## Title
Orientation information in encoding facial expressions



## Authors
Yu, Deyue
Chai, Andrea
Chung, Susana TL

## Publication Date
2018-09-01

## DOI
10.1016/j.visres.2018.07.001

Peer reviewed

HHS Public Access
Author manuscript
*Vision Res*. Author manuscript; available in PMC 2019 September 01.

Published in final edited form as:
*Vision Res*. 2018 September ; 150: 29–37. doi:10.1016/j.visres.2018.07.001.

# Orientation Information in Encoding Facial Expressions

**Deyue Yu**[1], **Andrea Chai**[2], and **Susana T.L. Chung**[2]

[1]College of Optometry, The Ohio State University, Columbus, Ohio

[2]School of Optometry, University of California, Berkeley, CA

## Abstract

Previous research showed that we use different regions of a face to categorize different facial expressions, e.g. mouth region for identifying happy faces; eyebrows, eyes and upper part of nose for identifying angry faces. These findings imply that the spatial information along or close to the horizontal orientation might be more useful than others for facial expression recognition. In this study, we examined how the performance for recognizing facial expression depends on the spatial information along different orientations, and whether the pixel-level differences in the face images could account for subjects' performance. Four facial expressions—angry, fearful, happy and sad—were tested. An orientation filter (bandwidth=23°) was applied to restrict information within the face images, with the center of the filter ranged from 0° (horizontal) to 150° in steps of 30°. Accuracy for recognizing facial expression was measured for an unfiltered and the six filtered conditions. For all four facial expressions, recognition performance (normalized $d'$) was virtually identical for filter orientations of −30°, horizontal and 30°, and declined systematically as the filter orientation approached vertical. The information contained in mouth and eye regions is a significant predictor for subject's response (based on the confusion patterns). We conclude that young adults with normal vision categorizes facial expression most effectively based on the spatial information around the horizontal orientation which captures primary changes of facial features across expressions. Across all spatial orientations, the information contained in mouth and eye regions contributes significantly to facial expression categorization.

## Introduction

Recognizing facial expressions is an important skill in social interactions. Many previous studies have focused on evaluating the role of spatial frequencies in facial expression recognition. It has been shown that subjects perform the task of categorizing facial expressions based on low spatial frequency information contained within the face images, although the detection of an expression and the strength of an expression engages the use of high spatial frequency information (Calder, Young, Keane, & Dean, 2000; Schyns & Oliva, 1999; Vuilleumier, Armony, Driver, & Dolan, 2003). In early visual processing, retinal input such as face image was decomposed not only along the dimension of spatial frequency but also along the dimension of orientation (e.g., (De Valois, Albrecht, & Thorell, 1982; Hubel & Wiesel, 1968). We have also learnt that to precisely categorize facial expressions, subjects tend to use different configurations of facial regions (e.g. mouth region for happy face; eyebrows, eyes and upper part of nose for angry face) (Smith, Cottrell, Gosselin, & Schyns, 2005). Most of the facial elements such as eyebrows, eyes, mouths are more horizontally oriented. When generating facial expressions, there also seems to be more variations in configurations among facial elements oriented horizontally. As shown in a later section (Figures 7, 8 and 9), image comparisons between different facial expressions demonstrate that the majority of the configural differences occur near horizontal orientation. These findings indicate that information conveyed by channels along or near horizontal orientation might be more important than the others for facial expression recognition.

Several studies explored how the orientation of spatial information could affect face identification (e.g., Dakin & Watt, 2009; Yu & Chung, 2011). By evaluating face identification using images filtered along various bands of orientation, these studies showed that subjects performed best when viewing images containing information close to horizontal orientation, with performance declining gradually as the orientation of the reserved information approached vertical. Goffaux and Dakin (2010) further examined the impact of horizontally orientated facial information on several key behavioral signatures of face perception: inversion effect, identity after-effect, matching across viewpoints, and interactive processing of parts. They found that preferential processing of information around the horizontal orientation provides a significant account of the behavioral measures of face processing. While the invariant aspects of faces encode face identity, the changeable aspects of faces construct emotional expressions (Haxby, Hoffman, & Gobbini, 2000). It has been suggested that separate functional and neural pathways are involved in the perception of invariant aspects of faces and of changeable aspects of faces (Bruce & Young, 1986; Hasselmo, Rolls, & Baylis, 1989; Haxby et al., 2000; Winston, Henson, Fine-Goulden, & Dolan, 2004), implying that identifying faces and categorizing facial expressions could depend on different input information. On the other hand, both face identification and facial expression categorization have been shown to rely on the configural information of facial components (Calder et al., 2000; Leder, Candrian, Huber, & Bruce, 2001). Therefore, it remains unclear whether the spatial information most crucial for categorizing facial expressions is the same as that for recognizing face identities.

A recent study utilized orientation bubbles to reveal the diagnostic information for facial expressions and found a strong link between the horizontal information and the successful categorization of several facial expressions (anger, disgust, fear, happy and sad) but not for the surprise expression (Duncan et al., 2017). These authors further showed that individual differences in the reliance of horizontal information were best predicted by the utilization of eye region alone. However, facial regions other than the eyes have been shown to be important for expression categorization. In fact, Smith et al (2005) showed that the facial regions diagnostic of a certain emotion expression are different for different expressions and share very little overlapping in their locations on a face image. Also, by examining only happy and sad expressions, Huynh and Balas (2014) found that the magnitude of the preference of horizontal orientation (compared to vertical) can be modulated by factors such as mouth openness.

In this study, we systematically evaluated the dependency of facial expression categorization on the orientation of spatial information. Specifically, we examined how the performance for recognizing facial expression depends on information restricted to different orientation bands. We asked whether categorizing facial expressions shows a similar orientation dependency on spatial information as that for recognizing face identities, i.e. primarily the horizontal structures. In addition, we examined the confusion patterns among different facial expressions for different filter orientations, and investigated how local facial regions (differences between facial expressions at the pixel level) may contribute to the categorization performance.

# Methods

## Subjects

Fifteen subjects (eight females and seven males) with normal or corrected-to-normal vision, aged between 18 and 39 years, participated in the study. All subjects were naïve to the purpose of the experiment, and performed the task binocularly. The research was conducted in accordance with the Declaration of Helsinki. Prior to the commencement of data collection, every subject signed a consent form approved by the Institutional Review Board at the University of California, Berkeley.

## Apparatus and Stimuli

We used custom-written software written in MATLAB (version 7.7.0) and Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) to control the experiment using a Macintosh computer (MacBook 5.1). Stimuli were presented on a gamma-corrected SONY color graphic display (model: Multiscan E540), at a pixel resolution of 1280 × 1024 (dimensions: 39.3 cm × 29.4 cm) and a refresh rate of 75 Hz.

Four facial expressions were tested: angry, fearful, happy and sad. We selected stimuli from the NimStim Set of Facial Expressions, a standardized database of naturally posed photographs of professional actors (Tottenham et al., 2009). As shown by Huynh and Balas (2014), the openness of the mouth can influence the emotion-dependent reliance on horizontally orientated face information. To examine the effect of filter orientation without the possible interfering effect of mouth openness, only closed-mouth versions were used in the study. To ensure none of the subjects viewed the same image more than once, we generated more test faces by morphing (Abrosoft FantaMorph 4 Deluxe) between two persons (of the same gender) with the same facial expression (a total of 118 source images were used). There were a total of 140 different faces (morphed and original) obtained for each facial expression, with 55 female faces and 85 male faces. To create a morphing image, about 100 key dots were placed on the major elements (such as eyes, eyebrows, nose, mouth, the outline of the face, and creases induced by facial expressions) of both source images. Each key point on one face image was automatically matched to its corresponding key point on the other face. The two source face images were then linearly interpolated by the software to produce a morphed image. Only one morphing level, 50%, was used so that the facial features from both faces were equally presented. Additionally, for each image (morphed and original), two reference points were defined, one at the center of the mouth and the other at the midpoint between the eyes. Rotation was then made to each image until the two reference points fell on a vertical line. The mean distance between the two reference points was 153 pixels for happy, 167 pixels for sad, 162 pixels for angry, and 167 pixels for fearful.

An orientation filter (wrapped Gaussian distributions with a bandwidth $\sigma = 23°$) was applied to restrict information contained in the stimuli, with the center of the filter ranged from 0° (horizontal) to 150° in steps of 30°, as in Yu and Chung (2011). For each filtered condition, information within the filter orientation ± the bandwidth was retained, and the rest was filtered out. All face images were converted to gray scale and cropped to an oval shape (minor and major diameters are 273 and 405 pixels). Across all conditions, images were normalized to equate the root mean square (RMS) contrast (0.12) and luminance (0.5). Stimuli were presented on a gray background (29 cd/$m^2$). Subjects were tested binocularly. At our viewing distance of 40 cm, the angular subtense of the images was 8° horizontally and 11.9° vertically. Figure 1 shows examples of the four facial expressions in the unfiltered and the six filtered conditions. Accuracy for recognizing facial expressions filtered with each of these filters, as well as for the unfiltered condition, was measured.

### Procedures

There were a total of 28 conditions (four facial expressions × seven filter orientations). Twenty trials per filter orientation were tested for each facial expression and each subject. For each subject, trials were divided into four blocks with 140 trials per block (testing conditions were completely randomized within each block). No subject viewed the same image more than once. Prior to testing, each subject completed a practice block using a different set of face images to familiarize themselves with the task.

Before each trial, a white fixation dot was presented at the center of the display. Subjects were instructed to press a mouse button to initiate a trial. Each face image was presented for 53ms, which was selected based on pilot data to avoid any ceiling or floor effect in performance. Immediately after the stimulus disappeared, a white-noise post-mask was presented for 500ms, followed by a response screen displaying four choices in words—angry, fearful, happy and sad. Using the mouse, subjects selected the response for each trial. Figure 2 illustrates a schematic diagram of the experimental paradigm.

## Results

The proportion correct of recognition, averaged across the 15 subjects, was plotted as a function of the orientation of the spatial filter for each facial expression in Figure 3. Given that the task was a four alternative forced choice task, the chance performance is 0.25. Values lower than 0.25, such as the accuracy of recognizing angry faces filtered along the vertical orientation, may be accounted for by sampling error or response bias.

With the exception of fearful expression, subjects' proportion-correct performance demonstrated a tuning relationship with respect to the orientation of filter. In other words, the orientation of the spatial information contained within the stimuli is an important determinant of subjects' performance, at least for the angry, happy and sad expressions.

What about fearful expression? Subjects' performance toward face images containing fearful expressions appeared to be invariant with the orientation of filter — subjects' proportion-correct performance was between 0.66 and 0.82 across all filter orientation for stimuli with the fearful expression. The main difference between the pattern of results for the fearful expression versus other expressions is that the proportion correct performance was much higher for filter orientations close to vertical for the fearful expression than for other expressions. One explanation for this finding is that subjects often chose "fearful" as their responses when the stimuli contained only the information close to the vertical orientation. In the following, we examined the stimulus-response confusion matrix, as well as the miss and false-alarm rate. These analyses reveal a fuller picture of how subjects' responses depend on the spatial information within the face stimuli.

Figure 4 shows the confusion matrix with data accumulated across all subjects. The rows are stimuli (targets presented) and the columns are subject responses. Each cell shows the proportion of response X given a target Y. The diagonal elements show the proportion correct when a given target was presented (hit). The off-diagonal elements show the pattern of confusions (false alarms in columns and misses in rows). We also constructed the average face images based on the stimuli presented in the corresponding cells (see Figure A1).

In general, subjects showed high performance when recognizing facial expressions for unfiltered face images or face images filtered at orientations near the horizontal (i.e., −30°, 0°, and 30°). Happy expression seemed to be the most resistant to information loss due to filtering across orientations except when the stimuli were filtered along the vertical orientation (i.e., stimuli containing information near the vertical orientation only). As the filter orientation approached vertical, more misses and false alarms occurred, regardless of the facial expressions. Recognition of fearful expression was least affected by the filter orientation near the vertical (also see Figure 3), which can be accounted for by subjects' strong bias toward classifying an expression as fearful for stimuli filtered with the vertical orientation. We also observed that subjects tended to produce many false alarms for sad expression regardless of the filter orientations. Overall, the miss rate was especially high for angry and sad expressions for the filtered conditions. As will be shown in the discussion, the information content conveyed at the pixel level of the stimulus images can partially account for the variability of the subjects' response (Figures 7, 8 and 9).

To examine the strength of signal relative to noise, we transformed accuracy data into d-prime ($d'$) values. $d'$, indicating the ability to distinguish target-present from target-absent, was calculated for each combination of facial expression and filter orientation using the equation $d' = \phi^{-1}(H) - \phi^{-1}(FA)$, where $\phi^{-1}(H)$ and $\phi^{-1}(FA)$ denote the $z$ scores (i.e. the inverse Gaussian distribution) of the hit rate and the false-alarm rate, respectively.[1] We found higher $d'$ value for happy expression (4.52±0.08 (SE)) than for angry, fearful and sad expressions (3.32±0.14, 3.44±0.14, and 3.22±0.16) for the unfiltered condition ($ps \leq$

[1]Due to the lack of a widely accepted method for calculating d′ for multiple-alternative forced-choice situations, we decided to calculate d′ using the standard formula for a two-alternative forced-choice task. Our task has four alternative choices. When performing d′ calculation for each target facial expression, we assembled the other three expressions into the same group (target-absent group). Although the validity of this approach may be debatable, we would like to include the analyses and results here as an alternative or reference. The same is true for the calculation of response bias, *c*.

0.0005). Normalized $d'$ was then calculated by subtracting the $d'$ of the unfiltered condition from the $d'$ of each filtered condition. This was performed for each facial expression and subject. Based on the same hit and false-alarm rates, we also computed response bias, $c = -(\phi^{-1}(H) + \phi^{-1}(FA))/2$, the distance between the criterion and the neutral point ($c_{neutral\ point} = 0$). A large absolute value of $c$ signifies that subject biases towards one type of response regardless of the stimulus (Stanislaw & Todorov, 1999).

Since the testing trials were divided into four blocks, we assessed the possible incidental learning or fatigue (i.e. whether normalized $d'$ changes across blocks) by comparing the four testing blocks. Repeated measures ANOVA (4 facial expressions × 7 filter orientations × 4 testing blocks) were employed. Neither the main nor the interaction effects involving testing block were significant, justifying our pooling of data over the four blocks for analysis.

Figure 5 plots normalized $d'$ (averaged across subjects) as a function of the orientation of filter for each facial expression. Clearly, a major difference between Figure 5 and Figure 3 is that we observe a tuning function for all four facial expressions, including the fearful one. The change in the shape of the data (from a flat function in Figure 3 to a tuning function in Figure 5) is mainly due to the fact that we examined both the target-present and target-absent trials (both signal and noise).

As demonstrated in Figure 5, there was an effect of filter orientation when comparing all six filtered conditions ($F(5,70) = 131.04$, $p < 0.0005$) but not when considering only the three orientations near the horizontal (i.e., −30°, 0°, and 30°) or only the −60° and 60° filtered conditions. For the near horizontal orientations, all except happy expression showed significantly lower performance compared to the unfiltered condition ($ps \leq 0.002$). Normalized $d'$ was similar for angry, fearful and happy expressions, but lower for sad expression ($F(3,42) = 8.66$, $p < 0.0005$). In other words, subjects could not recognize sad expression as well as the other three at these three filtered conditions. At the −60° and 60° filtered conditions, subjects performed equally well on the four expressions. At the vertical filter orientation, recognition performance varied across expressions ($F(3,42) = 6.96$, $p = 0.001$), and was worst for happy expression ($ps < 0.05$).

Response bias is plotted as a function of the orientation of the filter for each facial expression in Figure 6. For the unfiltered condition (dashed lines), there was no response bias for happy expression, a small preference toward responding “yes” (value < 0) for sad expression ($c = -0.19$, $t(14) = -2.90$, $p = 0.01$), and some biases toward the “no” response (values > 0) for angry ($c = 0.48$, $t(14) = 5.43$, $p < 0.0005$) and fearful expressions ($c = 0.48$, $t(14) = 6.19$, $p < 0.0005$). When assessing the response biases occurred under the filtered conditions, we examined each bias relative to its corresponding baseline (i.e. the bias under the unfiltered condition). Interestingly, for all filtered conditions, sad expression consistently showed a bias toward “no” response ($p \leq 0.003$), which explains the high miss rates observed under the six filtered conditions in Figure 4. Across the four facial expressions, the largest biases occurred at the 90° filtered condition—relative to the unfiltered condition, bias was strongly toward “yes” response for fearful expression and “no” for the other three expressions. These results are consistent with, and can account for, the observation in Figure

3 that the proportion-correct for categorizing face images filtered with the 90° filter was much higher for fearful expression than for other facial expressions; as well as the observation in Figure 4 that over half of the responses were toward fearful expression for images filtered with the 90° filter.

## Discussion

Consistent with previous work (e.g. Huynh & Balas, 2014; Goffaux & Greenwood, 2016; Duncan et al, 2017), we found that the spatial information that lies around the horizontal orientation captures primary changes of facial features across expressions and is the most important information for recognizing facial expressions for young adults with normal vision. In addition, we further showed that for all four facial expressions (angry, fearful, happy and sad), recognition performance was virtually identical for filter orientations of −30°, horizontal (0°) and 30°. Beyond ±30° filter orientation, performance declined systematically as the filter orientation approached vertical.

Processing facial expression information is supported by both low and high spatial frequencies (Calder et al., 2000; Schyns & Oliva, 1999; Vuilleumier et al., 2003). Both the individual features (i.e., fine-grained information contained in high spatial-frequency components) (Ellison & Massaro, 1997; Smith et al., 2005) and the configural information (i.e., global information contained in low spatial-frequency components) play an important role in facial expression categorization (Calder et al., 2000). Our findings show that signals in the horizontal and nearby orientation channels contain more diagnostic information when compared to the other orientations, which is possibly due to the fact that the local features and the configural information in faces are most reliably reserved in the horizontal and nearby oblique orientation channels. In other words, it is not the absolute amount of energy contained within a given orientation channel that makes the signal within that orientation diagnostic; instead, it is the variations (the relative changes) of this signal across different expressions that are important.

By analyzing confusion matrices constructed based on subjects' performance, we found that the vertical filter orientation produced the largest amount of confusions among the four expressions while the reporting rate for fearful expression was substantially high (accounting for half of all responses at the vertical orientation). Our finding suggests that the information carried by the vertical and its neighboring orientations primarily contains signature for fearful expression, and has little variation across expressions, as demonstrated in Figure 7. Therefore, it is reasonable to observe the substantial biases presented at the vertical filter orientation (Figure 6). In addition, as indicated by our data (Figure 3), the diagnostic features for fearful expression are not exclusively carried by the vertical and nearby orientation channels. When the vertical information was removed (e.g., when face images were filtered at orientations near the horizontal), subjects were still able to recognize fearful expression.

Besides neutrality, facial expressions are generally classified into six basic categories: angry, fearful, happy, sad, disgusted, and surprised (Ekman & Friesen, 1975). In the present study, we tested four of the facial expressions: angry, fearful, happy and sad. It is possible that the horizontal tuning and the pattern of confusions among different facial expressions may change if we included all expressions. A recent study showed that when all six facial expressions plus neutrality were tested, a considerable degree of confusion was found between surprised and fearful expressions, and that the information around horizontal orientation is not diagnostic for surprised expression (Duncan et al., 2017).

A feed-forward neural network, EMPATH model, has been developed to model facial expression recognition (Dailey, Cottrell, Padgett, & Adolphs, 2002). Li and Cottrell (2012) used the same model to model our findings. The model consists of three layers—a set of model neurons based on the magnitude of Gabor filters, principal components generated from Gabor filter outputs to capture the distinguishing features of each facial expression, and a linear perceptron with the principal components mapped to different emotion categories. The model showed a very similar response pattern to that of our human subjects.

### Orientation selectivity for facial expression categorization vs. face identification

Facial expression and face identity represent changeable and invariant aspects of faces, respectively. Despite the suggested dissociation between the representations for processing the two types of aspects of faces (Bruce & Young, 1986; Hasselmo et al., 1989; Haxby et al., 2000; Winston et al., 2004), facial expression categorization and face identification (Dakin & Watt, 2009; Yu & Chung, 2011) showed similar dependency on information restricted to different orientation bands. To further compare the orientation sensitivity profiles for face identification and facial expression categorization, we fit our expression categorization data and face identification data from Goffaux & Greenwood (2016, data for the masked outline and upright face condition) using a Gaussian function and compare the shape and position of the profiles. We found that both facial expression processing and identity processing are strongly tuned to horizontal orientation (peaks of the curves reside within 3° from horizontal). Subjects consistently benefit most from the presence of spatial information around the horizontal orientation. The tuning bandwidth (full-width at half-maximum) is broader for the sensitivity to facial expression (93°, 94°, 98° and 95° for angry, fearful, happy and sad expressions, respectively) than for the sensitivity to facial identity (62°), which indicates that a broader range of orientations contribute more similarly to processing facial expression than processing facial identity. It has been shown that in face identification, the face inversion effect is essentially related to a decreased sensitivity to horizontal information (Goffaux & Greenwood, 2016). The modeling results by Li and Cottrell (2012) indicate that performance on categorizing facial expression is largely driven by stimulus images (information contained in stimulus image should be invariant regardless of image orientation). Although we have not tested face inversion effect on facial expression, it is possible that we may not observe the same amount of disruption on the horizontal selectivity for facial expression processing as it does for facial identity processing, that is, facial expression categorization may not be as susceptible to face inversion as face identification.

### Pixel-Level Face Image Analysis

A recent study showed that in the categorization of facial expressions, horizontal tuning is best predicted by eye diagnosticity (Duncan et al., 2017). Here, to examine whether subject’s response was driven by the energy content in a specific facial region, we performed further analysis on the stimulus images. First, we performed pixel-level stimulus comparison between each pair of facial expressions by subtracting the average face image of one expression from the average face image of another expression. The comparison was done for all facial expression pairs and all seven conditions (unfiltered, −60°, −30°, 0°, 30°, 60°, and 90° filtered conditions). Figures 7, 8 and 9 show the image subtraction for vertical and horizontal filtered conditions and the unfiltered condition, respectively. Difference was shown in all the off-diagonal pairs. The larger the difference (i.e. the larger the deviations of pixel values from the mean gray background), the easier it could be to differentiate the two facial expressions. RMS contrast was calculated to quantify the amount of information contained in each subtracted image, and was found to peak near the horizontal orientation. We also performed linear regression modeling to evaluate the contributions of different facial regions to subjects' response. Specifically, we divided each subtracted image into four regions (eye, mouth, nose, and the rest; see examples in Figure 9), and calculated RMS contrast for each region. The outcome variable was the proportion of subjects' response (i.e., the cells of confusion matrices in Figure 4). Only off-diagonal values were considered. Since we only have 12 data points (i.e., 12 off-diagonal values) for each filtered condition, we compiled the data across all filtered conditions to model the proportion of subjects' response as a function of $RMS_{eye}$, $RMS_{nose}$, $RMS_{mouth}$ and $RMS_{rest}$. By examining both the linear and two-way interaction terms in the model, we found that only the mouth and eye regions showed significant associations with the outcome measure. The model (proportion of response = $0.65 - 9.80 \times RMS_{mouth} - 7.74 \times RMS_{eye} + 113.76 \times RMS_{mouth} \times RMS_{eye}$) explains 49% ($R^2_{adj}$) of the variability of the proportion of subjects' response. In other words, 49% of the variance of subjects' response can be explained by the information content contained in the mouth and eye regions at the pixel level. The unique contribution to explaining the total variance in the proportion of response is the largest for $RMS_{mouth}$ ($sr^2$ = 0.33), followed by $RMS_{eye}$ ($sr^2$ = 0.22), and the smallest for the interaction term ($sr^2$ = 0.05).[2] The results suggest that energy contents in different facial regions have different weights in facial expression recognition. It is possible that this weight distribution varies for different tasks and/or testing conditions.

## Conclusions

In this study, we showed that for normally-sighted young adults, the spatial information around the horizontal orientation is the most important for recognizing facial expressions. It is important to ask whether similar results can be found in older adults and people with visual impairment. If crucial orientation(s) can be identified in these populations, selective enhancement of face images along the orientation(s) would possibly improve the ability of these people to recognize facial expression. When analyzing the face image stimulus on a pixel level, we further showed that human performance for categorizing facial expressions is primarily driven by the information contained around the mouth and eye regions, supporting the importance of these regions for the task of facial expression categorization.

[2]The sum of the three $sr^2$ is larger than $R^2_{adj}$ (0.49) because the two variables, $RMS_{eye}$ and $RMS_{mouth}$, are mutually enhancing (due to their negative correlation ($r = -0.24$, $p = 0.03$)), that is, each variable accounts for more of the variance in the outcome measure (the proportion of subjects' response) in the presence of the other variable than it does alone (Cohen & Cohen, 1975).

## Acknowledgments

This work was presented at the 2011 VSS conference (Yu, Chai, & Chung, 2011). The research was supported by NIH grants EY012810 and EY025658. We would like to thank Valerie Goffaux and John Greenwood for sharing their data (Goffaux & Greenwood, 2016).

## References

Balas B, Huynh C, Saville A, Schmidt J. Orientation biases for facial emotion recognition during childhood and adulthood. Journal of experimental child psychology. 2015; 140:171–183. [PubMed: 26247810]

Bruce V, Young A. Understanding face recognition. British journal of psychology. 1986; 77(3):305–327. [PubMed: 3756376]

Calder AJ, Young AW, Keane J, Dean M. Configural information in facial expression perception. Journal of Experimental Psychology: Human perception and performance. 2000; 26(2):527. [PubMed: 10811161]

Cohen J, Cohen P. Applied multiple regression/correlation for the behavioral sciences. New York, NY: Wiley; 1975.

Dailey MN, Cottrell GW, Padgett C, Adolphs R. EMPATH: A neural network that categorizes facial expressions. Journal of cognitive neuroscience. 2002; 14(8):1158–1173. [PubMed: 12495523]

Dakin SC, Watt RJ. Biological "bar codes" in human faces. Journal of vision. 2009; 9(4):2–2.

De Valois RL, Albrecht DG, Thorell LG. Spatial frequency selectivity of cells in macaque visual cortex. Vision research. 1982; 22(5):545–559. [PubMed: 7112954]

Duncan J, Gosselin F, Cobarro C, Dugas G, Blais C, Fiset D. Orientations for the successful categorization of facial expressions and their link with facial features. Journal of vision. 2017; 17(14):7–7.

Ekman P, Friesen WV. Unmasking the face: A guide to recognizing emotions from facial cues. Englewood Cliffs, NJ: Prentice Hall; 1975.

Ellison JW, Massaro DW. Featural evaluation, integration, and judgment of facial affect. Journal of Experimental Psychology: Human perception and performance. 1997; 23(1):213. [PubMed: 9090153]

Goffaux V, Dakin S. Horizontal information drives the behavioral signatures of face processing. Frontiers in Psychology. 2010; 1:143. [PubMed: 21833212]

Goffaux V, Greenwood JA. The orientation selectivity of face identification. Scientific reports. 2016; 6:34204. [PubMed: 27677359]

Hasselmo ME, Rolls ET, Baylis GC. The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey. Behavioural brain research. 1989; 32(3): 203–218. [PubMed: 2713076]

Haxby JV, Hoffman EA, Gobbini MI. The distributed human neural system for face perception. Trends in cognitive sciences. 2000; 4(6):223–233. [PubMed: 10827445]

Hubel DH, Wiesel TN. Receptive fields and functional architecture of monkey striate cortex. The Journal of physiology. 1968; 195(1):215–243. [PubMed: 4966457]

Huynh CM, Balas B. Emotion recognition (sometimes) depends on horizontal orientations. Attention, Perception, & Psychophysics. 2014; 76(5):1381–1392.

Leder H, Candrian G, Huber O, Bruce V. Configural features in the context of upright and inverted faces. Perception. 2001; 30(1):73–83. [PubMed: 11257979]

Li R, Cottrell G. A new angle on the EMPATH model: Spatial frequency orientation in recognition of facial expressions. Paper presented at the Proceedings of the Cognitive Science Society; 2012.

Schyns PG, Oliva A. Dr. Angry and Mr. Smile: When categorization flexibly modifies the perception of faces in rapid visual presentations. Cognition. 1999; 69(3):243–265. [PubMed: 10193048]

Smith ML, Cottrell GW, Gosselin F, Schyns PG. Transmitting and decoding facial expressions. Psychological Science. 2005; 16(3):184–189. [PubMed: 15733197]

Stanislaw H, Todorov N. Calculation of signal detection theory measures. Behavior Research Methods, Instruments, & Computers. 1999; 31(1):137–149.

Vuilleumier P, Armony JL, Driver J, Dolan RJ. Distinct spatial frequency sensitivities for processing faces and emotional expressions. Nature neuroscience. 2003; 6(6):624–631. [PubMed: 12740580]

Winston JS, Henson R, Fine-Goulden MR, Dolan RJ. fMRI-adaptation reveals dissociable neural representations of identity and expression in face perception. Journal of neurophysiology. 2004; 92(3):1830–1839. [PubMed: 15115795]

Yu D, Chai A, Chung S. Orientation information in encoding facial expressions. Journal of Vision. 2011; 11(11):604–604.

Yu D, Chung ST. Critical orientation for face identification in central vision loss. Optometry and vision science: official publication of the American Academy of Optometry. 2011; 88(6):724. [PubMed: 21399554]

## Appendix A

Figure A1 shows the average confusion matrix with data accumulated across all subjects and the six filtered conditions. In each cell, an average face image was constructed based on all the stimulus faces presented in the corresponding cell.

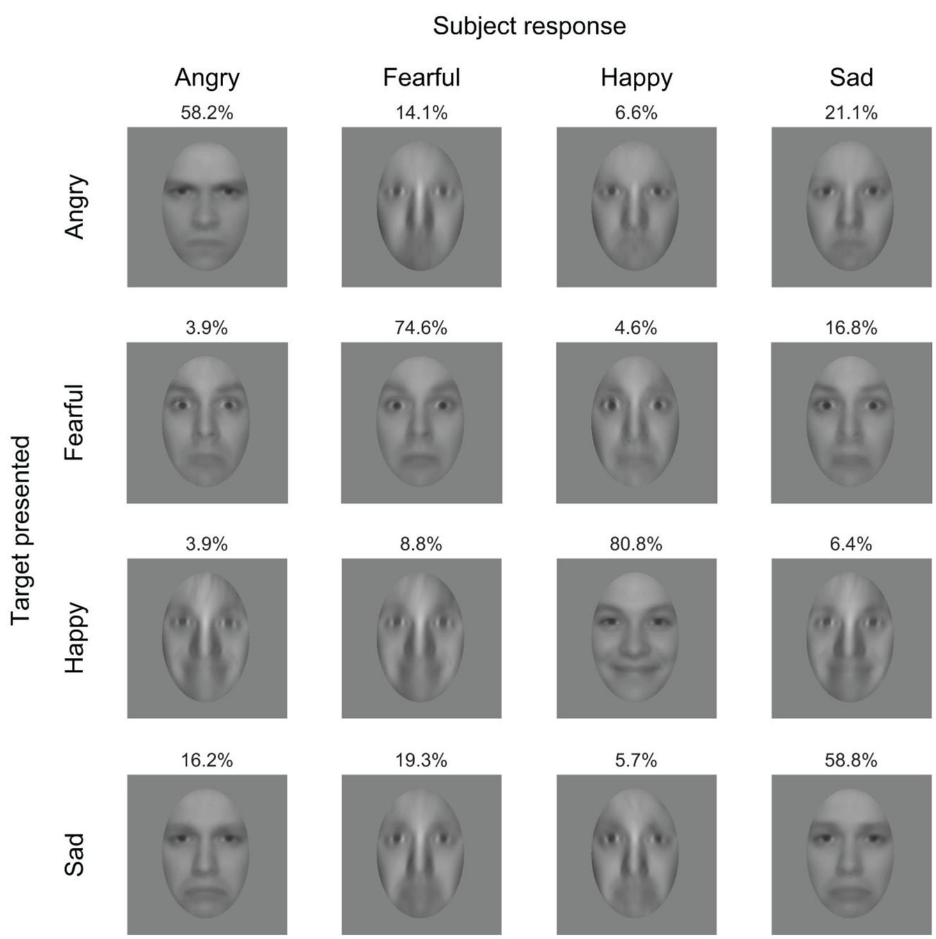


**Figure A1.**

The average confusion matrix with data accumulated across all subjects and the six filtered conditions. The rows are stimuli (targets presented) and the columns are subject response.

Each cell shows the proportion of response X given a target Y, and the face image which was constructed based on all the stimulus faces presented in the corresponding cell.

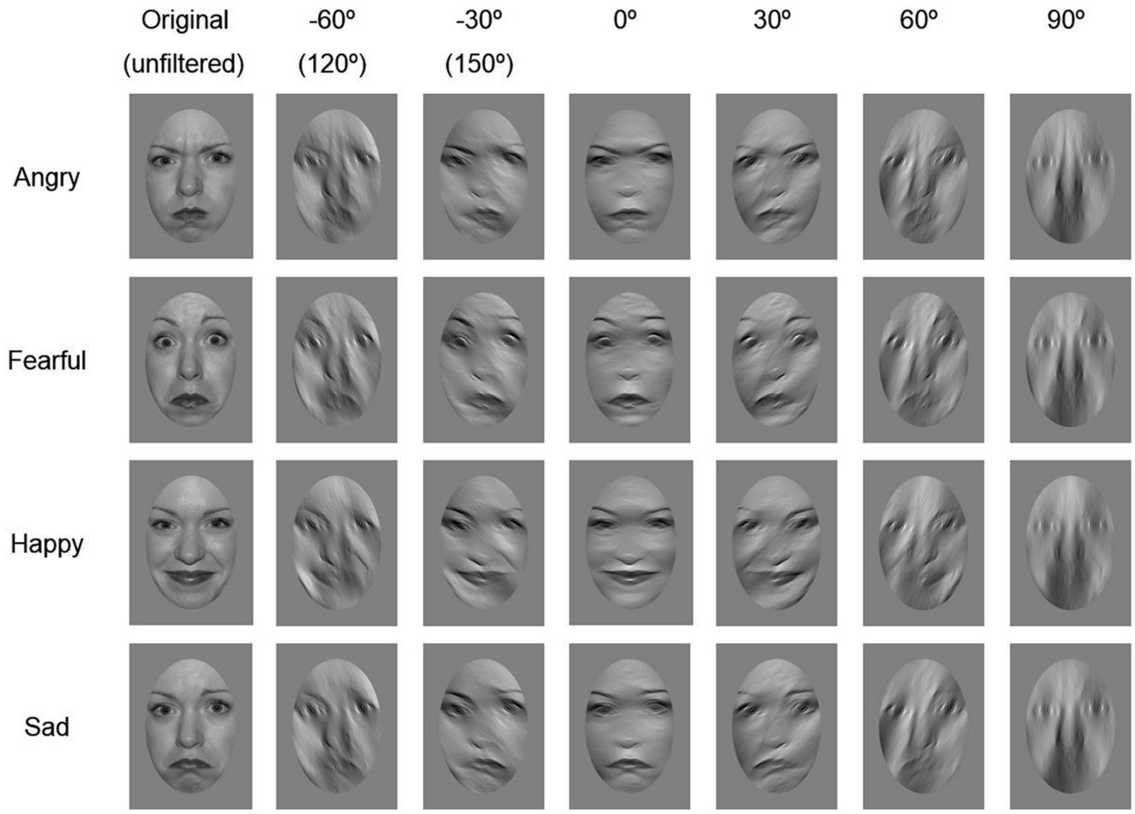


**Figure 1.**

Examples of the four facial expressions—angry, fearful, happy and sad—in an unfiltered and the six filtered conditions. For the filtered conditions, an orientation filter was applied to images along one of the six orientations (−60° (120°), −30° (150°), 0°, 30°, 60°, 90°) with a bandwidth of 23°. Image RMS contrast and luminance were normalized.

Author Manuscript
Author Manuscript
Author Manuscript



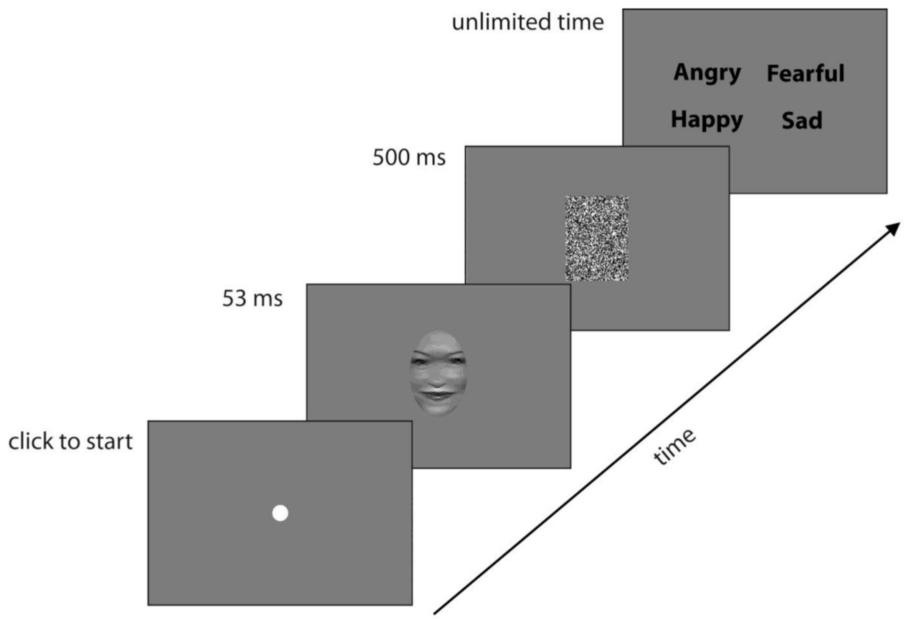


**Figure 2.**
A schematic diagram of the experimental paradigm.

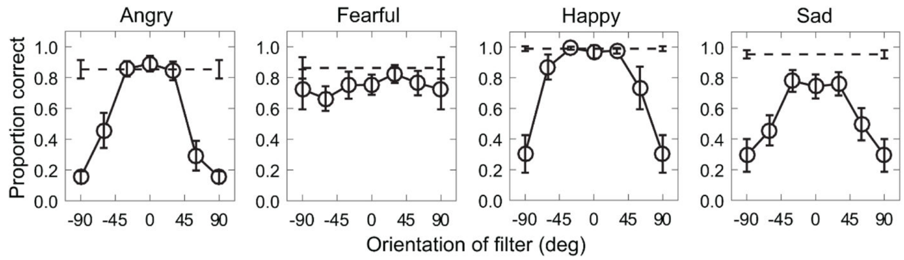


**Figure 3.**

Proportion correct of recognizing facial expression as a function of filter orientation. Both 90° and −90° refer to the vertical filter orientation. Dashed lines denote the performance for recognizing unfiltered images. Error bars indicate 95% confidence intervals.

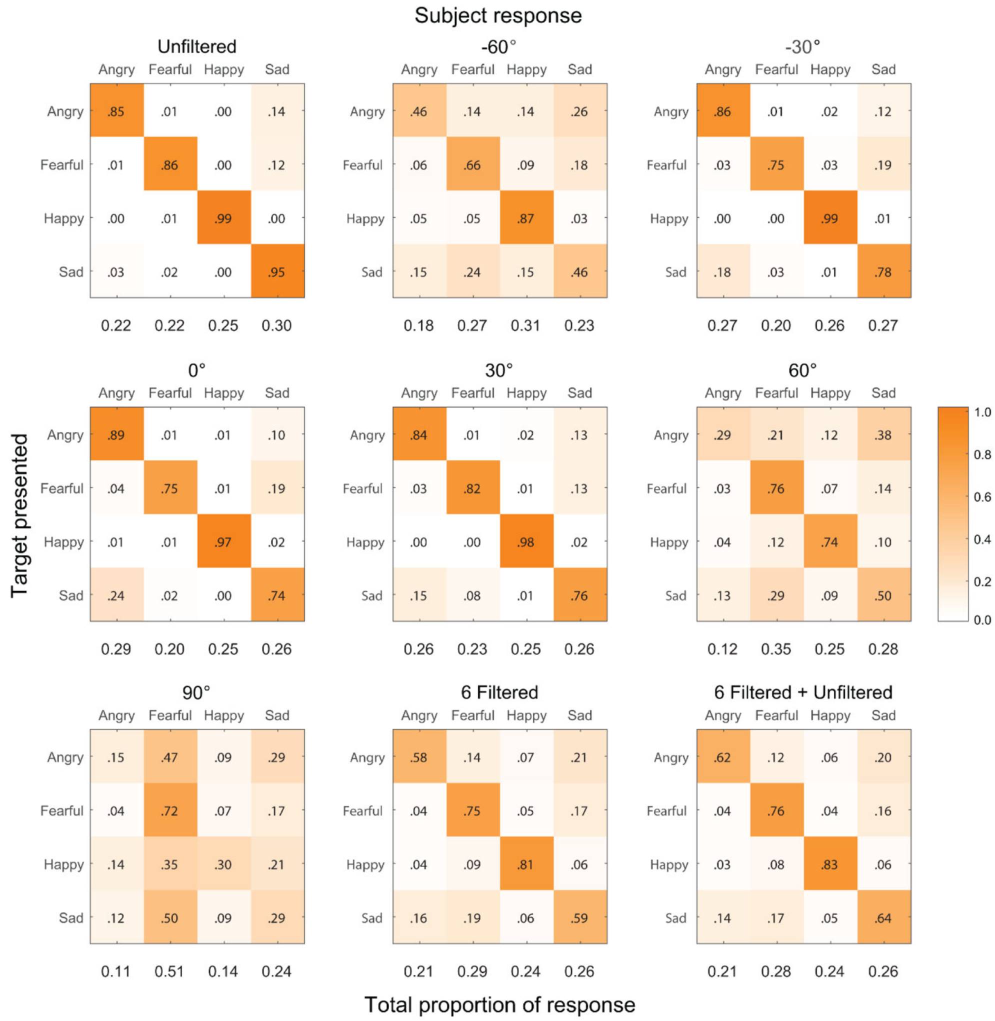


**Figure 4.**

Confusion matrices with data accumulated across all subjects for different conditions (unfiltered, −60°, −30°, 0°, 30°, 60°, and 90° filtered conditions). We also constructed the average confusion matrices across the six filtered conditions ("6 Filtered") and across all conditions ("6 Filtered + Unfiltered"). In each matrix, the rows are stimuli (targets presented) and the columns are subject response. Each cell shows the proportion of response X given a target Y. The diagonal elements show the proportion correct when a given target was presented. The off-diagonal elements show the pattern of confusions (false alarms in columns and misses in rows). The value below each column represents the total proportion of response X for the given filtered condition. The colored key represents the different accuracy.

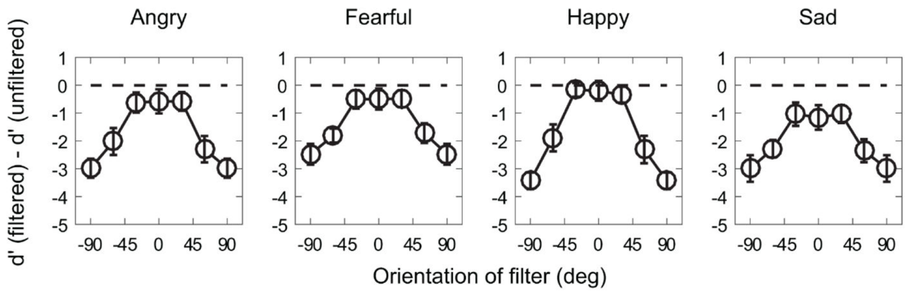


**Figure 5.**

Normalized $d'$ ($d'$ for each filtered condition subtracted by the $d'$ of the unfiltered condition) as a function of filter orientation for each facial expression. Both 90° and −90° refer to the vertical filter orientation. Dashed lines denote a normalized $d'$ of zero (i.e., no difference between performance for a filtered condition and the unfiltered condition). Error bars indicate 95% confidence intervals.

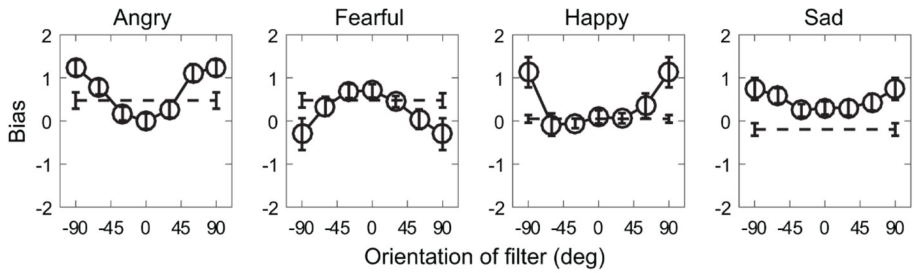


**Figure 6.**

Response bias, *c*, as a function of filter orientation for each facial expression. Both 90° and −90° refer to the vertical filter orientation. Dashed lines denote *c* of the unfiltered condition. Negative value of *c* indicates a bias toward responding "yes", whereas positive value indicates a preference toward the "no" response. Error bars indicate 95% confidence intervals.

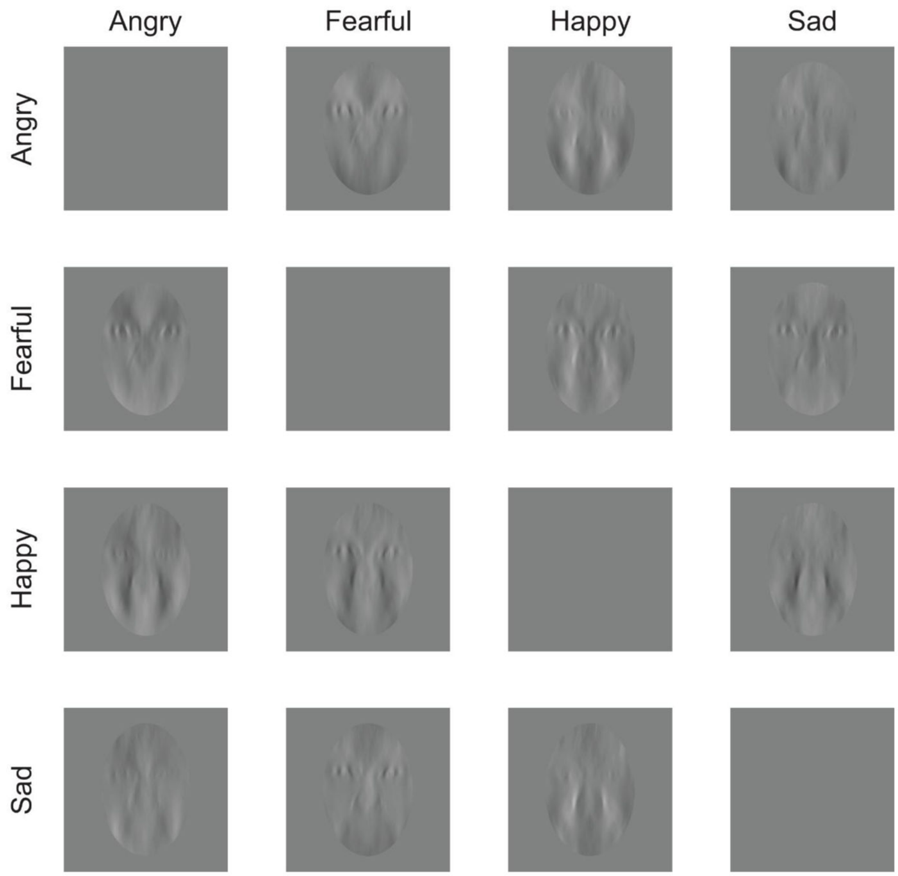


**Figure 7.**

Image subtraction (pixel by pixel) between each pair of facial expressions for the vertical filtered condition.

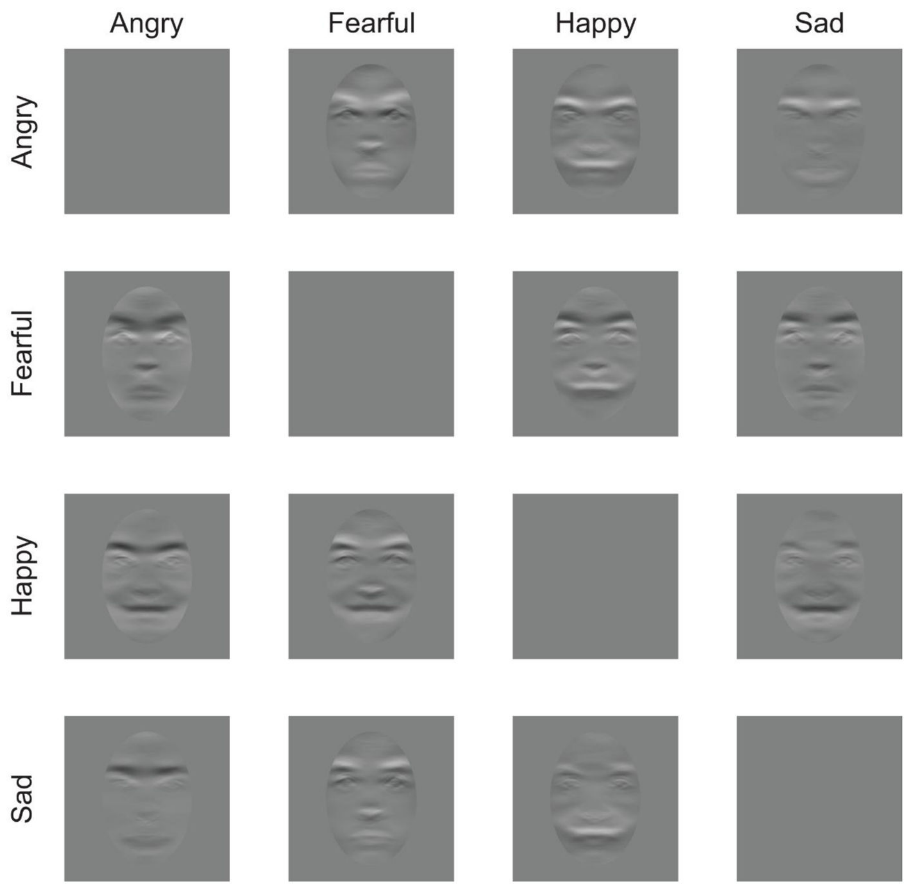


**Figure 8.**
Image subtraction (pixel by pixel) between each pair of facial expressions for the horizontal filtered condition.

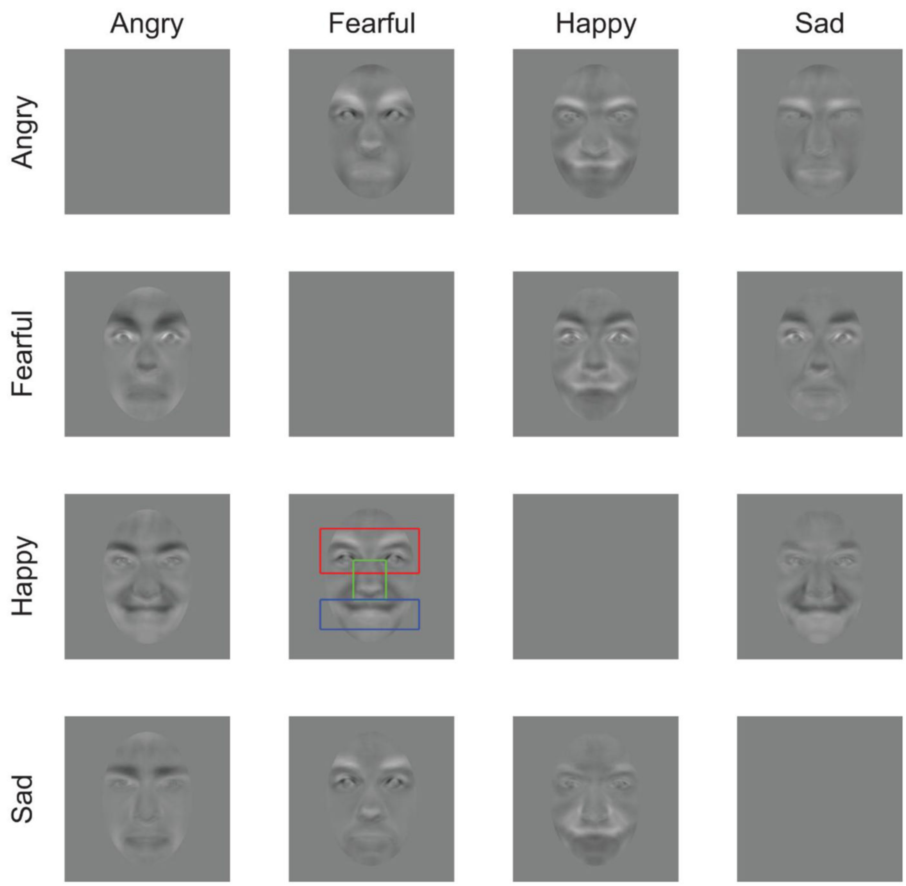


**Figure 9.**
Image subtraction (pixel by pixel) between each pair of facial expressions for the unfiltered condition. In the happy-fearful subtracted image, we showed an example on how the eye, nose and mouth regions are defined. The eye region is bounded by the red box. The mouth region is defined by the blue box. The nose region is bounded by green lines. All the leftover regions are categorized as "the rest".