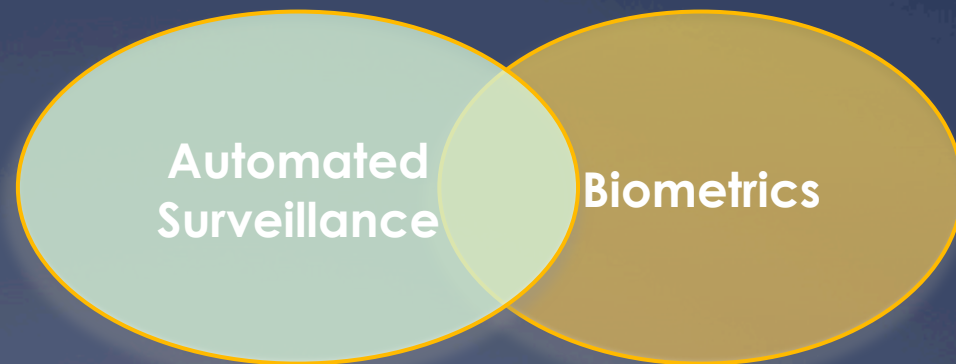


Hugo Proença

University of Beira Interior
Department of Computer Science
Covilhã, Portugal

Ocular Biometrics in Non-Cooperative Environments



1 - Biometrics Fundamentals

- 1.1 – Definitions
- 1.2 - Existing Systems
- 1.3 – Non-cooperative recognition

2 - Our approach: QUIS-CAMPI

- 2.1 - Why ocular region?
- 2.2 - Main ocular degradation factors
- 2.3 – Periocular recognition
- 2.4 - Hardware framework / scene overview
- 2.5 - Processing chain

3 - Low-level Vision Problems

- 3.1 - Background subtraction, motion detection
- 3.2 - Human Detection
- 3.3 – Tracking
- 3.4 - Path prediction
- 3.5 - Camera calibration
- 3.6 - Target selection
- 3.7 - Face detection

4 - Pose Estimation + Soft labels

- 4.1 - 3D Head Meshes
- 4.2 - Pose Hypotheses
- 4.3 - Joint Indexing
- 4.4 - SOM Labels

5 - Data Segmentation

- 5.1 - Periocular Segmentation
- 5.2 - Iris Segmentation

6 - Feature Extraction / Encoding, Recognition

- 6.1 - Ordinal Measures
- 6.2 - Uncorrelated Weak / Strong Experts

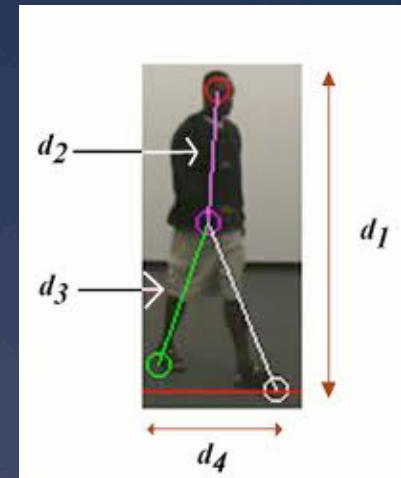
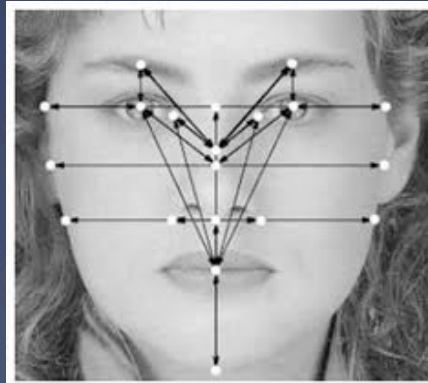
7 - State-of-the-Art Performance

- 7.1 - Multiple Systems
- 7.2 - Current challenges

8 – ICBRW16: Competition in ICB 16

* Biometrics

* “Biometrics consists of methods for uniquely recognizing humans based upon one or more intrinsic physical or behavioral traits”. **(Wikipedia)**



Emerging biometric technologies are providing more choice and increased accuracy

- * **Current state:** the deployed systems obtain **remarkable low error rates**, provided that **users cooperate with the data acquisition system and are willing to be recognized.**



NEC Face Recognition, VeriLook Surveillance: (indoor, exhaustive enrollment)



Iris-in-the-move, FaceIN: (indoor, relatively constrained)



Biometrics National IDs (heavily constrained)



* Main Goal

* To develop an automaton able to robustly recognize human beings, without requiring them any cooperation in the acquisition process.

* Perhaps contrary to usual belief, such automata are still confined to science Fiction.



(<http://www.prisonplanet.com/>)

* Book “*Nineteen Eighty-Four*” by George Orwell (1948) gave birth to the idea of “*Big Brother*”.

1 - Biometrics Fundamentals

- 1.1 – Definitions
- 1.2 - Existing Systems
- 1.3 – Non-cooperative recognition

2 - Our approach: QUIS-CAMPI

- 2.1 - Why ocular region?
- 2.2 - Main ocular degradation factors
- 2.3 – Periocular recognition
- 2.4 - Hardware framework / scene overview
- 2.5 - Processing chain

3 - Low-level Vision Problems

- 3.1 - Background subtraction, motion detection
- 3.2 - Human Detection
- 3.3 – Tracking
- 3.4 - Path prediction
- 3.5 - Camera calibration
- 3.6 - Target selection
- 3.7 - Face detection

4 - Pose Estimation + Soft labels

- 4.1 - 3D Head Meshes
- 4.2 - Pose Hypotheses
- 4.3 - Joint Indexing
- 4.4 - SOM Labels

5 - Data Segmentation

- 5.1 - Periocular Segmentation
- 5.2 - Iris Segmentation

6 - Feature Extraction / Encoding, Recognition

- 6.1 - Ordinal Measures
- 6.2 - Uncorrelated Weak / Strong Experts

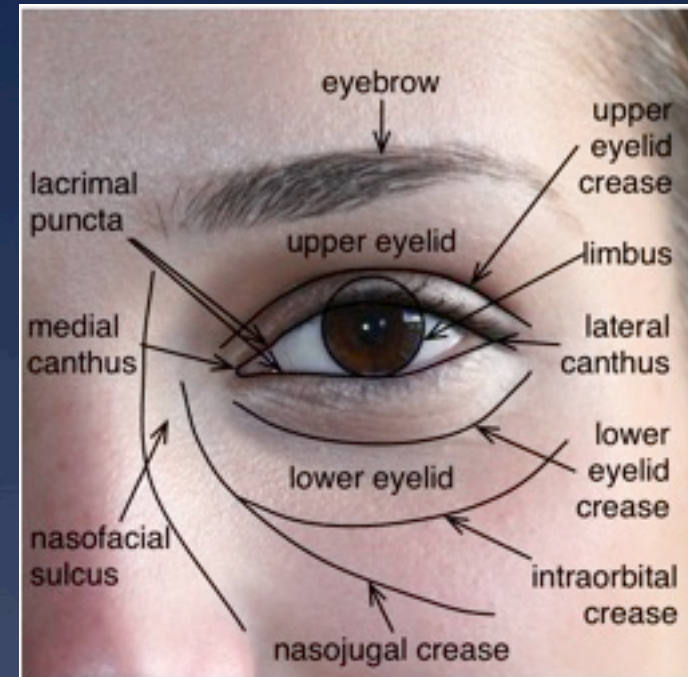
7 - State-of-the-Art Performance

- 7.1 - Multiple Systems
- 7.2 - Current challenges

8 – ICBRW16: Competition in ICB 16

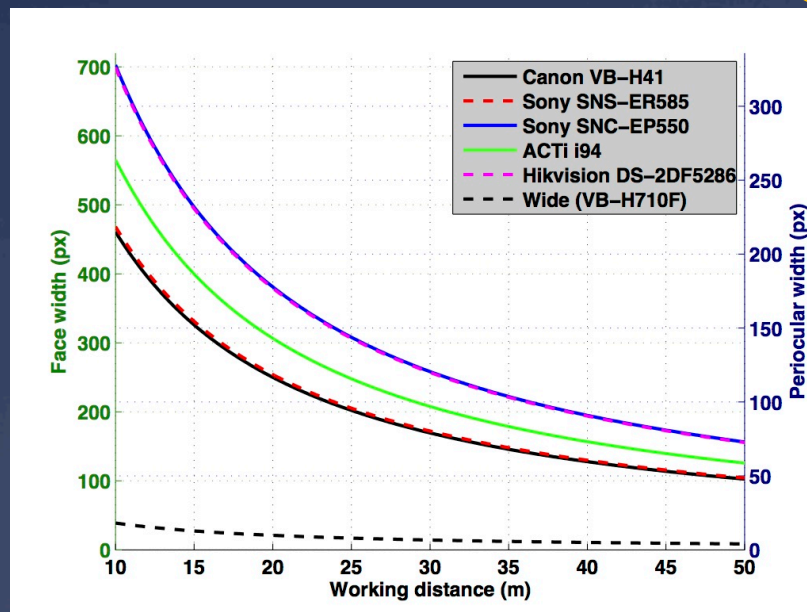
* Why use the ocular region?

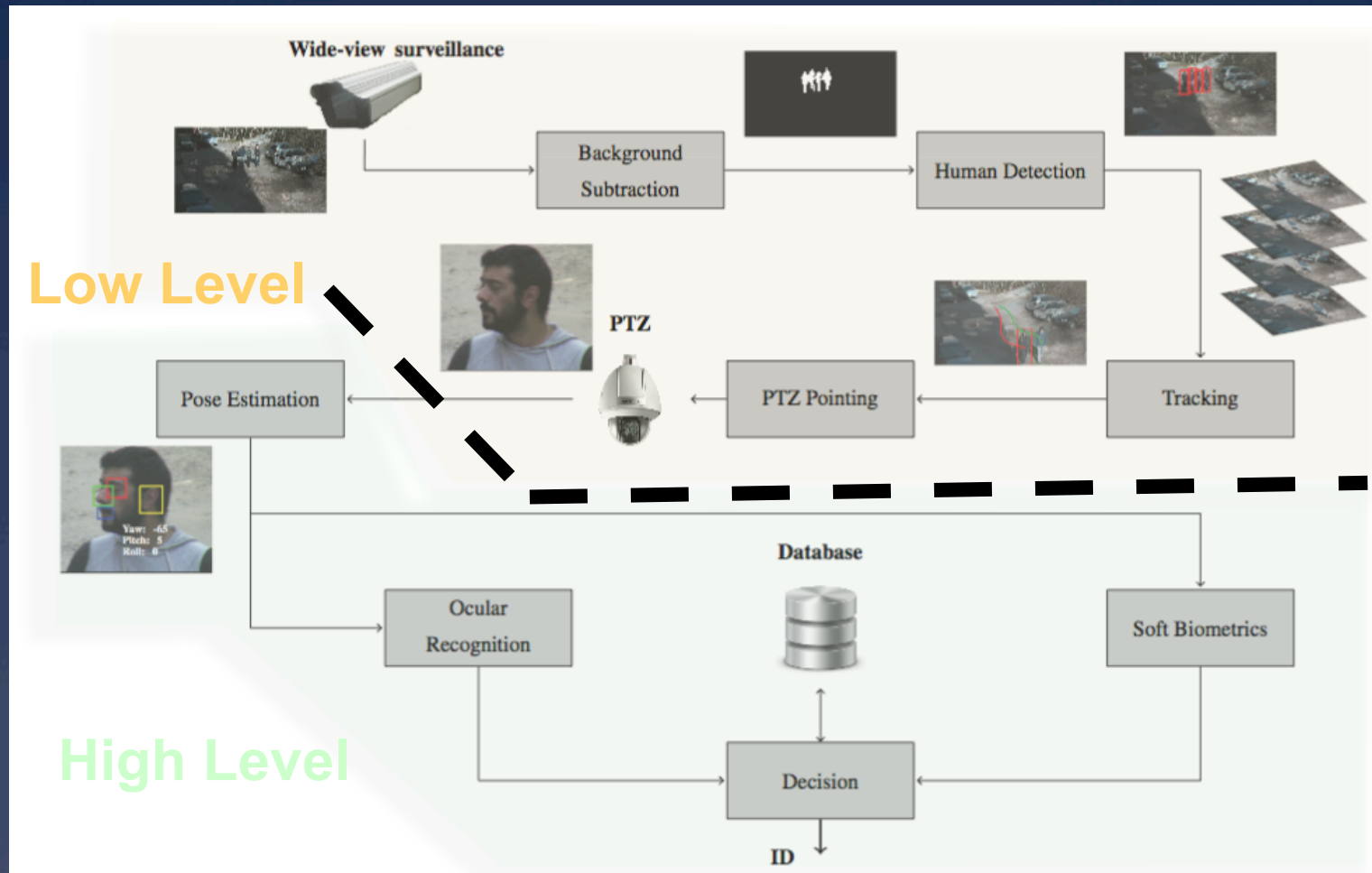
- The iris is a naturally protected internal organ;
- Supports **contactless** data acquisition;
- The iris uses **the lowest and middle-low frequency components** to perform recognition.
- The iris has a **regular shape** turns easier its detection and segmentation;
- The iris is **not deformable**, and does not change with expressions;
- The remaining components in the periocular region are less sensitive to expressions than the remaining face components
- The ocular region is relatively **planar**, which makes easy to compensate for off-angle acquisition.



* QUIS-CAMPI Recognition Framework

- * Outdoor environment
- * Typical surveillance Scenario
- * Two cameras:
 - * Wide view Canon VB-H710F;
 - * PTZ Hikvision DS-2DF5286





1 - Biometrics Fundamentals

- 1.1 – Definitions
- 1.2 - Existing Systems
- 1.3 – Non-cooperative recognition

2 - Our approach: QUIS-CAMPI

- 2.1 - Why ocular region?
- 2.2 - Main ocular degradation factors
- 2.3 – Periocular recognition
- 2.4 - Hardware framework / scene overview
- 2.5 - Processing chain

3 - Low-level Vision Problems

- 3.1 - Background subtraction, motion detection
- 3.2 - Human Detection
- 3.3 – Tracking
- 3.4 - Path prediction
- 3.5 - Camera calibration
- 3.6 - Target selection
- 3.7 - Face detection

4 - Pose Estimation + Soft labels

- 4.1 - 3D Head Meshes
- 4.2 - Pose Hypotheses
- 4.3 - Joint Indexing
- 4.4 - SOM Labels

5 - Data Segmentation

- 5.1 - Periocular Segmentation
- 5.2 - Iris Segmentation

6 - Feature Extraction / Encoding, Recognition

- 6.1 - Ordinal Measures
- 6.2 - Uncorrelated Weak / Strong Experts

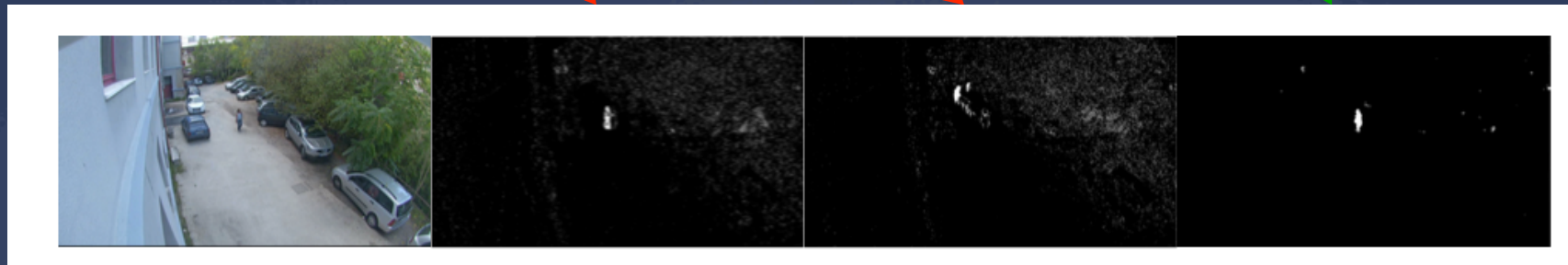
7 - State-of-the-Art Performance

- 7.1 - Multiple Systems
- 7.2 - Current challenges

8 – ICBRW16: Competition in ICB 16



- * Background subtraction (BS) methods aim to divide the scene into two disjoint parts: 1) **background**, that contains the static regions in the input data, which is usually disregarded from subsequent processing; and 2) **foreground**, that contains the regions-of-interest of all objects that the system should care about.
- * We evaluated the most relevant previously published methods for BS, divided into three families: 1) **basic**; 2) **Gaussian-based**; and 3) **machine-learning** based, ordered by their level of complexity.
- * Even though, BS in surveillance conditions is a far from solved problem.



Input Data

Frame Differencing [BS1]

Mixture of Gaussians [BS2]

SOBS [BS3]

[BS1] Y. Kameda and M. Minoh. A human motion estimation method using 3-successive video frames. In Proceedings of the International Conference on Virtual Systems, 1996.

[BS2] C. Stauffer and W. Grimson. Adaptive background mixture models for real-time tracking. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'99), vol. 2. pag. 246–252, 1999.

[BS3] L. Maddalena and A Petrosino. A Self-Organizing Approach to Background Subtraction for Visual Surveillance Applications. IEEE Transactions on Image Processing, vol. 17, no. 7, pag. 1168–1177, 2008.

Main Problems Observed: 1) Long-standing objects; 2) Low-frequency background changes; 3) Background Initialization

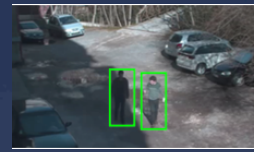
* Background subtraction In-the-wild has several data variation factors that increase the challenges in discriminating accurately between the foreground / background components:



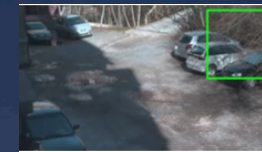
Global lighting variability



Scale



Shadows



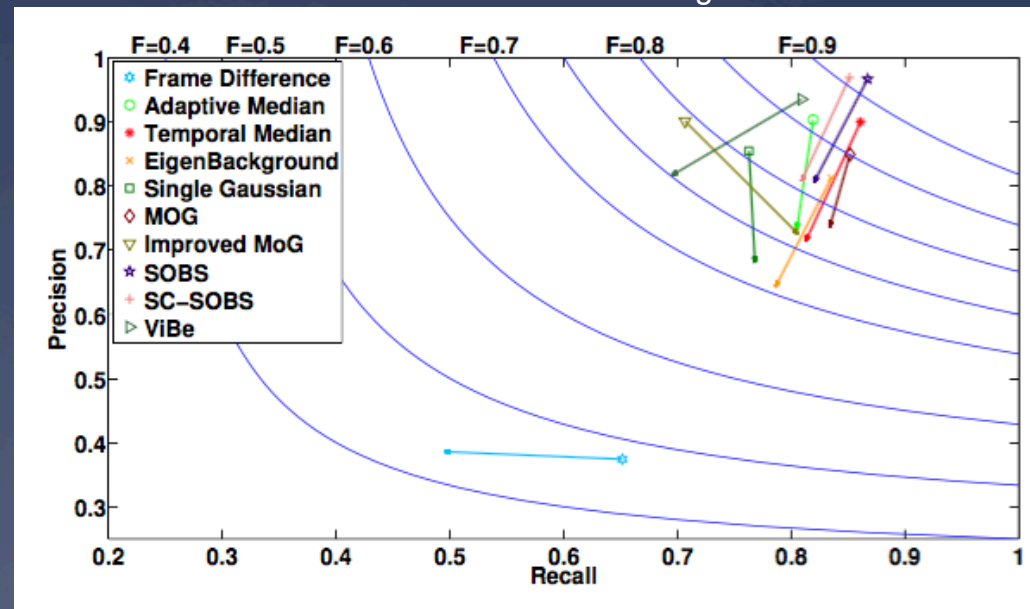
Moving /
complex
background



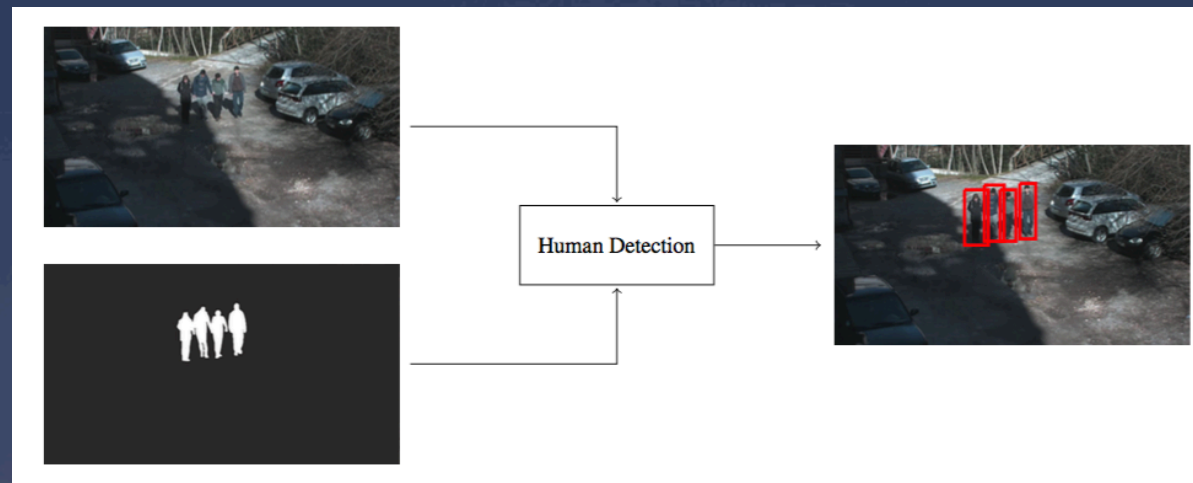
Reflections
(water)

Average decrease in BS performance, with respect to the “hardness” of environments (i.e., from constrained indoor → unconstrained outdoor environments)

SOBS was observed to be the **outperforming method**, both for constrained and unconstrained scenarios.



- * Two families of Human Detection (HD) methods were evaluated: 1) **holistic** methods, where the whole body is searched in the image; and 2) **part-based** methods, where each part of the body is detected independently and information is further fused for consistency purposes.
- * Usually, holistic methods learn a discriminative model, being well represented by the popular Viola and Jones' method, adapted to detect humans using motion patterns [H1].
- * Regarding part-based methods, Mikolajczyk et al. [H2] was considered a representative method, using a probabilistic model to assemble all parts of the body, each one detected in a coarse-to-fine strategy.

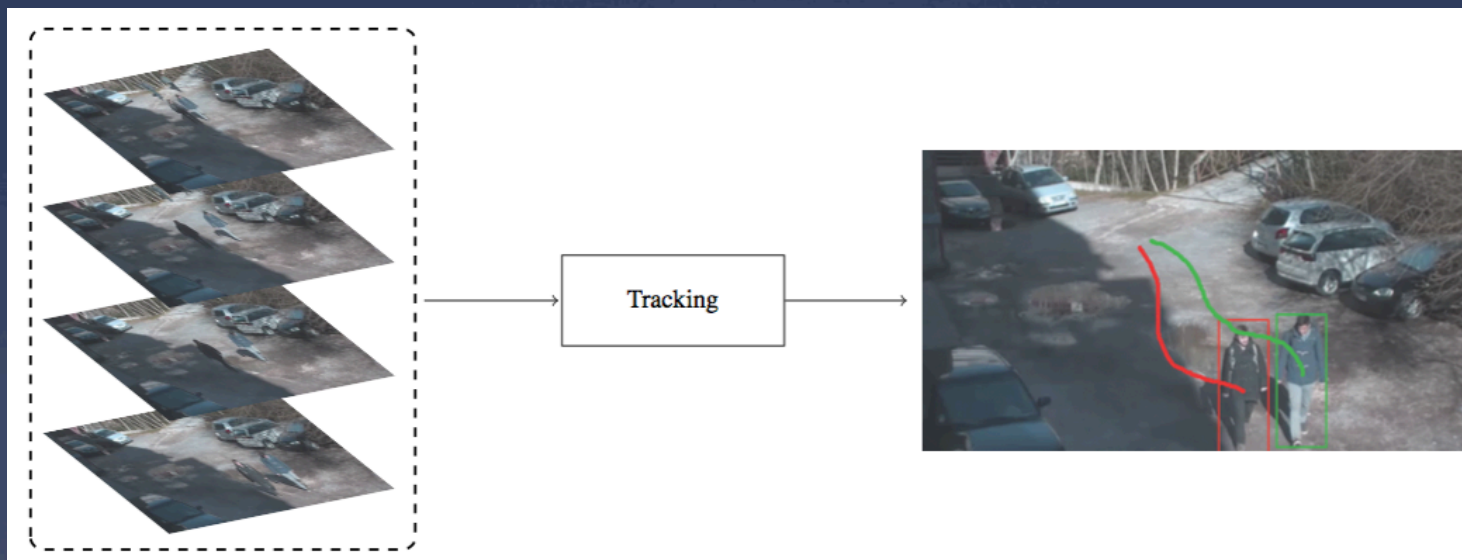


[H1] P. Viola, M. Jones and D. Snow. Detecting pedestrians using patterns of motion and appearance. in Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV'03), vol. 2, pag. 734–741, 2003.

[H2] K. Mikolajczyk, C. Schmid and A. Zisserman. Human Detection Based on a Probabilistic Assembly of Robust Part Detectors. In Proceedings of the European Conference on Computer Vision (ECCV 2004), LNCS, vol. 3021, pag. 69–82, 2004.

Main Problems Observed: 1) Superimposed / occluded regions; 2) Deformable shape; 3) Local lighting variations; 4) shadows.

- * Given an initial detection of one object, the tracking phase determines the positions of that object in the subsequent frames, having in our case two major goals:
 - 1) by perceiving the object path, accurate predictions of the location of the object in forthcoming frames can be made, allowing to timely point the PTZ device for a specific position;
 - 2) once the high-resolution data of a subject is acquired, that element can be ignored of any subsequent processing.



Main Problems Observed: 1) Real-time multiple object tracking; 2) Objects cross-paths; 3) Local changes in image features (brightness, entropy, . . .)

- * Having defined the trajectory of one subject in the scene, the task is to point the PTZ device to that region of the space, to get the high-resolution image.
- * We should take into account that the mechanical operations (rotational speed + mechanical zoom) introduce a delay in the system (approx 750ms in our hardware).
- * If cannot simply point the PTZ device to the region in the scene where the face was detected (wide-view data).
- * Instead, we record the set of positions of the subject in previous frames, to infer its position 750 ms. ahead.



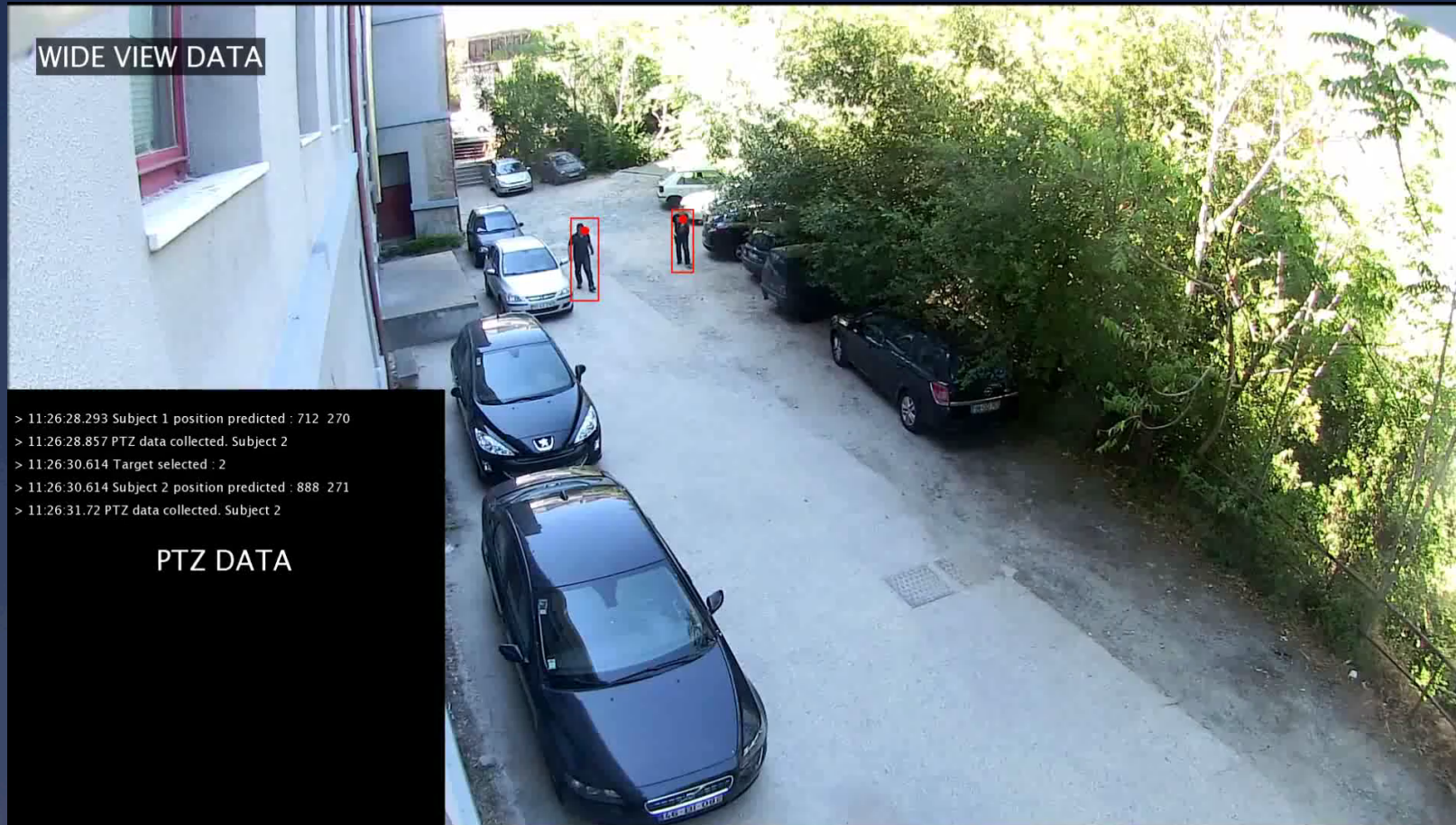
Without prediction ✗



With prediction ✓



* Linear models do not perform well to predict human walk paths.



 Linear prediction model
 Generalized regression neural network

* The relation between the 2D coordinates of images from the wide-view and PTZ devices is an ill-posed problem.

Height 0

Height 175 cm.

Wide-view projection: (1188, 840)

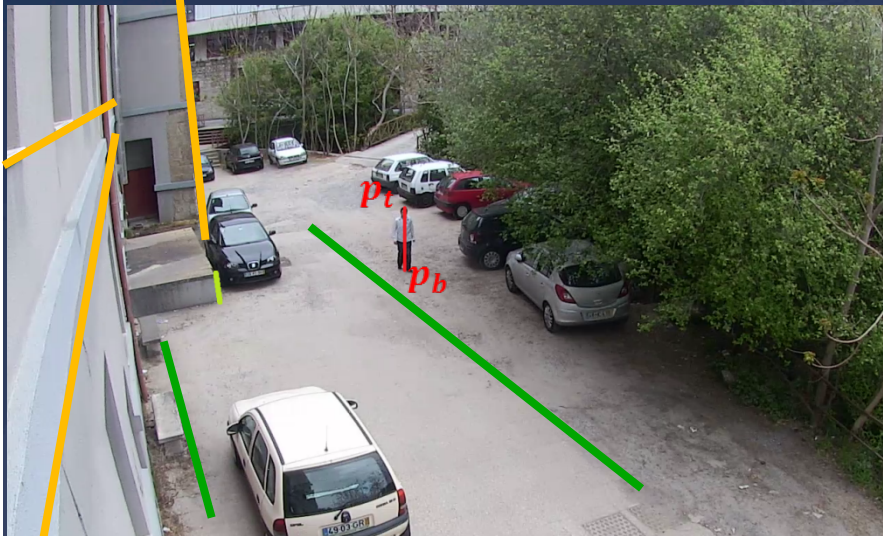
In practice, there are multiple positions in the PTZ coordinate system that correspond to the same pixel in the wide-view image






To solve such ambiguity, 3D information is required. Stereo reconstruction is a classical solution, which will demand one additional wide-view camera

* Our current solution uses the concept of “single-view metrology”

* Criminisi et al. IJCV 40(2), 123-148, 2000.



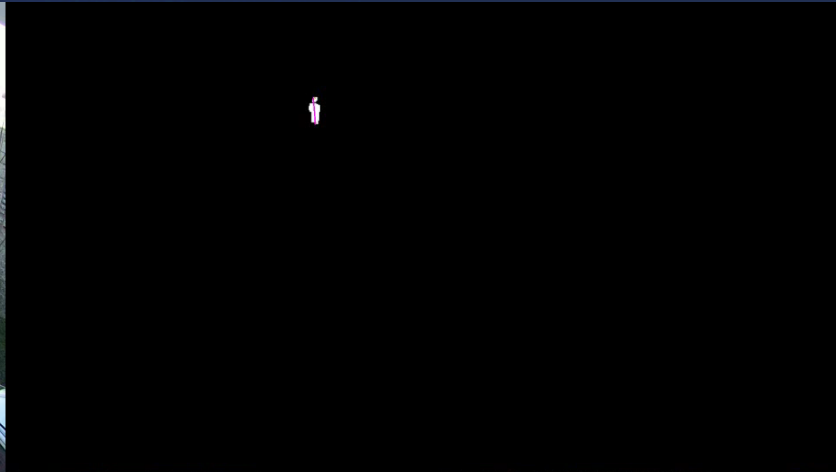
-  Reference height
-  Human height
-  Horizontal / vertical parallel lines

1. The intersection of two pairs of lines in the reference plane (—) enable to obtain two **vanishing points** (v_x, v_y).
2. The line passing through v_x, v_y is the **vanishing line** “l”
3. Similarly, the intersection of two vertical lines gives the **vertical vanishing point** v_z .
4. Given l, v_z , the head and feet positions in the scene (p_t and p_b), the height of the subject is given by:

$$h = -\frac{\|\mathbf{p}_b \times \mathbf{p}_t\|}{\alpha(\mathbf{l}, \mathbf{p}_b)\|\mathbf{v}_z \times \mathbf{p}_t\|}$$

with $\alpha = -\|\mathbf{p}_{rb} \times \mathbf{p}_{rt}\| / (h_r(\mathbf{l}, \mathbf{p}_{rb})\|\mathbf{v}_z \times \mathbf{p}_{rt}\|)$
 being p_{rb}, p_{rt} the bottom / top positions of a reference object in the scene (with height h_r)

- * If the height estimation is repeated across several frames, we use the average height as an estimate of height, which was proved to be quite close to the true subjects height.



CONSUMED TIME :67 ms

HEIGHT :159,954827 cm



J. Neves, J. Moreno, S. Barra, H. Proença. Acquiring High-resolution Face Images in Outdoor Environments: A master-slave Calibration Algorithm. In Proceedings of the IEEE Seventh International Conference on Biometrics: Theory, Applications and Systems – BTAS 2015, Washington DC, U.S.A., September 8-11, 2015, ISBN: 978-1-4244-1597-7.

- * Assuming the pin-hole model, the projective transformation of 3D scene points into the 2D image plane is given by:

$$\lambda \begin{pmatrix} x_t \\ y_t \\ 1 \end{pmatrix} = \underbrace{\mathbf{K}[\mathbf{R}|\mathbf{T}]}_{\mathbf{P}} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

with λ being the scale factor, \mathbf{K} the intrinsic matrix, and \mathbf{R} , \mathbf{T} the extrinsic matrices.

- * Having one point in the wide-view image: (x_t, y_t) = the (x, y) coordinates of the head center of mass, the above equality yields an under-determined system, i.e., infinite 3D positions in the scene are projected into x_t, y_t .
- * However, if the Z component (subject height) is also known, the equation above reduces to:

$$\lambda \begin{pmatrix} x_t \\ y_t \\ 1 \end{pmatrix} = [\mathbf{p}_1 \quad \mathbf{p}_2 \quad h\mathbf{p}_3 + \mathbf{p}_4] \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix}$$

being \mathbf{p}_i the column vectors of “ \mathbf{P} ”.

- * Let $(0, \rho \sin \theta_t, \rho \sin \theta_t)$ be the center of rotation of the PTZ device.
- * ρ is the displacement between the mechanical rotation axis and the image plane (approximated by camera focal distance “f”)
- * Having the (X, Y, Z) coordinates of a point in the world coordinate system, the corresponding location in the PTZ coordinate system is given by:

$$\begin{pmatrix} X_p \\ Y_p \\ Z_p \end{pmatrix} = [\mathbf{R} | \mathbf{T}] \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

- * The corrected coordinates are given by:

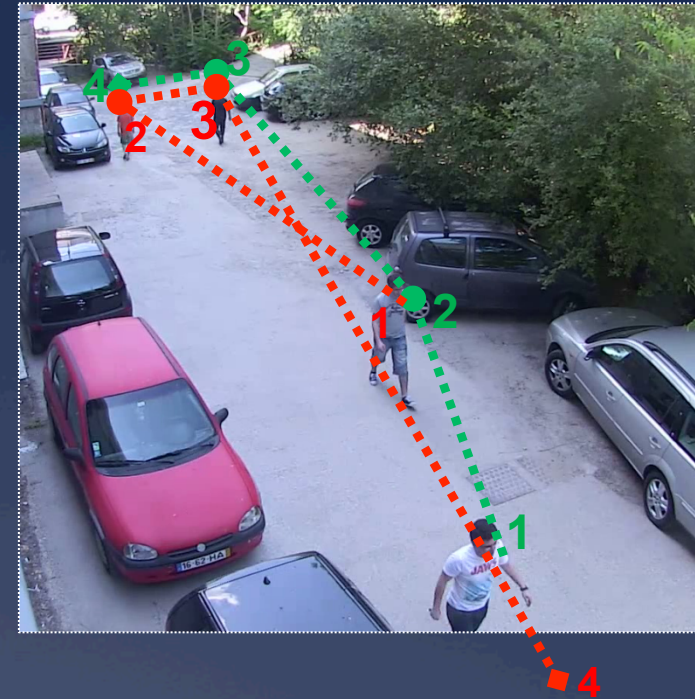
$$\begin{pmatrix} X'_p \\ Y'_p \\ Z'_p \end{pmatrix} = \begin{pmatrix} X_p \\ Y_p - \rho \sin \theta_t \\ Z_p + \cos \theta_t \end{pmatrix}$$

- * Finally, the pan (θ_{pan}) and tilt (θ_{tilt}) angles applied to the PTZ device are:

$$\theta_{\text{pan}} = \arctan \left(\frac{X'_p}{Z'_p} \right)$$

$$\theta_{\text{tilt}} = \arcsin \left(\frac{Y'_p}{\sqrt{(X'_p)^2 + (Y'_p)^2 + (Z'_p)^2}} \right)$$

- * If we consider multiple subjects in a scene, the order by which the system attempts to acquire high resolution images from each subject is also a problem.
- * The 3D position of each subject in the scene should be considered...
- * Together with each subject path, pose and velocity...
 - * *“Frontal subjects almost leaving the scene field-of-view should be privileged”*
- * And the number of data acquisition attempts performed for each target...
 - * *“Subjects with less number of acquisition attempts should be privileged”*
- * There are some basic strategies for this task:
 - * *Random, first-seen, last-seen,...*



Random



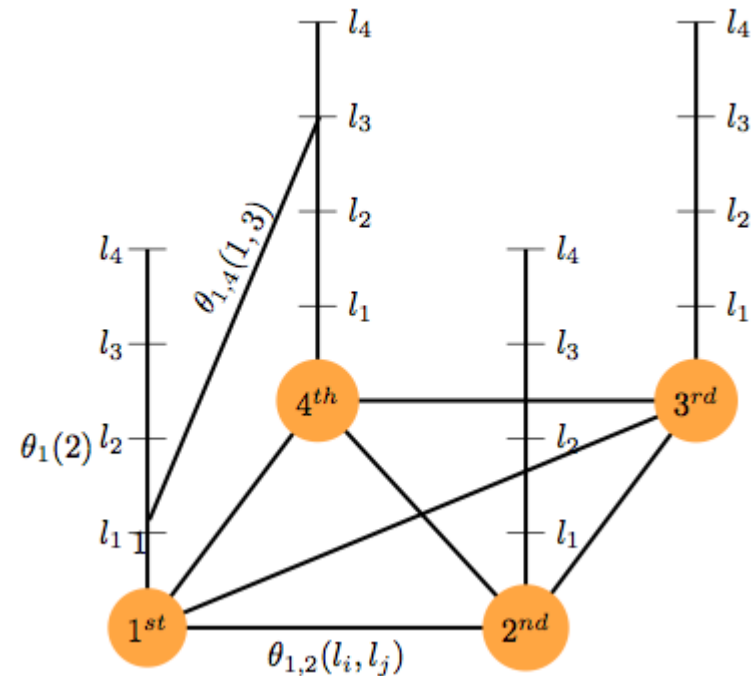
Time : 2.2 s
Targets : 3

Our Solution



Time : 1.9 s
Targets : 4

- * We model the dynamic scheduling of PTZ cameras as a graph-based problem whose solution can be approximated by the energy minimization of a MRF;
- * Detecting “N” subjects in the scene, we build a fully connected MRF of “N” vertices, each one with “N” labels:
- * Each vertex can be assigned to “N” different labels, corresponding to the “N” targets in the scene.
- * This structure allows to determine the order that each target will be observed by taking into account both the temporal constraints (vertex information) and the transition costs (pairwise relations between vertices)



- * The energy of a configuration “ l ” of the MRF is the sum of the unary $\theta_i(l_u)$ and pairwise $\theta_{ij}(l_u, l_v)$ costs.

$$E(l) = \sum_{i \in V} \theta_i(l_u) + \sum_{(i,j) \in E} \theta_{i,j}(l_u, l_v).$$

- * Under this formulation, finding the best tour is equivalent to infer the random variables in the MRF by minimizing its energy:

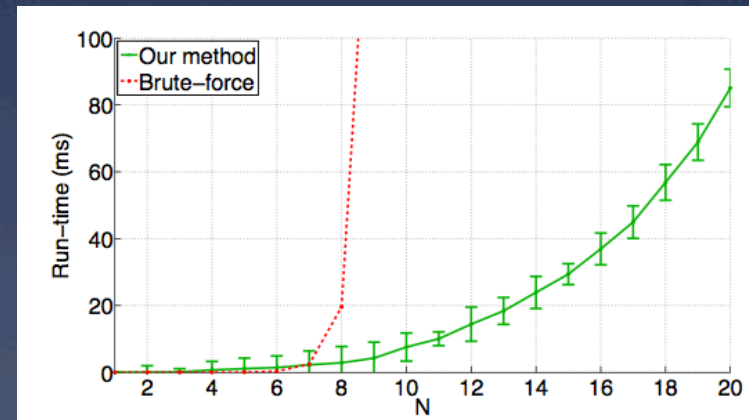
$$\hat{l} = \arg \min_l E(l),$$

- * In practice, it is impossible to evaluate all possibilities for a large number of subjects in the scene (over 6/7).

* The Loopy Belief Propagation was used in energy minimization. Even though it is not guaranteed to converge to global minimums on loopy non sub-modular graphs (such as our MRF), we observed that the algorithm provides in general good approximations.

- * We consider the following information:

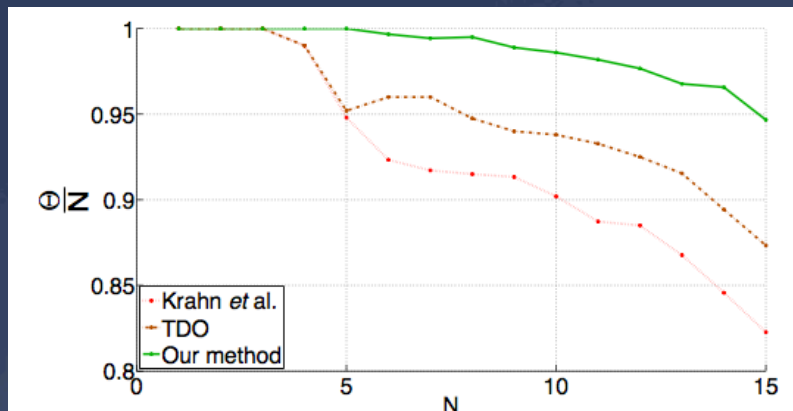
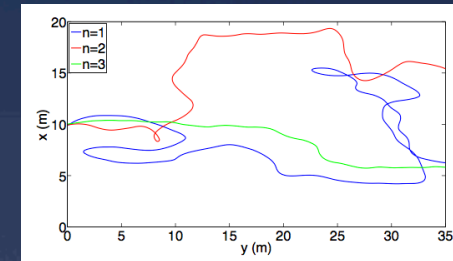
- * $p_u(t) = (x_u(t); y_u(t); z_u(t))$: the 3D position of the u^{th} target at time t ;
- * $\alpha(p_u)$: the pan angle required to move the PTZ to p_u ;
- * $\beta(p_u)$: the tilt angle required to move the PTZ to p_u ;
- * $\Lambda_u(t)$: expected time to target p_u leave the scene;
- * τ : average time required to acquire a target.



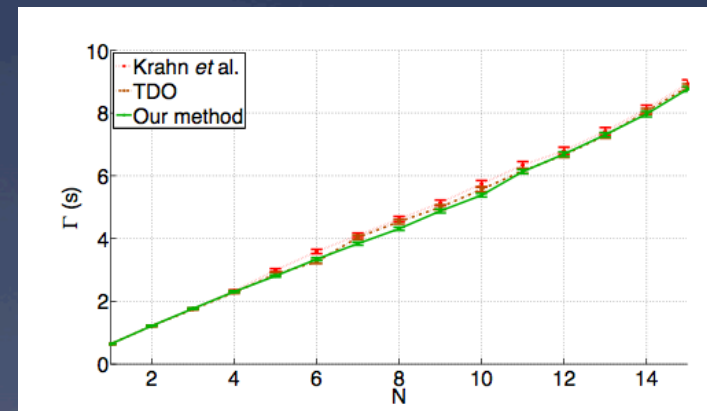
Average Required Time to Observe N subjects

* Using randomly generated simulated paths, and comparing the effectiveness of this strategy against the state-of-the-art [T1] [T2], we conclude that the proposed solution [T3] offers consistent advantages, particularly by augmenting the probability of observing a large number of subjects.

* Results were validated both for linear (more realistic) and chaotic paths.



Average Observation Rate (proportion of subjects visited)



Average Required Time to Observe N subjects

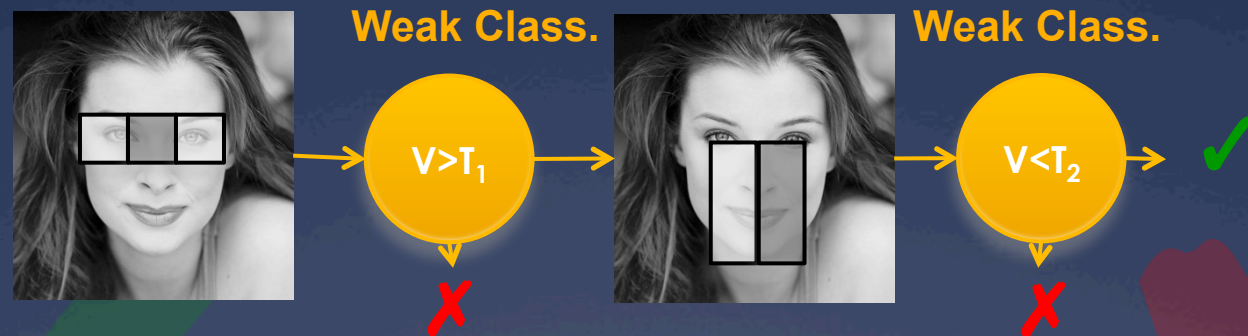
[T1] (TDO) A. D. Bimbo and F. Pernici. Towards on-line saccade planning for high-resolution image sensing. Pattern Recognition Letters, 27:1826 – 1834, 2006.

[T2] (Krahn et al.) N. Krahnstoever, T. Yu, S.-N. Lim, K. Patwardhan, and P. Tu. Collaborative Real-Time Control of Active Cameras in Large Scale Surveillance Systems. In Workshop on Multi-camera and Multi-modal Sensor Fusion Algorithms and Applications - M2SFA2, 2008.

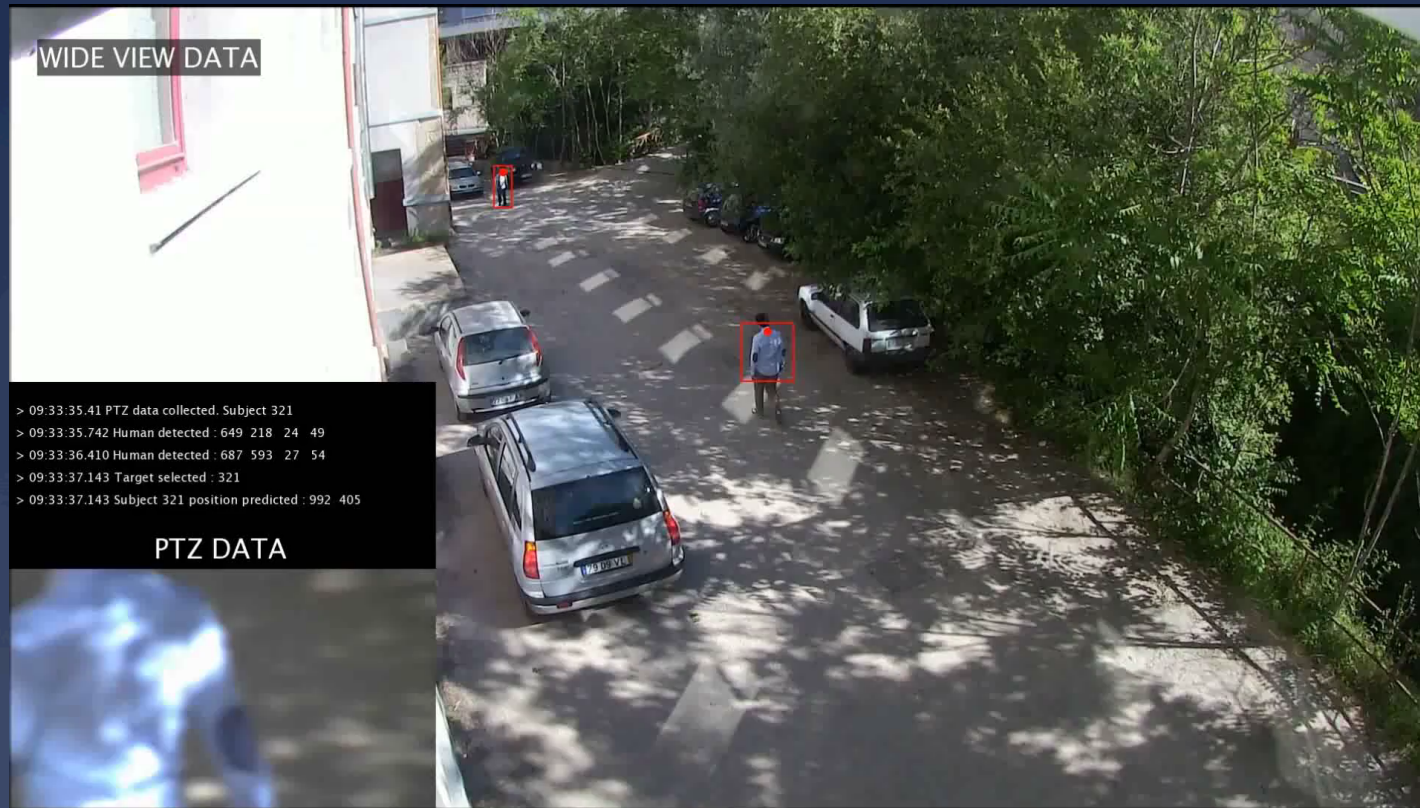
[T3] João C. Neves, Hugo Proença. Dynamic Camera Scheduling for Visual Surveillance in Crowded Scenes using Markov Random Fields. In Proceedings of the 12th IEEE International Conference on Advanced Video and Signal based Surveillance - AVSS 2015, Karlsruhe, Germany, August 25-28, 2015.

* Face Detection: Cascade of Weak Classifiers.

- * Weak classifiers are based in a single feature, that analyzes the relative intensity of two image patches
- * In order to be classified as a “Face”, an instance must pass over “N” weak classifiers.

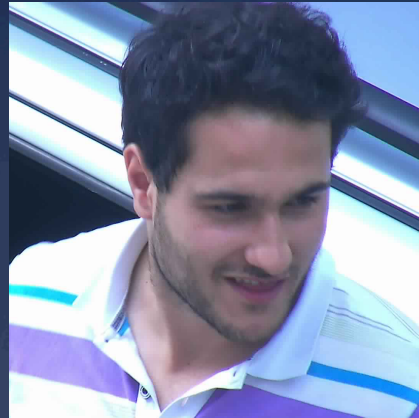


- * This video illustrates the current state of our data acquisition framework, able to acquire data from the ocular region of moving subjects, located over 40m. away from the cameras.

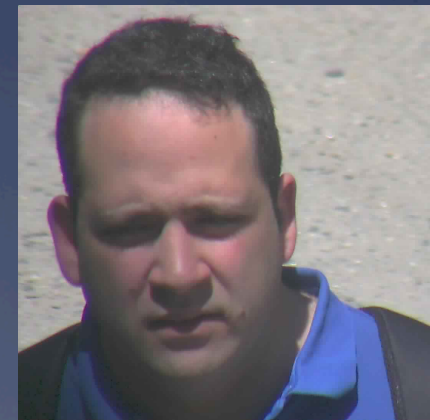
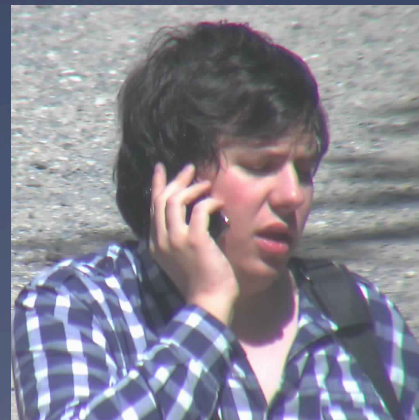


- * Avoiding moving background (tree leaves)
- * Avoiding non-human moving objects (cars, ...)
- * Selecting automatically the best targets

- * Using the afore described strategy, it is possible to get sharp images of the ocular region, taken from moving subjects, up to 45 meters away from the acquisition system.



320 x 115 px.



1 - Biometrics Fundamentals

- 1.1 – Definitions
- 1.2 - Existing Systems
- 1.3 – Non-cooperative recognition

2 - Our approach: QUIS-CAMPI

- 2.1 - Why ocular region?
- 2.2 - Main ocular degradation factors
- 2.3 – Periocular recognition
- 2.4 - Hardware framework / scene overview
- 2.5 - Processing chain

3 - Low-level Vision Problems

- 3.1 - Background subtraction, motion detection
- 3.2 - Human Detection
- 3.3 – Tracking
- 3.4 - Path prediction
- 3.5 - Camera calibration
- 3.6 - Target selection
- 3.7 - Face detection

4 - Pose Estimation + Soft labels

4.1 - 3D Head Meshes

4.2 - Pose Hypotheses

4.3 - Joint Indexing

4.4 - SOM Labels

5 - Data Segmentation

5.1 - Periocular Segmentation

5.2 - Iris Segmentation

6 - Feature Extraction / Encoding, Recognition

6.1 - Ordinal Measures

6.2 - Uncorrelated Weak / Strong Experts

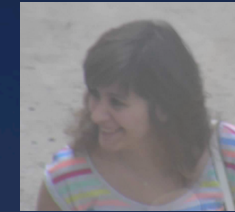
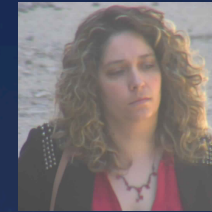
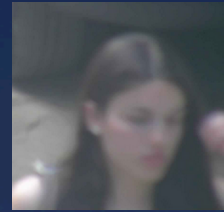
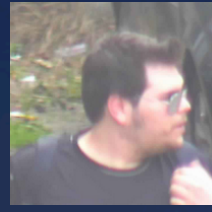
7 - State-of-the-Art Performance

7.1 - Multiple Systems

7.2 - Current challenges

8 – ICBRW16: Competition in ICB 16





* Pose estimation / soft labeling phases are based in the 3D morphology of the human head.

- * The problem in most soft biometric labels is that they are **relatively easy to forge** by anyone not willing to be recognized.
- * Anyone can dramatically change the hair style / length and the eyes color.

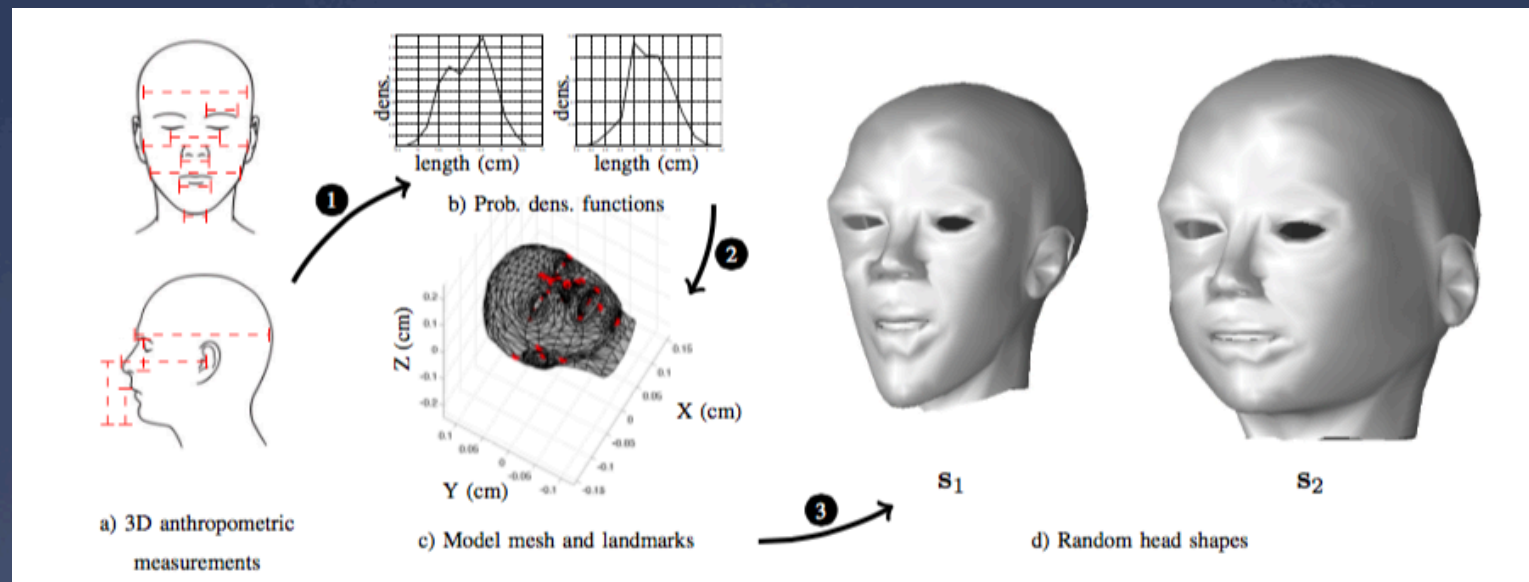
2- Propose an algorithm to find the most likely joint pose / head shape hypotheses for a query.



(http://www.prettyyourworld.com/ColorRevival_Review-ezine-issue-006.html)

1- We generate a set of pose and head shape hypotheses.

- * Based on head anthropometric surveys, we generate a large set of random head shapes.
- * Deform one “basis” head mesh, according to a vector of random deformations:
 - * Based on a set of 3D vertices (x_i) and their normal to the surface (n_i), we apply a displacement α_{ij} , such that the length between two vertices is approximately l_{ij}



- * Displacements are found by solving a constrained quadratic system.
- * Each desired length between two vertices (l_{ij}) is given by:

$$\|\mathbf{n}_{ij}^T \mathbf{x}_{ij} \alpha_{ij}^2 + 2\mathbf{x}_{ij}^T \mathbf{n}_{ij} \alpha_{ij} + \mathbf{x}_{ij}^T \mathbf{x}_{ij} - l_{ij}\|_2 = 0,$$

$$\|[\mathbf{n}_{ij}^T \mathbf{x}_{ij}, 2\mathbf{x}_{ij}^T \mathbf{n}_{ij}, \mathbf{x}_{ij}^T \mathbf{x}_{ij}] [\alpha_{ij}^2, \alpha_{ij}, 1]^T - l_{ij}\|_2 = 0,$$

(Matrix form)

- * Considering $\mathbf{c}_{ij} = [\mathbf{n}_{ij}^T \mathbf{x}_{ij}, 2\mathbf{x}_{ij}^T \mathbf{n}_{ij}, \mathbf{x}_{ij}^T \mathbf{x}_{ij}]$, $\alpha_{ij} = [\alpha_{ij}, \sqrt{\alpha_{ij}}, 1]^T$ we have:

$$\mathbf{C} = \left. \begin{bmatrix} \mathbf{c}_{ij} & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{c}_{i'j'} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{c}_{i''j''} \end{bmatrix} \right\} n \times 3n$$

$$\alpha = \left. \begin{bmatrix} \alpha_{ij} \\ \alpha_{i'j'} \\ \vdots \\ \alpha_{i''j''} \end{bmatrix} \right\} 3n \times 1, \mathbf{l} = \left. \begin{bmatrix} l_{ij} \\ l_{i'j'} \\ \vdots \\ l_{i''j''} \end{bmatrix} \right\} n \times 1,$$

- * The deformation model is given by:

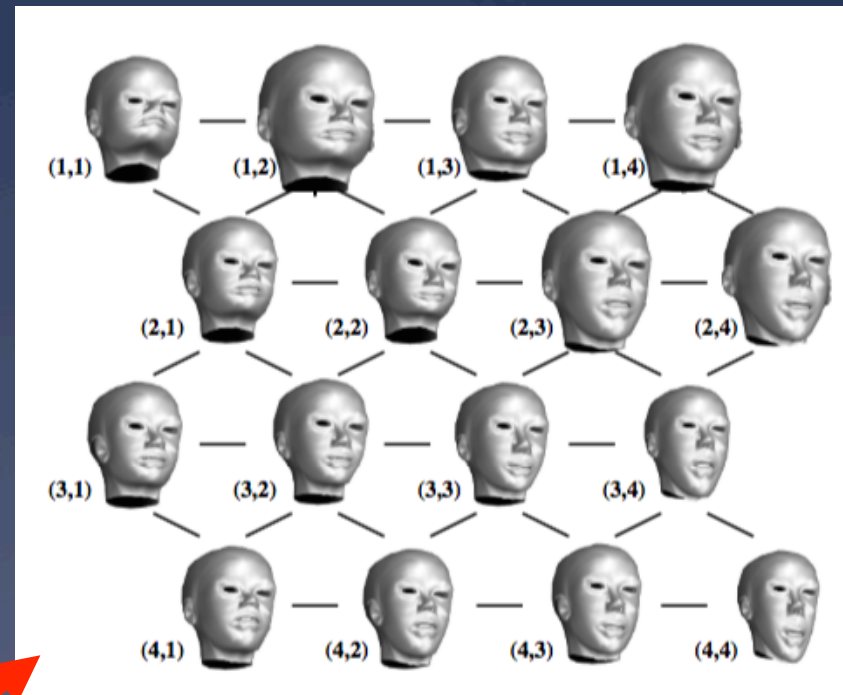
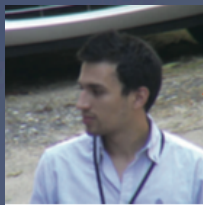
$$\hat{\alpha} = \arg \min_{\alpha} (\mathbf{C}\alpha - \mathbf{l})^T (\mathbf{C}\alpha - \mathbf{l}),$$

$$\text{s.t. } \|[\alpha_{ij}, \dots, \alpha_{i''j''}]\|_{\infty} \leq \kappa_1,$$

- * This stochastic enables to generate a large number of random head shapes, each one regarded as a point in a high-dimensional space (concatenation of the 3D coordinates of its vertices).
- * These elements are the input of a SOM, that provide a set of prototypes topologically ordered (**key concept: head shape alikeness**).

Representation of the 3D head prototypes resulting of a 4×4 SOM. Note the similarity in size / shape between adjacent elements, rooted in the preservation of the topological properties of the input space that this kind of maps offers.

Goal: To associate a query image of poor quality to one pose / head shape joint hypothesis, obtaining a simultaneous estimate of head pose and one soft label (the head shape)



4 – Pose Estimation + Soft labels

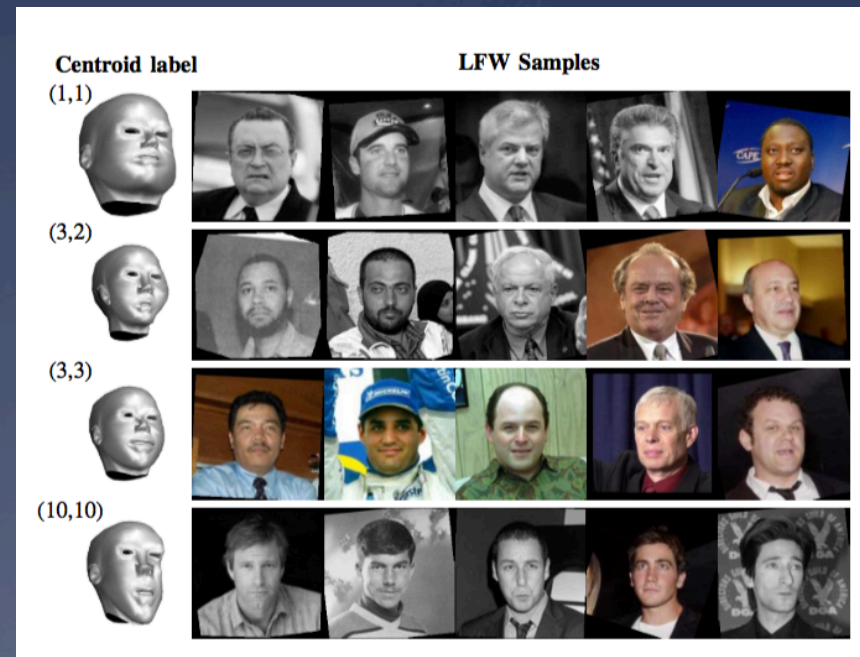
- * For every combination of head shape prototype / pose hypotheses (typically 25 (head) x 1,000 (pose) = 25,000), we obtain the projection of the all head vertices (\mathbf{x}) according to pose ($\mathbf{p}=\{\mathbf{R}, \mathbf{t}\}$, \mathbf{A} : intrinsic matrix):

$$f_{w \rightarrow i}(\mathbf{x}, \mathbf{p}) = \frac{1}{v} \mathbf{A}[\mathbf{R}|\mathbf{t}] \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix}$$
- * Next, having a query represented by a set of head landmarks, we rank the joint hypotheses, according to their likelihood of matching the query.
- * Iteratively, we consider a joint pose / head shape hypothesis and (neglecting local minima), use a convex (unconstrained) optimization algorithm to obtain a solution.
 - * If the solution is not considered acceptable, the next joint pose / head shape hypothesis is considered.
 - * Otherwise, the process stops.

Soft labeling results



Pose estimation results



- * The pose estimation + soft labeling module is used to discard some of the images ignored in further processing phases:
 - * Feature encoding / matching, recognition.



1 - Biometrics Fundamentals

- 1.1 – Definitions
- 1.2 - Existing Systems
- 1.3 – Non-cooperative recognition

2 - Our approach: QUIS-CAMPI

- 2.1 - Why ocular region?
- 2.2 - Main ocular degradation factors
- 2.3 – Periocular recognition
- 2.4 - Hardware framework / scene overview
- 2.5 - Processing chain

3 - Low-level Vision Problems

- 3.1 - Background subtraction, motion detection
- 3.2 - Human Detection
- 3.3 – Tracking
- 3.4 - Path prediction
- 3.5 - Camera calibration
- 3.6 - Target selection
- 3.7 - Face detection

4 - Pose Estimation + Soft labels

- 4.1 - 3D Head Meshes
- 4.2 - Pose Hypotheses
- 4.3 - Joint Indexing
- 4.4 - SOM Labels

5 - Data Segmentation

- 5.1 - Periocular Segmentation
- 5.2 - Iris Segmentation

6 - Feature Extraction / Encoding, Recognition

- 6.1 - Ordinal Measures
- 6.2 - Uncorrelated Weak / Strong Experts

7 - State-of-the-Art Performance

- 7.1 - Multiple Systems
- 7.2 - Current challenges

8 – ICBRW16: Competition in ICB 16

* The first generation of periocular recognition algorithms were **holistic**.

* They define a ROI around the iris, with dimensions proportional to the iris diameter (d)

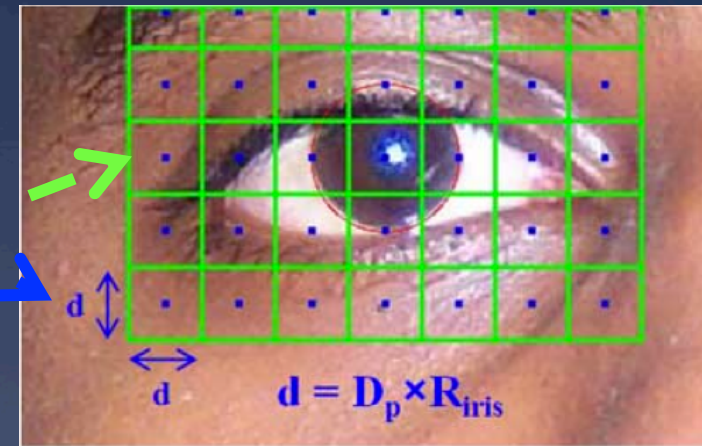
* Invariance to scale

* Next, the ROI is divided into blocks, and each **block**

described in terms of texture descriptors (HOGs, LBPs, SIFT)

* Such descriptors are used independently of the biological component inside each block

* Should the eyelids / eyebrows information be encoded similarly to the skin texture?



http://www2.cse.msu.edu/~rossarun/pubs/ParkPeriocularBiometrics_TIFS2011.pdf

- * The goal here is to obtain a (at least) coarse estimate of each biological component in a periocular image.
- * (1) Hair; (2) Eyebrows; (3) Eyelashes; (4) Sclera; (5) Iris; (6) Glasses and (7) Skin.
- * Based in such information, feature encoding / matching strategies specific to each component can be developed.



* **Our solution is composed of two major phases:**

1. Seven non-linear binary classifiers (one per component) run at the pixel level provide the posterior probabilities for each image position / class of interest.

* **They provide the local data appearance information**

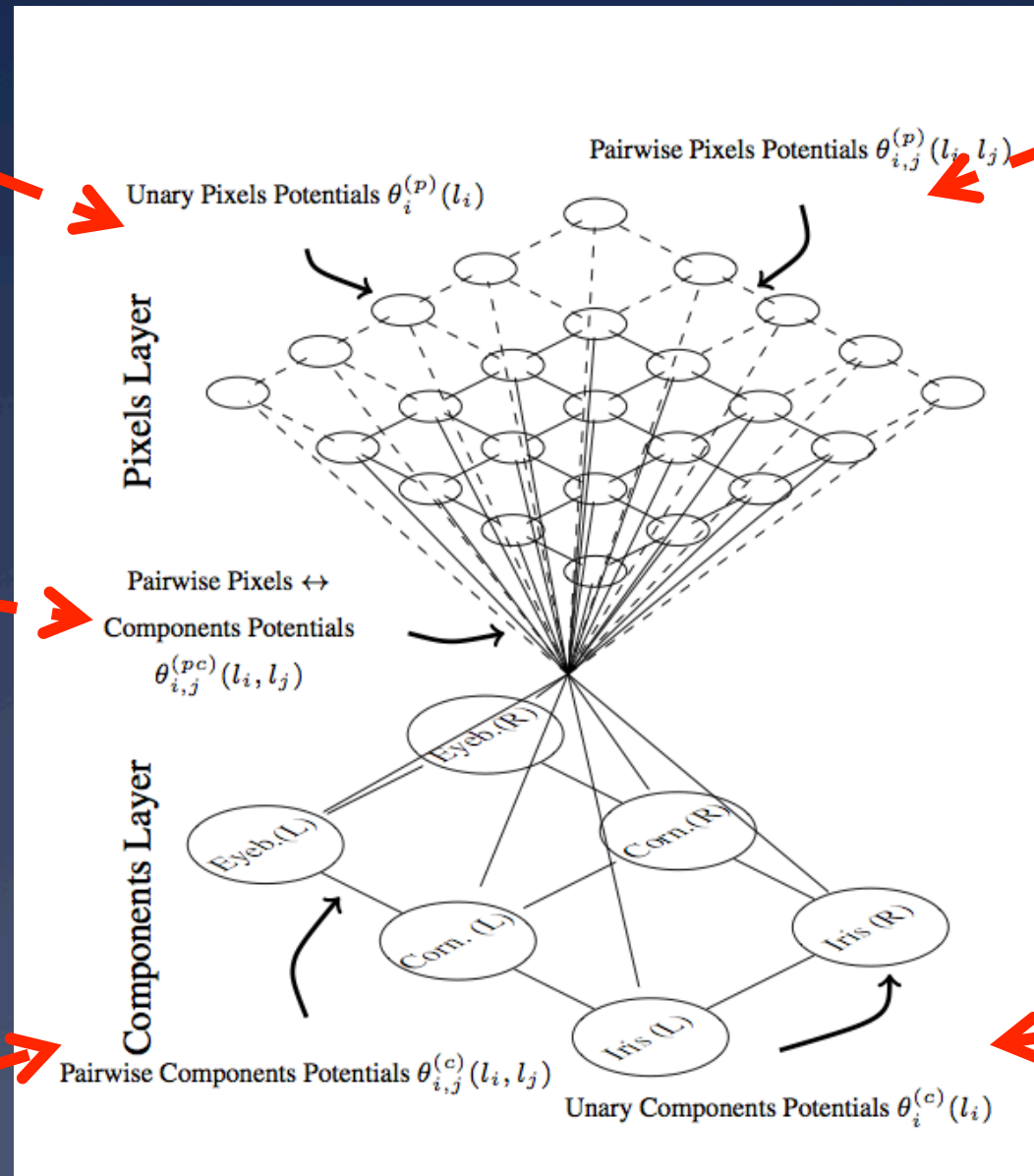
2. A hierarchical MRF combines the data local appearance to geometric / biological constraints and neighborhood information.

* **By minimizing the energy of the MRF, we obtain an acceptable segmentation solution.**

Data appearance information (provided by NNs)

Consistency between data appearance and components parameterization

Geometric constraints (e.g., the cornea surround the iris)



Probability that a “skin” pixel is adjacent to an “iris” pixel

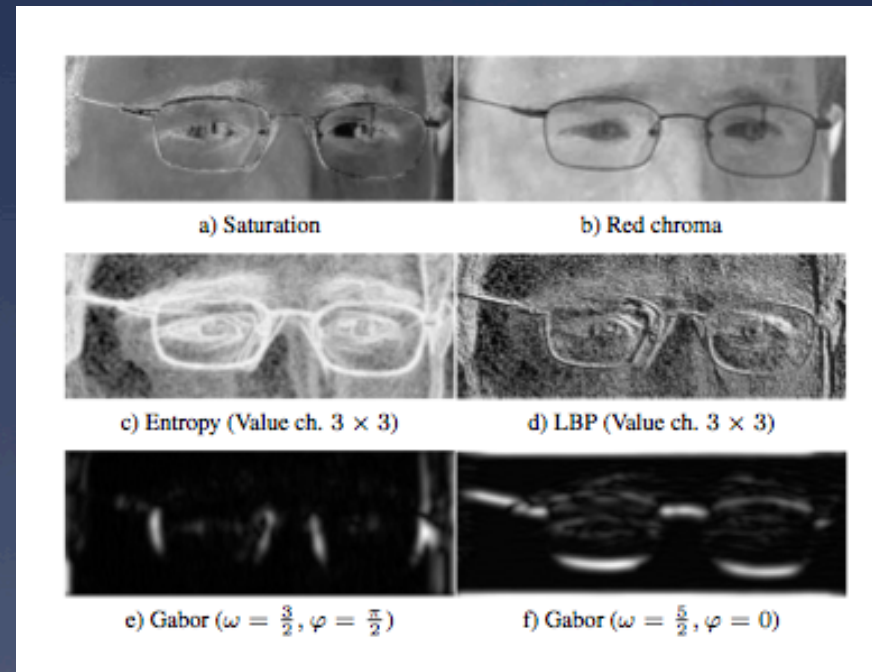
Scores of a randomized elliptical Hough transform

* Extract a set of (well known) local texture descriptors. Function: $\gamma(x, y) \in \mathbb{R}^{34}$

1. Saturation
2. Red Chroma
3. Entropy (intensity)
4. LBP (intensity)
5. Gabor response

* All descriptors extracted at three different scales.

* RGB, HSV and YCbCr color spaces used.



* Learn seven binary classification models, each one specialized in detecting one component. ω_i

* Learning set of a relative reduced number of images (<100), manually labeled. $\eta_i : \mathbb{R}^{34} \rightarrow [0, 1]$

* Use the response of the classification models to obtain the likelihood of each class.

* Kernel density estimation $p(\eta_i(\gamma(x, y)) | \omega_i)$

* According to the Bayes rule, the posterior probability functions are given by:

$$P(\omega_i | \eta_i(\gamma(x, y))) = \frac{P(\eta_i(\gamma(x, y)) | \omega_i)}{\sum_{j=1}^7 P(\eta_j(\gamma(x, y)) | \omega_j)}$$



Glasses posterior



Skin posterior

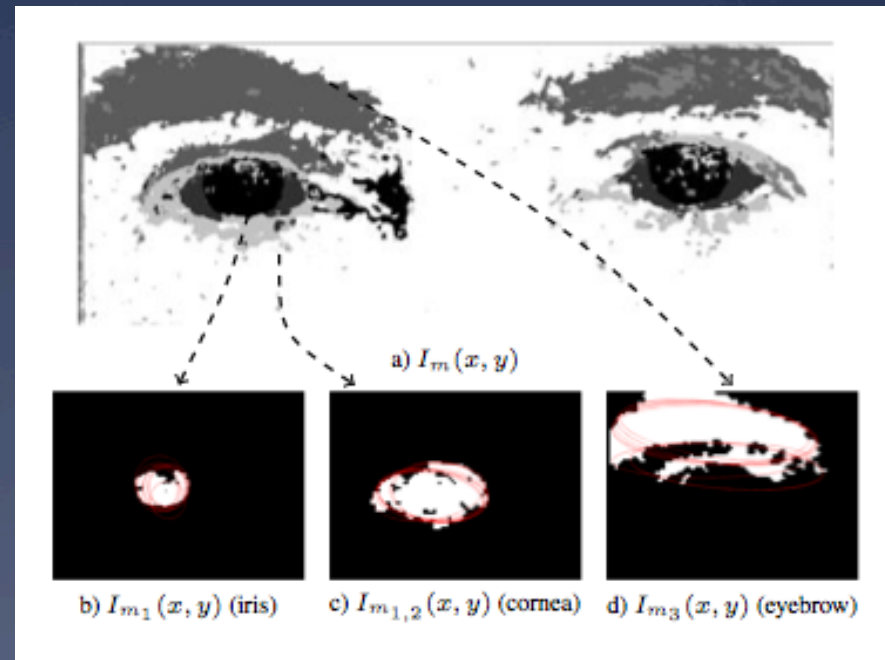
- * Obtain the image labeled by the index of the maximum posterior probability (among all classes) at each position:

$$I_m(x, y) = \arg \max_j p(\omega_j | \eta_j(\gamma(x, y)))$$

- * Based in this data, a set of binary images can be obtained by:

$$I_{m_i}(x, y) = \begin{cases} 1 & , \text{if } I_m(x, y) = i \\ 0 & , \text{otherwise} \end{cases}$$

- * For the “iris”, “cornea” and “eyebrow” images, run the Randomized Elliptical Hough Transform, to obtain a set of candidate parameterizations (per component).



* **Privilege consistency between data appearance and components parameterization:**

Inside ellipse test

$$\theta_{i,j}^{(pc)}(l_i, l_j) = \begin{cases} \min_k \|(x_i, y_i) - (x_{jk}, y_{jk})\|_2, & \text{if } l_i \in \mathcal{I}_j \\ \text{and } \psi(x_i, y_i, x_j, y_j, a_j, b_j, \varphi_j) = 0 \\ 0, & \text{if } l_i \notin \mathcal{I}_j \\ \text{and } \psi(x_i, y_i, x_j, y_j, a_j, b_j, \varphi_j) = 0 \\ 0, & \text{if } l_i \in \mathcal{I}_j \\ \text{and } \psi(x_i, y_i, x_j, y_j, a_j, b_j, \varphi_j) = 1 \\ \max_k \|(x_i, y_i) - (x_{jk}, y_{jk})\|_2, & \text{if } l_i \notin \mathcal{I}_j \\ \text{and } \psi(x_i, y_i, x_j, y_j, a_j, b_j, \varphi_j) = 1 \end{cases}$$

Pixels in this region should be **penalized** to be “iris” pixels, for any parameterization

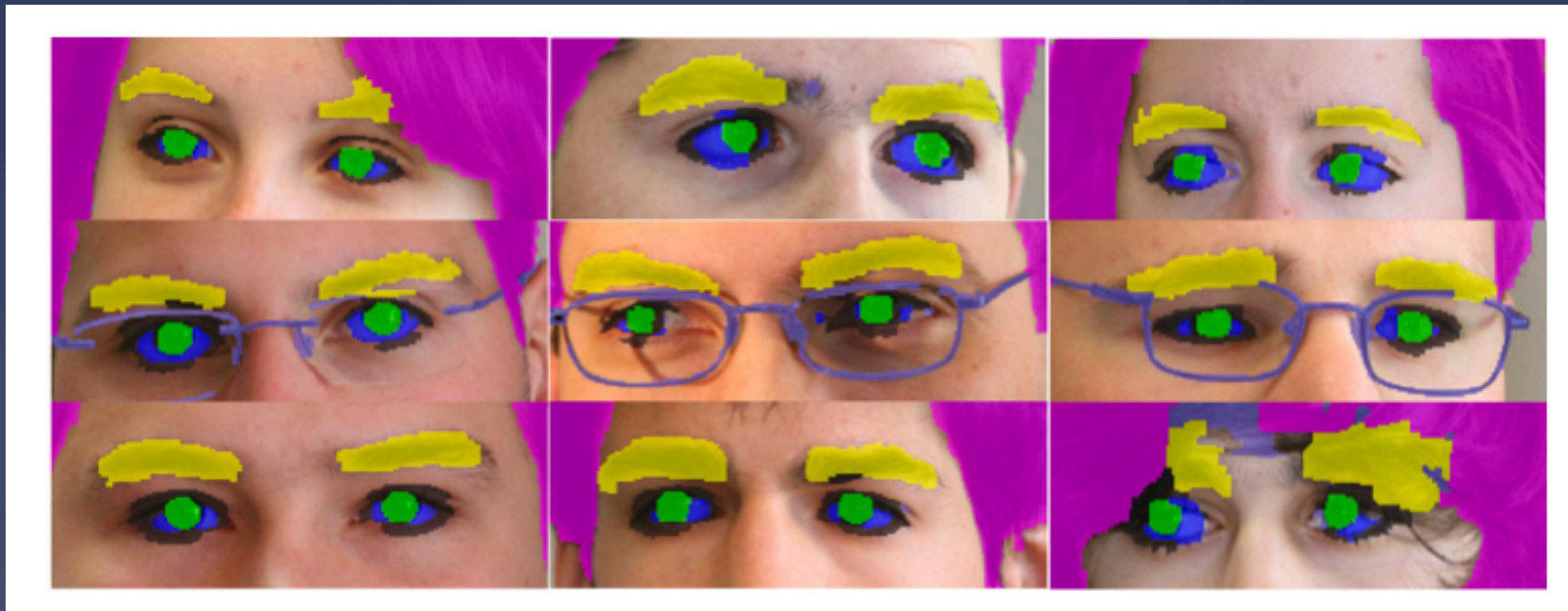


Pixels in this region should be **penalized** to be “iris” pixels, if parameterization “2” is chosen

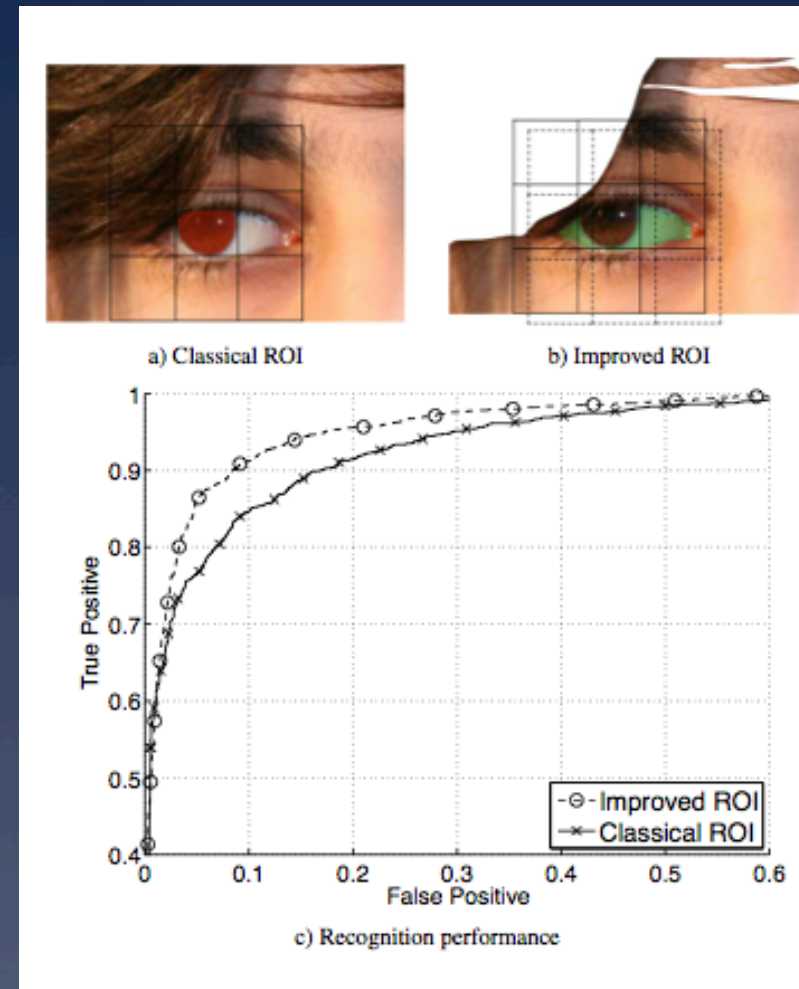
Pixels in this region should be **encouraged** to be “iris” pixels, for any parameterization

Pixels in this region should be **encouraged** to be “iris” pixels, if parameterization “1” is chosen

- * A particularly interesting performance was observed for glasses, where the algorithm attained remarkable results for various types of frames. Glasses were the unique non-biological component
- * In opposition, the most concerning cases happened when the eyebrows and the hair were overlapped.
- * Also, for heavily deviated gazes, the sclera was sometimes under-segmented (typically, by non-detecting the less visible side).
- * Eyelashes tended to be over-segmented, with isolated eyelashes being grouped in large eyelash regions.



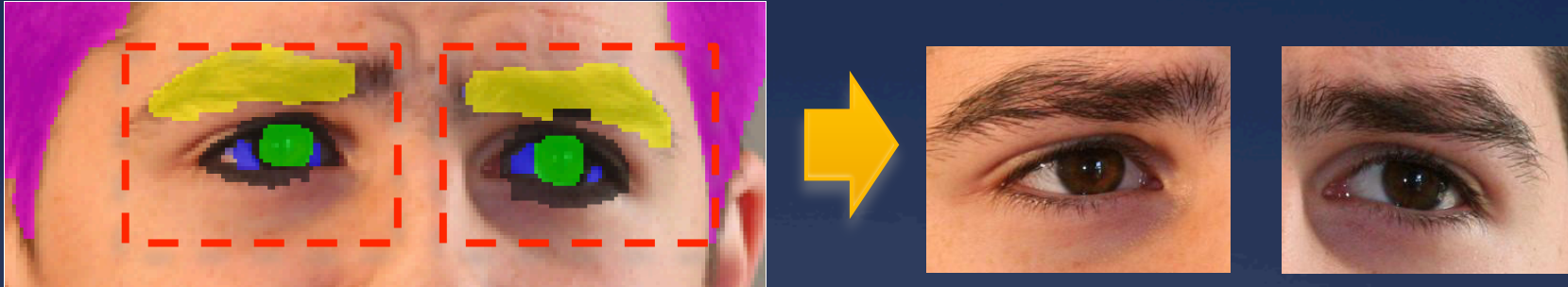
- * Segmenting the periocular region substantially **increases performance** by improving the way the ROI is defined:
 - Excluding hair;
 - Excluding glasses;
 - Using the center of the cornea as reference point, instead of the iris center.



Recognition performance of Park et al. [P1] algorithm, when using the classical and improved ROI

[P1] U. Park, R. Jillela, A. Ross and A. Jain. Periocular Biometrics in the Visible Spectrum. IEEE Transactions on Information Forensics and Security, vol. 6, no. 1, pag. 96–106, 2011.

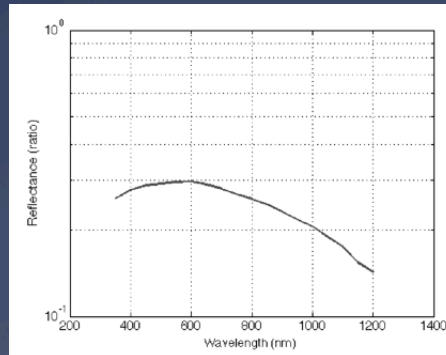
- * Given a (roughly) segmented periocular image, the next task comprises the accurate segmentation of the iris region (**Region-based methods**)



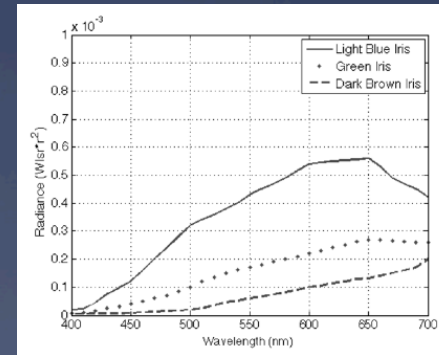
- * The classical “**boundaries-based**” algorithms fail completely in this type of images, due to the data variability and degradation factors:

$$\max(r, x_0, y_0) \left| G_\sigma(r) \approx \frac{\partial}{\partial r} \oint_{r, x_0, y_0} \frac{I(x, y)}{2\pi r} ds \right|$$

- * In visible-light data, the sclera is the most easily distinguishable component:



Spectral reflectance of the sclera in visible wavelengths [R1]

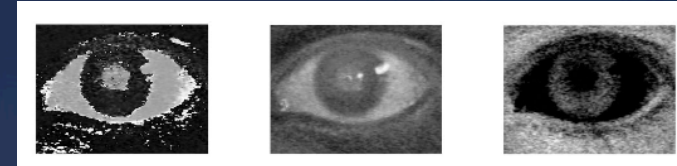


Spectral reflectance of the iris in visible wavelengths with respect to the levels of iris pigmentation

[R1] B. Nemati, H. Grady Rylander III, and A.J. Welch, “Optical Properties of Conjunctiva, Sclera, and the Ciliary Body and Their Consequences for Transscleral Cyclophotocoagulation,” Applied Optics, vol. 35, no. 19, pp. 3321-3327, July 1996.

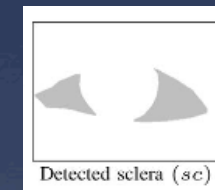
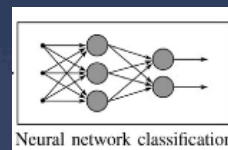
[R2] F. Imai, “Preliminary Experiment for Spectral Reflectance Estimation of Human Iris Using a Digital Camera,” technical report, Munsell Color Science Laboratories, Rochester Inst. of Technology, 2000.

- * The blue and red chroma, together with the hue channel are particularly effective in discriminating between the sclera and the remaining components



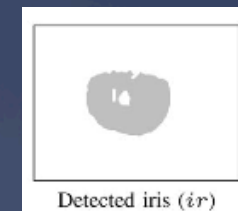
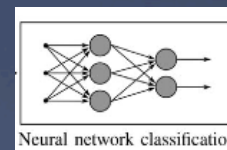
a) hue, b) blue-chroma and c) red-chroma

- * It is relatively straightforward to develop a local classifier that detects the sclera pixels.



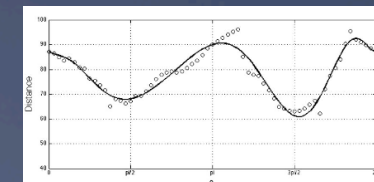
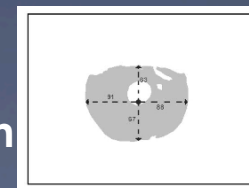
- * As the sclera surrounds the iris, this “sclera mask” is an essential information in the development of a local iris classifier

- * Using also the saturation channel.

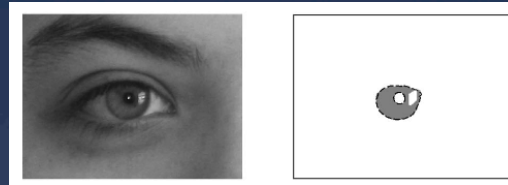
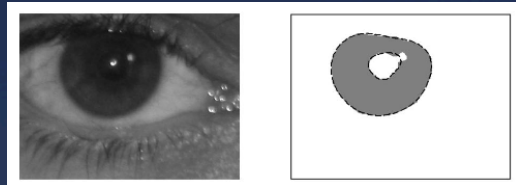


- * Obtaining a rough estimate of the iris center, the inner and outer boundaries of the iris are parameterized

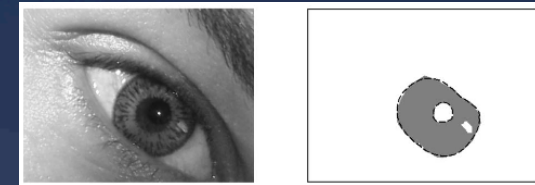
- * Constrained polynomial regression problem



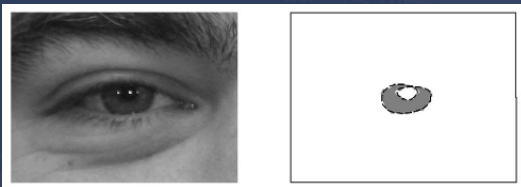
* Examples of visible light iris images classified and parameterized:



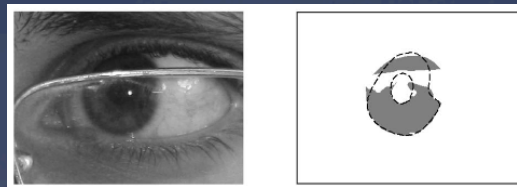
Poor resolution



Roll rotations



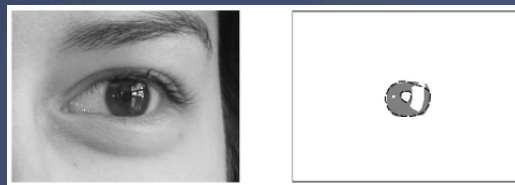
Eyelids occlusions



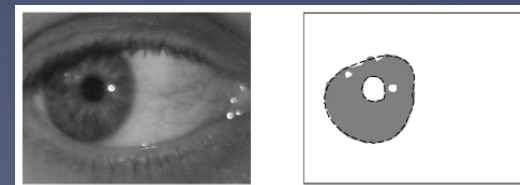
Glasses



Heavy pigmented irises



Varying pose



Deviated gaze

1 - Biometrics Fundamentals

- 1.1 – Definitions
- 1.2 - Existing Systems
- 1.3 – Non-cooperative recognition

2 - Our approach: QUIS-CAMPI

- 2.1 - Why ocular region?
- 2.2 - Main ocular degradation factors
- 2.3 – Periocular recognition
- 2.4 - Hardware framework / scene overview
- 2.5 - Processing chain

3 - Low-level Vision Problems

- 3.1 - Background subtraction, motion detection
- 3.2 - Human Detection
- 3.3 – Tracking
- 3.4 - Path prediction
- 3.5 - Camera calibration
- 3.6 - Target selection
- 3.7 - Face detection

4 - Pose Estimation + Soft labels

- 4.1 - 3D Head Meshes
- 4.2 - Pose Hypotheses
- 4.3 - Joint Indexing
- 4.4 - SOM Labels

5 - Data Segmentation

- 5.1 - Periocular Segmentation
- 5.2 - Iris Segmentation

6 - Feature Extraction / Encoding, Recognition

- 6.1 - Ordinal Measures
- 6.2 - Uncorrelated Weak / Strong Experts

7 - State-of-the-Art Performance

- 7.1 - Multiple Systems
- 7.2 - Current challenges

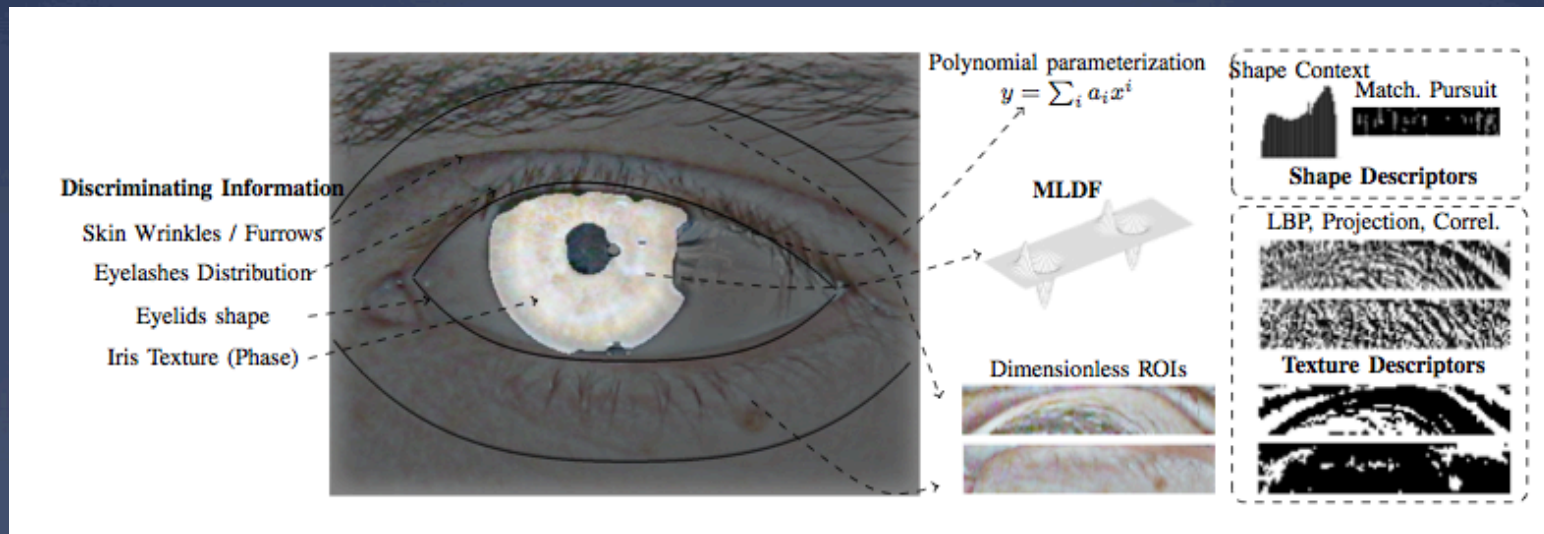
8 – ICBRW16: Competition in ICB 16

* Segmenting (labeling) the periocular region offers an interesting possibility:

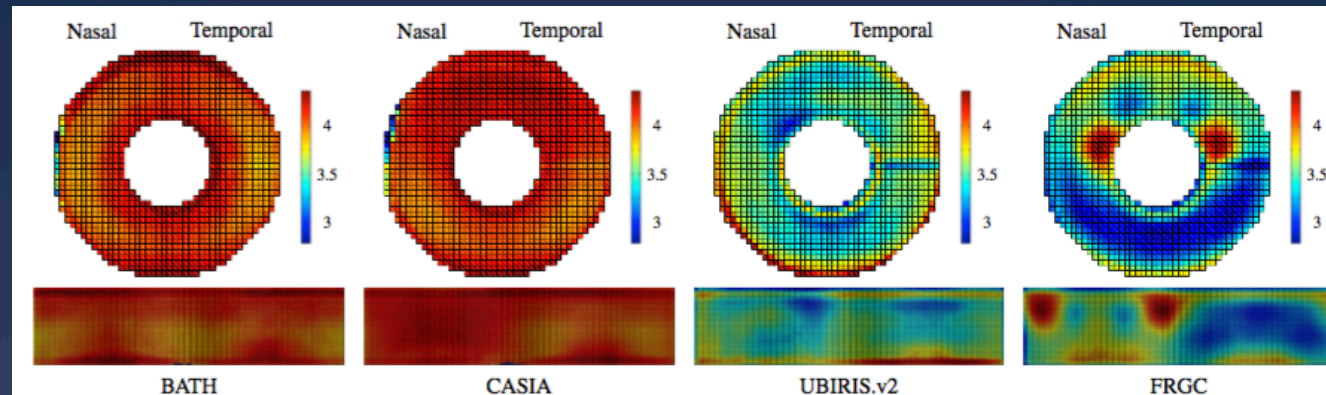


* Develop periocular recognition methods made of disparate components (two in our case):

- * One expert analyses the iris texture and exhaustively exploits the multi-spectral information in visible-light data;
- * Another expert analyses the shape of eyelids and defines a surrounding dimensionless region-of-interest, from where statistics of the eyelids, eyelashes and skin wrinkles / furrows are encoded.



- * **Not surprisingly, images acquired under visible-light unconstrained protocols actually have less information than NIR data acquired in constrained setups:**



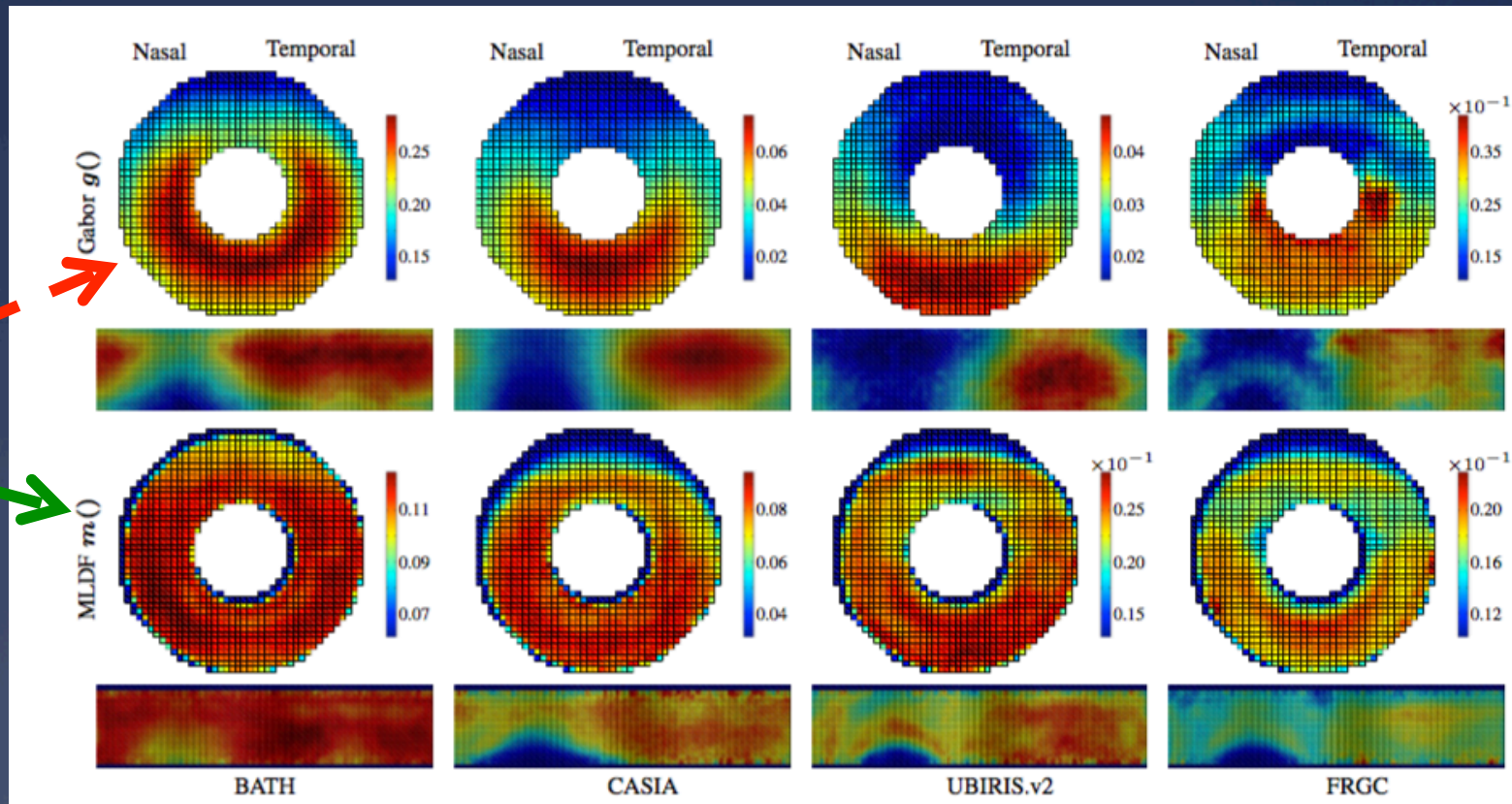
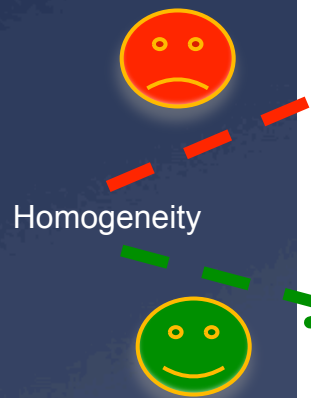
Average amount of information (Shannon entropy in 9 x 9 patches of the normalised images) across the different regions of the irises

- * Also, apart from the known **fragility** (intra-class consistency) of some bits in the iris codes, there is another phenomenon, particularly evident in visible-light data acquired under non-homogenous lighting conditions:
 - * Some bits in the codes have predominant values (either “0” or “1”), even considering images of different subjects.
 - * This have led to the concept of **bit discriminability** (intra-class consistency + inter-class variability).
- * Feature encoding / matching strategies specialized for visible-light data should be particularly useful.

[Q1] H. Proença. Iris Recognition: What's Beyond Bit Fragility? IEEE Transactions on Information Forensics and Security, volume 10, issue 2, pag. 321-332.

* Let $x_k = P(C_k^{(p)} \oplus C_k^{(q)} = 0 | H_0)$ and $y_k = P(C_k^{(p)} \oplus C_k^{(q)} = 0 | H_a)$ be 2D representation of the intra-class consistency (x_k) and inter-class variability (y_k) of the k^{th} bit in an iris code.

* A quantitative measure of bit **discriminability** is given by: $\tau(k) = |y_k - x_k|$

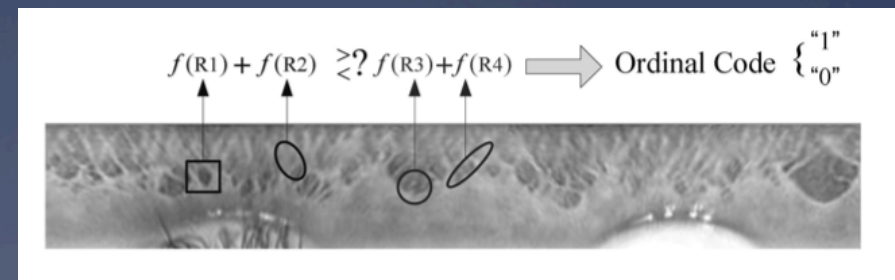
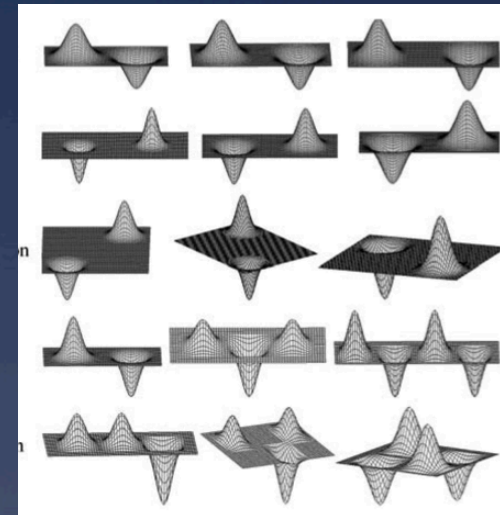


Average bit discriminability across the iris, in four well known iris data sets.

* **Multi-lobe Differential filters** were first applied to the iris recognition problem, as Ordinal Measures [O1], which are important for the extraction of robust discriminating features from degraded data.

- * **Extremely simple to extract.**
- * **Computationally inexpensive (Difference of Gaussian blurred images)**
- * **Analyze non-adjacent patterns**
- * **High number of parameters, making hard to obtain the optimal configurations**

$$MLDF = C_p \sum_{i=1}^{N_p} \frac{1}{\sqrt{2\pi\delta_{pi}}} \exp \left[\frac{-(X - \mu_{pi})^2}{2\delta_{pi}^2} \right] - C_n \sum_{j=1}^{N_n} \frac{1}{\sqrt{2\pi\delta_{nj}}} \exp \left[\frac{-(X - \mu_{nj})^2}{2\delta_{nj}^2} \right],$$



[O1] Z. Sun and T. Tan; "Ordinal Iris Measures for Iris Recognition", IEEE Transactions on Pattern Analysis and Machine Intelligence, 31 (12), 2211-2226, 2009.

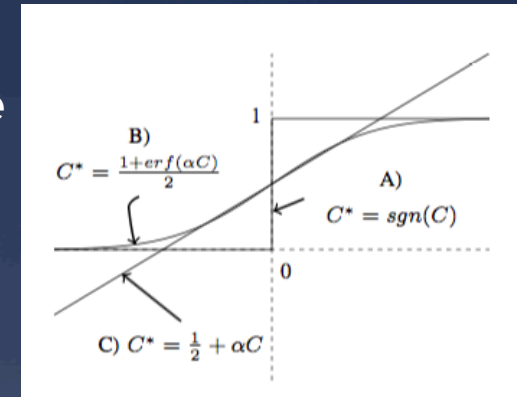
- * The most acknowledged iris recognition algorithm **only considers signal phase**, as amplitude is considered unreliable.
- * However, due to the lack of discriminating information in visible-light unconstrained recognition, it is important to maximize the potential amount of information sources.

* Three different code quantization strategies can be compared:

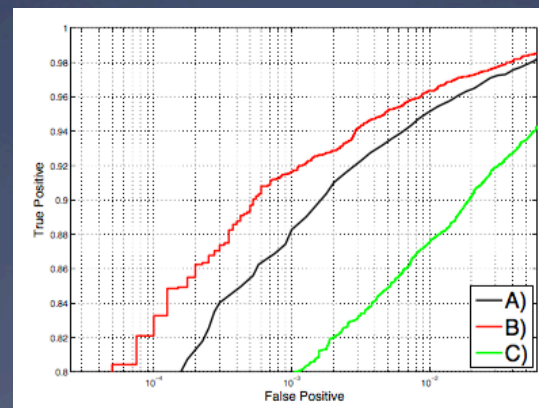
A) signal

B) sigmoid

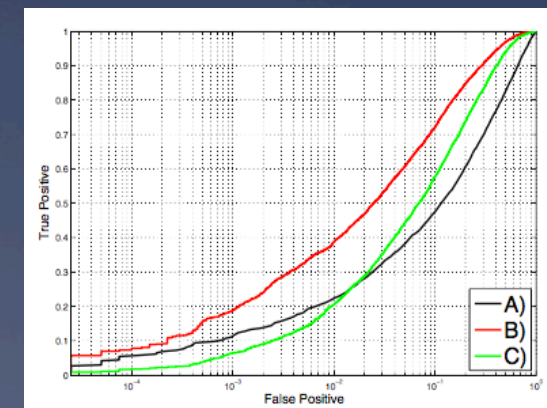
C) No quantization



BATH, NIR data (constrained)



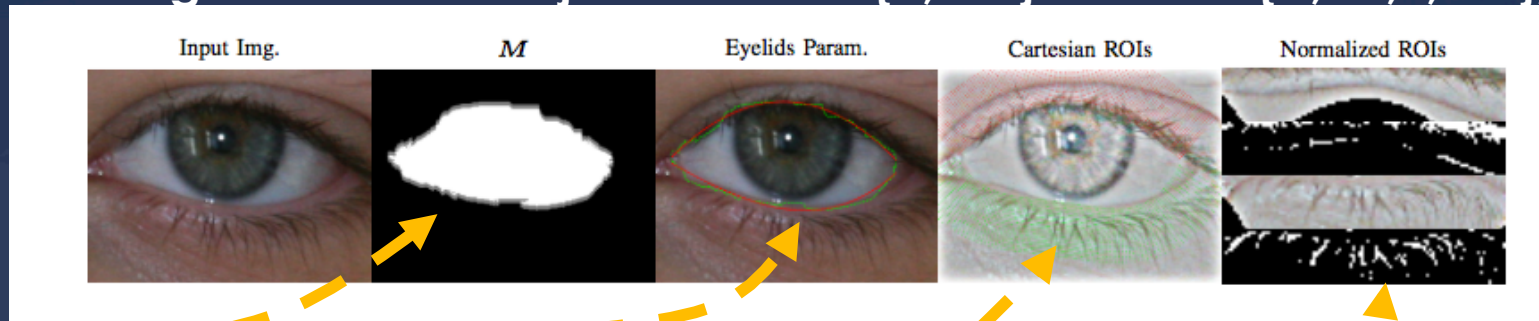
UBIRIS.v2 VW data (less constrained)



A sigmoid-based quantization is always better than signal-based, both in NIR and VW data.

[Q1] H. Proença. Iris Recognition: What's Beyond Bit Fragility? IEEE Transactions on Information Forensics and Security, volume 10, issue 2, pag. 321-332.

- * We use the Iris “I” and sclera “S” masks, defining $M=I \vee S$.
- * Let $E=\{(x_i, y_i)\}$ be the set of edge pixels in “M”, ordered clockwise.
- * Let $l^* = \arg \min_i x_i$ and $r^* = \arg \max_i x_i$ be the deemed positions of eye corners (assuming moderate “roll” variations)
- * Dividing “E” into two disjoint sets $e^{(1)}=\{l^*, \dots, r^*\}$ and $e^{(2)}=\{r^*, \dots, t, 1, \dots, l^*\}$,



M= Iris \vee Sclera Mask

→ First generation eyelids boundaries

→ Second generation eyelids boundaries (inliers only)

→ Lower Eyelid sampling points

→ Upper eyelid sampling points

Dimensionless normalized ROIs and corresponding eyelashes masks

- * The following system finds the coefficients “ a_i ” of an interpolating polynomial of degree “n”. using RANSAC, a subset of the $e^{(\cdot)}$ elements are selected (as inliers) and the process repeated until convergence.

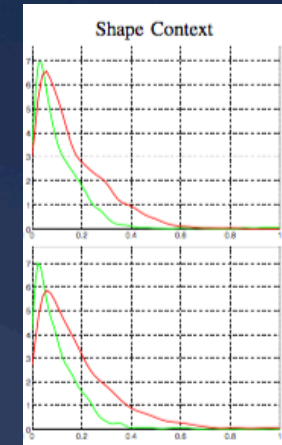
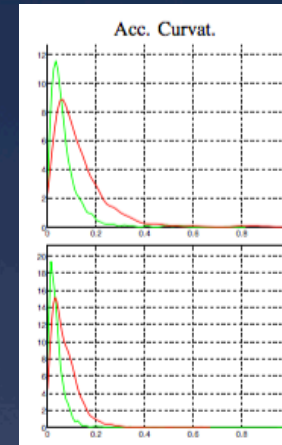
$$\begin{bmatrix} x_{e_0}^n & x_{e_0}^{n-1} & \dots & x_{e_0}^0 \\ x_{e_1}^n & x_{e_1}^{n-1} & \dots & x_{e_1}^0 \\ \vdots & \vdots & \ddots & \vdots \\ x_{e_n}^n & x_{e_n}^{n-1} & \dots & x_{e_n}^0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{bmatrix},$$

* However, note that each shape / texture descriptor can only be considered a “weak” biometric expert:

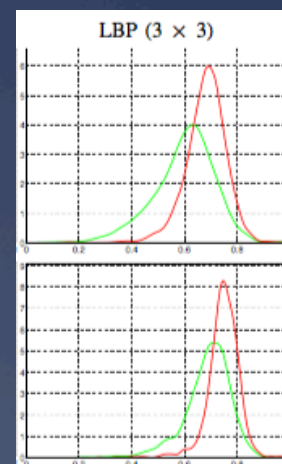
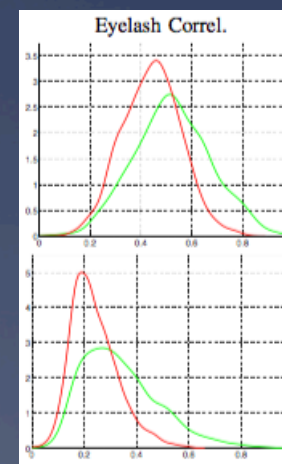
* There is a small separability between the “genuine” / “impostor” scores distribution.

* Finally, the “strong” / “weak” scores are fused according to a Bayesian paradigm, having tested the classical “sum”, “product”, “min” and “max” rules.

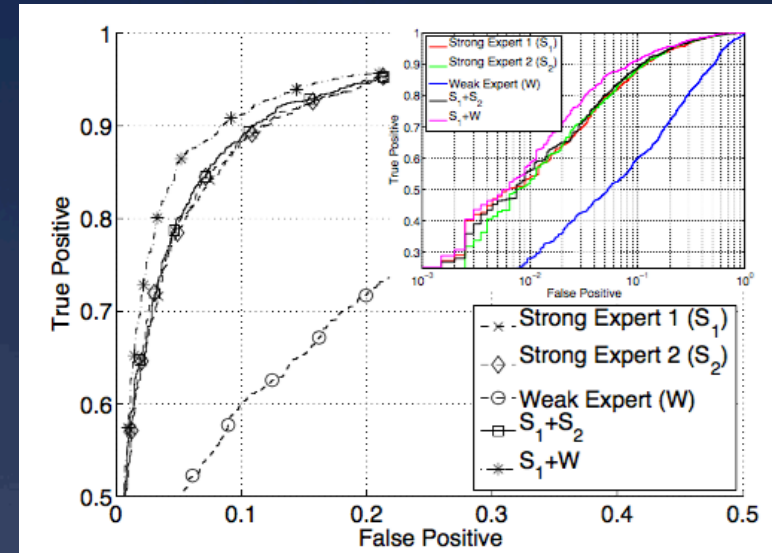
Texture Descriptors
Upper ROI
Lower ROI



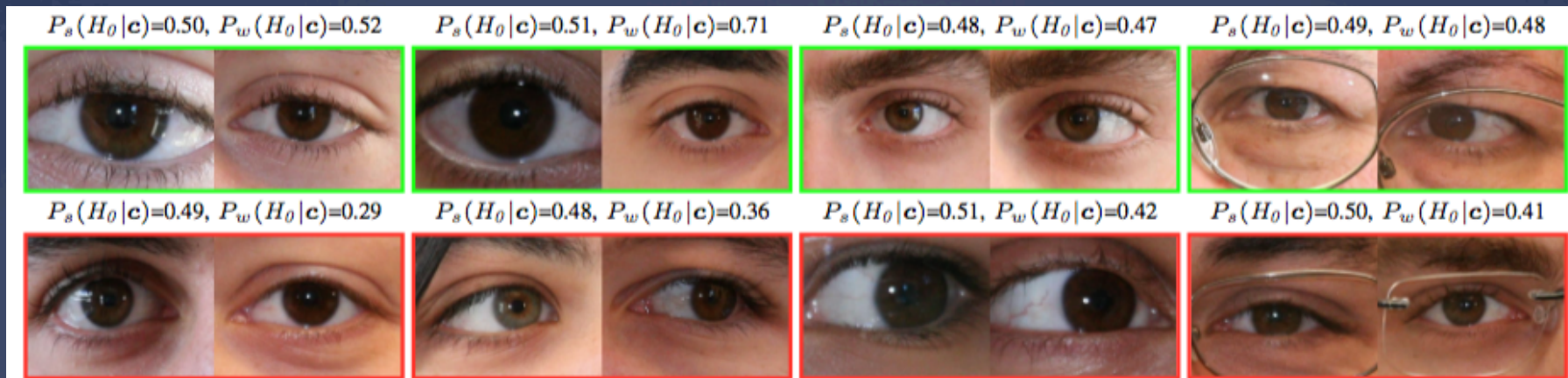
Shape Descriptors
Upper Eyelid
Lower Eyelid



* The interesting point is that it is better to fuse a strong (iris) and a weak (periocular) experts, than two strong experts (even if they extract completely disjoint types of feature sets).



* The weak ensemble was designed so to make it particularly useful in the “uncertain region” of the iris expert:



Examples of pair wise comparisons that fall in the uncertain region $P_s(H_0 | x) \approx 0.5$ of the strong biometric expert

1 - Biometrics Fundamentals

- 1.1 – Definitions
- 1.2 - Existing Systems
- 1.3 – Non-cooperative recognition

2 - Our approach: QUIS-CAMPI

- 2.1 - Why ocular region?
- 2.2 - Main ocular degradation factors
- 2.3 – Periocular recognition
- 2.4 - Hardware framework / scene overview
- 2.5 - Processing chain

3 - Low-level Vision Problems

- 3.1 - Background subtraction, motion detection
- 3.2 - Human Detection
- 3.3 – Tracking
- 3.4 - Path prediction
- 3.5 - Camera calibration
- 3.6 - Target selection
- 3.7 - Face detection

4 - Pose Estimation + Soft labels

- 4.1 - 3D Head Meshes
- 4.2 - Pose Hypotheses
- 4.3 - Joint Indexing
- 4.4 - SOM Labels

5 - Data Segmentation

- 5.1 - Periocular Segmentation
- 5.2 - Iris Segmentation

6 - Feature Extraction / Encoding, Recognition

- 6.1 - Ordinal Measures
- 6.2 - Uncorrelated Weak / Strong Experts

7 - State-of-the-Art Performance

- 7.1 - Multiple Systems
- 7.2 - Current challenges

8 – ICBRW16: Competition in ICB 16

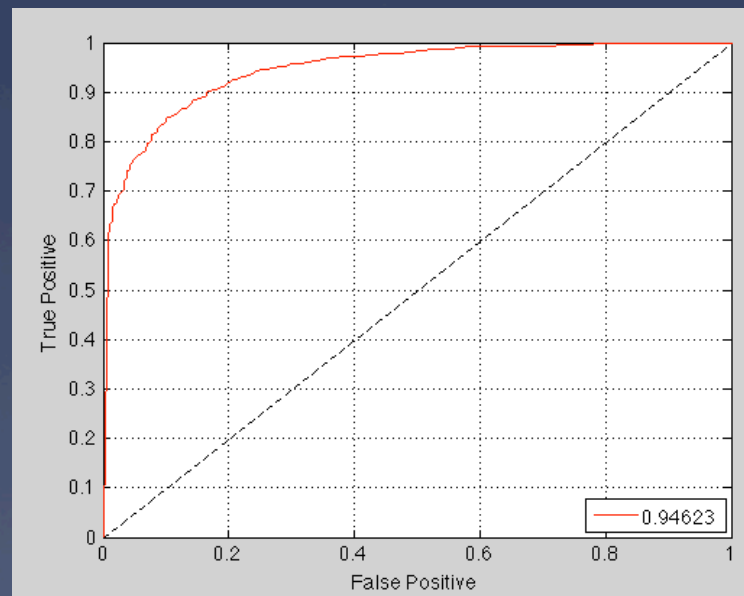
* Ocular Recognition: State-of-the-Art Performance

$$d' = \frac{|\mu_E - \mu_I|}{\sqrt{\frac{\sigma_I^2 + \sigma_E^2}{2}}}$$

* Decidability index:

- * Tan et al: $d'=2.57$ (also considers the eyes' surrounding).
- * $d' > 3$ is considered a “good” pattern recognition system.
- * Using exclusively the iris ($d'=2.12$).

* Receiver Operating Characteristic Curve (AUC=0.94623)



T. Tan, X. Zhang, Z. Sun and H. Zhang. “Noisy Iris Image Matching by Using Multiple Cues”, Pattern Recognition Letters, vol. 33, pag. 970-977, 2012.

- * Let X and Y represent the genuine / impostor distance distributions given by a state-of-the-art ocular recognition expert.

$$X \sim \mathcal{N}(\mu_x, \sigma_x)$$

$$Y \sim \mathcal{N}(\mu_y, \sigma_y)$$

- * Let $Y_{(k)}$ be the k^{th} order statistic of a sample $\{Y_1, \dots, Y_n\}$. Its distribution is given by, with $F(y)$ representing the cumulative distribution function of Y :

$$\begin{aligned} F_{(k)}(y) &= P(Y_{(k)} \leq y) \\ &= \sum_{i=k}^n \binom{n}{i} [F(y)]^i [1 - F(y)]^{n-i} \end{aligned}$$

- * Let $f_{(k)}$ be the density of $Y_{(k)}$:

$$\begin{aligned} f_{(k)}(y) &= n \binom{n-1}{k-1} [F(y)]^{k-1} [1 - F(y)]^{n-k} f(y) \\ &= \frac{n!}{(k-1)!(n-k)!} [F(y)]^{k-1} [1 - F(y)]^{n-k} f(y). \end{aligned}$$

- * The probability that a genuine score X returns a cumulative rank- k is obtained from the previously found density function:

$$\begin{aligned} P(\text{rank}(X) \leq k) &= P(X \leq Y_k) \\ &= \int_{-\infty}^{\infty} f_{(k)}(y) \int_{-\infty}^y f(x) dx dy. \end{aligned}$$

- * Also, the probability that the genuine score has exact rank- k is given by:

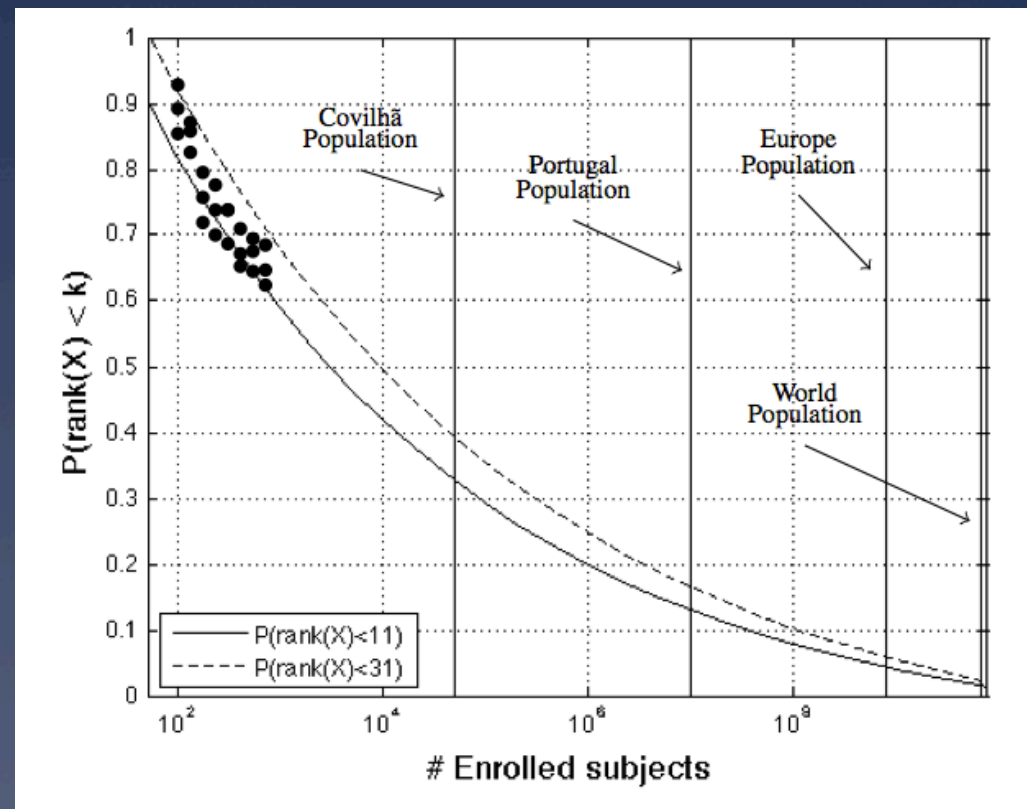
$$P(\text{rank}(X)=k) = P(X \leq Y_{(k)}) - P(X \leq Y_{(k-1)})$$

- * Using the above equations, we can estimate the probability that a sample has a cumulative rank- k , as function of the gallery size “ n ”.

* **Probability that the true identity of a query is returned in the first 10 / 30 results, with respect to the number of subjects enrolled in the system.**

Only **2%** of the queries would return the true identity in top-10 results, if the entire **world population** is considered.

Values rise to **4%** and **12%** if the number of gallery identities is reduced to **Continental** (Europe) or **National** (Portuguese) population.



- * Suppose that we are able to install several consecutive recognition systems.
- * Assuming the independence of the responses provided by these systems, the probability that a subject is screened by k recognition systems and correctly recognized (identity in top-10) in k' of these corresponds to the probability mass function of a binomial distribution:

$$P(R_{k'}) = \frac{k!}{k'! (k - k')!} \alpha^{k'} (1 - \alpha)^{k - k'}$$

- * Assuming that a match is reported only if at least k' (out of k) report a match:

$$\begin{aligned} P(R_{\geq k'}) &= \sum_{j=k'}^k P(R_j) \\ &= \sum_{j=k'}^k \frac{k!}{j! (k - j)!} \alpha^j (1 - \alpha)^{k - j} \end{aligned}$$

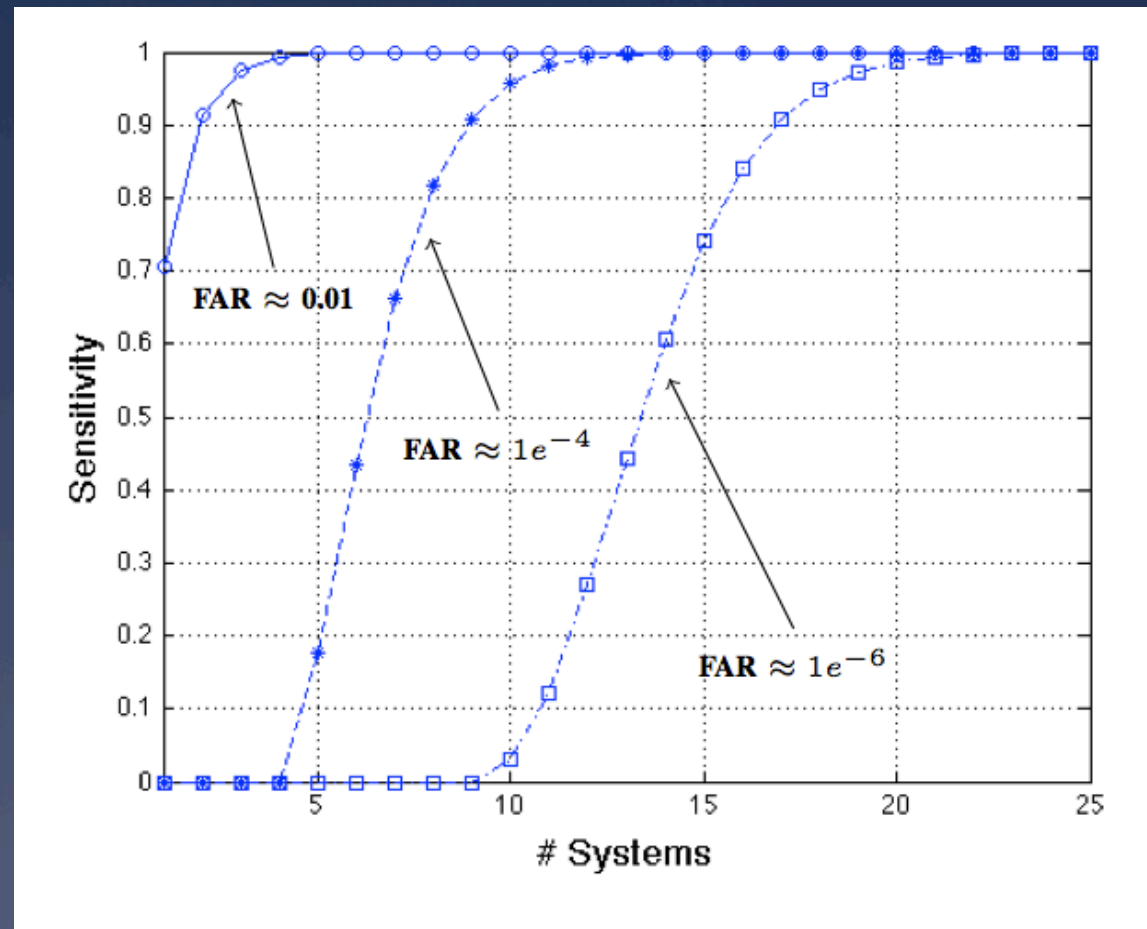
- * We can now obtain the expected sensitivity of such a **multi-point biometric recognition system**.

- * Upper bound for the expected sensitivity of a set of ocular recognition systems, at 3 different levels of FAR. (k'=1, National (Portuguese, 10,000,000 inhabitants) context)

5 recognition systems will be enough to attain full sensitivity, if FAR ≈ 0.01 is tolerated

The number of systems required raises considerably if lower FARs are required.

13 systems for FAR $\approx 1e^{-4}$ and 22 systems for FAR $\approx 1e^{-6}$



1 - Biometrics Fundamentals

- 1.1 – Definitions
- 1.2 - Existing Systems
- 1.3 – Non-cooperative recognition

2 - Our approach: QUIS-CAMPI

- 2.1 - Why ocular region?
- 2.2 - Main ocular degradation factors
- 2.3 – Periocular recognition
- 2.4 - Hardware framework / scene overview
- 2.5 - Processing chain

3 - Low-level Vision Problems

- 3.1 - Background subtraction, motion detection
- 3.2 - Human Detection
- 3.3 – Tracking
- 3.4 - Path prediction
- 3.5 - Camera calibration
- 3.6 - Target selection
- 3.7 - Face detection

4 - Pose Estimation + Soft labels

- 4.1 - 3D Head Meshes
- 4.2 - Pose Hypotheses
- 4.3 - Joint Indexing
- 4.4 - SOM Labels

5 - Data Segmentation

- 5.1 - Periocular Segmentation
- 5.2 - Iris Segmentation

6 - Feature Extraction / Encoding, Recognition

- 6.1 - Ordinal Measures
- 6.2 - Uncorrelated Weak / Strong Experts

7 - State-of-the-Art Performance

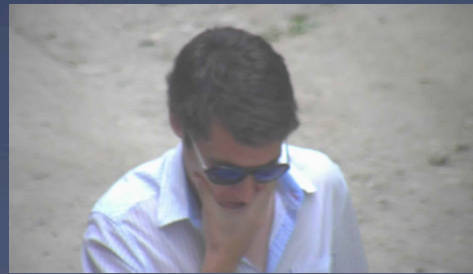
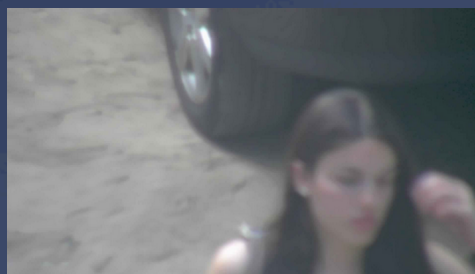
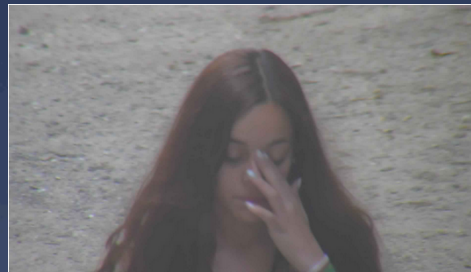
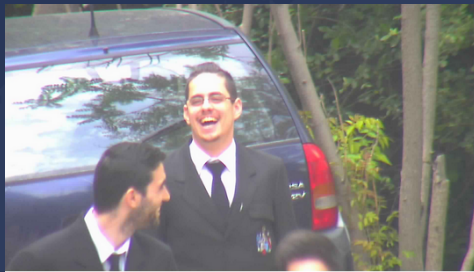
- 7.1 - Multiple Systems
- 7.2 - Current challenges

8 – ICBRW16: Competition in ICB 16

- * The described wide-view + PTZ data acquisition framework has been running continuously in our university campus, since May 2015.
- * We asked to students / teachers in our Department to offer as volunteers for the “QUIS-CAMPI dataset”.
- * The idea is to collect high resolution data in indoor, controlled conditions, asking subjects collaboration to get (as much as possible) noise-free data, in order to simulate the enrollment conditions / protocols.
- * The enrollment data is composed by:
 - * High resolution frontal / profile face images
 - * Full body frontal / profile images
 - * 0°, 45° and 90° gait sequences
 - * 3D face models
 - * Soft biometric information (age, gender, race)



- * The competition “**ICBRW16: International Challenge on Biometric Recognition in the Wild**” will run in the scope of ICB16.
 - * The main task consists in matching the images acquired outdoor in uncontrolled conditions to the watch list elements (the enrollment images).
 - * Given a (degraded) sample, participants in this competition are required to produce a similarity score for each of the watch list elements.
 - * The required result will be an ordered list (with the similarity between the sample and each of the enrolled identities)



- * Additional details can be found at: <http://icbrw.di.ubi.pt>

?