Vision Research 50 (2010) 1814–1823

Contents lists available at ScienceDirect

Vision Research

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

The contribution of the upper and lower face in happy and sad facial expression classification

Mei-Yen Chen [a], Chien-Chung Chen [a,b,*]

[a] Department of Psychology, National Taiwan University, Taipei, Taiwan
[b] Center for Neurobiology and Cognitive Science, National Taiwan University, Taipei, Taiwan

ARTICLE INFO

Article history:
Received 24 December 2009
Received in revised form 4 June 2010

Keywords:
Multidimensional signal detection
Classification
Eccentricity
Holistic processing
General recognition theory

ABSTRACT

We used a happy/sad classification task and a psychophysical model to study the tuning properties of facial expression processors across viewing conditions. Using morphed faces, in this study we measured the extent to which classification of facial expressions depends on the intensity of a particular expression on either the upper or lower face. In the fovea, the upper and lower parts of the test image were either aligned or had a lateral shift of 44′ visual angle. In the periphery, the aligned test image was placed at a 6° visual angle to the left of the fixation. Observers were asked to classify a test image of a facial expression as happy or sad. We discovered that the alignment of the upper and lower halves of the face had no effect on happy/sad classification in the fovea, suggesting that the classification of facial expressions is an analytic process. The model also showed no interaction between the two halves of the face in foveal facial expression classification. In addition, the poor performance of observers in recognizing happiness in the periphery manifests a computational complexity, suggesting a model in which the happy-face processor relies on both facial features and the interaction between them to recognize happiness in the periphery.

1. Introduction

The ability to correctly interpret one another's emotions is an essential part of social interaction. After generations of evolution, facial expressions, which are combinations of the movements and states of facial muscles, have become perhaps the most efficient means of communicating emotion (Darwin, 1872/1965; Ekman & Friesen, 1975). For the successful communication of emotion through facial expressions, not only does the sender need to move his facial muscles correctly, but the visual system of the perceiver also needs to decode the apparent muscle movement in the right way. That is, the visual system has to extract important information from the states of the sender's facial muscle movements and turn it into a percept of the emotion represented by the facial expression (Etcoff & Magee, 1992; Goren & Wilson, 2006; Young et al., 1997).

However, in the literature, controversy remains as to what information is crucial in order for the visual system to classify a facial expression. The results from the reverse correlation technique (Kontsevich & Tyler, 2004; Mangini & Biederman, 2004) and the *Bubbles* method (Gosselin & Schyns, 2001; Smith, Cottrell, Gosselin, & Schyns, 2005) do not agree. Kontsevich and Tyler (2004) applied white noise to the lower and upper parts of Leonardo's painting of the Mona Lisa. They discovered that the perceived expression of the Mona Lisa was affected by noise applied to the lower part of the face but not the upper part. Further analysis showed that noise applied only to the corners of the mouth could affect the perceived expression. Using the *Bubbles* method, Smith et al. (2005) had their observers identify the expressions of faces that were partially visible through Gaussian apertures. By summing the stimuli weighted by the observers' responses, they were able to identify features that are critical to expression identification. These critical features were very local and varied from one expression to another. However, there is an inconsistency between the results derived from the reverse correlation method (Kontsevich & Tyler, 2004) and the *Bubbles* method (Smith et al., 2005). The former suggests that happy/sad discrimination can be achieved by a small change at the corners of the mouth alone. In contrast, the latter indicates that the identification of a sad face requires both the mouth and the areas around the eyes, even though the identification of a happy face can be achieved through the mouth alone. This inconsistency could result from various reasons, from the nature of the noise (e.g., additive versus multiplicative) to the difference in the underlying hypothesis of the two methods (Gosselin & Schyns, 2004; Murray & Gold, 2004).

In addition, the contribution of each facial feature to the perception of facial expressions may depend on context. Kontsevich and Tyler (2004) discovered the long-range effect that the perception

* Corresponding author at: Department of Psychology, National Taiwan University, 1, Sec. 4, Roosevelt Rd., Taipei 106, Taiwan.
E-mail address: c3chen@ntu.edu.tw (C.-C. Chen).

0042-6989/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2010.06.002

of expression in the eyes was dependent on the mouth. This context dependency was also illustrated in the composite face paradigm (Calder, Young, Keane, & Dean, 2000; Young, Hellawell, & Hay, 1987), where the observers were presented with faces whose upper and lower halves could have the same, or different, expressions. Calder et al. (2000) showed that it took more time to identify the facial expression on one part of the composite face when the other part showed a different expression than when they were the same. This effect disappeared when the two parts were misaligned. Notice that they used this result as evidence of holistic processing of expression. While this reference to holistic processing is not surprising, given that there is ample evidence of holistic processing of face perception (Carey & Diamond, 1977; Chen, Kao, & Tyler, 2007; Fantz, 1961; Leder & Bruce, 2000; O'Toole, Deffenbacher, Valentin, & Abdi, 1994; Valentine, 1988), Calder et al. (2000) may overstate their case. In their experiment, the participants only attended to one part of a face in a given trial. Hence, what they studied was how the response to one feature can be affected by its context, rather than facial expression as a whole. This kind of context effect has been demonstrated with a variety of stimuli. For instance, the detection threshold of a Gabor patch can be affected by the presence of another Gabor patch projected onto another part of the retina (Chen & Tyler, 2001, 2008; Polat & Sagi, 1993). This context effect depends heavily on the relative distance and location (e.g., side or collinear) of the target and context Gabor patches (Chen & Tyler, 2008). In general, one would not consider the effect of one Gabor patch on the other as an evidence of holistic processing. Thus, one should restrain from calling the effect of face feature on the other as an evidence of holistic processing. The result of Calder et al. (2000) may just suggest that an interaction exists between the upper and lower face, but does not necessarily prove holistic processing. Thus, to better study this issue of holistic versus analytic processing, we instructed the observers to attend to the whole face. This made it more likely that the observer would use all available information on a face.

Whatever information the visual system might use to decode a facial expression, there is evidence suggesting that perception of the basic emotions may be categorical (Ekman & Friesen, 1975; Etcoff & Magee, 1992; Young et al., 1997). That is, the visual system is relatively insensitive to a change in the image within a categorical boundary, but is sensitive to a change across that boundary (Angeli, Davidoff, & Valentine, 2008; Beale & Keil, 1995; Etcoff & Magee, 1992; Rotshtein, Henson, Treves, Driver, & Dolan, 2004). This categorical perception implies that the visual system contains certain mechanisms, or processors, that respond to particular expressions better than others. In this study, we attempt to resolve the controversy of holistic versus analytic process in the perception of facial expressions by modeling the tuning properties of facial expression processors.

Our approach was to study how the classification of facial expressions can be affected by changes in intensity of a particular expression in either the upper or the lower face. Our experiment paradigm used image manipulation in a way similar to the composite face paradigm (Calder et al., 2000; Young et al., 1987). We morphed the upper and lower halves of the face from sadness to happiness, and randomly combined these half images into test faces. We then measured the probability of each image being classified as happiness. We applied a model based on the multidimensional signal detection theory (Ashby, 1992) to estimate the tuning properties of the expression processors. The assumption was that the probability of an observer classifying a face into a particular expression category would depend on the relative intensities between facial expression processors. The more sensitive a processor is to a particular feature, the greater its response to that feature would be. Accordingly, an observer would be more likely to put the input facial image into the facial expression category represented by that processor. In addition, this model provides an estimation of response covariance between features in a processor. If there is no interaction between features in determining the facial expression, the response covariance matrix should be a diagonal matrix (see the Section 2 for details). Hence, this paradigm not only allows us to identify the contributions of the upper and lower face but also any possible interaction between them. Such information may resolve the controversy in the literature (Kontsevich & Tyler, 2004; Smith et al., 2005) regarding the role of eyes and eyebrows in happy/sad classification, as discussed above.

With this paradigm, we also explored other issues regarding perception of facial expressions. Here, we focused on two issues. The first was whether perception of facial expression is holistic. There is much evidence for holistic processing in face identification and recognition (Yin, 1969; Carey & Diamond, 1977). That is, face recognition may be based on an analysis of the spatial relationship between facial features rather than the properties of the features themselves (Carey & Diamond, 1977; Chen et al., 2007; Fantz, 1961; Leder & Bruce, 2000; O'Toole, Deffenbacher, Valentin, & Abdi, 1994; Valentine, 1988). However, we cannot conclude that the processing of facial expressions is holistic; there is also evidence that localized face features are essential for face recognition. Hence, face recognition may be an analytic process (Bradshaw & Wallace, 1971; Konar, Bennett, & Sekuler, 2010; Sekuler, Gaspar, Gold, & Bennett, 2004). It is suggested that holistic processing may be interrupted if the upper and the lower parts of a face are not aligned (Calder et al., 2000). Hence, if the classification of facial expressions is indeed holistic, we should observe different classification responses for aligned and misaligned faces. In addition, the interaction terms in the response covariance matrix should be nonzero for aligned faces and should be about zero for misaligned faces.

The second issue we were interested in was whether there are different mechanisms of facial expression perception at different eccentricities. Neuroimaging results have shown that a face with negative emotions presented in the periphery produces a much greater response in the amygdala than one with positive emotions (Silvert et al., 2007; Vuilleumier, Armony, Driver, & Dolan, 2001). This effect was not found for stimuli presented in the fovea. On the other hand, Goren and Wilson (2006) showed that it was more difficult to recognize negative emotions when a face was presented in the periphery than in the fovea. However, for happy faces, recognition performance was about the same at different eccentricities. These results suggest that different mechanisms might be involved in analyzing facial expressions in the fovea than in the periphery. If this is the case, then, we should observe a change in the covariance matrix.

2. Methods

2.1. Apparatus

The visual stimuli were presented on a 17-in. LCD monitor controlled by a PC-compatible computer. The LCD monitor was calibrated with an International Light PRS380 radiometer and Lightmouse photometer (Tyler & McBride, 1997) for both luminance and chromaticity. The LCD monitor had a $1024(H) \times 768(V)$ spatial resolution and a 75 Hz temporal refresh rate. The viewing distance was 127.5 cm. At this distance, each pixel occupied one minute of visual angle. A chin rest was used to restrain observers' head movements. The experiment software was written in MATLAB with the Psychophysics Toolbox (Brainard, 1997).

2.2. Stimuli

The grayscale images of happiness and sadness, using the same model, were chosen from Ekman's POFA (Pictures of Facial Affect)

database (Ekman & Friesen, 1976). The images were morphed, or interpolated, between a happy and sad face of the same person by FantaMorph 4.0 (Abrosoft, www.fantamorph.com). The morphing levels were defined by the proportion of the 'happy' face used in the morphed image. The morphing levels we used were 0%, 20%, 40%, 50%, 60%, 80%, and 100% (see Fig. 1). There were images of four models (two male and two female) in the stimulus set. All of the expressions on the chosen images could be correctly recognized by local observers (Cho, 2001).

For each morphed image, we first normalized the contrast and the luminance of each image by equating the mean and the standard deviation of the luminance distribution of each image. We then cut the normalized image at the midpoint to separate it into upper and lower halves, and randomly joined the upper and lower halves from different morphing levels to create the test images. For each test image, the upper and lower halves were always from the same model. Combining seven morphing levels for both the upper and lower face, we had 49 test images for each model. Those images constituted the stimulus space, as shown in Fig. 1.

2.3. Procedure and observers

There were three test conditions in the experiment: aligned images in the fovea and in the periphery, and misaligned images in the fovea. In the foveal viewing conditions, the images were centered on a fixation point placed at the center of the display. In the peripheral viewing condition, the images were placed at six degrees to the left of the fixation point. It has been shown that performance in facial expression recognition is the same regardless of whether the images are placed to the left or the right of the visual field (Goren & Wilson, 2006).

The image sizes were determined by the measured cortical magnification factor for facial expression classification (Appendix A). The distance between the midpoint of the two eyes and the center of the mouth was 27 min for the foveal and 103 min for the peripheral viewing condition, as determined by the cortical magnification factor experiment. Accordingly, the image size was 58 × 44 min in the fovea and 216 × 162 min in the periphery. Notice that, our foveal stimuli were smaller than most studies on facial expression perception. Therefore, due to the limit resolution on the display, our observer might not access details of facial features for their tasks. In the foveal viewing condition, the upper and lower parts of the test image were either aligned or had a lateral shift of 44 min visual angle. Misaligned faces served as a test of the effect of spatial configuration on facial features in the fovea. Regardless of viewing conditions, the duration of each test image was 200 ms. In the experiment, there were 40 blocks for every observer. Within each block, the 49 test images were presented randomly.

A training session was given to prepare all observers for consistent identification of happy and sad faces before the experimental session. In the training session, observers practiced discriminating happy (100% morph level) and sad (0% morph level) faces in both foveal and peripheral viewings. Auditory feedback was delivered until their performance reached at least 90% correct. In the experimental session, these well-trained observers were asked to press a button to indicate whether a presented image was happy or sad while staring at a fixation point.

Four observers (two male and two female, all in their early 20s) participated in the experiment. One observer was the author and the others were paid observers, who were naïve to the purpose of the experiment. All observers had experience in psychophysics experiments and had normal or corrected-to-normal visual acuity (20/20).

2.4. Data analysis

We used the multidimensional signal detection theory (Ashby, 1992) to model our result. Our stimuli consisted of images varied in a two dimensional space—the upper and lower face. In our classification paradigm, the task of the observer was equivalent to dividing this space into two parts. Suppose that a sample, x, whose coordinates (a, b) denote the a% morphing level in the lower face and b% morphing level in the upper face, produced a response f_H in the happy-face processor and f_S in the sad-face processor. The response in each processor follows a bivariate normal distribution:

$$f_H(x) \sim N(x; \mu_H, \Sigma_H) \text{ and } f_S(x) \sim N(x; \mu_S, \Sigma_S), \quad (1)$$

where $N(x; \mu, \Sigma)$ denotes the Gaussian probability density function with location vector μ and covariance matrix Σ; and the subscript H denotes the happy-face processor and S denotes the sad-face

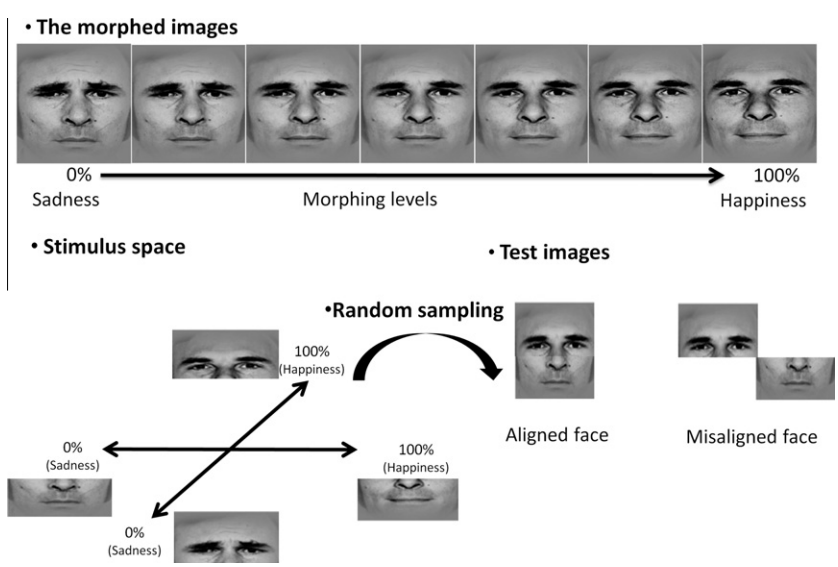

Fig. 1. *The stimuli.* The top panel presents the morphed images. The morphing levels are defined as the proportion of the happy face used in the morphed image. The images we used were morphed from 0% to 100% in 20% increments and at 50% morphing levels. We cut the image at the midpoint to separate it into upper and lower halves. These half-face images were used to create the stimulus space. The stimulus space, presented in the bottom panel, has two dimensions: one for the upper face, and another for the lower face. The coordinates of the stimulus space are the morphing levels. The randomly-sampled upper and lower faces, at different morphing levels, were joined to create the test images. The misaligned face had a lateral shift of 44′ visual angle.

processor. In the context of our experiment μ is a 2×1 vector for the preferred stimulus of the processor in the two-dimensional stimulus space and Σ is a 2×2 matrix that represents the covariance matrix for each processor's response to the upper and lower faces. The upper and lower faces were presented independently throughout the experiment. Hence, if there is no interaction between the contributions of the upper and lower faces, the off-diagonal elements of Σ should be zero and the covariance matrix should be reduced to a diagonal matrix. Otherwise, it should be possible to quantify the interaction between the upper and the lower faces in classification of expressions by the covariance.

Happy/sad classification depends on the likelihood ratio, $l(x)$, between the responses of the two processors,

$$l(x) = f_H(x)/f_S(x), \tag{2}$$

The image x is classified as happy if $l(x) \geqslant 1$, and sad if $l(x) < 1$. For convenience of computation, we took the logarithm transform of the likelihood ratio in Eq. (2),

$$h(x) = \ln(l(x)) = ln(f_H(x)) - \ln(f_S(x)), \tag{3}$$

where $h(x)$ is the likelihood of x to be a happy face. A face is considered to be happy if $h(x) \geqslant 0$, and sad if otherwise.

Plugging the probability density function in Eq. (1) into Eq. (3), the likelihood function $h(x)$ is:

$$h(x) = [-(x-\mu_H)'\Sigma_H^{-1}(x-\mu_H) + (x-\mu_S)'\Sigma_S^{-1}(x-\mu_S)], \tag{4}$$

The probability of an observer classifying a stimulus x as a happy face is a function of the likelihood ratio, $h(x)$. The cumulative distribution function of the standard normal distribution accounts for the observer's performance. That is, the probability of an image, x, being considered as happiness, $P_H(x)$, is:

$$P_H(x) = \Phi(h(x), m, s), \tag{5}$$

where $\Phi(z, m, s$ is the Gaussian cumulative distribution function of the dummy variable z with location parameter (mean), m, and scale parameter (standard deviation), s. Empirically, we found that $m = 0$ and $s = 1$ gave a good fit to the data. The observers' performance conforms to the cumulative standard normal distribution.

The decision boundary is the zero-crossing of the likelihood function. It has been shown that the general form of the decision boundary is quadratic, unless the covariance matrices of both processors, Σ_H and Σ_S, are the same (Ashby, 1992, p. 28). That is, if the noise structures of the two processors are equal, a linear decision boundary could be expected. Otherwise, we would expect a quadratic decision boundary.

A Chi-square goodness-of-fit test (Wackerly, Mendenhall, & Scheaffer, 2002) served as the model fitting indicator. The F-test was used to compare a full model with a reduced model (Kirk, 2003). That is,

$$F = [(\mathrm{SSE}_r - \mathrm{SSE}_f)/(df_f - df_r)]/[\mathrm{SSE}_f/df_f] \tag{6}$$

where SSE_r and SSE_f were the sum of squared error of the reduced and full model fits respectively and df_r and df_f were the degree of freedom of the corresponding models. The degree of freedom of the F-test was $df_f - df_r$ for the numerator and df_f for the denominator.

3. Results

Fig. 2 plots the probability of classifying a test image as happiness as a function of morphing levels in the lower face. Each row in Fig. 2 shows the data from one observer. The left column shows the data on the aligned faces presented in the fovea; the middle column, the misaligned faces presented in the fovea; the right column, the aligned faces presented in the 6° periphery. Different curves in each panel denote different morphing levels of the upper faces. The light blue curves and circles denote a sad upper face (0% morphing level); the red curves and circles denote a happy upper face (100% morphing level). A color bar on the side of the figure denotes that the greater degree of red in the curve and circles corresponds to the greater morphing level on the upper face. The smooth curves are the fits of the model discussed below.

In general, the probability of happiness classification increased monotonically with the morphing levels of the lower face in all conditions, as shown in Fig. 2. The functions of the increment of the 'happiness' response have a sigmoid shape. The probability of the expression being judged to be happy also increased monotonically with the upper face morphing level, suggesting the contribution of the upper face in happy/sad classification. However, the magnitude of change produced by the happy and sad upper faces was different. Here, for simplicity, we used the classification performance for images with a 50% lower face morphing level to illustrate this difference. As plotted as solid lines in Fig. 3, when both the upper and the lower faces were at 50% morphing level, the response of the observers to the images was at the level of chance (.51). We used this point as a reference point for the following comparisons. When the morphing level in the upper face increased toward the happy face (from 50% to 100%), the probability of a 'happiness' response reached a plateau of .60. That is, there was a maximum of .09 increment ($t(11) = -2.98$, $p = .02 > \alpha = .01$) in the probability of a 'happiness' response that can be produced by a happy upper face. That is, there is practically no effect of increasing "happiness" proportion in the image beyond 50%. On the other hand, when the morphing level moved toward the sad face (from 50% to 0%), the probability of a 'happiness' response reduced to .32. That is, the change in the probability of a 'happiness' response produced by a sad upper face (.19) was statistically significant ($t(11) = -3.18$, $p = .008 < \alpha = .01$) and was about twice that of the happy upper faces.

On the other hand, the change of lower face, while produced a more pronounced effect in expression classification, did not show such bias toward either happy or sad. Increasing the happiness proportion in the image from 50% to 100% increased the probability of a 'happiness' response to 84%. That is, the probability of a 'happiness' response increased by 33%. On the other hand, decreasing the happiness proportion in the image from 50% to 0% decreased the probability of a 'happiness' response to 16%, or a 35% decrease. That is, moving the morphing level toward either expression in the lower face had a symmetric effect.

Fig. 4 plots the averaged probability of a 'happiness' response for an image with a morphing level of 50% in both the lower and the upper face in the three experimental conditions. Classification performance for the fovea was different from the peripheral viewing condition ($t(3) = 4.37$, $p = 0.0220 < \alpha = .05$). It was more difficult to perceive a face presented in the periphery as happy. Classification performance was similar, however, for both the aligned and misaligned face conditions ($t(3) = -2.24$, $p = 0.1112 > \alpha = .05$). That is, spatial configuration played little role in classification of facial expressions.

The data were fit with the model presented in Section 2. This model provided a good fit for the data ($\chi^2 = 21–41 < \chi^2_{(40)} = 55.76$, $\alpha = .05$). Table 1 lists the best fit parameter values and the F-test results of the goodness-of-fit for each observer's data. Notice that the parameters are the same for both the aligned and misaligned conditions. Even though we allowed the parameters for these two datasets to be different, the goodness-of-fit did not improve ($F(10, 126) = 0.46–1.74$, $p = .07–.94 > \alpha = .05$).

The covariance between the upper and lower facial features was constrained to be zero for the two foveal viewing conditions. This constraint was applied to both happy- and sad-face processors.

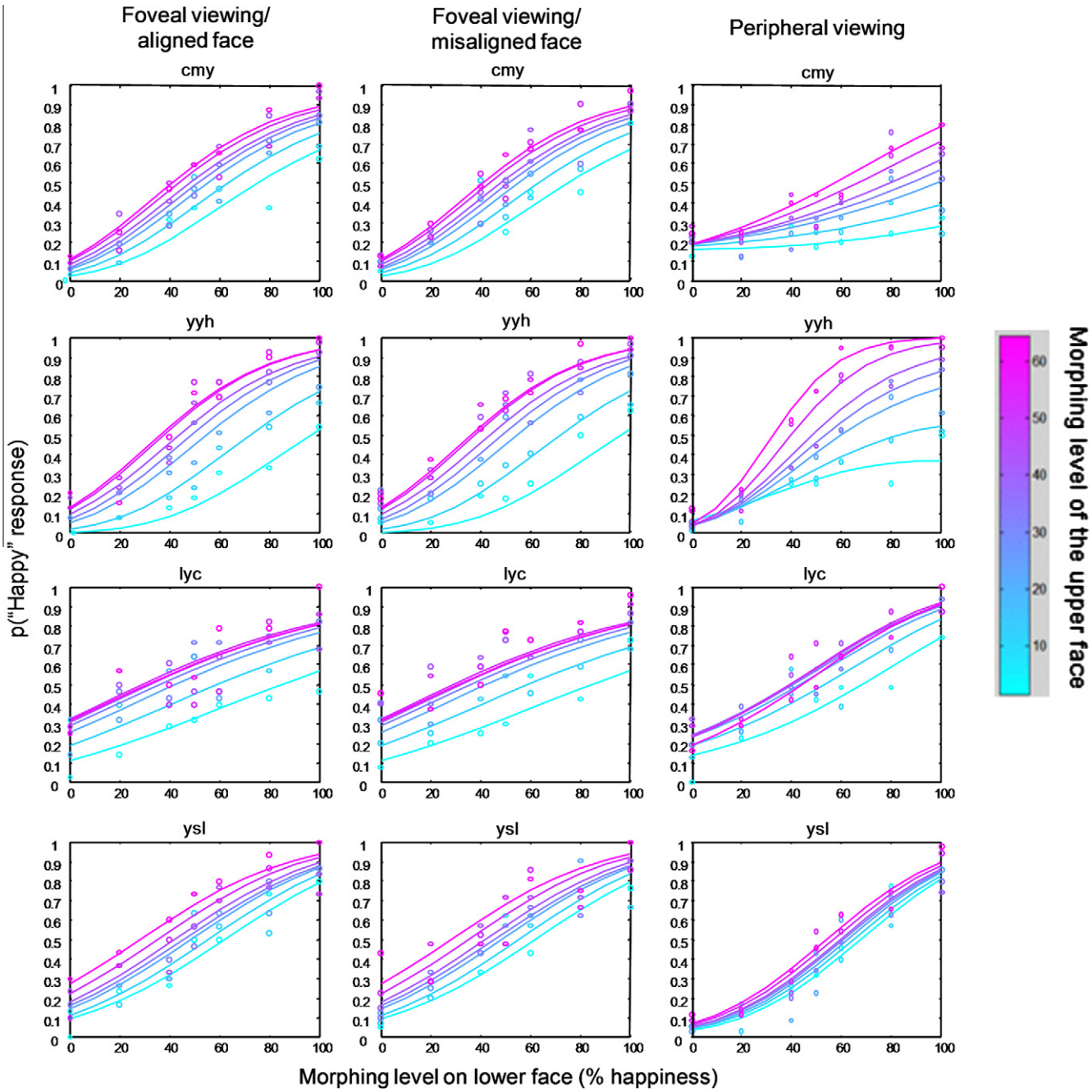

Fig. 2. The data and the model prediction of the three conditions. The columns are the experimental conditions; the rows are the observers. In each plot, the vertical axis represents the probability of a test image being classified as happiness; the horizontal axis represents the morphing levels on the lower face. The morphing level of the upper face is differentiated by colored lines. As the color of the lines changes from blue to red, the morphing levels in the upper face changes from sad to happy. The circles and lines represent the empirical data and the model prediction respectively. The probability of classification of happiness increased monotonically with morphing levels of the lower face in all conditions.

Allowing the two covariant parameters to be free provided little, if any, improvement in the model fit ($F(2, 38) = 0.59–2.67$, $p = 0.08–0.55 > \alpha = .05$). This suggests that there was little interaction between upper and lower facial features in expression classification. This result is consistent with the idea discussed above, that spatial configuration has little effect in classification of facial expressions.

The parameters in Table 1 suggest that the mechanisms of facial expression classification differ across eccentricities. We first constrained the parameters for the foveal and peripheral viewing conditions to be the same. This constraint significantly deteriorated the goodness-of-fit of the model ($F(10, 126) = 2.05–4.35$, $p = 0–0.03 < \alpha = .05$). The difference between foveal and peripheral classification results from the difference in the covariance structure of the contribution of the upper and lower faces to the happy- and sad-face processors. Contrasted with the foveal conditions, the covariance parameters for the happy-face processor in the peripheral viewing conditions could not be reduced to zero. Such constraint would significantly deteriorate the goodness-of-fit ($F(1, 185) = 39.27$, $p = 0.0000 < \alpha = .05$). The same constraint for the sad-face processor, on the other hand, had little, if any, influence on the goodness-of-fit ($F(1, 185) < 0.0001$, $p > 0.9999 > \alpha = .05$).

Fig. 5 shows the decision boundary derived from the parameters in Table 1. The decision boundary is the zero-crossing of the likelihood function. That is, it denotes the point where the stimuli

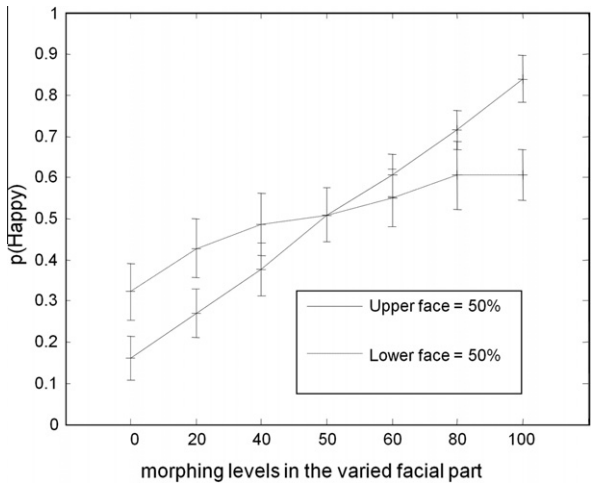

Fig. 3. The different contribution of the upper and the lower face in happy/sad classification. When the upper face was fixed at 50% (solid lines), an increment of morphing level in the lower face significantly increased the probability of happiness classification. However, when the lower face was fixed at 50% (dashed lines), an increment of morphing level in the upper face beyond 50% had little effect on classification. The horizontal axis represents the morphing levels, or the proportion of happiness, of the other half face. The vertical axis represents the probability of observers classifying an image as a happy face. The data points were averaged across observers and test conditions. Error bar indicates one standard error.

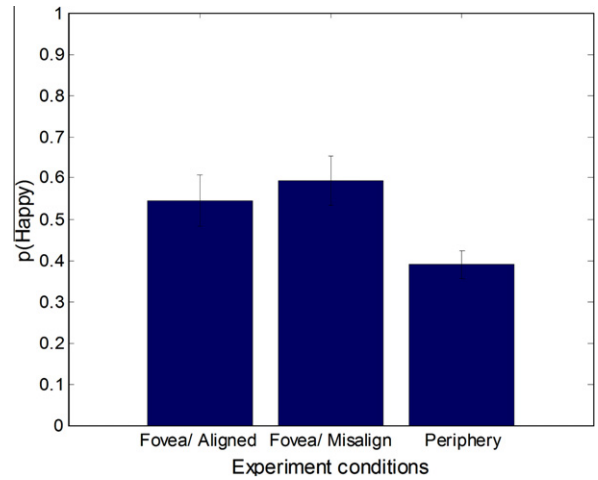

Fig. 4. A comparison of the three experimental conditions. The horizontal axis represents the three experimental conditions. The vertical axis shows the averaged probability of a 'happiness' response in the four observers. The observers' performance in the two foveal viewing conditions is the same (one-tailed paired *t*-test, $p = 0.9444 > \alpha = .05$), but is different from the peripheral viewing condition (one-tailed paired *t*-test, $p = 0.0110 < \alpha = .05$). Error bar indicates one standard error.

produced the same amount of response in both happy- and sad-face processors. The stimuli sampled from the left of the decision boundary would produce more response in the sad-face processor than in the happy-face processor. Therefore, these stimuli would be classified as sadness, and vice versa. Compared with the foveal viewing conditions, the decision boundaries for the peripheral viewing condition shifted in the stimulus space. This is consistent with the finding in Fig. 4 that the sensitivity of the happy-face processor is lower than that of the sad-face processor in the periphery.

The form of the decision boundaries is determined by the covariance matrices of the expression processors. A decision boundary is linear when the covariance matrices of the two processors, Σ_H and Σ_S, are the same (Ashby, 1992; Duda & Hart, 1973); otherwise, it is quadratic. As can be seen from Table 1 and Fig. 5, all except one observer showed quadratic decision boundaries.

4. Discussion

We can conclude the following points from the data and the model fits. First, there is no difference between aligned and misaligned faces, at least in the fovea. The same set of parameters can well fit classification performance in both conditions. Second, there is no interaction between features on the upper and lower face in determining the facial expression in the fovea. The model still fits well even though the covariance matrices of both happy- and sad-face processors are diagonal matrices. Third, the interaction between features on the upper and lower face in the happy-face processor is, however, pronounced in the periphery. Fourth, the decision boundary is always quadratic in most of the observers, regardless of viewing conditions. This implies an unequal variance in the processors, or a different bandwidth of expression tuning, in the happy- and sad-face processors. Fifth, the decision boundary shifts sideways in the peripheral viewing condition. This implies that the happy-face processor is relatively insensitive in the periphery.

Table 1
The result of model fitting and parameters in both viewing conditions.

Observation	Foveal viewing condition				Peripheral viewing condition			
	cmy	ysl	yyh	lyc	cmy	ysl	yyh	lyc
Parameters (f_H)								
μ_{upper}	72.60	73.14	71.32	75.26	97.14	94.22	166.68	70.83
μ_{lower}	48.07	39.07	43.61	31.28	60.31	56.57	20.11	40.80
σ_{upper}	20.16	16.68	19.72	29.14	21.33	28.86	18.08	29.73
σ_{lower}	20.73	20.15	22.82	23.86	18.56	22.16	19.03	17.26
Covariance	0.00	0.00	0.00	0.00	10.41	5.55	18.11	6.29
Parameters (f_S)								
μ_{upper}	71.46	72.12	69.64	74.55	95.54	91.13	163.11	70.98
μ_{lower}	42.54	36.08	36.19	27.00	57.33	49.14	6.50	37.85
σ_{upper}	20.37	16.68	20.35	30.64	21.41	28.34	17.83	31.28
σ_{lower}	21.14	20.15	23.24	24.05	18.44	22.58	19.49	17.15
Covariance	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Goodness-of-fit								
SSE	0.2277	0.3005	0.2122	0.4770	0.2620	0.2274	0.1546	0.2587
χ^2	21.9694	27.3288	35.1949	40.4143	30.7040	41.2263	29.2949	26.5237
p-value	0.9908	0.9610	0.686	0.4520	0.8260	0.3735	0.8705	0.9360

Note: The critical value of the goodness-of-fit test is ($\chi^2_{(40)} = 55.7585$, $\alpha = .05$ in foveal viewing condition and $\chi^2_{(39)} = 54.5722$, $\alpha = .05$ in peripheral viewing condition.

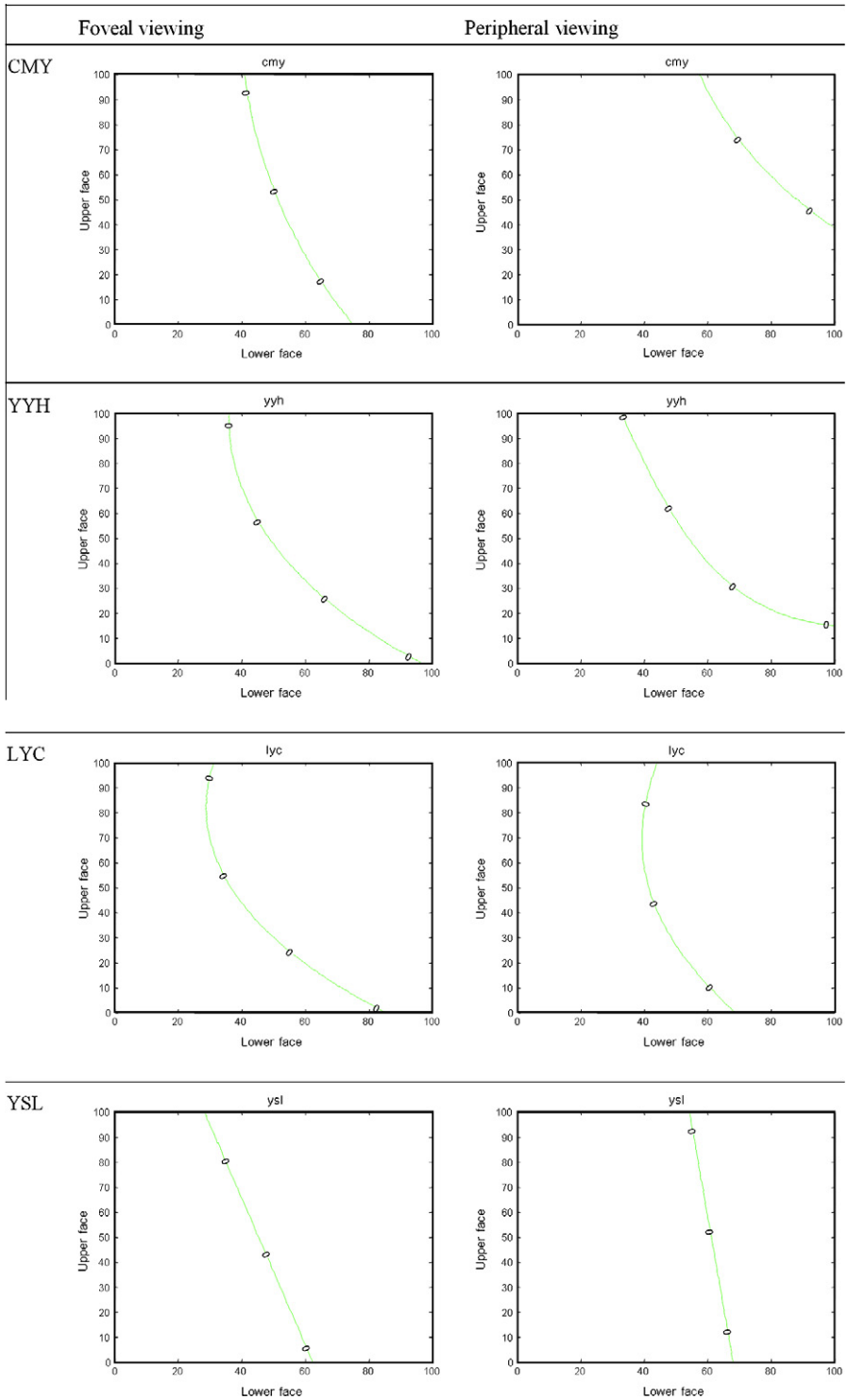

Fig. 5. The decision boundary. Each column represents the viewing conditions; each row represents one observer. Each panel represents a stimulus space. The *x*-axis shows the morphing level of the lower face; the *y*-axis the morphing level of the upper face. The dark line shows the decision boundary, the zero-crossing of the likelihood function. The decision boundaries shift sideways in the periphery. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Our results support the assumption that happy/sad classification is an analytic process, at least in the fovea. There is no significant difference in classification performance between the aligned and misaligned conditions, even though misaligned faces change the relative positions of facial features in the upper and lower face, and in turn change the spatial configuration of the face. Our result is different from results in studies of face recognition, which is considered to be a holistic process (Carey & Diamond, 1977; Chen et al., 2007; Yin, 1969). We suggest that the visual system may use different computational rules to deal with different types of information, such as identity and expression, in a face (Calder & Young, 2005).

The significant difference between the foveal and peripheral viewing conditions implies that there is more than one mechanism operating at different eccentricities. In our experiment, we controlled the impact of spatial frequency on classification of facial

expressions (Livingstone, 2000). The image size was scaled to equalize the sensitivity to a face in the foveal and peripheral viewing conditions. The shift of the decision boundaries can be explained by a change in sensitivity in the expression processors, rather than a difference in sensitivity to faces in general. Such a difference occurs in the covariance matrices of the happy- and sad-face processors. In the periphery, the happy-face processor relies on the interaction between the upper and lower parts of a face. This computational complexity may account for the difficulty of recognizing happiness in the periphery.

Previous studies have also shown a difference in brain activation caused by emotional faces presented in the fovea and in the periphery. Those results showed that fearful faces presented in the periphery produce more BOLD signals in the amygdala than those presented in the fovea (Morris et al., 1996; Silvert et al., 2007; Vuilleumier, Armony, Driver, & Dolan, 2003; Winston, Vuilleumier, & Dolan, 2003). A common explanation for this effect is that it is due to the different roles of the parvo- and magno-cellular pathways in carrying information about expressions. Our result is consistent with this notion. However, our result differs from Goren and Wilson's (2006), which shows that the geometry change threshold for negative emotion is greater in the periphery than in the fovea. In their study the judgment was based on three negative emotions and one positive emotion, while in our experiment the judgment was based on one positive and one negative emotion. It is likely that observers had difficulty in discriminating among the three negative emotions in Goren and Wilson's task. Thus, their results may reflect a difficulty in discriminating facial features rather than in classification of positive versus negative emotions.

Our results also show a difference in the contribution of the upper face to the happy- and sad-face processors which may solve the inconsistency in the literature (i.e., Kontsevich & Tyler, 2004; Smith et al., 2005). A happy upper face did not increase the judgment of happiness by much when the lower face was sad. However, a sad upper face dramatically decreased the probability of the expression being judged as happy, even when the lower face was happy. This effect was also reflected in the quadratic decision boundary. The decision boundary at high morphing levels of the upper face (i.e., happier) tended to be near the center of the stimulus space, while at lower face morphing levels (i.e., sadder) it tended to shift rightward. That is, much stronger happiness was required in the lower face for a happy-face processor to cancel the effect from a sad upper face. These results are consistent with the locations of the behavioral receptive fields revealed by the *Bubbles* method (Smith et al., 2005). That is, while the lower face is informative in recognizing a happy expression, the upper face is necessary for recognition of a sad one. These regions have also been demonstrated to be perceptually necessary and sufficient to recognize happiness and sadness in facial expressions on video (Nusseck, Cunningham, Wallraven, & Bulthoff, 2008) and in reaction time (Calder et al., 2000). The finding of Kontsevich and Tyler (2004), the upper face contributes little to classification of facial expressions, may be due to the stimuli used. The weak expression around the Mona Lisa's eyes cannot produce enough effect to manifest itself in the classification task.

While our image manipulation, in which the upper and lower faces were combined using different morphing levels, was similar to that used in a composite face paradigm (Calder et al., 2000; Young et al., 1987), our paradigm differed from the composite face paradigm in two important ways. First, the composite face paradigm measured reaction time while our paradigm measured classification behavior. It is unclear whether the same neural mechanism is tagged in the measurement of reaction time and classification performance. Hence, it may not be appropriate to make a direct comparison between our paradigm and the composite face paradigm. Second, in our experiment, the observers were instructed to make their decisions based on the information of a whole face, while in a typical composite face paradigm observers are instructed to attend to one half of a face. In other words, the observers are encouraged to use localized information, but their performance tends to be affected by the unattended half. Thus, it is suggested that the composite face effect can be considered as evidence of holistic processing (Young et al., 1987) for face recognition. In our paradigm, the observers were instructed to attend to the whole face. That is, we encouraged the observers to use global, or holistic information. In a sense, our paradigm was designed to amplify the interaction, if there is any, between facial features. Yet, the observers' performance in our paradigm showed that there is little interaction between upper and lower faces. Since our result was in the opposite direction from the bias of our instruction, it shows the robustness of our result.

In conclusion, the present study has revealed the tuning properties of the facial expression processor in happy/sad classification in the fovea and periphery. We showed that perception of facial expressions is an analytic process, at least in the fovea. There is no interaction between facial features in happy/sad classification in the fovea. In addition, our results suggest different mechanisms across eccentricities. The interaction between facial features becomes valuable in recognizing a happy face in the periphery.

Acknowledgment

Supported by National Science Council (Taiwan) NSC 97-2410-H-002-158-MY2 to CCC.

Appendix A

This study was designed to measure the cortical magnification factor (Beard, Levi, & Klein, 1997) of happy/sad facial expression classification. This information was used to determine the image size of our stimuli when presented at different eccentricities. The size of a face was measured as the distance between two anchoring points: (1) the middle point of a line that links the two pupils and (2) the center of the mouth. The relationship between eccentricity and size spatial threshold is:

$$\mathrm{Th}_i = \mathrm{Th}_0 \times (1 + E/E_2) \tag{A1}$$

where E is the eccentricity of the stimulus center, Th_i is the size threshold at eccentricity i, Th_0 is the size threshold at the fovea, and E_2 is the cortical magnification factor. In the present study, we measured the size threshold at different eccentricities. Thus, we can estimate the cortical magnification factor with the size threshold and Eq. (A1).

Three observers participated in the experiment (one female; two male). One was the author and the other two were naïve to the purpose of the experiment. The task of the observer was to determine whether the presented face was happy or sad. Observers were asked to fixate on the fixation point. Feedback was given according to the response in every trial.

The PSI (Kontsevich & Tyler, 1999) dynamic threshold seeking algorithm was used to measure size threshold at 86% correct response level at the five different eccentricities from 0° to 10° visual angle with 100-min step. These eccentricities were randomized in five threshold measurement runs. Each run contained 40 trials. In each trial, either a happy face (100% morphing level) or a sad face (0% morphing level) was presented at a pre-designated eccentricity for 200 ms. See Section 2 for detailed information on the images. The images were placed to the left of the fixation point.

Fig. A1 plots size threshold (blue open circle) as a function of eccentricity in three observers. Each panel represents data from

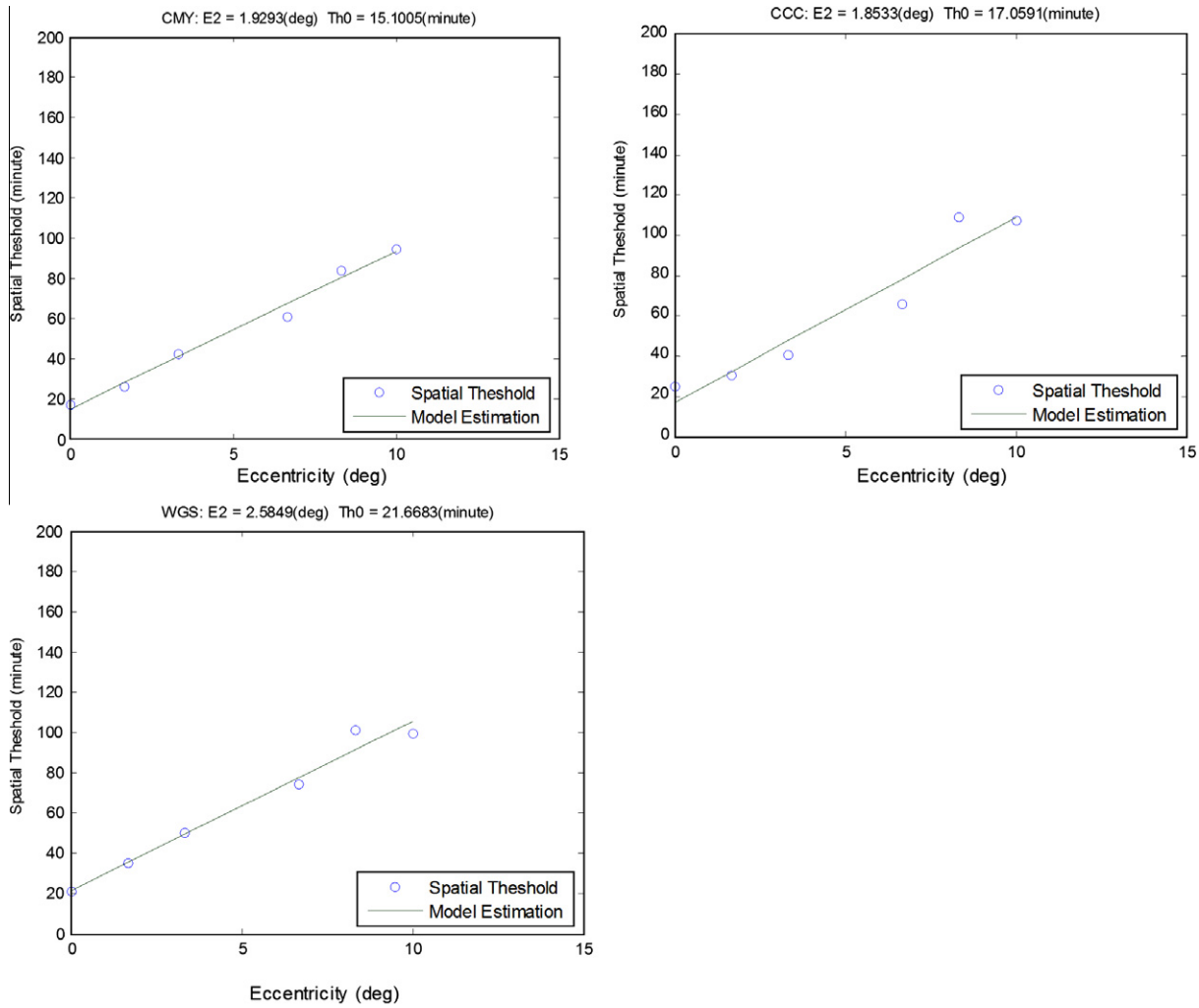

Fig. A1. The empirical spatial threshold and model prediction. The result of each observer is plotted in three separate panels. The vertical axis shows the spatial threshold; the horizontal axis shows eccentricity. The blue circle represents the empirical spatial threshold. The green line represents the model prediction. The cortical magnification factor, E_2, and the spatial threshold at the fovea, Th_0, are listed above each plot. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

one observer. The line in each panel is the fit of Eq. (A1). At the fovea, the averaged minimum image size for an observer to correctly discriminate happiness from sadness is 17.9 min visual angle. The size threshold increased with eccentricity. From the fovea to 10° eccentricity, the size threshold increased four to fivefold. This result was well fit by Eq. (A1). The mean cortical magnification factor, E_2, is 2.12° visual angle. Hence, the size threshold at 6° eccentricity is 68.6 min visual angle. With this information, in the main experiment, we used an image size that was 1.5 times the size threshold at a designated eccentricity. These image sizes yielded at least 90% correct performance.

References

Angeli, A., Davidoff, J., & Valentine, T. (2008). Face familiarity, distinctiveness, and categorical perception. *Quarterly Journal of Experimental Psychology, 61*, 690–707.

Ashby, F. G. (1992). Multidimensional models of categorization. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 449–483). Hillsdale, NJ: L. Erlbaum.

Beale, J. M., & Keil, F. C. (1995). Categorical effects in the perception of faces. *Cognition, 57*, 217–239.

Beard, B. L., Levi, D. M., & Klein, S. A. (1997). Vernier acuity with non-simultaneous targets: The cortical magnification factor estimated by psychophysics. *Vision Research, 37*, 325–346.

Bradshaw, J. L., & Wallace, G. (1971). Models for the processing and identification of faces. *Perception & Psychophysics, 9*, 443–448.

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

Calder, A. J., & Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nature Reviews Neuroscience, 6*, 641–651.

Calder, A. J., Young, A. W., Keane, J., & Dean, M. (2000). Configural information in facial expression perception. *Journal of Experimental Psychology: Human Perception & Performance, 26*, 527–551.

Carey, S., & Diamond, R. (1977). From piecemeal to configurational representation of faces. *Science, 195*, 312–314.

Chen, C. C., Kao, K. L. C., & Tyler, C. W. (2007). Face configuration processing in the human brain: The role of symmetry. *Cerebral Cortex, 17*, 1423–1432.

Chen, C. C., & Tyler, C. W. (2001). Lateral sensitivity modulation explains the flanker effect in contrast discrimination. *The Proceedings of the Royal Society of London, Series B, 268*, 509–516.

Chen, C. C., & Tyler, C. W. (2008). Excitatory and inhibitory interaction fields of flankers revealed by contrast-masking. *Journal of Vision, 8*, 1–14.

Cho, S. L. (2001). The emotion intensity rating results of images in Ekman and Friesen's Picture of Facial Affect (POFA; 1976) from Taiwanese sample. Unpublished raw data.

Darwin, C. (1872). *The expression of the emotions in man and animals*. Chicago: University of Chicago Press.

Duda, R. O., & Hart, P. E. (1973). *Pattern classification and scene analysis*. New York: Wiley. pp. 24–31.

Ekman, P., & Friesen, W. V. (1975). *Unmasking the face: A guide to recognizing emotions from facial clues*. Englewood Cliffs, NJ: Prentice-Hall.

Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press.

Etcoff, N. L., & Magee, J. J. (1992). Categorical perception of facial expressions. *Cognition, 44*, 227–240.

Fantz, R. L. (1961). The origin of form perception. *Scientific American, 204*, 66–72.

Goren, D., & Wilson, H. R. (2006). Quantifying facial expression recognition across viewing conditions. *Vision Research, 46*, 1253–1262.

Gosselin, F., & Schyns, P. G. (2001). Bubbles: A technique to reveal the use of information in recognition tasks. *Vision Research, 41*, 2261–2271.
Gosselin, F., & Schyns, P. G. (2004). No troubles with bubbles: A reply to Murray and Gold. *Vision Research, 44*, 471–477.
Kirk, R. E. (2003). *Experimental design: procedures for the behavioral sciences*. Belmont, CA: Brooks/Cole. pp. 236–239.
Kontsevich, L. L., & Tyler, C. W. (1999). Bayesian adaptive estimation of psychometric slope and threshold. *Vision Research, 39*, 2729–2737.
Kontsevich, L. L., & Tyler, C. W. (2004). What makes Mona Lisa smile? *Vision Research, 44*, 1493–1498.
Konar, Y., Bennett, P., & Sekuler, A. (2010). Holistic processing is not correlated with face-identification accuracy. *Psychological Science, 21*, 38–43.
Leder, H., & Bruce, V. (2000). When inverted faces are recognized: The role of configural information in face recognition. *The Quarterly Journal of Experimental Psychology A, 53*, 513–536.
Livingstone, M. (2000). Is it warm? Is it real? Or just low spatial frequency? *Science, 290*, 1299.
Mangini, M., & Biederman, I. (2004). Making the ineffable explicit: Estimating the information employed for face classifications. *Cognitive Science: A Multidisciplinary Journal, 28*, 209–226.
Morris, J. S., Frith, C. D., Perrett, D. I., Rowland, D., Young, A. W., Calder, A. J., et al. (1996). A differential neural response in the human amygdala to fearful and happy facial expressions. *Nature, 383*, 812–815.
Murray, R. F., & Gold, J. M. (2004). Troubles with bubbles. *Vision Research, 44*, 461–470.
Nusseck, M., Cunningham, D., Wallraven, C., & Bulthoff, H. (2008). The contribution of different facial regions to the recognition of conversational expressions. *Journal of Vision, 8*, 1–23.
O'Toole, A. R., Deffenbacher, K. A., Valentin, D., & Abdi, H. (1994). Structural aspects of face recognition and the other-race effect. *Memory and Cognition, 22*, 208.
Polat, U., & Sagi, D. (1993). Lateral interactions between spatial channels: Suppression and facilitation revealed by lateral masking experiments. *Vision Research, 33*, 993–999.
Rotshtein, P., Henson, R. N. A., Treves, A., Driver, J., & Dolan, R. J. (2004). Morphing Marilyn into Maggie dissociates physical and identity face representations in the brain. *Nature Neuroscience, 8*, 107–113.
Sekuler, A., Gaspar, C., Gold, J., & Bennett, P. (2004). Inversion leads to quantitative, not qualitative, changes in face processing. *Current Biology, 14*, 391–396.
Silvert, L., Lepsien, J., Fragopanagos, N., Goolsby, B., Kiss, M., Taylor, J. G., et al. (2007). Influence of attentional demands on the processing of emotional facial expressions in the amygdala. *Neuroimage, 38*, 357–366.
Smith, M. L., Cottrell, G. W., Gosselin, F., & Schyns, P. G. (2005). Transmitting and decoding facial expressions. *Psychological Science, 16*, 184–189.
Tyler, C. W., & McBride, B. (1997). The Morphonome image psychophysics software and a calibrator for Macintosh systems. *Spatial Vision, 10*, 479.
Valentine, T. (1988). Upside-down faces: A review of the effect of inversion upon face recognition. *British Journal of Psychology, 79*, 471–491.
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. *Neuron, 30*, 829–841.
Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2003). Distinct spatial frequency sensitivities for processing faces and emotional expressions. *Nature Neuroscience, 6*, 624–631.
Wackerly, D. D., Mendenhall, W., & Scheaffer, R. L. (2002). *Mathematical statistics with applications*. Boston, Mass: Duxbury Press. pp. 680–691.
Winston, J. S., Vuilleumier, P., & Dolan, R. J. (2003). Effects of low spatial-frequency components of fearful faces on fusiform cortex activity. *Current Biology, 13*, 1824–1829.
Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81*, 141–145.
Young, A. W., Rowland, D., Calder, A. J., Etcoff, N. L., Seth, A., & Perrett, D. I. (1997). Facial expression megamix: Tests of dimensional and category accounts of emotion recognition. *Cognition, 63*, 271–313.
Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception, 16*, 747–759.