Vision Research 48 (2008) 2817–2826

Contents lists available at ScienceDirect

Vision Research

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

Spatial frequency tuning of upright and inverted face identification

Carl Gaspar [a], Allison B. Sekuler [b,c], Patrick J. Bennett [b,c,*]

[a] Centre for Cognitive Neuroimaging, Department of Psychology, University of Glasgow, G12 8QB, UK
[b] Department of Psychology, Neuroscience, and Behaviour, McMaster University, 1280 Main Street West, Hamilton, Ont., Canada L8S 4K1
[c] Centre for Vision Research, York University, Toronto, Ont., Canada L8S 4K1

ARTICLE INFO

Article history:
Received 1 July 2008
Received in revised form 9 September 2008

Keywords:
Face perception
Inversion effect
Spatial frequency
Channels
Noise
Masking
Identification

ABSTRACT

Previous research suggests that observers use information near the eyes and eyebrows to identify both upright and inverted faces [Sekuler, A. B., Gaspar, C. M., Gold, J. M., & Bennett, P. J. (2004). Inversion leads to quantitative, not qualitative, changes in face processing. *Current Biology*, *14*(5), 391–396]. Here we ask whether more significant differences between upright and inverted face processing exist in the spatial frequency domain. Thresholds were measured in a 1-of-10 identification task with upright and inverted faces presented in no noise, white Gaussian noise, and in low-pass and high-pass filtered noises with various cutoff frequencies. In Experiment 1, all faces were presented in fronto-parallel view; in Experiment 2, viewpoint varied across trials. Thresholds were higher for inverted faces, but the magnitude of the inversion effect did not vary across conditions or experiments. Moreover, the shapes of the noise-masking functions obtained with low-pass and high-pass noise were the same for upright and inverted faces, did not vary between experiments, and revealed that identification was based on information carried by a 1.5 octave wide band of spatial frequencies centered on approximately 7 cycles per face width. Finally, individual differences in the magnitude of the inversion effect were not related to individual differences in the frequency selectivity of face identification. The results indicate that the face inversion effect for identification judgments is not due to subjects using different bands of spatial frequencies to identify upright and inverted faces.

1. Experiment 1

Aside from the written word, there is no class of objects that the adult human has greater experience identifying than the human face. Of necessity, we become expert face identifiers and can recognize thousands of faces at a single glance (Bahrick, Bahrick, & Wittlinger, 1975). However, if those same faces are turned upside-down, recognition becomes much more difficult (Yin, 1969; for a review, see Valentine, 1988). This result, known as the inversion effect, is interesting because an inverted face contains the same amount of information as an upright face. Thus, inverted faces must be more difficult to identify because the way we use that information varies as a function of orientation.

Previous results suggest that observers are less efficient at extracting the relevant information from inverted faces compared to upright faces (Gaspar, Bennett, & Sekuler, 2008; Sekuler, Gaspar, Gold, & Bennett, 2004), but the orientation-related changes in processing are subtle. For example, using the response classification technique, Sekuler et al. (2004) showed that observers used similar, localized regions of the face (i.e., near the eyes and eyebrows) to discriminate pairs of faces, regardless of face orientation. Hence, the difference in performance measured with upright and inverted faces reflected the fact observers used information around the eyes more optimally when the faces were upright, and not that different parts of the stimulus were used at different orientations. In this paper we examine whether larger effects of orientation exist in the spatial frequency domain than in the space domain.

Faces are broadband patterns, containing information at all spatial scales (Gold, Bennett, & Sekuler, 1999). Nonetheless, upright face identification appears to rely most heavily on a narrow band of spatial frequencies near 10 cycles per face (Boutet, Collin, & Faubert, 2003; Costen, Parker, & Craw, 1996; Gold et al., 1999; Nasanen, 1999; Peli, Lee, Trempe, & Buzney, 1994; Tanskanen, Nasanen, Montez, Paallysaho, & Hari, 2005). Is the perception of inverted faces mediated by a similar narrow band of spatial frequencies? Perceptual learning alters the spatial frequencies that are used to do some visual tasks (Busey, Schneider, & Wyatte, 2008; Dosher & Lu, 1998), so perhaps the development of expertise with upright faces (Diamond & Carey, 1986) is associated with a change in the spatial frequencies that are used to discriminate and identify them. Results reported by Goffaux and Rossion (2006) are consistent with this view. Based on studies that measured the Composite

* Corresponding author. Address: Department of Psychology, Neuroscience, and Behaviour, McMaster University, 1280 Main Street West, Hamilton, Ont., Canada L8S 4K1.

E-mail addresses: carl@psy.gla.ac.uk (C. Gaspar), sekuler@mcmaster.ca (A.B. Sekuler), bennett@mcmaster.ca (P.J. Bennett).

0042-6989/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2008.09.015

Face Effect (Young, Hellawell, & Hay, 1987) and the Whole-Part advantage effect (Tanaka & Farah, 1993), Goffaux and Rossion argued that low spatial frequencies are particularly important for perceiving the holistic and/or configural properties of faces, whereas high spatial frequencies are important for perceiving features. Some researchers have suggested that holistic and configural processing is especially important for the perception of upright faces (Maurer, Le Grand, & Mondloch, 2002), and therefore Goffaux and Rossion's findings suggest that low spatial frequencies may be more important for the identification of upright faces. Other evidence, however, is inconsistent with the view that different spatial frequencies are used to identify upright and inverted faces: Boutet et al. (2003), for example, found than manipulating the spatial frequency content of upright and inverted stimuli had the same effect on response accuracy in a face discrimination task. Also, varying the degree of spatial frequency overlap between two faces has similar effects on performance in a matching task when the stimuli are upright and inverted (Collin, Liu, Troje, McMullen, & Chaudhuri, 2004).

The present study examined the spatial frequency selectivity of upright and inverted face identification using critical band masking (Patterson, 1976). In two experiments, we measured identification thresholds for faces embedded in low-pass and high-pass noise that varied in terms of cutoff frequency. By measuring how threshold varied as a function of the spatial frequency content of the noise, it was possible to determine which spatial frequencies were used by observers to identify faces. Our results indicate that observers used the same narrow band of spatial frequencies to identify both upright and inverted faces.

1.1. Methods

1.1.1. Observers

Six observers (4 male, 2 female; average age = 20 years) participated in the experiment. Four were experienced psychophysical observers, but all were naïve about purpose of the experiment. All observers had an uncorrected or corrected binocular Snellen Acuity of 20/20 or better. The observers were not familiar with the models who were used to create the face stimuli.

1.1.2. Stimuli and apparatus

Stimuli were generated by a Power Mac G4 computer, and presented on a Sony Trinitron GDM-F520 monitor (frame rate = 75 Hz, non-interlaced) using MATLAB and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). Face stimuli were based on digitized photographs of 10 faces (5 male and 5 female) cropped to an oval window, excluding areas showing the chin and hair, including the hairline (see Gold et al., 1999, for details). From the viewing distance of 100 cm, the height and width of each face subtended 3.3 deg (96 pixels) and 2.3 deg (67 pixels), respectively. Faces were centered within a 4.4 $\times$ 4.4 deg square (128 $\times$ 128 pixels), and the amplitude spectrum of each individual face image was replaced with the average spectrum across all 10 images. Faces were presented to observers on a uniform background of average luminance (i.e., no noise) or embedded within filtered or unfiltered white Gaussian noise (see Fig. 1). Filtered noise was either low- or high-pass with cutoff frequencies of 1, 2.1, 4.2, 8.4, 16.8, and 33.5 cycles per face width (cpf). The contrast variance of the unfiltered noise was 0.08. The faces themselves were not filtered.

1.1.3. Procedure

The sequence of events on a given trial was as follows: a small, high-contrast (Weber contrast = ±0.5) fixation square (8.3 $\times$ 8.3 arc min) appeared at the center of the screen. To reduce adaptation, the contrast polarity of the square – black or white – varied randomly across trials. After 100 ms, the fixation square disappeared and a face was displayed for 500 ms. After the target face disappeared, observers were presented with a selection window that contained all ten faces presented without noise at an rms contrast of 0.32. Each face in the selection window was resized to 2.6 $\times$ 1.8 deg. The observer identified the target face by using the computer mouse to click on an item in the selection window. Auditory feedback, in the form of 600 and 200 Hz tones, was presented after correct and incorrect responses, respectively. After the response, the selection window disappeared, and the fixation square reappeared signaling the start of the next trial. Observers were informed that each face had an equal probability of being shown on each trial.

Observers participated in 3–6 one-hour sessions held on different days. Each session consisted of 12 or 14 blocks of trials, with a different combination of noise and stimulus orientation presented in each block. All sessions contained the following stimulus condi-

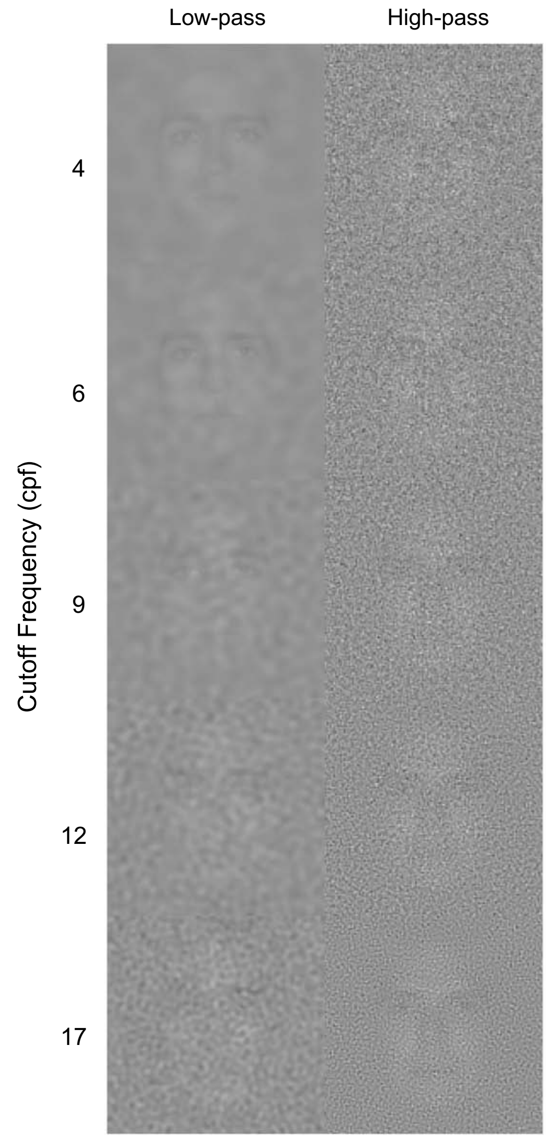

Fig. 1. An illustration of the filtered noise used in the experiments. Two faces are embedded in low- and high-pass filtered noise, with cutoff frequency varying from 4 to 17 cycles per face width (cpf). Face contrast variance is the same in every image. With low-pass noise, the face becomes difficult to identify when the cutoff frequency is above 9 cpf; with high-pass noise, the face is difficult to identify when the cutoff is below 9 cpf. In the experiments, cutoff frequency varied from 1 to 34 cpf. In addition, faces were presented in all-pass (i.e., white) noise and in no noise (i.e., a uniform background).

tions: no-noise, all-pass (white) noise, and low- and high-pass noise with cutoff frequencies of 1, 2.1, 4.2, 8.4, and 16.8 cpf. Some sessions included two additional blocks of all-pass noise and blocks of low-pass and high-pass noise with a cutoff frequency of 33.5 cpf. The order of conditions was randomized within a session.

In each block of trials, contrast variance of the face stimuli varied according to a 2-down, 1-up staircase procedure. Contrast variance was defined as $c^2 = \langle c^2(x,y)\rangle$, where $\langle\cdot\rangle$ indicates expected value taken over all pixel locations, $c(x,y) = (L(x,y) - L_{\text{background}})/L_{\text{background}}$, and $L(x,y)$ is the luminance at location (x,y). Each staircase was terminated after 50 trials. The contrast variance needed to produce 71% correct responses was estimated from a psychometric (Weibull) function that was fit to the results from each staircase. Final estimates of thresholds were calculated by averaging thresholds obtained in the last three sessions.

1.1.4. Noise-masking functions

Noise-masking functions were analyzed quantitatively by fitting different sigmoidal curves to thresholds obtained with low-pass and high-pass noise. Thresholds in low-pass noise conditions were analyzed with the function

$$t_{(L,C)} = k + a\int_{-\infty}^{C} g(\mu_L, \sigma_L)\,\mathrm{d}f \tag{1}$$

where $t_{(L,C)}$ is threshold, expressed in terms of contrast variance, measured with low-pass noise with a cutoff frequency of C, parameters k and a are constants, and $g(\mu_L, \sigma_L)$ is a log-Gaussian probability density with a mean of μ_L and a standard deviation of σ_L:

$$g(\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\log_{10}(f/\mu)^2/2\sigma^2} \tag{2}$$

Eq. (1) describes a sigmoidal function that has a lower asymptotic value of k and increases monotonically with C to an upper asymptotic value of $(k+a)$. When the cutoff spatial frequency of the noise, C, is plotted on a logarithmic axis, the sigmoidal function is centered on the spatial frequency μ_L and has a slope that is related inversely to σ_L. Thresholds in high-pass noise conditions were analyzed with the function

$$t_{(H,C)} = k + a\int_{C}^{\infty} g(\mu_H, \sigma_H) \tag{3}$$

where $t_{(H,C)}$ is threshold, expressed in terms of contrast variance, measured with high-pass noise with a cutoff frequency of C. Eq. (3) describes a sigmoidal function that has an upper asymptotic value of $(k+a)$ and *decreases* monotonically with increasing values of C to an lower asymptotic value of k.

Eq. (2) represents the relative sensitivity of face identification to the presence of noise at different spatial frequencies: greater masking will be produced when the noise contains spatial frequencies that are near the peak of Eq. (2). In the no-noise condition, the value of the integral in Eqs. (1) and (3) was assumed to be zero, and therefore threshold equals k. In the all-pass noise condition, the integral equals one and threshold therefore is $k+a$.

Eqs. (1) and (3) were fit to all thresholds from each subject by adjusting the parameters $(k, a, \mu_L, \sigma_L, \mu_H, \sigma_H)$ to minimize the sum

$$\sum_i \left[\log_{10}\left(\frac{t_i}{\hat{t}_i}\right)\right]^2$$

where t_i and $\hat{t}_i$ are, respectively, the observed and predicted thresholds in condition i.

1.2. Results

Statistical analyses were done with R (R Development Core Team, 2007). When appropriate, the Huynh–Feldt estimate of sphericity $(\tilde{\epsilon})$ was used to adjust p values of F tests conducted on within-subject variables (Maxwell & Delaney, 2004). The strength of association between the dependent and independent variables was expressed as partial omega-squared (ω_p^2), which was calculated using formulae described by Kirk (1995).

Face inversion effects were defined as $\log_{10}(c_i/c_u)$, where c_i and c_u are thresholds, expressed as rms contrast, obtained with inverted and upright faces. Inversion effects measured in the various conditions were very similar (Fig. 2). The mean inversion effect—averaged first across conditions and then subjects—was 0.241, 95%CI = (0.12, 0.37), which was significantly greater than zero, $t(5) = 4.9, p = .002$ (one-tailed), and did not differ from the mean value of 0.204 reported by Martelli, Majaj, and Pelli (2005) in a meta-analysis of 16 face recognition studies, $t(5) = 0.76, p = .48$. Schneider, DeLong, and Busey (2007) suggested that white external noise has a greater effect on the processing of inverted faces than the processing of upright faces. This hypothesis predicts that inversion effects measured with all-pass noise should be larger than inversion effects measured in the absence of noise, but a comparison of the effects measured in those conditions failed to find any difference, 95%CI = $(-0.15, 0.1)$, $t(5) = -0.54, p = .61$.

Inversion effects measured with low-pass and high-pass noise were analyzed in a 5 (cutoff frequency) $\times$ 2 (noise type) within-subject analysis of variance (ANOVA). Inversion effects measured with a cutoff frequency of 33.5 cpf were not included in this ANOVA because only two subjects were tested in that condition. The main effects of cutoff frequency, $F(4,20) = 0.82, \tilde{\epsilon} = 0.74, p = .50, \omega_p^2 = 0$, and noise type, $F(1,5) = 0.01, p = .92, \omega_p^2 = 0$, were not significant, nor was the cutoff $\times$ noise type interaction, $F(4,20) = 0.4$, $\tilde{\epsilon} = 0.63, p = .72, \omega_p^2 = 0$. Next, inversion effects measured in the no-noise and all-pass noise conditions were combined with the data from the low-pass and high-pass noise conditions (except for the 32 cpf condition) and analyzed in a one-way within-subjects ANOVA. The effect of condition was not significant, $F(11,55) = 0.51, \tilde{\epsilon} = 0.93, p = .88, \omega_p^2 = 0$. Hence, these analyses found no evidence that inversion effects varied across conditions.

Fig. 3 depicts average thresholds, expressed as contrast variance, plotted as a function of noise cutoff frequency, as well as the best-fitting noise-masking functions. The noise-masking func-

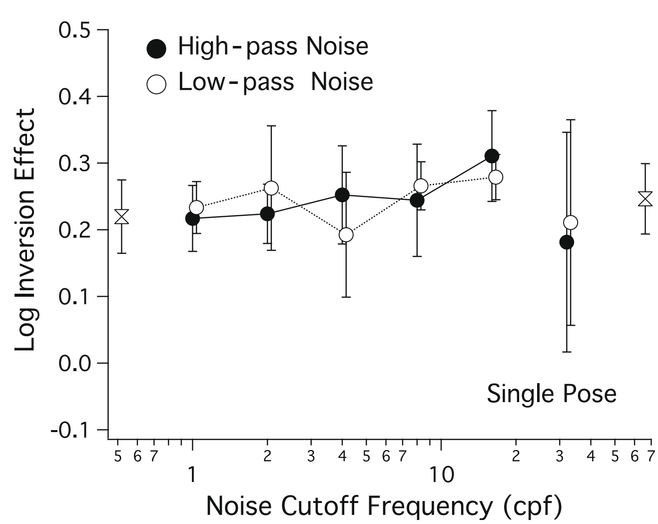

Fig. 2. Average inversion effects measured in Experiment 1 plotted as a function of the cutoff frequency of the noise. The inversion effect was defined as the logarithm of the ratio of thresholds, expressed as rms contrast, obtained with inverted and upright faces. The leftmost and rightmost symbols represent inversion effects measured in the no-noise and all-pass noise conditions. The points at 33.5 cpf represent the mean of the two subjects tested in that condition; the remaining points represent the mean of six subjects. Error bars represent ±1 SEM. Inversion effects did not differ significantly across conditions.

tions (top panels) fit the threshold data reasonably well in both the upright and inverted face conditions. Note that the masking functions in the inverted face conditions are shifted upwards relative to the functions obtained with upright faces, but otherwise have similar shapes, a result that is consistent with the finding that inversion effects were approximately constant across conditions (Fig. 2). The log-Gaussian functions (e.g., Eq. (2)) used to derive the masking functions are shown in the lower panels of Fig. 3. The log-Gaussian functions, which are the derivatives of the corresponding masking functions, indicate the relative sensitivity of masking to noise power at each spatial frequency. For inverted faces, the functions derived from the high-pass and low-pass noise conditions were virtually identical ($\mu = 9.34$ cpf; $\sigma = 0.20$), and indicate that thresholds increased significantly when frequency components within the range of 5–20 cpf were added to the masking noise. For upright faces, the functions derived from the high-pass and low-pass conditions differed slightly: although the functions had similar widths ($\sigma = 0.26$), peak sensitivity occurred at a slightly higher frequency in the low-pass noise condition ($\mu = 11.1$ cpf) than in the high-pass noise condition ($\mu = 6.8$ cpf).

Masking functions were fit to the data obtained from individual subjects in each condition, and the mean values of μ and σ are shown in Figs. 4 and 5, respectively. Inspection of Fig. 4 suggests that μ did not vary significantly with face orientation, but was slightly (i.e., ≈ 0.1 log units) greater in the low-pass noise conditions than in the high-pass noise conditions. However, a 2 (face orientation) $\times$ 2 (noise type) within-subjects ANOVA on the log-transformed values of μ failed to find any significant effects (noise: $F(1,5) = 2.56, p = .17, \omega_p^2 = 0.21$; orientation: $F(1,5) = 0.01$, $p = .92, \omega_p^2 = 0$; noise $\times$ orientation: $F(1,5) = 0.23, p = .65$, $\omega_p^2 = 0$). The value of σ, shown in Fig. 5, did not vary significantly with stimulus orientation but was approximately 0.1 log units greater in the low-pass noise condition than in the high-pass noise condition. However, as was the case with the μ parameter, a 2 (face orientation) $\times$ 2 (noise type) within-subjects ANOVA on the log-transformed values of σ failed to find any significant effects (noise: $F(1,5) = 1.94, p = .22, \omega_p^2 = 0.14$; orientation: $F(1,5) = 1.18$, p=.33, $\omega_p^2 = 0.03$; noise $\times$ orientation: $F(1,5) = 0.01, p = .91, \omega_p^2 = 0$). Hence, the statistical analyses suggest that the parameters governing the shapes of the masking functions did not vary significantly across conditions.

1.3. Discussion

The current findings agree with previous reports that observers rely on a narrow band of spatial frequencies to identify upright faces (e.g., Gold et al., 1999; Nasanen, 1999): identification thresholds for upright faces were influenced most strongly by masking components that were within 1.5 octaves of 8–10 cpf (Figs. 4 and 5). We hypothesized that face inversion effects might reflect a dif-

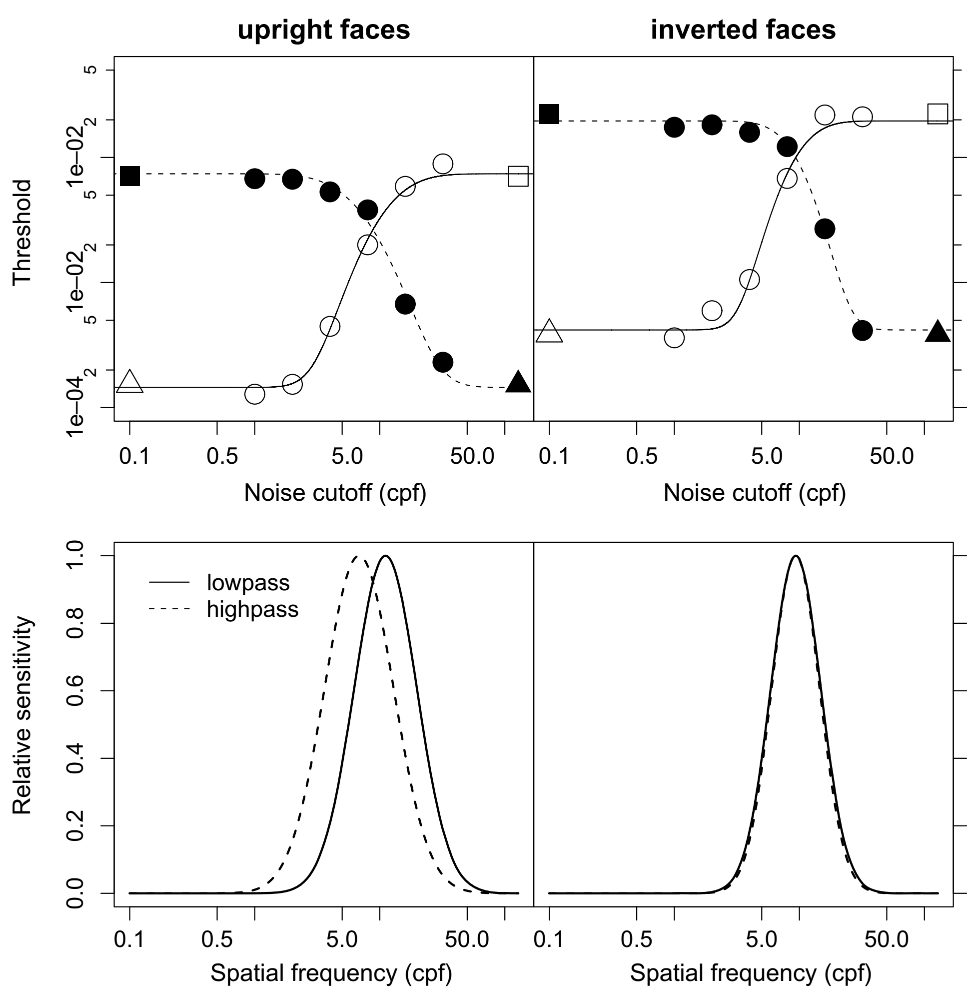

Fig. 3. Mean results obtained in Experiment 1 with upright (left column) and inverted faces (right column). The upper panels show thresholds, expressed as contrast variance, obtained in high-pass (filled circles) and low-pass (unfilled circles) noise conditions, as well as the no-noise (triangles) and all-pass noise (squares) conditions. The solid and dashed lines are the best-fitting noise-masking functions (Eqs. (1) and (3)). The lower panels show the log-Gaussian functions (Eq. (2)) that were used to create the masking functions. The Gaussian functions have been normalized to have a peak of one. The two functions in the lower-right panel nearly overlap.

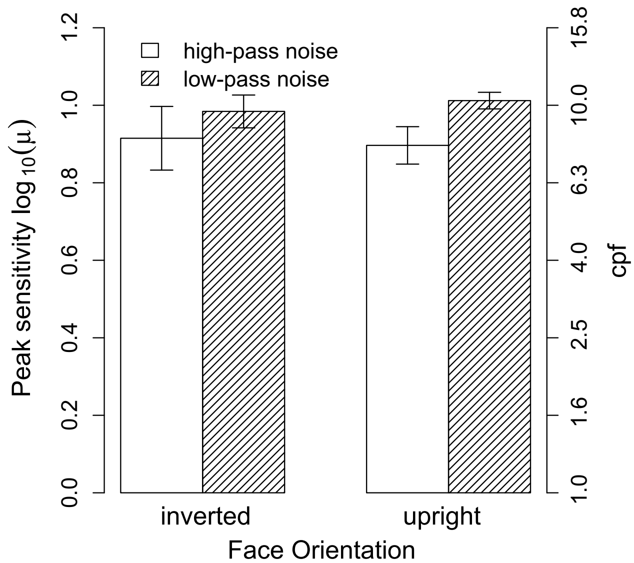

Fig. 4. The values of μ taken from the masking functions that were fit to data from individual subjects in Experiment 1. The bars show the mean of six subjects. The *y*-axis on the left shows log-transformed μ (see Eq. (2)). The axis on the right shows the same values transformed to units of cpf. Error bars represent ±1 SEM.

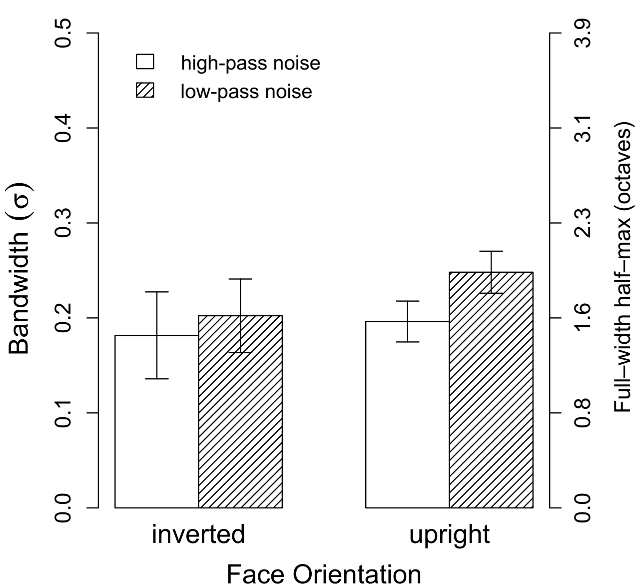

Fig. 5. The values of σ taken from the masking functions that were fit to data from individual subjects in Experiment 1. The bars show the mean of six subjects. Error bars represent ±1 SEM. The *y*-axis on the left shows σ (see Eq. (2)). The axis on the right shows bandwidth expressed as full-width-at-half-maximum.

ference in the spatial frequencies chosen to identify faces, but the current results are inconsistent with this idea. Neither the peak frequency (μ) nor the bandwidth (σ) of masking differed for upright and inverted faces. Furthermore, the magnitude of the inversion effect did not depend on the spatial frequency content of the masking noise. In summary, the current results suggest that observers rely on the same narrow band of spatial frequencies to identify upright and inverted faces.

2. Experiment 2

In naturalistic viewing conditions, faces can be identified despite changes in viewpoint. However, Experiment 1 used only one view per face, and therefore did not require observers to somehow ignore changes in low-level stimulus characteristics that are produced by a change in viewpoint. In these conditions, observers may have identified stimuli on the basis of an image matching process rather than processes that identify faces in more naturalistic conditions (e.g., Bruce et al., 1999). In Experiment 2, this idea was examined by measuring identification thresholds for faces presented at multiple viewpoints.

2.1. Methods

2.1.1. Observers

Fifteen observers (8 males, 7 females; average age = 22 years) participated in the experiment. All but two participants were experienced psychophysical observers, and all but one were naïve about the purpose of the experiment. All observers had an uncorrected or corrected binocular Snellen Acuity of 20/20 or better. The observers were not familiar with the models who were used to create the face stimuli.

2.1.2. Stimuli

Photographs of the faces of five male and five female Caucasian models (average age = 24 years) were taken with a Kodak DX3600 digital camera (1800×1200 pixels) in normal room lighting against a black matte background. The models had no facial hair, no visible piercings, and no eye glasses. Before the photography session began, subjects were positioned so that their eyes and heads faced directly into the camera, and they were told to maintain this head-relative eye position, and a neutral facial expression, during all photographs. Photographs were then taken as the models rotated and tilted their heads to fixate each one of 14 different points located on a wall behind the camera. The fixation points were arranged in two rows of seven, with the middle point in each row falling on an imaginary line that intersected the center of the subject's head and the camera. Within each row, the fixation points were separated by 4.5 deg of visual angle; the top and bottom rows were separated by 7.7 deg. At least two photographs were taken from each viewpoint.

After excluding photographs in which the model blinked or smiled, it was not possible to construct full sets of 14 images for each model. However, every model did yield a usable image in the full-frontal viewpoint, and it was possible to construct sets of images of each model in which left and right diagonal views were nearly balanced. For female models, there were 25 usable images in which the face was rotated leftward, and 24 images in which the face was rotated rightward. For male models, there were 26 usable images for both left and right diagonal views. The experimental stimuli were constructed from these 111 images (i.e., 101 diagonal views and 10 frontal views).

Photographs were stored using the JPEG format and imported into Adobe Photoshop 7.0. An oval aperture with a height:width ratio of 1.5 was superimposed on each image and adjusted so that the height of the oval matched the distance between the highest and lowest parts of the face. The horizontal and vertical coordinates of the oval were adjusted to exclude as much of the non-face background as possible. Images were imported into MATLAB, where they were converted to gray scale by averaging intensity values across each RGB channel. Next, the pixels within the aperture were converted into contrast units, and pixels lying outside the oval aperture were set to zero. Each image was then cropped so that the width of the margins (i.e., the distance between the edge of the aperture and the edge of the image) was equal to one-eighth of the aperture's height. Finally, the edge of the aperture was blurred by convolving it with a 5×5 Gaussian kernel. The resulting image was scaled and centered within a 372×372 pixel matrix. Examples of the stimuli constructed from two models are shown in Fig. 6.

The images were presented on a Sony Trinitron GDM-F520 monitor set to a resolution of 2048×1536 pixels (53 pixels/cm). The frame rate was 75 Hz (non-interlaced). Observers viewed the display binocularly from a distance of 95 cm. The height and width of each face subtended 3.4 and 2.3 deg of visual angle, which were similar to the dimensions of the faces used in Experiment 1. The entire stimulus (372×372 pixels) subtended 4.3×4.3 deg. Faces were presented to observers on a uniform background of average luminance (i.e., no noise) or embedded within filtered or unfiltered white Gaussian noise. Filtered noise was either low- or high-pass filtered with cutoff frequencies of 2.2, 3.2, 5.4, 8.1, 13.4, 21, and 34 cpf. The contrast variance of the unfiltered noise was 0.08, which was the same value used in Experiment 1. However, the spectral density of the noise was lower in Experiment 2 because the increased spatial resolution of the display caused each stimulus pixel to subtend only $\approx 11\%$ of the pixel area used in Experiment 1. As in Experiment 1, the faces themselves were not filtered.

2.1.3. Procedure

The procedure was the same as the one used in Experiment 1, with one important difference. On each trial, the subject was shown a diagonal-view of one face for 100 ms. After the target face disappeared, observers were presented with a selection window that contained all ten frontal-view faces presented without noise at an rms contrast of 0.22. The target faces were always diagonal views; frontal views were shown only in the response selection window. Observers were informed that each facial identity had an equal probability of being shown on each trial, and that the particular diagonal view that was presented was selected randomly from all diagonal views for that face. All other aspects of the testing procedure were the same as in Experiment 1. Noise type and stimulus orientation were blocked, and the order of blocks was randomized in each test session. For each subject and condition, a single psychometric function was fit to the combined data from the last four testing sessions. Threshold was defined as the rms face contrast needed to produce an accuracy rate of 71%.

2.2. Results

Face inversion effects measured in the various conditions were very similar (Fig. 7). The mean inversion effect – averaged first across conditions and then subjects – was 0.208, 95%CI = (0.17,0.25), which was significantly greater than zero, $t(14) = 10.7, p < .0001$ (one-tailed), but did not differ significantly from the mean value of 0.204, $t(14) = 0.22, p = .83$, reported by Martelli et al. (2005). As was found in Experiment 1, the inversion effects measured in the no-noise and all-pass noise conditions did not differ, 95%CI = $(-0.04, 0.06)$, $t(14) = 0.44, df = 14, p = .66$. Finally, the inversion effects measured in Experiment 2 did not differ significantly from those obtained in Experiment 1, 95%CI = $(-0.12, 0.06)$, $t(19) = -0.77, p = .45$.

Inversion effects measured with low-pass and high-pass noise were analyzed in a 7 (cutoff frequency) $\times$ 2 (noise type) within-subject analysis of variance (ANOVA). The main effects of cutoff frequency, $F(6, 84) = 0.8, \tilde{\epsilon} = 0.98, p = .57, \omega_p^2 = 0$, and noise type, $F(1, 14) = 1.95, p = .18, \omega_p^2 = 0.004$, were not significant, nor was the cutoff $\times$ noise type interaction, $F(6, 84) = 1.38$, $\tilde{\epsilon} = 0.94, p = .23, \omega_p^2 = 0.01$. Next, inversion effects measured in the no-noise and all-pass noise conditions were combined with the data from the low-pass and high-pass noise conditions and analyzed in a one-way within-subjects ANOVA. Again, the effect of condition, $F(15, 210) = 1.26, \tilde{\epsilon} = 1, p = .23, \omega_p^2 = 0.01$, was not significant. Hence, as was the case in Experiment 1, our analyses found no evidence that inversion effects varied across conditions.

Fig. 6. Examples of the stimuli used in Experiment 2. The left and middle panels show a subset of the diagonal views for one male and one female model. The right panel shows the frontal views that were used to create the response selection window.

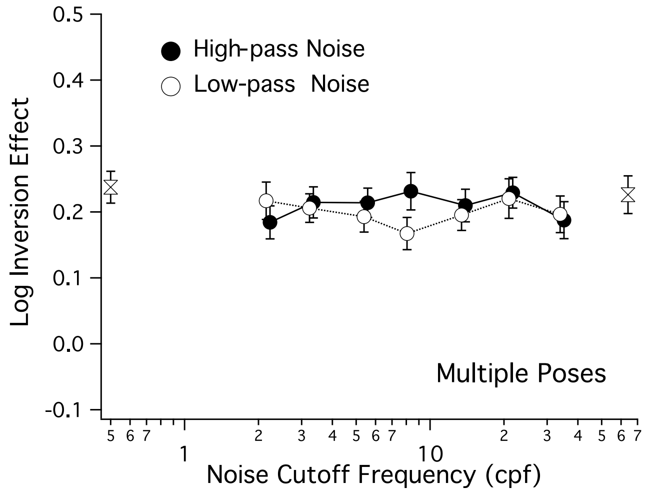

Fig. 7. Average inversion effects measured in Experiment 2 plotted as a function of the cutoff frequency of the noise. The inversion effect was defined as the logarithm of the ratio of thresholds, expressed as rms contrast, obtained with inverted and upright faces. The leftmost and rightmost symbols represent inversion effects measured in the no-noise and all-pass noise conditions. Error bars represent ±1 SEM. Inversion effects did not differ significantly across conditions.

Furthermore, the magnitude of the inversion effect did not differ significantly across experiments.

Fig. 8 depicts average thresholds plotted as a function of noise cutoff frequency, as well as the best-fitting noise-masking functions. A comparison of Figs. 3 and 8 indicates that the amount of masking – i.e., the difference between thresholds measured in the no-noise and all-pass noise conditions – was considerably less in Experiment 2. This reduction in masking likely is due to the lower spectral density of the noise used in Experiment 2 (see Section 2.1.2). Nevertheless, significant masking was observed: thresholds in the all-pass and no-noise conditions differed by 0.51 and 0.49 log units in the upright and inverted conditions, respectively, and these differences were statistically significant (upright faces: $t(14) = 8.89, p < .0001$; inverted faces: $t(14) = 9.87, p < .0001$). Moreover, the noise-masking functions fit the average data, as well as data from individual subjects, reasonably well. Thus, although the masking was reduced relative to the amount obtained in the first experiment, it was sufficiently strong for us to estimate the parameters of the noise-masking functions. The log-Gaussian functions (e.g., Eq. (2)) used to derive the masking functions are shown in the lower panels of Fig. 8. For both upright and inverted faces, the functions derived with low-pass and high-pass noise had nearly equal bandwidths (upright: $\sigma \approx 0.18$; inverted: $\sigma \approx 0.16$),

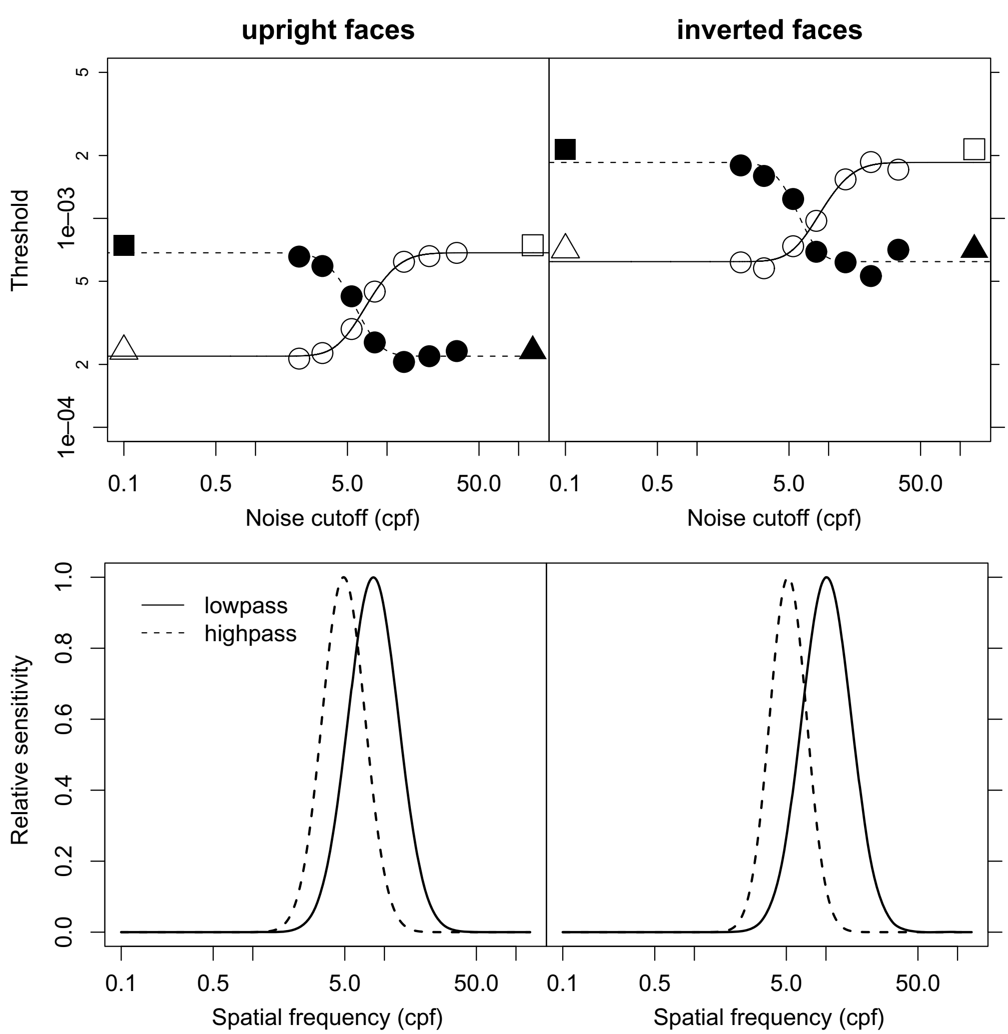

Fig. 8. Mean results obtained in Experiment 2 with upright (left column) and inverted faces (right column). Plotting conventions are the same as in Fig. 3.

but appeared to differ in peak sensitivity. For upright faces, peak sensitivity, or μ, was 4.9 and 8.2 cpf in the high- and low-pass noise conditions, respectively. For inverted faces, μ was 5.2 and 10.1 cpf in the high-pass and low-pass conditions. The peaks and bandwidths of the log-Gaussian functions did not appear to depend significantly on stimulus orientation.

Masking functions were fit to the data obtained from individual subjects in each condition. The mean values of μ (averaged across subjects) are shown in Fig. 9. As was observed in Experiment 1, μ was slightly higher in the low-pass noise condition than in the high-pass condition. A 2 (face orientation) $\times$ 2 (noise type) within-subjects ANOVA on the log-transformed values of μ found a significant main effect of noise type, $F(1,14) = 6.91$, $p = .02$, $\omega_p^2 = 0.09$, which reflected the fact that μ was greater in the low-pass noise condition. Hence, unlike what was found in Experiment 1, the effect of noise type on μ was statistically significant. This difference between experiments likely reflects the fact that Experiment 2 used 15 subjects, rather than 6 as in Experiment 1, and therefore had higher statistical power. Despite this increase in power, the effect of orientation on μ, $F(1,14) = 0.82, p = .38, \omega_p^2 = 0$, and the noise $\times$ orientation interaction, $F(1,14) = 2.07, p = .17, \omega_p^2 = 0.02$, were not significant. Mean values of σ are shown in Fig. 10. A 2 (face orientation) $\times$ 2 (noise type) within-subjects ANOVA on the log-transformed values of σ failed to find any significant effects (noise: $F(1,14) = 0.051$, $p = .83$, $\omega_p^2 = 0$; orientation: $F(1,14) = 0.01, p = .93, \omega_p^2 = 0$; noise $\times$ orientation: $F(1,14) = 0.01, p = .91, \omega_p^2 = 0$).

Finally, we examined whether individual differences in the size of the inversion effect were correlated with individual differences in the shapes of the masking functions measured with upright and inverted faces. For each subject, we calculated the difference between peak sensitivity ($\triangle\mu$) and bandwidth ($\triangle\sigma$) that were estimated with upright and inverted faces. Also, a single estimate of the inversion effect was computed for each subject in both experiments by averaging the inversion effects obtained in the various conditions. The average inversion effect was not correlated with $\triangle\mu$, $r = -.16, t(19) = -0.7, p = .49$, or $\triangle\sigma$, $r = .09, t(19) = 0.39, p = .70$. Hence, there was no evidence that individual differences in the size of the inversion effect were correlated with subtle changes in the masking functions obtained with upright and inverted faces.

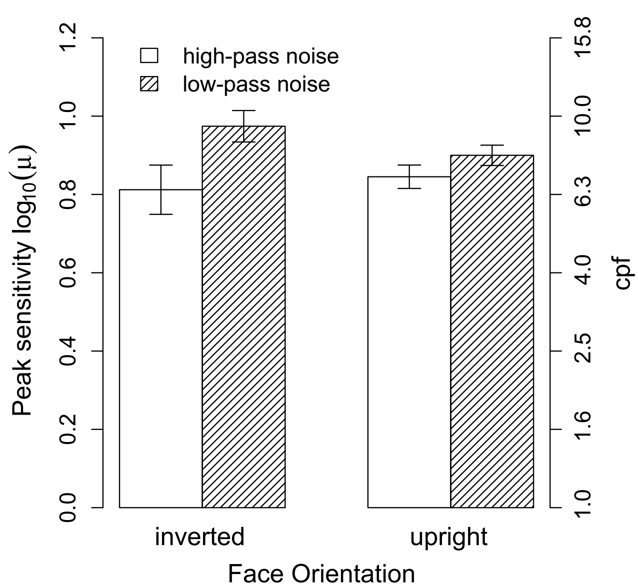

Fig. 9. The values of μ taken from the masking functions that were fit to data from individual subjects in Experiment 2. The bars show the mean of 15 subjects. The y-axis on the left shows log-transformed μ (see Eq. (2)). The axis on the right shows the same values transformed to units of cpf. Error bars represent ±1 SEM.

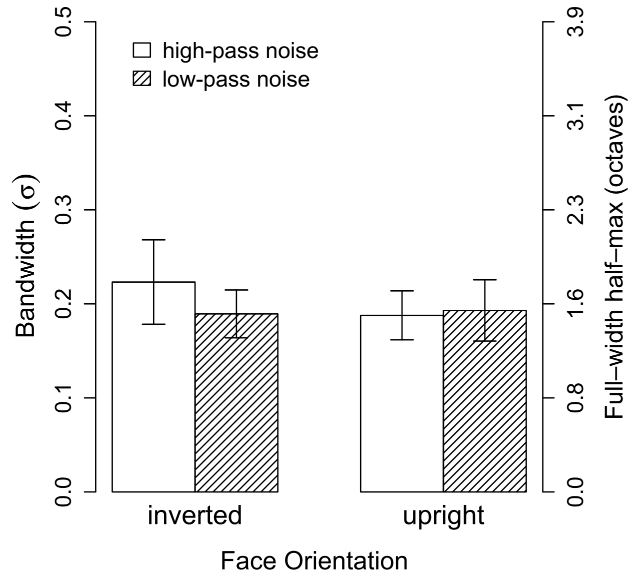

Fig. 10. The values of σ taken from the masking functions that were fit to data from individual subjects in Experiment 2. The bars show the mean of 15 subjects. Error bars represent ±1 SEM. The y-axis on the left shows σ (see Eq. (2)). The axis on the right shows bandwidth expressed as full-width-at-half-maximum.

2.3. Discussion

Experiment 2 examined the effect of increasing the number of viewpoints on the magnitude of the face inversion effect and the spatial frequency selectivity of face identification. Surprisingly, we found that increasing the number of viewpoints had virtually no effect on the face inversion effect or on the frequency selectivity of identification. As was found in Experiment 1, the face inversion effect did not vary across conditions that used white noise, no noise, and different types of low-pass and high-pass noise. Furthermore, the magnitude of the face inversion effect was essentially the same in the two experiments. Also, we found that varying the number of viewpoints had little effect on the spatial frequency selectivity of masking: as in Experiment 1, identification thresholds for upright and inverted faces were influenced most strongly by masking components that were within ≈ 1.5 octaves of 6–8 cpf (Figs. 9 and 10). Finally, across both experiments, individual differences in the inversion effect were not correlated with individual differences in $\triangle\mu$ or $\triangle\sigma$.

Experiment 2 found that μ was slightly higher when faces were embedded in low-pass filtered noise. Although this effect was not significant in Experiment 1, the trend was similar: the difference in μ was ≈ 0.1 log units in both experiments, and the association strength between noise type and μ actually was greater in Experiment 1 than Experiment 2. Therefore, the different results probably reflect a difference in statistical power across experiments. Parameter μ governs the lateral position of the masking function along the log spatial frequency axis in Fig. 8. Hence, differences in μ mean that masking functions obtained with low-pass noise were shifted to higher spatial frequencies relative to those obtained with high-pass noise. These shifts in the masking function may have been caused by a visual analog of off-frequency listening (Patterson & Nimmo-Smith, 1980; Solomon & Pelli, 1994). If an observer could select a channel that maximized the signal-to-noise ratio, then it would be advantageous to select a filter tuned to higher spatial frequencies when the faces were embedded in low-pass noise, and select a low frequency filter when the faces were embedded in high-pass noise. The current results are consistent with the hypothesis that observers selected channels with peak frequencies that differed by 0.1 log units in response to the noise. However, this flexibility in channel selection did not vary across face orientations, nor were individual differences in $\triangle\mu$ correlated with individual differences in the inversion effect. It is unlikely, therefore, that off-frequency looking contributed to differences between upright and inverted faces.

3. General discussion

We measured face identification thresholds for upright and inverted faces embedded in different types of noise. We found that the face inversion effect did not vary significantly across a wide range of conditions that used white noise, no noise, and low- and high-pass noise that differed in cutoff frequency. Quantitative modeling of the masking functions revealed that subjects used information conveyed by similar narrow bands of spatial frequencies to identify upright and inverted faces. The current estimate of the critical band of spatial frequency for identification – roughly 1.5 octaves wide and centered on ≈ 7 cpf – is similar to several previous estimates that were derived for upright faces using different methods (e.g., Boutet et al., 2003; Costen et al., 1996; Gold et al., 1999; Nasanen, 1999; Peli et al., 1994). Hence, the spatial frequency selectivity of face identification is reasonably robust to changes in stimuli and experimental procedure. Finally, varying the number of viewpoints in the stimulus set did not alter the size of the face inversion effect or the spatial frequency selectivity of masking. The current findings are consistent with a recent study by Willenbockel et al. (2008), who used the Bubbles method (Gosselin & Schyns, 2001) to show that face inversion does not alter the spatial frequencies that are correlated with responses in a face identification task.

We did not find any significant differences between the frequency selectivity of upright and inverted face identification. However, other aspects of face perception – for example, judgements of gender, emotional state, or attractiveness – may rely on information carried by distinct bands of spatial frequency (Smith, Gosselin, & Schyns, 2004, 2007) which may vary with stimulus orientation. Alternatively, subjects may use information carried by different frequency bands in a flexible manner to make judgments about upright faces (Schyns & Oliva, 1999; Sowden & Schyns, 2006) but not inverted faces. In other words, larger differences between upright and inverted faces may emerge if subjects are asked to make different facial judgments.

Goffaux, Hault, Michel, Vuong, and Rossion (2005) used low-pass and high-pass spatial filtering to measure the effect of spatial frequency on face discrimination. When the faces were changed by substituting the original eyes with eyes from another face, discrimination was easier with faces that contained high spatial frequencies, but when the faces were changed by moving the eyes horizontally or vertically, discrimination was easier with faces that contained low spatial frequencies. Goffaux et al. (2005) concluded that their experiments provided evidence for "a functional dissociation between LSFs [low spatial frequencies] and HSFs [high spatial frequencies] in supporting the extraction of configural and featural cues for face processing" (p. 83). Holistic and configural processing are thought to be especially important for the perception of upright faces (Maurer et al., 2002), and therefore Goffaux et al.'s (2005) findings suggest that there ought to be differences between the frequency selectivity of upright and inverted face orientation. Our results do not support that prediction. However, this dissociation between low and high spatial frequencies has not been found in other experiments (Boutet et al., 2003; Wenger & Townsend, 2000) that used methods that were similar to those used by (Goffaux et al., 2005), and so it is unclear how Goffaux et al.'s findings bear on the current experiments.

Different evidence for a dissociation between the roles played by low and high spatial frequencies in face processing comes from a study by Goffaux and Rossion (2006). They found that the magnitudes of the Composite Face Effect (Young et al., 1987) and the Whole-Part advantage effect (Tanaka & Farah, 1993) were larger for faces that were low-pass filtered than faces that were high-pass filtered. Both the Composite Face Effect and the Whole-Part advantage effect are larger for upright than inverted faces, and have been interpreted as providing evidence for greater holistic processing of upright faces (Maurer et al., 2002). If the holistic processing that is measured by these effects is important for face identification, then embedding faces in low frequency noise should impair identification of upright faces more than inverted faces. We did not find such an effect: the spatial frequency selectivity of masking did not vary with face orientation. A potential explanation for our failure to find the predicted result is that holistic processing, *as indexed by the Composite Face and Whole-Part effects*, does not influence face identification significantly. Consistent with this view, Konar, Bennett, and Sekuler (2007, 2008) found that the magnitude of the Composite Face Effect is not correlated with accuracy on face identification tasks. It is plausible, therefore, to suggest that face identification is not constrained by the processes that give rise to the Composite Face Effect. (To our knowledge, no similar study has examined the correlation between face identification accuracy and the Whole-Part effect.) Thus, although the current results suggest that similar bands of spatial frequencies contribute to the identification of upright and inverted faces, they are not necessarily inconsistent with the claim that the Composite Face Effect and the Whole-Part effect are produced by mechanisms that are more sensitive to low spatial frequencies.

Using the classification image technique, Sekuler et al. (2004) found that subjects used information conveyed by pixels near the eyes and brows to identify both upright and inverted faces, but that the information was used less efficiently when faces were inverted (see also, Gaspar et al., 2008). The current results suggest that the relative inefficiency of inverted face identification is not caused by subjects using a different, less-informative spatial frequency band. It is possible, however, that the same spatial frequencies are used in different ways to identify upright and inverted faces. For example, subjects might use linear and/or non-linear filters that are most sensitive to 8 cpf to encode local structure near an eye brow in an upright face, but use similar filters to construct a "coincidence detector" (Morgan & Regan, 1987; Morgan, Ward, & Hole, 1990) that encodes the distance between the eyes in an inverted face. If the former strategy yields more information about face identity than the latter one, then efficiency would be lower for inverted faces even though subjects use the same spatial frequencies for both orientations. It is also possible that the same computations are performed on upright and inverted faces, but that other factors—differences in spatial uncertainty, for example—place greater constraints on sensitivity when faces are inverted. The current experiments do not allow us to distinguish among these possibilities because they show only which spatial frequencies are used to identify faces, not how they are used. Some constraints on how processing of upright and inverted faces can differ come from previous studies: the results of Sekuler et al. (2004), for example, suggest that non-linear mechanisms make similar contributions to upright and inverted face identification, and Gaspar et al. (2008) showed that processing differences affect efficiency but not internal noise. The current findings impose an additional constraint, namely that identification processes do not differ in spatial frequency selectivity.

We conclude by considering why face identification is based on spatial frequencies near 6–10 cpf (Gold et al., 1999; Nasanen, 1999). Keil (2008) analyzed the spatial frequency spectra of 1700 gray scale images of human faces. After removing external features (e.g., background objects, hair, shoulders) by applying a spatial window to the original image, the average, whitened amplitude spectrum exhibited a conspicuous peak near 10 cpf. Further analyses indicated that the 10 cpf peak occurred primarily along the horizontal orientation, and Keil suggested that it was caused by the eyes and/or the mouth. Gaspar et al. (2006) conducted an analysis of the spatial frequency content of Thatcherized faces (Thompson, 1980) that is consistent with this view. Specifically, Gaspar found that information that distinguishes an eyes-only Thatcherized face from the original face is conveyed by a narrow range of spatial frequencies centered on 10 cpf, which suggests that structure near the eyes is carried by spatial frequencies near 10 cpf. Hence, the spatial frequency selectivity of face identification may simply reflect the fact that subjects rely strongly on the eyes and brows to identify upright and inverted faces.

References

Bahrick, H., Bahrick, O., & Wittlinger, R. (1975). Fifty years of memory for names and faces: A cross-sectional approach. *Journal of Experimental Psychology: General, 104*(1), 54–75.

Boutet, I., Collin, C., & Faubert, J. (2003). Configural face encoding and spatial frequency information. *Perception & Psychophysics, 65*(7), 1078–1093.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*(4), 433–436.

Bruce, V., Henderson, Z., Greenwood, K., Hancock, P., Burton, A., & Miller, P. (1999). Verification of face identities from images captured on video. *Journal of Experimental Psychology: Applied, 5*, 339–360.

Busey, T. A., Schneider, B. L., & Wyatte, D. (2008). Expertise and the width of the visual filter in fingerprint examiners. *Journal of Vision, 8*(6), 178a.

Collin, C. A., Liu, C. H., Troje, N. F., McMullen, P. A., & Chaudhuri, A. (2004). Face recognition is affected by similarity in spatial frequency range to a greater degree than within-category object recognition. *Journal of Experimental Psychology: Human Perception and Performance, 30*(5), 975–987.

Costen, N. P., Parker, D. M., & Craw, I. (1996). Effects of high-pass and low-pass spatial filtering on face identification. *Perception & Psychophysics, 58*(4), 602–612.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General, 115*(2), 107–117.

Dosher, B. A., & Lu, Z. L. (1998). Perceptual learning reflects external noise filtering and internal noise reduction through channel reweighting. *Proceedings of the National Academy of Sciences of the United States of America, 95*(23), 13988–13993.

Gaspar, C. M., Bennett, P. J., & Sekuler, A. B. (2008). The effects of face inversion and contrast-reversal on efficiency and internal noise. *Vision Research, 48*(8), 1084–1095.

Gaspar, C. M. (2006). *The effects of stimulus information and orientation on face processing*. Unpublished doctoral dissertation, McMaster University, Hamilton, Ont., Canada.

Goffaux, V., Hault, B., Michel, C., Vuong, Q. C., & Rossion, B. (2005). The respective role of low and high spatial frequencies in supporting configural and featural processing of faces. *Perception, 34*(1), 77–86.

Goffaux, V., & Rossion, B. (2006). Faces are "spatial"–holistic face perception is supported by low spatial frequencies. *Journal of Experimental Psychology: Human Perception and Performance, 32*(4), 1023–1039.

Gold, J., Bennett, P. J., & Sekuler, A. B. (1999). Identification of band-pass filtered letters and faces by human and ideal observers. *Vision Research, 39*(21), 3537–3560.

Gosselin, F., & Schyns, P. G. (2001). Bubbles: A technique to reveal the use of information in recognition tasks. *Vision Research, 41*(17), 2261–2271.

Keil, M. (2008). Does face image statistics predict a preferred spatial frequency for human face processing? *Proceedings of Biological Science, 275*(1647), 2095–2100.

Kirk, R. E. (1995). *Experimental design: Procedures for the behavioral sciences* (3rd ed.). Pacific Grove, CA: Brooks/Cole.

Konar, Y., Bennett, P. J., & Sekuler, A. B. (2007). The composite face effect is not correlated with face identification accuracy. *Journal of Vision, 7*(9), 501a.

Konar, Y., Bennett, P. J., & Sekuler, A. B. (2008). The composite face effect is still not correlated with face identification accuracy. *Journal of Vision, 8*(6), 891a.

Martelli, M., Majaj, N. J., & Pelli, D. G. (2005). Are faces processed like words? A diagnostic test for recognition by parts. *Journal of Vision, 5*(1), 58–70.

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences, 6*(6), 255–260.

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Morgan, M. J., & Regan, D. (1987). Opponent model for line interval discrimination: Interval and vernier performance compared. *Vision Research, 27*(1), 107–118.

Morgan, M. J., Ward, R. M., & Hole, G. J. (1990). Evidence for positional coding in hyperacuity. *Journal of Optical Society of America A, 7*(2), 297–304.

Nasanen, R. (1999). Spatial frequency bandwidth used in the recognition of facial images. *Vision Research, 39*(23), 3824–3833.

Patterson, R. D. (1976). Auditory filter shapes derived with noise stimuli. *Journal of the Acoustical Society of America, 59*, 640–654.

Patterson, R. D., & Nimmo-Smith, I. (1980). Off-frequency listening and auditory-filter asymmetry. *Journal of the Acoustical Society of America, 67*(1), 229–245.

Peli, E., Lee, E., Trempe, C. L., & Buzney, S. (1994). Image enhancement for the visually impaired: The effects of enhancement on face recognition. *Journal of Optical Society of America: A Optics, Image Science, and Vision, 11*(7), 1929–1939.

Pelli, D. G. (1997). The video toolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*(4), 437–442.

R Development Core Team. (2007). *R: A language and environment for statistical computing*. Vienna, Austria. http://www.R-project.org.

Schneider, B. L., DeLong, J. E., & Busey, T. A. (2007). Added noise affects the neural correlates of upright and inverted faces differently. *Journal of Vision, 7*(4), 4.

Schyns, P. G., & Oliva, A. (1999). Dr. Angry and Mr. Smile: When categorization flexibly modifies the perception of faces in rapid visual presentations. *Cognition, 69*(3), 243–265.

Sekuler, A. B., Gaspar, C. M., Gold, J. M., & Bennett, P. J. (2004). Inversion leads to quantitative, not qualitative, changes in face processing. *Current Biology, 14*(5), 391–396.

Smith, M. L., Gosselin, F., & Schyns, P. G. (2004). Receptive fields for flexible face categorizations. *Psychological Science, 15*(11), 753–761.

Smith, M. L., Gosselin, F., & Schyns, P. G. (2007). From a face to its category via a few information processing states in the brain. *Neuroimage, 37*(3), 974–984.

Solomon, J., & Pelli, D. (1994). The visual filter mediating letter identification. *Nature, 369*(6479), 395–397.

Sowden, P. T., & Schyns, P. G. (2006). Channel surfing in the visual brain. *Trends in Cognitive Sciences, 10*(12), 538–545.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *The Quarterly Journal of Experimental Psychology A, 46*(2), 225–245.

Tanskanen, T., Nasanen, R., Montez, T., Paallysaho, J., & Hari, R. (2005). Face recognition and cortical responses show similar sensitivity to noise spatial frequency. *Cerebral Cortex, 15*(5), 526–534.

Thompson, P. (1980). Margaret thatcher: A new illusion. *Perception, 9*(4), 483–484.

Valentine, T. (1988). Upside-down faces: A review of the effect of inversion upon face recognition. *British Journal of Psychology, 79*(4), 471–491.

Wenger, M. J., & Townsend, J. T. (2000). Spatial frequencies in short-term memory for faces: A test of three frequency-dependent hypotheses. *Memory & Cognition, 28*(1), 125–142.

Willenbockel, V., Fiset, D., Chauvin, A., Blais, C., Arguin, M., Tanaka, J., et al. (2008). The face inversion effect is nothing "spatial". *Journal of Vision, 8*(6), 153a.

Yin, R. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81*(1), 141–145.

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception, 16*(6), 747–759.