Vision Research 77 (2013) 1–9

Contents lists available at SciVerse ScienceDirect

Vision Research

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

Perceptual learning of second order cues for layer decomposition

Dicle N. Dövencioğlu, Andrew E. Welchman, Andrew J. Schofield [*]

School of Psychology, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

ARTICLE INFO

Article history:
Received 1 August 2012
Received in revised form 6 November 2012
Available online 28 November 2012

Keywords:
Second-order
Layer-decomposition
Perceptual-learning

ABSTRACT

Luminance variations are ambiguous: they can signal changes in surface reflectance or changes in illumination. Layer decomposition—the process of distinguishing between reflectance and illumination changes—is supported by a range of secondary cues including colour and texture. For an illuminated corrugated, textured surface the shading pattern comprises modulations of luminance (first order, LM) and local luminance amplitude (second-order, AM). The phase relationship between these two signals enables layer decomposition, predicts the perception of reflectance and illumination changes, and has been modelled based on early, fast, feed-forward visual processing (Schofield et al., 2010). However, while inexperienced viewers appreciate this scission at long presentation times, they cannot do so for short presentation durations (250 ms). This might suggest the action of slower, higher-level mechanisms. Here we consider how training attenuates this delay, and whether the resultant learning occurs at a perceptual level. We trained observers to discriminate the components of plaid stimuli that mixed in-phase and anti-phase LM/AM signals over a period of 5 days. After training, the strength of the AM signal needed to differentiate the plaid components fell dramatically, indicating learning. We tested for transfer of learning using stimuli with different spatial frequencies, in-plane orientations, and acutely angled plaids. We report that learning transfers only partially when the stimuli are changed, suggesting that benefits accrue from tuning specific mechanisms, rather than general interpretative processes. We suggest that the mechanisms which support layer decomposition using second-order cues are relatively early, and not inherently slow.

1. Introduction

Interpreting the luminance variations in an image in terms of their underlying physical cause poses a significant challenge to the visual system. Specifically, luminance variations in an image can have two distinct causes: (i) they might arise from variations in 3D surface geometry so that different portions of a surface are differentially illuminated by the light source(s) and/or (ii) they might arise from variations in the surface albedo, such as different textures or paint on the surface. Somehow, the visual system should parse changes caused by the illumination with respect to the 3D surface (shape-from-shading) from changes in surface reflectance properties. This process is known as layer decomposition (Kingdom, 2008) or intrinsic image extraction (Barrow & Tanenbaum, 1978) and it can be achieved by considering the relationship between luminance variations and a range of other cues including colour (Kingdom, 2003) and, as we review below, second-order cues that arise in objects with a textured surface.

A potentially informative cue to layer decomposition is provided by the spatial relationship between changes in local mean luminance (LM) and local variations in the range of luminance values that arise from an albedo texture: local luminance amplitude (AM; Schofield et al., 2006). In particular, when the illumination varies across an albedo textured surface, changes in local mean luminance (LM) are positively correlated with changes in local luminance amplitude (AM). Adding an albedo texture to a shaded surface, such that LM and AM correlate positively (in-phase; LM + AM), enhances the impression of depth (Schofield et al., 2006, 2010; Todd & Mingolla, 1983; compare Fig. 1A and C). Further, if LM and AM are negatively correlated (anti-phase; LM − AM) the impression of depth is reduced (compare Fig 1D with A). If both relationships are present in a plaid configuration, the in-phase paring appears as a shaded undulating surface whereas the anti-phase paring appears as a flat material change (Schofield et al., 2006, 2010; Fig. 1E). The enhanced shape-from-shading in the in-phase case may be due to improved layer decomposition due to the information provided by the relative phase of the AM cue.

The changes in local luminance amplitude described above are, mathematically, closely related to the contrast modulations typically used to study second-order vision. The human visual system is known to be sensitive to second order signals and it is thought that they are detected separately from first order cues (Baker, 1999; Dakin & Mareschal, 2000; Ellemberg, Allen, & Hess, 2006; Fleet & Langley, 1994; Schofield & Georgeson, 1999, 2003). First- and second-order information are correlated in natural

---

[*] Corresponding author.
E-mail address: A.J.Schofield@bham.ac.uk (A.J. Schofield).

0042-6989
http://dx.doi.org/10.1016/j.visres.2012.11.005

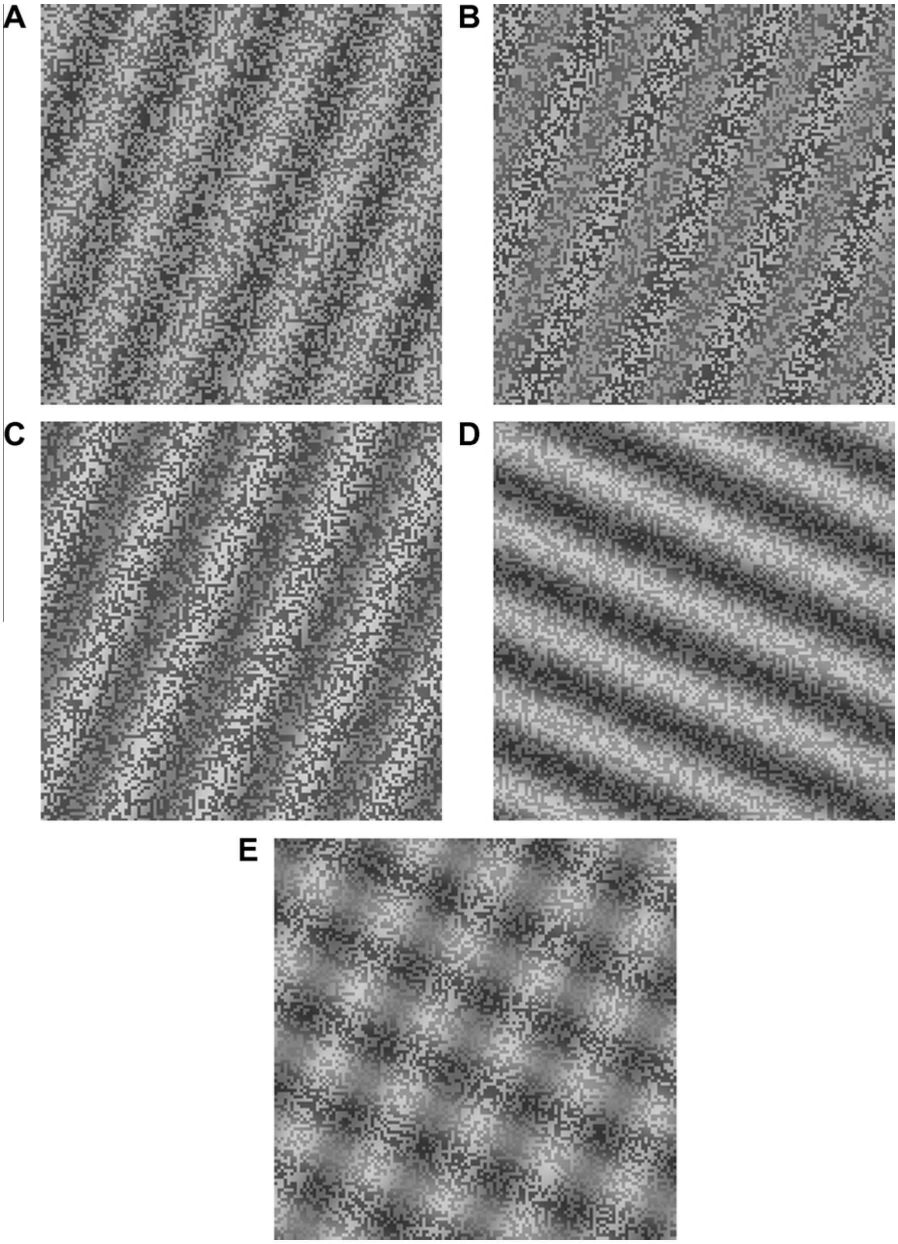

Fig. 1. Stimulus examples (A) LM-only: A 45 deg oriented sine wave luminance grating added to binary noise (B) AM-only: Amplitude modulated binary noise pattern (modulation depth = 0.40). (C) An in-phase composite grating where peaks of LM (highest luminance) and AM (highest amplitude) are superimposed. (D) An anti-phase grating where LM troughs are superimposed with AM peaks. (E) A plaid consisting of an in-phase grating on the right diagonal (LM + AM) and an anti-phase grating on the left diagonal (LM − AM).

images (Johnson & Baker, 2004) but the sign of this correlation varies (Schofield, 2000), suggesting that contrast/amplitude modulations are informative by virtue of their relationship with luminance variations.

Building on the physiological work of Zhou and Baker (1996), Schofield et al. (2010) developed the shading channel model in order to explain the role of AM in layer decomposition. In this model LM and AM are initially detected separately and then recombined in an orientation/frequency specific additive sum that broadly mimics Zhou and Baker's (1996) envelope neurons. In-phase parings sum to produce an enhanced output (greater perceived depth); whereas anti-phase parings subtract, weakening the output/depth percept. However AM components are given a relatively low weighting at the summation stage such that their effect on single LM components is marginal. A competitive gain control mechanism working across orientations produces the dramatic scission found for plaid stimuli. The model has been used to describe a range of psychophysical results (Schofield et al., 2010; Sun & Schofield, 2011), has been applied directly to natural images (Schofield et al., 2010), and has been used as the basis for a machine vision system for layer decomposition (Jiang, Schofield, & Wyatt, 2010).

The shading channel model relies on relatively low-level mechanisms, which might be considered comparable to envelope neurons found in area 17/18 of cat visual cortex (Zhou & Baker, 1996). Therefore we would expect layer decomposition based on LM and AM mixtures to be automatic and fast acting. Indeed, there are many examples of fast processing of textured stimuli for similarly complex tasks such as: estimating shape from texture (Gurnsey et al., 2006); detecting and discriminating second-order