Vision Research 49 (2009) 3001–3012

Contents lists available at ScienceDirect

Vision Research

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

How do amplitude spectra influence rapid animal detection?

Carl M. Gaspar *, Guillaume A. Rousselet

Centre for Cognitive Neuroimaging (CCNi), Department of Psychology, University of Glasgow, G12 8QB Glasgow, UK

ARTICLE INFO

Article history:
Received 16 April 2009
Received in revised form 23 September 2009

Keywords:
Natural scenes
Amplitude spectra
Phase spectra
Object detection
Categorization
Scene gist
Image statistics

ABSTRACT

Amplitude spectra might provide information for natural scene classification. Amplitude does play a role in animal detection because accuracy suffers when amplitude is normalized. However, this effect could be due to an interaction between phase and amplitude, rather than to a loss of amplitude-only information. We used an amplitude-swapping paradigm to establish that animal detection is partly based on an interaction between phase and amplitude. A difference in false alarms for two subsets of our distractor stimuli suggests that the classification of scene environment (man-made versus natural) may also be based on an interaction between phase and amplitude. Examples of interaction between amplitude and phase are discussed.

1. Introduction

Animal detection in natural scenes is rapid (Kirchner & Thorpe, 2006; Thorpe, Fize, & Marlot, 1996), orientation invariant (Guyonneau, Kirchner, & Thorpe, 2006), and can be supported by mainly feedforward mechanisms (VanRullen & Koch, 2003). Animal detection can also be performed efficiently in conditions that demand a high degree of parallel processing (Li, VanRullen, Koch, & Perona, 2002; Rousselet, Fabre-Thorpe, & Thorpe, 2002; Rousselet, Thorpe, & Fabre-Thorpe, 2004). More generally, the human visual system appears to be well suited to encode the information that is required for animal detection. However it is not yet clear what that information is. Most animal detection tasks, including the one used in this study, employ large sets of highly variable natural scenes, and a wide variety of animal species that vary in location and size. The complicated nature of animal detection suggests that there may be different sources of information available.

Many of the studies that have attempted to characterize the information underlying natural scene perception have made use of Fourier analysis to help distinguish between two different sources of information: amplitude and phase spectra. While the amplitude spectrum has a noticeable effect on image appearance, the phase spectrum determines most of the recognizable image structure in natural scenes (Oppenheim & Lim, 1981; Piotrowski & Campbell, 1982; Rousselet, Pernet, Bennett, & Sekuler, 2008). Many models of natural scene classification focus on information provided by the phase spectrum (Loschky & Larson, 2008; Loschky et al., 2007; Serre, Oliva, & Poggio, 2007; Thorpe, Delorme, & Van Rullen, 2001; Wichmann, Braun, & Gegenfurtner, 2006). However, the high speed of animal detection suggests that this task may also benefit from the processing of abstract image structures that are not immediately recognizable, nor obviously related to the semantic content of a scene (Honey, Kirchner, & VanRullen, 2008; Johnson & Olshausen, 2003; Oliva, 2005). Some theories have suggested that the amplitude spectrum may provide this type of abstract information, especially during rapid stimulus presentations (Guyader, Chauvin, Peyrin, Herault, & Marendaz, 2004; Honey et al., 2008; Johnson & Olshausen, 2003; Joubert, Rousselet, Fize, & Fabre-Thorpe, 2009; Kaping, Tzvetanov, & Treue, 2007; Torralba & Oliva, 2003). Consistent with this idea, rapid recognition supports a hierarchy of scene information, from broad categories very early on (animate versus inanimate objects), to specific object identities later on (a Labrador versus other breeds of dogs) (Fei-Fei, Iyer, Koch, & Perona, 2007). Early visual cortices may provide information about broad scene categories by signalling specific patterns in the amplitude spectrum (Guyader et al., 2004; Joubert et al., 2009; Kaping et al., 2007; Torralba & Oliva, 2003).

Animal detection requires us to distinguish between two very broad classes of natural scenes: scenes that contain animals and scenes that do not. It is difficult to imagine a single cue that could distinguish these two classes of scenes. If a single cue were useful on its own then it would have to be both present in the majority of animal scenes and absent in the majority of non-animal scenes. And yet simple image statistics applied to our own stimuli (shown in Fig. 1) suggest that the amplitude spectra of non-animal scenes

* Corresponding author.
E-mail addresses: carl@psy.gla.ac.uk (C.M. Gaspar), g.rousselet@psy.gla.ac.uk (G.A. Rousselet).

0042-6989/$ - see front matter
doi:10.1016/j.visres.2009.09.021

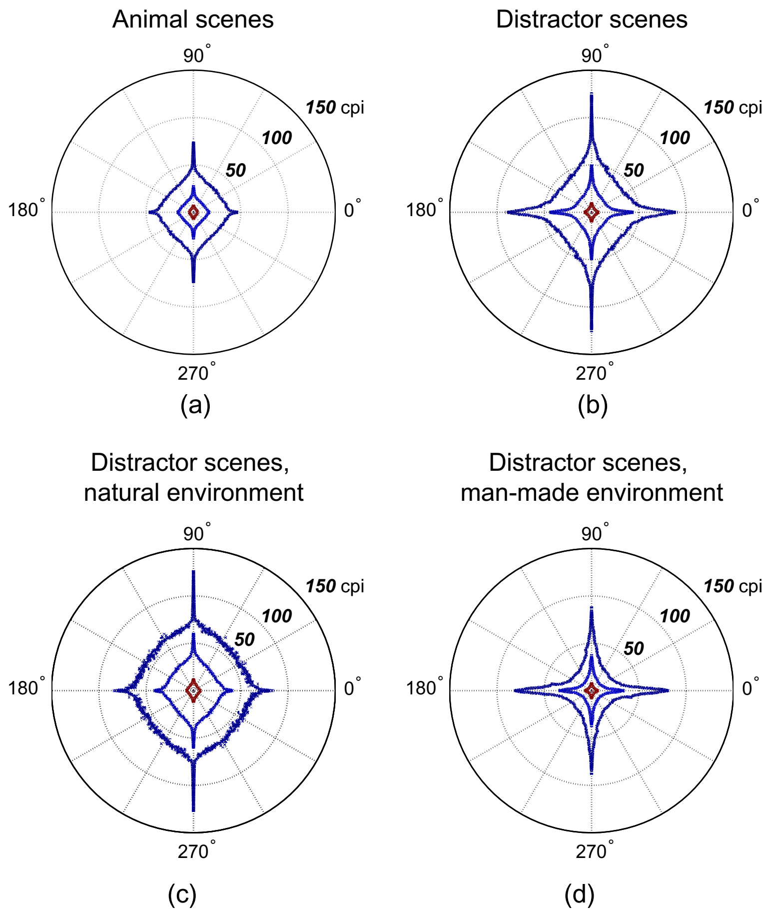

Fig. 1. Spectral-energy contours of four different image categories: (a) scenes of natural environments containing animals (600 images); (b) scenes without animals (600 images); (c) scenes of natural environments without animals (300 images); and (d) scenes of man-made environments without animals (300 images). Spectral-energy contours are obtained by first averaging the power spectra (e.g., squared amplitude) of images in each category. The contour plots represent 60%, 80% and 90% of energy (e.g., summed power) in the spectrum. The contour is selected so that the sum of the power inside the contour represents 60% (innermost contour), 80% and 90% (outermost contour) of the total power. Radius is plotted in units of cycles per image width (cpi), and angle is in degrees.

possess at least two properties that are significantly attenuated in the amplitude spectra of animal scenes: concentration of amplitude at cardinal orientations (especially scenes of natural environments), and a more even distribution of amplitude between high and low frequencies (especially for scenes of man-made environments). Both Drewes, Wichmann, and Gegenfurtner (2005) and Torralba and Oliva (2003) have already made this observation, and have devised simple computational models that can achieve high levels of animal detection accuracy by using only the amplitude spectrum. Whether or not human observers can use the amplitude spectrum in the same manner as these models is, however, an entirely separate matter.

If human observers can perform animal detection by extracting information from the amplitude spectrum, then one can simply measure detection accuracy in a condition where phase information has been completely randomized. In such conditions, the only available cue would be the amplitude spectrum. According to computational models, classification by human observers can still be up to 75% correct. However, Wichmann et al. (2006) found that animal detection by human observers is just barely above chance level when phase is completely randomized. This is strong evidence that humans cannot extract information from the amplitude spectrum with any significant efficiency. One might also be tempted to conclude that variations in the amplitude spectrum are completely irrelevant for animal detection. However, recent results obtained by Drewes (2006) shows that this is not the case. Drewes (2006) assessed the contribution of the amplitude spectrum to animal detection by measuring accuracy in a condition where all amplitude spectra are replaced by the mean amplitude spectrum across all images, both animal and non-animal scenes. In this amplitude-normalized condition, variations in the global amplitude spectra are completely obliterated. If the amplitude spectrum matters at all, then accuracy in the amplitude-normalized condition should be significantly lower than for scenes with their original amplitude spectra. Indeed, this is exactly what Drewes (2006) found (see Joubert et al. (2009) for a similar observation using a different task). This result raises an interesting issue: if variations in amplitude spectra do not provide human observers with any substantial cues for animal detection, why should the removal of such variations have any effect on accuracy?

One way to reconcile the results of Drewes (2006) with the results of Wichmann et al. (2006) is to suppose that successful animal detection is at least partially dependent on some interaction between phase and amplitude spectra. In other words, the amplitude spectrum may not be a significant source of information on its own; nonetheless, natural-scene phase processing might be enhanced by having the correct amplitude spectrum. This may seem like a surprising proposition because the natural scene literature almost always treats amplitude and phase spectra as though they were independent sources of information. However, the dichotomy between the amplitude and phase spectra is a mathematical distinction that is used to simplify the analysis of information content in images. Human observers may use amplitude and phase content independently, or they may be using information that is