Article

PERCEPTION

Image Valence Modulates the Processing of Low-Resolution Affective Natural Scenes

Natasha Stevenson and Kun Guo (iD)
University of Lincoln, UK

Abstract
In natural vision, noisy and distorted visual inputs often change our perceptual strategy in scene perception. However, it is unclear the extent to which the affective meaning embedded in the degraded natural scenes modulates our scene understanding and associated eye movements. In this eye-tracking experiment by presenting natural scene images with different categories and levels of emotional valence (high-positive, medium-positive, neutral/low-positive, medium-negative, and high-negative), we systematically investigated human participants' perceptual sensitivity (image valence categorization and arousal rating) and image-viewing gaze behaviour to the changes of image resolution. Our analysis revealed that reducing image resolution led to decreased valence recognition and arousal rating, decreased number of fixations in image-viewing but increased individual fixation duration, and stronger central fixation bias. Furthermore, these distortion effects were modulated by the scene valence with less deterioration impact on the valence categorization of negatively valenced scenes and on the gaze behaviour in viewing of high emotionally charged (high-positive and high-negative) scenes. It seems that our visual system shows a valence-modulated susceptibility to the image distortions in scene perception.

Our visual inputs in natural vision are often noisy and distorted (e.g., bad weather conditions, degraded images, and videos taken by closed-circuit television). Although our visual system can tolerate the degradation in visual signal quality to certain degree (Castelhano & Henderson, 2008; Röhrbein et al., 2015; Torralba, 2009; Watson & Ahumada, 2011), image

Corresponding author:
Kun Guo, School of Psychology, University of Lincoln, Lincoln LN6 7TS, UK.
Email: kguo@lincoln.ac.uk

distortion can affect our perceptual scene judgement and associated image-viewing gaze allocation that is essential to select and extract informative local visual information for scene understanding. For instance, adding white noise or blurring images could reduce the perceived image quality rating, and viewing of these noisy images was accompanied with increased fixation durations, shortened saccade distances, and concentrated central fixation bias (Loschky & McConkie, 2002; Nuthmann, 2013; Röhrbein et al., 2015; van Diepen & d'Ydewalle, 2003). Interestingly, these distortion effects on scene perception and gaze distribution were less evident for man-made scenes than for natural scenes (Röhrbein et al., 2015), indicating a scene category-dependent susceptibility to the image distortions in our scene perception.

Considering that natural visual signals are often associated with different affective meanings (e.g., spring flowers are pleasant, but snakes are frightening visual inputs for most of us), they can be grouped into different scene categories according to the embedded emotional valences and intensities. By using high-quality visual stimuli, recent psychophysical, neurophysiological, and brain imaging studies have revealed that the affective value of stimuli has significant impact on our visual processing capabilities (e.g., visual attention allocation, target detection speed and accuracy, and perceptual field size) and neuronal activation in a range of cortical regions in the visual pathway, starting as early as primary visual cortex (e.g., Guo et al., 2020; Li et al., 2019; Phelps, 2006; Vuilleumier, 2005). This is usually reflected by relatively quicker detection time, faster processing speed, and enhanced neural responses for negatively valenced relative to neutral stimuli (Markovic et al., 2014; Phelps, 2006; Pourtois et al., 2013). The extensive interactions between visual and emotional information along the visual pathway suggest that vision and emotion are less decomposable, and hence, it could be hypothesized that our susceptibility to the image distortions in scene perception could be further modulated by the scene emotional value.

Indeed, recent studies have observed that local semantic regions in affective scenes tended to attract more early attention or fixations than in emotionally neutral scenes, even when these scene images were masked by pink noise to only allow a gist understanding of the scene (Pilarczyk & Kuniecki, 2014). These findings are broadly consistent with the observations from event-related potential studies. When overlaid with colour noise pixels or filtered through low- or high-pass spatial filters, those degraded but still identifiable high-arousing emotional pictures elicited larger components of early posterior negativity (Schupp et al., 2008) and late positive potential (De Cesarei & Codispoti, 2011) than low-arousing or neutral pictures, indicating less vulnerable attentional capture and motivational significance for high-arousing emotional cues in noisy visual inputs. Emotional scenes, especially unpleasant ones, also needed a wider range of spatial frequencies and longer response time to be identified than neutral scenes (De Cesarei & Codispoti, 2011). Furthermore, when degrading fine-grained image details by Gaussian noise masking, image size reduction or low-pass spatial filtering, emotionally arousing pictures with either positive or negative valence attracted more fixations (Todd et al., 2012) or attentional resources to interfere with unrelated cognitive tasks (De Cesarei & Codispoti, 2008) and were judged as less noisy and more perceptually vivid than emotionally neutral pictures (De Cesarei & Codispoti, 2008; Todd et al., 2012). Taken together, it seems that emotional cues can enhance or facilitate our perception of degraded scenes. However, it is unclear whether different categories and/or intensities of scene valence (e.g., high-positive vs. high-negative, or high-positive vs. medium-positive) will show different degree of modulation on our processing of degraded natural scenes.

A few studies on recognizing human facial expressions of varying image qualities have shed light on addressing this question. Increasing face image blur (e.g., >15 cycles/image) or

decreasing image resolution (e.g., $<48 \times 64$ pixels) would typically reduce expression categorization accuracy and expression intensity rating, increase fixation duration, and strengthen central fixation bias (Guo et al., 2019). However, these distortion effects were expression-dependent with less deterioration impact on relatively positive happy and surprise expressions in comparison with negative expressions such as angry, sad, fear, and disgust (Du & Martinez, 2011; Guo et al., 2019), suggesting that the emotional valence could influence our perceptual judgement and gaze behaviour in viewing of (at least) degraded facial expression images. Furthermore, recent computational studies have reported that images with negative and neutral valence could be classified with $\sim$70% accuracy by using Fourier amplitude spectrum (Rhodes et al., 2019), and different combinations of global image properties (e.g., colour values, symmetry, edge density, self-similarity, Fourier slope and sigma, and first-order and second-order edge-orientation entropies) could even predict valence and arousal ratings of affective scenes at above chance level (Redies et al., 2020), suggesting that different categories of affective scenes may have systematic differences in image statistics. Given that image distortion would disturb image statistics (Sheikh et al., 2005) and natural images of different local and global properties subsequently show different susceptibilities to the same distortion process (Röhrbein et al., 2015), it is plausible that our perception of degraded affective scenes would vary according to their emotional valence and arousal.

In this eye-tracking experiment by presenting natural animal scenes with different categories and levels of emotional valence (high-positive, medium-positive, neutral/low-positive, medium-negative, and high-negative), we systematically investigated our perceptual sensitivity (including image valence categorization and arousal rating, which reflect both qualitative and quantitative analysis of emotional cues) and image-viewing gaze behaviour to the changes of image resolution. Based on the aforementioned research on facial expression perception and affective scene statistics, we hypothesized that those resolution-induced image distortion effects could be further modulated by the scene emotional valence.

Methods

Thirty-five undergraduate students (30 women), age between 18 and 38 years old with the mean of 21.43 ± 4.21 (Mean $\pm$ SEM), took part in this study in exchange of course credit. This sample size is compatible with those published studies in this field (e.g., Guo et al., 2019; Röhrbein et al., 2015). All participants had normal or corrected-to-normal visual acuity. The Ethical Committee in School of Psychology, University of Lincoln approved this study (PSY181995). Written informed consent was obtained from each participant, and all procedures complied with the British Psychological Society Code of Ethics and Conduct and were in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki).

Twenty affective natural scene images with comparable semantic meaning (i.e., animals in their natural environment) were selected from the animal category in the international affective picture system (IAPS; Lang et al., 2008) and divided into five categories based on their IAPS valence ratings (four images per category): high-positive (8.16 ± 0.06; including image ID [valence rating] such as 1440 [8.19], 1441 [7.97], 1460 [8.21], 1750 [8.28]), medium-positive (6.91 ± 0.03; including 1410 [7], 1603 [6.9], 1740 [6.91], 1812 [6.83]), neutral or low-positive (5.74 ± 0.12; including 1333 [6.11], 1670 [5.82], 1908 [5.28], 1947 [5.85]), medium-negative (4.53 ± 0.1; including 1080 [4.24], 1390 [4.5], 1726 [4.79], 1945 [4.59]), and high-negative (3.35 ± 0.07; including 1052 [3.5], 1111 [3.25], 1220 [3.47], 1271 [3.19]). The differences in image valence ratings between adjacent categories were $\sim$1.2. Each image

contained more than one informative cue or animal to promote more distributed image-viewing gaze allocation. The selected images were further controlled to be comparable on both low level and global image properties (Lakens et al., 2013; Redies et al., 2020) across the valence categories. Overall, images of different valence ratings had no significant difference in image brightness, $F(4, 19) = 0.68$, $p = .62$; root mean square contrast, $F(4, 19) = 3.13$, $p = .05$; hue-saturation-value colour space, $F(4, 19) < 0.75$, $p > .57$; first-order edge-orientation entropy, $F(4, 19) = 1.61$, $p = .22$; and second-order edge-orientation entropy, $F(4, 19) = 1.35$, $p = .30$; symmetry, $F(4, 19) = 1.71$, $p = .20$; and self-similarity, $F(4, 19) = 1.79$, $p = .18$.

The procedure for image manipulation is the same as in Guo et al. (2019), from which the following details are reproduced. The original high-quality images (1,024 × 768 pixels) were processed in Adobe Photoshop to downsize to 512 × 384 pixels (referred to as resolution 1). For each of these "resolution 1" images, four subsequent images were constructed by further downsizing to 64 × 48 pixels (resolution 1/8), 32 × 24 pixels (resolution 1/16), 16 × 12 pixels (resolution 1/32), and 8 × 6 pixels (resolution 1/64; see Figure 1 for examples). To provide a constant presentation size for all images, the four downsized images were scaled back to 512 × 384 pixels (19 × 14°) using bilinear interpolation, which preserves most of the spatial frequency components. As a result, 100 images were generated for the testing session (20 natural scene images × 5 image resolutions). These images were displayed once in a random order during the testing, and those images with the same IAPS ID and only differing in resolution were not displayed in consecutive trials to minimize the potential practice or carryover effects.

The general experimental setup and procedure for data collection and analysis here are similar to Röhrbein et al. (2015) and Green and Guo (2018), from which the following details are reproduced. All digitized colour images were presented through a ViSaGe graphics system (Cambridge Research Systems, UK) and displayed on a noninterlaced gamma-corrected colour monitor (30 cd/m^2 background luminance, 100 Hz frame rate, Mitsubishi Diamond Pro 2070SB) with the resolution of 1,024 × 768 pixels. At a viewing distance of 57 cm, the monitor subtended a visual angle of 40 × 30°.

During the free-viewing experiments, the participants sat in a chair with their head restrained by a chin rest and viewed the display binocularly. To calibrate eye movement signals, a small red fixation point (FP, 0.3° diameter, 15 cd/m^2 luminance) was displayed randomly at one of nine positions (3 × 3 matrix) across the monitor. The distance between adjacent FP positions was 10°. The participant was instructed to follow the FP and maintain fixation for 1 s. After the calibration procedure, the trial was started with an FP displayed 10° left or right to the screen centre to minimize central fixation bias. If the participant maintained fixation for 1 s, the FP disappeared, and a testing image was presented at the centre of the screen for 3 s. The participant was instructed to "view the image as you normally do" and after image presentation verbally report the perceived image valence (positive, neutral, or negative) and arousal rating on a 9-point scale (1 represents *not aroused at all* and 9 represents *extremely aroused*). No reinforcement was given during this procedure.

Horizontal and vertical eye positions from the self-reported dominant eye (determined through the Hole-in-Card test or the Dolman method if necessary) were measured using a Video Eyetracker Toolbox with 250 Hz sampling frequency and up to 0.25° accuracy (Cambridge Research Systems). The software developed in MATLAB computed horizontal and vertical eye displacement signals as a function of time to determine eye velocity and position. Fixation locations were then extracted from the raw eye-tracking data using velocity (less than 0.2° eye displacement at a velocity of less than 20°/s) and duration (greater than 50 ms) criteria (Guo, 2007). All the statistical analyses on perceptual sensitivity and image-viewing gaze distribution data were conducted in SPSS v26 (IBM SPSS Statistics).
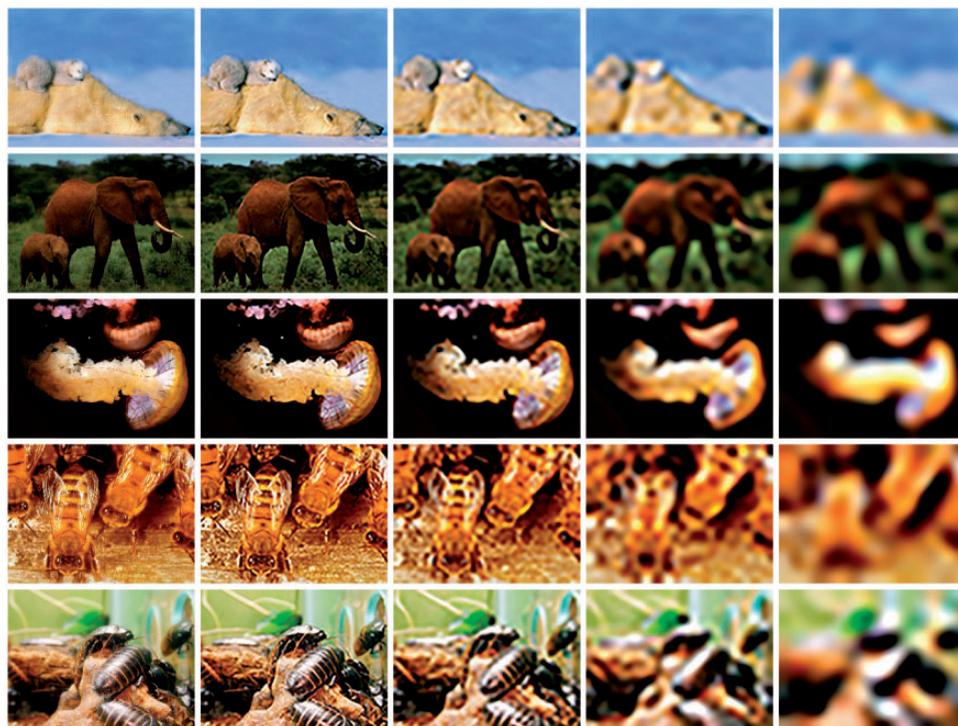
Figure 1. Examples of IAPS Natural Animal Scene Images With Different Valence Ratings (From Top to Bottom: High-Positive, Medium-Positive, Neutral/Low-Positive, Medium-Negative, and High-Negative) at Varying Image Resolutions (From Left to Right: Resolution 1, Resolution 1/8, Resolution 1/16, Resolution 1/32, and Resolution 1/64).

Note. Please refer to the online version of the article to view the figure in colour.

Results

Analysis of Behavioural Responses in Perceiving Affective Natural Scenes

To examine the extent to which image resolution affected participants' ability of judging affective natural scenes, we conducted 5 (Image Resolution [1, 1/8, 1/16, 1/32, and 1/64]) × 5 (Scene Valence Category [high-positive, medium-positive, neutral, medium-negative, high-negative]) repeated-measures analyses of variance (ANOVAs) with the image valence categorization accuracy (positive, neutral, or negative) and arousal rating as the dependent variables. For each ANOVA, Greenhouse–Geisser corrections were applied where sphericity was violated.

For valence categorization accuracy (Figure 2A), the analysis revealed significant main effect of image resolution, $F(2.62, 89.16) = 136.12$, $p < .001$, $\eta^2_p = 0.8$, with similar categorization accuracy across all scene categories at resolution 1, 1/8, and 1/16 (Bonferroni correction for multiple comparisons, resolution 1 vs. 1/8 vs. 1/16, all $ps > .05$) but a monotonically reduced accuracy at resolution 1/32 and 1/64 (resolution 1/32 or 1/64 vs. others, all $ps < .001$; linear trend $F(1, 34) = 276.51$, $p < .001$, $\eta^2_p = 0.89$), significant main effect of scene valence category, $F(2.28, 77.58) = 5.29$, $p = .005$, $\eta^2_p = 0.14$, with higher accuracy for high-positive and high-negative scenes than for medium-negative scenes (all $ps < .01$), and significant Image Resolution × Scene Valence Category, $F(3.94, 133.99) =$
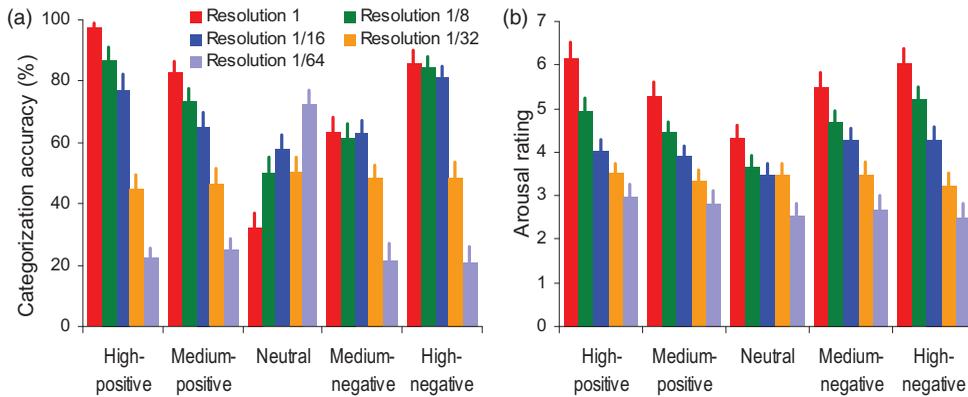
Figure 2. Mean Valence Categorization Accuracy (a) and Arousal Rating (b) for Judging IAPS Images of Varying Valence Categories (High-Positive, Medium-Positive, Neutral, Medium-Negative, and High-Negative) as a Function of Image Resolution. Error bars represent SEM.
Note. Please refer to the online version of the article to view the figure in colour.

21.4, $p < .001$, $\eta^2_p = 0.39$. Specifically, reducing image resolution from 1 to 1/64 increased the categorization accuracy for images of neutral valence monotonically except at resolution 1/32 (all other $ps < .01$; linear trend $F(1, 34) = 28.84$, $p < .001$, $\eta^2_p = 0.46$) but tended to decrease the categorization accuracy for images of positive and negative valence, especially at resolution 1/32 and 1/64 (all $ps < .01$). Interestingly, scene valence further affected our categorization of the degraded scenes. Reducing image resolution from 1 to 1/16 had detrimental effect on recognizing positively valenced (high- and medium-positive) scenes (all $ps < .05$) but had no impact on negatively valenced (high- or medium-negative) scenes (all $ps > .05$).

For image arousal rating (Figure 2B), the analysis revealed significant main effect of image resolution, $F(1.41, 48.04) = 40.92$, $p < .001$, $\eta^2_p = 0.55$, with image degradation leading to monotonically decreased arousal rating (all $ps < .001$; linear trend $F(1, 34) = 50.37$, $p < .001$, $\eta^2_p = 0.6$), significant main effect of scene valence category, $F(2.53, 85.91) = 7.68$, $p < .001$, $\eta^2_p = 0.19$, with lower arousal rating for neutral scenes than for all other types of scenes (all $ps < .002$), and significant Image Resolution × Scene Valence Category, $F(8.28, 281.62) = 8.67$, $p < .001$, $\eta^2_p = 0.2$. Specifically, at resolution 1, scenes with valence of high-positive and high-negative attracted the highest arousal rating, followed by medium-positive and medium-negative, and then by neutral (all $ps < .001$); at resolution 1/8 or 1/16, high-positive, high-negative, medium-positive, and medium-negative scenes attracted indistinguishable arousal rating (all $ps > .05$) that was higher than neutral scenes (all $ps < .05$); at resolution 1/32 or 1/64, however, all images attracted the same arousal rating regardless their original IAPS valence ratings (all $ps > .05$). Clearly, although reducing image resolution has decreased the averaged arousal rating across all the tested images monotonically, image degradation had different degrees of impact on different scene categories at each image resolution.

Analysis of Gaze Behavioural in Perceiving Affective Natural Scenes

To examine the extent to which image resolution affected participants' gaze behaviour in viewing of affective natural scenes, we conducted 5 (Image Resolution [1, 1/8, 1/16, 1/32,

and 1/64]) $\times$ 5 (Scene Valence Category [high-positive, medium-positive, neutral, medium-negative, high-negative]) ANOVAs with averaged number of fixations directed at each image, averaged fixation duration across all fixations directed at each image (i.e., quantifying the differences in the amount of information processed from the fixated region across different image resolutions and scene categories), and averaged fixation distance from image centre across all fixations directed at each image (i.e., quantifying the differences in the spread of the fixations across different image resolutions and scene categories) as the dependent variables.

For the number of fixations (Figure 3A), the analysis revealed significant main effect of image resolution, $F(4, 136) = 5.31$, $p = .001$, $\eta^2_p = 0.14$, with images at resolution 1 attracting slightly more fixations (9.34 ± 0.37) than images at other resolutions (8.78 ± 0.38, 8.52 ± 0.4, 8.64 ± 0.38, and 8.53 ± 0.4 at resolution 1/8, 1/16, 1/32, and 1/64, respectively; all $ps < .05$), and significant main effect of scene valence category, $F(4, 136) = 14.39$, $p < .001$, $\eta^2_p = 0.3$, with more fixations directing at neutral (8.93 ± 0.39), medium- (9.4 ± 0.36), and high-negative (8.89 ± 0.36) scenes than at medium- (8.3 ± 0.39) and high-positive (8.3 ± 0.4) scenes (all $ps < .05$). The Image Resolution $\times$ Scene Valence Category was nonsignificant, $F(8.88, 301.85) = 1.31$, $p = .23$, $\eta^2_p = 0.04$.

For fixation duration (Figure 3B), the analysis revealed significant main effect of image resolution, $F(2.37, 80.55) = 3.6$, $p = .03$, $\eta^2_p = 0.1$, with a tendency of longer fixation duration for images at resolution 1/32 (292 ± 26) and 1/64 (312 ± 31), significant main effect of scene valence category, $F(2.73, 92.68) = 10.56$, $p < .001$, $\eta^2_p = 0.24$, with longer fixation at medium- (312 ± 30) and high-positive (310 ± 29) scenes than at neutral (281 ± 24), medium- (264 ± 20), and high-negative (286 ± 26) scenes (all $ps < .05$), and significant Image Resolution $\times$ Scene Valence Category, $F(6.15, 209.14) = 2.23$, $p = .04$, $\eta^2_p = 0.06$. Specifically, high-positive or high-negative scenes at different resolutions attracted the same length of fixation duration (all $ps > .05$), but medium-positive, medium-negative, or neutral scenes at low resolution (1/32 and/or 1/64) attracted longer fixation duration than the same scene at high resolution 1 (all $ps < .05$).

For fixation distance from image centre (Figure 3C), the analysis revealed significant main effect of image resolution, $F(2.92, 99.13) = 31.32$, $p < .001$, $\eta^2_p = 0.48$, with shorter distance from centre in lower resolution images than in higher resolution images, significant main effect of scene valence category, $F(2.95, 100.29) = 17.75$, $p < .001$, $\eta^2_p = 0.34$, with longer distance from centre in neutral and medium-negative scenes than in other scene categories (all $ps < .01$), and significant Image Resolution $\times$ Scene Valence Category, $F(6.86, 233.06) = 3.81$, $p = .001$, $\eta^2_p = 0.1$. Specifically, changes of image resolution in high-positive and high-negative scenes had no impact on fixation distance from the image centre (all $ps > .05$), but reducing image resolution in medium-positive, neutral, and medium-negative scenes gradually decreased fixation distance from the image centre, especially at resolution 1/32 and 1/64 (all $ps < .05$).

Discussion

In this study, we investigated the extent to which the emotional information embedded in low-resolution natural scenes would affect our scene perception and associated gaze behaviour. In general, reducing image resolution led to decreased image valence categorization accuracy and arousal rating, decreased number of fixations in image-viewing but increased fixation duration, and stronger central fixation bias. Interestingly, these distortion effects were valence-modulated with less deterioration impact on the perceptual valence categorization of negatively valenced (high- and medium-negative) scenes and on the gaze behaviour
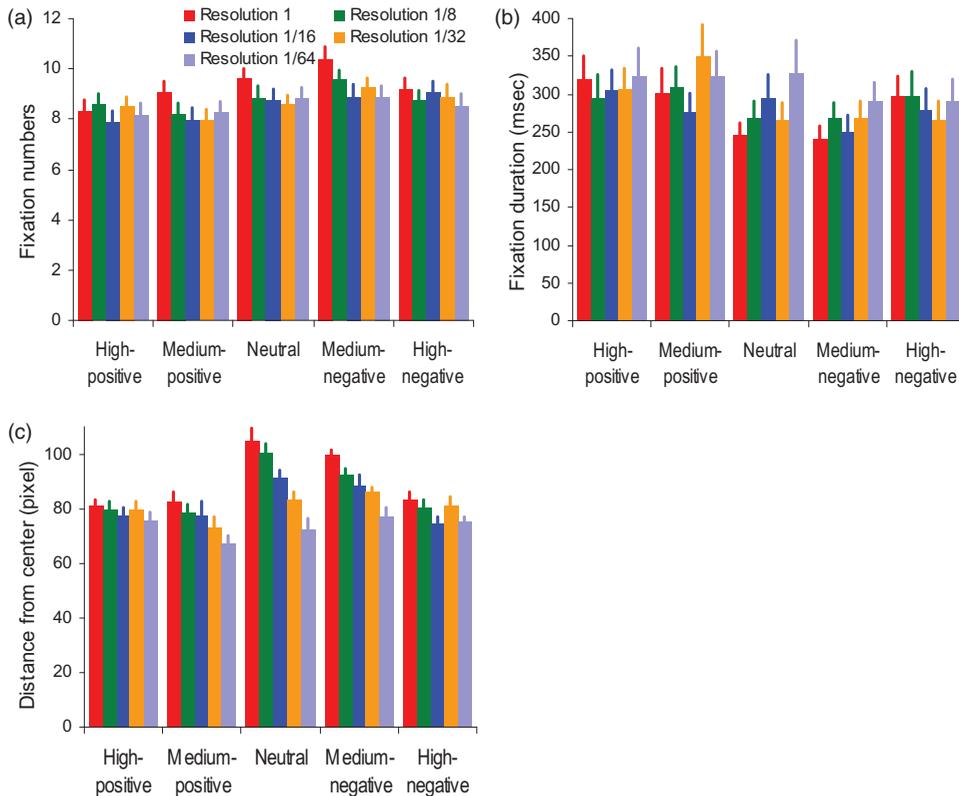
Figure 3. Average Number of Fixations (a), Average Fixation Duration Across All Fixations (b), and Average Fixation Distance From Image Centre (c) for Viewing IAPS Images of Varying Valence Categories (High-Positive, Medium-Positive, Neutral, Medium-Negative, and High-Negative) as a Function of Image Resolution. Error bars represent SEM.
Note. Please refer to the online version of the article to view the figure in colour.

in viewing of highly charged (high-positive and high-negative) emotional scenes. It seems that our visual system shows a valence-dependent susceptibility to image distortion in scene perception.

In comparison with positively valenced scenes, our capability of categorizing negatively valenced scenes is more tolerant to image distortion. Reducing image resolution up to 32×24 pixels (resolution 1/16) induced a steady decline in image valence categorization accuracy for positive scenes but had little impact on categorizing negative scenes (Figure 2A). This valence-sensitive susceptibility to image degradation could be due to negativity bias in human cognition (Baumeister et al., 2001), in which negative emotional stimuli are more salient and have higher impact on a person's cognition than neutral or positive stimuli. For instance, in comparison with neutral or happy facial expressions, angry and fearful expressions tend to pop out more easily, capture and hold attention automatically, and amplify perceptual process (Anderson, 2005; Öhman et al., 2001; Yang et al., 2007). This negativity bias may enable our participants to still recognize degraded negative scenes and/or to be biased to interpret those scenes of ambiguous valence as negative which leads to fewer "missed" trials in the presentation of degraded negative

scenes. Consequently, the valence categorization for negative scenes is less susceptible to image degradation.

Further image degradation to 16×12 pixels (resolution 1/32) or 8×6 pixels (resolution 1/64) led to decreased but comparable categorization accuracy across scenes of varying valence categories and levels, indicating a similar sensitivity in recognizing positive and negative affective cues in these low-resolution images. Interestingly, the categorization accuracy for neutral scenes was increased with decreasing image resolution, indicating a tendency to judge the valence of highly distorted images as neutral rather than positive or negative.

In this experiment, the data for image valence categorization accuracy were recorded along with the perceived arousal intensity. It appears that low image resolution-induced reduction in arousal rating showed similar changes in direction and speed as reduction in valence categorization accuracy for positive scenes, but not for negative scenes (Figure 2). Specifically, reducing image resolution from 1 to 1/16 monotonically decreased both valence categorization accuracy and arousal rating for positive scenes but only decreased arousal rating for negative scenes with little impact on negative valence recognition. Considering that the assessment of valence category and arousal intensity are related to qualitative and quantitative analysis of affective information embedded in the natural scenes, respectively, image resolution may influence differently upon qualitative and quantitative evaluation of negative affective cues. Similar viewing condition-modulated changes in qualitative and quantitative assessment have also been reported in other tasks. For instance, changing viewing perspective of expressive faces (e.g., from frontal to profile view) had negligible influence on expression recognition accuracy but considerably decreased expression intensity rating (Guo & Shaw, 2015).

We have further compared image-viewing gaze behaviour across different image resolution levels. Overall, the viewing of low-resolution images is accompanied by prolonged fixation duration and more concentrated fixation bias towards the image centre (Figure 3). These observed changes in image-viewing gaze behaviour were broadly consistent with previous research on face and scene perception. Compared with face, man-made or landscape scene images in high-quality, low-quality version of the same images (distorted with low image resolution, image blur, additive Gaussian white noise, etc.) induced longer individual fixation durations, shorter saccades, and stronger central fixation bias (Guo et al., 2019; Judd et al., 2011; Röhrbein et al., 2015). As the changes in gaze behaviour were largely driven by the degree/level of distortion (e.g., image noise intensity) rather than distortion types or image categories, this may reflect an adapted change in oculomotor strategy in processing noisy visual inputs (Guo et al., 2019). The ambiguity in the degraded scenes can increase visual task difficulty, reduce visual saliency of those objects/features in peripheral visual field, and increase the difficulty to select next fixation target and initiate saccadic eye movement (Nuthmann et al., 2010; Reingold & Loschky, 2002). This subsequently leads to prolonged fixation duration, shortened saccadic distance, and concentrated central fixation bias (Röhrbein et al., 2015).

Interestingly, the emotional strength of these affective natural scenes could modify the impact of image resolution on gaze behaviour. Reducing image resolution from 1 to 1/64 induced evident changes in gaze distribution when viewing scenes of neutral/low-positive, medium-positive, and medium-negative emotional valence (determined via IAPS valence ratings). On the other hand, the gaze distribution in viewing high-positive and high-negative affective scenes was more tolerant to the image resolution reduction. It seems that informative local visual cues in highly charged emotional scenes could attract similar gaze distribution in both high- and low-resolution images. Judd et al. (2011) have also reported when viewing images that are easy to understand or have relatively simple image complexity (e.g., only

containing one clear salient object), the location of fixations on high-resolution images tended to be similar to and predictive of fixations on low-resolution images. Our study has extended this previous observation from image complexity to image affective strength.

Although in this study the presented images of different emotional categories (high-positive, medium-positive, neutral, medium-negative, and high-negative) had comparable low level and global image properties (please see the Methods section for details), it could be argued that the comparable gaze behaviour in viewing those highly charged (high-positive and high-negative) emotional scenes of varying resolutions might be due to the relatively higher low-level local image saliency (e.g., brighter colour contrast, higher luminance contrast) in the fixated region. However, it has been reported that in comparison with visual scenes of neutral valence, our overt visual attention (or gaze allocation at the early stage of image-viewing) is more likely to be captured by unpleasant or pleasant emotional content (or semantic emotional features) rather than low-level image features in affective scenes (Humphrey et al., 2012; Markovic et al., 2014; Nummenmaa et al., 2006; Pourtois et al., 2013; Vuilleumier, 2005). In addition, the tendency to allocate early fixations at emotionally charged rather than visually salient local regions persists even when these affective images are masked by pink noise (Pilarczyk & Kuniecki, 2014) and is positively correlated with image arousal ratings (Niu et al., 2012). Taken together, it seems that local affective cues in highly charged emotional scenes are more effective in attracting visual attention and are less susceptible to image distortions than those affective cues in low-arousal scenes.

Furthermore, these low image resolution-induced changes in gaze behaviour did not always match with the changes in image valence categorization performance. Specifically, reducing image resolution from 1 to 1/16 in positive scenes showed no impact on fixation duration and central fixation bias but monotonically decreased valence categorization accuracy. Similar dissociation between image-viewing gaze behaviour and the related perceptual performance has been observed in facial expression recognition, in which low-quality happy face images led to stronger central fixation bias in gaze allocation but hardly reduced categorization accuracy for happiness (Guo et al., 2019). Future research could further examine how commonly this dissociation exists in scene perception.

It is worth mentioning that in this study we used a within-subject design and a relatively small set of well-controlled natural animal scenes with comparable low level and global image properties across image categories. This approach would generate more comparable data between testing conditions but has inevitably restricted variability in scene structure within and across scene categories. As natural scenes typically show a large between-image variation in complexity and low-level physical image salience (e.g., local image structure, shape, luminance intensity, and contrast), it remains to be seen to what extent the current findings can be generalized to other types of scene categories. Furthermore, given that affective images, especially pleasant ones, commonly contain more reddish-yellow hues, colour could provide diagnostic information about image emotional content and subsequently modify attention allocation (McMenamin et al., 2013). Although colour is unlikely to be a confounding factor in this study as we have controlled image brightness, root mean square contrast, and hue–saturation–value colour space in images across different valence categories (please see the Methods section for details), it would be interesting to examine whether the current findings can be extended to black and white affective images. Nevertheless, this study represents a step forward in our understanding of how affective meaning in natural scenes influences our scene understanding and associated gaze behaviour. It seems that our visual system shows a valence-modulated susceptibility to image distortion in scene perception.

Declaration of Conflicting Interests

Funding

ORCID iD

Kun Guo ![ORCID] https://orcid.org/0000-0001-6765-1957

References

Anderson, A. K. (2005). Affective influences on the attentional dynamics supporting awareness. *Journal of Experimental Psychology*, *134*, 258–281.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, *5*, 323–370.

Castelhano, M. S., & Henderson, J. M. (2008). The influence of color on perception of scene gist. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 660–675.

De Cesarei, A., & Codispoti, M. (2008). Fuzzy picture processing: Effects of size reduction and blurring on emotional processing. *Emotion*, *8*, 352–363.

De Cesarei, A., & Codispoti, M. (2011). Scene identification and emotional response: Which spatial frequencies are critical? *Journal of Neuroscience*, *31*, 17052–17057.

Du, S., & Martinez, A. M. (2011). The resolution of facial expressions of emotion. *Journal of Vision*, *11*, 24. https://doi.org/10.1167/11.13.24

Green, C., & Guo, K. (2018). Factors contributing to individual differences in facial expression categorization. *Cognition and Emotion*, *32*, 37–48.

Guo, K. (2007). Initial fixation placement in face images is driven by top-down guidance. *Experimental Brain Research*, *181*, 673–677.

Guo, K., Calver, L., Soornack, Y., & Bourke, P. (2020). Valence-dependent disruption in processing of facial expressions of emotion in early visual cortex – A transcranial magnetic stimulation study. *Journal of Cognitive Neuroscience*, *32*, 906–916.

Guo, K., & Shaw, H. (2015). Face in profile view reduces perceived facial expression intensity: An eye-tracking study. *Acta Psychologica*, *155*, 19–28.

Guo, K., Soornack, Y., & Settle, R. (2019). Expression-dependent susceptibility to face distortions in processing of facial expressions of emotion. *Vision Research*, *157*, 112–122.

Humphrey, K., Underwood, G., & Lambert, T. (2012). Salience of the lambs: A test of the saliency map hypothesis with pictures of emotive objects. *Journal of Vision*, *12*, 22. https://doi.org/10.1167/12.1.22

Judd, T., Durand, F., & Torralba, A. (2011). Fixations on low-resolution images. *Journal of Vision*, *11*, 14. https://doi.org/10.1167/11.4.14

Lakens, D., Fockenberg, D. A., Lemmens, K. P. H., Ham, J., & Midden, C. J. H. (2013). Brightness differences influence the evaluation of affective pictures. *Cognition and Emotion*, *27*, 1225–1246.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual* (Technical Report A-8). University of Florida.

Li, Z., Yan, A., Guo, K., & Li, W. (2019). Fear-related signals in the primary visual cortex. *Current Biology*, *29*, 4078–4083.

Loschky, L. C., & McConkie, G. W. (2002). Investigating spatial vision and dynamic attentional selection using gaze-contingent multi-resolutional display. *Journal of Experimental Psychology: Applied*, *8*, 99–117.

Markovic, J., Anderson, A. K., & Todd, R. M. (2014). Tuning to the significant: Neural and genetic process underlying affective enhancement of visual perception and memory. *Behavioural Brain Research*, *259*, 229–241.

McMenamin, B. W., Radue, J., Trask, J., Huskamp, K., Kersten, D., & Marsolek, C. J. (2013). The diagnosticity of color for emotional objects. *Motivation and Emotion*, *37*, 609–622.

Niu, Y., Todd, R. M., & Anderson, A. K. (2012). Affective salience can reverse the effects of stimulus-driven salience on eye movements in complex scenes. *Frontiers in Psychology*, *3*, 336.

Nummenmaa, L., Hyönä, J., & Calvo, M. G. (2006). Eye movement assessment of selective attentional capture by emotional pictures. *Emotion*, *6*, 257–268.

Nuthmann, A. (2013). On the visual span during object search in real-world scenes. *Visual Cognition*, *21*, 803–837.

Nuthmann, A., Smith, T. J., Engbert, R., & Henderson, J. M. (2010). Crisp: A computational model of fixation durations in scene viewing. *Psychological Review*, *117*, 382–405.

Öhman, A., Lundqvist, D., & Esteves, F. (2001). The face in the crowd revisited: A threat advantage with schematic stimuli. *Journal of Personality and Social Psychology*, *80*, 381–396.

Phelps, E. A. (2006). Emotion and cognition: Insights from studies of the human amygdala. *Annual Review of Psychology*, *57*, 27–53.

Pilarczyk, J., & Kuniecki, M. (2014). Emotional content of an image attracts attention more than visually salient features in various signal-to-noise ratio conditions. *Journal of Vision*, *14*, 4. https://doi.org/10.1167/14.12.4

Pourtois, G., Schettino, A., & Vuilleumier, P. (2013). Brain mechanisms for emotional influences on perception and attention: What is magic and what is not. *Biological Psychology*, *92*, 492–512.

Redies, C., Grebenkina, M., Mohseni, M., Kaduhm, A., & Dobel, C. (2020). Global image properties predict ratings of affective pictures. *Frontiers in Psychology*, *11*, 953.

Reingold, E. M., & Loschky, L. C. (2002). Saliency of peripheral targets in gazecontingent multi-resolutional displays. *Behavior Research Methods, Instruments, & Computers*, *34*, 491–499.

Rhodes, L. J., Ríos, M., Williams, J., Quiñones, G., Rao, P. K., & Miskovic, V. (2019). The role of low-level image features in the affective categorization of rapidly presented scenes. *PLoS One*, *14*, e0215975.

Röhrbein, F., Goddard, P., Schneider, M., James, G., & Guo, K. (2015). How does image noise affect actual and predicted human gaze allocation in assessing image quality? *Vision Research*, *112*, 11–25.

Schupp, H. T., Stockburger, J., Schmälzle, R., Bublatzky, F., Weike, A. I., & Hamm, A. O. (2008). Visual noise effects on emotion perception: Brain potentials and stimulus identification. *NeuroReport*, *19*, 167–171.

Sheikh, H. R., Bovik, A. C., & de Veciana, G. (2005). An information fidelity criterion for image quality assessment using natural scene statistics. *IEEE Transactions on Image Processing*, *14*, 2117–2128.

Todd, R. M., Talmi, D., Schmitz, T. W., Susskind, J., & Anderson, A. K. (2012). Psychophysical and neural evidence for emotion-enhanced perceptual vividness. *Journal of Neuroscience*, *32*, 11201–11212.

Torralba, A. (2009). How many pixels make an image? *Visual Neuroscience*, *26*, 123–131.

van Diepen, P., & d'Ydewalle, G. (2003). Early peripheral and foveal processing in fixations during scene perception. *Visual Cognition*, *10*, 79–100.

Vuilleumier, P. (2005). How brains beware: Neural mechanisms of emotional attention. *Trends in Cognitive Sciences*, *9*, 585–594.

Watson, A. B., & Ahumada, A. J. (2011). Blur clarified: A review and synthesis of blur discrimination. *Journal of Vision*, *11*, 10. https://doi.org/10.1167/11.5.10

Yang, E., Zald, D. H., & Blake, R. (2007). Fearful expressions gain preferential access to awareness during continuous flash suppression. *Emotion*, *7*, 882–886.