Vision Research 211 (2023) 108281

Contents lists available at ScienceDirect

Vision Research

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

Detection of emotional faces: The role of spatial frequencies and local features

Léa Entzmann [a,b,c,*], Nathalie Guyader [b], Louise Kauffmann [a], Carole Peyrin [a], Martial Mermillod [a]

[a] Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, LPNC, 38000 Grenoble, France
[b] Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, 38000 Grenoble, France
[c] Icelandic Vision Lab, School of Health Sciences, University of Iceland, Reykjavík, Iceland

ARTICLE INFO

Keywords:
Facial expressions
Eye movements
Spatial frequencies
Contrast
Visual saliency

ABSTRACT

Models of emotion processing suggest that threat-related stimuli such as fearful faces can be detected based on the rapid extraction of low spatial frequencies. However, this remains debated as other models argue that the decoding of facial expressions occurs with a more flexible use of spatial frequencies. The purpose of this study was to clarify the role of spatial frequencies and differences in luminance contrast between spatial frequencies, on the detection of facial emotions. We used a saccadic choice task in which emotional-neutral face pairs were presented and participants were asked to make a saccade toward the neutral or the emotional (happy or fearful) face. Faces were displayed either in low, high, or broad spatial frequencies. Results showed that participants were better to saccade toward the emotional face. They were also better for high or broad than low spatial frequencies, and the accuracy was higher with a happy target. An analysis of the eye and mouth saliency of our stimuli revealed that the mouth saliency of the target correlates with participants' performance. Overall, this study underlines the importance of local more than global information, and of the saliency of the mouth region in the detection of emotional and neutral faces.

1. Introduction

In everyday life, the ability to detect a face is crucial, as faces convey essential information for appropriate social behaviour, especially through their expression. Thereby, the human visual system developed specific mechanisms to efficiently process faces (Haxby et al., 2000; Liu et al., 2002; Zhao et al., 2018) and several behavioural experiments highlighted the preferential processing of face stimuli (Cerf et al., 2009; Coutrot & Guyader, 2014; Farah et al., 1998; Langton et al., 2008). For example, some studies using eye-tracking showed that faces can be detected as early as 100 ms, whereas more time is needed for the detection of other stimuli, such as vehicles or animals (Crouzet, 2010; Entzmann et al., 2021; Kauffmann et al., 2019, 2021).

This efficient and fast processing of face stimuli might be explained by the fact that face detection relies on coarse more than fine information (Awasthi et al., 2011; Goffaux et al., 2011; Goffaux & Rossion, 2006; Guyader et al., 2017; Peters et al., 2018; Quek et al., 2018). In fact, visual perception is thought to be based on the parallel extraction of different visual features on different spatial frequencies and to follow a default, predominantly coarse-to-fine processing sequence (Bar, 2003; Hegde, 2008; Kauffmann et al., 2014; Kauffmann, Chauvin, et al., 2015; Musel et al., 2012; Petras et al., 2019; Peyrin et al., 2010; Schyns & Oliva, 1994). Low spatial frequencies (LSF; carrying coarse information) would be extracted first and rapidly processed allowing to form a coarse representation of the visual input, while high spatial frequencies (HSF), which convey finer information (e.g., details and edges), would be carried more slowly and used to refine the first LSF-based analysis. Several studies highlighted the primary role of LSF information in face processing (Awasthi et al., 2011; Goffaux et al., 2011; Goffaux & Rossion, 2006; Guyader et al., 2017; Peters et al., 2018; Quek et al., 2018). For example, Goffaux and Rossion (2006) showed that the whole-part advantage (i.e., superior recognition of the eyes when it is presented in the context of a whole face rather than isolated) in an identity matching task was larger with LSF and broad spatial frequency (BSF; i.e., unfiltered images) than HSF images. Moreover, Guyader et al. (2017) presented simultaneously a face and a vehicle image, both presented

* Corresponding author at: LPNC, 1251 Avenue Centrale, Batiment Michel Dubois, 38400 Saint-Martin-d'Heres, France.
E-mail address: leaentzmann@hi.is (L. Entzmann).

https://doi.org/10.1016/j.visres.2023.108281
Received 14 November 2022; Received in revised form 18 June 2023; Accepted 28 June 2023
Available online 6 July 2023

either in BSF, HSF or LSF, and asked participants to make a saccade toward the face. They observed that within 130–140 ms participants were able to make more correct than incorrect saccades to faces in BSF or LSF, whereas they required more time for images presented in HFS.

Therefore, the processing of faces may primarily rely on the extraction of LSF information. Previous studies have demonstrated that face identification is dependent on spatial frequencies between 7 and 16 cycles per face (e.g., Gaspar et al., 2008; Gold et al. 1999; Näsänen, 1999; Willenbockel et al., 2010). However, the role of spatial frequencies in the detection of facial expressions is not as well-established. Facial expressions provide indications on the emotional state of others, and consequently, on the identification of potential threats. Classical views of the visual processing of emotional stimuli in the brain suggest that LSF play a crucial role in the perception of threat-related stimuli, such as fearful faces. This hypothesis relies on the existence of a short and direct superior colliculus-pulvinar pathway to the amygdala which enables rapid detection of threats (LeDoux, 2000; Méndez-Bértolo et al., 2016; Morris, 1998; Öhman, 2005; Tamietto & de Gelder, 2010; Vuilleumier et al., 2003). This subcortical pathway is supposed to transmit LSF information in parallel to a slower cortical pathway that transmits finer information. This view is notably supported by an fMRI study, showing overall higher activation in the amygdala when fearful rather than neutral faces were presented to participants (Vuilleumier et al., 2003). Specifically, this effect was only significant if the image presented was in LSF or BSF. Similarly, using intracranial EEG recordings in the amygdala fearful faces were found to evoke an early activity (75 ms post-stimulus onset), but only when they were in BSF or LSF (Méndez-Bértolo et al., 2016). Finally, Burra et al. (2019) demonstrated that presenting fearful faces in the blind visual field of a patient with cortical blindness did not elicit an amygdala response when the images were presented in HFS. Overall, these neuroimaging studies suggest that facial expressions can be detected rapidly, through LSF information only. In this sense, computational studies showed that the information carried by LSF is sufficient to discriminate facial expressions (Mermillod et al., 2009, 2010). Using fast periodic visual stimulation and EEG, Van der Donck et al. (2020) found that at least spatial frequencies higher than 5.93 cycles per image were needed to differentiate fearful from neutral faces. But, their results also suggest that HSF plays a key role in fear discrimination, as the oddball EEG signal was strongest in HSF (between 94.8 and 189 cycles per image).

The existence of a subcortical pathway for the rapid detection of fearful faces, as well as the crucial role of LSF, is still a matter of debate (e.g., Pessoa & Adolphs, 2010). For example, McFadyen et al. (2017) used magnetoencephalography to measure neural activity while participants performed a gender discrimination task of neutral and fearful faces in LSF or HSF. They demonstrated through dynamic causal modeling that the most likely underlying subcortical neural network consisted of a pulvinar-amygdala connection that was neither influenced by spatial frequencies nor by emotions, in line with other dynamic causal modeling results. Overall, studies in the literature did not consistently find differences between neutral and fearful faces in LSF or unfiltered images (Corradi-Dell'Acqua et al. 2014; Fitzgerald et al., 2006; Garvert et al., 2014; Ottaviani et al. 2012). And, some studies suggested that responses in the visual cortex can occur with latencies similar to those observed in subcortical areas (e.g., Krolak-Salmon et al., 2004). Also, several studies underlined a more flexible use of spatial frequencies to decode facial expressions, occurring later in the time course of visual processing (Schyns et al., 2009). Different spatial frequency bands would be used depending on the facial expression (Cassidy et al., 2021; Kumar & Srinivasan, 2011; Morrison & Schyns, 2001; Oliva & Schyns, 1997; Smith & Schyns, 2009) and the task (Schyns & Oliva, 1999; Smith & Merlusca, 2014). For example, fearful face categorisation might rely mostly on the wide-opened eyes and therefore on the extraction of HSF, whereas a larger scale would be used for other expressions (Adolphs et al., 2005; Smith et al., 2005; Stein et al., 2014). Also, Schyns and Oliva (1999) found that when participants were asked to categorise facial expressions as angry, happy, or neutral, LSF were used more often. In contrast, when the task was to indicate whether the face was expressive or neutral, HSF were used more often. However, such behavioural studies on the role of spatial frequencies in the processing of facial expressions used manual responses and may reflect processes occurring at a later stage of visual processing than what is described in neuroimaging studies on fast fear detection.

In a previous study, we tested the detection of emotional and neutral faces using a saccadic choice task. More precisely, emotional-neutral pairs of faces were presented to participants, who were asked to make a saccade as fast and as accurately as possible toward the emotional (happy or fearful) or the neutral face (Entzmann et al., 2021). The use of saccadic eye movements as a behavioural response provides a more precise measure of visual processing speed than what allows manual responses (Kirchner & Thorpe, 2006). Also, in our opinion, this paradigm is ecological since in everyday life faces rarely appear in central vision, and a natural response to emotional stimuli is to orient the gaze towards it. But, more importantly, using more peripherally presented stimuli allows us to test the subcortical processing of emotional faces in a more appropriate setting. Indeed, the subcortical pathway is supposed to allow the rapid detection of a threat in peripheral vision based on LSF information provided by rod photoreceptors (e.g., LeDoux, 2000; Tamietto & de Gelder, 2010). At the neural level, the subcortical pathway is assumed to reach the amygdala faster, and from there the information could reach the superior colliculus to trigger saccadic responses (Mulckhuyse, 2018). We aimed to test the possibility that saccadic responses (which cannot be assessed with centrally presented stimuli) could provide a better assessment of the fast detection of LSF stimuli (e.g., fearful faces). Previously, we found that participants were faster and made fewer errors when they had to saccade toward the emotional than the neutral face. This was especially the case when the emotional face was happy (Entzmann et al., 2021). Using an artificial neural network, we showed that this advantage for happy faces can, at least partly, be explained by perceptual factors. More precisely, the network was trained and tested on its ability to discriminate a neutral from an emotional face, and performed better when the emotional face was happy rather than fearful. We also analysed the saccade endpoints and we observed that saccades landed lower on happy than fearful faces, suggesting that local features like the mouth or the eyes attracted the gaze differently according to the emotional expression. Indeed, the eyes and the mouth are the most useful parts to discriminate facial expressions (Blais et al., 2012; Eisenbarth & Alpers, 2011; Smith et al., 2005; Smith & Schyns, 2009; Wegrzyn et al., 2017). It is likely that the saliency of these regions, or their respective diagnosticity, varies according to the expression, and shifts the endpoints of the saccades.

In the present paper, we reproduced the eye-tracking experiment introduced by Entzmann et al. (2021), but with different spatial frequency conditions. This can be seen as an equivalent to what Guyader et al. (2017) did for face detection, as they use a saccadic choice task with face-vehicle pairs presented in HSF, LSF, and BSF. In our previous study participants were required to switch target categories, the emotional or the neutral face, between sessions. Despite that in each session the instruction requires the participants to distinguish between the same emotional and neutral faces, we found that participants were faster and made fewer errors when saccading towards emotional faces. This result aligns with existing literature showing that emotional targets are detected more efficiently than neutral ones in visual search (see for a review, Frischen et al., 2008) or saccadic choice tasks (Bannerman et al., 2009; D'Hondt et al., 2016). Then, even when participants are required to distinguish between two images, the more salient stimulus is detected first, capturing their attention. This heightened processing may arise from either the emotional value of the stimulus, its salient perceptual features, or both (e.g., Lundqvist et al., 2015; Mermillod et al., 2009; Öhman, 2005). At the brain level, it would be associated with increased activation in the amygdala or the visual cortex (although activation in the visual cortex may be modulated by signals from the amygdala;

Vuilleumier, 2015). Indeed, fearful faces are strongly linked with amygdala activation, but the amygdala is generally more active in response to emotional than neutral faces (Fusar-Poli et al., 2009; Sander et al., 2003). Therefore, we suggest that emotional faces may have a higher signal in the saccade map, particularly for short-latency saccades, which would bias attention toward them (Bisley & Mirpour, 2019; Klink et al., 2014). Overall, the superior-colliculus-pulvinar pathway could be involved in this effect by transmitting rapidly the LSF information about facial expressions to the amygdala (Mulckhuyse, 2018).

Our purpose was first to clarify the role of spatial frequencies in the detection of emotional faces. Several methods can be employed to investigate the use of spatial frequency content in a task. Most of the studies that we have cited in the previous paragraphs used fixed high and low-pass filters to isolate the high or low spatial frequencies of an image (e.g., Goffaux & Rossion 2006; Guyader et al., 2017; Kumar & Srinivasan, 2011; McFadyen et al., 2017; Méndez-Bértolo et al., 2016; Vuilleumier et al., 2003). A high-pass filter retains only the spatial frequencies above a chosen cutoff frequency, while a low-pass filter retains only the spatial frequencies below a chosen cutoff frequency. With this method, cutoff frequencies are fixed beforehand, and can vary between studies. Other methods, such as spatial frequency *Bubbles* (Willenbockel et al., 2010; Plouffe-Demers et al., 2019; Charbonneau et al., 2021), band masking (Gaspar 2008), and bandpass filtering (Gold et al., 1999), can also be used to manipulate the spatial frequency content of stimuli. Spatial frequency *Bubbles*, for example, involves presenting stimuli with randomly sampled patches of spatial frequencies to determine which frequencies are important for a particular task. This method requires more trials, but it is particularly useful for identifying the most diagnostic spatial frequencies for a task, with the computation of precise spatial frequency tuning curves. Band masking involves masking specific frequency bands of a stimulus to determine the contribution of those bands, and bandpass filtering involves filtering out all frequencies except a specific range to selectively examine the contribution of those frequencies. In our study, our goal was not to assess diagnostic frequency bands, but rather to dissociate the contribution of low and high spatial frequencies related to the contribution of magnocellular versus parvocellular pathways, respectively (e.g., DeValois et al., 1990; McFadyen et al., 2017; Méndez-Bértolo et al., 2016; Tamietto & de Gelder, 2010; Vuilleumier et al., 2003). In connection with previous neurophysiological and neuroimaging studies we hypothesise that the rapid processing of LSF in the subcortical pathway could induce an early gaze capture. But, we don't assume these frequencies are the most diagnostic, and previous studies showing evidence for an early LSF processing of emotional faces used high and low pass filters. For these reasons, we decided to use high and low pass filtered images to create HSF and LSF stimuli.

Here, we also considered methodological issues about the filtering procedure that may explain discrepancies between studies (Kauffmann, Chauvin, et al., 2015; Perfetto et al., 2020; Vlamings et al., 2009). Filtered images, like LSF and HSF images, differ not only in their spatial frequency content but also in their luminance contrast (i.e., the magnitude of luminance variation in a stimulus relative to its mean luminance, more simply referred to as contrast; Shapley & Enroth-Cugell, 1984). This is because the luminance contrast in scenes decreases as spatial frequency increases (Field, 1987), leading LSF images to have higher luminance contrast than HSF. Such luminance contrast differences between HSF and LSF can influence the processing of the spatial frequency content in visual stimuli. Indeed, previous studies have shown that high contrast stimuli can be detected faster than low-contrast stimuli (e.g., Ludwig et al., 2004). Actually, in studies using visual scenes as stimuli, the temporal advantage of LSF on HSF processing, was attenuated by the equalisation of luminance contrast between LSF and HSF (Kauffmann, Chauvin, et al., 2015; Kauffmann, Ramanoël, et al., 2015). This suggests that the higher luminance contrast of LSF may contribute to the classical effect of the predominance of LSF on HSF information. Importantly, such methodological issues may also explain discrepancies observed in past studies investigating the role of spatial frequencies on emotional face processing. For example, McFadyen et al. (2017) used LSF and HSF faces that were equalised in luminance contrast, unlike previous neuroimaging studies (e.g., Méndez-Bértolo et al., 2016; Vuilleumier et al., 2003). This may explain the absence of selectivity of the amygdala to LSF fearful faces and suggests that higher response in the amygdala for LSF faces may be due to their higher contrast.

We used a saccadic choice task with emotional-neutral face pairs, in which we manipulated the spatial frequency content of images. Participants were asked to make a saccade as rapidly as possible toward the emotional (happy or fearful) or the neutral face. A first group of participants was tested on stimuli for which the luminance contrast across faces was not equalised (NonEQ condition) whereas a second group was tested on stimuli for which this luminance contrast was equalised (using a root-mean-square contrast normalization; EQ condition). We chose this paradigm of opposing two faces because it allowed us to show in our previous study that, for unfiltered images, emotional faces are better detected than neutral faces (Entzmann et al., 2021). Analysing differences between target conditions (the emotional or the neutral face) allows us to investigate the general advantage in processing emotional faces, and analysing differences between emotion conditions allows us to compare two types of expressions (happy and fearful) against each other. This effect is not systematically found with other paradigms (Devue & Grimshaw, 2017). For example, if we oppose a face and a vehicle and ask participants to make a saccade towards the face, they perform equally whether the face is emotional or neutral (Entzmann et al., 2021).

We expected better performances (higher accuracy, shorter latencies) when participants were asked to saccade toward the emotional compared to the neutral face. In addition, we supposed that, if LSF are sufficient to discriminate facial expressions and are processed faster than HSF, latencies should be shorter for LSF than HSF stimuli. We also expected, when the target was an emotional face, an interaction between the spatial frequency content and the emotional facial expression. More specifically, we expected better performance for happy than fearful face targets for BSF images, as previous studies using unfiltered images found an advantage in the detection of happy faces (Calvo & Nummenmaa, 2009, 2011; Entzmann et al., 2021). However, for LSF images, we expected the opposite (i.e., an advantage in the detection of fearful faces), as LSF may trigger a fast brain response that would facilitate the detection of fearful faces. We also expected performance to be modulated according to the contrast condition (NonEQ vs. EQ group), in the form of an interaction between the spatial frequency content and the contrast group. More precisely, we expected that if the fast detection of LSF fearful faces is also explained by their higher contrast, the better detection of LSF over HSF faces should be enhanced in the NonEQ group. However, for the EQ group, the difference between LSF and HSF faces should be attenuated, or could even be inverted toward an advantage of HSF. We also expected, when the target was an emotional face, an interaction between the spatial frequency content, the contrast group and the emotional facial expression, as we supposed that the better detection of fearful than happy faces in LSF would be reduced or disappear in the EQ group.

In the second part of this paper, we present a saliency analysis whose aim was to better understand the results obtained with the saccadic choice task. Specifically, we computed the saliency maps of our face stimuli (Borji & Itti, 2013; Foulsham & Underwood, 2008; Itti & Koch, 2000; Marat et al., 2009) and tested the hypothesis that the conditions in which the mouth or the eyes of the target are the most salient, are the conditions in which the task was the easiest.

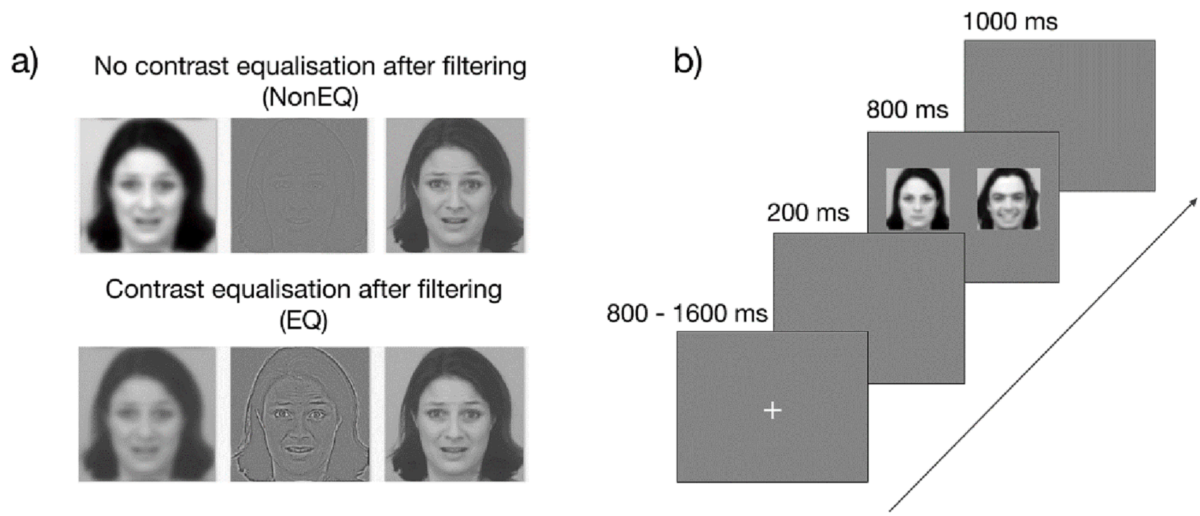

Fig. 1. (a) Example of an image in the different contrast and spatial frequency conditions: In the first raw, the non-equalised (NonEQ) contrast condition with LSF, HSF, and BSF images. In the second raw, the contrast equalised (EQ) condition with LSF, HSF, and BSF images. In the EQ condition, all images have a mean RMS contrast of 46, for pixel intensity values comprised between [0,255]. In the NonEQ condition, LSF, HSF and BSF images have a mean RMS contrast of 68, 8, and 46, respectively. Mean luminance was set to 125 and was the same in all conditions. (b) Time course of a trial: A central fixation cross is displayed during 800 to 1600 ms, followed by a 20-ms gap. Then two images, an emotional and a neutral face are displayed for 800 ms, followed by a 1000 ms inter-stimulus interval.

2. Saccadic choice task

2.1. Materials and method

2.1.1. Participants

Eighty-one participants recruited from Grenoble Alpes University took part in the experiment. Based on our previous study (Entzmann et al., 2021), we were able to estimate a sample size that is sufficient enough to find the effect of the target (the emotional or the neutral face) on saccadic responses. Using the G*Power software (Faul et al., 2007) with a power of 0.95 at the standard 0.05 alpha error probability the estimated sample size was 16 (the expected effect size was a 0.87 Cohen's d, taken from the accuracy analysis in our previous study). However, in the present study, we introduced new variables, with manipulation of the spatial frequency content and contrast of our stimuli. We did not find any study that allows a precise estimate of the strength of the expected effects (i.e., a study with facial expressions, saccadic responses, and spatial frequencies or contrast). Therefore, we choose to consider a larger sample size (approximately 40 participants per group) to increase the statistical power. Three participants were removed from statistical analysis due to a low proportion of correct responses (below 50% in each session), or a high proportion of invalid trials (above 50%), leading to a group of seventy-eight participants included in our data analysis (39 females; mean age ± SD: 21.39 ± 0.98 years; age range: 18–33 years). They all had a normal or corrected-to-normal vision and gave their informed written consent before the experiment. Ethical approval was not required for this study in accordance with the local legislation and institutional requirements (confirmation that the study did not require ethical approval was obtained from the ethics committee of the University Grenoble Alpes, CER-Grenoble Alpes, COMUE University Grenoble Alpes, IRB00010290). However, we followed the Code of Ethics of the World Medical Association (Declaration of Helsinki) for experiments involving humans.

2.1.2. Stimuli

Stimuli consisted of 60 grayscale face photographs portraying 20 different individuals (10 women) with 3 emotions (happy, fearful or neutral); images had a resolution of 300 × 300 pixels and were chosen among the Karolinska Directed Emotional Faces database (Lundqvist et al., 1998). The mean luminance (i.e., mean pixel intensity), as well as the root mean squared contrast (RMS contrast; corresponding to the standard deviation of pixel intensity), were equalised among all the images prior to any further manipulation to obtain a mean luminance of 125 and a mean RMS contrast of 46 (for pixel intensity values comprised between [0,255]; such values corresponded to the mean luminance and contrast values of all the stimuli). Each image was then filtered with a low- or a high-pass filter with cutoff frequencies respectively set to 1 cycle per degree (11 cycles per image, 7 cycles per face for LSF) and 6 cycles per degree (66 cycles per image, 40 cycles per face for HSF). These values were chosen to match a previous study (Guyader et al., 2017) that measured the influence of spatial frequencies on fast saccades toward faces using a saccadic choice task.[1] Therefore, each image was viewed under three different spatial frequency conditions: a HSF condition, a LSF condition and a BSF condition. Because the amplitude spectrum of natural images decreases as spatial frequency increases (Field, 1987; Kauffmann, Chauvin, et al., 2015; Van der Schaaf & van Hateren, 1996), LSF images have a higher RMS contrast than HSF images. Hence, to dissociate the respective contributions of spatial frequency and luminance contrast, two sets of stimuli were built. In one stimulus dataset, a RMS contrast equalisation was applied after the filtering process (EQ condition), whereas in the other stimulus dataset, there was no contrast equalisation after filtering (NonEQ condition). In the EQ dataset, all images, filtered and unfiltered, had the same mean luminance contrast (set to the value of 46, which corresponds to the contrast of BSF images), whereas in the NonEQ dataset the mean luminance contrast of HSF, LSF, and BSF images was left unchanged (mean RMS contrast for HSF images: 8; mean RMS contrast for LSF images: 69; mean RMS contrast for BSF images: 46; all values comprised between [0,255]; see Fig. 1a for examples of stimuli).

2.1.3. Materials

Stimuli were displayed on a 24-inch screen with a spatial resolution of 1360 × 768 pixels and a refresh rate of 60 Hz. Eye movements were

[1] It should be noted, however, that selecting an appropriate cutoff frequency is not straightforward, as it varies across studies. For instance, Goffaux and Rossion (2006) and Kumar and Srinivasan (2011) used cutoffs below 8 and above 32 cycles per face, Vlamings et al. (2009) used cutoffs below 12 and above 36 cycles per face, and McFadyen et al. (2017), Méndez-Bértolo et al. (2016), and Vuilleumier et al. (2003) used cutoffs below 6 and above 24 cycles per face.

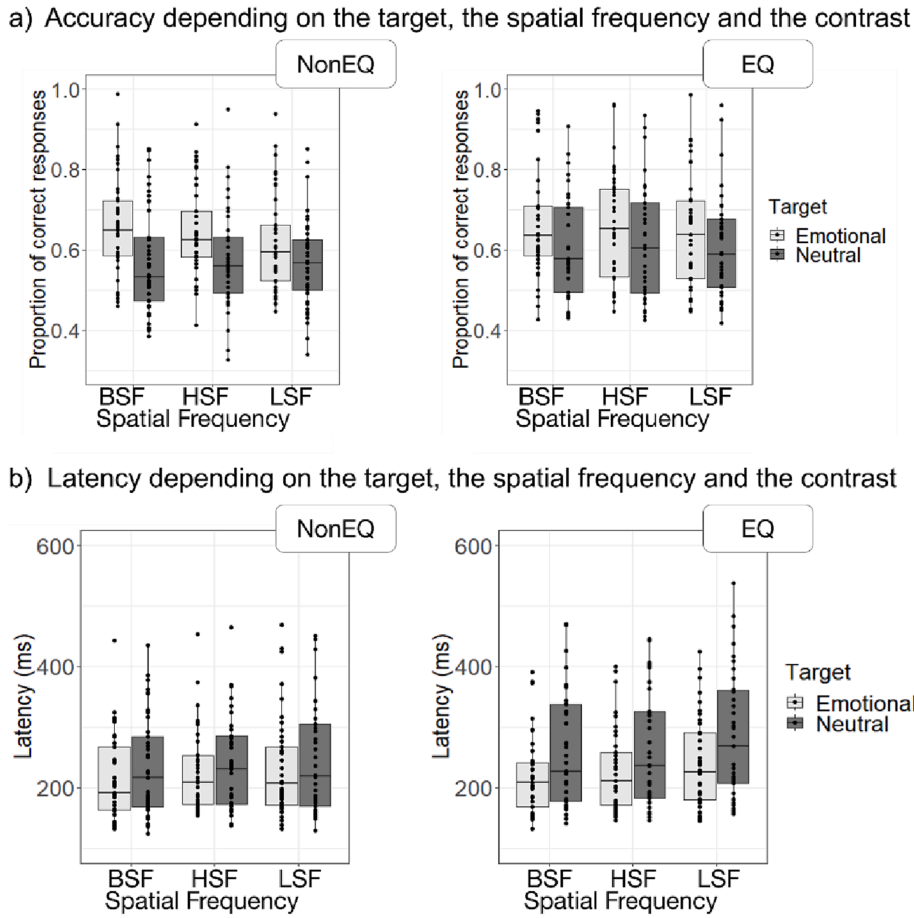

Fig. 2. Boxplots for (a) mean accuracy (proportion of correct responses) and (b) mean latency of correct saccadic responses (in ms) for emotional (black) and neutral (gray) targets, for the three spatial frequency conditions (BSF, HSF, and LSF) and the two contrast conditions (NonEQ and EQ).

recorded with an Eyelink 1000 Plus (SR Research) eye-tracker with a 1000 Hz sampling frequency. Viewing was binocular, but only the position of the dominant eye was recorded (the left eye for 19 participants, and the right eye for 59 participants). Saccades were automatically detected by the Eyelink software based on a minimum velocity of 30°/s, a minimum acceleration of 8000°/s2 and a minimum motion of 0.15°. Blinks were detected when the pupil was partially or totally occluded, and fixations were detected when there was no blink and no saccade in progress.

2.1.4. Procedure

The experiment was divided into two sessions of 240 trials each, whose order was counterbalanced between participants. In one session, the target stimulus was the emotional face (happy or fearful; the distractor was the neutral face) while in the other session, the target stimulus was the neutral face (the distractor was the emotional face, which was either happy or fearful). Each session was divided into blocks of 80 trials each, corresponding to the different spatial frequency conditions (HSF, LSF, and BSF) presented in a randomised order. In each block, each of the 40 different emotional faces was displayed, in a random order, once on the left and once on the right side of the screen. Each emotional face was paired with a neutral face from the opposite gender, therefore, there was never a trial with a neutral and an emotional face from the same individual. Since there were twice as many emotional faces as neutral faces, the neutral faces were each displayed twice on each side. Once with a happy emotional face and once with a fearful emotional face. As the pairs are chosen in a semi-random way, they are not the same in each block and for each participant. Out of the 800 possible combinations (40 emotional faces × 10 neutral faces from the opposite gender × 2 sides), we selected 80 pairs for each block. One group of participants (n = 37) performed the task with contrast equalisation of stimuli whereas another group (n = 41) performed the task without contrast equalisation. A calibration was performed at the beginning of each session and in each session, at the beginning of each block (every 80 trials). A drift correction was applied every ten trials (if the drift was larger than 1° a new calibration was done). During the calibration, participants were asked to gaze at 9 white dots appearing sequentially in a 3 × 3 grid covering the entire screen. Matlab (MathWorks, Natick, MA) and the Psychophysics Toolbox (Brainard, 1997) were used to control timing and stimulus display, as well as communication with the eye-tracker.

During the experiment, participants were seated in a semi-lighted room, with their head stabilized by a forehead-rest and a chin-rest at a fixed distance of 57 cm away from the screen in order to respect a stimulus size of 11 × 11 degrees of visual angle. On the images, faces sized approximately 7 degrees of visual angle (excluding hair). A session lasted approximately 20 min and the whole experimental procedure took approximately 50 min. For each trial, participants were asked to fixate a white cross presented during a pseudo-random time interval ranging between 800 and 1600 ms. After a 200 ms gap, two images (an emotional and a neutral face) were simultaneously displayed (one on each side of the screen) during 800 ms. Participants were asked to make a saccade as fast as possible toward the target image. The target was given to participants at the beginning of each session: the emotional or the neutral face. The center of each image was located at a fixed distance of 8° of eccentricity from the center of the screen. A trial ended with the presentation of a gray background during 1000 ms (Fig. 1b). Twelve practice trials (including different images than those used in the

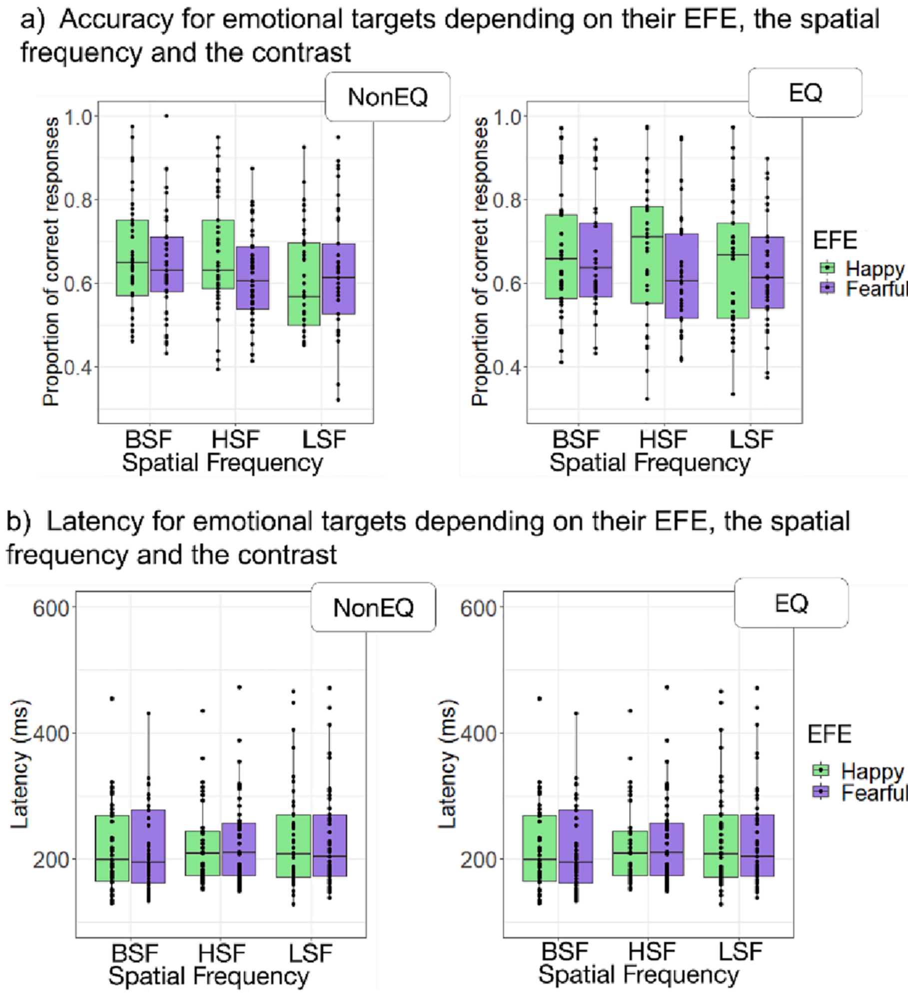

Fig. 3. Boxplots for (a) mean accuracy and (b) mean latency for emotional targets, with a fearful (purple) or happy (green) face, for the three spatial frequency conditions (BSF, HSF, and LSF) and the two contrast conditions (NonEQ and EQ). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

experiment) were set up in a training session at the beginning of the experiment to allow participants to be familiar with the task.

2.1.5. Data analysis

Preprocessing: A preprocessing was applied to eye movement data in order to eliminate non-valid trials. Valid trials were selected according to the following criteria. First, a saccade should be the first event after stimulus onset (i.e., there is no blink before the response). Second, this first saccade should have a latency greater than 50 ms (to remove anticipatory saccades), a starting point within a radius of 2° around the center of the screen, and a duration smaller than 100 ms. Moreover, the amplitude of the first saccade should be greater than 1° and should not go beyond the screen. This preprocessing led to a rejection of 13% of the initial number of trials. The first saccade was considered "correct" if it was directed toward the side of the display containing the target and "error" if it was in the opposite direction (i.e. directed toward the distractor).

Statistical analyses: Statistical analyses were carried out using the open-source software R with R Studio (Racine, 2012). Mean accuracy (in proportion of correct saccades) and mean latency of correct saccadic responses (in ms) were computed for each participant in each experimental condition and were analysed as dependent variables. First, we studied accuracy and latency differences according to the target, the spatial frequency and the contrast (i.e., we tested main effects and interactions between those factors). A mixed analysis of variance (ANOVA) was used, with the Target (Emotional, Neutral) and the Spatial Frequency (BSF, HSF, LSF) as within-subject factors, and the Contrast (EQ, NonEQ) as a between-subject factor.

Then, two other analyses were carried out, to test for the influence of the Emotional facial expression (EFE) of the target in one analysis, or the influence of the EFE of the distractor in another analysis. The first analysis was applied on trials in which the target was an emotional face (and the distractor was a neutral face). This analysis was set up to study the influence of the EFE of emotional face targets (i.e., to test for the main effect of the EFE, and interactions with spatial frequency and contrast). A mixed ANOVA was used, with the EFE of the target (Happy, Fearful) and the Spatial Frequency (BSF, HSF, LSF) as within-subject factors, and the Contrast (EQ, NonEQ) as a between-subject factor. The second analysis was then applied on trials in which the target was a neutral face (and the distractor was an emotional face). In this case, the EFE of the distractor (Happy, Fearful) and the Spatial Frequency (BSF, HSF, LSF) were defined as within-subject factors, and the Contrast (EQ, NonEQ) as a between-subject factor. An effect was considered significant if its p-value was below the threshold $\alpha = 0.05$, and effect sizes were estimated by calculating partial eta-squared (η_p^2). T-tests with Bonferroni corrections were used for pairwise comparisons. Detailed results with corresponding ANOVA tables are presented in Appendix A.

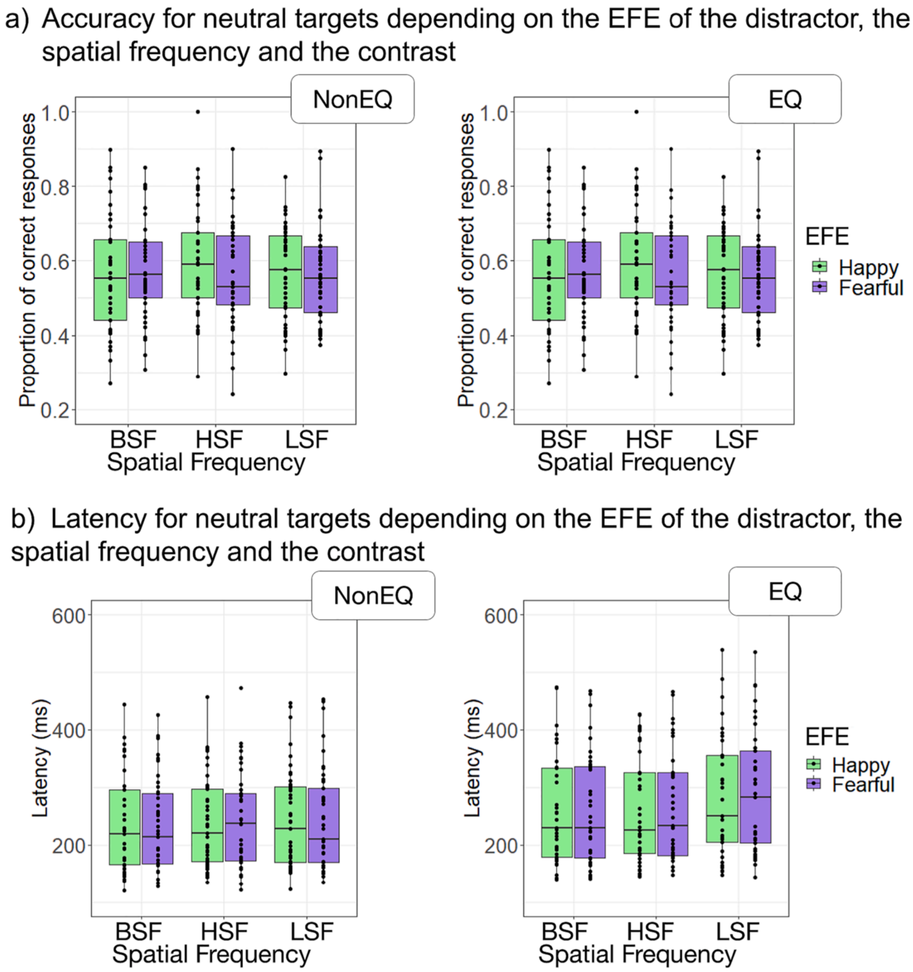

Fig. 4. Boxplots for (a) mean accuracy and (b) mean latency for neutral targets, with a fearful (purple) or happy (green) distractor, for the three spatial frequency conditions (BSF, HSF and BSF) and two contrast conditions (EQ and NonEQ). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.2. Results

2.2.1. Accuracy

Mixed ANOVA performed on mean accuracy (Fig. 2a) indicated that the Target, Spatial Frequency and Contrast interaction was not significant. The interactions between Target and Contrast and Spatial Frequency and Contrast were also non-significant and the interaction between Spatial Frequency and Target was marginally significant ($F(2,152) = 3.1$, $p = .05$, $\eta_p^2 = 0.04$). Pairwise comparisons showed that the accuracy was significantly higher for emotional than neutral faces in every spatial frequency condition. But, when the target was neutral there was no significant effect of spatial frequencies, whereas there was a difference between BSF ($M \pm SD$: 0.66 ± 0.13) and LSF ($M \pm SD$: 0.60 ± 0.13; $p_{corrected} = 0.009$) images when the target was emotional.

Overall, a significant effect of the Target ($F(1,76) = 67.1$, $p < .001$, $\eta_p^2 = 0.47$), and the Spatial Frequency ($F(2,152) = 3.89$, $p = .023$, $\eta_p^2 = 0.05$) were observed. The accuracy was higher when the target was emotional ($M \pm SD$: 0.65 ± 0.12) than neutral ($M \pm SD$: 0.59 ± 0.12). Also, it was higher for both BSF ($M \pm SD$: 0.62 ± 0.12; $p = .013$) and HSF ($M \pm SD$: 0.62 ± 0.12; $p = 0.016$) than LSF ($M \pm SD$: 0.61 ± 0.12) images. The main effect of Contrast was not significant.

2.2.2. Latency

Mixed ANOVA performed on mean latency (Fig. 2b) of correct saccadic responses indicated that the Target, Spatial Frequency and Contrast interaction was not significant. The interaction between Target and Spatial Frequency was also non-significant, and the interaction between Spatial Frequency and Contrast was only marginally significant ($F(1,152) = 2.6$, $p = .073$, $\eta_p^2 = 0.034$). Pairwise comparisons showed that latencies were shorter for HSF than LSF only in the EQ contrast condition (*M* ± *SD for HSF*: 244.1 ± 81.1 ms; *M* ± *SD for LSF*: 264 ± 87.3 ms; $p_{corrected} < 0.001$). In both contrast condition, latencies were shorter for BSF than LSF images, although it was only marginally significant in the NEQ condition (*M* ± *SD for LSF NEQ*: 237.9 ± 81.9 ms; *M* ± *SD for BSF EQ*: 241.1 ± 76.3 ms; $p_{corrected} < 0.001$ *in the EQ condition* and $p_{corrected} = 0.07$ *in the NEQ condition*). The ANOVA also revealed a significant interaction between Target and Contrast ($F(1,76) = 8.2$, $p = .005$, $\eta_p^2 = 0.1$). Pairwise comparisons showed that latencies were significantly shorter for emotional compared to neutral targets in the EQ condition (*M* ± *SD for emotional targets*: 229.5 ± 67.5 ms; *M* ± *SD for neutral targets*: 269.9 ± 96 ms; $p_{corrected} < 0.001$) but this effect was only marginally significant in the NEQ condition (*M* ± *SD for emotional targets*: 224.2 ± 71.2 ms; *M* ± *SD for neutral targets*: 238.5 ± 80.2 ms; $p_{corrected} = 0.097$) contrast condition.

Overall, a significant effect of the Target ($F(1,76) = 35.8$, $p < .001$, $\eta_p^2 = 0.31$) and the Spatial Frequency ($F(2,152) = 18.9$, $p < .001$, $\eta_p^2 = 0.2$) were observed. Saccades were elicited faster when the target was emotional ($M \pm SD$: 227 ± 69 ms) than neutral ($M \pm SD$: 253 ± 89 ms). Latencies were also shorter for BSF ($M \pm SD$: 232 ± 74 ms) than HSF ($M \pm SD$: 238 ± 76 ms; $p = .01$) or LSF ($M \pm SD$: 250 ± 85 ms; $p < .001$) images, and for HSF than LSF images ($p < .001$).

2.2.3. Analysis as a function of the EFE of emotional face targets

Accuracy: For emotional face targets, mixed ANOVA performed on mean accuracy (Fig. 3a) indicated that the EFE, Spatial Frequency and Contrast interaction was not significant. The interactions between EFE and Contrast as well as Spatial Frequency and Contrast were also non-significant. A significant interaction between Spatial Frequency and EFE ($F(2,152) = 7.2$, $p = .001$, $\eta_p^2 = 0.087$) was observed. Pairwise comparisons showed that a better accuracy for happy than fearful targets was only significant when images were filtered in HSF (*M ± SD for happy faces*: 0.68 ± 0.15; *M ± SD for fearful faces*: 0.62 ± 0.12; $p < .001$). Furthermore, accuracy for happy targets was higher when images were in HSF ($p_{corrected} = 0.001$) and BSF (*M ± SD*: 0.67 ± 0.14; $p_{corrected} = 0.011$) compared to LSF (*M ± SD*: 0.62 ± 0.14) images, whereas for fearful target accuracy was marginally higher for BSF (*M ± SD*: 0.66 ± 0.14) than HSF ($p_{corrected} = 0.08$) images. Overall, the ANOVA indicated a significant main effect of the EFE ($F(1,76) = 8.4$, $p = .004$, $\eta_p^2 = 0.1$), and the Spatial Frequency ($F(2,152) = 5.8$, $p = .003$, $\eta_p^2 = 0.07$). Performances were better when the target was happy (*M ± SD*: 0.66 ± 0.13) than fearful (*M ± SD*: 0.64 ± 0.12). They were also better for both BSF (*M ± SD*: 0.66 ± 0.13; $p < .001$) and HSF (*M ± SD*: 0.65 ± 0.13; $p = 0.036$) than LSF (*M ± SD*: 0.63 ± 0.13) images.

Latency: For emotional face targets, mixed ANOVA performed on mean latency of correct responses (Fig. 3b) indicated that all interactions were not significant. The main effects of Target and EFE were also not significant. Only a significant main effect of the Spatial Frequency ($F(2,152) = 12.26$, $p < .001$, $\eta_p^2 = 0.14$) was found. Latencies were shorter for BSF (*M ± SD*: 217 ± 67 ms) than HSF (*M ± SD*: 226 ± 68.4 ms; $p = .007$) or LSF (*M ± SD*: 236 ± 79.4 ms; $p < .001$) images, as well as for HSF than LSF ($p < .001$) images.

2.2.4. Analysis as a function of the EFE of the distractor for neutral face targets

Accuracy: For neutral face targets, mixed ANOVA performed on mean accuracy (Fig. 4a) indicated that the EFE of the distractor, Spatial Frequency and Contrast interaction was not significant. The interactions between EFE and Contrast as well as Spatial Frequency and Contrast were not significant. A significant interaction between the Spatial Frequency and the EFE of the distractor ($F(2,152) = 4.3$, $p = .016$, $\eta_p^2 = 0.053$) was observed. Pairwise comparisons showed a better accuracy with happy distractors only in HSF (*M ± SD for happy distractors:* 0.62 ± 0.15; *M ± SD for fearful distractors:* 0.57 ± 0.14; $p_{corrected} = 0.022$). Overall, a marginal effect of the EFE of the distractor ($F(1,76) = 3.8$, $p = .05$, $\eta_p^2 = 0.047$) was observed. The accuracy was marginally higher when the distractor was happy (*M ± SD:* 0.60 ± 0.13) than fearful (*M ± SD:* 0.58 ± 0.12).

Latency: For neutral face targets, mixed ANOVA performed on mean latency (Fig. 4.b) indicated that the EFE of the distractor, Spatial Frequency and Contrast interaction was not significant. The interactions between EFE and Contrast and Spatial Frequency and EFE were also non-significant. A significant interaction between the Spatial Frequency and the Contrast ($F(2,152) = 3.1$, $p = .047$, $\eta_p^2 = 0.04$) was observed. Pairwise comparisons showed shorter latencies for HSF and BSF than LSF only in the EQ contrast group (*M ± SD for HSF:* 262.2 ± 95.1; *M ± SD for BSF:* 261.2 ± 94.1; *M ± SD for LSF:* 286;2 ± 105; $p_{corrected} < 0.001$ for HSF and $p_{corrected} = 0.003$ for BSF). A significant effect of the Spatial Frequency ($F(2,152) = 10.2$, $p < .001$, $\eta_p^2 = 0.11$), and a marginal effect of the EFE of the distractor ($F(1,76) = 3.71$, $p = .057$, $\eta_p^2 = 0.04$). Latencies were shorter for images in BSF (*M ± SD:* 246.6 ± 88.5 ms) and HSF (*M ± SD:* 249.8 ± 87.6 ms) compared to LSF (*M ± SD:* 263.6 ± 97.9 ms; p <.001). Latencies were marginally shorter when the distractor was a happy (*M ± SD:* 252.1 ± 88.1 ms) than a fearful (*M ± SD:* 254.5 ± 90.24 ms) face.

3. Analysis of the eye and the mouth saliency

The purpose of this analysis was to see if local statistical differences in the saliency of the mouth or the eyes, which are usually considered diagnostic for expression decoding (Calvo & Nummenmaa, 2011; Eisenbarth & Alpers, 2011; Smith & Schyns, 2009; Sweeny et al., 2013; Wegrzyn et al., 2017), could explain participants' behaviour in the saccadic choice task. Calvo and Nummenmaa (2011) suggested in their study of happy face detection, that emotion detection may arise from a two-stage processing mechanism. The first stage would be purely perceptual (i.e., only relying on the physical attribute of the face), with an analysis of visually salient regions. In the second stage, the detection of salient features would be used for expression recognition and semantic retrieval. Visual saliency would therefore be important because it would be linked to the efficiency of the decoding. Here, we expected that the more salient the mouth or eye of the target face is, the better the performances were in the saccadic choice task. Therefore, we hypothesised that saliency might have played a role in the detection of emotional faces and the use of spatial frequencies. Specifically, the mouth of emotional faces may be more salient than that of neutral faces, especially with a happy face, which would capture attention and enhance decoding efficiency. Moreover, mouth saliency may be greater in HSF than LSF, particularly in the case of a happy face, which could account for the higher performance observed.

Visual saliency can be viewed as the intensity with which a region will attract attention, independently of task demands, and several computational models were proposed to compute saliency and made behavioural predictions (Borji & Itti, 2013; Foulsham & Underwood, 2008; Itti & Koch, 2000; Marat et al., 2009). The calculation of saliency differs depending on the model, but overall, a specific region is likely to be salient when it can be easily distinguished from its neighborhood, based on low-level visual attributes (Koehler et al., 2014). In their study, Calvo and Nummenmaa (2011) studied the role of perceptual (e.g., luminance, mouth or eye saliency differences), as well as higher level factors (e.g., valence) on the detection of happy faces in a saccadic choice task. They only found a contribution of the mouth saliency. The target was always a happy face, and they studied the difference between the mouth or eye saliency of the target and that of the distractor (a neutral, sad, angry, fearful, disgusted or surprised face). Here, we expected that the more salient the mouth or eye of the target face is, the better the performances were in the saccadic choice task. Indeed, even if Calvo and Numenmaa (2011) found no contribution of the eye saliency in their happy face detection task, we still analysed it, as it may still contribute to the detection of emotional or neutral faces. In fact, the eyes were shown to be an important feature for expression decoding (Eisenbarth & Alpers, 2011; Wegrzyn et al., 2017), especially with fearful faces, which were used in our study (Smith & Schyns, 2009).

3.1. Method

3.1.1. Computation of the mouth and the eye saliency

The eye and mouth saliency computations were performed in two steps. First, saliency maps were generated (i.e., on happy, neutral, or fearful faces, each in BSF, HSF, or LSF, with or without contrast equalisation), then rectangular boxes encompassing the eyes and mouth were used to compute the average saliency of the eyes and mouth for each saliency map. Maps were computed for each image using the computational saliency model proposed by Walther and Koch (2006). This model is inspired by the biology of the human visual system and considers intensities and orientations to attribute a saliency value to each pixel of an input image. The model was implemented in Matlab using the SaliencyToolbox (https://www.saliencytoolbox.net) and the *makeSaliencyMap* function. Note that we have changed the default settings for calculating the maps. First, the center-surround parameters were lowered to adapt to our stimuli resolution and allow the identification of more localised regions. Then, the maps were generated after a single iteration, as we were not interested in any inhibition of return mechanism (see Appendix B for an overview of the used parameters). At the end of this process, a saliency map was obtained for each image. Fig. 5a

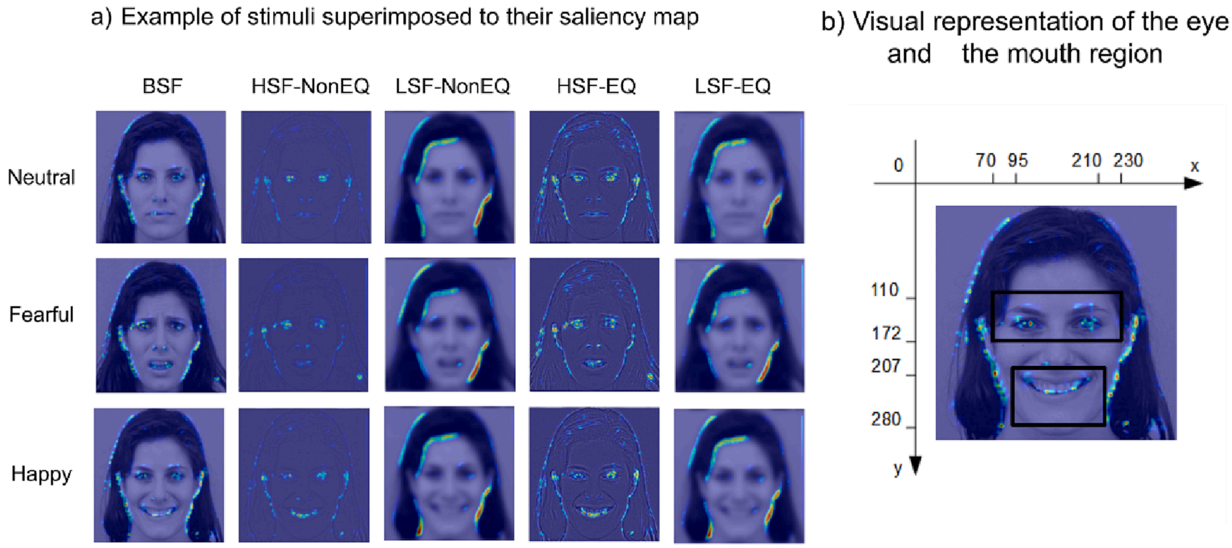

Fig. 5. (a) Example of saliency maps computed for one face viewed in its different spatial frequency and contrast conditions, with a neutral, happy or fearful expression with the model from Walther and Koch (2006). Experimental conditions are (from left to right) BSF, HSF NonEQ, LSF NonEQ, HSF EQ and LSF EQ. (b) Visual representation of the eye and the mouth regions added on a saliency map. These boxes were used to calculate the eye and mouth saliency, by taking the average saliency value of all the pixels included in a box. For illustration purposes, we have displayed the saliency maps overlaid on the stimuli from which they were computed.

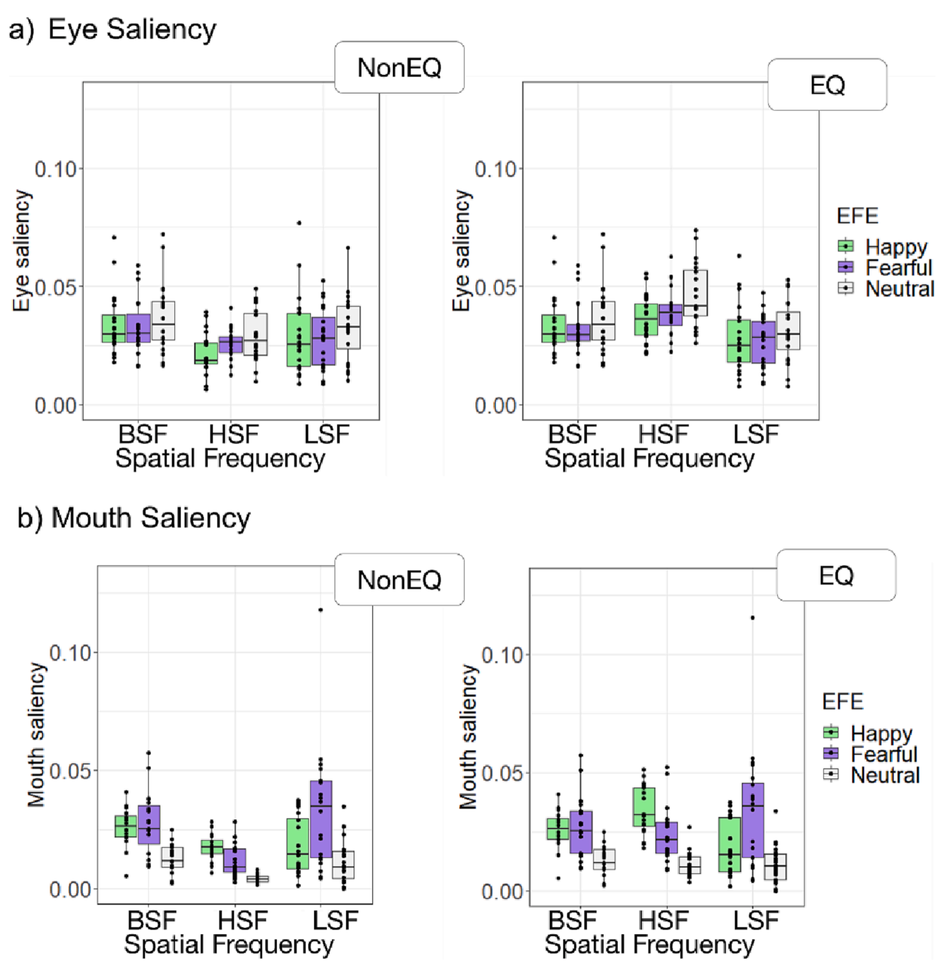

Fig. 6. Boxplots for mean (a) eye or (b) mouth saliency, for happy (green), fearful (purple) or neutral (gray) faces, for the three spatial frequency conditions (BSF, HSF, and LSF) and the two contrast conditions (NonEQ and EQ). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

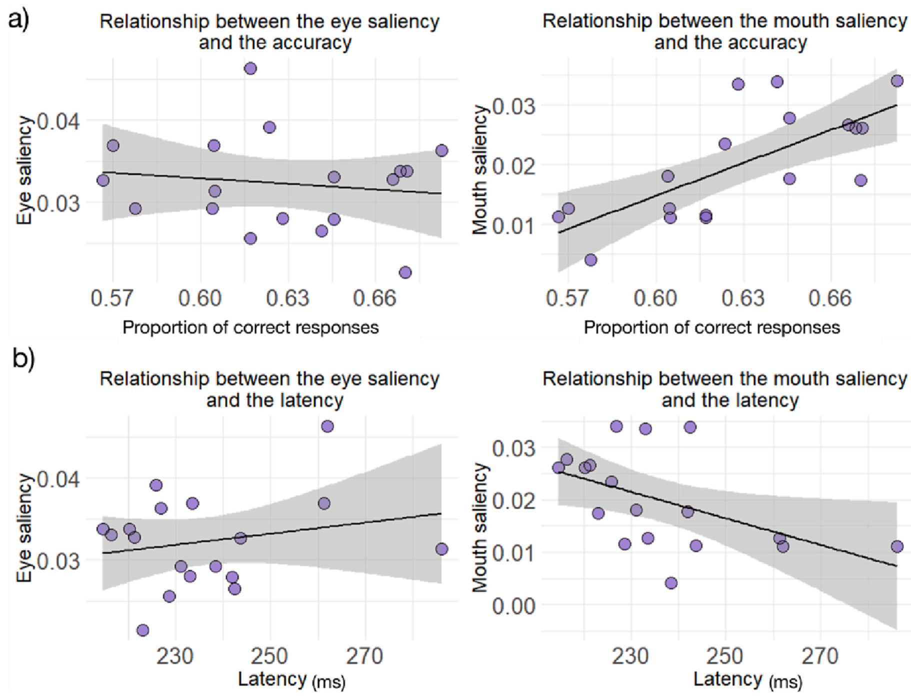

Fig. 7. Relationship between (a) the accuracy or (b) the latency and the eye (left) or mouth (right) saliency. One dot corresponds to one condition of the saliency analysis (a happy, fearful, or neutral target, each in BSF, HSF, or LSF, with or without contrast equalisation).

shows examples of saliency maps superimposed on their corresponding stimuli (i.e., a specific face in the different spatial frequency, contrast and emotion conditions).

The mouth and eye regions were then visually selected so that we obtained rectangular boxes that contain the mouth or the eyes (including the eyebrows) for all our individuals, with any expression. More precisely, the mouth region corresponded to pixels with an X-coordinate between 95 and 210, and a Y-coordinate between 207 and 280 (for images that sized 300x300 pixels; X_0 and Y_0 coordinates being the top-left corner). The eye region corresponded to pixels with an X-coordinate between 70 and 237, and a Y-coordinate between 110 and 172. Then the mouth and eye regions were the same for all the stimuli. Afterwards, we computed for each image the mean saliency value, by averaging all the saliency values of all the pixels contained in the box surrounding the mouth or the eyes. The mouth and eye region can be visualised in Fig. 5b.

3.1.2. Data analysis

First, a factorial 3x3x2 ANOVA with the EFE (Fearful, Neutral, Happy), the Spatial Frequency (BSF, HSF, LSF) and the Contrast (NonEQ, EQ) as between-image factors was applied on the mean saliency of each image, once for the eyes and once for the mouth. Note that BSF images are the same in both contrast conditions, as contrast equalisation only modulates the contrast of HSF and LSF images. We still used t-tests with Bonferroni corrections for pairwise comparisons.

Then, we also performed Pearson's correlations to evaluate the strength of the relationship between mean accuracy and latency in each condition in the saccadic choice task (independently of participants), and the mean mouth or eye saliency in the same condition (when the target was a happy, neutral or fearful face, each in BSF, HSF, or LSF, in the NonEQ or EQ contrast group). The goal was to determine if the conditions for which the mouth or the eyes are, on average, the most salient, were the conditions in which the performance was, on average, the highest. Therefore, for each correlation, we studied the relationship between 18 saliency values and 18 performance values (1 value for each condition: when the target was a happy, neutral or fearful face, each in BSF, HSF, or LSF, in the NonEQ or EQ contrast group).

3.2. Results

3.2.1. Eye saliency

The ANOVA performed on mean eye saliency (Fig. 6a) indicated that the EFE, Spatial Frequency and Contrast interaction was not significant. The interactions between EFE and Contrast and Spatial Frequency and EFE were also non-significant. There was a significant interaction between the Spatial Frequency and the Contrast ($F(2,342) = 15.9$, $p < .001$, $\eta_p^2 = 0.085$). Pairwise comparison revealed that eye saliency was higher for EQ than NonEQ stimuli filtered in HSF only (*M* ± *SD for the NonEQ condition: 0.025* ± *0.01; M* ± *SD for the EQ condition: 0.041* ± *0.012*). Furthermore, in the NonEQ condition, eye saliency was higher for BSF (*M* ± *SD: 0.034* ± *0.013*) than HSF (*$p_{corrected} < 0.001$)* images, whereas in the EQ condition it was higher for HSF than LSF (*M* ± *SD: 0.029* ± *0.013; $p_{corrected} < 0.001$*) images. The ANOVA also revealed a main effect of the EFE ($F(2, 342) = 6.08$, $p = .002$, $\eta_p^2 = 0.034$), the Spatial Frequency ($F(2, 342) = 5.44$, $p = .005$, $\eta_p^2 = 0.031$), and the Contrast ($F(1, 342) = 11.6$, $p < .001$, $\eta_p^2 = 0.033$). Overall, mean eye saliency was higher for neutral (*M* ± *SD: 0.0355* ± *0.014*) than happy (*M* ± *SD: 0.03* ± *0.014; p =.003*) and fearful (*M* ± *SD: 0.031* ± *0.012; p =.007*) faces. It was also higher for BSF (*M* ± *SD: 0.035* ± *0.013; p =.003*) and HSF (*M* ± *SD: 0.033* ± *0.013; p =.032*) than LSF (*M* ± *SD: 0.029* ± *0.013*) images. Finally, it was overall higher for the EQ (*M* ± *SD: 0.035* ± *0.014)* than NonEQ (*M* ± *SD: 0.03* ± *0.013)* contrast condition.

3.2.2. Mouth saliency

The ANOVA performed on mean mouth saliency (Fig. 6b) indicated that the EFE, Spatial Frequency and Contrast interaction, as well as the EFE and Contrast interaction, were not significant. There was a significant interaction between the Spatial Frequency and the EFE ($F(4,342) = 9.86$, $p < .001$, $\eta_p^2 = 0.1$) and between the Spatial Frequency and the Contrast ($F(2,342) = 9.8$, $p < .001$, $\eta_p^2 = 0.054$). For the interaction between the Spatial Frequency and the EFE, pairwise comparison showed that, for LSF images mouth saliency was higher for fearful (*M* ± *SD: 0.034* ± *0.026)* than happy (*M* ± *SD: 0.018* ± *0.012)* or neutral (*M* ± *SD: 0.011* ± *0.008; $p_{corrected} < 0.001$*) faces. For HSF images, mouth

Table A1
Results of mixed ANOVA performed on saccade accuracy and latency with the Target: (Emotional, Neutral) and the Spatial Frequency (BSF, HSF, LSF) as within-subject factors, and the Contrast (EQ, NonEQ) as a between-subject factor.

			Accuracy			Latency		
	Df	Df	F	η_p^2	p	F	η_p^2	p
Contrast	1	76	1.02	0.013	0.32	1.11	0.014	0.29
Spatial Frequency	2	152	3.89	0.049	0.022*	18.9	0.199	<0.001***
Target	1	76	67.1	0.47	<0.001***	35.8	0.320	<0.001***
Contrast × Spatial Frequency	2	152	0.43	0.006	0.65	2.66	0.033	0.072.
Contrast × Target	1	76	2.51	0.032	0.11	8.18	0.097	0.005**
Target × Spatial Frequency	2	152	3.05	0.039	0.05.	0.48	0.006	0.61
Contrast × Target × Spatial Frequency	2	152	0.34	0.004	0.72	0.66	0.008	0.51

Table A2
Results of mixed ANOVA performed on saccade accuracy and latency for trials for which the target was the emotional face. The EFE of the target (Happy, Fearful) and the Spatial Frequency (BSF, HSF, LSF) were defined as within-subject factors, and the Contrast (EQ, NonEQ) as a between-subject factor.

			Accuracy			Latency		
	Df	Df	F	η_p^2	p	F	η_p^2	p
Contrast	1	76	0.32	0.004	0.57	0.11	0.001	0.73
Spatial Frequency	2	152	5.82	0.071	0.003**	12.2	0.13	<0.001***
EFE	1	76	8.40	0.099	0.004 **	1.71	0.022	0.19
Contrast × Spatial Frequency	2	152	0.59	0.007	0.54	0.76	0.009	0.46
Contrast × EFE	1	76	0.08	0.001	0.77	1.38	0.017	0.24
EFE × Spatial Frequency	2	152	7.14	0.085	0.001 ***	0.11	0.001	0.87
Contrast × EFE × Spatial Frequency	2	152	1.04	0.013	0.351	0.79	0.01	0.44

Table A3
Results of mixed ANOVA performed on saccade accuracy and latency for trials for which the target was a neutral face. The EFE of the distractor (Happy, Fearful) and the Spatial Frequency (BSF, HSF, LSF) were defined as within-subject factors, and the Contrast (EQ, NonEQ) as a between-subject factor.

			Accuracy			Latency		
	Df	Df	F	η_p^2	p	F	η_p^2	p
Contrast	1	76	1.84	0.023	0.17	2.47	0.031	0.11
Spatial Frequency	2	152	1.03	0.013	0.35	10.22	0.11	<0.001***
EFE	1	76	3.81	0.04	0.054.	3.71	0.046	0.057.
Contrast × Spatial Frequency	2	152	0.036	0.0004	0.96	3.10	0.039	0.052.
Contrast × EFE	1	76	0.033	0.0004	0.85	2.53	0.032	0.11
EFE × Spatial Frequency	2	152	4.24	0.052	0.016*	1.31	0.016	0.27
Contrast × EFE × Spatial Frequency	2	152	0.18	0.002	0.82	0.28	0.003	0.74

Table A4
Results of a factorial ANOVA with the EFE (Fearful, Neutral, Happy), the Spatial Frequency (BSF, HSF, LSF) and the Contrast (NonEQ, EQ) as between-image factors, applied on the mean saliency of each image, once for the eyes and once for the mouth.

			Eye saliency			Mouth saliency		
	Df	Df	F	η_p^2	p	F	η_p^2	p
Contrast	1	462	13.7	0.028	<0.001***	13.9	0.029	<0.001***
Spatial Frequency	2	462	6.35	0.026	0.002 **	8.54	0.035	<0.001***
EFE	2	462	8.86	0.036	<0.001***	102.4	0.30	<0.001***
Contrast × Spatial Frequency	2	462	18.3	0.073	<0.001***	14	0.057	<0.001***
Contrast × EFE	2	462	0.1	0.0004	0.89	0.86	0.003	0.42
EFE × Spatial Frequency	4	462	0.86	0.007	0.48	12.2	0.095	<0.001***
Contrast × EFE × Spatial Frequency	4	462	0.095	0.0008	0.98	0.91	0.007	0.45

saliency was higher for happy (*M ± SD: 0.25 ± 0.012;* $p_{corrected} < 0.001$) and fearful (*M ± SD: 0.017 ± 0.011;* $p_{corrected} = 0.017$) than neutral (*M ± SD: 0.007 ± 0.005)* faces. Similarly, for BSF images, it was higher for happy (*M ± SD: 0.26 ± 0.008)* and fearful (*M ± SD: 0.027 ± 0.013)* than neutral (*M ± SD: 0.013 ± 0.006;* $p_{corrected} < 0.001$) faces. For happy and neutral faces, there was no difference between frequency conditions. For fearful faces, mouth saliency was higher for LSF ($p_{corrected} < 0.001$) and BSF ($p_{corrected} = 0.024$) than HSF images. Concerning the interaction between the Contrast and the Spatial Frequency, the mouth saliency was higher for the EQ than NonEQ condition for HSF filtered faces only (*M ± SD for the NonEQ condition: 0.011 ± 0.007; M ± SD for the EQ condition: 0.023 ± 0.013;* $p_{corrected} = 0.003$). In the NonEQ condition, mouth saliency was higher for BSF (*M ± SD: 0.022 ± 0.011;* $p_{corrected} < 0.001$) and LSF (*M ± SD: 0.021 ± 0.02;* $p_{corrected} = 0.003$) than HSF images, whereas there was no difference between spatial frequency conditions in the EQ group.

The ANOVA revealed a significant main effect of the EFE ($F(2, 342) = 55.9$, $p < .001$, $\eta_p^2 = 0.25$), the Spatial Frequency ($F(2, 342) = 5.77$, $p = .003$, $\eta_p^2 = 0.032$), and the Contrast ($F(1, 342) = 8.81$, $p = .003$, $\eta_p^2 = 0.025$). Overall, mouth saliency was higher for happy (*M ± SD: 0.023 ±*

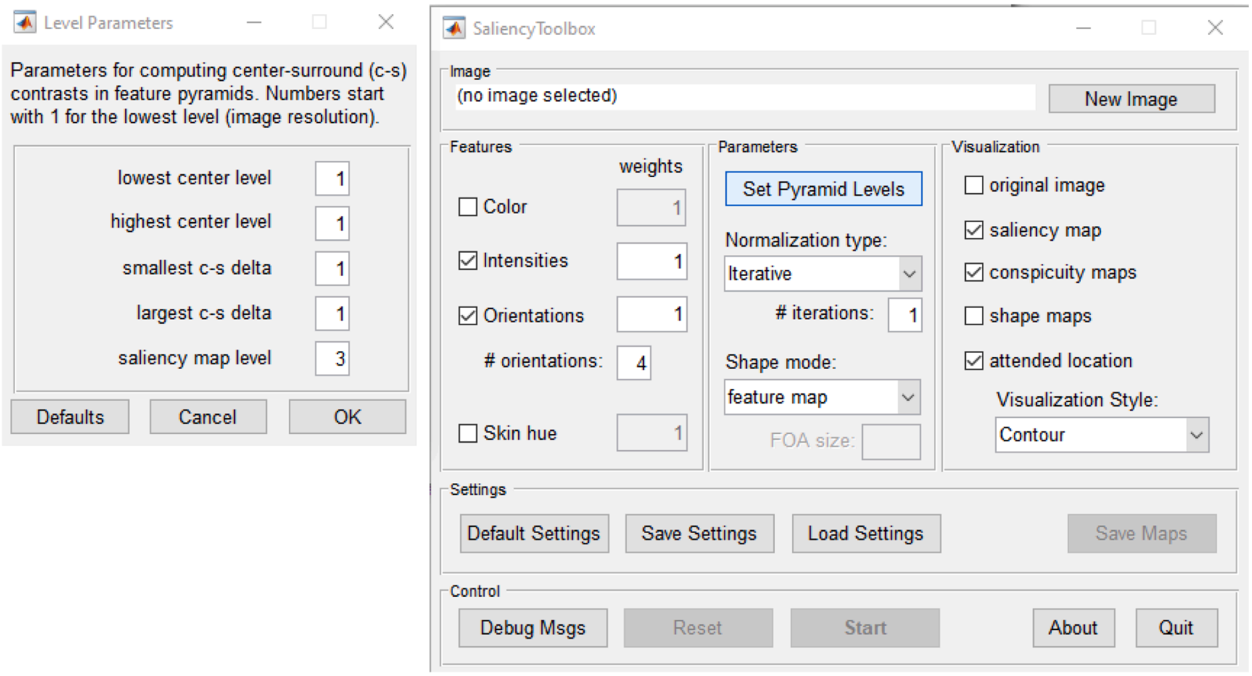

Fig. A1. Overview of the Saliency toolbox interface (Walther and Koch, 2006) with the parameters used to generate the saliency maps in this study.

Model Comparison

Models	P(M)	P(M|data)	BF_M	BF_{10}	error %
SF + Target + Contrast + SF ✻ Contrast + Target ✻ Contrast	0.053	0.435	13.866	1.000	
SF + Target + Contrast + Target ✻ Contrast	0.053	0.306	7.918	0.702	6.801
SF + Target	0.053	0.055	1.043	0.126	6.279
SF + Target + Contrast + SF ✻ Contrast	0.053	0.053	1.011	0.122	8.015
SF + Target + Contrast + SF ✻ Target + SF ✻ Contrast + Target ✻ Contrast	0.053	0.050	0.938	0.114	8.255
SF + Target + Contrast + SF ✻ Target + Target ✻ Contrast	0.053	0.033	0.607	0.075	10.765
SF + Target + Contrast	0.053	0.032	0.602	0.074	6.821
SF + Target + Contrast + SF ✻ Target + SF ✻ Contrast + Target ✻ Contrast + SF ✻ Target ✻ Contrast	0.053	0.023	0.419	0.052	88.718
SF + Target + Contrast + SF ✻ Target + SF ✻ Contrast	0.053	0.005	0.099	0.013	10.430
SF + Target + SF ✻ Target	0.053	0.005	0.095	0.012	6.215

Note. All models include subject, and random slopes for all repeated measures factors.
Note. Showing the best 10 out of 19 models.

Fig. A2. Jasp output. Model comparison obtained from the Bayesian repeated measures ANOVA for the Target × Spatial Frequency × Contrast analysis on saccade latencies.

Analysis of Effects

Effects	P(incl)	P(excl)	P(incl|data)	P(excl|data)	BF_{incl}
SF	0.737	0.263	1.000	1.036×10^{-6}	344851.936
Target	0.737	0.263	1.000	2.477×10^{-5}	14415.582
Contrast	0.737	0.263	0.940	0.060	5.593
SF ✻ Target	0.316	0.684	0.119	0.881	0.293
SF ✻ Contrast	0.316	0.684	0.566	0.434	2.827
Target ✻ Contrast	0.316	0.684	0.846	0.154	11.863
SF ✻ Target ✻ Contrast	0.053	0.947	0.023	0.977	0.419

Fig. A3. Jasp output. Analysis of effects obtained from the Bayesian repeated measures ANOVA for the Target × Spatial Frequency × Contrast analysis on saccade latencies.

0.011) and fearful (*M* ± *SD: 0.026* ± *0.019*) than neutral faces (*M* ± *SD: 0.011* ± *0.001; p* <*.001*). It also tended to be higher for fearful than happy faces (*p* = 0.09). Then, it was higher for BSF (*M* ± *SD: 0.022* ± *0.011; p* =*.008*) and LSF (*M* ± *SD: 0.021* ± *0.011; p* =*.038*) than HSF (*M* ± *SD: 0.017* ± *0.012)* images. Concerning the contrast equalisation, mean mouth saliency was higher in the EQ (*M* ± *SD: 0.022* ± *0.015)* than in the NonEQ (*M* ± *SD: 0.018* ± *0.15)* contrast condition.

3.2.3. Relationship between saliency and behavioural data

Accuracy: Mean accuracy in the saccadic choice task in each condition was found to be positively correlated with the mean mouth saliency of the target ($r = 0.72, p < .001$). The more the mouth of the target was salient, the better the accuracy. There was no significant correlation between mean accuracy in the saccadic choice task and the eye saliency ($r = 0.14$, $p = .57$). The proportion of correct saccades as a function of the mouth or eye saliency of the target is presented in Fig. 7a, in which one point corresponds to one condition.

Latency: Mean latency in the saccadic choice task in each condition was found to be negatively correlated with the mean mouth saliency of the target (r = -0.5, p =.033). The higher the mouth was salient, the shorter the latency. There was no significant correlation between mean latency in the saccadic choice task and the eye saliency (r = 0.22, p =.39). The saccade latency as a function of the mouth or eye saliency of the target is presented in Fig. 7b.

4. General discussion

The aim of this study was to clarify the role of spatial frequency and luminance contrast in the detection of emotional faces through an eye-tracking experiment and an analysis of the saliency of the diagnostic face features (the eyes and the mouth) of the stimuli. For the eye-tracking experiment, based on a saccadic choice task, we replicated findings from previous studies. Performances were greater (higher accuracy, shorter latency) when the target was an emotional than a neutral face[2] (Bannerman et al., 2009; Entzmann et al., 2021). In addition, they were better for HSF than LSF images. For emotional targets, the accuracy was overall higher when the emotional face was happy than fearful. Interestingly, the difference between happy and fearful faces was only significant for HSF images, and the difference between HSF and LSF images was only significant for happy faces. Altogether, this suggests that participants mainly rely on HSF information to detect emotions, and that HSF are even more useful when the emotional face is happy. Concerning the latencies, they were overall shorter for BSF than filtered images, and for HSF compared to LSF images. Finally, the impact of contrast equalisation on performances was not as high as we expected. Nevertheless, the marginal interaction between the spatial frequencies and the contrast that we observed on saccade latencies showed that the difference between HSF and LSF was significant for equalised images only (i.e., when the contrast of HSF faces is enhanced to match that of BSF and LSF images).

To explain our behavioural results, we analysed in the second part of the paper, the saliency maps associated with our stimuli. As several previous studies suggest that the eyes and the mouth are the most important regions for expression decoding (Eisenbarth & Alpers, 2011; Smith & Schyns, 2009; Sweeny et al., 2013; Wegrzyn et al., 2017), we focused in our analyses on the saliency of the mouth and the eye regions. This saliency analysis was also motivated by some of our previous findings (Entzmann et al., 2021). Indeed, using similar saccadic choice tasks (i.e., with emotional faces; either presented in emotional-neutral or face-vehicle pairs) we showed that saccades landed differently on the faces according to the expression (e.g., saccades were lower on emotional than neutral faces, especially happy ones). We supposed that the saliency of the face features differentiating each expression can modulate the attention, leading to differences in saccade endpoints. Here, we were specifically interested in how this can be linked with the performance of participants. Overall, in the present study, unlike the mean eye saliency, the mean mouth saliency in the different experimental conditions was found to correlate with both the mean accuracy, and the mean latency of participants.[3] Correlation analysis shows that the greater the mouth saliency of the target was in one condition, the better the performance was.

A primary use of HSF despite the statistical sufficiency of LSF?

According to the coarse-to-fine theory of visual processing, LSF information is processed faster than HSF information (Bar, 2003; Hegde, 2008; Kauffmann et al., 2014; Kauffmann, Chauvin, et al., 2015; Musel et al., 2012; Peyrin et al., 2010; Schyns & Oliva, 1994). However, although the results of our saccadic choice task showed that LSF information is sufficient to accurately detect an emotion, as accuracy with LSF images was above chance, we observed shorter latencies for HSF. A possible explanation for the HSF bias in the saccadic choice task is that participants favor HSF because LSF are not informative enough to rapidly disambiguate the emotional or neutral content of faces. In our task, we have two faces side by side, which probably have the same global structure (thus a similar LSF content). To differentiate between the two in terms of emotion, we would have to rely on more detailed information. Furthermore, the analysis of the effect of the expression of distractor allowed us to show that accuracy was higher in HSF, with happy distractors. Thus, happy HSF faces do not seem to attract attention more, otherwise, they would have caused more errors as distractors. Rather, it seems that these happy faces make the task easier, especially in HSF, by conveying the most useful information to dissociate the neutral face from the emotional face. Indeed, this is in line with the idea that there is a flexible use of spatial frequencies for facial expression decoding. Different frequencies would be extracted depending on the expression (Cassidy et al., 2021; Morrison & Schyns, 2001; Oliva & Schyns, 1997; Schyns et al., 2009; Smith et al., 2005) and the task (Schyns & Oliva, 1999; Smith & Merlusca, 2014). For example, Schyns and Oliva found that when participants were asked to categorise facial expressions as angry, happy or neutral, they relied more on LSF information whereas they relied on HSF when they had to indicate whether the face was expressive or neutral (Schyns & Oliva, 1999), a result that is consistent with our data. Therefore, the use of spatial frequency information would not be a fixed process, as it may depend on the scale of the local features that are diagnostic, considering both task constraints and the spatial configuration of the face.

In our task, we showed that the mouth saliency is correlated with performance, but there may be other processes that contribute to the HSF over LSF preference. Participants may attend for local more than global differences between the two face images, which are structurally similar. Several studies show that attention can flexibly be directed to both global and local levels depending on expectations. For example, studies using Navon display showed enhanced processing of HSF after attending to local structure, and enhanced processing of LSF after attending to global structure (Flevaris et al., 2011; Ivry & Robertson, 1998; Robertson & Ivry, 2000). Moreover, some previous studies also proposed that stimulus duration influences the use of spatial information, and, that there is a HSF over LSF bias for long stimulus duration (Schyns & Oliva 1994). A well-known study from Schyns and Oliva (1994) used hybrid stimuli with two natural scene images from different categories and frequency scales to test the influence of stimuli duration. They found that a very short presentation time (30 ms) elicited a categorisation based on the LSF content whereas a longer presentation time (150 ms) elicited a categorisation based on the HSF content. In our task stimulus were displayed for a long presentation time (800 ms), there

[2] To verify that the repetition of the neutral faces cannot explain the emotional face advantage, we performed an additional analysis keeping only the first occurrence of each neutral face on each side (then each emotional and neutral face was repeated once on each side). Results show that there is still an effect of the Target, with higher accuracy ($t(77) = 7.9$, $p < .001$, $d = 0.9$) and shorter latencies ($t(77) = -5.6$, $p < .001$, $d = 0.64$) for emotional ($M \pm SD$ *for the accuracy*: 0.64 ± 0.12; *for latencies:* 228.1± 69.7 ms) than neutral faces ($M \pm SD$ *for the accuracy:* 0.58 ± 0.12; *for latencies:* 255.1 ± 89.8 ms). Thus, we conclude that the emotional face advantage cannot be attributed to the repetition of neutral faces.

[3] Note that we also tested whether the overall image saliency (i.e., the sum of the value of all the pixels of the saliency map, divided by the number of pixels) could significantly correlate with performance. Our results showed no significant correlation between overall saliency and performance for both saccade accuracy (r = -0.13, p =.61) and saccade latency (r = 0.12, p =.38) according to Pearson's correlation results.

was, therefore, no strong time constraint, which could explain the HSF bias. Nevertheless, in saccadic choice tasks with face-vehicle pairs, a LSF bias was found even with a relatively long presentation time (400 ms; Guyader et al., 2017). This could suggest that if the LSF content is distinct enough between the target and distractor (like with faces and vehicles that would be easy to distinguish in LSF) it could be favored regardless of the presentation time.

Overall, contrary to our hypothesis, we found no evidence for a better detection of LSF fearful faces, suggesting that they are not automatically (i.e., quickly and unintentionally) prioritised compared to other emotional faces competing for attention. Such a hypothesis (i.e., better detection of LSF fearful faces) was mainly based on neurophysiological data showing enhanced, or earlier, amygdala activation for LSF fearful faces, and on the existence of a subcortical pathway for rapid threat detection (LeDoux, 2000; Méndez-Bértolo et al., 2016; Morris et al., 1999; Öhman, 2005; Tamietto & de Gelder, 2010; Vuilleumier et al., 2003). However, it is important to clarify that the objective of the present study was not to provide evidence for such a pathway, a behavioural study alone would not be conclusive without being coupled with neuroimaging techniques. Rather, the objective was to test if, at the behavioural level, these stimuli could be detected more efficiently even when opposed to other faces. In this sense, we cannot exclude that LSF fearful faces activated the amygdala earlier in this task, but we showed that such an effect was not reflected here on eye movements. Then, even if a strict automaticity of the prioritization of LSF fearful faces is unlikely considering our results, it is still possible that it exists in other tasks. For example, while neurophysiological data diverge on the effect of the orientation of the attention on amygdala activity (Bayle & Taylor, 2010; Habel et al., 2007; Pessoa et al., 2002; Whalen et al., 1998; Vuilleumier et al., 2001), a prioritization of LSF fearful faces could still arise for implicit detection, as here we only explicitly assessed emotion detection. For example, this could be assessed if we replicate the experiment but without instructions on the target to see where the saccades go first.

Although cutoff frequency can vary across studies, we do not believe that our choices can explain the HSF advantage that we found here. Indeed, our cutoff frequency for HSF images is relatively high (40 cycles per face), compared to other studies (e.g., 24 cycles per face in McFadyen et al., 2017, Méndez-Bértolo et al. 2016, and Vuilleumier et al., 2003), therefore ensuring the activation of HSF parvocellular pathways specifically. Alternatively, our cutoff for LSF images (7 cycles per face) is closer to other studies in general (e.g., 6 cycles per face in McFadyen et al., 2017, Méndez-Bértolo et al. 2016, and Vuilleumier et al., 2003). Therefore, the HSF in our study are disadvantaged, since they contain less information (compared to other studies). With a lower HSF cutoff, we can suppose that the HSF advantage that we observe here would be even higher. One methodological choice that may have influenced our results and benefited HSF use is the shape of the filter used. In a study by Perfetto et al (2020), it was shown that the shape of the filter (Gaussian, Heaviside or Butterworth) can influence performance in a scene categorisation task. Performance was better in LSF than in HSF with a Heaviside filter, while with a Gaussian filter the opposite effect was observed. Given that we used Gaussian filters in our study, this may have also contributed to the lower performance for HSF images. However, the choice of Gaussian filter is allowing to compare our current study to the majority of scientific papers in the field (e.g., Goffaux & Rossion, 2006; Guyader et al., 2017; Kumar & Srinivasan, 2011; Vlamings et al., 2009; McFadyen et al., 2017; note that for some studies the shape of the filter used is not given, e.g., Méndez-Bértolo et al., 2016; Vuilleumier et al., 2003). Finally, although presenting stimuli in parafoveal vision may interfere with HSF processing, results clearly indicate that, at least for our chosen cutoff frequency for HSF images, HFS are perceived in parafoveal vision with the same accuracy as LSF. This is in line with studies that have shown that at 7.5° or 6.8° of eccentricity, the highest perceived spatial frequency is about 8–10 cycles per degree (Rovamo et al., 1978; Carrasco et al., 2001). Carrasco et al. (2001) also found a visual field inhomogeneity, as the detection of Gabor patches was stronger in the horizontal axis. Overall, it was expected that at least half of our HSF stimuli closest to the center is well perceived.

The role of contrast equalisation and local face features on the detection of emotions

The particularity of this study was also to consider contrast differences between HSF and LSF, by comparing performance in the saccadic choice task and in the mouth or eye saliency across two contrast conditions: one in which the RMS contrast was equalised after the filtering process, and one in which it was not (i.e., LSF images had a higher luminance contrast than HSF images). In the behavioural experiment, the expected interaction between contrast and spatial frequency was marginally significant on saccade latencies. More precisely, the difference between HSF and LSF conditions was only significant with equalised images. Even if the interaction did not reach the significance level, the direction of the effect is in line with previous findings (Vlamings et al., 2009; Perfetto et al., 2020; Kauffmann, Chauvin, et al., 2015; Kauffmann, Ramanoël, et al., 2015). For example, Kauffmann, Chauvin, et al. (2015) observed that scene categorisation as indoor or outdoor in HSF was slower without contrast equalisation than with contrast equalisation. In another study performed by Vlamings et al. (2009) using LSF or HSF faces, with or without contrast equalisation, participants were asked to decide whether the presented stimulus was a fearful or a neutral face. Results showed that LSF faces were categorised more rapidly than HSF faces, for both equalised and non-equalised stimuli. However, this observed difference was stronger when contrast was not equalised between LSF and HSF faces. Overall, this better processing of HSF after contrast equalisation is also congruent with the saliency analysis showing enhanced mouth or eye saliency in HSF when the contrast is equalised. The fact that the spatial frequency and contrast interaction was only marginal and not significant interaction on saccade latencies could be due to a lack of power. Indeed, using a Bayesian ANOVA we showed that this effect is anecdotal but still worth taking into account (see Appendix C for detailed results). Also, the effect size may be smaller in this study because the spatial frequency conditions were presented in blocks, allowing the participant to get used to the current contrast and take it as a baseline. Saccade accuracy was not sensitive to different contrast conditions, suggesting that overall contrast is not important for the accuracy of a discrimination, which may be mostly based on within-pairs differences.

Both saccade accuracy and latency correlate with the mouth saliency, supposing that local contrasts and orientations in the diagnostic regions contribute to the efficiency of the detection of emotional and neutral faces Our findings indicated that greater mouth saliency in one condition resulted in better performance, with higher accuracy and shorter latency. Overall the mouth saliency was higher for emotional than neutral faces. This is likely to be attributed to the fact that the emotional faces that we used provide salient information in the mouth area. Indeed, all emotional faces, with the exception of one happy face, have their mouths open. The mouth saliency is also higher for happy than fearful faces in HSF, and for HSF than LSF for happy faces, which could be explained by the fact that happy faces have more visible teeth, that are predominantly transmitted through HSF. The mouth saliency was higher for HSF in the EQ than NonEQ condition which is consistent with the results on the latency of participants. We suggest that higher mouth saliency in those conditions attracted attention, thereby making detection faster and easier. As proposed by Calvo and Nummenmaa (2011), emotion detection may arise from a two-stage processing mechanism. The mouth saliency would be analysed in a first, purely perceptual stage, and it will then be used for expression recognition and semantic retrieval.

However, it should be noted that the differences are not always systematically the same in the performance and saliency analysis. For example, the mouth of LSF fearful faces in the NonEQ contrast condition is particularly salient, probably because it shows fewer teeth than that of happy faces, and would therefore be almost exclusively transmitted through LSF. And, this is not associated with particularly high performance. Thus, even if the mouth saliency is strongly linked to the participant's performance, it is likely that other mechanisms are involved. For example, participants may choose to focus on HSF which would tune their perception toward HSF information. Or, despite this LSF information being salient in fearful faces, it may not be diagnostic enough to distinguish the two faces. Taken together, our results indicate that our initial hypotheses, based on the neural pathway potentially involved in threat detection in peripheral vision based on LSF, is actually less important than the perceptual saliency of the diagnostic features for a specific task.

Interestingly, whereas the mouth saliency was important for participants, the eye saliency didn't influence performance. This idea that the eye region plays a limited role in the detection of neutral and emotional faces contrasts with studies showing that both the eye and the mouth are important (Dailey et al., 2002; Eisenbarth & Alpers, 2011; Smith et al., 2005; Smith & Schyns, 2009; Wegrzyn et al., 2017). However, such studies focused on the categorisation of emotions. They showed that the mouth is particularly useful for the categorisation of happy faces and the eyes for fearful faces. Here, the task is different, and closer to a categorisation of faces as emotional or neutral. Therefore, it is not so surprising that the features are used differently than when the task is to categorise precisely the expression as happy or fearful. Moreover, several studies highlighted a higher reliance on the mouth region (Blais et al., 2012; Calvo & Nummenmaa, 2011; Calvo et al., 2014). For example, Blais et al. (2012) found that, across all expressions, the mouth area is the most important cue for categorising both static and dynamic facial expressions. The authors suggest that this area is prioritised because it is more informative (especially for dynamic stimuli), and that the brain focuses on the most informative regions due to the limited capacity of the visuo-cognitive system. Also, results from Calvo et al. (2014) showed a happy face recognition advantage with faces presented either at fixation or extrafoveally, which remained even when only the mouth region was presented. The smiling mouth was the most salient and distinctive feature of all expressions, which, according to the authors, explains the happy face advantage. Overall, it is possible that participants make the strategy of focusing on the mouth because they consider that it is the most informative region. Going further, we can suppose that such a strategy would favour the detection of emotional faces when they are happy. Indeed, for happy faces, the mouth region is highly diagnostic, whereas for fearful faces it is less stereotypical. Finally, our results are limited to the use of happy and fearful faces. We can suppose that with other emotional faces, the results could have been different, especially if the emotional faces convey useful information through the eye region (e.g., fearful and angry faces; Smith et al., 2005).

5. Conclusion

This study replicated findings from previous studies showing a better detection of emotional than neutral faces, especially with a happy emotional face (Entzmann et al., 2021). Also, we found that the discrimination between an emotional and a neutral face was overall easier for HSF and BSF than LSF images. On saccade latency, the HSF over LSF advantage was significant for equalised images only, suggesting that the processing of spatial frequencies can be dependent on contrast differences. The saliency analysis of our stimuli revealed that the mean mouth saliency in the different experimental conditions, unlike the eye saliency, correlated with both the mean accuracy and the mean latency of participants. Overall, our results go against the idea that there is an automatic (quick and unintentional) prioritization of low spatial frequency fearful faces. Rather, we suggest that participants favoured the use of high spatial frequencies in this task because low spatial frequencies are not informative enough to rapidly disambiguate the emotional or neutral content of faces. This would be consistent with the idea that there is a flexible use of spatial frequencies for facial expression decoding, depending on the scale of the useful information for a specific task (e.g., Smith & Merlusca, 2014). Also, as suggested in previous papers, the saliency of diagnostic features, the mouth in this study, may be used as a shortcut for efficient expression decoding (e.g., Calvo & Nummenmaa, 2016).

CRediT authorship contribution statement

Léa Entzmann: Conceptualization, Methodology, Software, Investigation, Formal analysis, Visualization, Writing – original draft. **Nathalie Guyader:** Conceptualization, Methodology, Writing – review & editing. **Louise Kauffmann:** Conceptualization, Methodology, Writing – review & editing. **Carole Peyrin:** Conceptualization, Methodology, Writing – review & editing. **Martial Mermillod:** Conceptualization, Methodology, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data and analysisc ode that support the findings of this study are available in the Open Science Framework repository at https://osf.io/hyr52.

Acknowledgments

This work was supported by NeuroCoG IDEX UGA in the framework of the "Investissements d'avenir" program (ANR-15-IDEX-02). This work has been partially supported by MIAI @ Grenoble Alpes, (ANR-19-P3IA-0003).

Appendix A

A. ANOVA tables

See Tables A1-A4

B. Saliency toolbox interface

See Figs. A1-A3

C. Bayesian repeated measures ANOVA for the Target × Spatial Frequency × Contrast analysis on saccade latencies

Here we present results from a Bayesian repeated measures ANOVA corresponding to the Target × Spatial Frequency × Contrast analysis on saccade latencies. Therefore, this analysis serves as a Bayesian equivalent to the frequentist analysis presented in section 2.2.2, which led to a marginally significant Spatial Frequency × Contrast interaction. The analysis was performed using Jasp (JASP Team, 2023). The advantage of this method is that it allows the strength of each effect to be quantified, without going through a significance threshold. First, this analysis provides a model comparison, whose results are presented in Fig. A.2 and show the support that the data offer for each possible model. The left-most column lists the tested models (here only the best 10 out of 19 models), ordered by their predictive performance relative to the best model. The winning model suggests that the data are influenced by each factor (Target, Spatial Frequency, and Contrast) as well as the Spatial Frequency × EFE interaction and, more interestingly, the Spatial Frequency × Contrast interaction. The BF_{10} associated with each model can inform us about the strength of the support for the winning model compared to the model at hand. For example, the difference between the first two models is anecdotal (BF_{10} between 0.33 and 1) and the difference between the first and the third model is moderate (BF_{10} between 0.01 and 0.33). This model comparison approach does not provide a separate Bayes Factor for each effect. But an analysis of effects can be made to look specifically at the likelihood ratio of models containing each effect. Results are presented in Fig. A.3. We found evidence for the Spatial Frequency × Contrast interaction, although the BFincl value was anecdotal. For a more in-depth discussion of Bayesian ANOVA and Bayes factor interpretation, see Jeffreys (1998), Rosenfeld and Olson (2021), and van den Bergh et al. (2020).

References

Adolphs, R., Gosselin, F., Buchanan, T. W., Tranel, D., Schyns, P., & Damasio, A. R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature, 433*(7021), 68–72. https://doi.org/10.1038/nature03086

Awasthi, B., Friedman, J., & Williams, M. A. (2011). Faster, stronger, lateralized : Low spatial frequency information supports face processing. *Neuropsychologia, 49*(13), 3583–3590. https://doi.org/10.1016/j.neuropsychologia.2011.08.027

Bannerman, R. L., Milders, M., & Sahraie, A. (2009). Processing emotional stimuli: Comparison of saccadic and manual choice-reaction times. *Cognition & Emotion, 23* (5), 930–954. https://doi.org/10.1080/02699930802243303

Bar, M. (2003). A Cortical Mechanism for Triggering Top-Down Facilitation in Visual Object Recognition. *Journal of Cognitive Neuroscience, 15*(4), 600–609. https://doi.org/10.1162/089892903321662976

Bayle, D. J., & Taylor, M. J. (2010). Attention inhibition of early cortical activation to fearful faces. *Brain Research, 1313*, 113–123.

Bisley, J. W., & Mirpour, K. (2019). The neural instantiation of a priority map. *Current opinion in psychology, 29*, 108–112.

Blais, C., Roy, C., Fiset, D., Arguin, M., & Gosselin, F. (2012). The eyes are not the window to basic emotions. *Neuropsychologia, 50*(12), 2830–2838.

Borji, A., & Itti, L. (2013). State-of-the-Art in Visual Attention Modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 35*(1), 185–207. https://doi.org/10.1109/TPAMI.2012.89

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Burra, N., Hervais-Adelman, A., Celeghin, A., De Gelder, B., & Pegna, A. J. (2019). Affective blindsight relies on low spatial frequencies. Neuropsychologia, 128, 44–49.

Calvo, M. G., Fernández-Martín, A., & Nummenmaa, L. (2014). Facial expression recognition in peripheral versus central vision: Role of the eyes and the mouth. *Psychological research, 78*, 180–195.

Calvo, M. G., & Nummenmaa, L. (2009). Eye-movement assessment of the time course in facial expression recognition: Neurophysiological implications. *Cognitive, Affective, & Behavioral Neuroscience, 9*(4), 398–411. https://doi.org/10.3758/CABN.9.4.398

Calvo, M. G., & Nummenmaa, L. (2011). Time course of discrimination between emotional facial expressions : The role of visual saliency. *Vision Research, 51*(15), 1751–1759. https://doi.org/10.1016/j.visres.2011.06.001

Calvo, M. G., & Nummenmaa, L. (2016). Perceptual and affective mechanisms in facial expression recognition: An integrative review. Cognition and Emotion, 30(6), 1081-1106.

Carrasco, M., Talgar, C. P., & Cameron, E. L. (2001). Characterizing visual performance fields: Effects of transient covert attention, spatial frequency, eccentricity, task and set size. *Spatial vision, 15*(1), 61–75.

Cassidy, B., Wiley, R., Sim, M., & Hugenberg, K. (2021). Spatial frequency and valence interact in complex emotion perception. *Cognition and Emotion, 35*(8), 1618–1625.

Cerf, M., Frady, E. P., & Koch, C. (2009). Faces and text attract gaze independent of the task : Experimental data and computer model. *Journal of Vision, 9*(12), 10. https://doi.org/10.1167/9.12.10

Charbonneau, I., Guérette, J., Cormier, S., Blais, C., Lalonde-Beaudoin, G., Smith, F. W., & Fiset, D. (2021). The role of spatial frequencies for facial pain categorization. *Scientific Reports, 11*(1), 1–12.

Corradi-Dell'Acqua, C., Schwartz, S., Meaux, E., Hubert, B., Vuilleumier, P., & Deruelle, C. (2014). Neural responses to emotional expression information in high- and low-spatial frequency in autism: Evidence for a cortical dysfunction. *Frontiers in human neuroscience, 8*, 189.

Coutrot, A., & Guyader, N. (2014). How saliency, faces, and sound influence gaze in dynamic social scenes. *Journal of Vision, 14*(8), 5. https://doi.org/10.1167/14.8.5

Crouzet, S. M. (2010). Fast saccades toward faces : Face detection in just 100 ms. *Journal of Vision, 10*(4), 1–17. https://doi.org/10.1167/10.4.16

Dailey, M. N., Cottrell, G. W., Padgett, C., & Adolphs, R. (2002). EMPATH: A Neural Network that Categorizes Facial Expressions. *Journal of Cognitive Neuroscience, 14*(8), 1158–1173.

DeValois, R. L., Burke, P., De Valois, K. K., & DeValois, K. K. (1990). *Spatial vision, No. 14*. Oxford University Press on Demand.

Devue, C., & Grimshaw, G. M. (2017). Faces are special, but facial expressions aren't: Insights from an oculomotor capture paradigm. *Attention, Perception, & Psychophysics, 79*(5), 1438–1452. https://doi.org/10.3758/s13414-017-1313-x

D'Hondt, F., Szaffarczyk, S., Sequeira, H., & Boucart, M. (2016). Explicit and implicit emotional processing in peripheral vision: A saccadic choice paradigm. *Biological psychology, 119*, 91–100.

Eisenbarth, H., & Alpers, G. W. (2011). Happy mouth and sad eyes : Scanning emotional facial expressions. *Emotion, 11*(4), 860–865. https://doi.org/10.1037/a0022758

Entzmann, L., Guyader, N., Kauffmann, L., Lenouvel, J., Charles, C., Peyrin, C., Vuillaume, R., & Mermillod, M. (2021). The Role of Emotional Content and Perceptual Saliency During the Programming of Saccades Toward Faces. *Cognitive Science, 45*, e13042.

Farah, M. J., Wilson, K. D., & Drain, M. (1998). What Is « Special ». *About Face Perception?, 17*.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191.

Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A, 4*(12), 2379. https://doi.org/10.1364/JOSAA.4.002379

Fitzgerald, D. A., Angstadt, M., Jelsone, L. M., Nathan, P. J., & Phan, K. L. (2006). Beyond threat : Amygdala reactivity across multiple expressions of facial affect. *NeuroImage, 30*(4), 1441–1448. https://doi.org/10.1016/j.neuroimage.2005.11.003

Flevaris, A. V., Bentin, S., & Robertson, L. C. (2011). Attention to hierarchical level influences attentional selection of spatial scale. *Journal of Experimental Psychology: Human Perception and Performance, 37*(1), 12–22.

Foulsham, T., & Underwood, G. (2008). What can saliency models predict about eye movements? Spatial and sequential aspects of fixations during encoding and recognition. *Journal of Vision, 8*(2), 6. https://doi.org/10.1167/8.2.6

Frischen, A., Eastwood, J. D., & Smilek, D. (2008). Visual search for faces with emotional expressions. *Psychological bulletin, 134*(5), 662.

Fusar-Poli, P., Placentino, A., Carletti, F., Landi, P., Allen, P., Surguladze, S., … Politi, P. (2009). Functional atlas of emotional faces processing: A voxel-based meta-analysis of 105 functional magnetic resonance imaging studies. *Journal of psychiatry and neuroscience, 34*(6), 418–432.

Garvert, M. M., Friston, K. J., Dolan, R. J., & Garrido, M. I. (2014). Subcortical amygdala pathways enable rapid face processing. *NeuroImage, 102*, 309–316. https://doi.org/10.1016/j.neuroimage.2014.07.047

Gaspar, C., Sekuler, A. B., & Bennett, P. J. (2008). Spatial frequency tuning of upright and inverted face identification. *Vision Research, 48*(28), 2817–2826.

Goffaux, V., Peters, J., Haubrechts, J., Schiltz, C., Jansma, B., & Goebel, R. (2011). From Coarse to Fine? Spatial and Temporal Dynamics of Cortical Face Processing. *Cerebral Cortex, 21*(2), 467–476. https://doi.org/10.1093/cercor/bhq112

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. https://doi.org/10.1037/0096-1523.32.4.1023

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.

Guyader, N., Chauvin, A., Boucart, M., & Peyrin, C. (2017). Do low spatial frequencies explain the extremely fast saccades towards human faces? *Vision Research, 133*, 100–111. https://doi.org/10.1016/j.visres.2016.12.019

Habel, U., Windischberger, C., Derntl, B., Robinson, S., Kryspin-Exner, I., Gur, R. C., & Moser, E. (2007). Amygdala activation and facial expressions: Explicit emotion discrimination versus implicit emotion processing. *Neuropsychologia, 45*(10), 2369–2377.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4*(6), 223–233. https://doi.org/10.1016/S1364-6613(00)01482-0

Hegde, J. (2008). Time course of visual perception : Coarse-to-fine processing and beyond. *Progress in Neurobiology, 84*(4), 405–439. https://doi.org/10.1016/j.pneurobio.2007.09.001

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research, 40*(10–12), 1489–1506. https://doi.org/10.1016/S0042-6989(99)00163-7

Ivry, R. B., & Robertson, L. C. (1998). *The two sides of perception*. MIT press.

Jeffreys, H. (1998). *The theory of probability*. OuP Oxford.

Kauffmann, L., Chauvin, A., Guyader, N., & Peyrin, C. (2015). Rapid scene categorization : Role of spatial frequency order, accumulation mode and luminance contrast. *Vision Research, 107*, 49–57. https://doi.org/10.1016/j.visres.2014.11.013

Kauffmann, L., Khazaz, S., Peyrin, C., & Guyader, N. (2021). Isolated face features are sufficient to elicit ultra-rapid and involuntary orienting responses toward faces. *Journal of Vision, 21*(2), 4. https://doi.org/10.1167/jov.21.2.4

Kauffmann, L., Peyrin, C., Chauvin, A., Entzmann, L., Breuil, C., & Guyader, N. (2019). Face perception influences the programming of eye movements. *Scientific Reports, 9* (1). https://doi.org/10.1038/s41598-018-36510-0

Kauffmann, L., Ramanoël, S., & Peyrin, C. (2014). The neural bases of spatial frequency processing during scene perception. *Frontiers in Integrative Neuroscience, 8*. https://doi.org/10.3389/fnint.2014.00037

Kauffmann, L., Ramanoël, S., Guyader, N., Chauvin, A., & Peyrin, C. (2015). Spatial frequency processing in scene-selective cortical regions. *NeuroImage, 112*, 86–95. https://doi.org/10.1016/j.neuroimage.2015.02.058

Kirchner, H., & Thorpe, S. J. (2006). Ultra-rapid object detection with saccadic eye movements : Visual processing speed revisited. *Vision Research, 46*(11), 1762–1776. https://doi.org/10.1016/j.visres.2005.10.002

Klink, P. C., Jentgens, P., & Lorteije, J. A. M. (2014). Priority maps explain the roles of value, attention, and salience in goal-oriented behavior. *Journal of Neuroscience, 34* (42), 13867–13869.

Koehler, K., Guo, F., Zhang, S., & Eckstein, M. P. (2014). What do saliency models predict? *Journal of Vision, 14*(3), 14. https://doi.org/10.1167/14.3.14

Krolak-Salmon, P., Hénaff, M. A., Vighetto, A., Bertrand, O., & Mauguière, F. (2004). Early amygdala reaction to fear spreading in occipital, temporal, and frontal cortex: A depth electrode ERP study in human. *Neuron, 42*(4), 665–676.

Kumar, D., & Srinivasan, N. (2011). Emotion perception is mediated by spatial frequency content. *Emotion, 11*(5), 1144.
Langton, S. R. H., Law, A. S., Burton, A. M., & Schweinberger, S. R. (2008). Attention capture by faces. *Cognition, 107*(1), 330–342. https://doi.org/10.1016/j.cognition.2007.07.012
LeDoux, J. E. (2000). Emotion Circuits in the. *Brain., 31*.
Liu, J., Harris, A., & Kanwisher, N. (2002). Stages of processing in face perception : An MEG study. *Nature Neuroscience, 5*(9), 910–916. https://doi.org/10.1038/nn909
Ludwig, C. J., Gilchrist, I. D., & McSorley, E. (2004). The influence of spatial frequency and contrast on saccade latencies. *Vision research, 44*(22), 2597–2604.
Lundqvist, D., Flykt, A., & Ohman, A. (1998). The Karolinska directed emotional faces (KDEF).
Lundqvist, D., Bruce, N., & Öhman, A. (2015). Finding an emotional face in a crowd: Emotional and perceptual stimulus factors influence visual search efficiency. *Cognition and Emotion, 29*(4), 621–633.
Marat, S., Ho Phuoc, T., Granjon, L., Guyader, N., Pellerin, D., & Guérin-Dugué, A. (2009). Modelling Spatio-Temporal Saliency to Predict Gaze Direction for Short Videos. *International Journal of Computer Vision, 82*(3), 231–243. https://doi.org/10.1007/s11263-009-0215-3
McFadyen, J., Mermillod, M., Mattingley, J. B., Halász, V., & Garrido, M. I. (2017). A Rapid Subcortical Amygdala Route for Faces Irrespective of Spatial Frequency and Emotion. *The Journal of Neuroscience, 37*(14), 3864–3874. https://doi.org/10.1523/JNEUROSCI.3525-16.2017
Méndez-Bértolo, C., Moratti, S., Toledano, R., Lopez-Sosa, F., Martínez-Alvarez, R., Mah, Y. H., Vuilleumier, P., Gil-Nagel, A., & Strange, B. A. (2016). A fast pathway for fear in human amygdala. *Nature Neuroscience, 19*(8), 1041–1049. https://doi.org/10.1038/nn.4324
Mermillod, M., Bonin, P., Mondillon, L., Alleysson, D., & Vermeulen, N. (2010). Coarse scales are sufficient for efficient categorization of emotional facial expressions : Evidence from neural computation. *Neurocomputing, 73*(13–15), 2522–2531. https://doi.org/10.1016/j.neucom.2010.06.002
Mermillod, M., Vuilleumier, P., Peyrin, C., Alleysson, D., & Marendaz, C. (2009). The importance of low spatial frequency information for recognising fearful facial expressions. *Connection Science, 21*(1), 75–83. https://doi.org/10.1080/09540090802213974
Mermillod, M., Vermeulen, N., Lundqvist, D., & Niedenthal, P. M. (2009). Neural computation as a tool to differentiate perceptual from emotional processes: The case of anger superiority effect. *Cognition, 110*(3), 346–357.
Morris, J. (1998). A neuromodulatory role for the human amygdala in processing emotional facial expressions. *Brain, 121*(1), 47–57. https://doi.org/10.1093/brain/121.1.47
Morris, J. S., Öhman, A., & Dolan, R. J. (1999). A subcortical pathway to the right amygdala mediating "unseen" fear. *Proceedings of the National Academy of Sciences, 96*(4), 1680–1685.
Morrison, D. J., & Schyns, P. G. (2001). Usage of spatial scales for the categorization of faces, objects, and scenes. *Psychonomic Bulletin & Review, 8*(3), 454–469. https://doi.org/10.3758/BF03196180
Mulckhuyse, M. (2018). The influence of emotional stimuli on the oculomotor system: A review of the literature. *Cognitive, Affective, & Behavioral Neuroscience, 18*, 411–425.
Musel, B., Chauvin, A., Guyader, N., Chokron, S., & Peyrin, C. (2012). Is Coarse-to-Fine Strategy Sensitive to Normal Aging? *PLoS ONE, 7*(6), e38493.
Näsänen, R. (1999). Spatial frequency bandwidth used in the recognition of facial images. *Vision research, 39*(23), 3824–3833.
Öhman, A. (2005). The role of the amygdala in human fear : Automatic detection of threat. *Psychoneuroendocrinology, 30*(10), 953–958. https://doi.org/10.1016/j.psyneuen.2005.03.019
Oliva, A., & Schyns, P. G. (1997). Coarse Blobs or Fine Edges? Evidence That Information Diagnosticity Changes the Perception of Complex Visual Stimuli. *Cognitive Psychology, 34*(1), 72–107. https://doi.org/10.1006/cogp.1997.0667
Ottaviani, C., Cevolani, D., Nucifora, V., Borlimi, R., Agati, R., Leonardi, M., … Brighetti, G. (2012). Amygdala responses to masked and low spatial frequency fearful faces: A preliminary fMRI study in panic disorder. *Psychiatry Research: Neuroimaging, 203*(2–3), 159–165.
Perfetto, S., Wilder, J., & Walther, D. B. (2020). Effects of Spatial Frequency Filtering Choices on the Perception of Filtered Images. *Vision, 4*(2), 29. https://doi.org/10.3390/vision4020029
Pessoa, L., & Adolphs, R. (2010). Emotion processing and the amygdala : From a « low road » to « many roads » of evaluating biological significance. *Nature Reviews Neuroscience, 11*(11), 773–782. https://doi.org/10.1038/nrn2920
Rosenfeld, P. J., & Olson, J. M. (2021). Bayesian data analysis: A fresh approach to power issues and null hypothesis interpretation. *Applied Psychophysiology and Biofeedback, 46*, 135–140.
Pessoa, L., McKenna, M., Gutierrez, E., & Ungerleider, L. G. (2002). Neural processing of emotional faces requires attention. *Proceedings of the National Academy of Sciences, 99*(17), 11458–11463.
Peters, J. C., Goebel, R., & Goffaux, V. (2018). From coarse to fine: Interactive feature processing precedes local feature analysis in human face perception. *Biological psychology, 138*, 1–10.
Petras, K., Ten Oever, S., Jacobs, C., & Goffaux, V. (2019). Coarse-to-fine information integration in human vision. *NeuroImage, 186*, 103–112.
Peyrin, C., Michel, C. M., Schwartz, S., Thut, G., Seghier, M., Landis, T., Marendaz, C., & Vuilleumier, P. (2010). The Neural Substrates and Timing of Top-Down Processes during Coarse-to-Fine Categorization of Visual Scenes : A Combined fMRI and ERP Study. *Journal of Cognitive Neuroscience, 22*(12), 2768–2780. https://doi.org/10.1162/jocn.2010.21424
Plouffe-Demers, M. P., Fiset, D., Saumure, C., Duncan, J., & Blais, C. (2019). Strategy shift toward lower spatial frequencies in the recognition of dynamic facial expressions of basic emotions: When it moves it is different. *Frontiers in psychology, 10*, 1563.
Quek, G. L., Liu-Shuang, J., Goffaux, V., & Rossion, B. (2018). Ultra-coarse, single-glance human face detection in a dynamic visual stream. *NeuroImage, 176*, 465–476.
Racine, J. S. (2012). RStudio : A platform-independent IDE for R and Sweave. *Journal of Applied Econometrics.*
Robertson, L. C., & Ivry, R. (2000). Hemispheric Asymmetries: Attention to Visual and Auditory Primitives. *Current Directions in Psychological Science, 9*(2), 59–63.
Rovamo, J., Virsu, V., & Näsänen, R. (1978). Cortical magnification factor predicts the photopic contrast sensitivity of peripheral vision. *Nature, 271*(5640), 54–56.
Sander, D., Grafman, J., & Zalla, T. (2003). The human amygdala: An evolved system for relevance detection. *Reviews in the Neurosciences, 14*(4), 303–316.
Schyns, P. G., & Oliva, A. (1994). From blobs to boundary edges: Evidence for time-and spatial-scale-dependent scene recognition. *Psychological science, 5*(4), 195–200.
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. https://doi.org/10.1016/S0010-0277(98)00069-9
Schyns, P. G., Petro, L. S., & Smith, M. L. (2009). Transmission of Facial Expressions of Emotion Co-Evolved with Their Efficient Decoding in the Brain : Behavioral and Brain Evidence. *PLoS ONE, 4*(5), e5625.
Shapley, R., & Enroth-Cugell, C. (1984). Chapter 9 Visual adaptation and retinal gain controls. *Progress in Retinal Research*, 3, 263–346. 10.1016/0278-4327(84)90011-7.
Smith, F. W., & Schyns, P. G. (2009). Smile Through Your Fear and Sadness : Transmitting and Identifying Facial Expression Signals Over a Range of Viewing Distances. *Psychological Science, 20*(10), 1202–1208. https://doi.org/10.1111/j.1467-9280.2009.02427.x
Smith, M. L., Cottrell, G. W., Gosselin, F., & Schyns, P. G. (2005). Transmitting and Decoding Facial Expressions. *Psychological Science, 16*(3), 184–189. https://doi.org/10.1111/j.0956-7976.2005.00801.x
Smith, M. L., & Merlusca, C. (2014). How task shapes the use of information during facial expression categorizations. *Emotion, 14*(3), 478–487. https://doi.org/10.1037/a0035588
Stein, T., Seymour, K., Hebart, M. N., & Sterzer, P. (2014). Rapid Fear Detection Relies on High Spatial Frequencies. *Psychological Science, 25*(2), 566–574. https://doi.org/10.1177/0956797613512509
Sweeny, T. D., Suzuki, S., Grabowecky, M., & Paller, K. A. (2013). Detecting and categorizing fleeting emotions in faces. *Emotion, 13*(1), 76–91. https://doi.org/10.1037/a0029193
Tamietto, M., & de Gelder, B. (2010). Neural bases of the non-conscious perception of emotional signals. *Nature Reviews Neuroscience, 11*(10), 697–709. https://doi.org/10.1038/nrn2889
Van den Bergh, D., Van Doorn, J., Marsman, M., Draws, T., Van Kesteren, E. J., Derks, K., … Wagenmakers, E. J. (2020). A tutorial on conducting and interpreting a Bayesian ANOVA in JASP. *LAnnee psychologique, 120*(1), 73–96.
Van der Donck, S., Tang, T., Dzhelyova, M., Wagemans, J., & Boets, B. (2020). Pinpointing the optimal spatial frequency range for automatic neural facial fear processing. *NeuroImage, 221*, Article 117151.
Van der Schaaf, V. A., & van Hateren, J. V. (1996). Modelling the power spectra of natural images: Statistics and information. *Vision research, 36*(17), 2759–2770.
Vlamings, P. H. J. M., Goffaux, V., & Kemner, C. (2009). Is the early modulation of brain activity by fearful facial expressions primarily mediated by coarse low spatial frequency information? *Journal of Vision, 9*(5), 12. https://doi.org/10.1167/9.5.12
Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. 2001 Effects of Attention and Emotion on Face *Processing in the Human Brain*: An Event-Related fMRI Study, 13.
Vuilleumier, P. (2015). Affective and motivational control of vision. *Current opinion in neurology, 28*(1), 29–35.
Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2003). Distinct spatial frequency sensitivities for processing faces and emotional expressions. *Nature Neuroscience, 6*(6), 624–631. https://doi.org/10.1038/nn1057
Walther, D., & Koch, C. (2006). Modeling attention to salient proto-objects. *Neural Networks, 19*(9), 1395–1407. https://doi.org/10.1016/j.neunet.2006.10.001
Wegrzyn, M., Vogt, M., Kireclioglu, B., Schneider, J., & Kissler, J. (2017). Mapping the emotional face. How individual face parts contribute to successful emotion recognition. *PLOS ONE, 12*(5), e0177239.
Whalen, P. J., Rauch, S. L., Etcoff, N. L., McInerney, S. C., Lee, M. B., & Jenike, M. A. (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *Journal of neuroscience, 18*(1), 411–418.
Willenbockel, V., Fiset, D., Chauvin, A., Blais, C., Arguin, M., Tanaka, J. W., … Gosselin, F. (2010). Does face inversion change spatial frequency tuning? *Journal of Experimental Psychology: Human Perception and Performance, 36*(1), 122.
Zhao, Y., Zhen, Z., Liu, X., Song, Y., & Liu, J. (2018). The neural network for face recognition : Insights from an fMRI study on developmental prosopagnosia. *NeuroImage, 169*, 151–161. https://doi.org/10.1016/j.neuroimage.2017.12.023