

VMGDb (the CONTACT Visuo Motor Grasping Database)

autori e affiliazioni

July 24, 2009

Abstract

Data papers are short (circa 4 pages)

1 Introduction

- Introduction - Barbara motivations (novelty) and context (applicazioni ispirate dai neuroni mirror con oggetto l'analisi di dati multi-modali (video e sensoriali))- Barbara
- state of the art - Claudio

The paper is organized as follows.

2 Acquisition setup

setup di acquisizione : descrizione (hardware, guanto, software di controllo, ... questa parte un po' nelle mani di Claudio), qui di seguito un copiacolla da iciap. Figure da cambiare.

add a picture of the acquisition set-up (Fra+Ni at the liralab?) or a drawing

Data were collected using two Watec *WAT-202D* colour cameras for the images and a 22-sensors Immersion *CyberGlove* for the hand posture. An Ascension *Flock-Of-Birds* magnetic tracker mounted on the subject's wrist, and a standard force sensing resistor glued to the subject's thumb were used to determine the hand position and speed, and the instant of contact with the object.

The cameras return two video sequences, one placed laterally with focus on the object (the *spectator*) and one placed in front of the subject (observing the *actor*).

The video sequence is acquired at 25Hz by each camera, while the glove is sampled at 100Hz. Since the three devices are independent of one another a

	ball	pen	duck	pig	hammer	tape	lego brick
cylindric power				X			
flat					X		X
pinch		X	X			X	X
spherical	X					X	
tripodal	X	X	X			X	

Table 1: Mapping grasps-objects. Each actor performs grasping actions on 13 object-grasp type pairs.

system of common time-stamps was used in order to synchronise the data.

The CyberGlove returns 22 8-bit numbers linearly related to the angles of the subject’s hand joints. The resolution of the sensors is on average about 0.5 degree. The sensors describe the position of the three phalanxes of each finger (for the thumb, rotation and two phalanxes), the four finger-to-finger abductions, the palm arch, the wrist pitch and the wrist yaw.

3 The dataset

The dataset is built considering 7 different objects ((see Fig. 1), top) and 5 grasps ((see Fig. 1), bottom). 20 different subjects participated to the acquisition. Table X (**TO DO: Nicoletta** reports information - gendre, age, ... - on the subjects.

Each object can be grasped in one or more different ways, according to the many-to-many relationship reported in Table 1. In total we consider 13 different pairs grasp-object. Finally, for each triple (object, grasp, subject) we have 20 replicates of the experiment.

3.1 Visual-motion data

Each experiment (object, grasp, subject, expnum) we have different information stored: - 2 sequences of images (ppm) acquired from 2 cameras with different focus (object vs action). **FRA: For space reasons we built the corresponding video sequences (avi) and for the same reason I would make the video available to the public** - 2 data files (one for each video) that report the time-stamp of each image - a file containing the sensor data and the time-stamp of acquisition. By comparing these time-stamps with the ones included in the other 2 files we can synchronize visual and motor data.

More details on the sensor data file, if needed: for each row, sensor data include * a time-stamp starting from 0 * 6 reals from the FoB tracker (XXX Claudio, cosa significa FoB?) * 9 zeros (XXX Claudio, perch?) * 22 integers

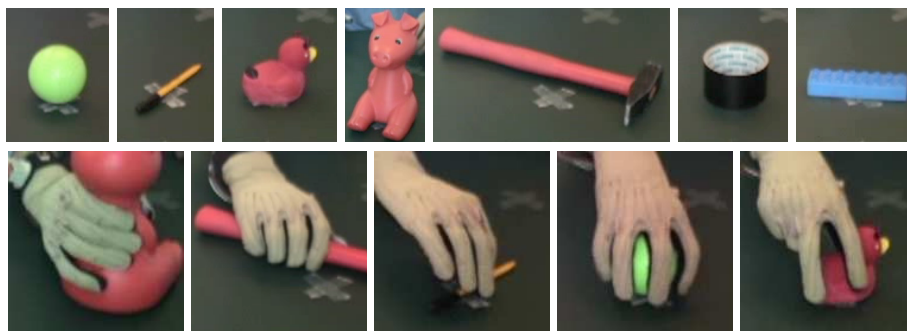


Figure 1: Top row: the objects used in our experiments. Bottom, the grasp types we consider: (*left to right*) cylindric power grasp, flat grasp, pinch grip, spherical and tripodal grip.

from the Data Glove * 4 integers from the pressure sensors (only the first one is meaningful XXX Claudio, perch?)

FRA: Prepare a figure showing frames from the two sequences of a given experiment.

3.2 Annotations and data indexing

TO DO Nicoletta: Describe how to access the various data files related to a given experiment

- Annotation: – labelling with respect to subject, object and grasp – we can extract the static glove configuration corresponding to the grasp by analysis the derivative of pressure sensor, which assumes lower values in those time instants (we thus implicitly extract also this temporal information).
- we could extract the the time instants when the action starts (by studying the position of the hand? TODO?)

- scripts for automatic reading? – We have different scripts that allow automatic data reading and extraction of all above information. Is it worth describing them, or else shall we just make them available?

forse si potrebbe mettere un grafico esplicativo che chiarisca cosa intendiamo quando affermiamo che il fermo immagine sulla presa pu essere ottenuto studiando l’andamenti del sensore di pressione, abbiamo lo script per l’estrazione

3.3 Data peculiarities

Report all peculiarities that make the data difficult and/or interesting:

- natural illumination - hands always present - actions variability among subjects (different subjects do the same grasp in quite different ways... Maybe some images to explain this...)
- different velocities - variability also among samples from the same subject (a subject is at ease as he repeats the same grasp many times)

4 Evaluation

NI: analisi? qualche risultato su object classification - grasp classification? abbiamo qualcosa tra le robe che ho fatto all'idiap (confronto tra feature diverse). FRA: aspettiamo di vedere come va il resto dell'articolo

5 Conclusions