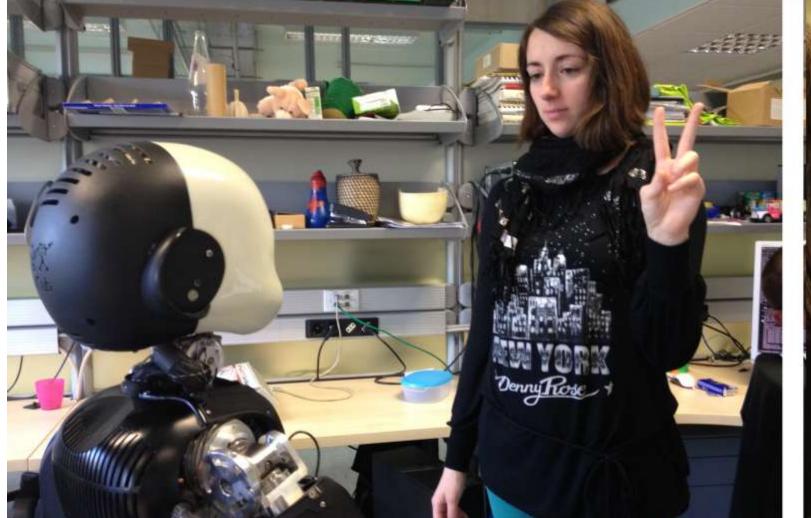


CAN YOU DEFEAT A ROBOT? All Gestures You Can: a Memory Game

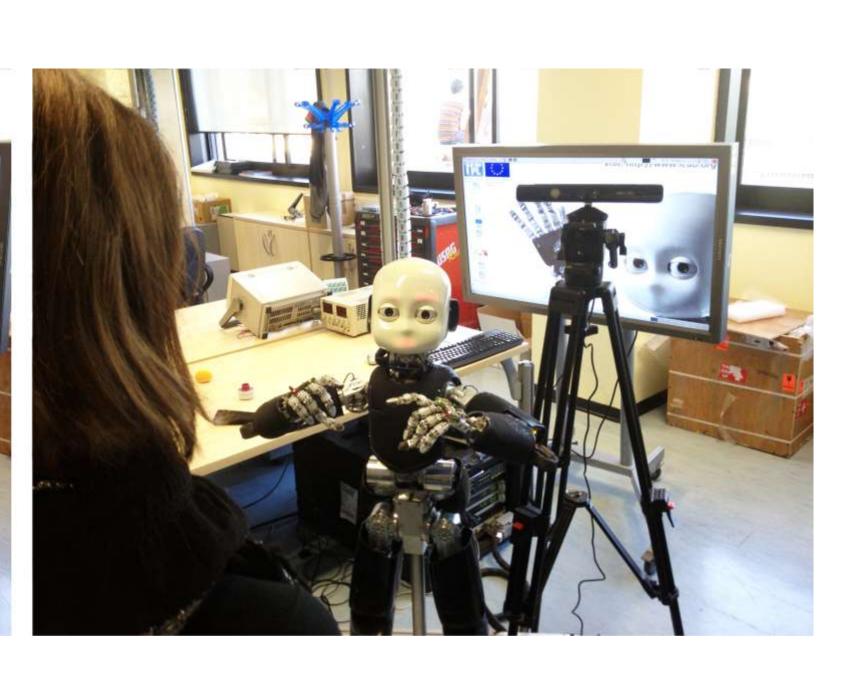


Ilaria Gori¹, Sean Ryan Fanello¹, Giorgio Metta¹, Francesca Odone²

- 1. Department of Robotics, Brain and Cognitive Sciences Istituto Italiano di Tecnologia, Genova Italy
- 2. Dipartimento di Informatica e Scienze dell'Informazione Università degli Studi di Genova, Genova Italy

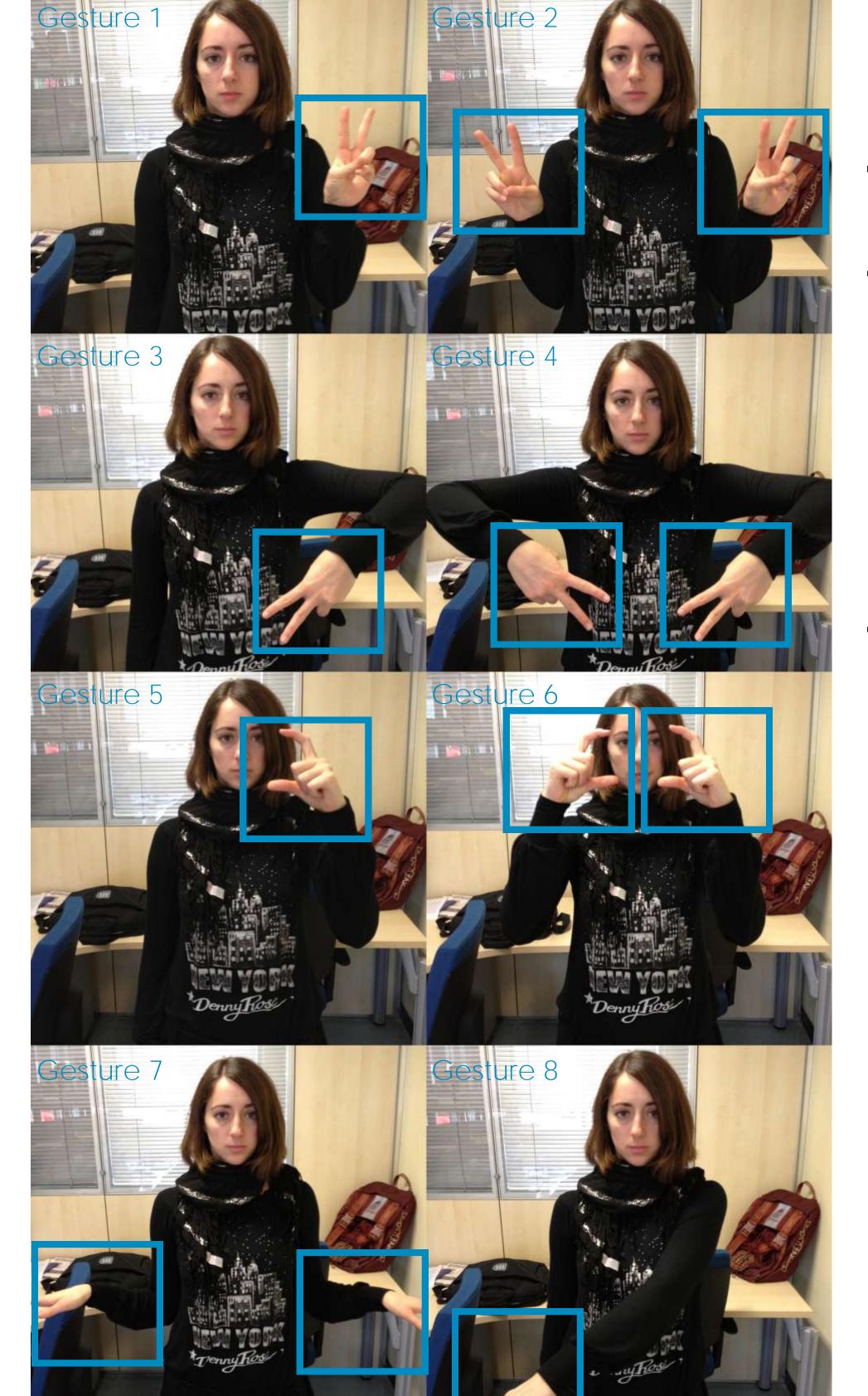






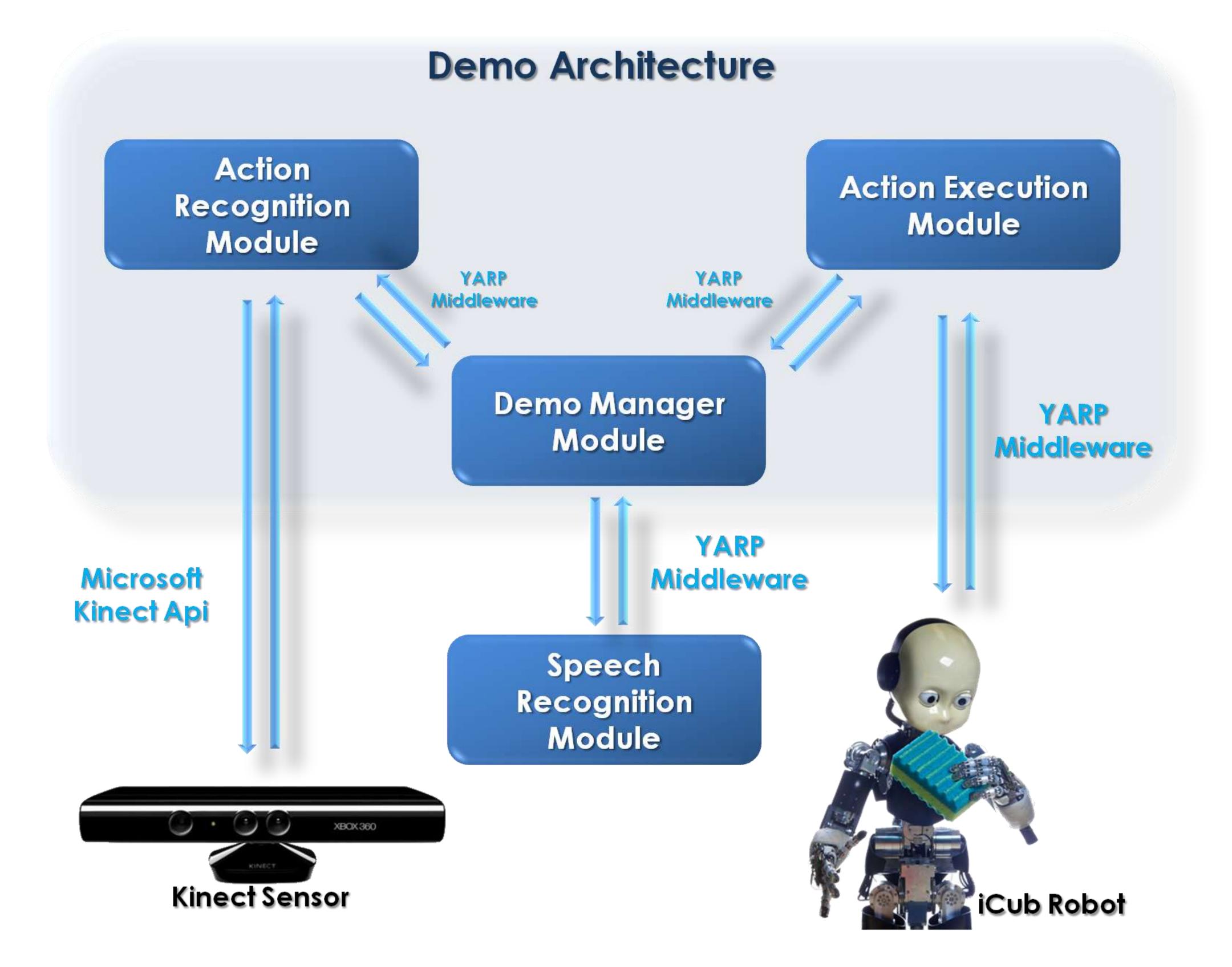
Abstract

"All gesture you can" is a real-time memory game: the goal is to perform the longest sequence of gestures that is possible to remember. The human player is challenged to confront a humanoid robot, iCub, that will never lose for lack of memory. The proposed game is considerably challenging but easy to play; it stresses player's memory, visual association and concentration.



Rules

The human and the robot take turns and perform the longest sequence of gestures by adding one gesture in each turn. The game starts with one player (e.g. the human) performing the first gesture. Then the iCub has to recognize it, to imitate it and perform another one randomly chosen. The game continues until one player loses: either the human because of his memory skills, or the robot because of the gesture recognition process.



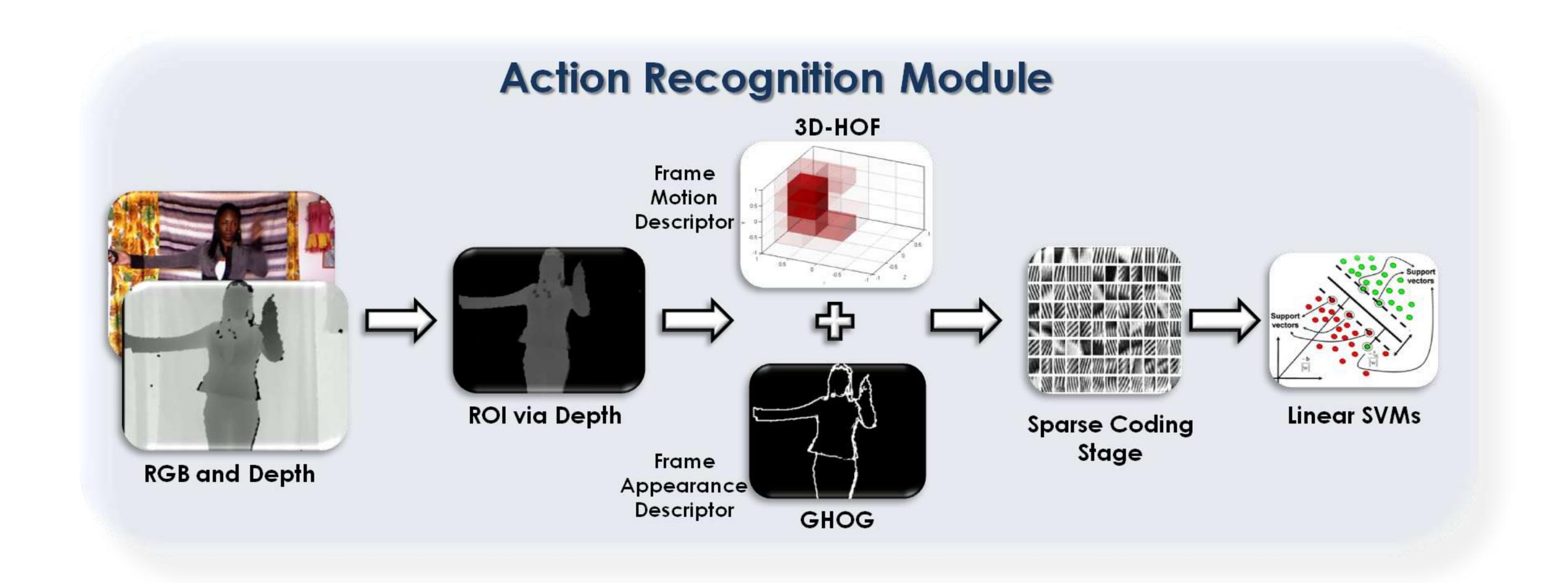
Cognitive Architecture

A simple and modularized architecture has been implemented by means of visual perception, speech recognition and robot motor capabilities.

Demo Manager Module keeps track of players' turns in the game; it takes care of the communication among the other modules and it offers an interface with the speech recognition module.

Action Recognition Module serves to classify gestures with Real-Time performances.

Action Execution Module exploits a simple position controller for the gesture execution.



Action Recognition Overview

The proposed action recognition system is modelled as a sequence of four layer:

- Region of Interest The ROI is detected via depth.
- 2. Frame Representation Each ROI is mapped into a feature space with a combination of 3D Histogram of Flow (3D-HOF) and Global Histogram of Oriented Gradient (GHOG).
- 3. Sparse Coding The combined 3D-HOF and GHOG descriptor is further processed via a Dictionary Learning step.
- 4. Learning and Recognition Linear SVMs are used on frame buffers. An online video segmentation algorithm isolates different actions.

