The Appearance of Shape in Visual Perception: Eye Movement Patterns During Recognition and Reaching

Filipe Cristino
f.cristino@bangor.ac.uk

Lina I. Conlan
pspa09@bangor.ac.uk

E. Charles Leek
e.c.leek@bangor.ac.uk

Wolfson Centre for Clinical and Cognitive Neuroscience, School of Psychology, **Bangor University**

ABSTRACT

One fundamental aspect of understanding appearance is the visual perception of shape, and how this is modulated by task demands. Here we examined how eye movement patterns relate to the perception of shape during tasks of object recognition and the planning of prehensile movement. Participants carried out either a recognition task (where they learned a set of novel objects and were then tested on recognition), or were asked to plan a reaching movement. The results show that eye movement patterns were linked to the perception of shape, and that these patterns varied between tasks: not only in terms of fixation and saccade parameters but also scan patterns.

Categories and Subject Descriptors

I.2.10 [Vision and Scene Understanding]: Shape

General Terms

Algorithms, Measurement, Experimentation, Theory

Keywords

Eye movements, Object Recognition, Reaching

1. INTRODUCTION

Shape is a fundamental property of visual appearance. Recent work has shown how eye movement patterns can be used to elucidate shape analysis strategies during visual perception [1], but surprisingly little is known about how the perception of shape varies between tasks of recognition and movement planning. Since the seminal work of Yarbus [2] on the interaction between eye movements and tasks when viewing pictures, numerous studies have attempted to understand the visual saccadic system. Most of these have focused on understanding what drives eye movements without taking into account the effects of task demands. Although the stimuli used in previous studies have varied greatly from dynamic natural scenes to fractals, only a few have used single objects. The goal of the current study was to extend the recent work of Leek et al. [1] who found that high curvature – particularly extreme concave minima – could predict eye movement patterns when performing an object recognition task. In this study we aimed to examine how eye movements could be used to elucidate differences in shape analyses strategies across tasks of object recognition, and the planning of prehensile (grasping) actions.

2. METHODS

2.1 Participants

30 students from Bangor University (21 female, mean age 22 years, SD = 7.07, 3 left handed) participated for course credit, testing procedures were in line with local ethics committee and

BPS guidelines. All participants had normal or corrected-to-normal visual acuity.

2.2 Stimuli

The stimuli were 24 novel objects (e.g. see Figure 1a) each comprising a unique spatial configuration of four volumetric parts. These object models were produced using Strata 3D CX software and rendered using Matlab with a single light source (top left) in a mustard-yellow color. On average, the stimuli subtended 18 degrees of visual angle horizontally. This scale was chosen to induce saccadic exploration during viewing. Renderings of each object were created from six different viewpoints at successive 60 degree rotations in depth around a vertical axis perpendicular to the sight line. The zero degree viewpoint was a 'canonical' three-quarter view. The 0, 120 and 240 degree versions served as training viewpoints in the recognition task, and the 60, 180 and 300 degree versions as novel test viewpoints. All the viewpoints were used for the reaching task.

2.3 Apparatus

A Tobii 1750 eye tracking system was used to record eye-movement data. Stimuli were presented on a TFT monitor running at a resolution of 1280 x 1024 pixels. A chin rest was used for head stabilization. Stimuli were viewed from a 60cm viewing distance. A standard USB keyboard and mouse were used to collect responses.

2.4 Design and Procedure

After a successful calibration, the study comprised two phases: planning and executing an imagined movement to novel objects (Task 1), and an object recognition task (Task 2). The recognition task was divided further into a memorization phase, and a test phase. All participants completed both phases and their eye movements were recorded during each task. For counterbalancing the task order, targets were split into two groups. The participants in Group 1 (15) used six of the stimuli (1 to 6) as targets in Task 1, and six stimuli (7 to12) were used as targets in Task 2. The remaining 12 stimuli were split between the two groups and served as distracters in the test phase. For Group 2 (15) this assignment was reversed: participants were randomly assigned to each group.

In the planning reaching phase (36 trials each), participants viewed six objects each at six different viewpoints (0, 60, 120, 180, and 240 degrees). Following fixation at either the left or right side of the screen, stimuli were presented centrally for 5 seconds and participants were asked to plan and imagine picking up the object on the screen by 'using' their thumb and forefinger. Subsequently, when asked, the participants were required to indicate (by two mouse clicks) where they would have placed firstly their thumb and secondly their forefinger to grab the object. The participants were told to prioritize accuracy over

speed. In the learning phase of the recognition task (18 trials each), participants viewed six (target) objects, each at three different viewpoints (0, 120 and 240 degrees).
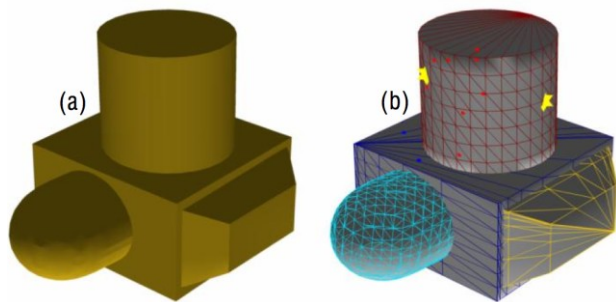

Figure 1. (a) Stimuli as seen by the participants (b) Eye movements (dots) and mouse clicks (stars) mapped onto the color coded mesh (single color per part)

The stimuli were presented sequentially for five seconds each at the center of the screen, following a starting fixation on either the left or the right side of the screen. In the test phase (72 trials each), participants were presented with previously seen target objects (from the six viewpoints), plus an additional set of six visually similar distracters presented from six viewpoints (60, 180 and 300 degree). Following fixation at either the left or right side of the screen, stimuli were displayed centrally until response. Participants were asked to determine whether the presented object was one of the previously learnt objects, and were invited to respond using a key-press (k for 'yes'/ d for 'no'). As in the reaching task, participants were told to prioritize accuracy over speed. The experiment lasted about 30 minutes.
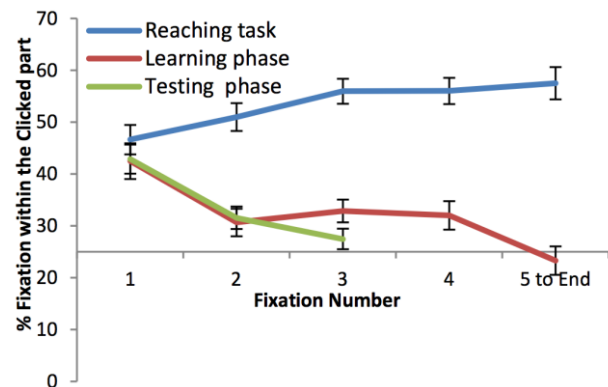
3. RESULTS

Figure 2. Time bin analysis of the % of fixations to clicked part

We first examined low level parameters of the eye movement data across the two tasks. Here we collapsed the data across groups as no significant differences were found between the two participant groups for any of the reported results. A one way ANOVA on dwell times $F_{(2, 87)} = 5.89$, $p = 0.004$ revealed a significant effect of task. Tukey post hoc tests indicated a significant difference between the reaching task (M=366ms, SD=105) and the memorizing phase of the recognition task (M=289ms, SD=57) $p < 0.001$ but not with the test phase (M=336ms, SD=90) $p > 0.5$. As expected, the fixation frequency data showed a similar pattern. A one way ANOVA revealed a significant effect of task $F_{(2, 87)} = 102.42$, $p < 0.001$. Tukey post hoc analyses showed a significant difference in the number of fixations made during the reaching

task (M=6.66, SD=1.37) and the memorizing phase of the recognition task (M=13.10, SD=3.25), but no significant difference with the test phase (M= 5.59, SD=1.42). Saccade amplitude was also analyzed. Here, no significant effect of task was found, but a Tukey post hoc test showed a significant amplitude effect between the reaching task (M= 4.47, SD=1.08) and the memorizing phase of the recognition task (M=4.98, SD=0.78) $p=0.048$. We developed a novel technique which enabled us to map fixations for a two dimensional rendered stimulus into the 3D mesh of the object itself (Figure 1b). With this method, we computed the object parts on which fixations were made. We were also able to map the mouse clicks made by participants during the reaching task and relate them to a given object part. In 92% of the trials, participants clicked twice on a single part. From this 92% of trials, we identified a clear preference for participants to fixate on the part of the object they planned to grab, where 56% of the fixations were made within the part clicked (M=56% SD=17.8). During the memorizing phase (M=23.26, SD=14.40) and test phase (M=23.83%, SD=11.33) of the recognition task, fixations to the clicked part were decreased to chance level (one in four – 25%). A time bin analysis was performed, where we computed the same analysis as previously reported, but using the first, second, third, fourth (etc) fixations (Figure 2). This shows that the first fixation is not predictive of the task but, from the second fixation onwards, the strategies clearly differ where participant performing the reaching task are much more likely to fixate on the single part with which they plan to interact. In contrast, during the recognition task they are equally likely to fixate any of the four parts.

4. DISCUSSION

Differences in eye movement patterns were found between the object recognition and reaching task. More saccades, with shorter dwell times and larger amplitudes, characterize the learning phase of an object recognition task. Fixations are also more dispersed across all parts of the object. When planning an interaction with an object – such as in our reaching task – eye movements are more localized to the object part where contact between the finger and the object is expected. This happens from the second fixation onwards. The first fixation does not seem to be indicative of the task ahead, but is probably based on the object's center of gravity [3]. These findings show how eye movement patterns can be used to elucidate the perceptual analysis underlying our perception of
shape appearance, and how this analysis differs between tasks. Additionally, the results show how important task differences are in any studies using eye movements. Our results suggest that when planning prehensile grasping actions observers quickly focus their analysis of shape on specific local parts that contain potential grasp locations. In contrast, during the encoding of object shape for recognition, fixation patterns are more spread indicative of a more global analysis of object configuration.

5. ACKNOWLEDGMENTS

6. REFERENCES
[1] Leek E. C., Cristino F., Conlan L. I., Patterson C., Rodriguez E., Johnston S. J. 2012. Eye movement patterns during the recognition of three-dimensional objects: preferential fixation of concave surface curvature minima. *Journal of Vision,* 12 (1) : 7
[2] Yarbus, A. L. 1967. *Eye Movements and Vision*. Plenum. New York.
[3] Vishwanath D., Kowler E. 2003. Localization of shapes: eye movements and perception compared. *Vision Research*, vol. 43, no. 15, pp. 1637-1653