Vision Research 117 (2015) 34–40

Contents lists available at ScienceDirect

Vision Research

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

Processing context: Asymmetric interference of visual form and texture in object and scene interactions

Matthew X. Lowe [a,b,*], Susanne Ferber [b,c], Jonathan S. Cant [a]

[a] Department of Psychology at Scarborough, University of Toronto, Toronto, ON M1C 1A4, Canada
[b] Department of Psychology at St. George, University of Toronto, Toronto, ON M5S 3G3, Canada
[c] Rotman Research Institute, Baycrest, Toronto, ON M6A 2E1, Canada

ARTICLE INFO

Article history:
Received 26 February 2015
Received in revised form 30 September 2015
Accepted 5 October 2015
Available online 7 November 2015

Keywords:
Object–scene perception
Context
Interference
Texture perception
Form perception
Global image features

ABSTRACT

Substantive evidence has demonstrated that scene-centered global image features influence the processing of objects embedded in complex visual scenes. Conversely, a growing body of work suggests that relevant object information may inherently influence diagnostic global scene statistics used in rapid scene categorization. Here, we investigate the potential effects of interference in object–scene perception when attending to form and texture in both simple figure-ground representations and more complex object–background scenes. Results reveal asymmetric interference in the perception of form and texture in object and scene processing: Inconsistent scene texture interfered with the classification of object texture, and inconsistent object form interfered with the classification of scene form, but not vice versa. These findings contribute to our understanding of the interactions between an object and its environment, and further inform our knowledge of the visual features which influence interactivity in object and scene perception.

1. Introduction

A remarkable aspect of the human visual system is the ability to draw on a broad range of cues to rapidly and efficiently identify and categorize objects embedded in a complex visual scene. In general, research has found that knowledge about which objects and scenes tend to co-occur facilitates the efficiency of both the search for and recognition of objects (Biederman, Mezzanotte, & Rabinowitz, 1982; Boyce & Pollatsek, 1992; Davenport, 2007; Davenport & Potter, 2004; De Graef, Christiaens, & d'Ydewalle, 1990; Henderson, Weeks, & Hollingworth, 1999; Joubert, Fize, Rousselet, & Fabre-Thorpe, 2008; Joubert, Rousselet, Fize, & Fabre-Thorpe, 2007; Palmer, 1975 for a review, see Oliva and Torralba (2007)). Conversely, a growing body of work has demonstrated evidence for the influence of object information on scene classification through a consistent-object advantage (Davenport, 2007; Davenport & Potter, 2004; Joubert et al., 2007), even without the need to activate semantic information from stored object representations (Mack & Palmeri, 2010). Such research suggests a dual-system, interactive account between scene and object processing. Nevertheless, we currently know very little about the visual features contributing to such an interactive system. Here, we investigate the extent to which common and relatively lower-level visual features (form and texture) influence the interactivity between object and scene processing through visual interference between object and background features.

Scene perception may be governed by general mechanisms that apply broadly across visual processing. For example, seminal work on global processing has suggested that the precedence of global image features is a prevailing property of visual perception, wherein global structure precedes the perception of local elements or fine-grained analyses (Navon, 1977). Navon presented compound letters representing larger figures (global configurations), which were spatially constructed from a suitable arrangement of smaller figures (local elements), and observed an advantage in the processing of global configurations over local elements (i.e., faster responses to global configurations compared with local elements), which he termed the 'global precedence effect'. Critically, when global configurations and local elements were inconsistent, responses to the local elements were subject to interference from the global configurations, but local features did not interfere with global perception. This result was subsequently referred to as the 'global interference effect'. In other words, involuntary attention

* Corresponding author at: Department of Psychology, University of Toronto Scarborough, 1265 Military Trail, Toronto, ON M1C 1A4, Canada.
E-mail address: matthew.lowe@mail.utoronto.ca (M.X. Lowe).

http://dx.doi.org/10.1016/j.visres.2015.10.010
0042-6989/

to the global level was observed when attention was directed to the local level, resulting in global inference in the perception of local elements.

Subsequent research on rapid scene identification has provided support for the primacy of global features over local region and object information, demonstrating that contextual information influencing object–scene interactivity is guided by global image features which direct attention early in the visual processing stream (Torralba, Oliva, Castelhano, & Henderson, 2006). That is not to say, however, that scene and object information is processed in a strictly scene-to-object hierarchical fashion. On the contrary, evidence has demonstrated an advantage for the classification of scenes that contain semantically consistent (compared with inconsistent) objects, suggesting that objects and scenes may be processed interactively and in parallel (Davenport, 2007; Davenport & Potter, 2004; Joubert et al., 2007). In fact, recent research has shown that rapid recognition of a scene's superordinate category of natural or man-made is modulated by the presence of a consistent or inconsistent object, even in the absence of explicit object recognition (Mack & Palmeri, 2010). Mack and Palmeri thus hypothesized that object–scene interference may be more simply explained by relatively low-level visual differences between objects and scenes, as opposed to relatively high-level semantic relationships between them.

Yet which visual features are utilized in such an interactive process? Previous work has shown that coarse, low-level global scene properties determine ultra-rapid scene recognition and categorization (Schyns & Oliva, 1994), and has drawn attention to the roles of form and texture in capturing the diagnostic structure necessary to perform these processes (for a review, see Oliva and Torralba (2006)). Indeed, investigations of scene processing using functional magnetic resonance imaging (fMRI) have demonstrated that the scene-selective parahippocampal place area (PPA), a region shown to respond selectively to scenes over individual objects or faces (Epstein & Kanwisher, 1998), represents scenes by processing global spatial layout (Epstein, Graham, & Downing, 2003). Similarly, recent evidence has revealed that PPA is also sensitive to processing visual cues such as material properties signaled by surface texture (Cant & Goodale, 2007, 2011), suggesting that diagnostic global statistics informing scene identity may incorporate both spatial structure and material properties.

In a similar vein, highly diagnostic visual cues such as surface reflectance properties, surface texture, and surface structure can cue stored knowledge of object material properties such as mass, compliance, and friction (Adelson, 2001; Buckingham, Cant, & Goodale, 2009; Motoyoshi, Nishida, Sharan, & Adelson, 2007). These cues not only aid in visual search and recognition, but also contribute to action planning (Gallivan, Cant, Goodale, & Flanagan, 2014), ultimately affecting how we physically engage with objects of various tactile qualities (e.g., rough vs. smooth), and the adjustment of our gait when moving through an environment containing different surface attributes (e.g., ice vs. grass). Evidence has demonstrated both independence and asymmetric interference in the perception of texture and form in object perception (Cant, Arnott, & Goodale, 2009; Cant, Large, McCall, & Goodale, 2008). In fact, visual texture may be especially important in defining edge and contour information used for finding partially occluded objects in complex and crowded environments (Biederman, 1987). While it has been argued that objects and scenes interact extensively, the influence of visual texture in such an interaction has yet to be explored, despite the importance of texture as a cue in both object and scene processing.

In the present study, we examine the extent to which form and texture consistency influence object–scene interactivity. We focused on global interference effects rather than global precedence effects, since the former capture interactions in visual processing across global and local levels, while the latter simply demonstrates that participants typically process global features faster than local features. In Experiment 1 we aim to initially replicate and extend a global interference effect of form (Navon, 1977) using modified Navon stimuli in simple figure-ground displays, and then investigate this effect for visual texture, predicting a similar interference effect in the perception of texture (i.e., slower judgments of local texture when global and local texture features are inconsistent). Thus, our motivation for Experiment 1 is to validate our stimuli and experimental paradigm by replicating well-established results of global interference in form perception and also to demonstrate novel results of global interference in texture perception. Having done so, we can then extend these findings to the study of more complex object–scene interactions, which we explore in Experiment 2. If form and texture are indeed important visual cues in scene and object perception, and scene perception proceeds from global properties to local elements, we expect to observe global scene interference in the perception of local object properties (i.e., a global interference effect of form and texture). However, as recent evidence has demonstrated that the perception of global scene statistics is modulated by inconsistent object information (Mack & Palmeri, 2010), we will also investigate the potential influence that both object form and object texture have on scene perception. Across both experiments, we elected to focus on speed of processing (reaction time) as a measure of interference, using accuracy only to ensure a constant level of attention across experimental tasks.

2. Experiment 1

Before investigating the interaction between object and scene information in the perception of form and texture, we first aimed to confirm that these visual features are processed in a global-to-local manner. Using modified classic Navon figures (1977), we incorporated both form and texture into simple figure-ground representations (see Fig. 1), predicting a replication of Navon's global interference effect for form (slower judgements of local form when local and global form were inconsistent, but not vice versa), and similarly expected a global interference effect to be found in the perception of texture.

2.1. Participants

Twelve participants (all female) between 20 and 32 years of age ($M = 21.50$) were recruited from the University of Toronto undergraduate community and received course credit for their participation. All participants had normal or corrected-to-normal visual acuity, were right-handed, and gave informed consent in accordance with the University of Toronto Ethics Review Board in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki).

2.2. Stimuli and apparatus

Sixty-four stimuli were generated using Adobe Photoshop CS3 software (Adobe Systems, San Jose, CA) and were presented electronically using E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA) on a ViewSonic 21-in. CRT monitor (1280×1024 resolution; 85-Hz refresh rate). Stimuli subtended $18.4° \times 18.4°$ of visual angle and were presented centrally against a white background following a black central fixation cross (subtending $1° \times 1°$) at a viewing distance of 52 cm. The stimuli were constructed so that visual features (form: heart versus star; texture: paint versus rock) could vary at both global and local levels of attention, and importantly, variations in each feature were manipulated across