Computer Vision & Multimedia Lab

University of Pavia
Industrial Engineering and Computer Science Department

Staff

- Virginio Cantoni, Professor - Director
- Luca Lombardi, Associate Professor
- Mauro Mosconi, Assistant Professor (part-time)
- Marco Porta, Assistant Professor
- Marco Piastra, Contract Professor
- Roberto Marmo, Contract Professor
- Alessandra Setti, System Administrator

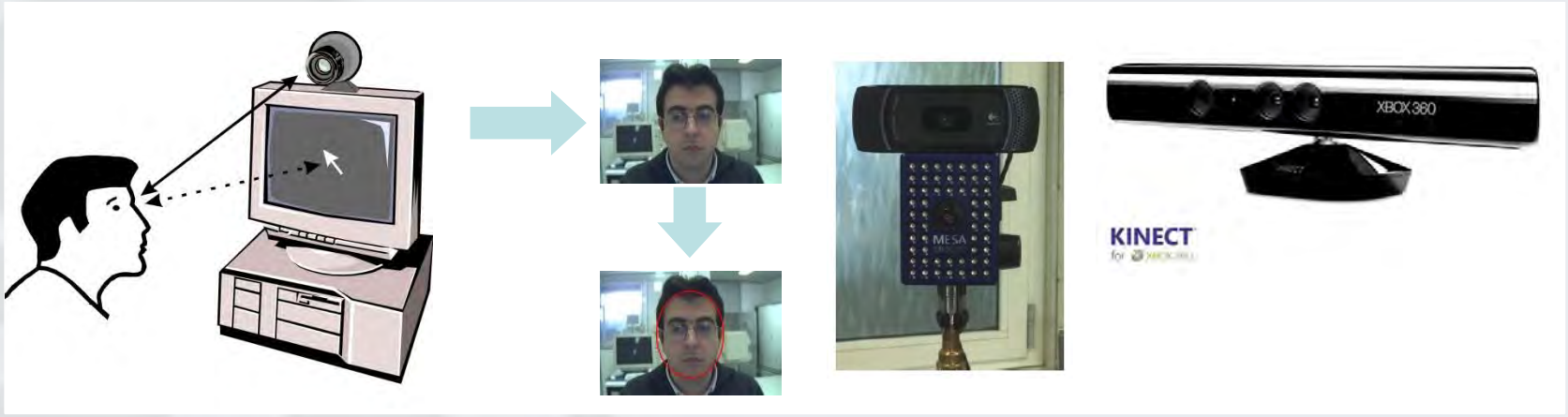
History – the beginnings

- The initial research activities of the group (early 70s) concentrated on the techniques of image enhancement and restoration, with particular regard for medical imagery
- Later, a broad background has been acquired on low level and intermediate level vision
- From the early 80s a new stream of research has been actively followed in the field of parallel architectures for vision and image processing
- ...

Current research areas

- New research areas are now activated on:
  - Pattern Recognition in Proteomics,
  - Human-Computer Interaction,
  - 3D Vision,
  - Multimedia,
  - E-learning,
  - Image Synthesis,
  - Visual Languages,
  - Pyramidal Architectures for Computer Vision

Vision-based perceptive interfaces



Exploiting vision to implement user interfaces based on **gesture recognition** and **head tracking**

- ordinary webcams
- TOF cameras
- Microsoft Kinect

Vision-based perceptive interfaces

Examples: hand gestures

Page scrolling

GEM: Gesture Enhanced Mouse

Human Computer Interaction

- "Traditional" activity: **design** of **visual interfaces** (i.e. drag-and-drop for e-commerce websites, new paradigms for browsing of images)
- Experience on **web accessibility**
- Experience on **usability evaluation** (cognitive walkthrough, thinking aloud)

Eye Tracking

- The Tobii 1750 Eye Tracker is integrated into a 17" TFT monitor.

  It is useful for all forms of eye **tracking studies with stimuli that can be presented on a screen**, such as websites, slideshows, videos and text
- The eye tracker is non-intrusive. Test subjects are allowed to move freely in front of the device

Eye Tracking

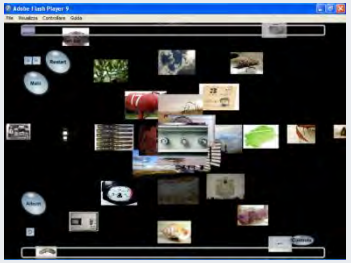
Application areas

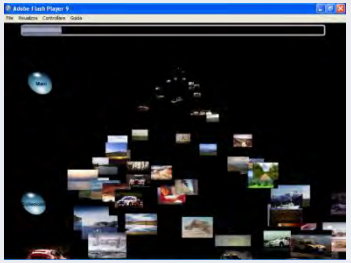
- **usability** and advertising **testing**
  what people watch reflects their thinking and cognitive processes – insights that cannot be obtained directly with other testing methods
- **eye control for accessibility**
  eye control enables users with special needs to communicate and interact using only their eyes

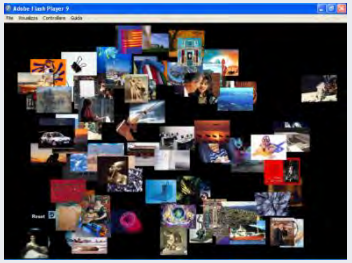
Eye Tracking

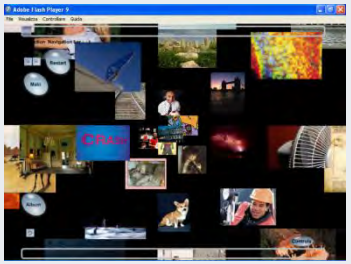
Advanced techniques for large image database browsing

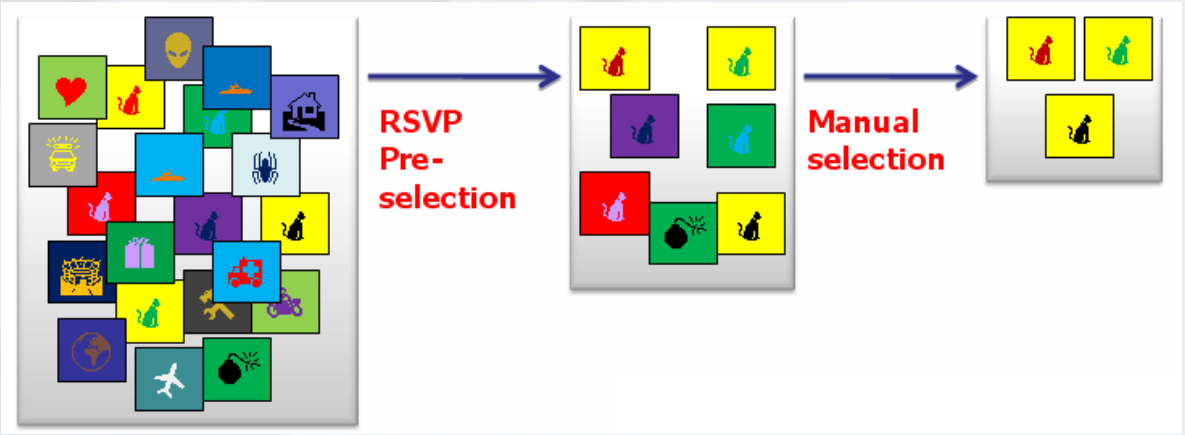
Eye-controlled **RSVP** (Rapid Serial Visual Presentation)







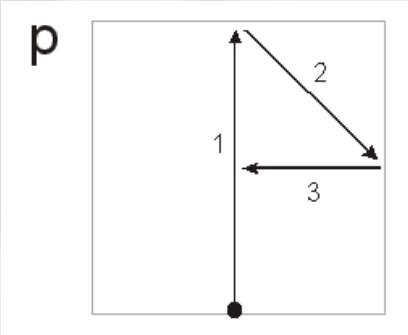


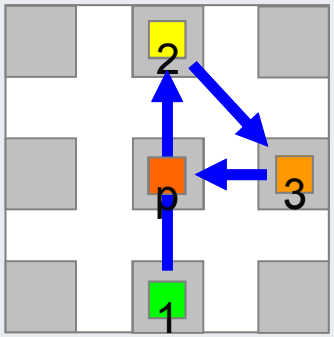


Eye-based communication
Eye-S, a system for pure eye-based communication

the user creates alphabet letters, as well as punctuation marks and commands definable according to the specific application to control, by means of sequences of fixations on nine predefined (and invisible) screen areas

E.g.

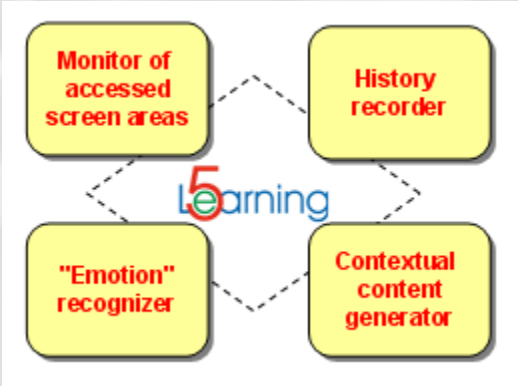




Eye Tracking

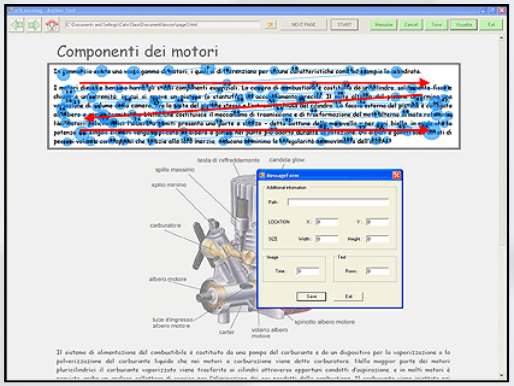
“Intelligent” e-learning systems

e5Learning: enhanced exploitation of eyes for effective eLearning



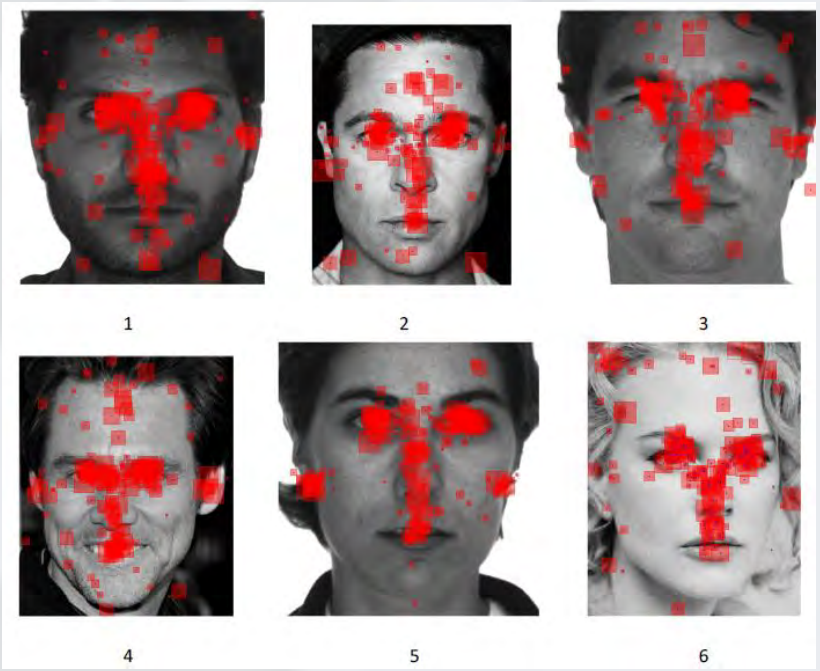
three main functionalities:

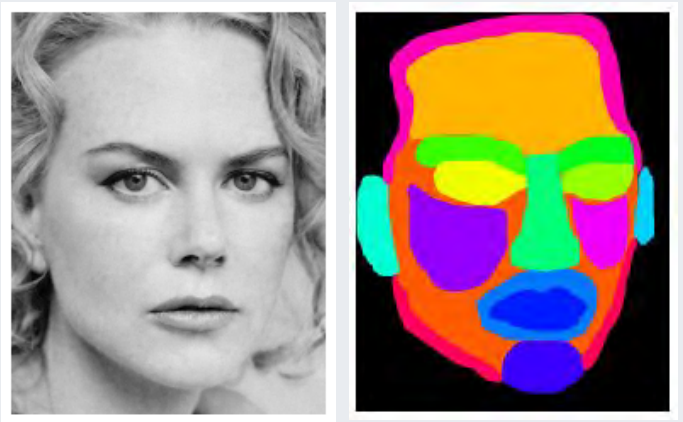
- detection of basic user activities, such as reading text and observing multimedia content
- contextual content generation
- recognition of stress, high workload and tiredness states in the user



Soft biometry techniques

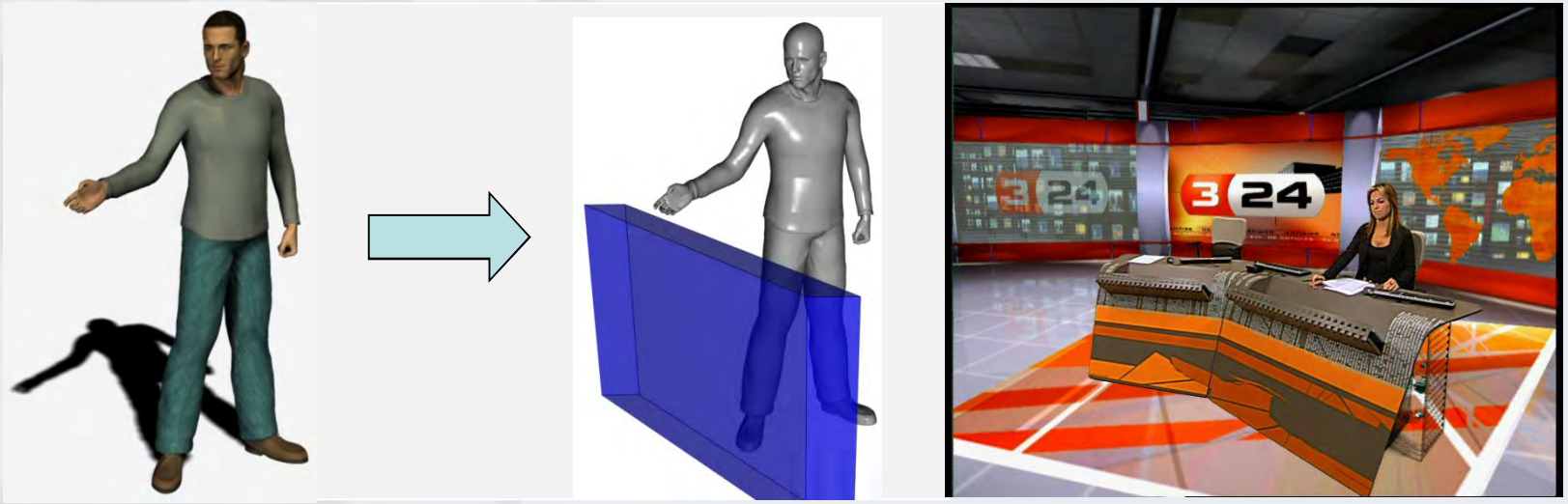
Are there any differences in the way specific kinds of images (e.g. faces) are watched by specific subjects?





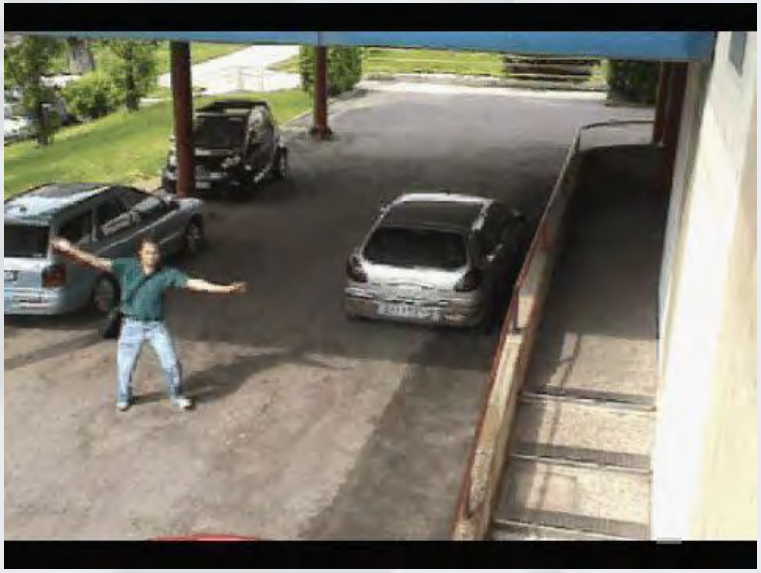
Collaboration with RAI multimedia
3D Vision – Virtual set

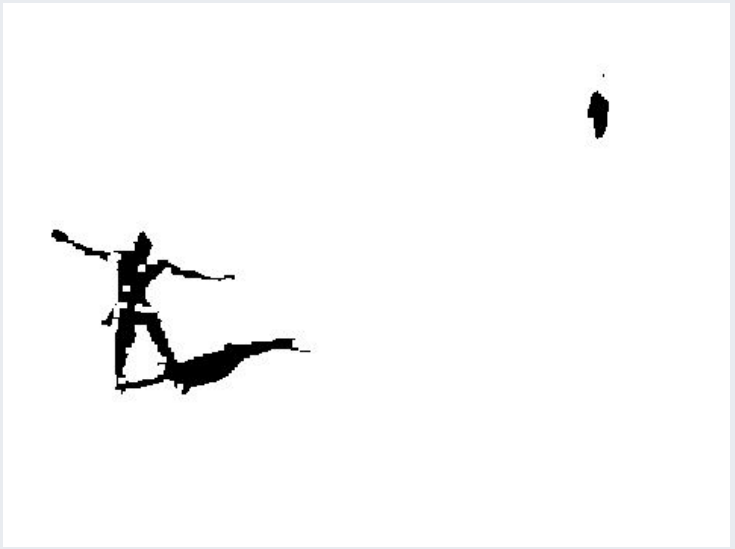
- Put virtual and real objects into a 3D computer generated space
- Problem: control control the occlusions between actors and virtual elements



Vision – Video Surveillance

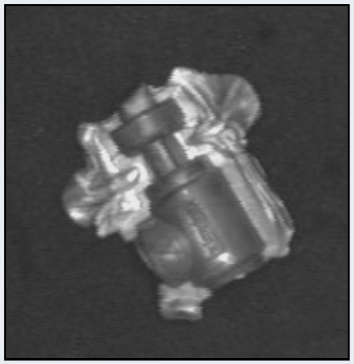
- Applications: outdoor, urban traffic, measurement of speed of cars





Vision – Object Recognition

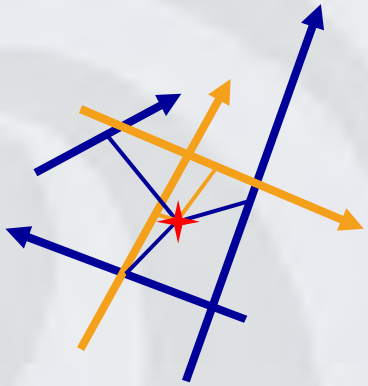
- Real time analysis



Road sign recognition, Railway sign detection

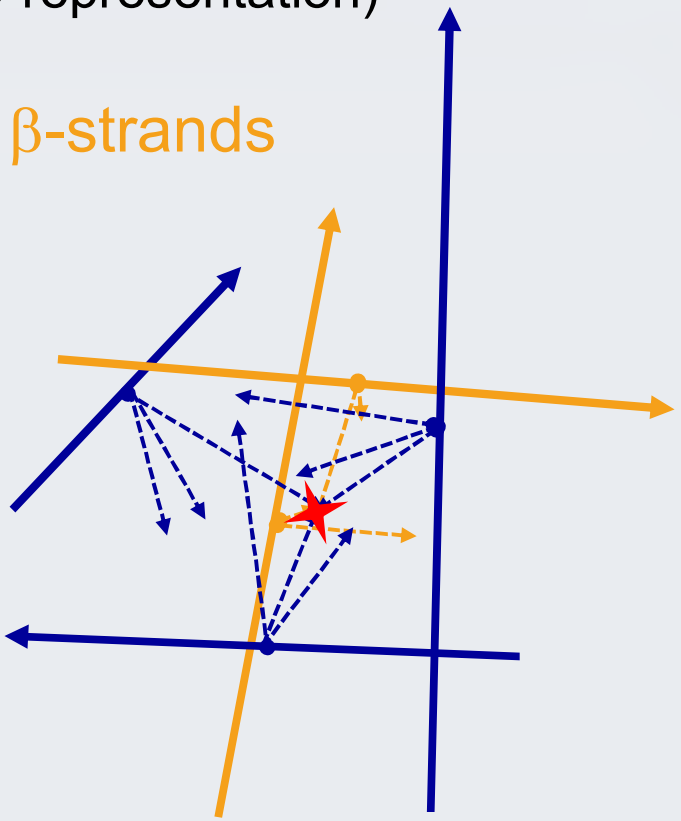
- Applications: intelligent vehicle, visual support to drivers

In the way of GHT (simplified 2D representation)



Query protein

α-helices and β-strands

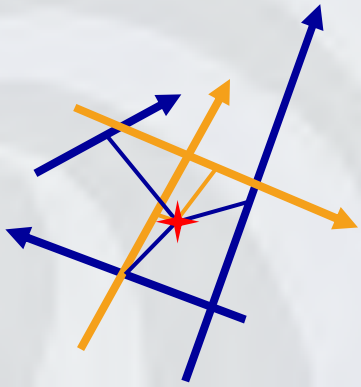


Mapping rule

Votes Space

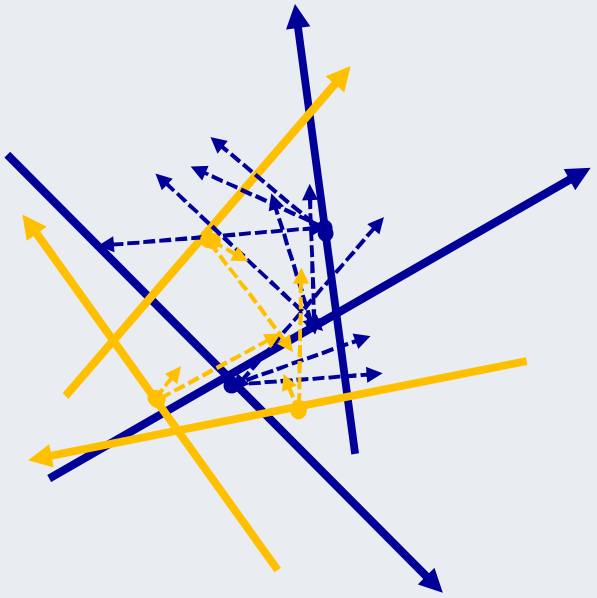
Pattern Recognition in Proteomix

In the way of GHT (simplified 2D representation)



Query protein

α-helices and β-strands

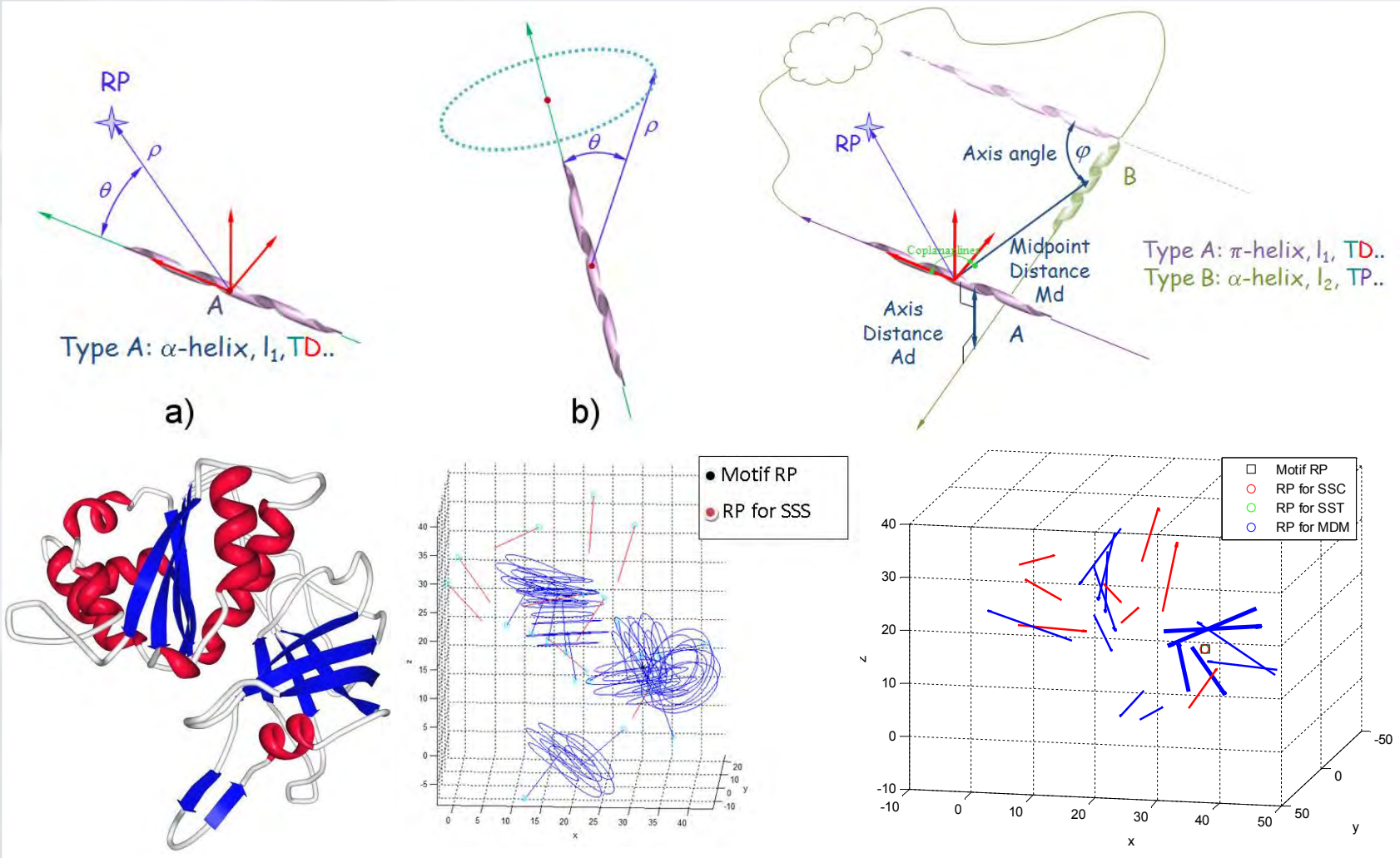


Votes Space

Mapping rule

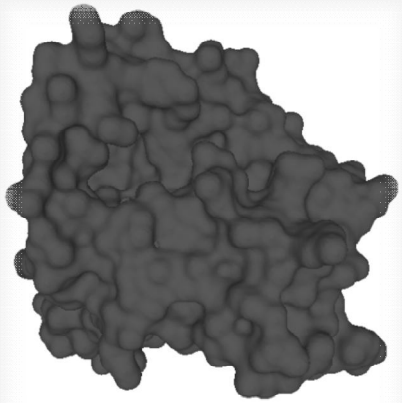
Pattern Recognition in Proteomix

- Search for structural motifs by the Hough Transform

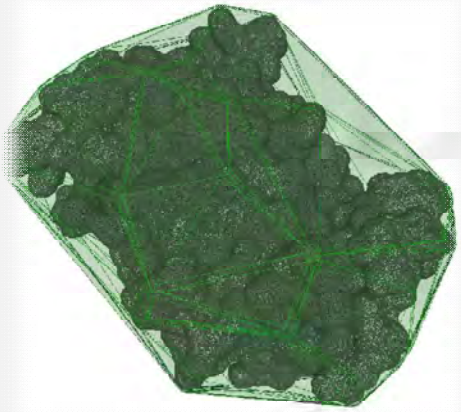


Pattern Recognition in Proteomix

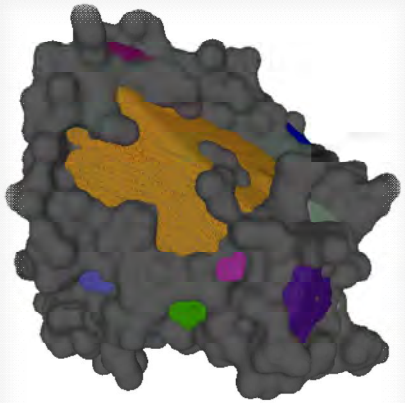
- Protein-ligand interaction



SAS

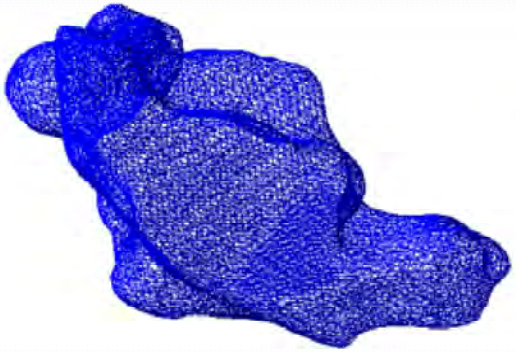


convex hull



pockets

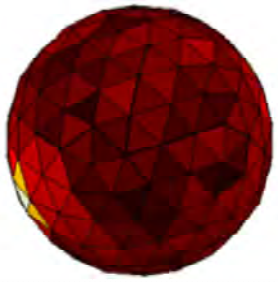
main pocket



ligand

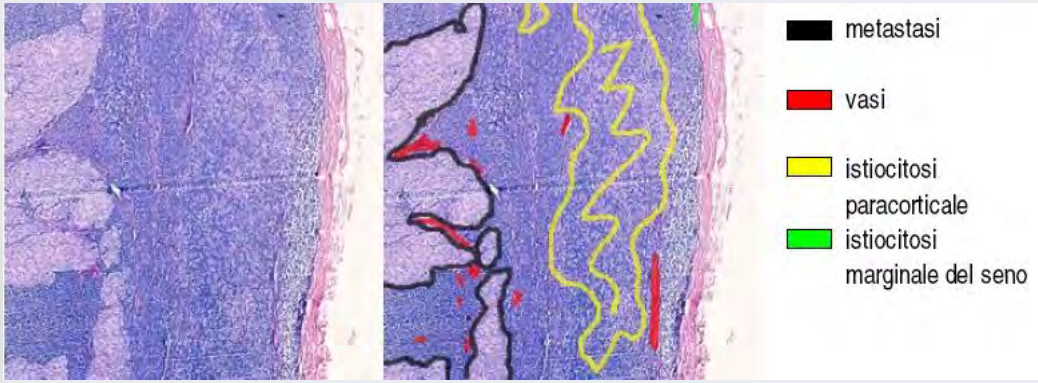
pocket EGI

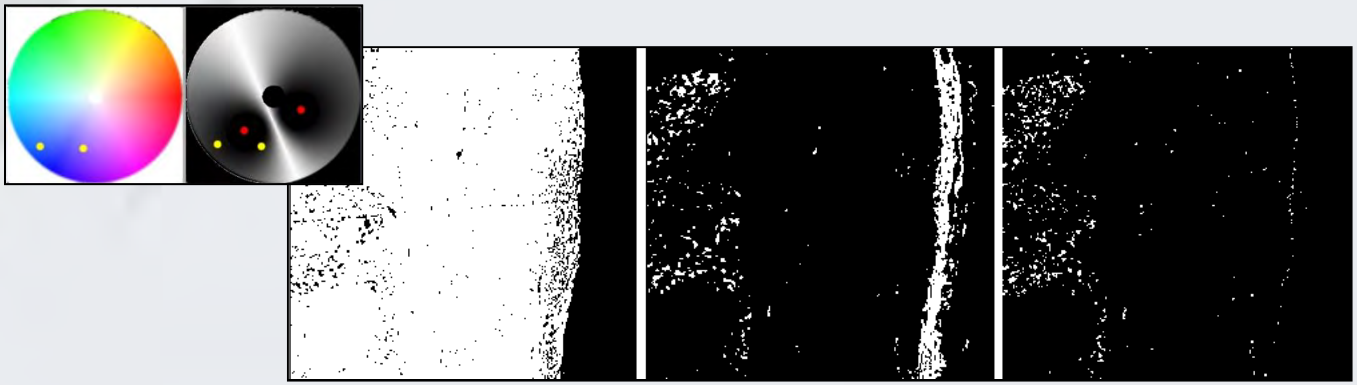
ligand EGI



Vision – Object Recognition

- Explorative screening systems for histologic preparates images





QCT/FEA Models of Proximal Femurs: Image-based Mesh Generation

Episode III : 2011-12-05

A joint effort with Mayo Clinic and STMicroelectronics

A Quick Recap (Episode I : 2010-09-13)

- Method Overview

3074R_NORMAL.003
57050L_OSTEOPOROTIC.003
908404R_NORMAL.003

↓

3D Image Pre-Processing

Dilation
Erosion
} *With a fixed-size structuring element (21 voxels in diameter)*
Gaussian Smoothing

↓

3074R_NORMAL.003.DE21.G1
57050L_OSTEOPOROTIC.003.DE21.G1
908404R_NORMAL.003.DE21.G1

↓

Selecting Level Iso-Surfaces: Water-tightness Analysis

SOAM Algorithm
(on a time-variant sweep over surfaces)

↓

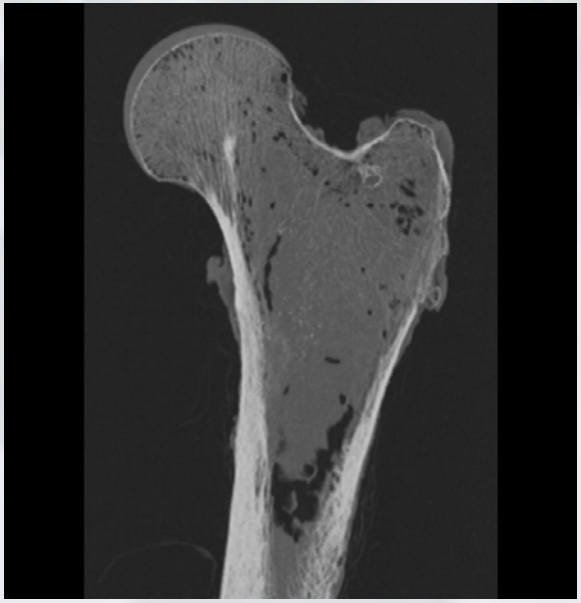
3074R_NORMAL.003.DE21.G1.L645.OuterShell.stl
57050L_OSTEOPOROTIC.003.DE21.G1.L400.OuterShell.stl
908404R_NORMAL.003.DE21.G1.L645.OuterShell.stl

Highest threshold level (HU) at which water-tightness is preserved

A Quick Recap (Episode I : 2010-09-13)

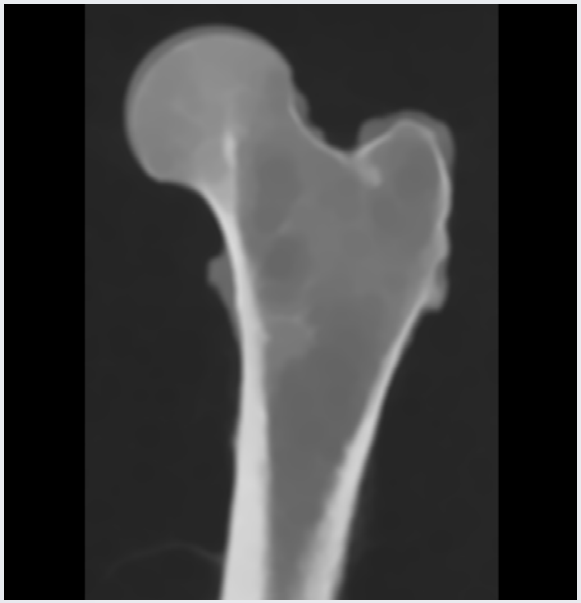
- Image Pre-Processing

Original



Dilation + Erosion

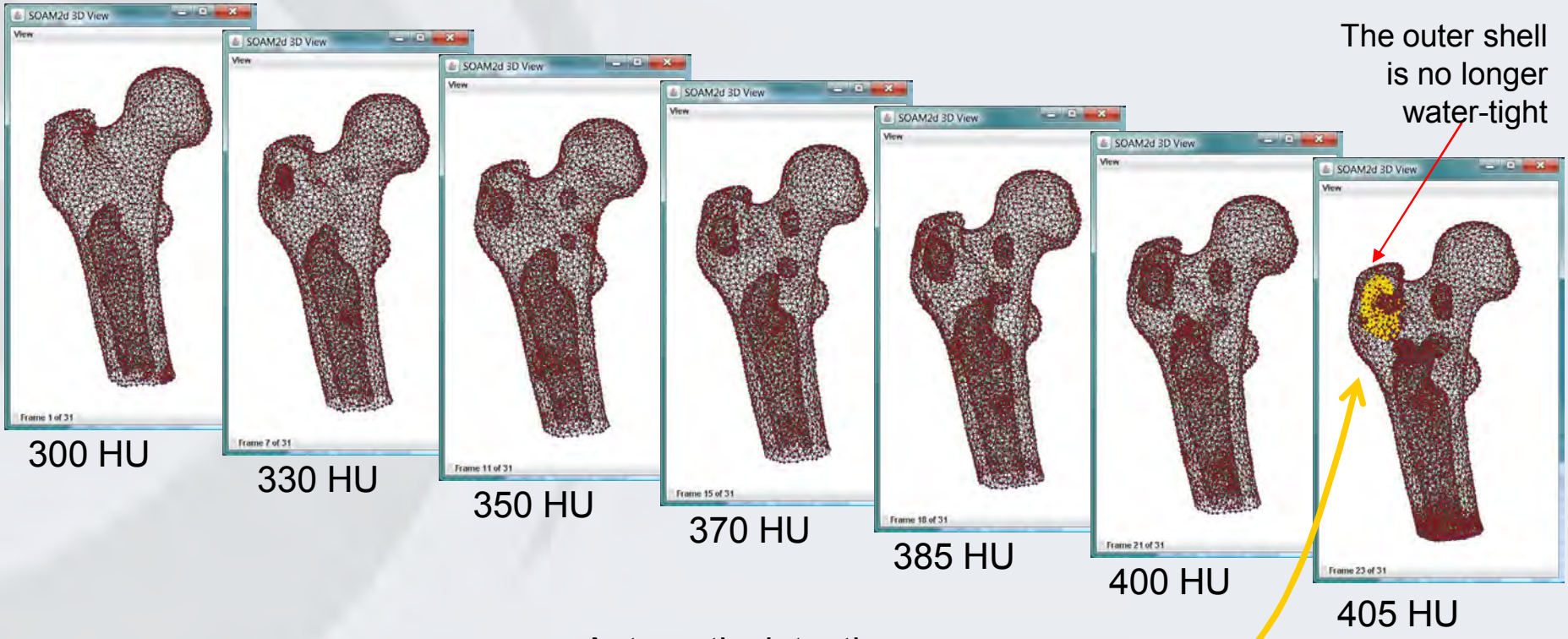
Gaussian Smoothing



A Quick Recap (Episode I : 2010-09-13)

- Selecting Iso-Surfaces

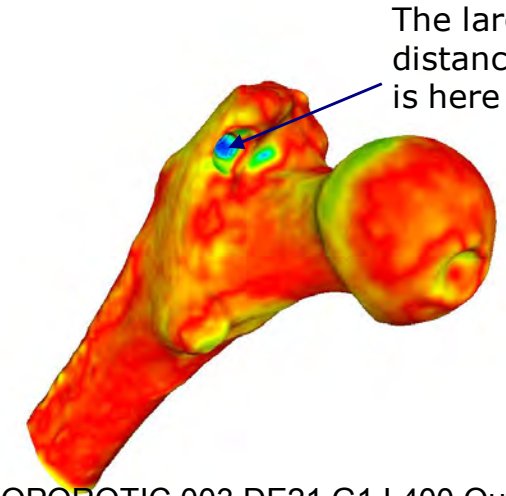
Time-variant 3D Triangle Mesh (SOAM algorithm)



A Quick Recap (Episode I : 2010-09-13)

- Results and Validation
  - Comparison of manually-segmented meshes with automatically-extracted ones
    - Using MeshLab and the METRO sampling algorithm (Cignoni et al. 1996)

3074R_NORMAL.003.DE21.G1.L645.OuterShell.stl



57050L_OSTEOPOROTIC.003.DE21.G1.L400.OuterShell.stl

908404R_NORMAL.003.DE21.G1.L645.OuterShell.st

What's New in Episode III (2011-12-05)

- Results

Data Set	SES (voxels)	Hausdorff Distance	Test Set: Christina	Ian	Rachel	Viorel	Vishwas
2675L_OSTEOPOROTIC	27	min	0.000083	0.000008	0.000033	0.000004	0.000043
		max	4.324452	4.389863	5.204681	4.679758	4.324683
		mean	0.207373	0.212788	0.388494	0.203164	0.228532
		RMS	0.302847	0.340244	0.513313	0.325063	0.347568
3075R_OSTEOPOROTIC	21	min	0.000005	0.000033	0.000017	0.000003	0.000017
		max	3.431049	3.556161	3.447876	3.972922	3.305173
		mean	0.210947	0.218107	0.228353	0.221990	0.205873
		RMS	0.304615	0.294806	0.336898	0.336424	0.284227
3082R_NORMAL	11	min	0.000000	0.000015	0.000026	0.000015	0.000019
		max	1.653172	2.164078	2.157052	2.893867	2.837292
		mean	0.187067	0.203129	0.205705	0.211246	0.191417
		RMS	0.224435	0.240362	0.246189	0.272521	0.235769
3086L_OSTEOPENIC	17	min	0.000033	0.000015	0.000007	0.000047	0.000005
		max	2.476942	2.343065	2.384681	2.537388	2.344930
		mean	0.193372	0.332697	0.194768	0.192596	0.181760
		RMS	0.264863	0.405104	0.246869	0.274163	0.245572
32471L_OSTEOPENIC	19	min	0.000028	0.000025	0.000016	0.000011	0.000012
		max	2.985517	3.311095	3.283424	3.212310	3.078978
		mean	0.175558	0.190040	0.192348	0.180661	0.175404
		RMS	0.226875	0.242556	0.252586	0.245195	0.226122
49545L_NORMAL	19	min	0.000023	0.000017	0.000008	0.000002	0.000008
		max	2.647166	2.849197	3.222083	3.377649	2.504953
		mean	0.217975	0.252744	0.265175	0.237507	0.226623
		RMS	0.281108	0.353121	0.362309	0.354369	0.296675
57050L_OSTEOPOROTIC	21	min	0.000056	0.000008	0.000020	0.000021	0.000057
		max	3.192670	3.334923	3.448210	3.658764	3.100749
		mean	0.273015	0.260918	0.277172	0.265849	0.257542
		RMS	0.376943	0.345752	0.368883	0.389208	0.350470
902514R_NORMAL	9	min	0.000055	0.000007	0.000006	0.000004	0.000011
		max	1.506914	1.782926	1.748946	1.922011	1.564114
		mean	0.160309	0.183127	0.182712	0.163791	0.160182
		RMS	0.187380	0.214271	0.216817	0.193727	0.186413
902893R_OSTEOPENIC	17	min	0.000006	0.000070	0.000006	0.000012	0.000012
		max	2.227875	3.701558	2.379014	2.365911	2.314178
		mean	0.185084	0.254320	0.189767	0.188475	0.182469
		RMS	0.259320	0.401195	0.261496	0.255980	0.259410

Data fusion segmentation

- Our approach uses these data sources:
  - A standard RGB camera
  - A TOF camera
- Possible applications:
  - People Tracking
  - Human – Machine Interaction (HCI)
  - 3D reconstruction
  - Augmented reality
  - Etc.

Time-of-Flight Cameras

- New kind of sensors which allow for depth measurements using a single device with no mechanical parts
- Use laser light in near infrared to measure distances between the camera and the objects in the scene
- Why to use TOF cameras:
  - Direct measure of the distance without additional computation
  - Can work at real-time
  - No need of external illumination
  - Can measure distance with any kind of background
  - No interaction with artificial illumination

Cameras Specifications

SwissRanger™ SR3000

Moduled light ToF camera

55 active leds that emits in near infrared spectrum (~850nm) at 20 Mhz

Max range without ambiguities: 7.5m

Field of view 47.5 x 39.6 degrees

Can reach 18-20 fps at QCIF resolution (176x144 pixels)

Supply two image per frames:

Distance Map

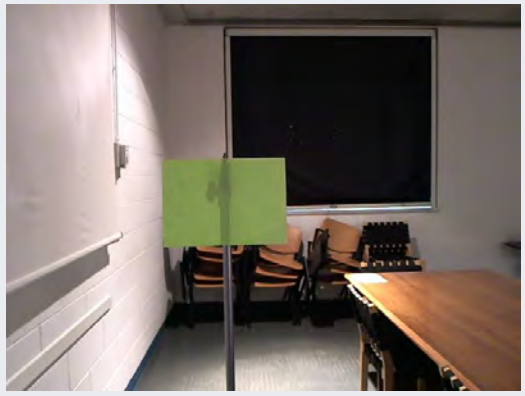
Intensity Map

Logitech HD Pro Webcam C910

RGB HD Camera

Max resolution 1080p, sets for the experiments at 640x480 pixels

Max frame rate 30fps, set to 18fps to syncronize with SR3000 speed



RGB

ToF based Segmentation

- Two main phases:
  1. A *thresholding* of the distance map based on the correspondent values of intensity map
  2. A *region growing* on the filtered intensity map: seeds planted in the peak of intensity map
- No learning phase
- No a priori knowledge of the background
- Shape of the objects does not influence the result

Intensity Map

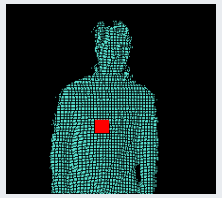
Depth Map

SR3000 Input

Intesity based Thresholding

Filtered Depth Map

Region Growing



Segmentation Result

Hand detection as an example

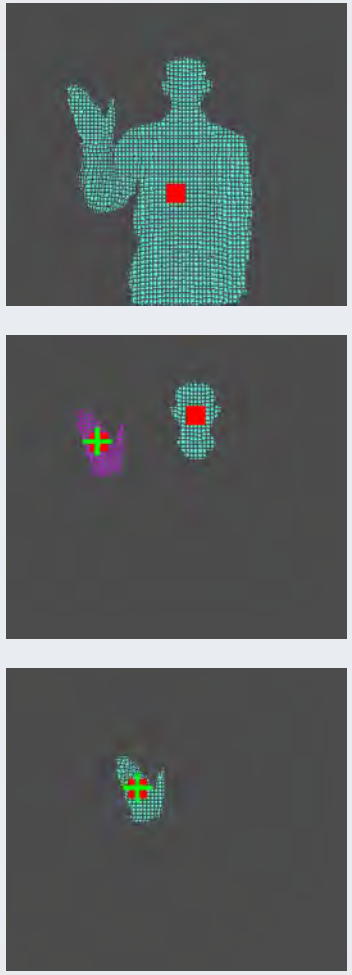
- The foreground segmentation significantly reduce the interest area
- Sub-segmentation is achieved filtering the found cluster with color data (converted in HSV) :

$$W = \{ y: -10° < H_y < 10°, S_y > TH_s, V_y > TH_v \}$$

- An initialization phase is use to set the min and max distance threshold δ_{min} and δ_{max} useful for excluding the head and clothes with skin like color
- New region growing formula:

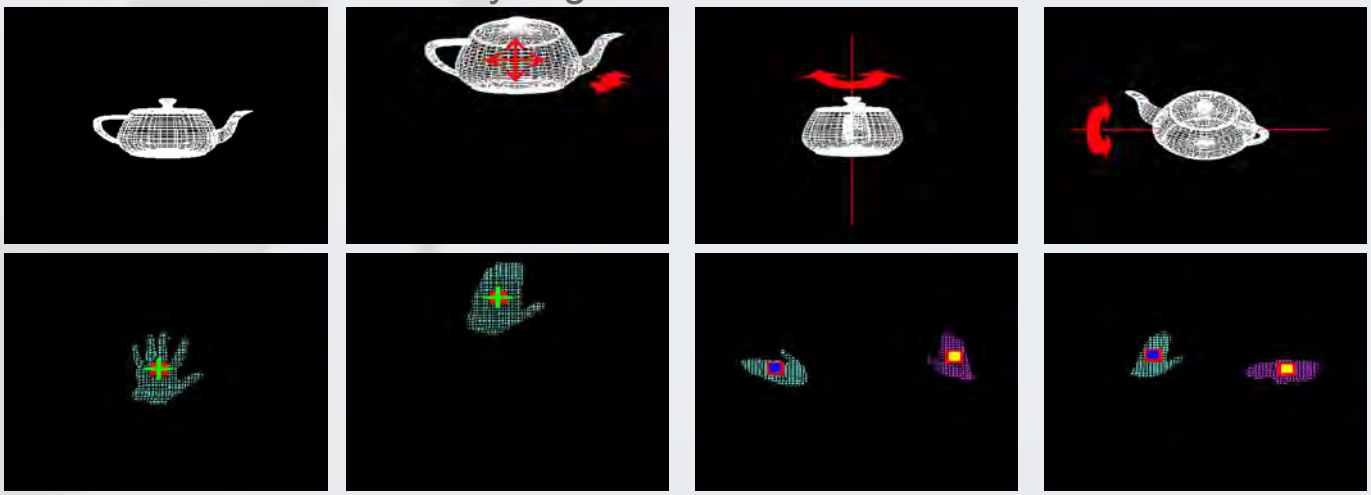
$$\{ x \in C, S(x, y) < \theta, I_y \in L, \ y \in W, \delta_{min} < D_y < \delta_{max} \} \rightarrow \{ y \in C \}$$

- Procedure totally automatic based only on chromatic charateristics of the image
- Totally indipendent by the shape or by the position of the hand



Interaction with 3D objects

- Two type of movements:
  - Translation – triggered when the hands are not alligned or there is only one hand (a cross mark the active cluster)
  - Rotation – triggered when the hands are alligned (two squares mark the hand: blue for Y-axis controller; yellow for X-axis controller)
- Gesture recognition based only on geometrical constraints (no need of a learning phase):
  - Translation – direct mapping of the hand coordinate to object coordinate
  - Rotation – achieved analyzing the moment of inertia of the hand/s



Gestures and correspondent object movement