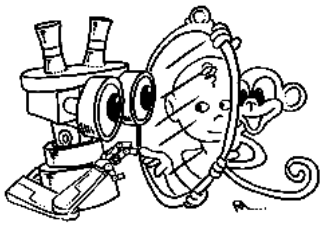


MIRROR

IST–2000-28159

Mirror Neurons based Object Recognition

Deliverable Item 3.1

Biological data acquisition set-up specifications

Delivery Date: April 30, 2002

Classification: Public

Responsible Person: Prof. Giulio Sandini – DIST University of Genova

Partners Contributed: Matteo Brunettini (DIST)

Short Description: This deliverable describes the experimental set-up being developed for the acquisition of visual and motor data during grasping actions performed by humans. The motivation for this set-up is to start experimenting with low-level algorithms based on visual and motor data that could be used to extract, code and recognize grasping actions. Visual data are acquired through two video cameras in a binocular stereo arrangement positioned so that the acquired video stream is very close to the "subjective" view of a person during manipulative actions. The motor data is acquired by means of a data glove measuring the evolution in time of the hand posture (22 sensor on palm and fingers) position and orientation of the wrist (6 more sensors). Visual and motor data are acquired synchronously and stored on disk for off-like processing.

Project funded by the European Community under the "Information Society Technologies" Programme (1998-2002)

Content list

1. Introduction

This deliverable describes the experimental set-up for the acquisition of visuomotor grasping data. In particular what we want to record simultaneously are sequences of images and kinesthetic information during grasping actions with the purpose of starting to investigate which visuomotor low-level processing are necessary to learn how to discriminate, from visual information only, between different grasping actions. In order to do that we need to acquire information as close as possible to the real visual and proprioceptive information available to humans during the initial years of life. Considering that it is impossible to record from a human the motor command elicited by the motor areas of the brain, we decided to record something similar to the proprioceptive information namely the evolution in time of the position and orientation of the hand and fingers during a grasping action. Furthermore, visual information will be acquired by a pair of stereo cameras mounted in such a way that they give a “subjective” view of the arm-hand during motion execution.

2. The Hardware

The grasping information we need to acquire are of two different kinds: images, showing the subject’s arm reaching and catching an object, and information about the hand kinematics, describing its position during the motion as well as the fingers posture.

The main task of this setup is to acquire and store information, without doing any on-line processing. As the system does not require any special computational speed, we decided the to use a commercial PC equipped with a Pentium IV class processor at 1.7 GHz and 256 Mbytes of RAM. The machine is equipped with an Ultra-SCSI Hard disk drive, that allows to record large amount of data, such as data required by images, in a quick and efficient way.

The image acquisition is achieved by using two Watec WAT202D digital cameras (see figure 1) with PAL standard (768x576 pixels, 25 Hz of frame rate, color) acquired by two Picolo Industrial Frame grabbers. These cameras are mounted behind the subject, up to his shoulders, one on each side to acquire images from a “subjective” point of view (see figure 2).

Figure 1: one of the two cameras used. The size of the camera is approximately 4x4x4cm

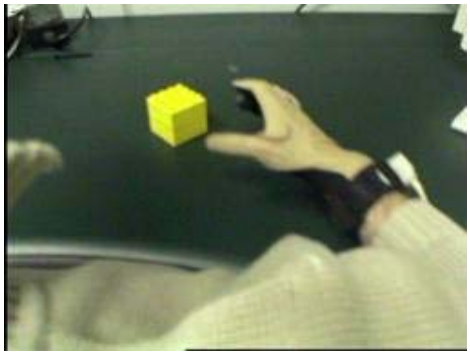


Figure: subjective sample image acquired by one of the two cameras.

To obtain the hand kinematics we use two devices: a Dataglove, which reads the fingers joints angle, and a 6 degrees of freedom tracker, to determine the hand position and orientation in space.

The Dataglove, a “CyberGlove” model produced by Immersion, consist of a glove mounting 22 sensors reading the hand joints angle and transmitting the date to the PC through a serial

interface. There are three sensors for each finger to read the phalanges relative angles, four sensors for the finger abduction, two sensor to measure palm inclination and orientation, and a sensor for the palm arch (see figure 3). Data from these sensors can be acquired up to a frequency of 112 Hz.

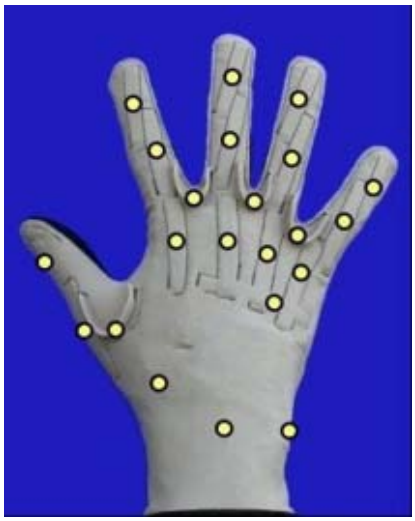
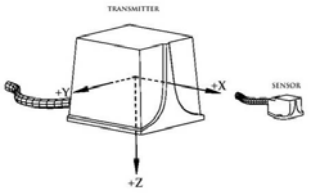


Figure 3: positioning of the 22 sensors of the data glove

The tracker is a “Flock of Bird” produced by Ascension (see figure 4), it determines the position of a sensor in space, using a magnetic field emitted by a transmitter. The sensor and the field emitter are wire-connected to a device that reads the sensor position and orientation, relative to the emitter, and send them to PC serial port. The tracker sensor is attached onto the wrist, using a specific housing slot on the CyberGlove. The glove is light weighted, also with the tracker sensor mounted on, and don’t limit in any way the hand motion and manipulation.



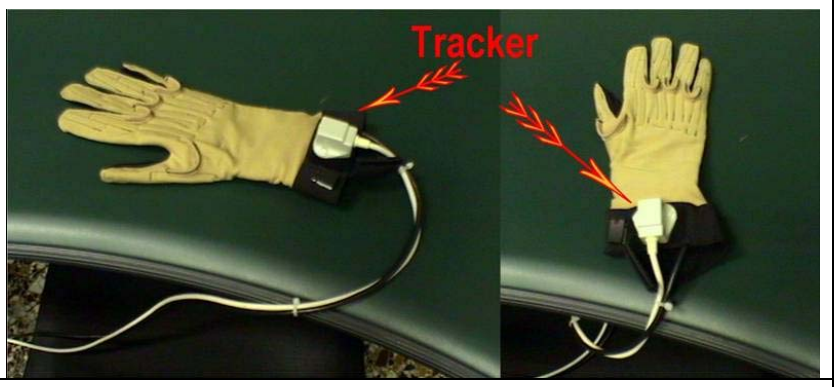


Figure 4: *Flock of bird* sensor (left) and its position on the dataglove.

The set-up is assembled so that the fixation point of the stereo pair will be on the target object.

3. The Software

The software that acquires and saves grasping data is written in C++ and compiled with Microsoft Visual C++, on Windows 2000 platform.

It gather information from the peripherals and write them on disk: the images are saved in Windows Bitmap format, with a size of 384x288, 24 bits per pixel of color depth, without any compression; the hand joint angles are written in a text file, using the windows Comma Separated Value standard, together with the hand position and orientation.

The program works with a frame rate of 25 Hz, this means that every 40 msec it acquires two images from the cameras and read the serial ports to record the Dataglove and tracker sensor data; this is the maximum frame rate allowed by the PAL standard, in order to have a complete new frame available in memory.

Considering that the purpose of the set-up is to acquire quantitative behavioral data, it is critical that, once the recording has started, no frames are lost and, furthermore, the video and kinematic data are synchronized. The kind of processing we will perform, in fact, may require the computation of image velocity information as well as the possibility of performing cross-correlation between visual and motor data. Of course the most demanding requirements are relative to the video signal. For this reason video acquisition is based on the use of two buffers (besides the double buffering mechanism used by the frame-grabber board). The acquisition buffer is accessed by the frame grabber through its DMA channel and by the main acquisition thread, which, every 40 msec, transfer the image pair in to a storage buffer. The main acquisition thread, besides the actual transfer, also adds the time-stamp information required for synchronization (see below). The storage buffer is accessed by the storage thread, which saves the acquired data on a swap file on disk. This operation is performed asynchronously and because of the fast disk access allowed by the Ultra-SCSI Hard disk drive, the storage buffer never overflows. An example of a sequence of images is shown in figure 5 (the actual sequence was composed of about 50 stereo frames).

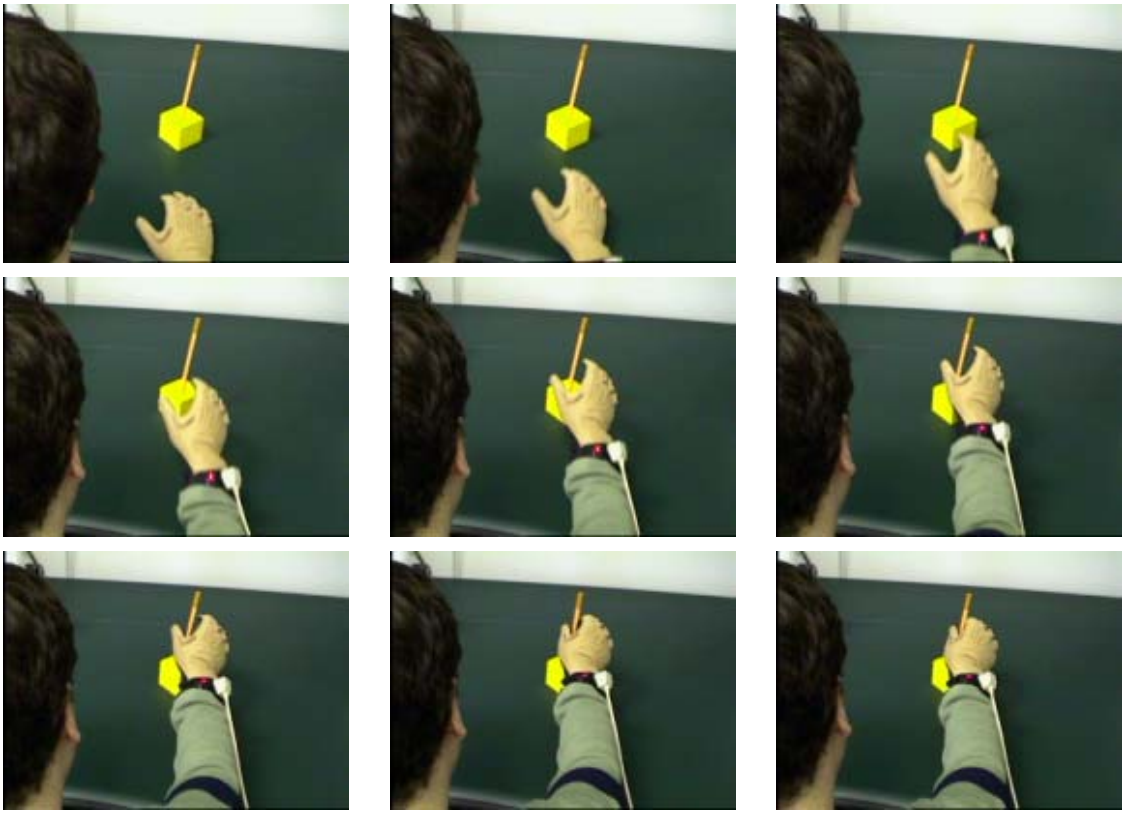
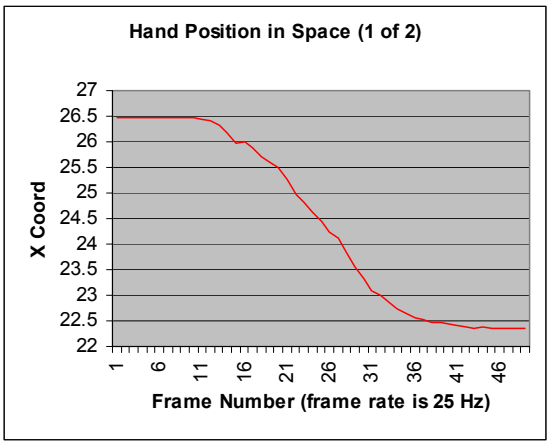


Figure 5: sample sequence from the right camera of a grasping action.

Another requirement is to maintain the synchrony between video and kinematics information, to determine which frame corresponds to a particular sensors reading and vice-versa. This synchrony assures that a given image pair (right one and left one) is grabbed in the same time interval (within the same 40 milliseconds) of the corresponding sensor data reading. To do this the program use time stamping (see above): every image pair is "time stamped" to determine to what time interval it belongs and the same thing is done with the kinematic data. Then, once the recording has been stopped, the data are correlated using the time stamps to reconstruct the correct data flow and the bitmap images and the text file are available in the hard disk.

In the following figure 6, sample data from the wrist position sensor are shown. Similar data are acquired from the 22 angular sensors of the data glove.



Hand-X	Hand-Y	Hand-Z
26.477859	12.379772	-21.0065
26.477859	12.362194	-21.01529
26.473464	12.370983	-21.019684
26.469069	12.366588	-21.019684
26.482253	12.353404	-21.01529
26.464675	12.379772	-21.024079
.....		
.....		

Figure 6: Numerical (right) and plotted (left) data from the position sensor at the wrist.

4. The protocol

Until now we have not started a the systematic acquisition of data but only few recoding trials have been made to test the performance of the apparatus. In the meantime we have defined the experimental protocol that will be used. The protocol will required the acquisition of at least 10 subjects each repeating for 10 times grasping actions toward 3 different objects in 3 different positions (left, right, in front). Each object will have a different shape so that it will elicit 3 different grasping actions. For example: a) an elongated object lying flat on the table (pencil); b) a cylinder of about 8 cm in diameter with its axis vertical (cup), c) a small sphere 1 cm in diameter (candy). The subject will be required to lay his/her hand flat on a table always in the same position and approximately in front of the subject on the midline. The elbow flexed and the fingers close to each other. An object will be placed in specified positions at a distance requiring the full extension of the arm to be performed to reach the object. In total there are nine possible conditions (3 grasps and 3 positions) to be repeated 10 times for each subject. For the pencil it is also possible to change the orientation on the plane which, initially, will be kept constant. The experimenter will initiate each trial by placing one of the objects in one position (this will initiate the start of the recording). The subject will be required to grasp the object and to bring it to the initial position of the hand with no particular instructions. Besides the visual and kinematic data, the time of contact between the hand and the object will be marked.