A VIRTUAL ENVIRONMENT TOOL FOR BENCHMARKING FACE ANALYSIS SYSTEMS

BEFIT 2012, October 13, Firenze, Italy

Mauricio Correa, Javier Ruiz-del-solar, Rodrigo Verschae

Department of Electrical Engineering,
Advanced Mining Technology Center,
Universidad de Chile

(our) Motivation

Robocup competitions

- Standard Platform League
- Robocup @home League

(our) Motivation

Robocup @home League competition
- Who Is Who Challenge

(our) Motivation

Face Analysis in Service Robots

- Testing is very time consuming when using robots in real conditions.
- Temporal and Spatial information might be useful.
- Active vision methods need to be evaluated.

Motivation

Current datasets and benchmarking frameworks:

- are design for static environments
- in general are used for analyzing methods under controlled conditions
- They consider different
  - number of persons,
  - camera sensors and image acquisition conditions,
  - aspects such us illumination invariance, aging, expression invariance, etc.

But they do not :

- provide real-world testing conditions that include spatiotemporal context,
- allow the sensing agent to actively change its point of view.

Proposed Virtual Enviroment

Key elements:

- a simulator
- new database with real face and background images,
- the images are taken under real-world conditions

Some applications:

- Robotics
- Security / video, e.g. select best view for recognition
- Method analysis

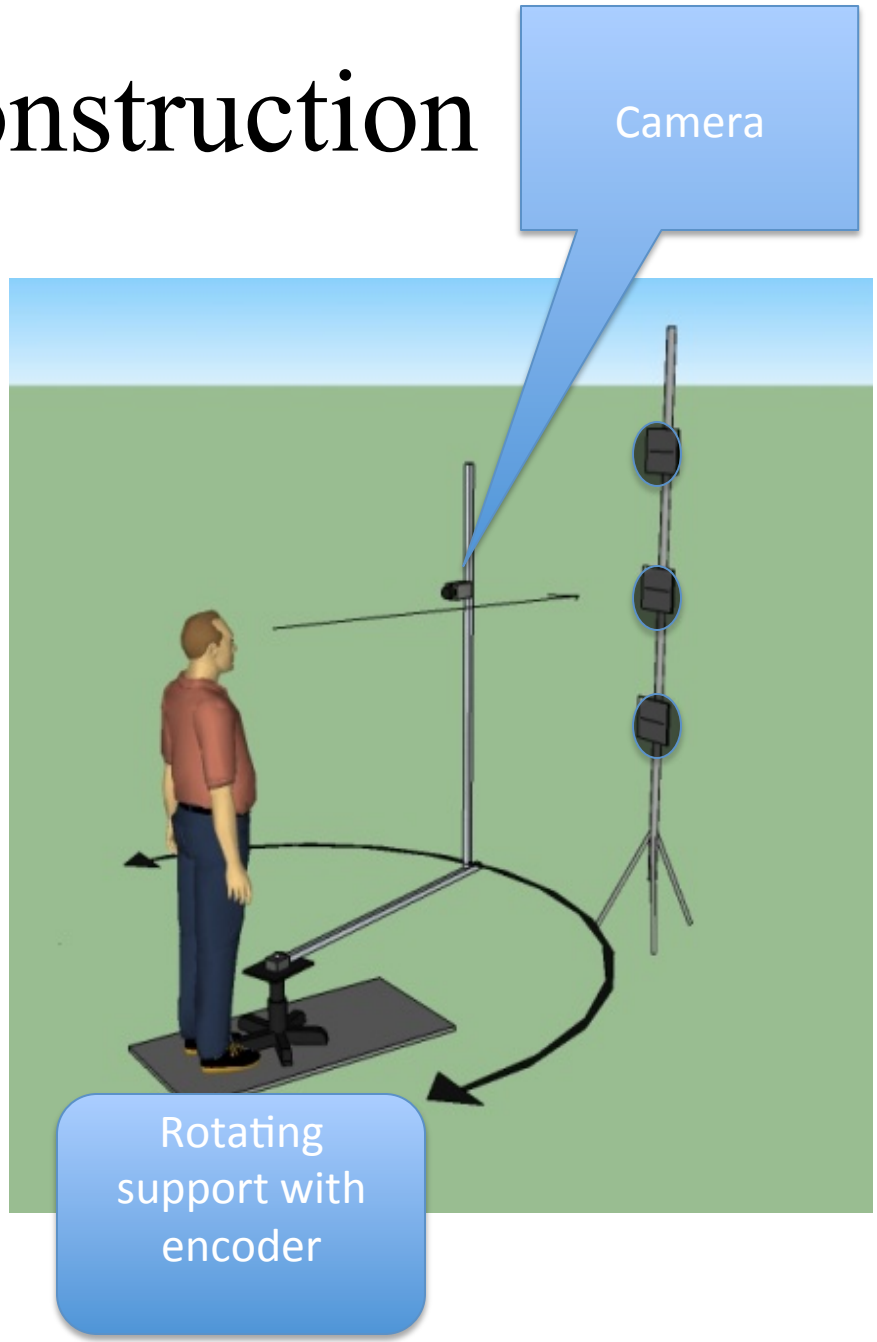
Outline

- Database Construction
- Virtual Environment
- Usage Example
- Conclusion & Future Work

Database Construction

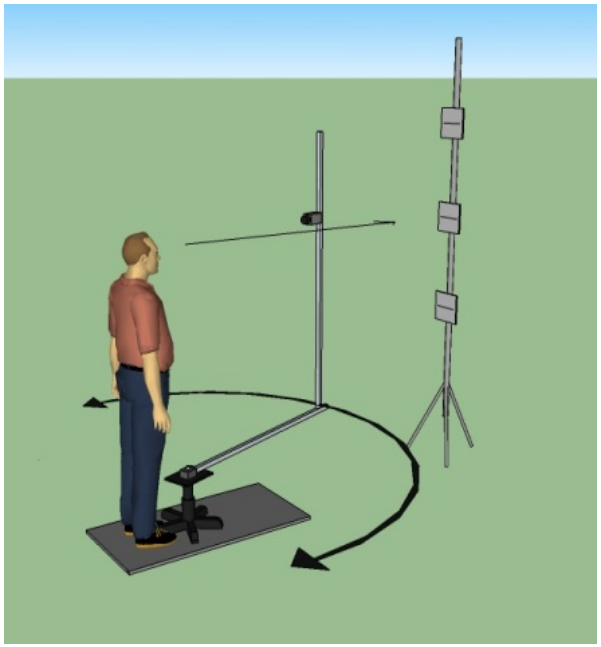
- The person under scan is in a still position
- The camera is placed at the same height than the person's face and at a fixed distance (140 cm)

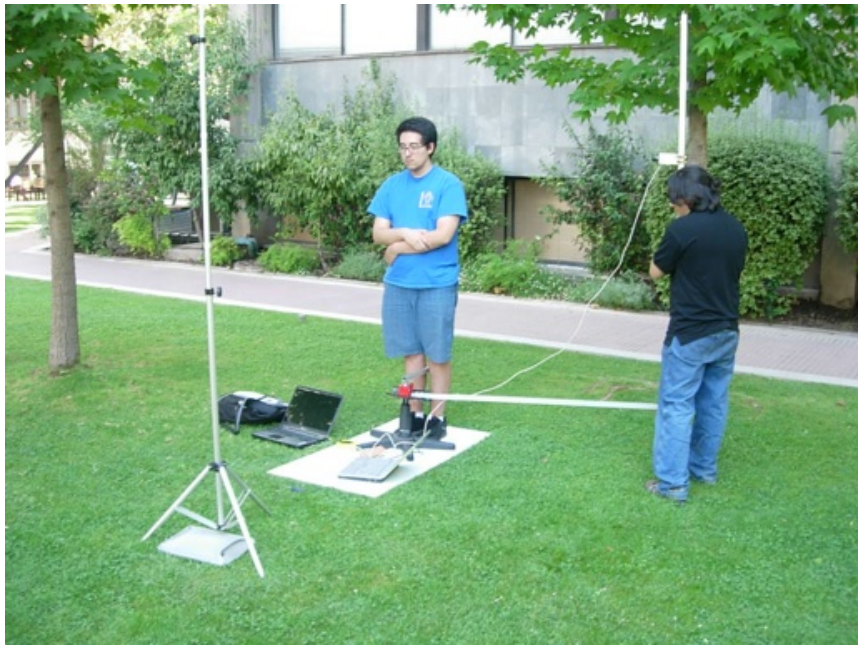
- Face images are acquired at different yaw and pitch angles
  - The yaw angle range is -120° to 120°, with a resolution of 2°, which gives 121 images.
  - For each different yaw, 3 pitch angles are considered (-15°, 0°, and 15°).
- We also capture the background



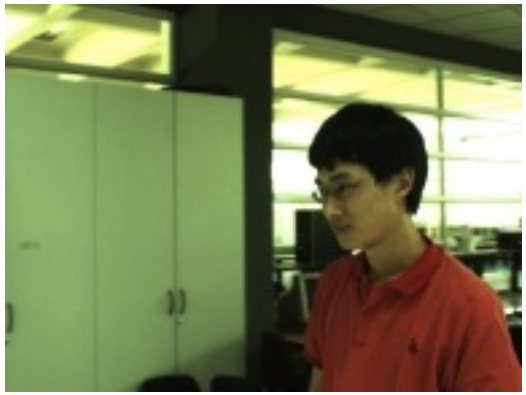
Database Construction

- The DB contain 50 persons in two different locations (*indoor* y *outdoor*).
- 726 images per person are registered (121 yaw x 3 pitch x 2 locations).
- We use a 1280 x 960 pixels CCD camera (DFK 41BU02 model).
- In the frontal image, the face's size is about 200x250 pixels.
- The scanning process takes 25 seconds.



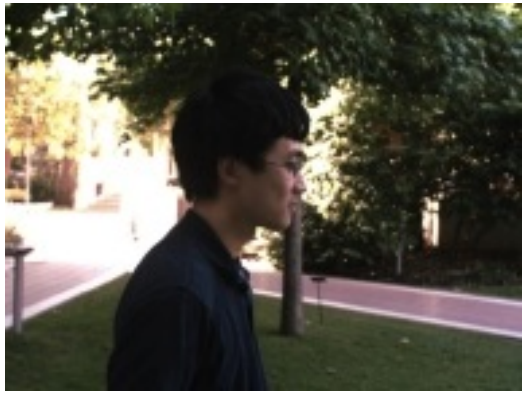


Database Examples



Yaw: 50° Pitch: -15

Yaw: 0° Pitch: 15



Yaw: 90° Pitch: -15

Yaw: 30° Pitch: 0

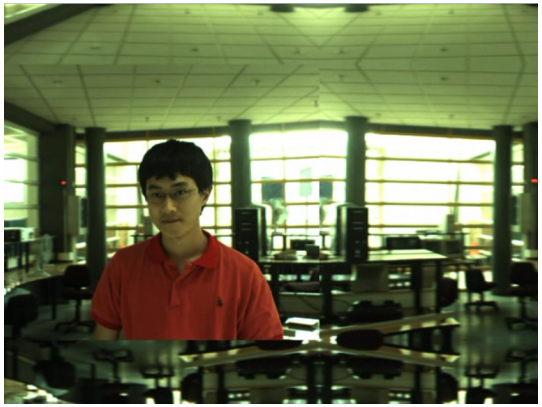
Database Video Example

Database construction

- Ground truth
  - For each captured Image we know the *yaw* angle as obtained from the encoder and *pitch* angle
  - We annotated the position of the eyes and nose of all frontal faces
  - We annotated the positions of the eyes and nose for some rotated faces and we interpolated the missing annotations

Virtual Enviroment

- The virtual environment composes the current view of the agent over time.
- The agent could move, approach the subject, change its view point, etc.
- They may be one or more subjects in the environment.



Virtual Enviroment

It consist of

- an agent (e.g. a robot)
- a simulator, which composes images to be seen by the agent
- a navigation module, which moves the agent in the global map
- a trajectory module, which defines the kinds of trajectories the agent can follow.

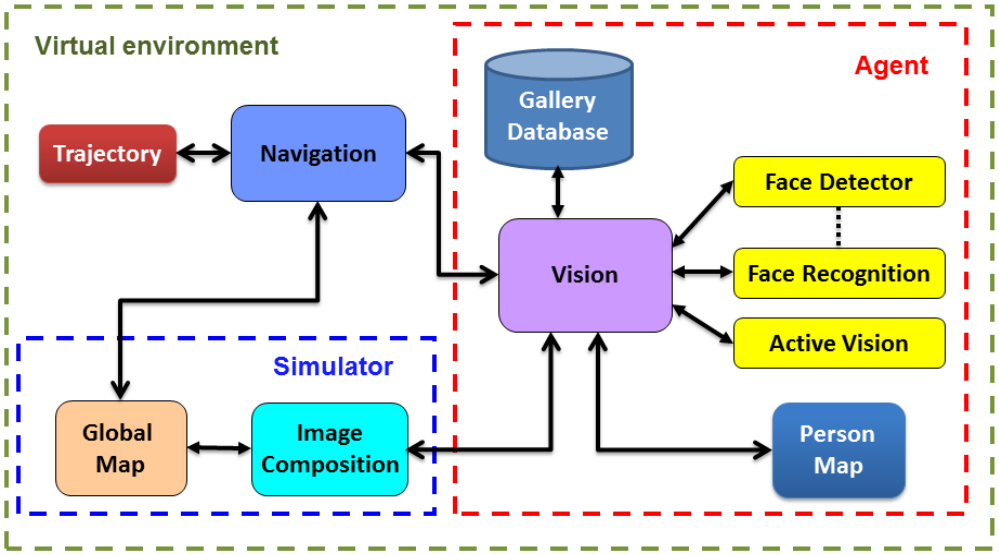
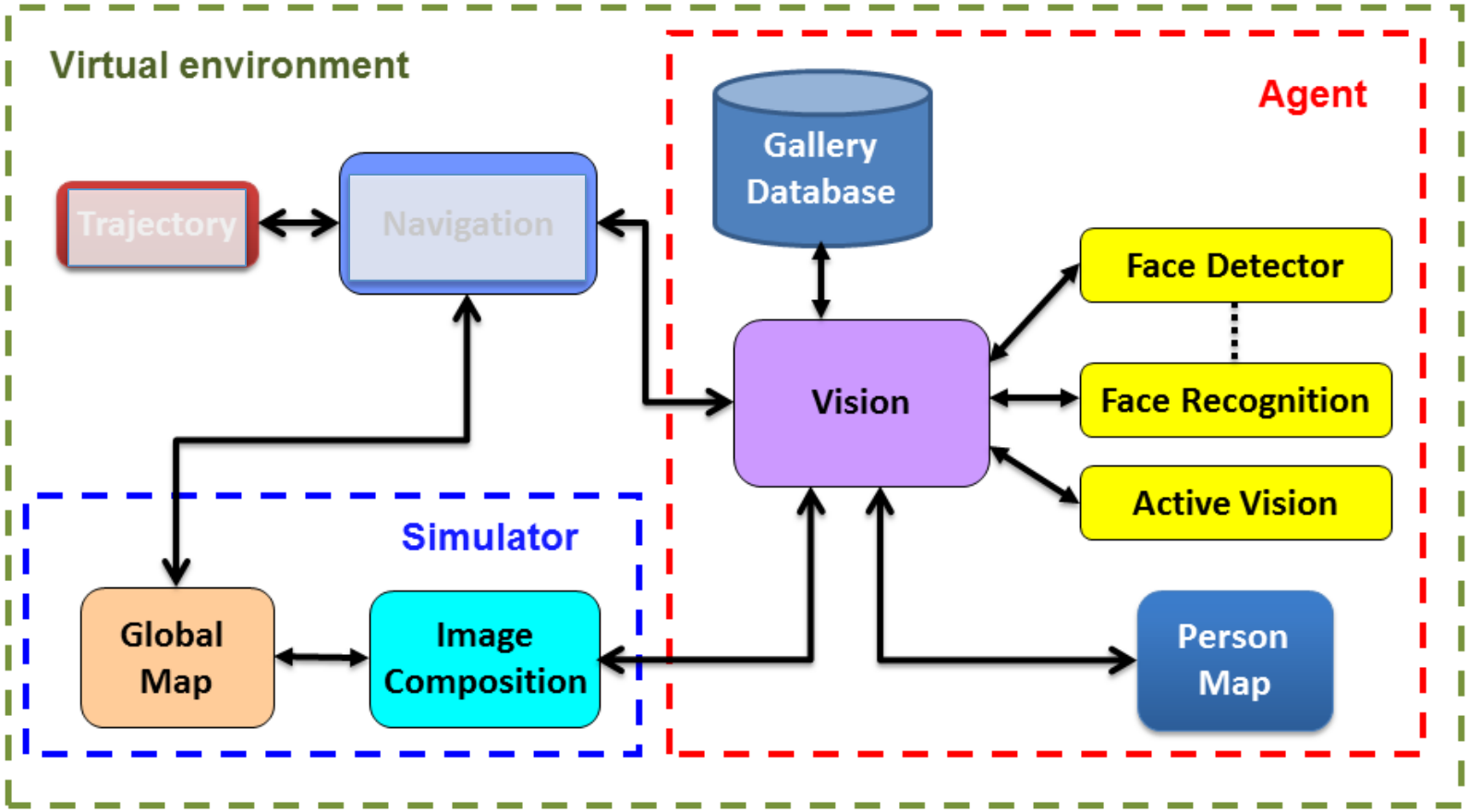


Diagram of the virtual environment



Global Map

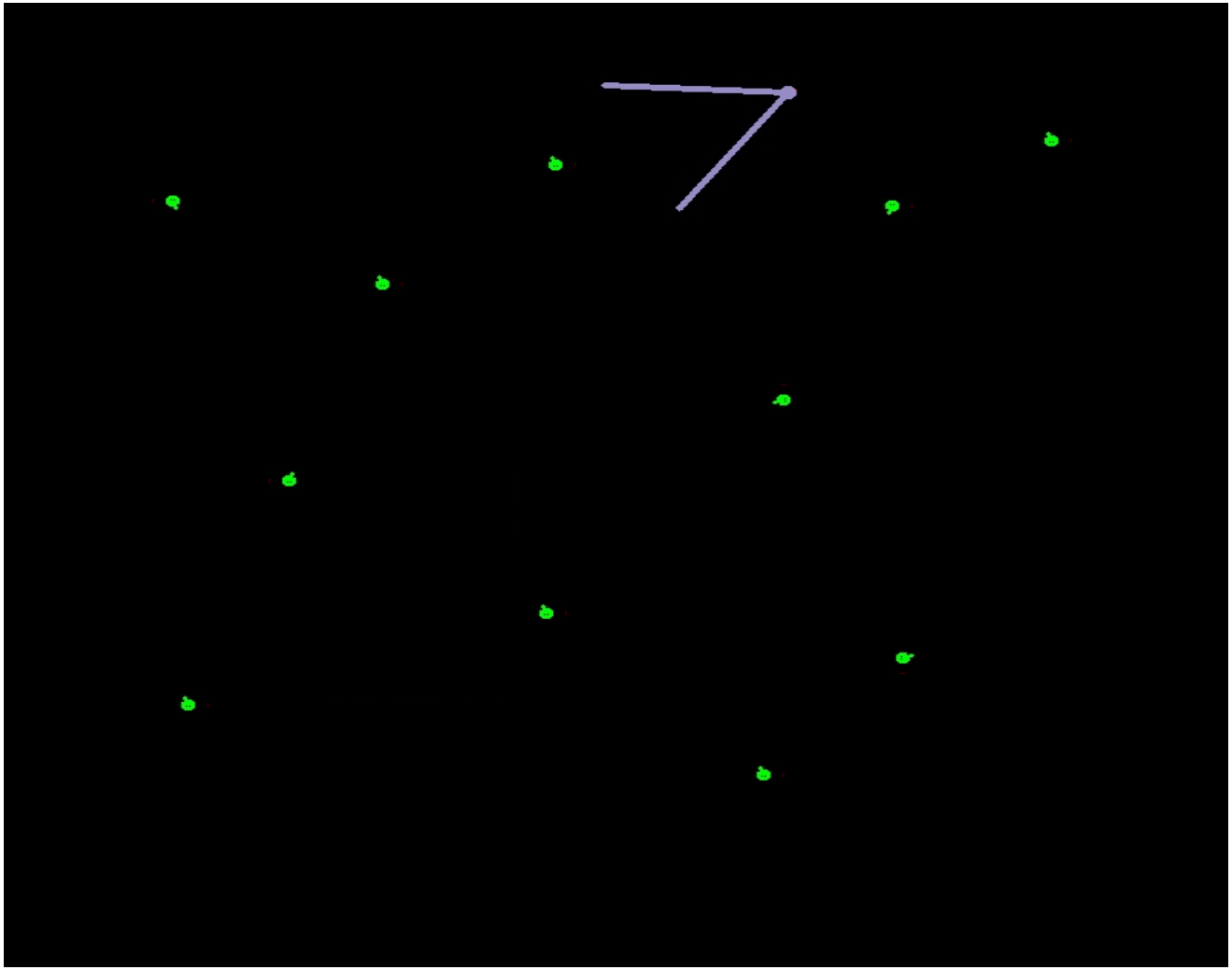


Image composition

The current view is composed using images from the database and the relative location of the subject and the agent.

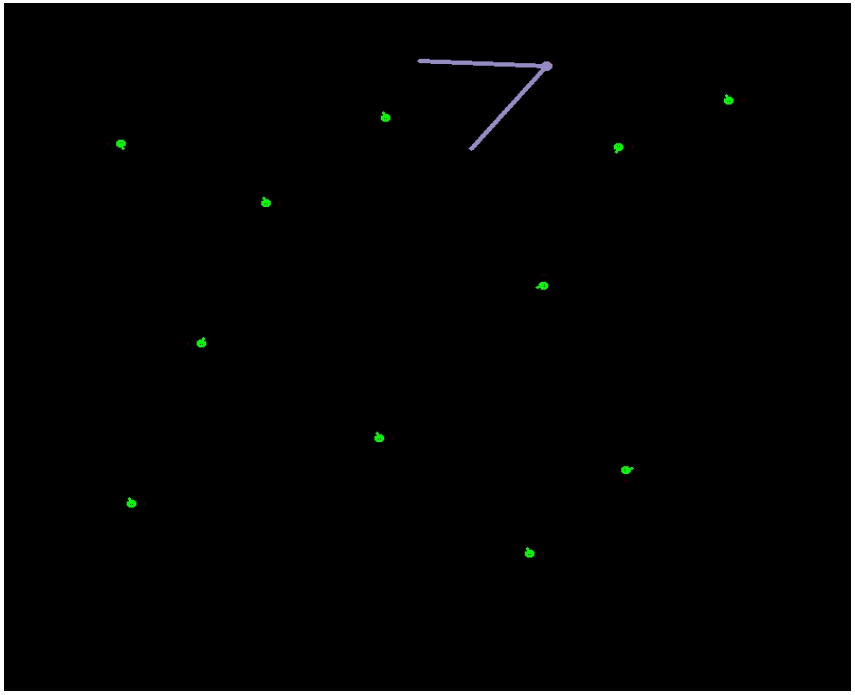
We consider at most one subject per view (the closest one).

Trajectories

- The virtual environment provides three different kinds of trajectories:
  - **Free navigation**
  - **Constrained navigation**
  - **Predefined navigation**

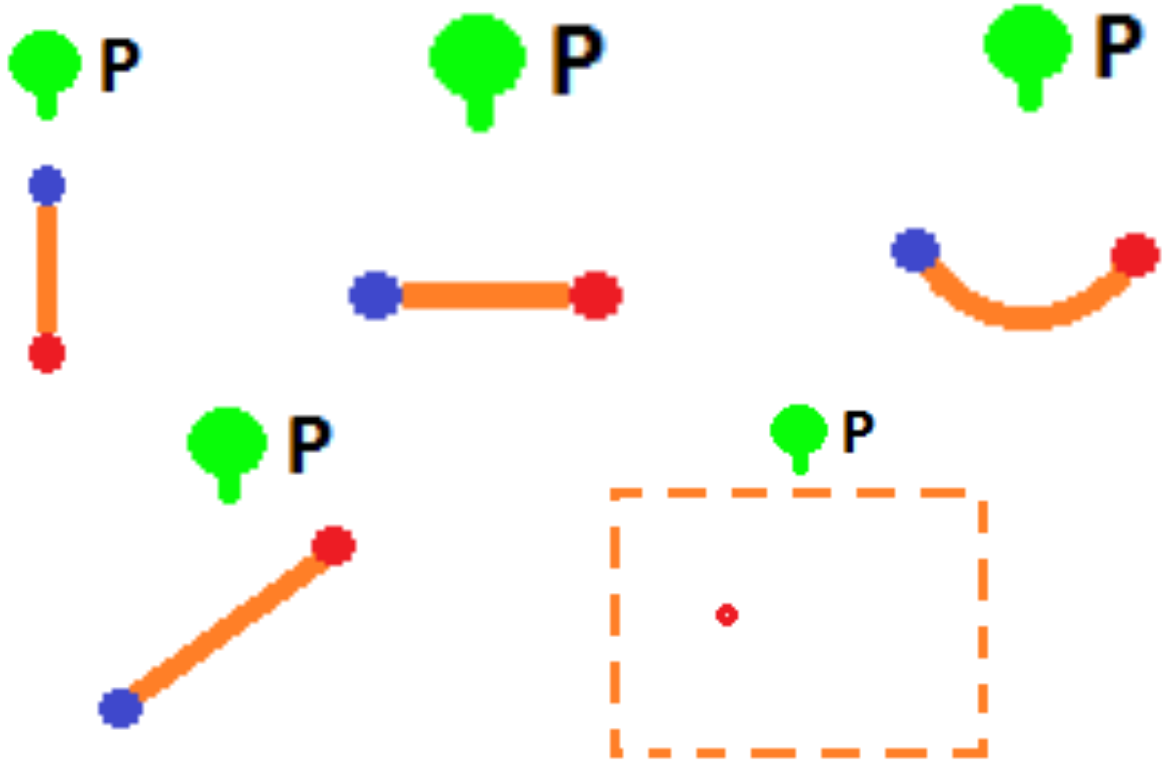
Trajectories

- The virtual environment provides three different kinds of trajectories:
  - **Free navigation**



Trajectories

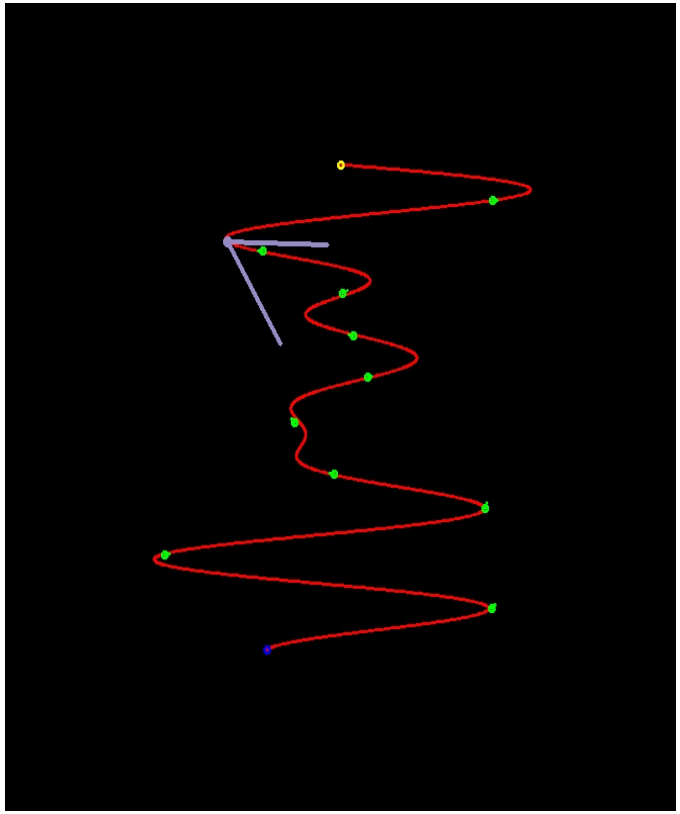
- The virtual environment provides three different kinds of trajectories:
  - **Constrained**



Example

Trajectories

- The virtual environment provides three different kinds of trajectories:
  - **Predefined navigation**



Example

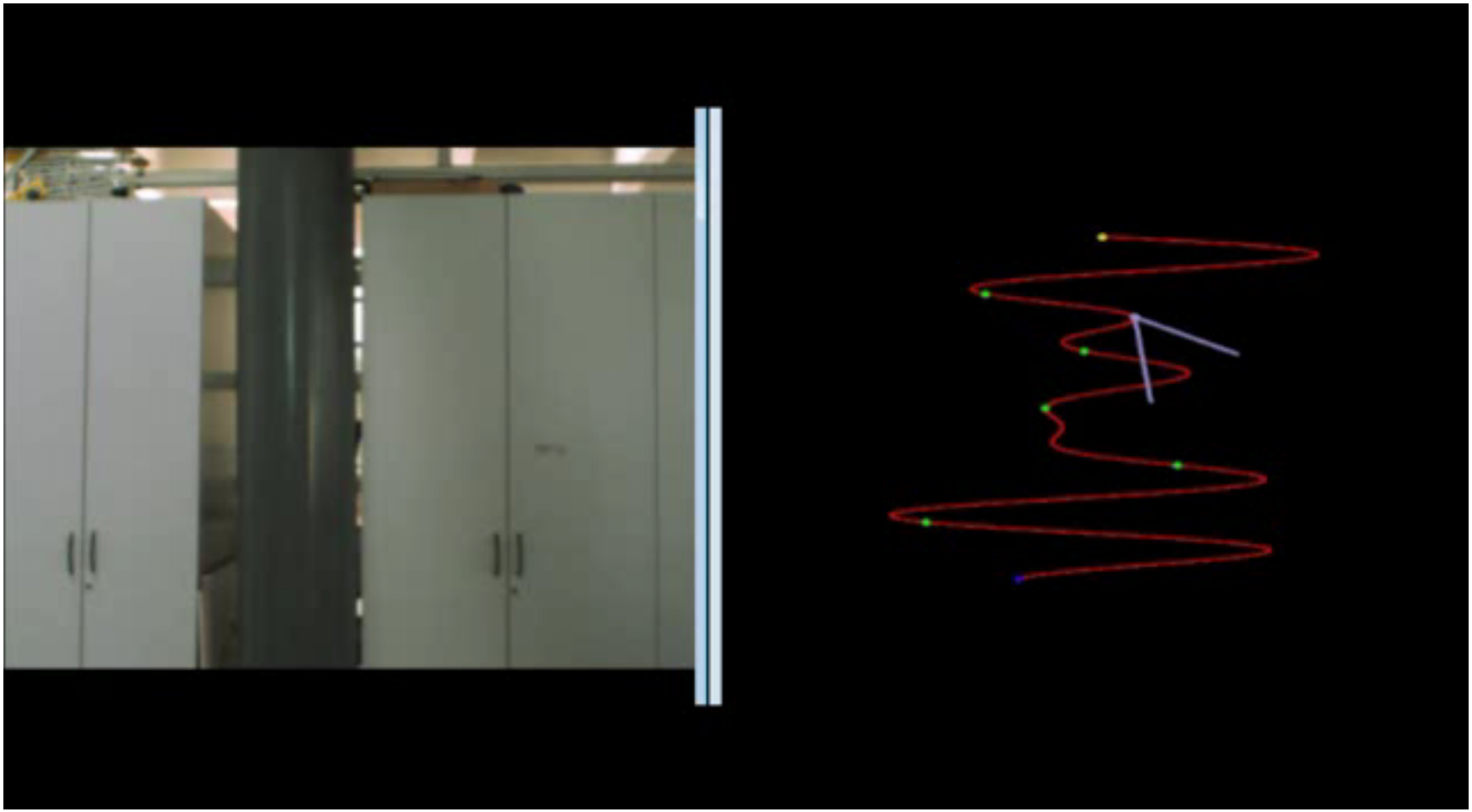
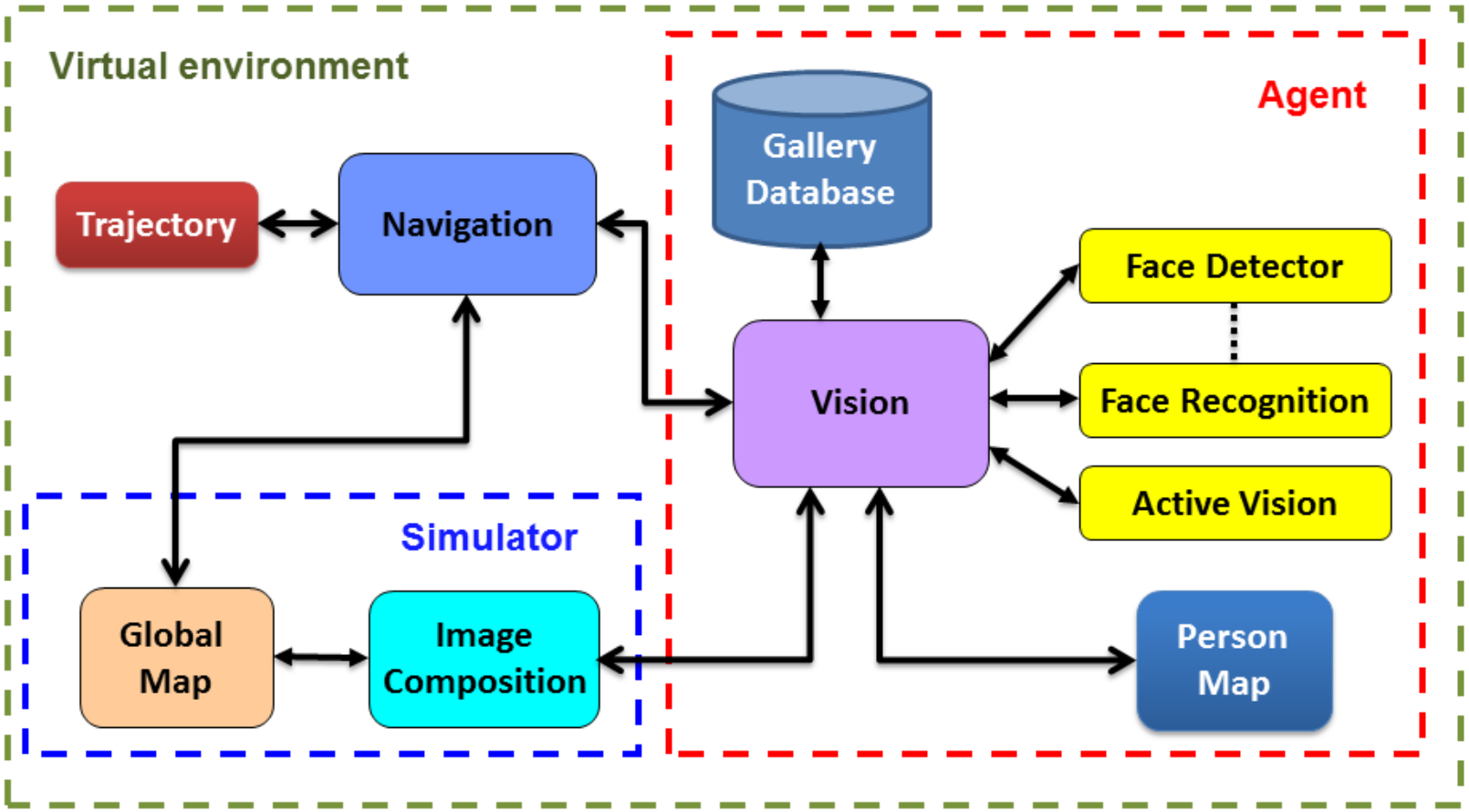


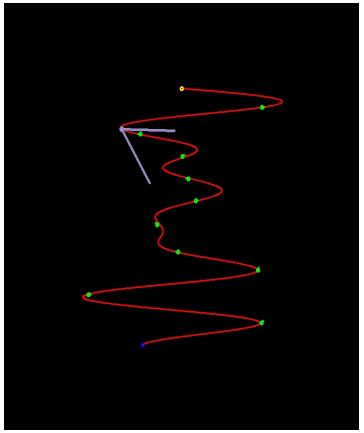
Diagram of the virtual environment



Usage Example 1: Face Detection Evaluation

	$\theta^p_{\max} = 0$															
	$\theta^y_{\max} = 0°$				$\theta^y_{\max} = 20°$				$\theta^y_{\max} = 40°$				$\theta^y_{\max} = 60°$			
	Indoor		Outdoor		Indoor		Outdoor		Indoor		Outdoor		Indoor		Outdoor	
	DR	FP	DR	FP	DR	FP	DR	FP	DR	FP	DR	FP	DR	FP	DR	FP
OpenCV Cascade	97.4	2	78.3	5	92.1	3	87.0	3	71.1	4	65.2	8	47.4	1	82.6	4
Cascade1	89.5	4	100	0	86.8	5	100	0	71.1	6	100	0	55.3	9	69.6	3
Cascade2	100	0	100	0	97.4	0	100	0	73.7	1	95.7	0	50.0	2	73.9	1
	$\theta^p_{\max} = \pm 15$															
	$\theta^y_{\max} = 0°$				$\theta^y_{\max} = 20°$				$\theta^y_{\max} = 40°$				$\theta^y_{\max} = 60°$			
	Indoor		Outdoor		Indoor		Outdoor		Indoor		Outdoor		Indoor		Outdoor	
	DR	FP	DR	FP	DR	FP	DR	FP	DR	FP	DR	FP	DR	FP	DR	FP
OpenCV Cascade	94.7	2	95.7	1	84.2	4	69.6	7	65.8	6	78.3	5	60.5	6	65.2	8
Cascade1	94.7	2	100	0	84.2	4	100	0	60.5	11	100	0	42.1	12	69.6	4
Cascade2	97.4	0	100	0	94.7	0	100	0	63.2	1	95.7	0	52.6	3	82.6	1

Usage Example 2: Face Recognition Evaluation



- Automatic Face Detection vs Ground Truth
- With active vision vs without active vision
- Gallery *Online* vs Gallery *Offline*

- Predifined trajectories
- 10 runs, and 10 persons per run

Experiment	Face Detector	Active Vision	Gallery	Persons added to the gallery	Persons correctly detected [%]	Recognition [%] (out of all subjects in the scene)	Recognition [%] (out of the detected subjects)
1	Cascade	No	Offline	10.0	84.0%	----	78.4% (*)
1	Cascade	No	Offline	10.0	84.0%	73.0%	86.8%
2	Cascade	No	Online	14.8	84.0%	59.0%	70.2%
3	Cascade	Yes	Offline	10.0	84.0%	**78.0%**	**92.9%**
4	Cascade	Yes	Online	12.2	84.0%	73.0%	86.9%
5	Ground Truth	Yes	Offline	10.0	**100.0%**	**92.0%**	**92.0%**
6	Ground Truth	Yes	Online	10.0	**100.0%**	90.0%	90.0%

Conclusion

- The testing tool combines the use of a simulator with real face and background images taken under real-world conditions.
- This tool is useful for testing face analysis systems, in particular for comparing different face detection and recognition systems under similar conditions.
- The applicability of the proposed tool is validated with an example.

Future Work

- Other face analysis subsystems can be also evaluated using this tool: age, gender, etc., but it would require extending the database.
- We are currently optimizing the code, but we will be open source it in future.
- We will provide the dataset (or a subset of it)

Thank you for you attention

http://vision.die.uchile.cl

mauricio.knight@gmail.com, jruizd@ing.uchile.cl, rodrigo@verschae.org