Biological Psychology 88 (2011) 204–214

Contents lists available at ScienceDirect

Biological Psychology

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

$1/f^p$ Characteristics of the Fourier power spectrum affects ERP correlates of face learning and recognition

M. Blickhan[a], J.M. Kaufmann[b,d], J. Denzler[c], S.R. Schweinberger[b,d], C. Redies[a,*]

[a] Institute of Anatomy I, Jena University Hospital, School of Medicine, Friedrich Schiller University, Jena, Germany
[b] Department of General Psychology and Cognitive Neuroscience, Friedrich Schiller University, Jena, Germany
[c] Department of Computer Science, Friedrich Schiller University, Jena, Germany
[d] DFG Research Unit Person Perception, Friedrich Schiller University, Jena, Germany

ARTICLE INFO

Article history:
Received 30 November 2010
Accepted 8 August 2011
Available online 19 August 2011

Keywords:
Face learning
Face perception
Event-related potential
N170
N250
Spatial frequency
Fourier power spectrum
Natural vision

ABSTRACT

We investigated the influence of Fourier power spectrum ($1/f^p$) characteristics on face learning while recording ERPs that are associated with the representation of faces. Two image sets with an altered $1/f^p$ characteristics were created. The first set consisted of stimuli with a STEEP SLOPE ($1/f^{3.5}$) and therefore enhanced low spatial frequencies (LSF) and attenuated high spatial frequencies (HSF). The second set consisted of stimuli with a SHALLOW SLOPE ($1/f^2$), similar to complex natural scenes and artwork, resulting in enhanced HSF and attenuated LSF. Faces with a SHALLOW SLOPE elicited larger N170 and N250 amplitudes and larger old/new effects for central positivity in comparison to unmodified faces. The opposite effect was observed for faces with a STEEP SLOPE that led to slower reaction times. This result suggests that diminishing the ratio of fine detail (HSF) to coarse structures (LSF) impairs face learning, whereas increasing it facilitates neurocognitive correlates of face learning.

1. Introduction

The human visual system uses spatial filtering to extract visual information (De Valois and De Valois, 1980; Westheimer, 2001). For example, visual information in the brain is processed through two separate pathways (Livingstone and Hubel, 1988): the magnocellular stream provides coarse information (low spatial frequencies) and the parvocellular stream provides fine information (high spatial frequencies). Several studies have demonstrated that the spatial frequency profile of face images has an effect on face learning and identity recognition. Typically, bandwidth frequency-filtered faces were used to study these effects (for a recent review, see Ruiz-Soler and Beltran, 2006).

For face recognition, the importance of the frequency range between 8 and 16 cycles per image (cpi) was implicated in behavioural studies (Collin et al., 2004; Costen et al., 1994; Fiorentini et al., 1983; Morrison and Schyns, 2001; Näsänen, 1999). A similar range is preferred for face identity discrimination (Costen et al., 1996; Parker and Costen, 1999), although other frequency ranges have also been suggested to contribute to face perception (Ginsburg, 1978; Halit et al., 2006; Harmon and Julesz, 1973).

Neurophysiological studies have also explored the processing of faces that were bandwidth frequency filtered. Results are inconsistent and favor different frequency ranges. In an fMRI study, Vuilleumier et al. (2003) measured the activation of fusiform cortex, which is thought to play an important role in face recognition (Kanwisher et al., 1997). They found a higher activation in fusiform cortex for unfiltered (broad spatial frequency, BSF) faces and high spatial frequency (HSF) faces in comparison to low spatial frequency (LSF) faces. Furthermore, they found long-lag repetition effects for HSF faces and BSF faces but not for LSF faces. They concluded that LSF faces lead to less stable face representation in fusiform cortex. In another fMRI study, Eger et al. (2004) found immediate repetition effects that where invariant to spatial frequency composition in fusiform cortex. By contrast, response adaption to repetition in occipital face regions only occurred with identical stimulus input with the same spatial frequency composition. In an event-related potential (ERP) study, Goffaux et al. (2003) found higher accuracy for LSF than for HSF faces in a familiarity task, in which bandwidth frequency-filtered faces had to be recognized. No significant differences between LSF and HSF faces were found for reaction times and amplitudes of the ERP component N170, which is widely assumed to reflect early face processing prior to face identity recognition (Eimer, 2000a,b). However there is some

* Corresponding author at: Institute of Anatomy I, University of Jena School of Medicine, Teichgraben 7, D-07743 Jena, Germany. Tel.: +49 3641 938 511; fax: +49 3641 938 512.
E-mail address: redies@mti.uni-jena.de (C. Redies).

0301-0511/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.biopsycho.2011.08.003

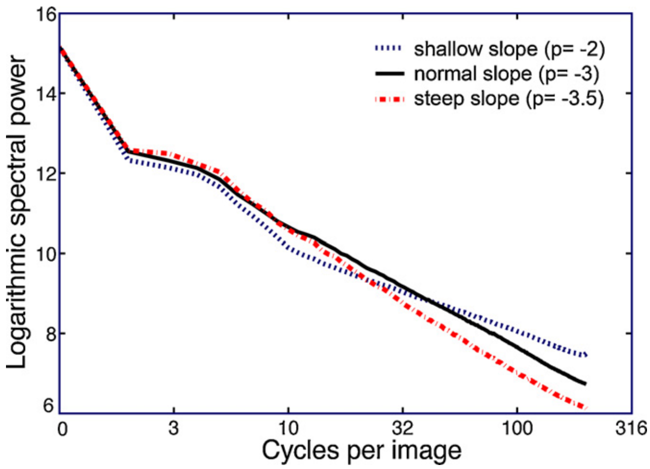

Fig. 1. Averaged Fourier spectral power of the 30 stimuli from each category. In the log-log plane, Fourier power (amplitude squared) was plotted as a function of spectral frequency (cpi, cycles per image). SHALLOW SLOPE – enhanced high spatial frequencies and diminished low spatial frequencies; NORMAL SLOPE – not manipulated; STEEP SLOPE – enhanced low spatial frequencies and diminished high spatial frequencies. *p*, slope of the curve. Spectral power was larger for the STEEP SLOPE than NORMAL SLOPE faces from 0 to 9 cpi and larger for STEEP SLOPE than SHALLOW SLOPE faces from 0 to 22 cpi. NORMAL SLOPE faces had larger power than SHALLOW SLOPE from 0 to 43 cpi and SHALLOW SLOPE faces had larger power above 43 cpi. A spatial frequency of 1 cpi is equivalent to 0.075 cycles per degree (cpd; 3 cpi = 0.23 cpd; 10 cpi = 0.75 cpd; 32 cpi = 2.41 cpd; 100 cpi = 7.52 cpd).

evidence that the N170 can be modulated by repeated face presentation (Caharel et al., 2009). Other ERP studies that focused on face detection or passive viewing tasks and not on identification showed larger N170 amplitudes for LSF faces than HSF faces (Halit et al., 2006), for HSF faces than LSF faces (Hsiao et al., 2005; Nakashima et al., 2008), or equal N170 amplitudes for HSF faces and LSF faces (Flevaris et al., 2008).

This inconsistency may be due to different tasks used in these studies (familiarity, gender and passive viewing task), or to different types of stimulus manipulation (contrast, luminance and different frequency cut offs). It has been shown that the face processing system uses frequency scales according to their diagnostic value for the respective task (Loftus and Harley, 2004; Morrison and Schyns, 2001; Schyns and Oliva, 1999).

Various ERP studies showed that bandwidth frequency-filtered faces did not elicit larger N170 amplitudes than unfiltered faces (Flevaris et al., 2008; Goffaux et al., 2003; Halit et al., 2006). Interestingly Nakashima et al. (2008) found a larger N170 for HSF faces than for BSF (unfiltered) faces and concluded that the N170 represents fine feature encoding of faces. One possible explanation for this discrepancy is that, by equalizing the contrast in their stimulus set, Nakashima et al. (2008) increased the spectral power of the HSF faces. Typically, a face has more spectral power in the lower frequencies than in the higher frequencies (Keil, 2008; Redies et al., 2007b). As a consequence, an HSF face has less overall spectral power than an LSF face. Possibly, the larger spectral power of the HSF faces, which was increased by contrast equalization, led to the larger N170 in comparison to BSF faces.

The above problem suggests that it may be advantageous to modulate the spatial frequency profile of faces content without "traditional" bandwidth filters, in order to avoid the possible confounding effect of differences in overall spectral power of the images. The radially averaged Fourier spectral power of most natural images, including face images, falls nearly linearly with increasing spatial frequency (f), according to a power law ($1/f^p$; Keil, 2008; Redies et al., 2007b). In log–log plots of Fourier power versus spatial frequency (Fig. 1), the value *p* corresponds to the slope of the plotted line and provides a measure for the ratio of high to low spatial frequencies in the image. The different stimuli, which our visual system perceives under normal everyday conditions, usually vary in their spectral slope and are not bandwidth frequency filtered. By manipulating the slope value, we changed the ratio of high and low spatial frequency power without affecting the overall spectral power of the face images. These images still contain the full frequency spectrum and may result in more realistic stimuli compared to bandwidth frequency-filtered images (Fig. 2).

The image manipulation carried out by us has further perceptual implications. It is well established that, in contrast to face images, the slope value of complex natural scenes is around -2 ($1/f^2$ characteristics), or -1 if Fourier amplitude is plotted instead of Fourier power (Burton and Moorhead, 1987; Tolhurst et al., 1992). This result implies that complex natural scenes are scale invariant (fractal-like), that is, when zooming in and out of a natural scene, the statistical properties of the Fourier spectral components remain constant. Moreover, it has been shown that the human visual system processes the statistical regularities of natural scenes with an efficient and sparse code (Hoyer and Hyvärinen, 2002; Olshausen and Field, 1996; Párraga et al., 2000; Vinje and Gallant, 2000, 2002). Scale invariance was also found in art images of different cultural provenance (Graham and Field, 2007; Redies et al., 2007a). Strikingly, it has been reported that artists portray human faces with the Fourier statistics of complex natural scenes, despite the fact that face photographs have significantly more negative slope values, indicative of less fine structure than natural scenes and art images (Keil, 2008; Redies et al., 2007b). A possible explanation is that artists intuitively endow their artworks with the statistical regularities that are processed optimally and efficiently in their visual system (Redies, 2007; Zeki, 1999).

Fig. 2. Example of the stimuli. NORMAL SLOPE – not manipulated; SHALLOW SLOPE – enhanced high spatial frequencies and diminished low spatial frequencies; STEEP SLOPE – enhanced low spatial frequencies and diminished high spatial frequencies.