Technical Report No. 55

Viewer-centered recognition of familiar faces

Nikolaus F. Troje[1] & Daniel Kersten[2]

January 1998

[1] AG Bülthoff, E–mail: niko@psyc.queensu.ca
[2] AG Bülthoff, E–mail: kersten@umn.edu

Viewer-centered recognition of familiar faces

Nikolaus F. Troje & Daniel Kersten

Abstract. The question of whether object representations in the human brain are object-centered or viewer-centered has motivated a variety of experiments with divergent results. A key issue concerns the visual recognition of objects seen from novel views. If recognition performance depends on whether a particular view has been seen before, it can be interpreted as evidence for a viewer-centered representation. Previous experiments used unfamiliar objects to provide the experimenter with complete control over the observers previous experience with the object. In this study, we tested whether human recognition shows viewpoint dependence for the highly familiar faces of well known colleagues and for the observer's own face. We found that observers are poorer at recognizing their own profile, whereas there is no difference in response time between frontal and profile views of other faces. This result shows that extensive experience and familiarity with one's own face is not sufficient to produce viewpoint invariance. Our result provides strong evidence for viewer-centered representation in human visual recognition even for highly familiar objects.

1 Introduction

Viewpoint dependent recognition has been reported for several classes of novel objects using a number of experimental paradigms (Rock & DiVita, 1987; Bülthoff & Edelman, 1992; Tarr, 1995). Recognition of unfamiliar faces also is clearly viewpoint dependent (Troje & Bülthoff, 1996; Hill & Bruce, 1996; Hill, Schyns, & Akamatsu, 1997). Other authors, in contrast, found viewpoint invariant recognition in studies using both novel and familiar objects (Biederman & Gerhardstein, 1982). Viewpoint dependence is often taken as evidence for viewer-centered as opposed to object-centered mental representations (Marr & Nishihara, 1978; for a review, see also: Ullman, 1996). However, even for object-centered representations such as feature spaces or 3D structural descriptions, one can expect to find some viewpoint dependence because particular views may provide more information about an object than others. This is in particular the case for viewpoint dependence measured in terms of view canonicality (Palmer, Rosch,

& Chase, 1981). Biederman and Gerhardstein (1982) used a priming paradigm to test for viewpoint dependence for familiar objects and found complete viewpoint invariance. However, like viewpoint canonicality this paradigm is not suitable to explore the nature of long term representations of familiar objects because it does not allow control about the information available to form the representation. Bruce, Valentine, and Baddeley (1987) reported viewpoint dependence for familiar faces. Response times to profile views were slightly longer than to frontal views. The reason might be that the attention towards another person's face is triggered if this person is facing towards us, resulting in an increased exposure to frontal views. If interpreted such, the findings of Bruce et al. would support a viewer-centered representation.

To explicitly test for viewer-centered versus object-centered representations, experiments have to be conducted that allow the experimenter to control which views of an object an observer has seen prior to testing. For this reason, the question of whether object representations are viewer-centered or object-

centered seems to be approachable only with novel, unfamiliar objects. Unfamiliar objects, however, that are experienced for the first time in the course of the experiment may be stored and represented in a different manner than familiar objects that have been known for a long time. Therefore, it is problematic to use results obtained from experiments using unfamiliar objects to draw conclusions about the representation of familiar objects.

To study the mental representation of familiar objects, one would like to be able to directly test the ability to generalize from familiar to novel views for familiar objects. The conundrum is how to find objects that have been identified many times in everyday life, but only from a restricted range of viewpoints. There are relatively few objects in the world that we are highly familiar with from certain views but not from others. One of them is our own face, which we experience from the daily glance into the mirror.

Typically, one has a disproportionately high exposure to near-frontal views of one's own face as seen in a mirror. The range of viewpoints is restricted by our ocular-motor system (ca. +/- 40 degs; Robinson, 1981), and apart from relatively rare situations (e.g. photographs and mirror arrangements at the barbershop), we do not see our own face in profile view. Nevertheless, the range of possible viewpoints is large enough in principal to provide enough information to reconstruct the full 3D structure of the head and thus all possible views (Ullman & Basri, 1991; Koenderink & van Doorn, 1991). If one's visual system relied on an object-centered representation for familiar faces, the difference between the distribution of views seen from one's own face compared with the distribution of views seen from other familiar faces should not influence recognition performance. If there is a difference for recognizing frontal and profile views of familiar faces of other people, the same difference should be observed for recognizing one's own face. If there is no difference between recognition of frontal and profile views for familiar faces, this would also be expected for recognizing one's own face. A view-based representation, on the other hand, predicts poorer recognition performance for the profile view of one's own face.

2 Material and Methods

2.1 Stimuli

We took color pictures of frontal and profile views of 26 members of our laboratory. The subjects were allowed to smile if they wanted to and they wore glasses if they did so normally. The pictures were all taken in front of a neutral grey wall. Images were digitized. The views ranged between 8 and 10 cm on a computer screen corresponding to a visual angle of about 6 degrees at the viewing distance of 85 cm.

2.2 Subjects

The same 26 people that served as models for the pictures also participated as observers in the experiment. All of them had worked in the lab for more than three months, saw each other at least once a week in the regular lab meeting (but usually much more often) and knew each other well by name.

2.3 Procedure

We used a naming paradigm measuring the time between stimulus onset and the beginning of the subject's response. Before the experiment, the procedure was explained in detail to the subjects while leaving them naive to the purpose of the experiment. We also prepared them not to be upset if they had a momentary memory block for a friend's name, an instance which happened to a few subjects once or twice. Before the experiment the subjects were shown a list of all occurring names, including their own, and were informed that they would be seeing images of themselves as well as the others. They were asked to call out the first name of the person shown in the image as quickly as possible. After starting the experimental run, all 54 images were shown, each separated by a blank screen (1000 msec) and a fixation cross (750 msec). Each image

remained on the screen until a microphone attached to the system registered an answer and response time was measured. The order of the images was randomized individually for each subject according to the following constraints: a.) Two successive images should not show the same face. b.) The subject's own face should not appear within the first 10 trials. c.) From the total of 26 subjects, 13 randomly chosen subjects saw first their frontal view and then their profile view, whereas the other 13 subjects saw first their profile view and then their frontal view.

2.4 Data analysis

Before the analysis we excluded all trials with response times longer than 2 standard deviations away from the mean response time calculated for each individual subject. This procedure excluded 53 outliers caused mainly by memory blocks. Not excluding these trials yields longer response times to the faces of the other people but still no difference between frontal and profile views. Errors were extremely infrequent (mean error rate: 0.7%) and were not analyzed.

3 Results

We measured response times for correct naming of frontal and profile views of 26 familiar faces including the subject's own face. Fig. 1 shows that response times to one's own face are faster than to other faces ($F(1,25) = 26.1$, $p < 0.01$), possibly due to greater familiarity with one's own name. There is also a main effect of face orientation ($F(1,25) = 13.2$, $p < 0.01$). Response times to frontal views are faster, than those to profile views. Most importantly, the results show an interaction between view and whether the image was that of the subject's own face or not ($F(1,25) = 6.4$, $p < 0.05$). Subjects can name profile views just as fast as frontal views when dealing with the familiar faces of their colleagues ($t(25) = 1.9$, $p > 0.05$), whereas they are significantly slower to recognize profiles than frontal views of themselves ($t(25) = 3.3$, $p < 0.005$). This result provides strong evidence that viewer-

centered representations are used even when processing highly familiar objects like one's own face.
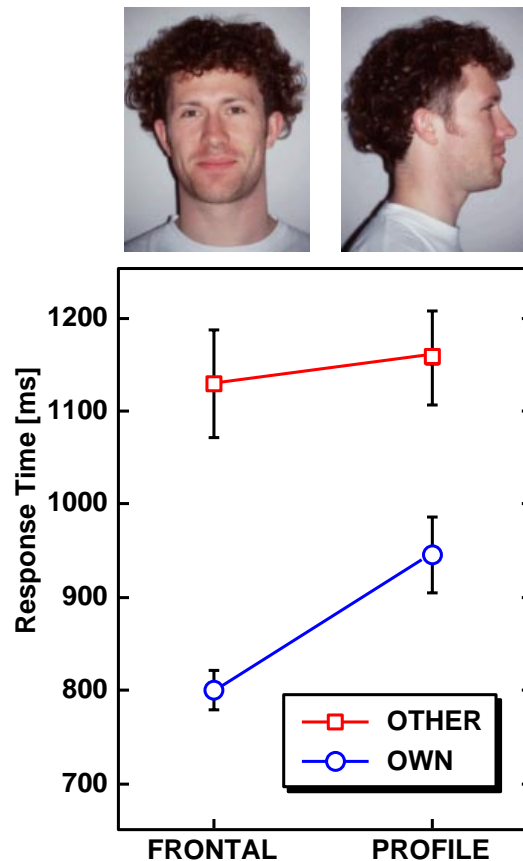

Figure 1: The graph shows the mean response times for naming frontal and profile views of a subject's own face and of all the other faces.

4 Discussion

The pronounced interaction between the two factors of interest provides strong evidence that viewer-centered representations are used even when processing highly familiar objects like one's own face. However there are two other outcomes that have to be mentioned.

Firstly, we observe that the mean response time for other faces is significantly longer than the one for one's own face. The task that subjects had to solve in this experiment contains actually two, conceptually different parts: the face has to be recognized first and subsequently a name has to be assigned to it. The difference in mean response times be-

tween other faces and one's own face is most likely due to greater familiarity to one's own name. Sometimes access to a name even of a good friend may be blocked, but never to one's own name. In fact, all the trials excluded as outliers were trials in which other faces were shown.

The second point refers to the slight trend suggesting that even with other familiar faces there is a small advantage for frontal views. This trend is not significant in our data but it is in accordance with findings by Bruce et al. (1987) as discussed above and it might reflect the fact that even for other faces the distribution of views we are exposed to is not completely homogenous.

There is still an ongoing discussion in the recent literature about the level on which differences between the processing of faces and the processing of other objects have to be described. We are aware that we have to be careful about extrapolating the conclusions of our results towards other object classes.

References

Biederman, I., & Gerhardstein, P. C. (1982). Recognizing depth-rotated objects: evidence and conditions for three-dimensional viewpoint invariance. *Journal of Experimental Psychology: Human Perception and Performance*, **19**, 1162 – 1182.

Bruce, V., Valentine, T., & Baddeley, A. (1987). The basis of the 3/4 view advantage in face recognition. *Applied cognitive psychology*, **1**, 109 – 120.

Bülthoff, H. H., & Edelman, S. (1992). Psychophysical support for a two-dimensional view interpolation theory of object recognition. *Proceedings of the National Academy of Science USA*, **89**, 60 – 64.

Hill, H., & Bruce, V. (1996). Effects of lighting on the perception of facial surfaces. *Journal of Experimental Psychology: Human Perception and Performance*, **22**, 986 – 1004.

Hill, H., Schyns, P. G., & Akamatsu, S. (1997). Information and viewpoint dependence in face recognition. *Cognition*, **62**, 201 – 222.

Koenderink, J. J., & van Doorn, A. (1991). Affine structure from motion. *Journal of the Optical Society of America*, **8**, 377 – 385.

Marr, D., & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three dimensional shapes. *Proceedings of the Royal Society of London B*, **200**, 269 – 294.

Palmer, S., Rosch, E., & Chase, P. (1981). Canonical perspective and the perception of objects. In J. Long & A. Baddeley (Eds.), *Attention and Performance IX* (pp. 135 – 151). Hillsdale, NJ: Lawrence Erlbaum.

Robinson, D. A. (1981). Control of eye movements. In V. B. Brooks (Ed.), *Handbook of Physiology Vol II Part 2* (pp. 1275 – 1320). Baltimore: Williams and Wilkins.

Rock, I., & DiVita, J. (1987). A case of viewer-centered object perception. *Cognitive Psychology*, **19**, 280 – 293.

Tarr, M. J. (1995). Rotating objects to recognize them: a case study on the role of viewpoint dependency in the recognition of three-dimensional objects. *Psychonomics Bulletin & Review*, **2**, 55 – 82.

Troje, N. F., & Bülthoff, H. H. (1996). Face recognition under varying poses: the role of texture and shape. *Vision Research*, **36**, 1761 – 1771.

Ullman, S. (1996). *High-level Vision*. Cambridge, MA: MIT Press.

Ullman, S., & Basri, R. (1991). Recognition by linear Combinations of Models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **13**, 992 – 1006.