# All Gestures you Can: a Memory Game

I. Gori[1], S.R. Fanello[1], G. Metta[1], F. Odone[2]
[1]Department of Robotics, Brain and Cognitive Sciences - Istituto Italiano di Tecnologia
[2]Dipartimento di Informatica e Scienze dell'Informazione - Universita' degli Studi di Genova
[1]{ilaria.gori, sean.fanello, giorgio.metta}@iit.it, [2]f.odone@disi.unige.it

## Abstract

*We propose a real-time, original memory game based on gesture recognition with Kinect. The object of the game is to perform the longest sequence of gestures that it is possible to remember. We explore the human-robot interaction field, letting the player confront a humanoid robot, iCub. The proposed game is easy to play and it is suitable for people of every age; it also allows improving cognitive skills, testing player's memory, visual association and concentration. The involvement of a robot makes the game more challenging, as it will always reproduce the sequence correctly (rote memorization is not a problem), but the system can misinterpret the observed action.*

## 1. Overview

Games that interact directly with the players have recently gained an increasing popularity. A satisfying interaction needs to rely on systems having reliable perception skills; to this end several devices have been developed, such as the Wii remote that senses movement, or Kinect, which recognizes the body posture of the player. Along with the diffusion of these interactive games, research in computational vision gained a boost and a growing interest. For this reason we thought that a game would be the best way to demonstrate the reliability and the effectiveness of our gesture recognition system based on Kinect.

Our belief is that every game leaves something to the person who plays it; in fact there are some games, especially those bonded to our childhood, which we remember for the rest of our lives. More important games are those increasing cognitive capabilities, that allow children to acquire specific skills: we propose a game aimed at improving memory skills, visual association and concentration. It takes inspiration from the classic Simon game, launched by Milton Bradley in 1978. Nevertheless, since the original version has been often defined as "visually boring", we developed a revisited version, based on gesture recognition, that involves a "less boring" opponent: a humanoid robot, the iCub [1]. The robot is supposed to imitate a sequence of recognized actions, and to perform another random gesture that is added to the sequence; the longer sequence so created will be executed by the player that will add another gesture. This routine is carried on until either the robot or the player makes a mistake and thus loses. The interaction with the robot makes the game more challenging but also more playful and fun.

### 1.1. Gesture Recognition System

We present here a brief description of our gesture recognition system. Related papers are still under review, therefore we cannot provide further references yet. Within the proposed system each action is described as a sequence of features based on motion and appearance cues; motion is fundamental in order to catch the variability of the gesture over time, whereas appearance provides information on the body posture, accounting for gestures that carry a specific meaning. The combination of descriptors on motion (3D Histogram of Flow) and appearance (Global Histograms of Oriented Gradient on the whole body and on the hands) represent our main features; this code is further improved by means of a sparse coding stage, needed to filter noise and to retrieve sparse descriptors. Each gesture is then modeled by a linear SVM, which guarantees constant complexity in testing and, thanks to the selected features, needs only one demonstration to build a robust classifier. One of the most challenging goal that every action recognition system has to achieve is segmentation, as it is necessary to detect the action starts and stops. We provide an effective online segmentation algorithm, which exploits variations of the SVMs scores in order to detect which action is being performed and when. In order to correctly isolate the player from the rest of the scene, we exploit depth information. In Fig. 1a) a block diagram of our complete system is shown. Importantly all the modules in our architecture run in realtime (25fps); this is fundamental to effectively recognize gestures in interactive environments. Moreover, thanks to Kinect and to the specific choice of features enriched with

[1] www.icub.org

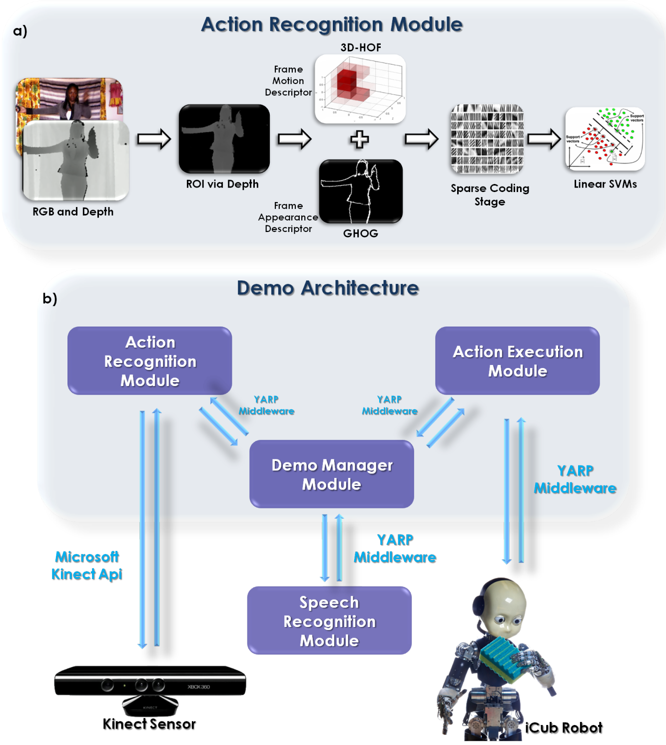


Figure 1. Fig. a): Overview of the recognition system. We propose a solution that cope with real-time requirements. Each frame is represented by 3DHOF and GHOG descriptors. The learning stage implicitly models the time dependency of actions concatenating descriptors of buffer of frames. Fig. b): Implementation of the demonstration. There are three main modules that take care of recognizing the action sequence, defining the game rules and making the robot perform gestures.

sparse coding, the system appears to be robust even in scenarios with additional noise.

### 1.2. How to play

We propose a game where the player confronts a humanoid robot, iCub. Both the human and the robot have to take turns and perform the longest possible sequence of gestures by adding one gesture in each turn. The game starts with one of the two opponents, e.g. the human, performing a gesture, that represents the first gesture of the sequence; the iCub has to recognize it, imitate it and perform another one, randomly chosen, that represents the second gesture of the sequence. Now the opponent must perform correctly the whole sequence and to choose randomly a third gesture that will be added to the sequence. The game continues until one of the two players loses: the human player can lose because of his limited memorization skills, whereas the robot can lose because of the gesture recognition process. The iCub keeps track of the sequence that is being performed, but it is asked, at each turn, to recognize again the whole sequence of gestures performed by the player: if the iCub finds out that the performed sequence is different from the saved one, it says "I think you are wrong". If the player is actually wrong, the iCub wins. Nonetheless, the robot could be wrong in recognition; in this last case the iCub loses. The system provides the possibility, during the game, to learn and recognize new gestures with a single demonstration; nevertheless, since the mapping between the recognized fingers position and the iCub's joint position is not trivial, we fix a set of predefined gestures that the players can choose.

## 2. Implementation

The proposed architecture is simple and modularized as it is organized in three meaningfully separated components; firstly, the Action Recognition Module, implementing our gesture recognition system, which has been briefly described in Sec. 1.1. Then the Action Execution Module, which deals with the iCub merely executing the gestures: it exploits a position controller that minimizes the error between the expected joint configuration and the actual one. Finally, the Demo Manager Module keeps track of players' turns in the game, it takes care of the communication between the action recognition and the action execution modules, and it offers an interface with an external speech recognition module that exploits the Microsoft speech recognition software and it is used for simple commands such as "Start" or "Win/Lose". All these programs communicates through the YARP Middleware, which allows exchanging information easily, hiding the underlying network structure. The efficiency of the proposed implementation is assured by its multithreading architecture, which also contributes to real-time performances. The code of the demonstration is available in the iCub repository [2].

## 3. Conclusion

We proposed a challenging and original memory game, which foresees a competition between a human player and a humanoid robot, iCub; the game challenges human rote memorization, in fact the iCub can only lose if the recognition process fails. The game is fast and playful since the modules run in real-time. The main contribution is a robust, complete and real-time gesture recognition system, which also allows for one-shot learning.
NOTE: due to shipping time and cost, we will not be able to bring iCub to the live demonstration. The system has been tested on the iCub, as it is possible to see from the related video, but for the demonstration we will rely on a simulator.

[2] Code available at https://robotcub.svn.sourceforge.net/svnroot/robotcub/trunk/iCub