Neuropsychologia 48 (2010) 2699–2708

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

Extrastriate cortical activity reflects segmentation of motion into independent sources

Gideon P. Caplovitz[a,*], Peter U. Tse[b]

[a] Princeton Neuroscience Institute, 3-N-13 Green Hall, Princeton University, Princeton, NJ 08544, United States
[b] Department of Psychological and Brain Sciences, HB 6207 Moore Hall, Dartmouth College, Hanover, NH 03755, United States

ARTICLE INFO

Article history:
Received 14 January 2010
Received in revised form 2 April 2010
Accepted 9 May 2010
Available online 15 May 2010

Keywords:
Rotational motion
Form motion interaction
V3v
V4v
V3A
V3B
Retinotopic cortex

ABSTRACT

Identical local image motion signals can arise from countless object motions in the world. In order to resolve this ambiguity, the visual system must somehow integrate motion signals arising from different locations along an object's contour. Difficulties arise, however, because image contours can derive from multiple objects and from occlusion. Thus, correctly integrating respective objects' motion signals presupposes the specification of what counts as an object. Depending on how this form analysis problem is solved, dramatically different object motion percepts can be constructed from the same set of local image motions. Here we apply fMRI to investigate the mechanisms underlying the segmentation and integration of motion signals that are critical to motion perception in general. We hold the number of image objects constant, but vary whether these objects are perceived to move independently or not. We find that BOLD signal in V3v, V4v, V3A, V3B and MT varies with the number of distinct sources of motion information in the visual scene. These data support the hypothesis that these areas integrate form and motion information in order to segment motion into independent sources (i.e. objects) thereby overcoming ambiguities that arise at the earliest stages of motion processing.

1. Introduction

The ability to encode and accurately represent motion is one of the fundamental functions of the visual system. Determining how this is accomplished is a primary goal of visual research. The problems that must be resolved in order to construct an accurate motion percept are non-trivial and arise at the earliest stages of visual processing. Due to the receptive field properties of neurons early in the visual processing stream, the local detection of motion is intrinsically ambiguous. An infinite number of possible real-world motions can produce identical local responses in neural populations that detect motion. The problem of solving this many-to-one mapping is commonly referred to as the 'aperture problem' (Adelson & Movshon, 1982; Fennema & Thompson, 1979; Marr, 1982; Nakayama & Silverman, 1988a, 1988b). It has been widely hypothesized that these ambiguities are resolved by processes that integrate local motion signals in the image together to produce percepts that, for the most part, accurately reflect the actual motion of an object in the visual scene (e.g. Adelson & Movshon, 1982; Bonnet, 1981; Burt & Sperling, 1981; Hildreth, 1984; Watson & Ahumada, 1985; Weiss & Adelson, 1998, 2000; Weiss, Simoncelli, & Adelson, 2002). Models such as these are based on the integration of motion signals arising along the contours of individually moving objects.

However, there is a problem with such models. It is often the case that the visual scene contains many moving objects, portions of whose contours may not be visible due to lighting, occlusion, or other environmental factors. As such, in order to properly integrate locally ambiguous motion signals, the visual system must first resolve which signals arise from the motion of which object. This process of segmentation defines the parameters within which mechanisms of motion integration may operate. Depending on the solution to the object segmentation problem, the manner in which motion signals are integrated together can lead to dramatically different perceptual outcomes that may or may not accurately reflect what is actually occurring in the world.

The motion of a rotating ellipse provides a simple perceptual framework for investigating these processes of form segmentation and motion integration. The spatio-temporal sequence of retinal images produced by a rotating ellipse is simultaneously consistent with rigid rotational and non-rigid deformational motion (Vallortigara, Bressan, & Bertamini, 1988; Wallach, Weisz, & Adams, 1956; Weiss & Adelson, 2000). However, despite this intrinsic ambiguity, observers generally perceive only one of these percepts for a given ellipse. Specifically, a high aspect-ratio 'skinny' ellipse rotating about its center in the 2D plane will most likely be perceived to rotate rigidly, whereas a low aspect-ratio 'fat' ellipse will most

---

[*] Corresponding author. Tel.: +1 215 495 5503; fax: +1 603 646 1419.
E-mail addresses: gcaplovi@princeton.edu (G.P. Caplovitz), Peter.Tse@dartmouth.edu (P.U. Tse).

0028-3932/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neuropsychologia.2010.05.017

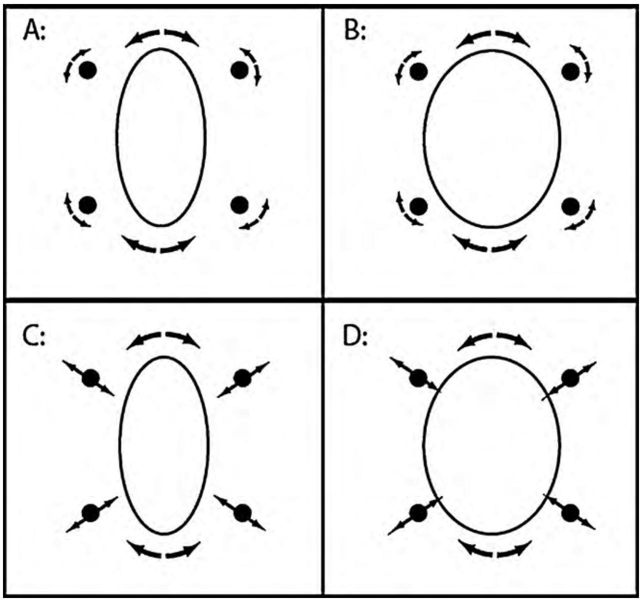

Fig. 1. Satellites and ellipses. When satellites rotate along with either a (A) skinny ellipse or (B) a fat ellipse, the percept of the ellipse's motion is that of rigid rotation. However, if the satellites translate in a radial manner, only the skinny ellipse (C) will appear to rotate rigidly. The fat ellipse (D) will be perceived to non-rigidly deform.

likely be perceived to deform non-rigidly, as if its contour were elastic, giving it a gelatinous appearance (Vallortigara et al., 1988; Wallach et al., 1956; Weiss & Adelson, 2000).

It has been hypothesized that this perceptual dissociation between rigid and non-rigid percepts reflects the mechanisms by which local motion signals are integrated along the elliptical contour (Grzywacz & Yuille, 1991; Hildreth, 1984; Horn & Schunck, 1981; Weiss & Adelson, 2000). For example, it has been hypothesized that locally detected motion signals are integrated across non-local regions in the image such that the resultant output will be optimized relative to certain constraints such as smoothness (Grzywacz & Yuille, 1991; Hildreth, 1984; Horn & Schunck, 1981; Weiss & Adelson, 2000). According to these models, local motion signals can be integrated such that the resultant global motion for a skinny ellipse is more consistent with rigid rotation, and the global motion for a fat ellipse is more consistent with non-rigidity or deformation.

Are the rigid and gelatinous cases distinct perceptual outcomes that are mediated by common neural mechanisms, as proposed by such models? Or are these two percepts mediated by distinct neural processes, contrary to such models? This is the first of two questions we seek to answer with the present research. Employing fMRI, we directly compare BOLD signals in response to stimuli that produce rigid and non-rigid responses. If these percepts are mediated by distinct neural mechanisms then we would expect to identify specific regions of visual cortex in which the responses differentiate between these two possible perceptual outcomes for the same stimulus.

Intriguingly, it has been shown that the motion of objects (satellites) distal to the elliptical contour can influence whether or not rigid rotation is perceived (Weiss & Adelson, 2000). In particular, a 'fat' ellipse can be made to appear to rotate rigidly by adding satellites that rotate with the same angular velocity, as if the dots were attached to the ellipse via invisible rods. If the satellites instead translate in a radial manner, maintaining a constant distance from the elliptical contour, as shown in Fig. 1, the same fat ellipse will appear to deform non-rigidly (Weiss & Adelson, 2000). In contrast, if the same satellite motion trajectories are added to a display in which a skinny ellipse is rotating, no such influence is observed. That is, no matter what the satellite trajectory is, the skinny ellipse will always be perceived to rotate rigidly. An example of the effects of satellites on ellipse rigidity can be observed in supplemental video 1.

This dissociation reflects the segmentation problem that the visual system must resolve in order to distinguish the motion of one object from another. Understanding the nature of these segmentation and integration processes and how they relate to percepts of rigid and non-rigid motion and motion perception in general has been the subject of both psychophysical research (e.g. Burr, Baldassi, Morrone, & Verghese, 2009; Caplovitz & Tse, 2006; Kohler, Caplovitz and Tse, 2009; Lorenceau & Shiffrar, 1992; Tse & Logothetis, 2002; Tse, 2006; Verghese & Stone, 1996; Weiss & Adelson, 2000) and computational modeling (Berzhanskaya, Mingola and Grossberg, 2007; Weiss & Adelson, 2000). What are the neural mechanisms that underlie these segmentation and integration processes? This is the second question we seek to address with the current research.

Existing neuroimaging research into the neuronal basis of motion segmentation has largely focused on the motion signals arising from translating dot fields (Muckli, Singer, Zanella, & Goebel, 2002) or drifting gratings (Castelo-Branco et al., 2002) that can either be integrated into a single coherently translating motion field or segmented into multiple transparent layers. Castelo-Branco et al. (2002) found increased activity in hMT+ and left V3/V3A when superimposed drifting gratings were segmented and perceived as transparent layers compared to when they were integrated into a coherently moving plaid. Similarly, Muckli et al. (2002) found increased activity in hMT+, the right posterior intraparietal sulcus and a portion of the fusiform gyrus when translating dot fields were segmented into transparent layers. Are the same mechanisms as those underlying the segmentation of translational motion signals into independent layers recruited for the case of moving objects and the segmentation of rotational motion into independent layers? The distinct challenges that rotational motion poses to the visual system suggest that additional processes may need to be recruited in order to resolve the ambiguities of rigid and non-rigid motion.

Here we employ fMRI to localize the mechanisms underlying the motion segmentation processes that mediate the influences of satellites on the disambiguation of the rigid and non-rigid percepts that can arise from a single continuously rotating ellipse. Specifically, we contrast the difference in BOLD signal activation between (1) conditions in which radial and rotational satellites move in conjunction with skinny ellipses, where radial satellite motion does not induce ellipse non-rigidity, with (2) conditions in which the same satellites move in conjunction with fat ellipses, where radial satellite motion does induce ellipse non-rigidity.

Hypotheses: Because the motions of the satellites and elliptical contours are only segmented into separate sources of motion in the radial skinny ellipse condition, we hypothesize that areas of visual cortex that are selectively involved in the processes associated with segmentation of motion into independent sources will show differential activity in the two skinny ellipse conditions (Fig. 2 top) and not in the two fat ellipse conditions. On the other hand, since the non-rigid percept is only present in the radial fat ellipse condition we hypothesize that areas of visual cortex selectively involved in dissociating rigid and non-rigid motion will show differential activity in the two fat ellipse conditions and not in the two skinny ellipse conditions (Fig. 2 bottom).

Summary of results: We find differential activity selective for the skinny ellipse conditions in several areas of extrastriate cortex. This activation likely reflects mechanisms underlying processes associated with the segmentation of image motions into independent sources or objects. In contrast, our analyses yield no evidence for regions of visual cortex that selectively dissociate rigid from non-rigid motion percepts.