Vision Research 108 (2015) 1–7

Contents lists available at ScienceDirect

Vision Research

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

Visual discomfort and the spatial distribution of Fourier energy

Olivier Penacchio [a,b,*], Arnold J. Wilkins [c]

[a] Computer Vision Centre, Universitat Autònoma de Barcelona, Spain
[b] School of Psychology and Neuroscience, University of St Andrews, United Kingdom
[c] Department of Psychology, University of Essex, United Kingdom

ARTICLE INFO

Article history:
Received 6 October 2014
Received in revised form 15 December 2014
Available online 6 January 2015

Keywords:
Visual discomfort
Fourier spectrum
Natural images
Scale invariance
Design

ABSTRACT

Quite independently of what they represent, some images provoke discomfort, and even headaches and seizures in susceptible individuals. The visual system has adapted to efficiently process the images it typically experiences, and in nature these images are usually scale-invariant. In this work, we sought to characterize the images responsible for discomfort in terms of their adherence to low-level statistical properties typically seen in natural scenes. It has been conventional to measure scale invariance in terms of the one-dimensional Fourier amplitude spectrum, by averaging amplitude over orientations in the Fourier domain. However, this loses information on the evenness with which information at various orientations is represented. We therefore fitted a two-dimensional surface (regular circular cone $1/f$ in logarithmic coordinates) to the two-dimensional amplitude spectrum. The extent to which the cone fitted the spectrum explained an average of 18% of the variance in judgments of discomfort from images including rural and urban scenes, works of non-representational art, images of buildings and animals, and images generated from randomly disposed discs of varying contrast and size. Weighting the spectrum prior to fitting the surface to allow for the spatial frequency tuning of contrast sensitivity explained an average of 27% of the variance. Adjusting the shape of the cone to take account of the generally greater energy in horizontal and vertical orientations improved the fit, but only slightly. Taken together, our findings show that a simple measure based on first principles of efficient coding and human visual sensitivity explained more variance than previously published algorithms. The algorithm has a low computational cost and we show that it can identify the images involved in cases that have reached the media because of complaints. We offer the algorithm as a tool for designers rather than as a simulation of the biological processes involved.

1. Introduction

Viewing certain patterns, typically of stripes, can sometimes result in a variety of bodily symptoms and perceptual distortions, often collectively referred to as visual stress (Wilkins, 1995). The bodily symptoms can include discomfort, malaise and nausea, and the perceptual symptoms can involve perceptual instability and hallucinatory colours and shapes. The patterns responsible are usually those that in patients with photosensitive epilepsy are capable of inducing seizures (Wilkins et al., 1984). The patterns are quite unlike those that occur in nature, as explained below.

Images from the natural world are largely scale invariant (Field, 1987; Pentland, 1984). The relative contrast energy of the image remains approximately the same at all spatial scales. In consequence the images have a particular spatial structure that can be characterized in terms of the Fourier amplitude spectrum. The amplitude decreases with increasing spatial frequency approximately as $1/f$, so that when plotted on log–log coordinates the slope is close to -1 (Field, 1987). Field (1994) argued that the bandwidths of channels tuned to spatial frequency are optimized for this type of amplitude spectrum. Their bandwidth remains constant when expressed on an octave scale, so that a similar amount of information is then carried by each channel. Atick and Redlich (1992) have argued that the shape of the contrast sensitivity function enables images with a $1/f$ spectrum to be coded efficiently. The contrast sensitivity is low (the channel has low gain) for low spatial frequencies that have a high amplitude, and this conserves metabolic energy. The sensitivity increases as the amplitude decreases to a point where the signal-to-noise ratio is low, and then falls at high spatial frequencies when the signal can no longer be discriminated from noise reliably. Psychophysical experiments on discrimination tasks show that performance is optimal when stimuli have a $1/f$ spectrum (Geisler et al.,

* Corresponding author at: School of Psychology and Neuroscience, University of St Andrews, St Andrews, Scotland KY16 9JP, United Kingdom.
E-mail address: op5@st-andrews.ac.uk (O. Penacchio).

http://dx.doi.org/10.1016/j.visres.2014.12.013
0042-6989/

2001; Knill, Field, & Kersten, 1990; Parraga, Troscianko, & Tolhurst, 2000). Baker and Graf (2009) found that under conditions of binocular rivalry the dominant eye was the one that received an image with amplitude spectrum closer to 1/*f*.

There is a possible relationship between image structure and energy metabolism. The receptive fields of neurons in the primary visual cortex are such that images with 1/*f* structure produce a sparse cortical response (Atick & Redlich, 1992; Olshausen & Field, 1997; Simoncelli & Olshausen, 2001). The defining characteristic of this sparse response is that the distribution of neural firing is such that few neurons are active and many are quiescent, thereby reducing metabolic demand (Attwell & Laughlin, 2001). It is of obvious importance to minimise energy metabolism given that the brain uses more than 15% of the body's energy, even though less than 1% of its neurons are active at any one time (Lennie, 2003). The metabolism involved in image processing can be explored using functional magnetic resonance imaging. The fMRI BOLD signal reflects neuronal activation via oxygen and glucose metabolism. The signal is large for patterns that are uncomfortable (e.g., gratings with mid-range spatial frequency, i.e., around 3 cycles per degree – cpd) and is larger in individuals who are particularly susceptible to discomfort *viz* those with migraine (Huang et al., 2003). Metabolism is also reflected in the near infrared oxyhaemoglobin signal, and this is also larger for coloured patterns that are uncomfortable (Haigh et al., 2013).

Haigh et al. (2013) proposed that the discomfort is homeostatic and acts to prevent hypermetabolism. If this is the case, one might expect images with 1/*f* spectra to be more comfortable than those with spectral slopes that depart from 1/*f*, irrespective of their content. Juricevic et al. (2010) demonstrated that this was indeed the case with artificial images generated from filtered random noise or randomly disposed overlapping rectangles. They asked participants to rate the discomfort each image evoked and showed that images with amplitude spectra having a slope in log–log coordinates of -1 were generally more comfortable than those with slopes that were greater or less than -1. Fernandez and Wilkins (2008) used a wider variety of images (art and photographs as well as filtered noise) and showed that those images with an excess of energy at mid-range spatial frequencies (around 3 cpd) relative to that expected from 1/*f* were particularly uncomfortable. Their findings were replicated by O'Hare and Hibbard (2011) who corrected for perceived contrast. Given that the visual system is generally most sensitive to mid-range spatial frequencies, these findings would be consistent in suggesting that discomfort results from an excess of energy at those spatial frequencies to which the visual system is in general most sensitive. Taken together, simple principles of adherence to natural statistics derived from a statistical ecological approach and basic knowledge of the energetics involved in visual processing provide access to simple computational measures of discomfort, as we will demonstrate.

All the above studies have investigated the regularity in the variation of energy across spatial frequencies, and they have done so by averaging across orientation. Yet there are regularities in the variation of energy with orientation that this approach cannot capture, for example the energy in the vertical and horizontal orientations is usually greater than in the obliques (Girshick, Landy, & Simoncelli, 2011; Switkes, Mayer, & Sloan, 1978; Torralba & Oliva, 2003). In log–log coordinates the two dimensional Fourier amplitude spectrum of a natural image is approximated by a cone with slope of -1, with small departures from radial symmetry that express the slightly greater energy in the horizontal and vertical orientations. We wished to take into account this regularity to see whether we could obtain a better computational correlate of discomfort than in previous studies, our purpose being to offer a tool for use by designers. By considering the whole (two-dimensional) spectrum, we grasp both irregularities across frequencies at each orientation and across orientations at each frequency. Our rationale for considering Fourier content in all orientations instead of averaged over orientation is that some uncomfortable patterns have specific features that cannot be grasped by a one-dimensional description. In a checkerboard, the energy is distributed widely with respect to orientation, whereas in stripes, which are generally more uncomfortable than checks (Wilkins et al., 1984), the energy has a single orientation. On the other hand, figures of radiating lines, such as the MacKay figure (MacKay, 1957), and concentric rings (MacKay, 1957), Figure 2, are perceptually more unstable than linear gratings, suggesting that the two dimensional distribution of energy is critical.

Our purpose in this paper is not to emulate or model the neurological processes involved in discomfort but to obtain an economic computational metric that predicts discomfort for practical purposes. We appeal to the simple logic that the visual system is optimized for the processing of natural images, which have a characteristic structure that enables efficient coding as a result of evolution and adaptation. Images that do not possess this structure are therefore likely to be processed relatively inefficiently, which may have consequences for metabolism, with possible associated homeostatic mechanisms that give rise to visual discomfort. We show that a simple measure of the departure from 1/*f* in the two-dimensional Fourier space is a better predictor of discomfort than previous measures that reduce the spatial frequency content to a single dimension by averaging over orientation (Fernandez and Wilkins (2008); Juricevic et al. (2010)). We propose that such a simple measure may help architects and designers prevent those aspects of design that have in the past given rise to complaints. We offer five case histories as examples of the complaints and the images that have given rise to them, and show that our metric would readily have identified these images as problematic.

2. Method

2.1. Images

From a variety of studies, we obtained 7 sets of images, 765 images in total, our choice dictated simply by their availability. The discomfort each image evoked had been rated on a Likert scale (the higher the rating, the higher the discomfort evoked). Set 1 and Set 2 consisted respectively of 45 images of non-representational art from Study 1 by Fernandez and Wilkins (2008), and 70 images of non-representational art from their Study 2. Set 3 was obtained from the 160 rural and 160 urban images used by Pretty, Peacock, Sellens, and Griffin (2005). Set 4 and Set 5 were two sets of 75 photographs of building frontages (all from the same camera) with ratings of discomfort from two independent groups of 10 young participants (Murphy, unpublished). Set 6 were the images used by Cole and Wilkins (2013 Experiment 3): 100 images of animals, 50 of them poisonous. Set 7 were 80 images of randomly arranged discs of variable size and contrast, with their associated ratings of discomfort, from an undergraduate thesis by An Le. With the exception of the images from Fernandez and Wilkins (2008) and Pretty et al. (2005) , the images were presented on the LCD screens (approximately 30 cm $\times$ 20 cm of laptop computers at a distance of about 60 cm. The images were rated individually by student volunteers; a minimum of 10 volunteers rated each image and their ratings were averaged.

2.2. Computational measure

The images were cropped to retain the central square section and resized to 256 $\times$ 256 using the Matlab nearest neighbour algorithm. On an LCD display we measured the luminance as separate