Vision Research 59 (2012) 45–63

Contents lists available at SciVerse ScienceDirect

Vision Research

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

Motion perception induced by dynamic grouping: A probe for the compositional structure of objects

Howard S. Hock [a,*], David F. Nichols [b]

[a] Department of Psychology, and the Center for Complex Systems and Brain Sciences, Florida Atlantic University, Boca Raton, FL 33431, United States
[b] Department of Psychology, Roanoke College, Salem, VA 24153, United States

ARTICLE INFO

Article history:
Received 31 March 2011
Received in revised form 14 November 2011
Available online 25 February 2012

Keywords:
Grouping
Perceptual organization
Affinity
State dependence
Nonlinearity
Super-additivity
Motion
High level motion
Line motion illusion

ABSTRACT

A new method is described for determining how the visual system resolves ambiguities in the compositional structure of multi-surface objects; i.e., how the surfaces of objects are grouped together to form a hierarchical structure. The method entails dynamic grouping motion, a high level process in which changes in a surface (e.g., increases or decreases in its luminance, hue or texture) transiently perturb its affinity with adjacent surfaces. Affinity is determined by the combined effects of Gestalt and other grouping variables in indicating that a pair of surfaces forms a subunit within an object's compositional structure. Such pre-perturbation surface groupings are indicated by the perception of characteristic motions across the changing surface. When the affinity of adjacent surfaces is increased by a dynamic grouping variable, their grouping is transiently strengthened; the perceived motion is away from their boundary. When the affinity of adjacent surfaces is decreased, their grouping is transiently weakened; the perceived motion is toward the surfaces' boundary. It is shown that the affinity of adjacent surfaces depends on the nonlinear, super-additive combination of affinity values ascribable to individual grouping variables, and the effect of dynamic grouping variables on motion perception depends on the prior, pre-perturbation affinity state of the surfaces. It is proposed that affinity-based grouping of an object's surfaces must be consistent with the activation of primitive three-dimensional object components in order for the object to be recognized. Also discussed is the potential use of dynamic grouping for determining the compositional structure of multi-object scenes.

1. Introduction

A major legacy of the Gestalt Psychology movement of the early 20th century was the determination that perceptual organization is based on laws of grouping. Originally delineated by Wertheimer (1923), the grouping laws characterize the effect of various stimulus attributes on perceptual organization. How the component surfaces of a stimulus are grouped together depends on such factors as closure, proximity, similarity, movement direction (common fate), and good continuation. However, despite a long history of perceptual research, these well-known grouping properties have not been incorporated into a framework that could form the basis for a theory of object recognition (Palmer, 1999; Palmer & Rock, 1994). The emphasis instead has been on the extraction of three-dimensional geometric primitives (Biederman, 1987; Marr & Nishihara, 1978; Pentland, 1987). This contrasts with research in artificial vision, for which grouping properties have been central to models of object recognition (e.g., Iqbal & Aggarwal, 2002; Lowe, 1987; McCafferty, 1990; Sarker & Boyer, 1993).

One reason for slow progress in grouping-based approaches to human object recognition is that with some exceptions (e.g., Adelson, 1993; Palmer & Rock, 1994), most studies of perceptual grouping have involved arrays of disconnected surfaces designed to isolate effects of particular grouping variables (e.g., Kubovy & Wagemans, 1995; Palmer, Neff, & Beck, 1996; Rush, 1937; Wertheimer, 1923). While these studies have been valuable, objects are not composed of disconnected surfaces. They are composed of adjacent, connected surfaces, with multiple grouping variables that compete or cooperate in determining the perceptual organization of the object's surfaces.

Another possible reason for slow progress in the development of grouping-based theories of object recognition is methodological. Previous methods typically *assume* intuitively reasonable ways in which surfaces are grouped together (usually based on Gestalt principles), and confirm the assumptions by assessing performance in a variety of information processing tasks. For example, grouping a target with distractors reduces spatial resolution in target detection (Banks & Prinzmetal, 1976), the time required to find a diagonal line segment in an array of vertical or horizontal line segments depends on whether the array is organized into horizontal rows or vertical columns (Carrasco & Chang, 1995), and the same-different

* Corresponding author. Fax: +1 561 297 2160.
E-mail address: hockhs@fau.edu (H.S. Hock).

0042-6989/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.visres.2011.11.015

comparison of two target stimuli is faster when both are grouped within connected line segments (Feldman, 2007).

In contrast, the methodology proposed in this article is aimed at *discovering* how an object's surfaces are grouped together, irrespective of whether the grouping is intuitive or otherwise, and does so directly rather than indirectly through performance in an information processing task. The method entails the perception of motions that are evident to individual, naïve observers, so their effects can be reliably established without extensive testing. Most importantly, the method can address a wide range of issues, enhancing prospects for a grouping-based theory of object recognition.

1.1. Compositional structure

How the surfaces of an object are grouped by the perceiver determines its compositional structure, by which we mean a hierarchical tree structure in which surfaces are combined into subunits, subunits are combined into larger groupings, and so on, depending on the complexity of the object. Such hierarchical structures have been described by Palmer (1977), Cutting (1986) and Feldman (1999). The empirical challenge stems from the ambiguity inherent in compositional structure. For example, when an object has been parsed into three surfaces (call them A, B, and C), a compositional hierarchy could group A with B, and the AB subunit with C (i.e., AB–C). Alternative groupings, (BC–A) and (AC–B), would constitute competing compositional structures. The proposed method determines how this ambiguity is resolved by the visual system; i.e., which of the alternative compositional structures is selected.[1]

Consistent with the primacy of surfaces (Gibson, 1954) and the potential sufficiency of two-dimensional surfaces and their boundaries for the formation of three-dimensional object representations (Marr, 1982), the stimuli tested at this initial stage of research are object-like to the extent that they are composed of adjacent surfaces whose organization depends on sometimes cooperating and sometimes competing grouping variables.[2] Grouping properties are characterized as variables because in most cases their contributions to perceptual organization can vary continuously.

1.2. Affinity and the perception of motion

The conceptual lynchpin for the reported experiments is *affinity*, which entails any variable affecting the likelihood of two surfaces being grouped together. It is derived from Ullman's (1978, 1979) account of how the visual system solves the motion correspondence problem, which arises when there are competing possibilities for the perception of apparent motion from an initially presented surface to one of two or more surfaces presented afterward. Such ambiguities are resolved by differences in the affinity of the initially presented surface with each of the subsequently presented surfaces.

The current study follows Ullman in that differences in affinity resolve ambiguities, but now for ambiguities in perceptual organization. It departs from Ullman in that changes in affinity result in the perception of motion *within* one of two adjacent surfaces, instead of motion *between non*-adjacent surface locations. In addition, the concept of affinity is extended to account for how multiple grouping variables combine to affect the grouping strength of pairs of surfaces. It is shown that overall affinity is determined by the nonlinear summation of affinity values ascribable to individual grouping variables.

[1] Although relevant, aspects of compositional structure entailing spatial relationships between an object's parts (Biederman, 1987; Barenholtz & Tarr, 2007) and global regularities like symmetry (Leeuenberg, 1971; Wagemans, 1997) are not addressed in this article.

[2] Because of interposition in the two-dimensional projection of three-dimensional scenes, surfaces that are retinally adjacent may not always belong to the same object.

In the experiments that follow, some grouping variables remain unchanged during the course of a trial while others take on different values, quantitatively increasing or decreasing the relative affinity of surface pairs without qualitatively changing their perceptual organization. We call the latter "dynamic grouping variables." We have found that changes in affinity created by dynamic grouping variables can be sufficient to elicit the perception of motion. Previous studies have provided evidence that grouping can affect motion perception (Kramer & Yantis, 1997; Martinovic et al., 2009), but to our knowledge this is the first to indicate that grouping processes themselves can be the source of perceived motion. It is from dynamic grouping motion that we determine how the visual system resolves ambiguities in the compositional structure of objects.

1.3. The line motion illusion and dynamic grouping

The objective of Experiment 1 is to empirically establish the dynamic grouping phenomenon for stimuli composed of two adjacent surfaces. Because there are only two surfaces, resolving ambiguity in compositional structure is not an issue, as it will be in Experiments 2–4. The starting point is the line motion illusion (Hikosaka, Miyauchi, & Shimojo, 1993), which previously was called polarized gamma motion (Kanizsa, 1978, 1979). The illusion is created by presenting one surface, then presenting another surface next to it. Although the entire second surface appears simultaneously, motion is perceived away from the initial surface (in Fig. 1a it looks as if the square is expanding into a horizontal bar). When the second surface is removed, perceived motion is in the opposite direction (in Fig. 1b it looks as if the bar is contracting back into a square).

Experiment 1 is based on a version of the line motion illusion in which two adjacent surfaces are always visible (Fig. 1c and d). The connectivity of the surfaces, the alignment of their horizontal edges (i.e., good continuation), and their luminance similarity are grouping variables that combine to determine the affinity of the two surfaces. Changing the lighter surface's luminance changes its similarity with the adjacent, unchanged surface. This is the dynamic grouping variable, which perturbs the affinity of the two surfaces and induces the perception of motion across the changing surface.

Dynamic grouping motion is phenomenologically similar to the line motion illusion (Movie 1). Although somewhat weaker, the perceived motion is consistently reported by naïve observers. When the lighter surface's luminance decreases for the stimulus in Fig. 1c, its similarity with the unchanged surface increases, transiently strengthening the grouping of the surfaces. The motion perceived across the changing surface is then away from its boundary with the unchanging surface. When the lighter surface's luminance increases (Fig. 1d), its similarity with the unchanged surface decreases, transiently weakening the grouping of the surfaces. Motion again is perceived across the changing surface, but now toward its boundary with the unchanging surface.

These perceived dynamic grouping motion directions – away from and toward the boundary of two surfaces – are characteristic for pairs of surfaces that are grouped together. The strength of the grouping depends on the surfaces' affinity. As indicated earlier (Section 1.2), it will be shown that: (1) the overall affinity for a pair of surfaces is determined by the nonlinear summation of affinity values ascribable to individual grouping variables (in Experiment 1, connectivity, good continuation and luminance similarity), and (2) the strength of the motion induced by perturbing the surfaces' affinity (i.e., by increasing or decreasing luminance similarity) depends on the surfaces' pre-perturbation affinity state.

1.4. Resolving ambiguities in compositional structure

The central premise of the current study is that perturbations in affinity that result in the perception of dynamic grouping motion