

$\frac{85}{100} \rightarrow 85$

BIOLOGICAL MOTION

CNS186 project

Domitillo: Interesting experiment if carried out a bit further
I am not convinced that you proved one of your conclusions
(i.e. that for the short displays the subjects were performing
at random) moreover you do not compare the
psychophysics with simple models of ~~what may be~~ the info ^{available}
~~going out~~, nor with psychophysics done on
synthetic motions to verify whether biological
motion is privileged -
We should discuss this briefly

Introduction

Biological motion is treated in a rather privileged way by human visual system so that those visual functions more or less directly concerned with motion processing may take advantage of this: they could work faster or better when an observed motion is in fact biological motion. Johansson (1973,1975) was one of the first people concerned with human ability in detecting and recognizing human motion and he run experiments which led him to think that people would percieve the movement of human body from just the motion of the joints.

Here we ask the following question:

What is the amount of information that the human visual system capture by looking at some activities?

More specifically we pose a *prediction* problem, i.e. if the human visual system can capture enough information by looking at a short human activity so to predict the body position K steps ahead. And in particular we ask:

To what accuracy can the visual system predict future human movements based on previous observations?

What is the amount of information that human visual system needs in order to Guess the final position of the human body involved in a complex activity?

In order to explore such questions we propose a set of experiments run on human subjects involved in a prediction task. We report the experimental results in figures which show the relation between prediction accuracy and amount of information available to take a decision.

Experiment Set up

The overall experiment is divided into two parts: in the first half we collect data and in the second part we show the properly manipulated data to 5 subjects in order to perform our psychophysic tests.

a) data acquisition

There were 14 equally spaced targets (10 cm one from the other) displaced along an horizontal line on a table at arm length from the subject who was sitting in front of the table. The top view is shown in Fig.1. The subject had to reach for each target one after another in arbitrary order and observing 3 seconds of pause at the base position and 2 seconds at the target.

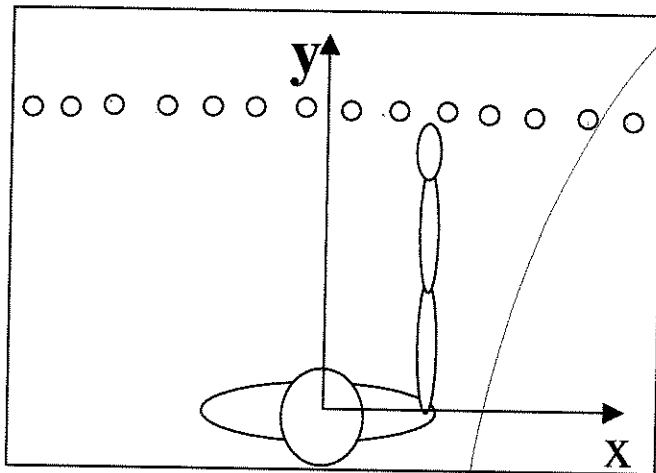


Fig 1. Top view of the experiment part a).

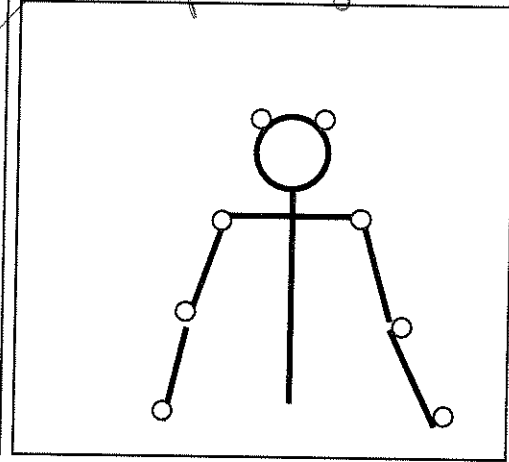


Fig 2. Position of the bulbs.

Since you had 14 targets you had different n. of reaches per target in why?

The motion capture system used tracks the position in time and space of salient points chosen by placing at their location bulbs that the machine sees as lights in a dark room. The bulbs were placed at the main joints and on the head as shown in Fig.2.

The data acquired consisted of 24 reach actions in total from which we constructed reach samples obtained by truncating each reach execution at proper time instants as explained in the following. As it is well known the velocity profile of a reach movement is characterized by a bell shape profile, an example from the data set is reported in Fig.3. Then from each reach example we construct three reach samples each one of them containing more information and obtained by truncating the reach execution in the three points shown in Fig.4. If we denote by d_L and d_R the width of the left part and the width of the right part respectively of the bell profile, then the sample with least information is the one which is truncated at $1/3 d_L$ before the peak, a richer sample is obtained by truncating the trajectory at the peak and the richest sample is obtained at $1/3 d_R$ over the peak.

from one or more people?

The interest is then in finding the accuracy to which a subject can predict which target is going to be picked given the amount of motion shown in the different sample obtained by truncating the trajectory.

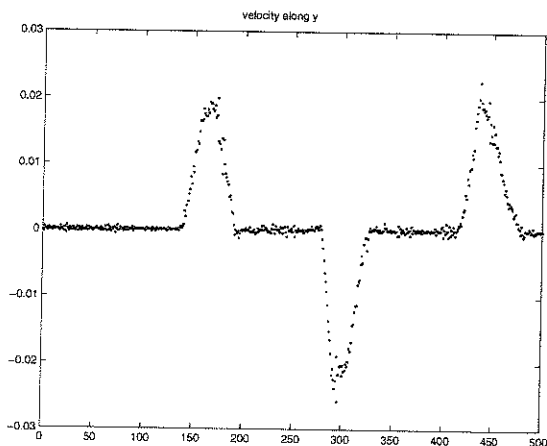


Fig. 3. Velocity profile in reach samples

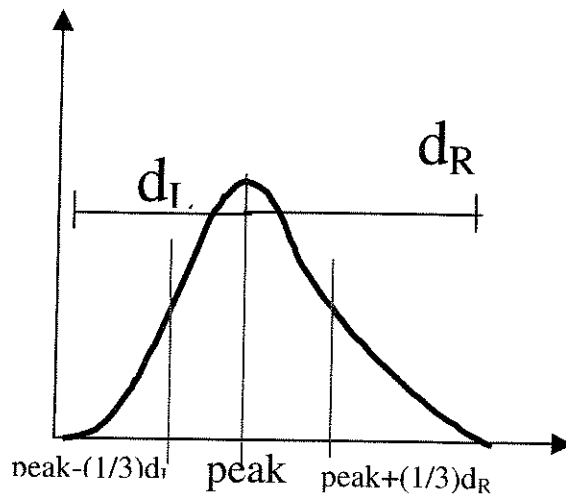


Fig. 4. Truncation points

b) psychophysics experiment

We performed the experiment on 5 different subjects. Each subject was shown a total of 72 truncated samples in which 24 were truncated at one third before the peak, 24 at the peak and 24 at one third over the peak. When the movie started the subject had to watch it until the truncation point occurred and the man as shown in Fig. 5 disappeared and the movie was frozen with just the targets shown. Then the subject had 3 seconds to click on the target he predicted the man was going to reach. After that the subject was shown the final part of the reach to realize if he was right.

We registered all the subjects responses by considering as measure unit for the prediction error the distance separating two close targets. The results are shown in Fig. 6, Fig.7 and Fig.8. Fig.6 shows the prediction errors for each of the five subjects for each of the three different truncation points. As a general result the prediction error decreases for each subject as an higher number of frames was shown (see also Fig.7 for the average result). Moreover some directions turn to be more difficult than others: the far left ones turn to be the trickiest together with the close right; instead the far right turn to be easier. There is no big variance across subjects in fact as we can see in Fig.6 they perform in very similar ways. In Fig.8 we show the percentage error as a function of the number of frames shown. As expected it is a decreasing quantity; moreover it turns out that to do a satisfactory prediction (less than 50% error) subjects need to see at least up to the point in which the person starts decreasing his speed otherwise they do not have enough information and tend basically to guess the position rather than to predict it. Anyway it is surprising that subject can still predict several positions correctly even if a poor amount of trajectory is shown: it could be that the biggest part of information is acquired in the first frames in which the movement is started.

The experiment had already been performed by freezing the person at the truncation point rather than letting it disappear. There are no big changes except for the fact that in the case the movie contained more information the performance turns now to be increased. This could be due to the fact that before people tended to predict the target also by looking at the closest one to the hand of the person when the movie was frozen: this was misleading some times.

Remarks: we concentrated here on prediction task not concerned with depth (y axis) : in such case the task would probably be harder.
The targets displacements is used also for determining a unit error measure: if the distance between targets was increased the error rates would probably go down since a less fine discrimination is required.

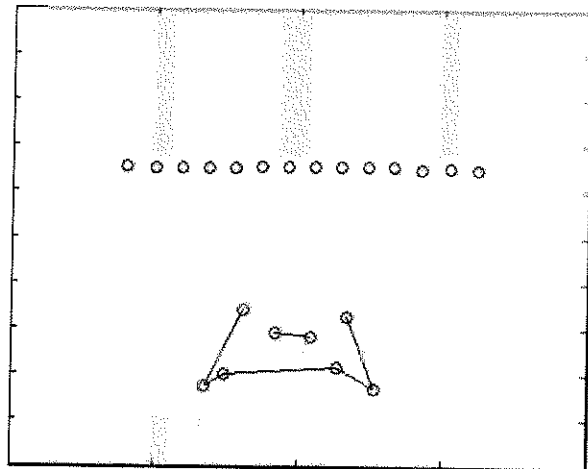


Fig. 5. Frame from one of the movie shown to the subjects

how do you know this?

Should we use some coordinates...

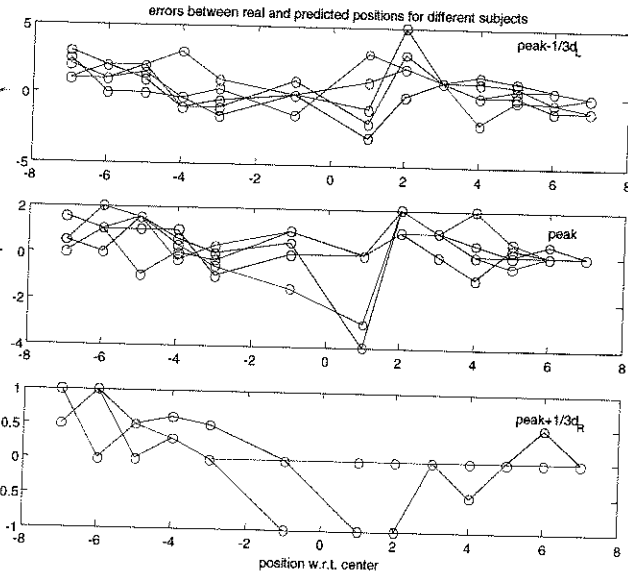


Fig.6: each subplot contains the prediction errors for the five different subjects for each one of the three different truncation points, i.e. at the peak-1/3dL (first subplot), at the peak (second subplot) and at the peak+1/3dR (third subplot). The x axis represents the positions of the targets according to the reference frame of Fig.1

Is this an interesting phenomenon?

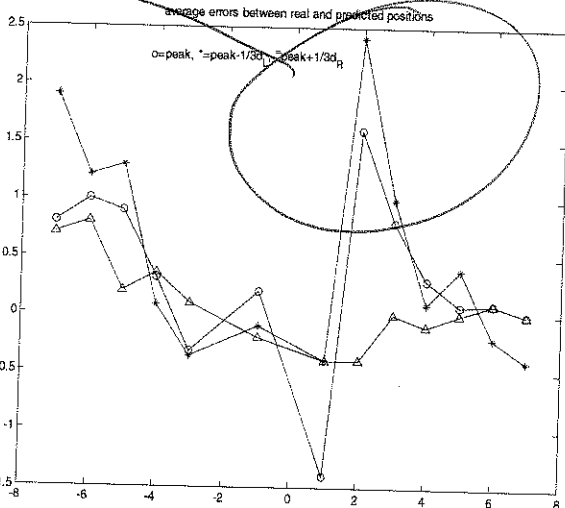


Fig.7: this figure contains the average between subjects of the prediction errors already reported in Fig.6.

Should have criterion normalizing by distance of hand in first frame.
 $\frac{1}{d}$ of hand in first frame
 $\frac{1}{d}$ of target

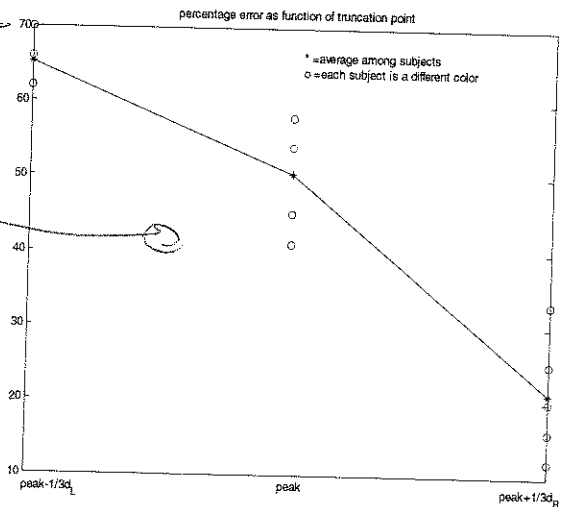


Fig.8: the black curve represents the average percentage error between subjects for each of the three truncation points (x axis). The circles represent the results for each of the five subjects.

Indicate performance at chance

It should be around 7% — why do you say that subjects performed at chance in the first condition?