Available online at www.sciencedirect.com
SCIENCE DIRECT®
Vision Research 45 (2005) 3189–3201
Vision Research
www.elsevier.com/locate/visres

Global shape coding for motion-defined radial-frequency contours

Stéphane J.M. Rainville [a,*,1], Hugh R. Wilson [b]

[a] Center for Visual Neuroscience, Department of Psychology, North Dakota State University, Fargo, ND 58105-5075, USA
[b] Center for Vision Research, York University, 4700 Keele Street, North York, Ont., Canada M1J 1P3

Received 4 December 2004; received in revised form 23 June 2005

Abstract

The visual system is highly skilled at recovering the shape of complex objects defined exclusively by motion cues. But while low-level and high-level mechanisms involved in shape-from-motion have been studied extensively, intermediate computational stages remain poorly understood. In the present study, we used motion-defined radial-frequency contours—or motion RFs—to probe intermediate stages involved in the computation of motion-defined shape. Motion RFs consisted of a virtual circle of Gabor elements whose carriers drifted at speeds determined by a sinusoidal function of polar angle. Motion RFs elicited vivid percepts of shape, and observers could detect and discriminate radial frequencies up to approximately five cycles. Randomizing Gabor speeds over a small contour segment impaired detection and discrimination performance significantly more than predicted by probability summation. Threshold comparisons between spatial-RF and motion-RF contours ruled out that motion-induced shifts in perceived position (i.e., the DeValois effect) determine shape perception in motion RFs. Together, results indicate that the shape of motion RFs is processed by synergistic mechanisms that perform a global analysis of motion cues over space. These results are integrated with data on perceptual interactions between motion RFs and spatial-RFs [Rainville & Wilson (2004). *Vision Research* 44(11), 1065–1077] and are discussed in terms of cue-specific and cue-invariant representations of object shape in human vision.

Keywords: Motion; Form-from-motion; Structure-from-motion; Shape-from-motion

1. Introduction

Human vision is highly skilled at recovering the shape of objects defined exclusively by motion cues. Classic examples of shape-from-motion (SFM) include dot displays where no shape is perceived when dots remain static but a vivid percept of shape arises when dots are animated (Gibson, 1986; Johansson, 1973; Todd, 1985; Ullman, 1979; Wallach & O'Connell, 1953). SFM stimuli are inherently ambiguous as the problem of recovering shape-from-motion is mathematically under-constrained and can lead to more than one interpretation. To recover shape-from-motion, the brain must constrain the problem using internal rules and assumptions about the visual environment. The solution that vision has found to the SFM problem is remarkably general, as it operates on a seemingly infinite set of shapes and handles object complexity, non-rigidity, and three-dimensionality with ease. How the brain computes SFM and constrains the solution remains a fundamental issue in visual neuroscience.

Psychophysical studies have relied on two major paradigms—henceforth labelled as the atomistic and holistic approaches—to probe neural mechanisms mediating SFM. The atomistic approach uses basic stimuli such as Gabors (i.e., Gaussian-windowed sinusoidal gratings) to infer the properties of simple mechanisms. These basic stimuli are then combined into more complex stimuli designed to test the rules by which simple mechanisms

* Corresponding author. Tel.: +1 701 231 8625; fax: +1 701 231 8426.
E-mail address: rainvill@yorku.ca (S.J.M. Rainville).
[1] Present address: Center for Vision Research, York University, 4700 Keele Street, North York, Ontario, Canada M3J 1P3. Tel.: +1 416 736 2100x33325; fax: +1 416 736 5857.

0042-6989/$ - see front matter
doi:10.1016/j.visres.2005.06.033

interact and build a representation of motion-defined shape. Atomistic studies have revealed that motion signals are computed locally (Anderson & Burr, 1989; Georgeson & Scott-Samuel, 2000; Qian, Andersen, & Adelson, 1994; Rainville, Scott-Samuel, & Makous, 2002) and pool their outputs to recover local motion speed and direction unambiguously (Adelson & Movshon, 1982; Wilson, Ferrera, & Yo, 1992). Evidence suggests local motion signals interact across space (Loffler & Orbach, 2003; Yuille & Grzywacz, 1988) and serve as input to further stages sensitive to more complex motion properties such as rotation and expansion (Bex & Dakin, 2002; Burr, Morrone, & Vaina, 1998; Lorenceau & Zago, 1999). Atomistic studies also suggest the existence of mechanisms sensitive to motion-defined contours and curvature (Bex, Simmers, & Dakin, 2003; Grzywacz, Watamaniuk, & McKee, 1995; Ledgeway & Hess, 2002; Loffler & Wilson, 2001). The main challenge for the atomistic approach is to probe vision with increasingly complex SFM stimuli whose structure can be represented by models that combine outputs from simple mechanisms in biologically plausible ways.

As an alternative to the atomistic approach, the holistic approach relies on SFM stimuli whose structure is highly complex and/or difficult to describe mathematically, as in the case of natural scenes. These stimuli are then manipulated (or "deconstructed") to probe the properties of visual mechanisms sensitive to SFM. Point-light walkers—stimuli defined only by dots attached to the joints of an animated human figure—are representative of holistic studies on SFM (Johansson, 1973). Even without a complete physical description of their structure, point-light walkers have been used to reveal spatial and temporal integration properties of mechanisms sensitive to biological motion (Giese & Lappe, 2002; Neri, Morrone, & Burr, 1998; Tadin, Lappin, Blake, & Grossman, 2002). The perception of complex motion-defined surfaces and objects has also been explored even if the properties of the retinal image were not fully understood or the stimulus set was too narrowly defined to allow generalization to other stimulus classes (Caudek & Rubin, 2001; Hildreth, Ando, Andersen, & Treue, 1995; Mukai & Watanabe, 1999; Norman & Lappin, 1992; Sperling & Landy, 1989; Watanabe, 1997). Due to the complexity and/or specificity of the stimuli it uses, the holistic approach faces difficulties in modeling and linking results with mechanisms identified by the atomistic approach.

Despite an abundance of data, the visual mechanisms mediating SFM remain poorly understood. Psychophysical data from atomistic and holistic studies are consistent with a visual hierarchy where lower-level signals are selectively combined to represent increasingly complex properties of motion-defined shape (Nakayama, He, & Shimojo, 1995). But while data from the two approaches converge, stimuli remain either too simple or too complex to investigate the processing stages between lower-level and higher-level mechanisms involved in SFM. What is required is a paradigm that investigates intermediate-level stages with a stimulus set that has enough complexity and variety to cover a large space of ecologically valid shapes and yet remains sufficiently simple and well-defined to allow modeling from an atomistic perspective.

In the present study, we investigated intermediate-level SFM using motion-defined radial-frequency stimuli (henceforth referred to as motion RFs). Motion RFs consist of drifting Gabor elements (i.e., sinewaves drifting behind static Gaussian windows) positioned and oriented such as to form a virtual circle (see Fig. 1). While the perceived shape of the stimulus is circular if the stimulus remains static, observers experience a dramatic distortion of the circular shape when Gabors are set in motion—for instance, the circle can be perceived as an ellipse, a square, or other shapes depending on the geometry of the velocity field. Motion RFs have a significant advantage over other SFM stimuli such as point-light walkers: due to the fact that Gabor position is fixed, the shape induced by the Gabor velocity field remains constant and can be studied in its steady state for arbitrarily long stimulus durations.

The geometry of a motion RF's velocity field is typically determined by a single sinusoidal function of polar angle which determines the speed at which each Gabor element is drifting. Through a linear combination of sinusoidal frequencies with appropriate amplitudes and phases (i.e., a Fourier synthesis), the velocity field of motion RFs can perceptually distort circles into single-centroid shapes of arbitrary complexity. The ensemble of motion RFs therefore defines a large and well-parameterized space of motion-defined shapes. The radial-frequency paradigm has been applied successfully to the study of shape perception in the spatial domain with contours defined by position rather than speed (Habak, Wilkinson, Zakher, & Wilson, 2004; Loffler, Wilson, & Wilkinson, 2003; Wilkinson, Wilson, & Habak, 1998) and lends itself well to modeling (Poirier & Wilson, 2004).

Psychophysical experiments reported herein measured the ability of human observers to detect, discriminate, and integrate motion RFs. Results from these experiments show that shape-from-motion is limited to stimuli whose velocity field varies smoothly over space, and that the coding of motion-defined shape is a global process that integrates local motion information synergistically over the extent of the stimulus. Control experiments ruled out the possibility that coding for motion RF exploits illusory positional artefacts reported in studies with drifting Gabor elements (De Valois & De Valois, 1991; Hayes, 2000). Together, results indicate this rich but well-parameterized stimulus set of motion RF contours can successfully probe intermediate-level