Human Body Segmentation with Multi-limb Error-Correcting Output Codes Detection and Graph Cuts Optimization

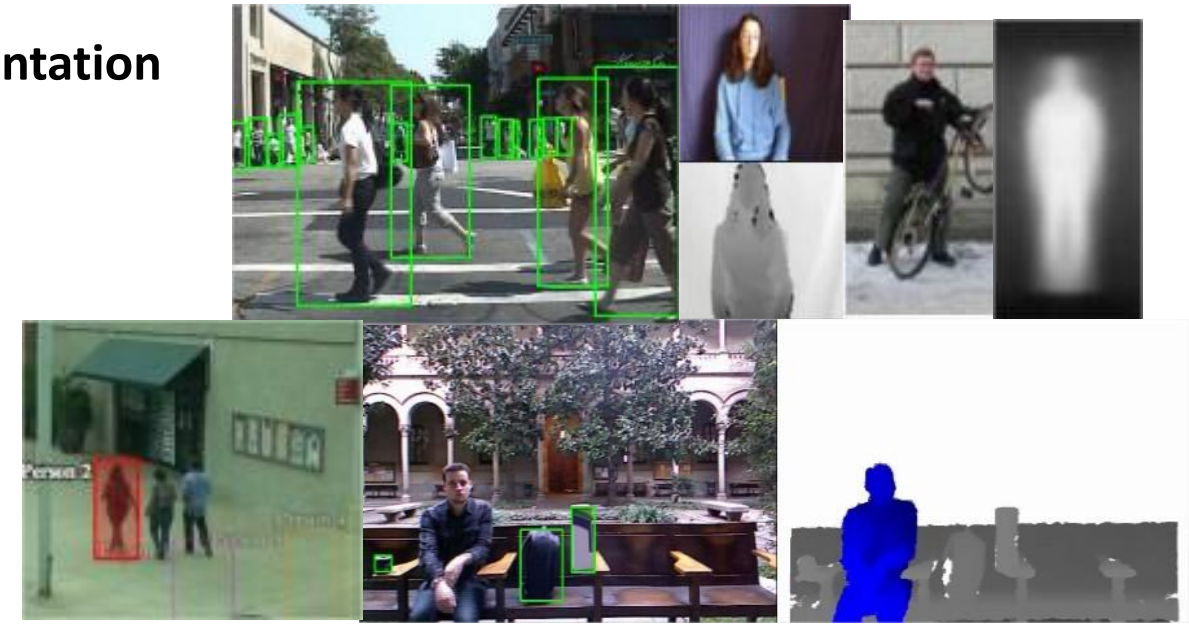
Daniel Sánchez, Juan Carlos Ortega,
Miguel Ángel Bautista & Sergio Escalera

Outline

1. Motivation
2. Proposal
3. Results
4. Conclusions

Motivation

- **User Detection/Segmentation**



- **Applications:** medicine, photography, sign language…

What we use

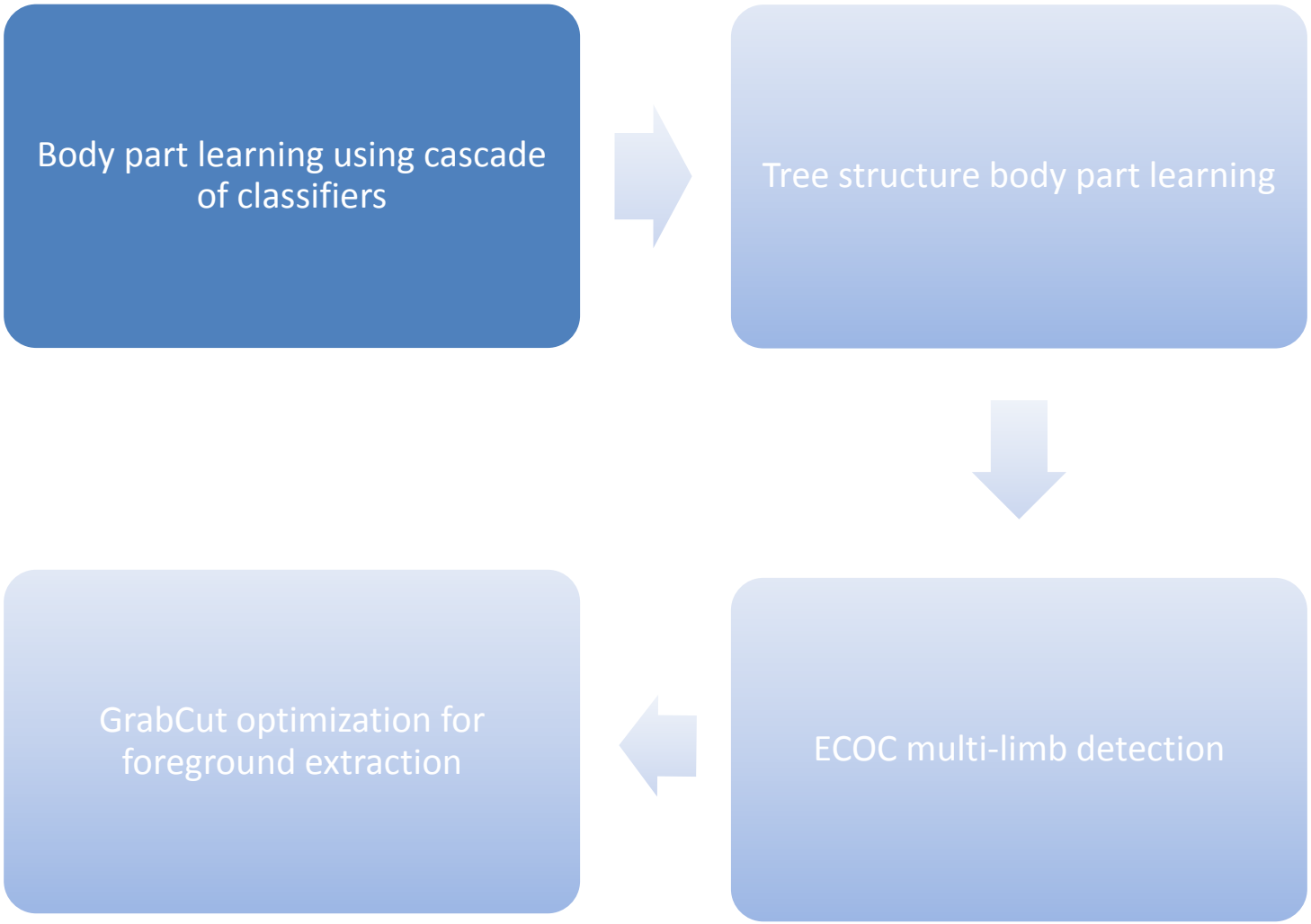
What we get

Body part learning using cascade of classifiers

Tree structure body part learning

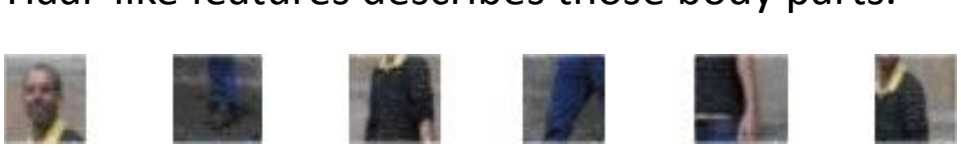
ECOC multi-limb detection

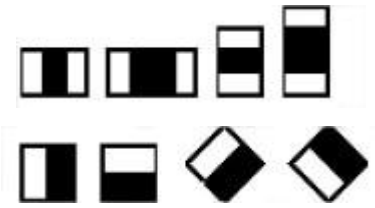
GrabCut optimization for foreground extraction


Body part learning using cascade of classifiers
Tree structure body part learning
ECOC multi-limb detection
GrabCut optimization for foreground extraction

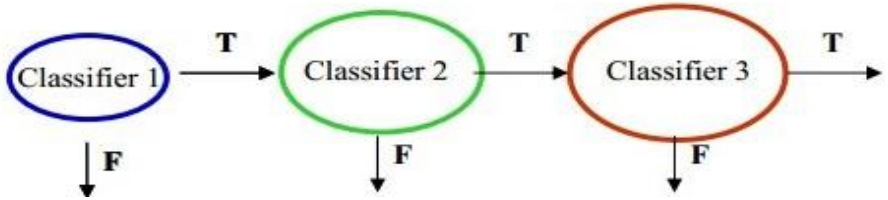
- Body parts rotational invariant by computing dominant orientation.

- Haar-like features describes those body parts.

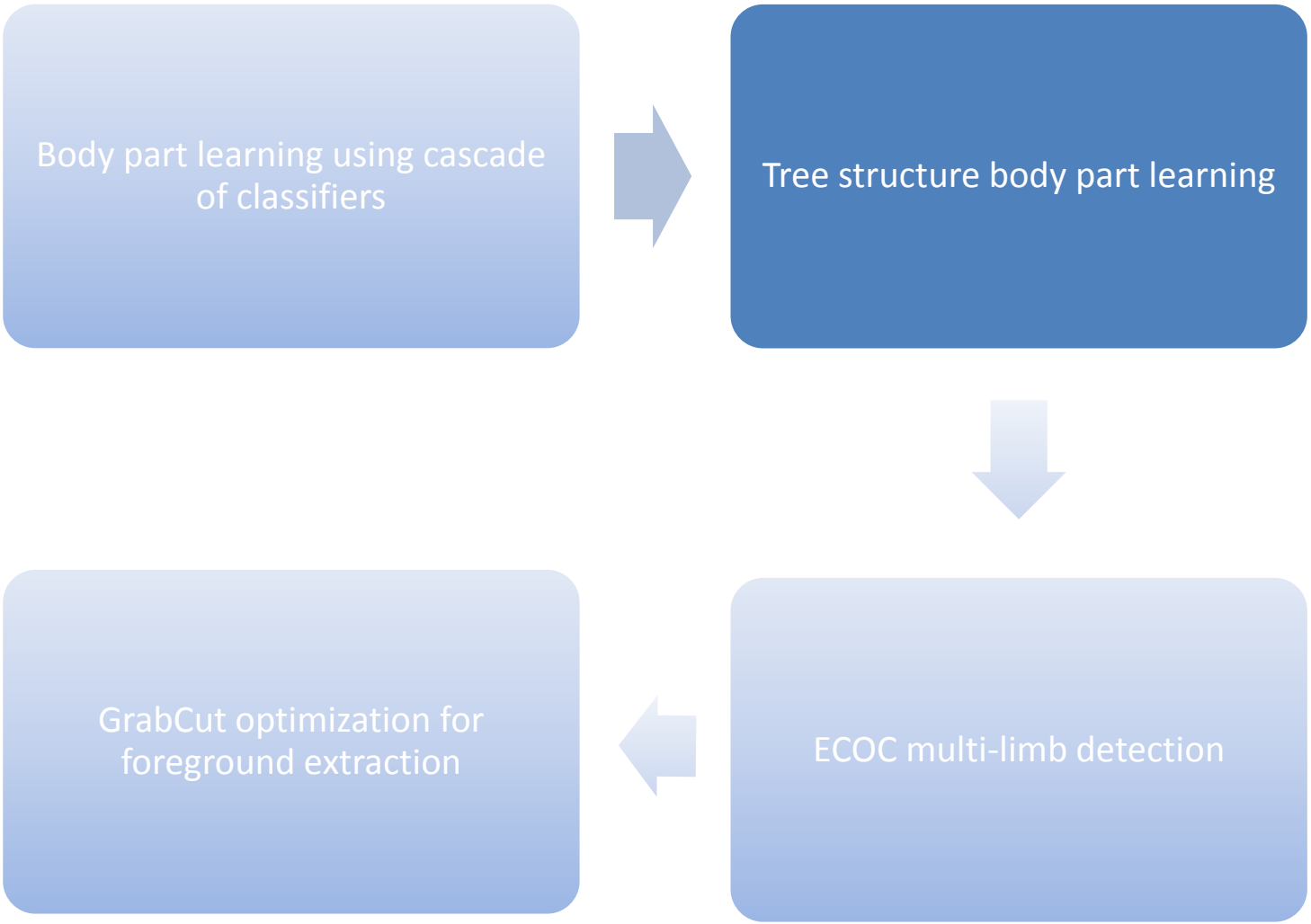




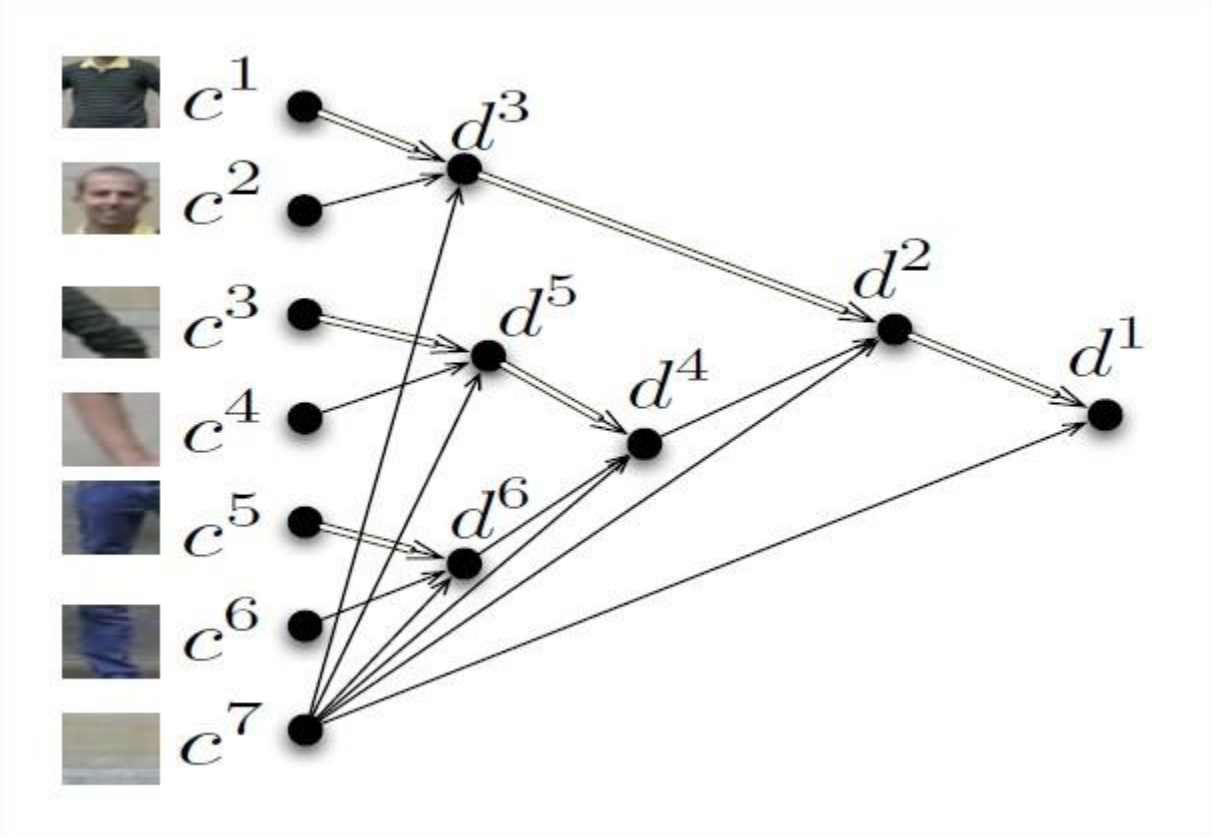
- Adaboost as the base classifier in the cascade architecture.



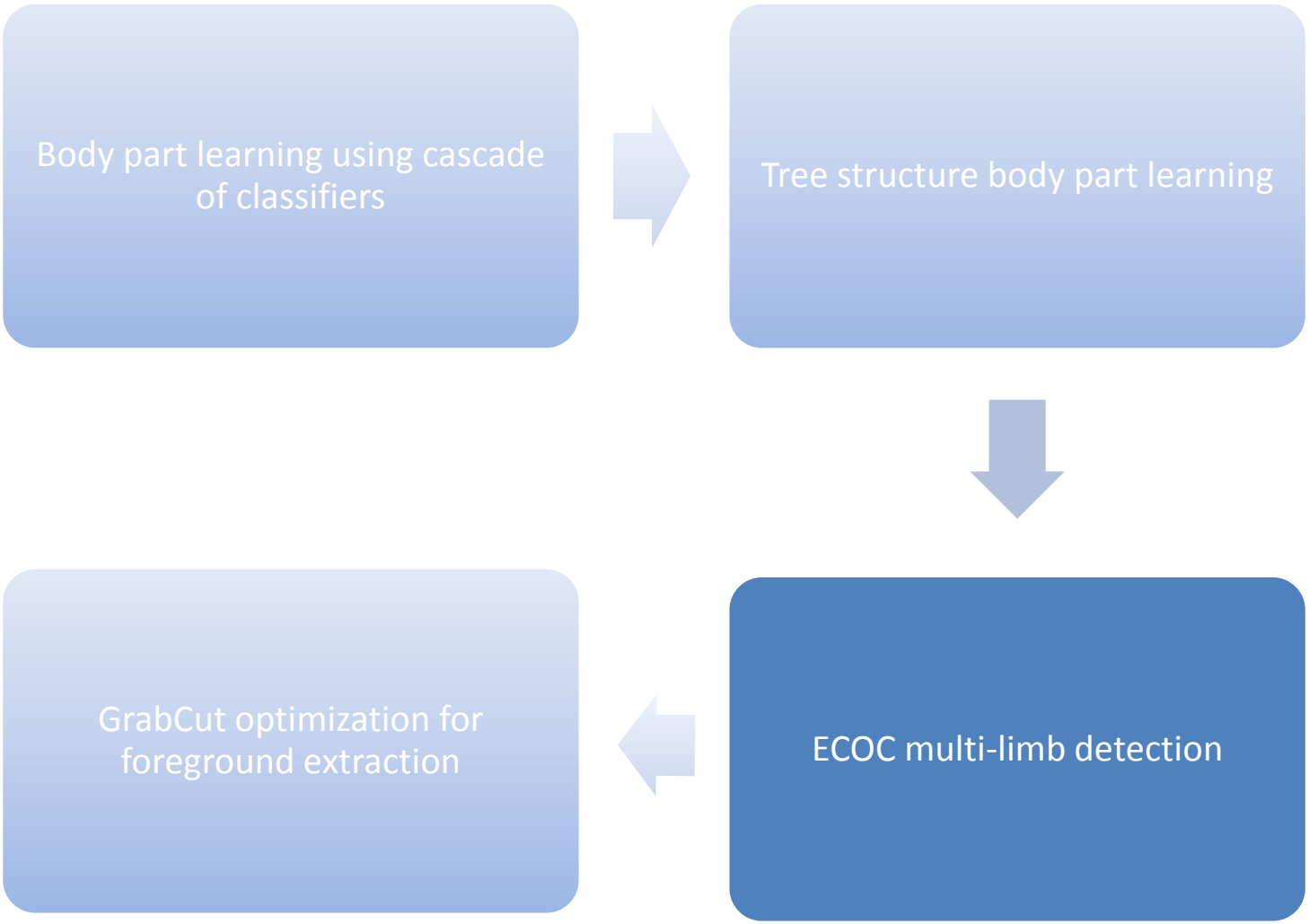
P. Viola, M. Jones, Rapid object detection using a boosted cascade of simple features, in: CVPR, Vol. 1, 2001.
Y. Freund, R. Schapire, A decision-theoretic generalization of on-line learning and an application to boosting, in: EuroCOLT, 1995, pp. 23-37.


Body part learning using cascade of classifiers
Tree structure body part learning
ECOC multi-limb detection
GrabCut optimization for foreground extraction

- Define the groups of limbs to be learnt by each individual cascade.

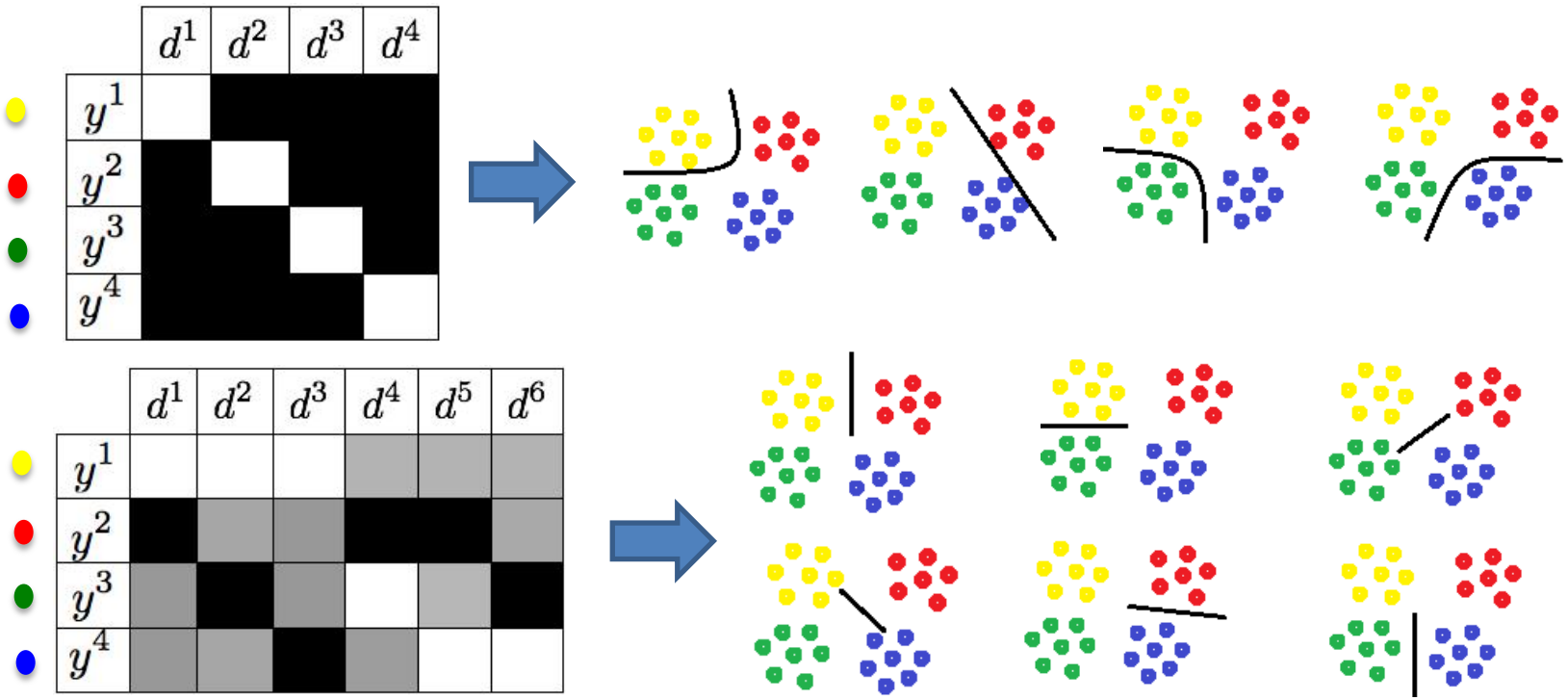


S. Escalera, O. Pujol, P. Radeva, On the decoding process in ternary error-correcting output codes, PAMI 32 (2010) 120-134.


Body part learning using cascade of classifiers
Tree structure body part learning
ECOC multi-limb detection
GrabCut optimization for foreground extraction

- In classification tasks, the goal is to classify an object among a certain number of possible categories.

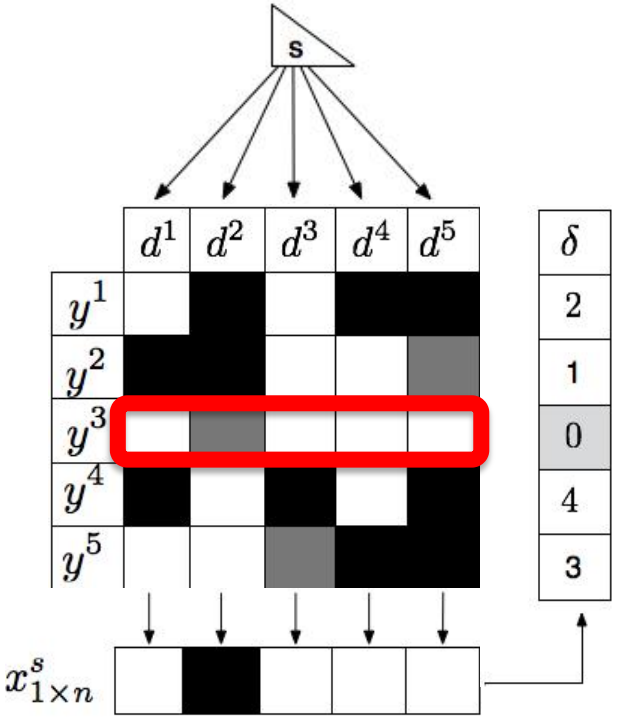
- This framework is composed of two different steps :
  - **Coding** : Decompose a given *N*-class problem into a set of *n* binary problems.
  - **Decoding** : Given a test sample *s*, determine its category.



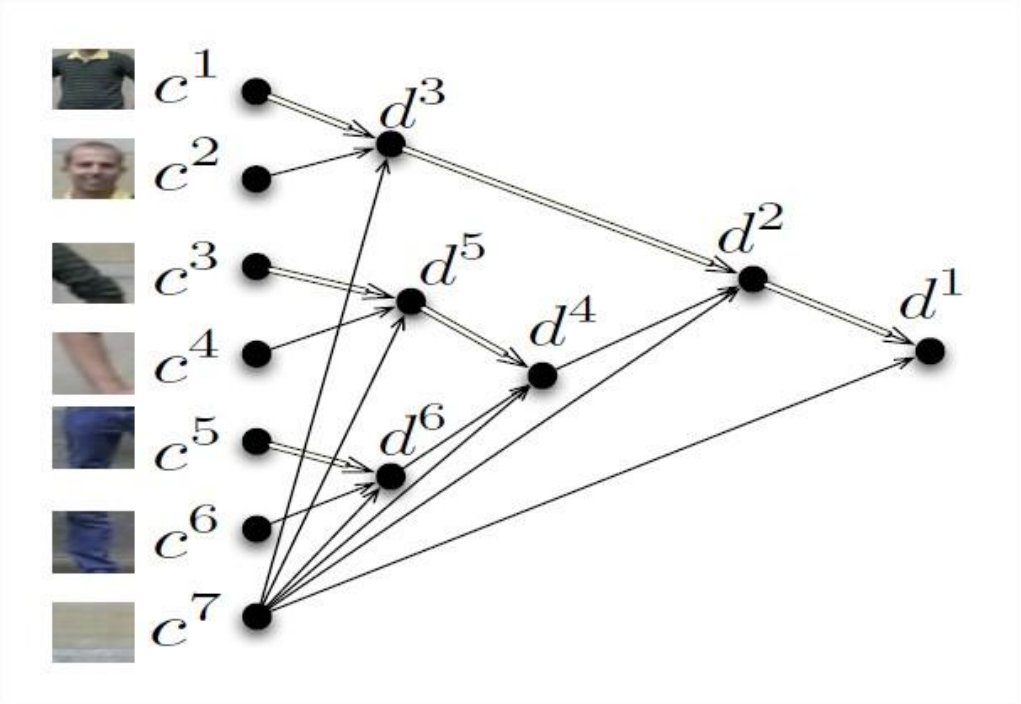
- At the decoding step a new sample s is classified by comparing the binary responses to the rows of M by means of a decoding measure δ.

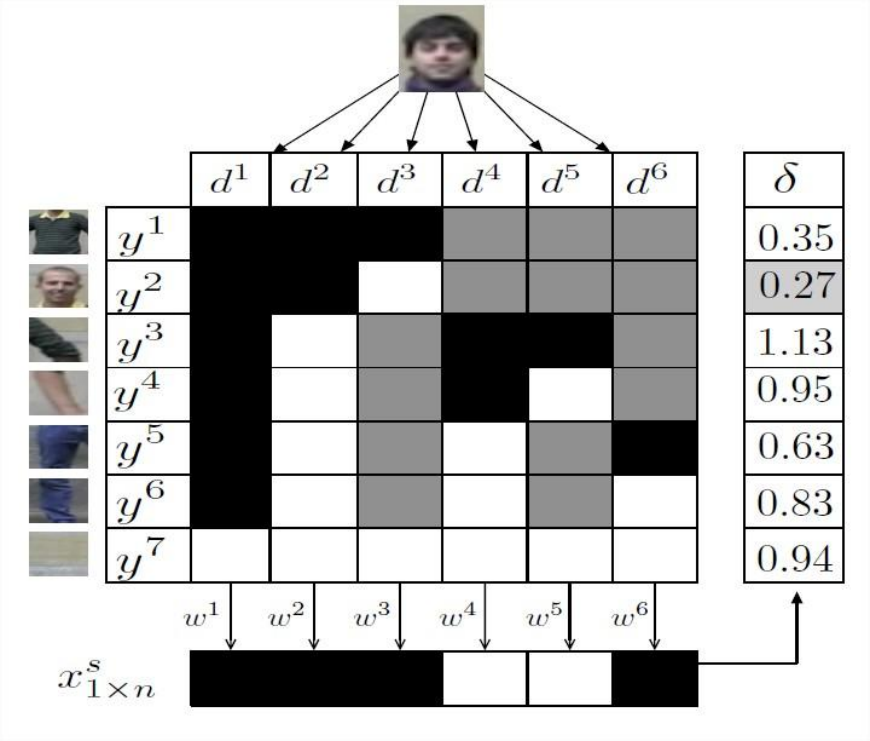
- Different types of decoding based on the distance used (i.e. Hamming, Euclidean, etc.)

$$\arg\min_i \delta(x^s, y^i)$$



- We propose to use a predefined coding matrix in which each dichotomy is obtained from the body part tree structure.

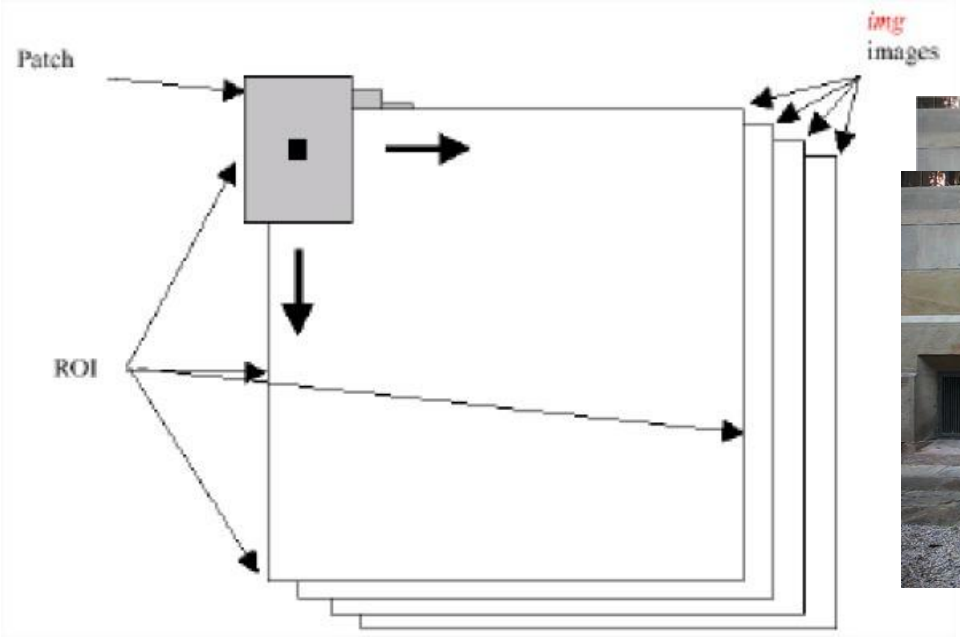




S. Escalera, D. Tax, O. Pujol, P. Radeva, R. Duin, Subclass problem-dependent design of error-correcting output codes, PAMI 30 (6) (2008) 1-14.

M. A. Bautista, S. Escalera, X. Baro, P. Radeva, J. Vitria, O. Pujol, Minimal design of error-correcting output codes, Pattern Recogn. Lett. 33 (6) (2012) 693-702.

- In order to classify a new sample we apply a sliding window over the image:

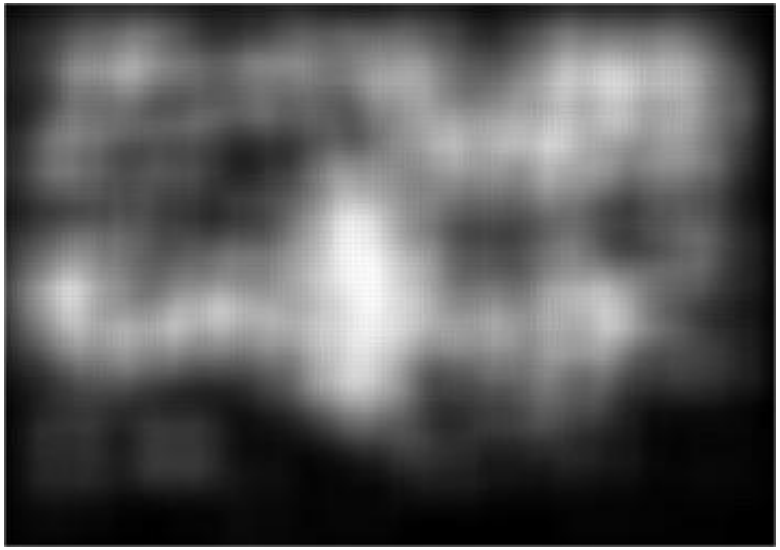


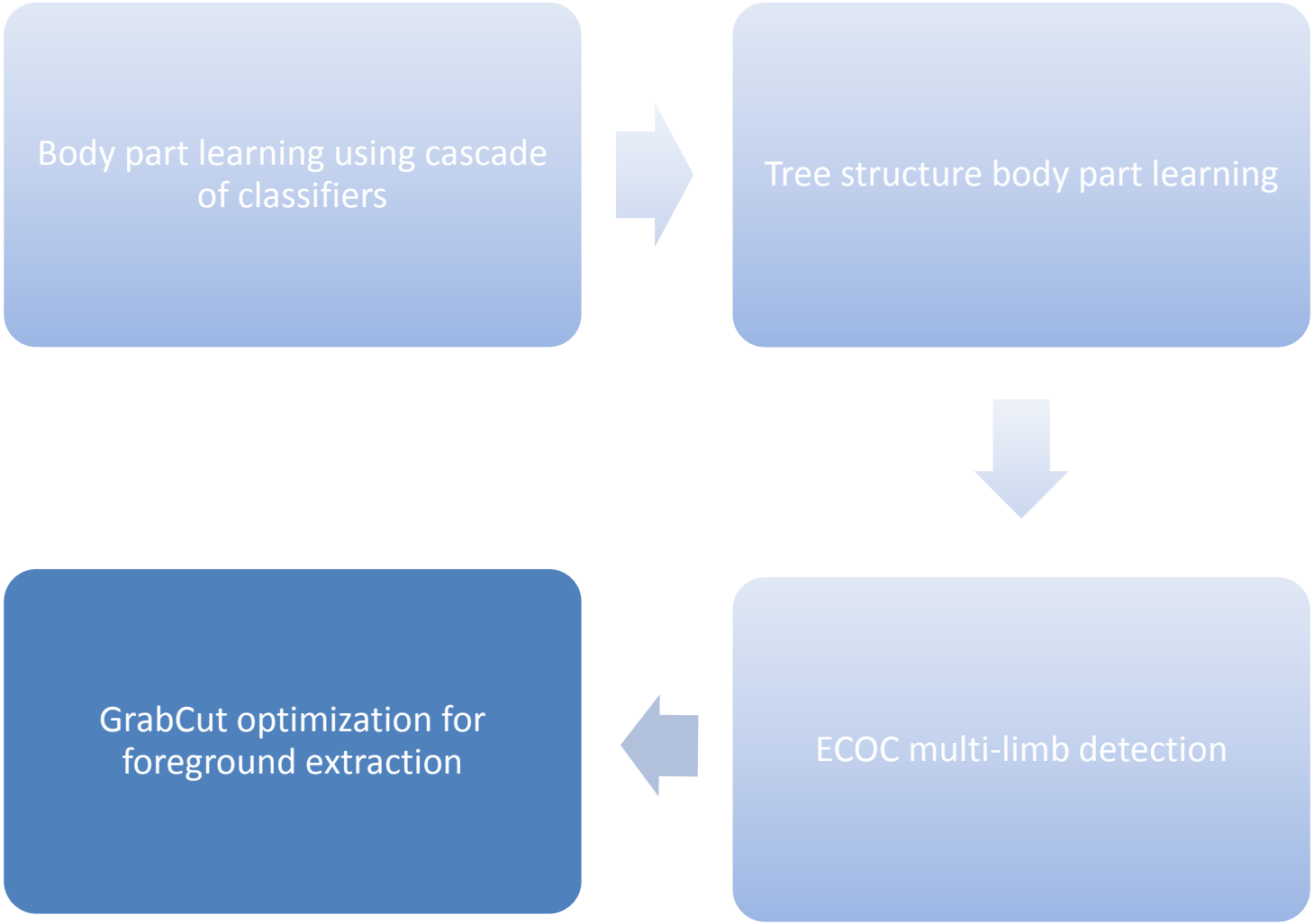
- Then, each cascade will give us its prediction and decoding ECOC step will be applied.
  - Loss-weighted decoding using cascade of classifier weights (takes into account classifier performances)

S. Escalera, D. Tax, O. Pujol, P. Radeva, R. Duin, Subclass problem-dependent design of error-correcting output codes, PAMI 30 (6) (2008) 1-14.
M. A. Bautista, S. Escalera, X. Baro, P. Radeva, J. Vitria, O. Pujol, Minimal design of error-correcting output codes, Pattern Recogn. Lett. 33 (6) (2012) 693-702.

ECOC multi-limb detection

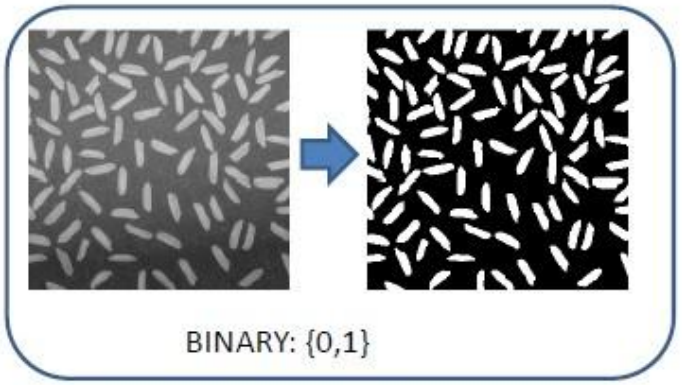
- A body-like probability map $P^{bl} \in [0,1]^{l \times w}$ is build




Body part learning using cascade of classifiers
Tree structure body part learning
ECOC multi-limb detection
GrabCut optimization for foreground extraction

GrabCut optimization for foreground extraction

- Image Segmentation == **Image labeling!**

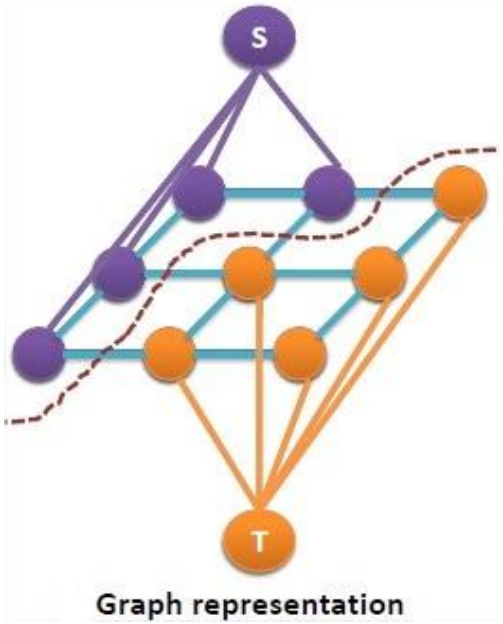


- Graph Cuts (Energy minimization)

$$\mathbf{E}(\alpha, \mathbf{u}, \theta, \mathbf{z}) = \mathbf{U}(\alpha, \mathbf{u}, \theta, \mathbf{z}) + \mathbf{V}(\alpha, \mathbf{z})$$

Unary Potential

Pair-wise Potential



Graph representation

Yuri Y. Boykov and Marie-Pierre Jolly, "Interactive Graph Cuts for Optimal Boundary & Region Segmentation of Objects in N-D Images", International Conference on Computer Vision, 2001

GrabCut optimization for foreground extraction

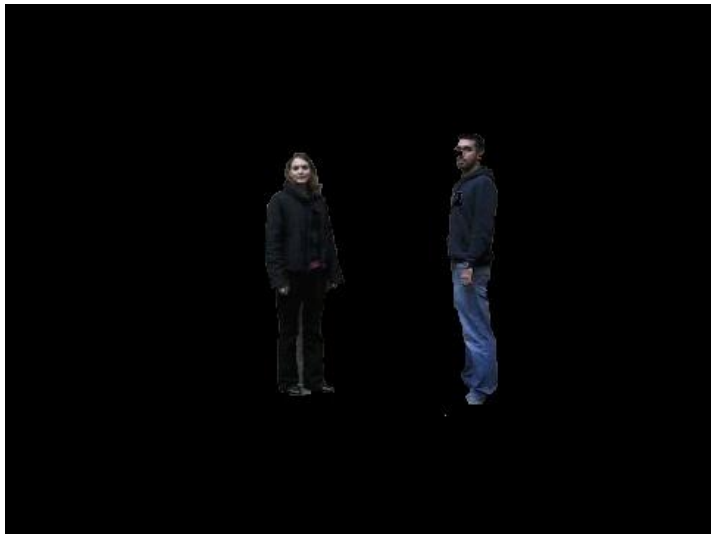
- User interaction by superimposed user input, background brush and so on.

- We propose to omit the classical interaction…

Yuri Y. Boykov and Marie-Pierre Jolly, “Interactive Graph Cuts for Optimal Boundary & Region Segmentation of Objects in N-D Images”, International Conference on Computer Vision, 2001

GrabCut optimization for foreground extraction

- Binary segmentation by means of background and foreground segmentation.
  - **Background**: Everything not related to body parts.
  - **Foreground**: Everything related to body parts.

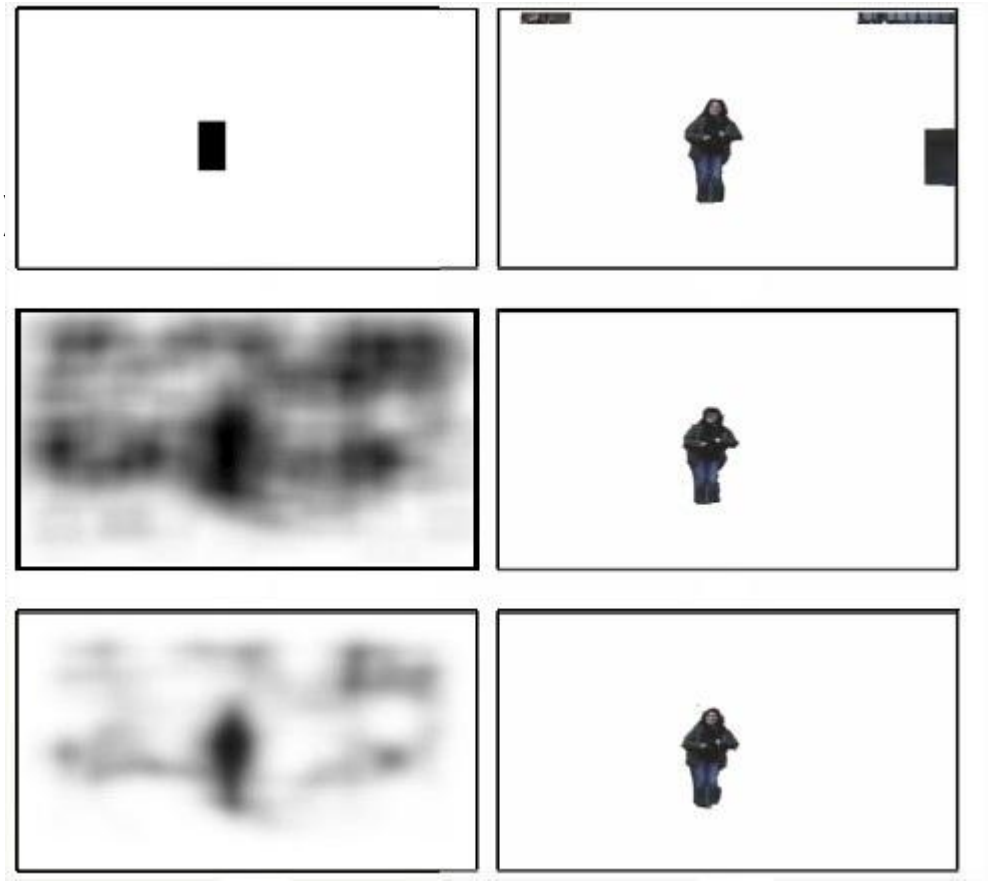


A. Hernandez-Vela, N. Zlateva, A. Marinov, M. Reyes, P. Radeva, D. Dimov, S. Escalera, Graph cuts optimization for multi-limb human segmentation in depth maps, in: CVPR, 2012, pp. 726-732.

- HuPBA-90(Human Pose Recovery and Behavior Analysis 90 images dataset) present a fully limb-labeled dataset:
  - Actors appear portraying a certain pose.
  - Point of view, lightning and background conditions remain invariant.
  - 14 limbs were manually tagged: Head, Torso, R-L Upper-arm, R-L Lower- arm, R-L Hand, R-L Upper-leg, R-L Lower-leg, R-L foot.
  - 90 images

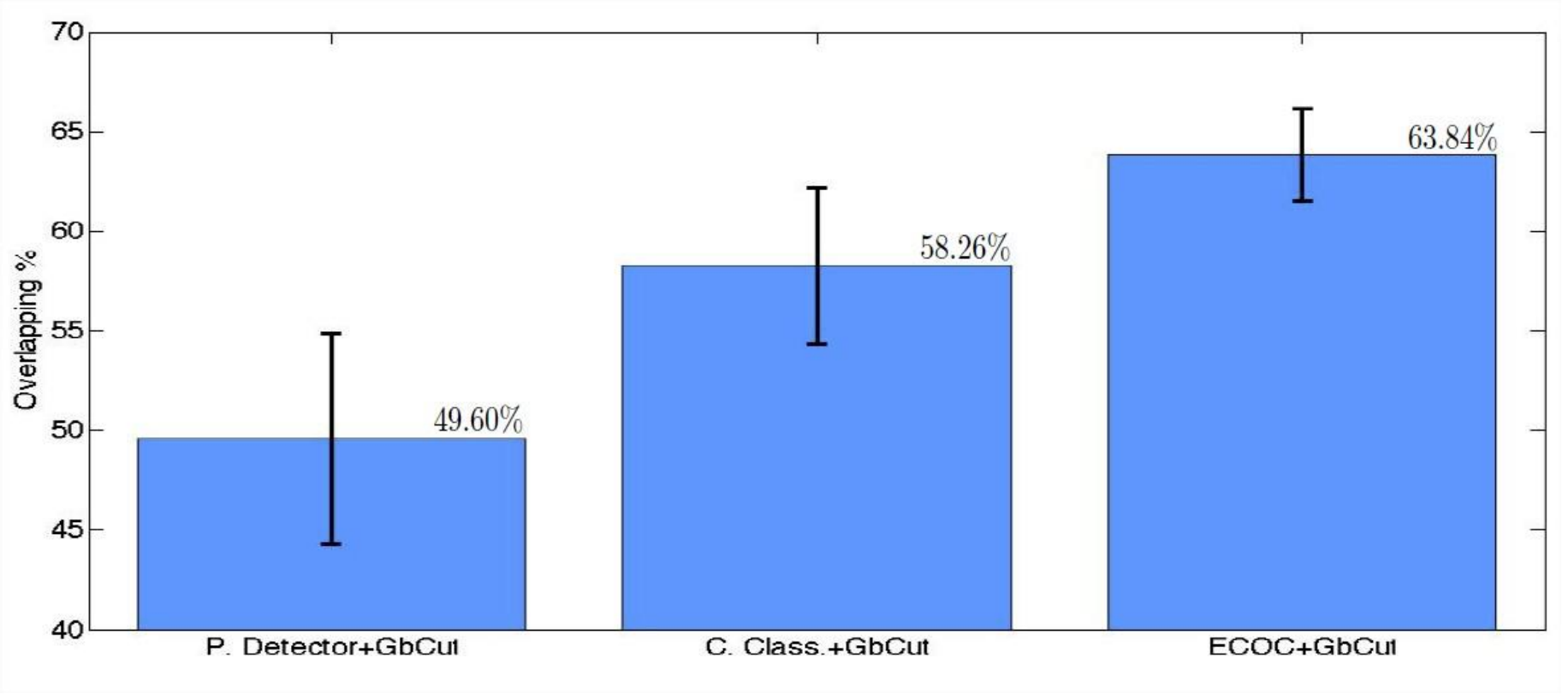
- 6 cascades of 8 levels each one were trained: 0.99 FP rate, 0.4 false alarm.
- Ten-fold applied to cascades.
- GrabCut: 5-fold for all methods.
- Segmentation is computed using overlapping with the Jaccard Index.

- We compare three methods:
  - Person Detector + GrabCut *
  - Cascade + GraphCut **
  - ECOC + GraphCut (Our proposal)



* N. Dalal, B. Triggs, Histograms of oriented gradients for human detection, in: CVPR, Vol. 1, 2005, pp. 886-893 vol. 1.
** P. Viola, M. Jones, Rapid object detection using a boosted cascade of simple features, in: CVPR, Vol. 1, 2001.

- Mean overlapping and standard deviation measures obtained on the 90 images of the dataset:



- We proposed a novel two-stage method for human segmentation in RGB images.
- First stage
  - Body parts trained in a body part tree structure architecture.
  - Cascade + ECOC.
  - Body-like probability map.
- Second stage
  - GraphCut segmentation procedure.
  - Novel limb-labeled dataset.
- Shows performance improvements in comparison to classical cascade of classifiers and human detector-based GraphCuts segmentation procedures.
- Robust results useful for posterior human pose and behavior analysis application.

Human Body Segmentation with Multi-limb Error-Correcting Output Codes Detection and Graph Cuts Optimization

Thank you!

Daniel Sánchez, Juan Carlos Ortega,
Miguel Ángel Bautista & Sergio Escalera