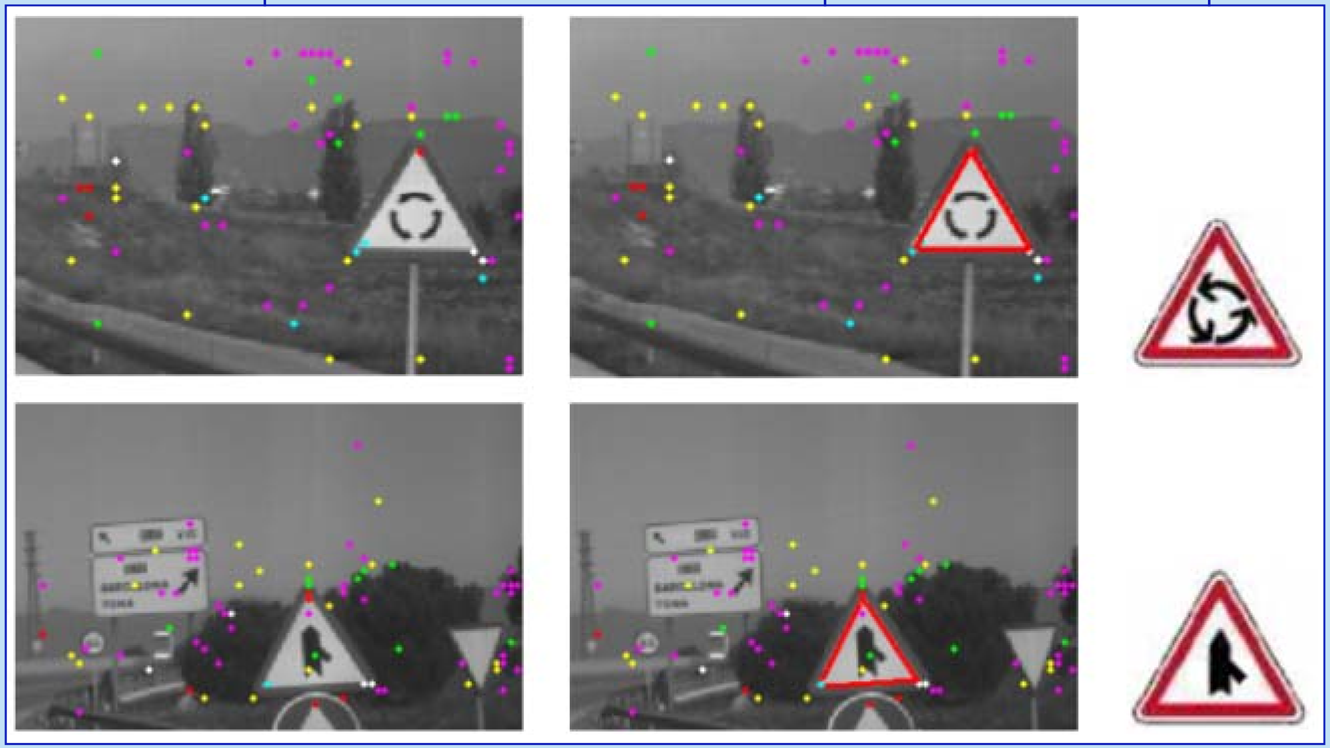
Fergus faces database. (a) Model. (b) Contour points information. (c) Correlogram for a detected landmark.

Category	Fergus	Boosting Context	Boosted Landmarks in Contextual Descriptors
Car (side)	88.50%	90.00%	**96.63%**
Face	96.40%	89.50%	**97.72%**
Motorbike	92.50%	**95.00%**	93.85%
car (rear)	90.30%	96.90%	**99.35%**
Plane	90.20%	**94.50%**	92.50%
Leaf	-	96.30%	**98.85%**
Spotted car	**90.00%**	86.50%	84.00%
Rank	2.50	1.86	**1.57**

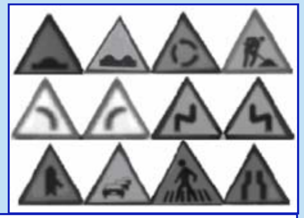
Detected contextual structure

Cascade Landmark candidates

Forest-ECOC classification



Forest-ECOC Classification



KNN – K-Nearest Neighbour

TD – Tangent distance

PCA-KNN – Principal Components Analysis

FLDA – Fisher discriminant analysis

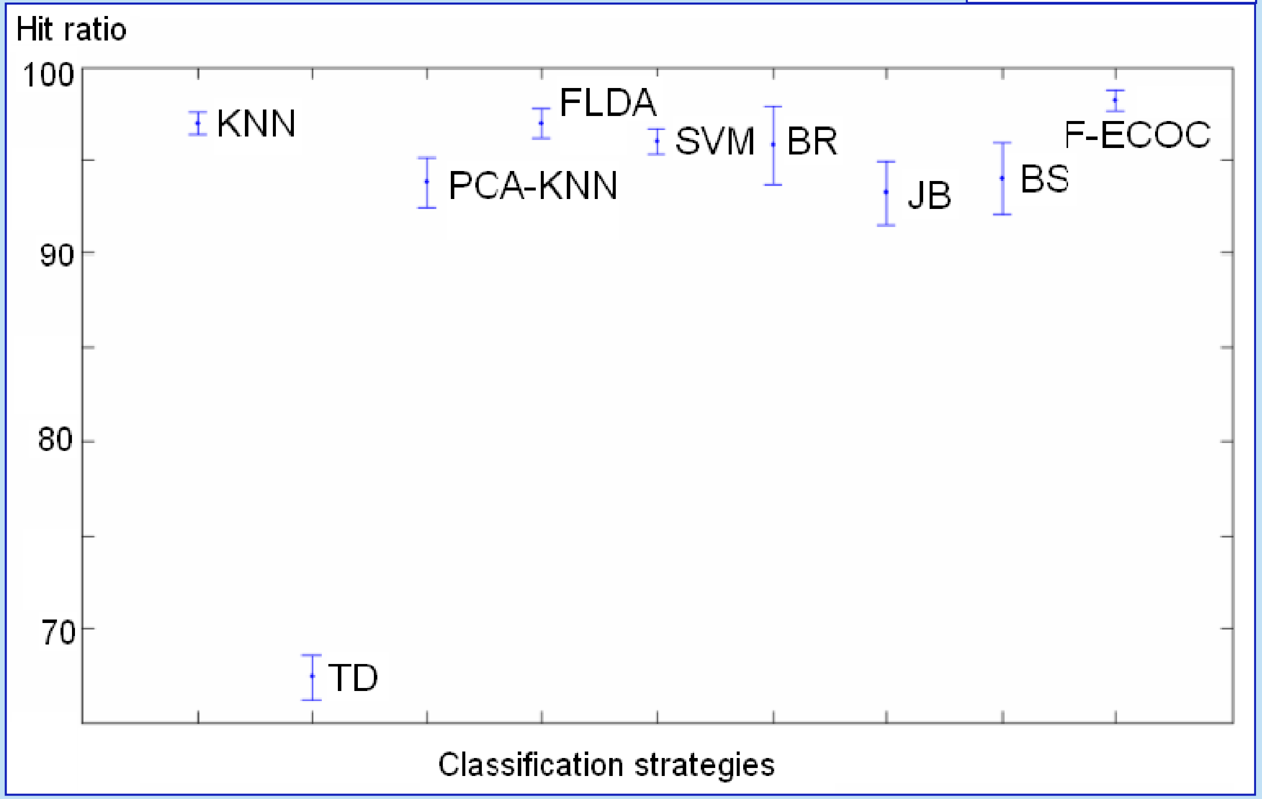
SVM – OSU Linear Support Vector Machine

BR – Gentle Adaboost with HaarLike features

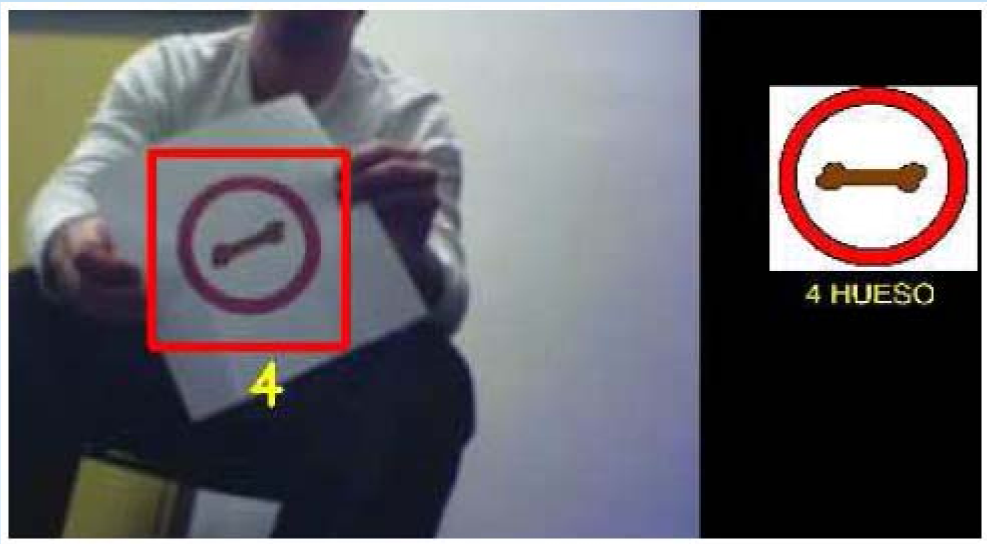
JB – Joint Boosting

BS – Gentle Adaboost samplinbg FLDA

F-ECOC – Forest ECOC



Aibo recognition demo



- Robust and invariant landmarks detection agains noise, slight affine deformations, illumination changes and partial occlusions.
- Robust detection.
- Competitive classification strategy with the state-of-art classification techniques.

Open

Changes of view

Deformable context

Thank you!

RMG