Robots and People

Vlad Estivill-Castro
Jordi Radev

Humanoids Learning who are Teammates and who are Opponents

Thanks for your interest

Vlad Estivill-Castro

School of Information and Communication Technology
Institute for Intelligent and Integrated Systems
Griffith University
Australia

Jordi Radev

Departament de Technologies de la Informacio
Universitat Pompeu Fabra
Spain

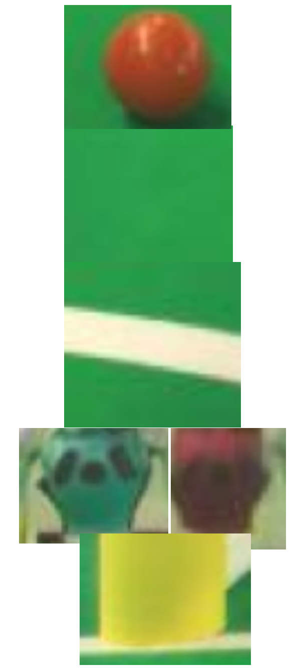
Context

"Given the low computational resources of the NAO robot, their perception is a rather difficult task, making the application of state-of- the-art computer vision approaches such as the *Histogram of Oriented Gradients (HOG)* impossible"

(A. Fabisch, T. Laue, and T. Röfer, "Robot recognition and modeling in the in the RoboCup standard platform league," in *5th Workshop on Humanoid Soccer Robots at Humanoids*, 2010)

Results:

- On board of the NAO we learn
  - the color of the ball
  - the color of the playing surface
  - the color of the line markings on the playing surface
  - the color of the teams shirts
  - the color of the goals

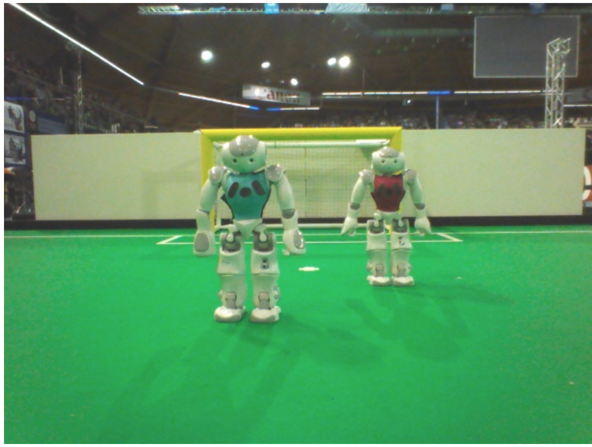


official colors

Robots and people

How do we do this

- Use shape
  - looking down
    - most background is the surface
    - looking for circles of certain established sizes is the ball
    - the rest are line markings
  - looking ahead
    - find other NAOs
      - HOG
    - find goals
      - HOG

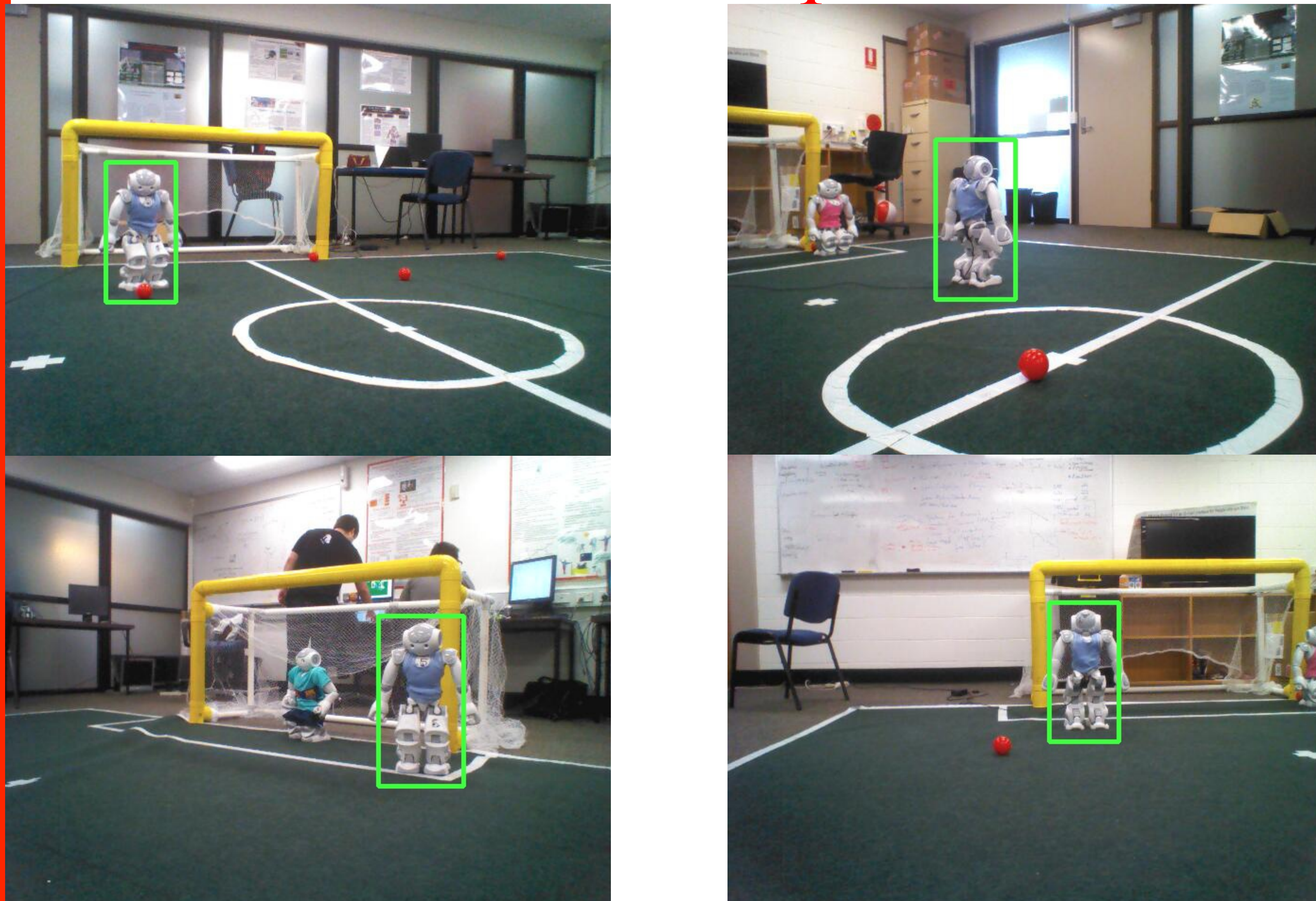


HOG (Histogram of Gradients)

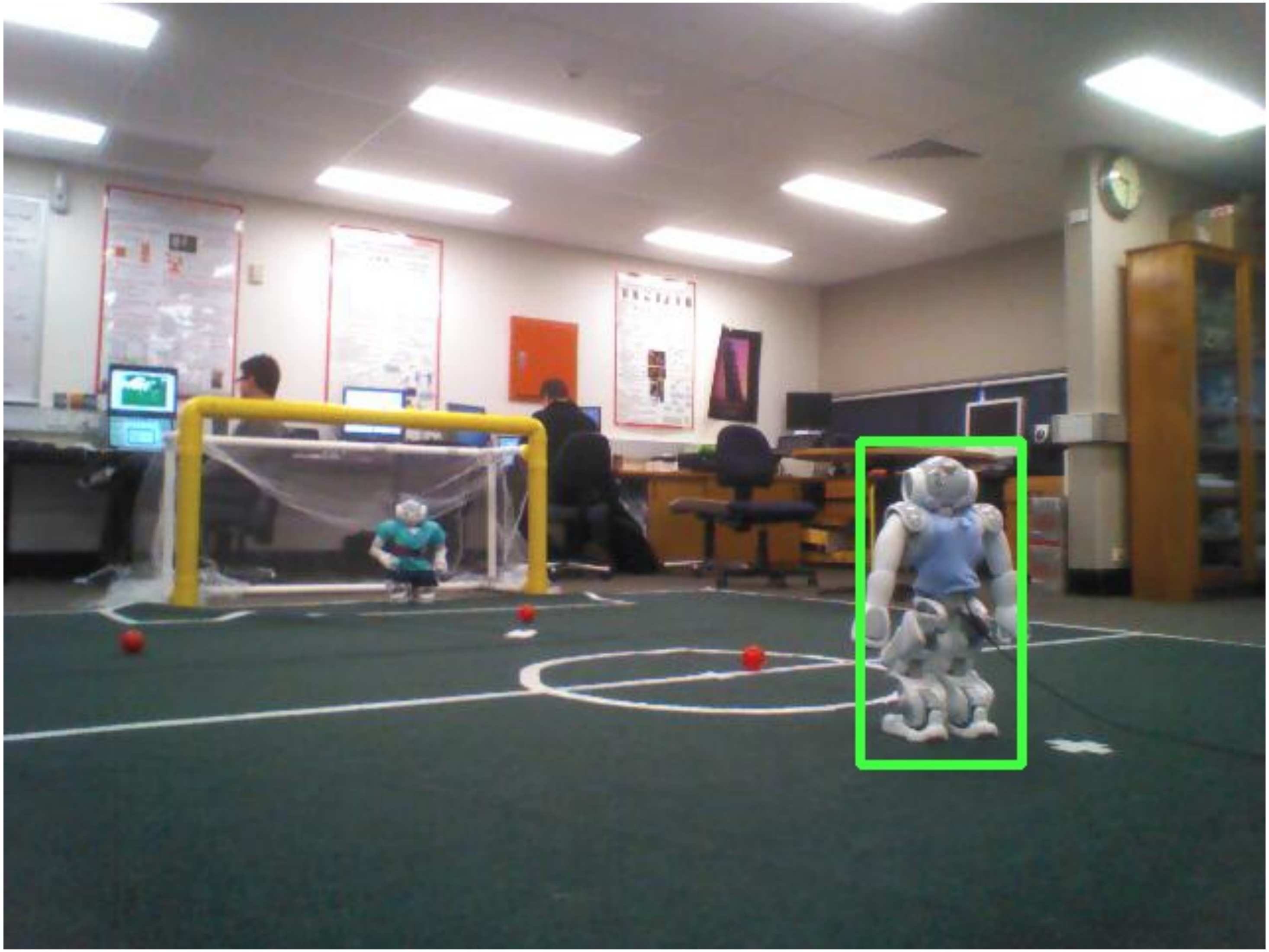
- Recognized as "the" technique for human form
- NAOs are humanoids, but simpler
- Thus, we can tailor the HOG

Robots and people

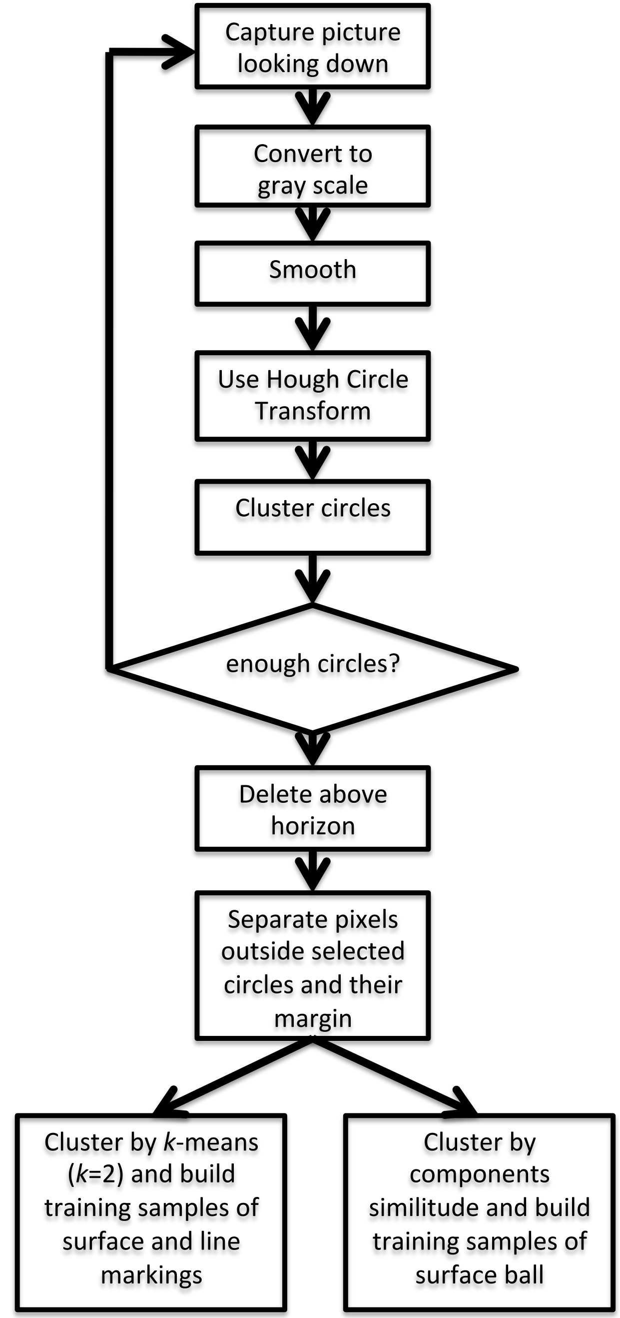
What is the HOG output



What is the HOG output



Environment Analysis



Environment Analysis

Original image

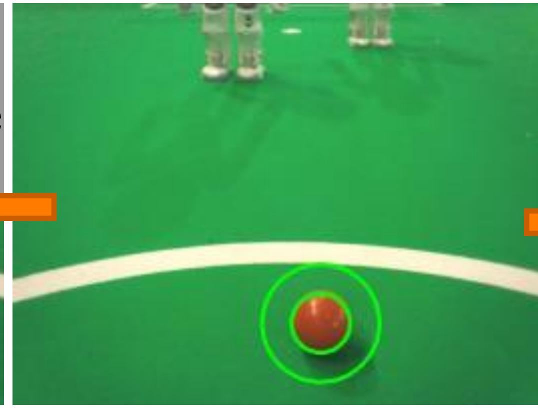
Hough Circle Transform

Draw circle, but separate on bigger circle (a margin)

Robots and people

Environment Analysis

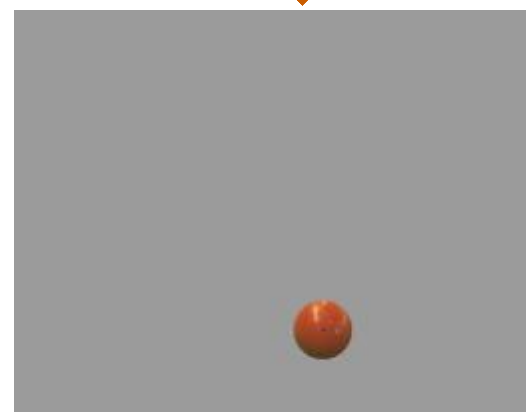
Take bottom half without the big circle part



Take ball circle part

Separate:
1) lines and field
vs
2) ball

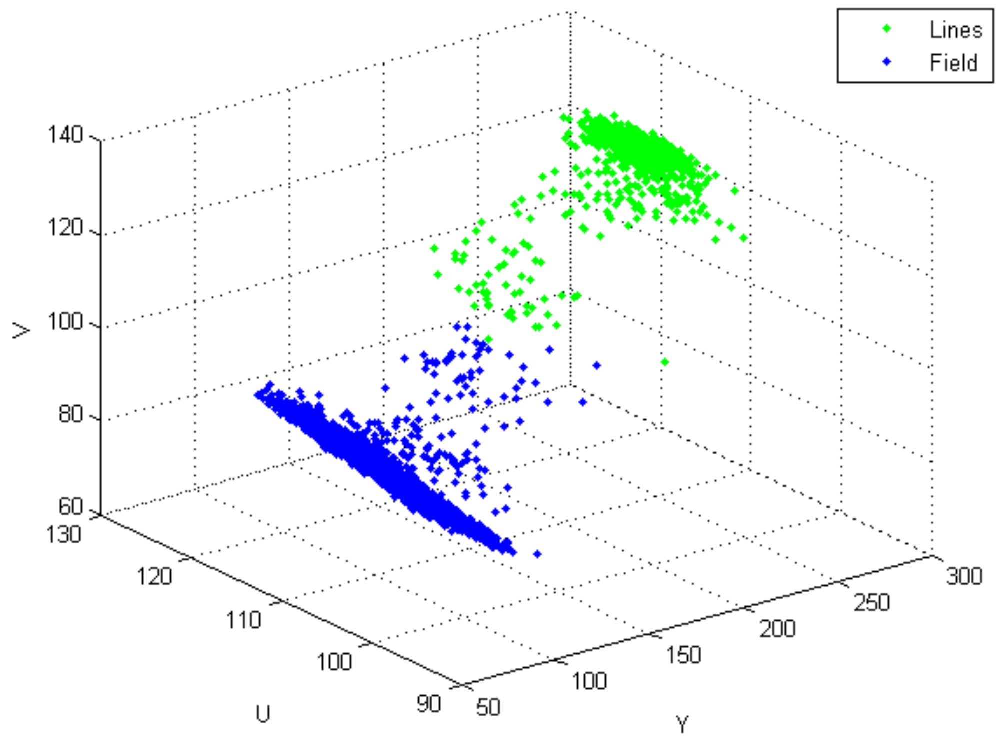
Cluster colours expecting two classes (k-means)



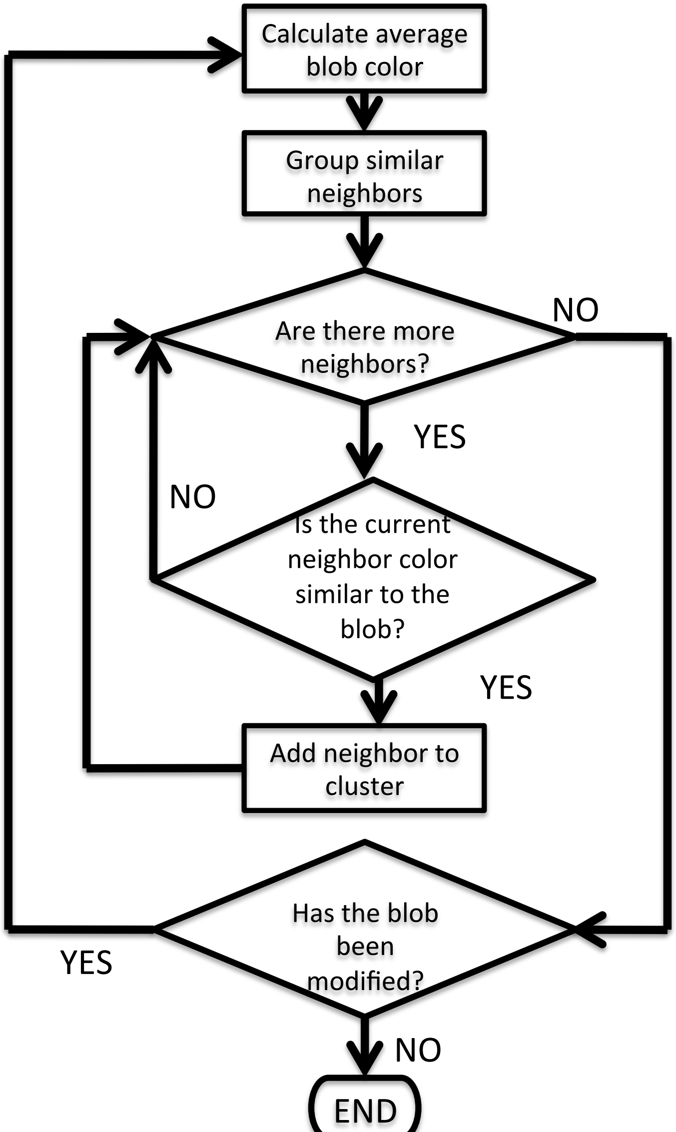
cluster (x,y,r) expecting one (histogram)

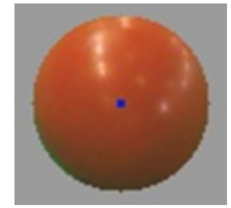
Environment Analysis

- *k*-means ($k = 2$) to separate the lines pixels from the field pixels

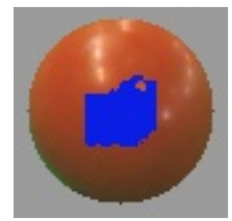


Environment Analysis

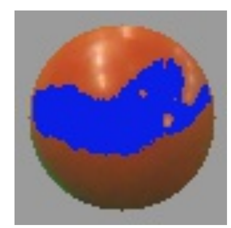




1 iteration
9 pixels



13 iterations
579 pixels

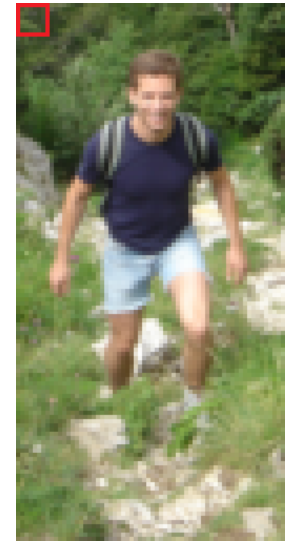


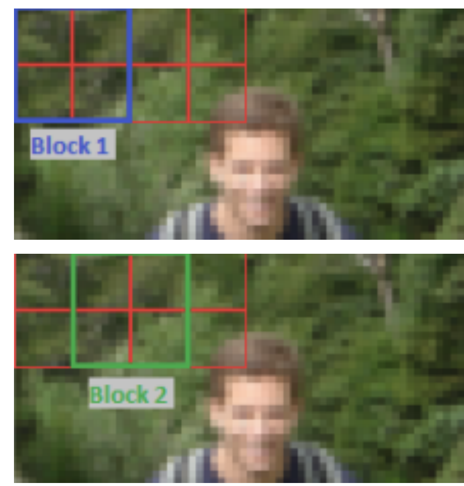
76 iterations
2546 pixels

Robots and people

Team colour detection

- Histogram of oriented gradients: feature descriptors used for object detection.
  - Image divided in cells
  - Gradients for each direction within a cell are quantified
  - Cells are grouped in blocks
  - Window is defined for detection

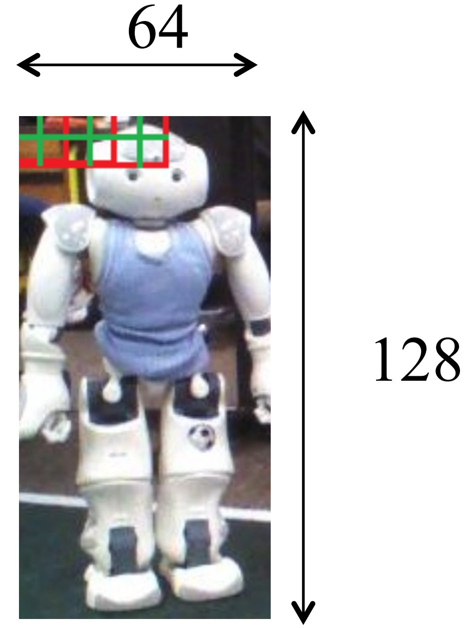


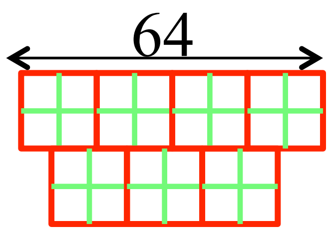


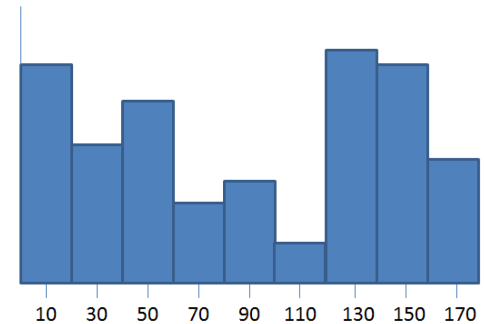
Robots and people

Training

- Window size: 64 x 128
- Block size: 16 x 16
- Block stride: 8 x 8
- Cell size: 8 x 8
- Number of bins: 9



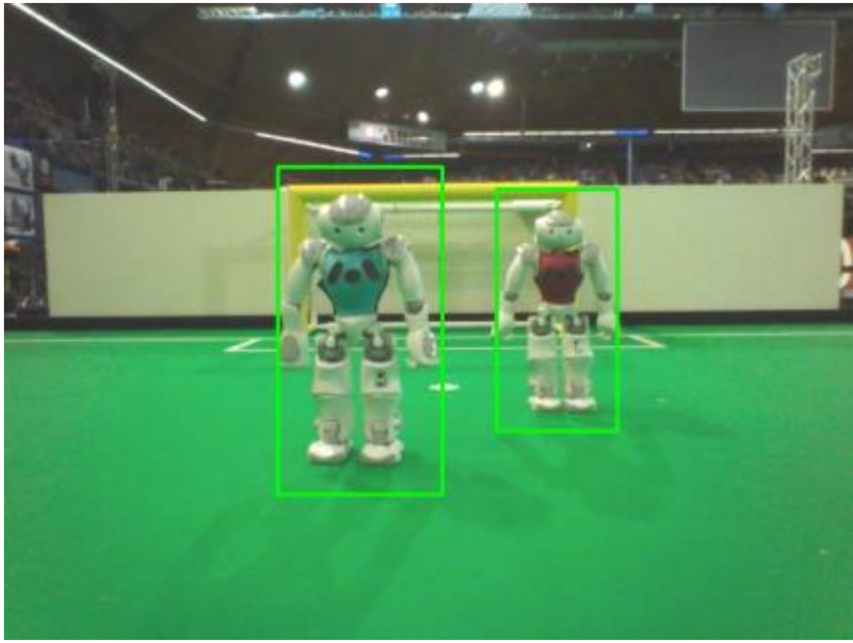




Feature vector size = (4+3) x (8+7) x 4 x 9 = 3780

Humanoid detection

- Shape-based NAO detection by SVM
- Training:
  - Number of positive training examples: 824
    - (412 images flipped horizontally)
  - Number of negative examples: 4621
    - (images flipped vertically and cut in windows of size 64x128)

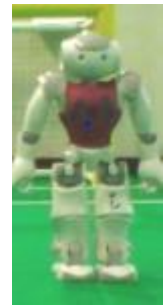


Robots and people

Torso detection

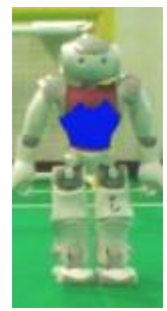
Pixel clustering for Torso detection

1 iteration
9 pixels



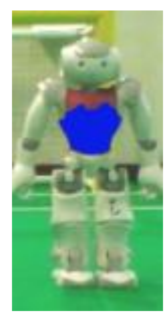
1 iteration
9 pixels

18 iterations
995 pixels



18 iterations
885 pixels

53 iterations
1347 pixels



108 iterations
2679 pixels

Goal localization

- Objectives:
  - Detect goals
  - Extract the goal colour
- Methods
  - Feature extraction using HOG
  - Detection by SVM

Goal colour learning

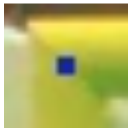
- Window size: 16 x 16
- Block size: 8 x 8
- Block stride: 4 x 4
- Cell size: 4 x 4
- Number of bins: 9

Feature vector size = 3 x 3 x 4 x 9 = 324

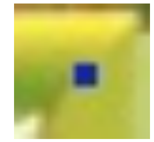
Goal learning parameters

- Shape-based goal corner detection by SVM
- Training:
  - Number of positive training examples: 344
    - (172 images flipped horizontally)
  - Number of negative examples: 316800
    - (images flipped vertically and cut in windows of size 16 x 16)

Yellow pixels selection



1 iteration
9 pixels



1 iteration
9 pixels

5
iterations
66 pixels

5
iterations
84 pixels

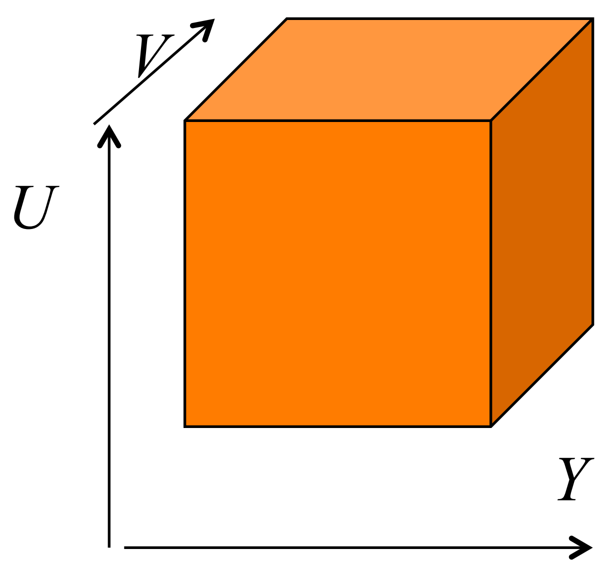
12
iterations
201 pixels

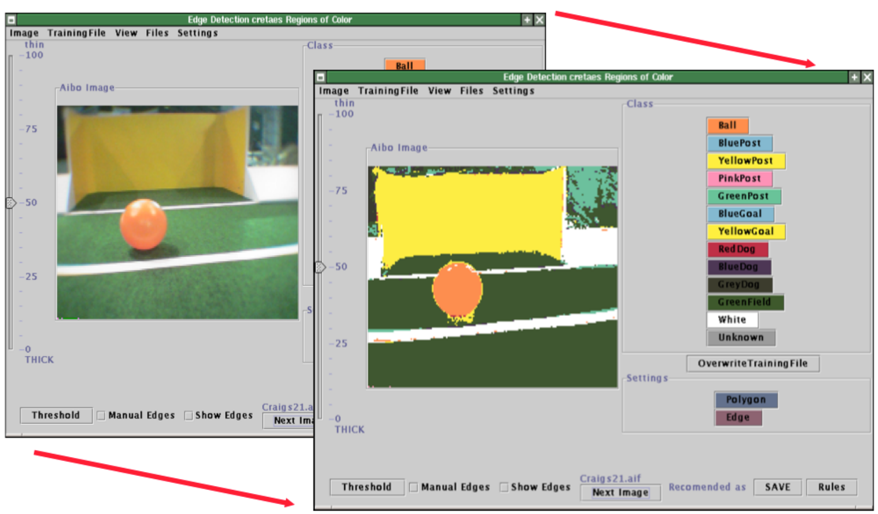
9
iterations
148 pixels

Robots and People

What we learn

- mapping (Y,U,V) to colour for segmentation.
  - Mapping is a classifier
  - *Colour_class*: $YxUxV \rightarrow$ *Colour*
  - *Colour_Class(y,u,v)=Orange*
  - $|Y|x|U|x|V|=256^3$

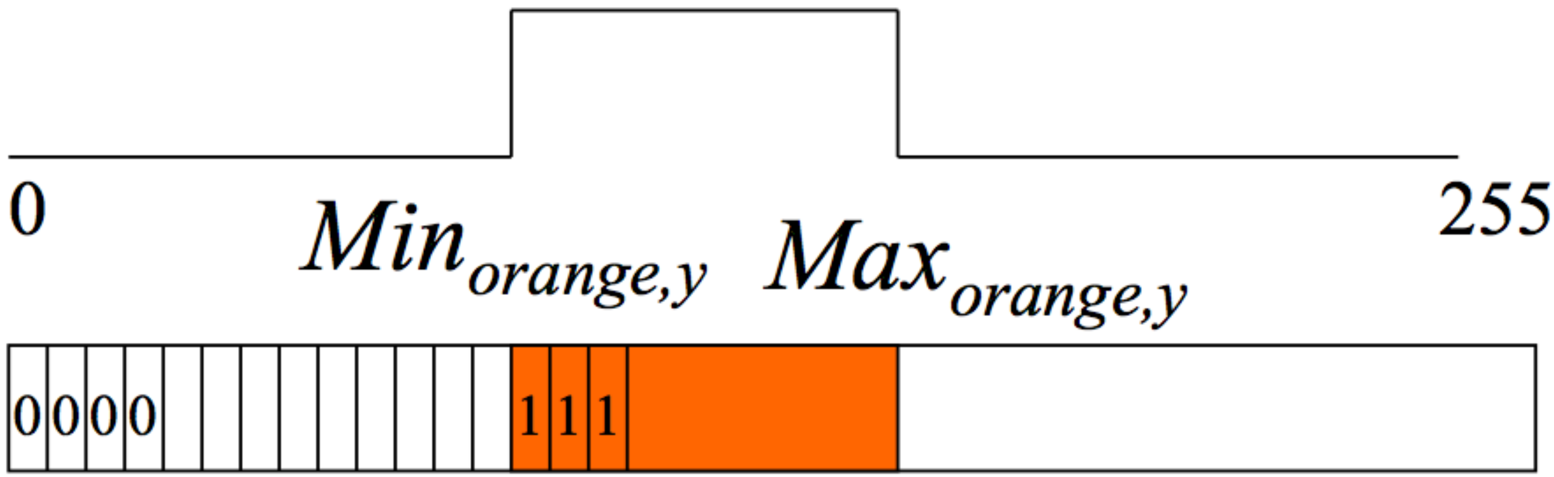




Robots and people

The bit-Map for a color projection

- *If* $Min_{range.v} \leq Y$ *AND* $Y \leq Max_{\ orange.v}$

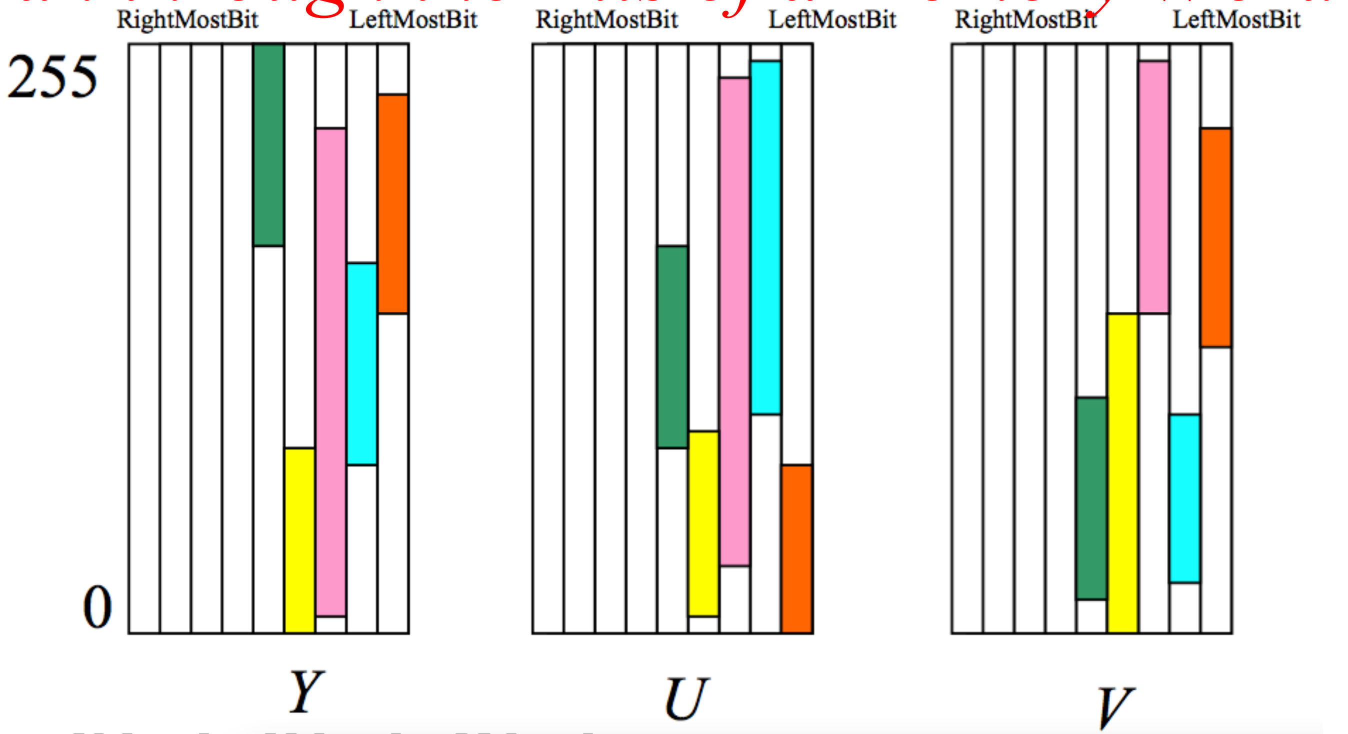


- A classifier is a table look up
  - C/C++bit-wise AND operation

```
Colour_Orange=Y[y]&U[u]&V[v]
```

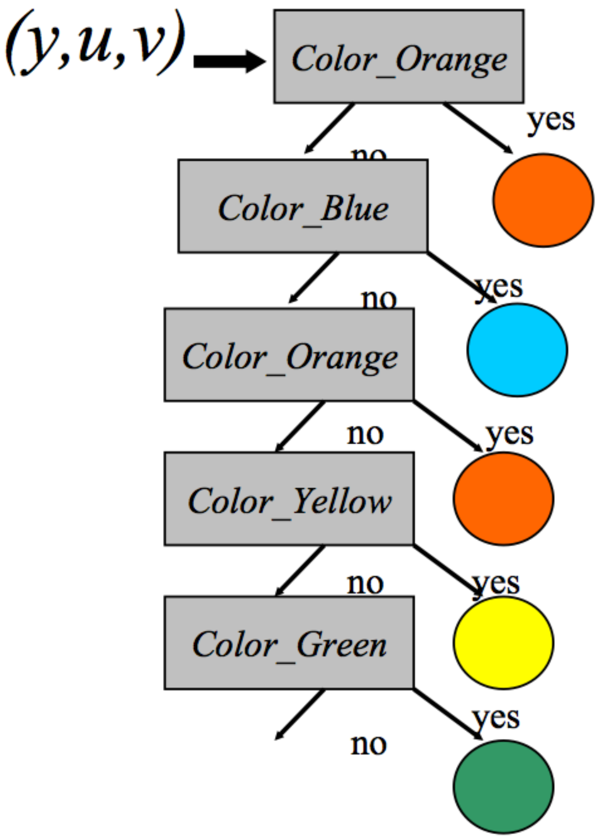
A Decision List

is a Scan through the Bits of a Memory Word




```
colour=Y[y]&U[u]&V[v];
colour_id=0 ;
while (colour & 1 == 0)
   { colour>>=1; colour_id++;}
```

List can repeat simple classifiers



Comparison of Decision List

- ANN using snns were 20K times slower
- *k*-NN with Quadtrees and Decision Trees (Weka) were 2K times slower

	DL	Look-up Table	Ratio
Maximum	2.87ms	2.46ms	1.16
Average	2.33ms	1.41ms	1.65
Minimum	2.08ms	1.27ms	1.63

Accuracy with Decision Lists is Marginally Better

Algorithm	10-fold accuracy	Lowest accuracy per class	Largest 2-class confusion	size	Learning time	Test set accurac y
PART	99.0%	96% (yellow goal)	10 blue dog Vs gray dog	26 Rules	1.15s	99.3%
k-NN	99.3%	97% (blue dog)	8 red dog Vs gray dog	*k*=3 6,226 Instanc es	0s	99.7%
DT	98.8%	95% (yellow goal)	10 red dog Vs gray dog	34 leaves 67 nodes	1.27s	99.6%
Look-up Table	71.6%	64% (yellow goal)	45 yellow goal vs orange ball	11 rules	manual	68.2%

How we learn

Merge of all training sets

Balancing of the classes

Building classifier

Loading the new classifier on the vision module

Qualitative results

- Environment analysis:
  - Performs well on different fields (adaptive smoothing)
- Teams detection
  - The trained SVM with HOG features detects NAOs in different positions and orientations
- Goal localization
  - The SVM with HOG features detects the corners of the goals in most cases

Robots and people

Quantitative results

TABLE I. RESULTS OF 6-FOLD CROSS VALIDATION TO ASSES ACCURACY OF HUMANOID ROBOT DETECTION.

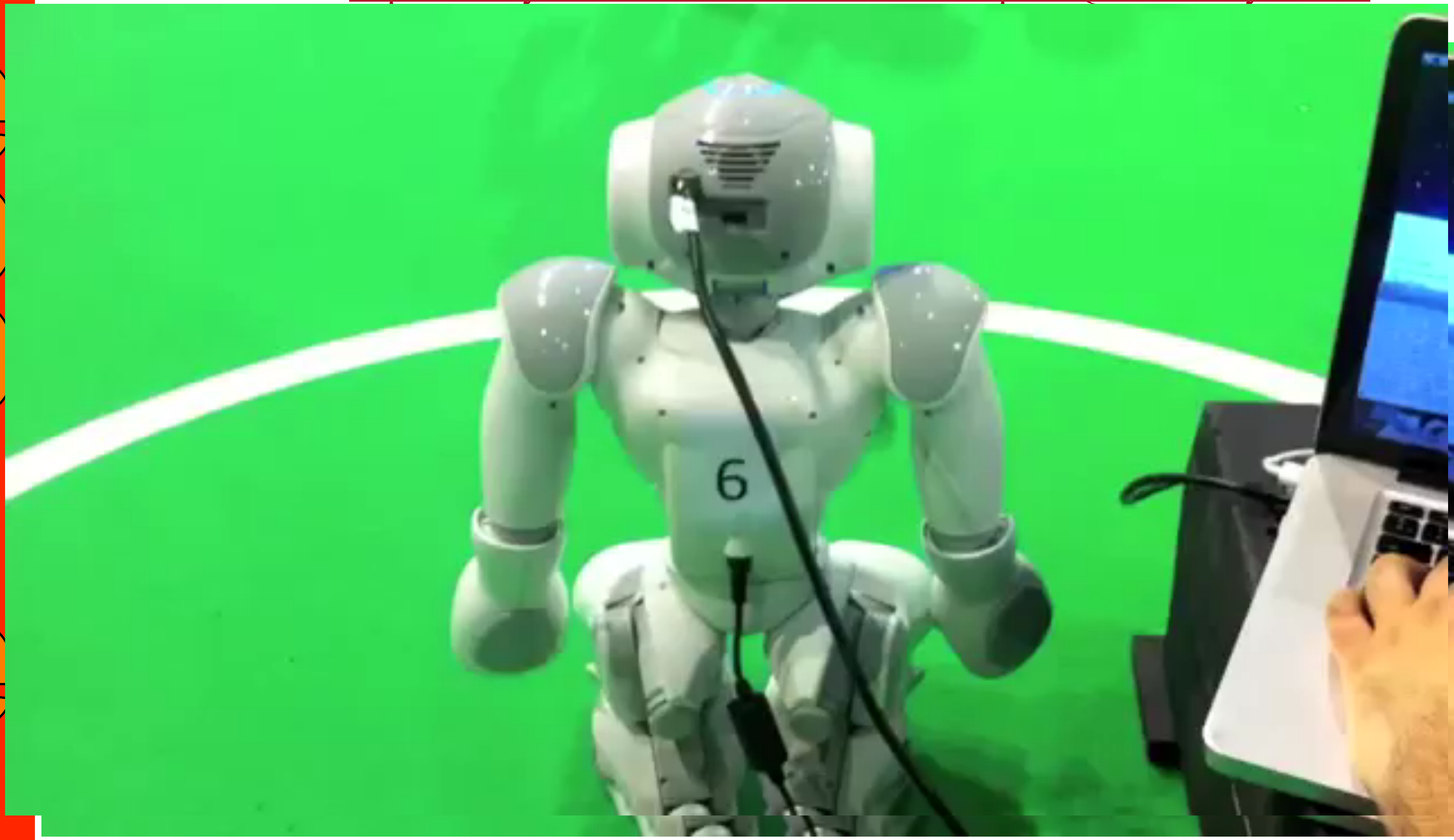
True positives	452
False positives	3
True negatives	2769
False negatives	60
Average precision	99.33%
Average recall	88.26%

TABLE II. RESULTS OF 6-FOLD CROSS VALIDATION TO ASSES ACCURACY OF GOAL CORNER DETECTION.

True positives	187
False positives	61
True negatives	64,406
False negatives	41
Average precision	75.55%
Average recall	82.57%

In one minute, at the RoboCup venue

- http://www.youtube.com/watch?v=DEMaRopZSrQ&feature=youtu.be



Conclusions

- The classifier built by the procedure is capable to segment the images and to recognize the important soccer elements
- The procedure is fast enough to be performed within a minute

Future lines

- Incorporate images and learning while game play
- Define strategies to integrate shape-based and colour-based detection
- Improve unbalanced classification of the goal

THANK YOU