Vision Research 126 (2016) 131–142

Contents lists available at ScienceDirect

Vision Research

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

Modeling spatiotemporal boundary formation

Gennady Erlikhman *, Philip J. Kellman

Department of Psychology, University of California, Los Angeles, United States

ARTICLE INFO

Article history:
Received 9 February 2015
Received in revised form 23 March 2015
Available online 12 April 2015

Keywords:
Boundaries
Illusory contours
Perceptual organization
Grouping
Perceptual model

ABSTRACT

Spatiotemporal boundary formation (SBF) refers to perception of continuous contours, shape, and global motion from sequential transformations of widely separated surface elements. How such minimal information in SBF can produce whole forms and the nature of the computational processes involved remain mysterious. Formally, it has been shown that orientations and motion directions of local edge fragments can be recovered from small sets of element changes (Shipley & Kellman, (1997). *Vision Research*, *37*, 1281–1293). Little experimental work has examined SBF in simple situations, however, and no model has been able to predict human SBF performance. We measured orientation discrimination thresholds in simple SBF displays for thin, oriented bars as a function of element density, number of element transformations, and frame duration. Thresholds decreased with increasing density and number of transformations, and increased with frame duration. An ideal observer model implemented to give trial-by-trial responses in the same orientation discrimination task exceeded human performance. In a second group of experiments, we measured human precision in detecting inputs to the model (spatial, temporal, and angular inter-element separation). A model that modified the ideal observer by added encoding imprecision for these parameters, directly obtained from Exp. 2, and that included two integration constraints obtained from previous research, closely fit human SBF data with no additional free parameters. These results provide the first empirical support for an early stage in shape formation in SBF based on the recovery of local edge fragments from spatiotemporally sparse element transformation events.

1. Introduction

A primary goal of the visual system is to use information in reflected light to perceive objects and surfaces. Crucial among the processes involved is detection of edges and surface boundaries, for which there are many cues, including discontinuities in luminance contrast, color, stereoscopic disparity, and texture. However, these cues may sometimes be insufficient, when depth differences are below threshold, in poorly lit environments, or where only sparse surface elements are visible. In such cases, surface boundaries can often be revealed by object or observer motion. Dynamic cues, especially the accretion and deletion of texture (Gibson et al., 1969), can provide sufficient information for the segmentation of similarly or sparsely textured surfaces and can result in the perception of boundaries, surfaces, and global motion (Kaplan, 1969; Andersen & Cortese, 1989; Gibson et al., 1969; Yonas, Craton, & Thompson, 1987; Stappers, 1989; Shipley & Kellman, 1993, 1994).

Although accretion and deletion of texture has been described primarily as a cue to relative depth (Gibson et al., 1969), it has also been noted that it produces perception of shape in the absence of any other cues to shape (Andersen & Cortese, 1989; Gibson et al., 1969; Shipley & Kellman, 1993, 1994). These latter phenomena pose a mystery. The perception of continuous illusory contours (and the shapes they delineate) across empty surface regions between elements does not obviously follow from the perception of occlusion of an element.

Shipley and Kellman (1993, 1994) found that gradual occlusion of elements was not even necessary, as discrete element disappearance also produces perceptions of boundaries and surfaces across gaps. Further, no form of accretion and deletion of texture elements, continuous or discrete, is needed. The visual system appears to use *any* abrupt change in local elements as inputs to a process that produces perceived edges, form, and global motion. Changes in element orientation, shape, color, or position all produced these effects, and they labeled this more general process of perception of continuous illusory boundaries and global form from sequential changes in local surface elements *spatiotemporal boundary formation* (SBF).

---

* Corresponding author.
E-mail address: gerlikhman@unr.edu (G. Erlikhman).

http://dx.doi.org/10.1016/j.visres.2015.03.016

How do local element changes produce the continuous boundaries seen in SBF? It has been proposed that shape in SBF depends on two processing stages (Shipley & Kellman, 1994, 1997). First, information from sets of element changes in small neighborhoods somehow produce local, oriented edge fragments. Second, these edge fragments connect to each other across gaps according to well-known interpolation processes that operate in the perception of illusory and occluded contours (Fantoni & Gerbino, 2003; Grossberg & Mingolla, 1985; Kanizsa, 1979; Michotte, Thines, & Crabbe, 1964; Kellman & Shipley, 1991; Palmer et al., 2006). Whereas the second stage involves processes that are well-understood, the first stage has remained mysterious. Shipley and Kellman (1994) showed mathematically that a local orientation could be derived from three sequential non-collinear element transformations. Little empirical research, however, has examined SBF with single edges and relatively few elements. Virtually all previous studies of SBF have used closed objects with smooth contours as stimuli (although see Barraza & Chen, 2006). Recently, we demonstrated that individual, oriented, illusory edge fragments can be recovered from sparse displays (Kellman et al., 2012). These results support the two-level theory of SBF, specifically in implicating a process that recovers local oriented edge fragments. These fragments are likely the basic units from which larger shapes are constructed in SBF.

Here we sought to develop and test a process model of how such edges are extracted. We implemented and tested an ideal observer model of edge extraction in SBF displays, based on the idea that triplets of sequential element transformations can provide an estimate of a local, oriented edge fragment. In Experiment 1, we measured orientation discrimination thresholds for SBF-defined edges across a variety of display properties. Human performance was much worse than the ideal observer model. Unlike the model, human performance may involve noise in registering relevant inputs as well as limits on information accumulation. In a second experiment, we used separate paradigms to measure noise in human registration of basic input features, such as inter-element separation. A model that incorporated simple information accumulation constraints and the measured spatial and temporal noise parameters in Experiment 2 was able to accurately predict human performance from Experiment 1 across all tested display conditions.

1.1. Background: SBF displays and models

In this section, we briefly review SBF phenomena and prior models. Fig. 1 shows an example of an SBF display. The dotted line defines the boundary of a virtual object. The elements are always stationary and the virtual object moves across the display. As the object moves, elements that fall within the boundary change in some property, such as color. The change is discrete, and the percept is of a moving figure with clear boundaries. In *unidirectional transformations*, elements initially have one value (e.g., white dots on a black background) and when they become encompassed within the virtual region, they change to a different value (e.g., white dots turn blue). Upon exiting the region, elements revert to their original value (e.g., blue dots revert to white). In *bidirectional transformations*, elements are randomly assigned one of two values and switch to the other upon entering or exiting the boundary of the moving object. For example, with blue and white dots on a black background, blue dots turn white upon entering the virtual region, and white dots turn blue. SBF occurs across a wide variety of parameters, with the precision of shape perception depending on element density, luminance differences between elements, the velocity of the virtual region, and frame duration (Andersen & Cortese, 1989; Cicerone et al., 1995; Shipley & Kellman, 1994).

In SBF, no single frame has visible edges of a shape. In some unidirectional transformation displays, there will be a region of elements having a different feature value from surrounding elements, but the shape of this region is not well specified. Other unidirectional transformations, such as local element motion, as well as all bidirectional transformations, offer no information in any static frame about shape or about any affected region. Because elements transform all at once, there is no oriented contour information as might be given by gradual occlusion of an object or texture element. Thus, in SBF, local edges are not given by any of the standard cues for edge perception. Moreover, even for a mechanism attempting to extract local edge fragments from local changes in element properties, SBF displays pose a difficult variant of the aperture problem (Adelson & Movshon, 1982; Wallach, 1935), what has been referred to as the “point aperture problem”, in which *neither* the orientation nor velocity of an edge are directly given in the stimulus (Prophet, Hoffman, & Cicerone, 2001; Shipley & Kellman, 1994, 1997). In the point aperture problem, there are no oriented edge fragments given in the stimulus. The visual system must simultaneously recover *both* the orientation and motion of a local edge from sparse and discrete element transformations.

A solution to the point aperture problem was proposed by Shipley and Kellman (1994, 1997). Given the positions and times of occurrence of three, non-collinear element transformations, the orientation of an edge that caused those transformations can be computed assuming a constant edge velocity and orientation (Shipley & Kellman, 1997). An intuition for the proof appears in Fig. 2. Fig. 2a depicts a sequence of element transformations (labeled 1, 2, and 3) caused by a moving edge. When two elements transform (in this case, disappear and reappear) in succession, a transformation vector, $\mathbf{v}_{12}$, is formed between them. The magnitude of the vector is determined by the spatial and temporal separation of the transformations. We use the term “transformation

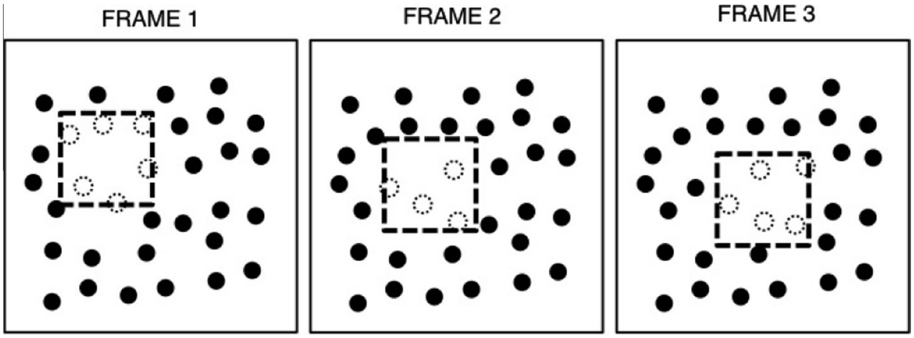

Fig. 1. Depiction of a square “virtual region” moving over a field of circular black elements. All elements inside the square region are in one state (white) and all those outside are in another (black). As the square moves (frames 2 and 3), elements entering and exiting the region change states. The resulting percept is of a moving region with crisply defined illusory contours.