Original scientific paper

UDK:
159.937.072
159.942.072

Received: August, 16.2021.
Revised: September, 22.2021.
Accepted: October, 17.2021.

doi: 10.23947/2334-8496-2021-9-3-359-368

Check for updates

Relationship Between Facial Areas With the Greatest Increase in Non-local Contrast and Gaze Fixations in Recognizing Emotional Expressions

Vitali Babenko[1], Denis Yavna[1], Elena Vorobyeva[2], Ekaterina Denisova[3], Pavel Ermakov[1], Ekaterina Kovsh[1]*

[1]Southern Federal University, Rostov-on-Don, Russian Federation,
e-mail: babenko@sfedu.ru, yavna@fortran.su, paver@sfedu.ru, katya-kovsh@yandex.ru
[2]Don State Technical University, Southern Federal University, Rostov-on-Don, Russian Federation,
e-mail: evorob2012@yandex.ru
[3]Don State Technical University, Rostov-on-Don, Russian Federation,
e-mail: denisovakeith@gmail.com

Abstract: The aim of our study was to analyze gaze fixations in recognizing facial emotional expressions in comparison with o the spatial distribution of the areas with the greatest increase in the total (nonlocal) luminance contrast. It is hypothesized that the most informative areas of the image that getting more of the observer's attention are the areas with the greatest increase in nonlocal contrast. The study involved 100 university students aged 19-21 with normal vision. 490 full-face photo images were used as stimuli. The images displayed faces of 6 basic emotions (Ekman's Big Six) as well as neutral (emotionless) expressions. Observer's eye movements were recorded while they were the recognizing expressions of the shown faces. Then, using a developed software, the areas with the highest (max), lowest (min), and intermediate (med) increases in the total contrast in comparison with the surroundings were identified in the stimulus images at different spatial frequencies. Comparative analysis of the gaze maps with the maps of the areas with min, med, and max increases in the total contrast showed that the gaze fixations in facial emotion classification tasks significantly coincide with the areas characterized by the greatest increase in nonlocal contrast. Obtained results indicate that facial image areas with the greatest increase in the total contrast, which preattentively detected by second-order visual mechanisms, can be the prime targets of the attention.

Keywords: face, emotion, eye movements, nonlocal contrast, second-order visual mechanisms.

Introduction

The ability to recognize a facial expression is considered as a component of emotional intelligence and plays important part in human communication, including educational communication (Kosonogov V. et al., 2019; Belousova and Belousova, 2020; Budanova I., 2021). In recent years, symptoms of a disruption of the ability to perceive facial expressions are often a special subject of therapeutic interventions (Skirtach et al., 2019). Contemporary research also acknowledges the genetic influence on functioning of the systems involved in the recognition and experiencing of emotions (Vorobyova et al., 2019).

Previous studies generally confirm that faces are detected and perceived faster than objects of other categories (Liu et al., 2000; Liu, Harris and Kanwisher, 2002; Crouzet, Kirchner and Thorpe, 2010; Crouzet and Thorpe, 2011). A face is not only categorized in a scene in less than 100 ms, but this time is enough to form a first impression of a person (Willis and Todorov, 2006; Cauchoix et al., 2014). MEG studies show the medial prefrontal cortex and amygdala activation in the first 95 ms during differentiating facial expressions (Liu and Ioannides, 2010). It is suggested that the ability to quickly recognize faces is mediated by a special "facial module", and the appearance of a face in the visual field automatically turns on this processing system (Fodor, 1983; 2000; Kanwisher, 2000; Rivolta, 2014).

The ultra-rapid saccades can be initiated to the face faster than to any other objects (Crouzet, Kirchner and Thorpe, 2010). At the same time, the movement of the eye always is the result of shifting attention to a new area in the visual field (Theeuwes, 2014). Moreover, because faces are simply more

*Corresponding author: babenko@sfedu.ru

© 2021 by the authors. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

effective at engaging observer's attention than other objects, they may already have a competitive advantage in the preattentive stage of processing (Reddy, Wilken and Koch, 2004). Prior research substantiates the belief that ultra-rapid face detection is largely determined by the information extracted preattentively (Vuilleumier, 2002; Honey, Kirchner and VanRullen, 2008; Allen, Lien and Jardin, 2017; see also the review by Tamietto and DeGelder, 2010).

Since selecting visual content is determined by the principle of maximizing information (Bruce and Tsotsos, 2005), then in face recognition the gaze distribution shows the most informative areas. These areas are proven to be the eyes, nose and mouth (Luria and Strauss, 1978; Mertens, Siegmund and Grüsser, 1993; Eisenbarth and Alpers, 2011). However, at the preattentive level of visual processing, there are no mechanisms that are selective to facial features. But the human visual system contains preattentive filters called second-order visual mechanisms (see the review by Graham, 2011). These filters are capable of highlighting areas of spatial heterogeneity in images. And it is namely these areas that can be the most informative (Itti, Koch and Niebur, 1998 ; Itti and Koch, 2001; Gao and Vasconcelos, 2007; Gao, Han and Vasconceloset, 2009; Hou et al., 2013).

The existence of second-stage filters at first has been predicted theoretically (Babenko, 1989; Chubb and Sperling, 1989; Sutter, Beck and Graham, 1989) and then was repeatedly attested by numerous experimental studies (e.g. Dakin and Mareschal, 2000; Landy and Oruç, 2002; Kingdom, Prins and Hayes, 2003; Reynaud and Hess, 2012; Babenko and Ermakov, 2015). These mechanisms combine the outputs of first-order visual filters (simple striate neurons) in a certain way and respond to spatial modulations of brightness gradients (their contrast, orientation, or spatial frequency).

Initially it has been assumed that the targets of attention can be local heterogeneities identified by first-order filters (Bergen and Julesz, 1983). However, more recent evidence reveals that higher-level traits have an advantage over lower-level traits in controlling overt attention (Frey, König and Einhäuser, 2007; Açık et al., 2009). So now, it is clear that the targets for attention are probably the extended areas of the image, which differ from the surroundings in their physical characteristics. Based on such differences various models of bottom-up saliency have been promoted over the past two decades (Hou and Zhang, 2007; Valenti, Sebe and Gevers, 2009; Perazzi et al., 2012; Wu et al., 2012; Marat et al., 2013; Xia et al., 2015).

The aim of our work is to analyze gaze fixations in recognizing facial emotional expressions in comparison with to the spatial distribution of the areas with the greatest increase in the total (nonlocal) contrast. The research hypothesis is that the most informative areas of the facial image that getting more of the observer's attention could be the areas with the greatest increase in nonlocal contrast.

Materials and Methods

Participants

The study sample consisted of 100 university students (Europeans, women 59%) aged 19 to 21 years (average age 20.4 ± 2.6). All participants had normal or normalized vision and no history of neurological or psychiatric illness. In the initial stage of the process all participants were informed about the study's purpose and procedure and gave written consent for voluntary participation. The study was approved by the local ethics committee and conducted in accordance with the ethical standards of The Code of Ethics of the World Medical Association (Declaration of Helsinki).

Stimuli

490 full-face photo images were used as stimuli, which were selected from open excess databases: MMI (Pantic et al., 2005), KDEF (Lundqvist, Flykt and Öhman, 1998), Rafd (Langner et al., 2010) and WSEFEP (Olszanowski et al., 2015). The number of male and female faces was equal (245 each). These were the faces of adult Caucasians. The images displayed faces of 6 basic emotions according to P. Ekman (anger, disgust, fear, happiness, sadness and surprise) (Ekman, 1992) and a neutral facial expression. We aligned the images by average brightness and RMS contrast and inscribed them into a conditional circle of 880 pixels in diameter (22.8 angular degrees).

Procedure

Participants were positioned in a head-chin rest at 60 cm distance from the center of the screen. The instruction did not require subjects to fixate gaze prior to the stimuli. The subjects were asked to recognize the emotional expression of the shown face. Images of male and female faces with different emotional expressions were presented in a random sequence. The duration of the stimulus exposure

was 700ms. Verbal labels of all possible facial expressions appeared following each faded stimulus The subjects responded by clicking a mouse button to indicate which emotion they thought was shown. Prior to the experiment, all subjects underwent training that helped to understand the task, procedure and allowed to actualize the names of emotional expressions. Since the differentiation of emotions is a common task for an adult, prolonged training was not required. At first, subjects in free viewing mode went through photographs of men and women showing different facial expressions. Each image was accompanied by a caption indicating the displayed emotion. Then, in order to familiarize the subjects with the procedure and make sure that they understood the task correctly, several training trials were carried out. The images used in the training were not used in the main experiment.

The duration of the main experiment did not exceed 20 minutes, and the experimental task was not tiring. However, since we recorded not only eye movements, but also the responses of the subjects, this allowed us to monitor the development of fatigue during the experiment. Comparing the percentage of correct answers in the first and last third of the experiment, we did not find a significant decrease in the performance efficiency.

Eye-tracking

Eye movements were recorded using the SMI Red-m tracker. The standard calibration procedure for the device was carried out prior to each experiment. The position of the eyes was recorded at a frequency of 60 Hz. The gaze localization accuracy was 30 arc minutes. For each stimulus, a fixation density map (FDM) was constructed by averaging over all subjects.

Digital image processing

Using software we developed that compares the total luminance contrast in the central operator window with the total contrast in the surrounding area, the face image areas with the highest (max), lowest (min), and intermediate (med) increases in the total contrast were established. The med areas were defined on a conditional straight line connecting the nearest min and max regions, while the degree of contrast increase in med was average between these min and max areas.

For digital image processing, we used a concentric operator. The operator included a central area (central window of the operator) and a surrounding ring (peripheral part of the operator). The width of the peripheral ring was equal to the diameter of the central window. First, in the center area of the concentric operator, we calculated the total energy of the image filtered at a frequency of 4 cycles per diameter of this central area. This filtering frequency was set based on the optimal ratio of carrier-envelope frequencies for human perception of contrast modulations (Babenko, Ermakov and Bozhinskaya, 2010; Sun and Schofield, 2011; Li et al., 2014). In the peripheral part of the operator, the spectral power of the entire range of spatial frequencies perceived by a person was calculated per 1 octave on average. The contrast modulation amplitude was equal to the difference in the spectral power calculated between the central and peripheral regions of the operator.

Changing the diameter of the operator's window while maintaining the filtering frequency (4 cycles per window diameter) made it possible to identify these areas in 5 different ranges of spatial frequencies 1 octave wide (with a center frequency of 4, 8, 16, 32 and 64 cycles per image). The relationship between the operator's diameter and the filtering frequency (the smaller the diameter, the higher the frequency) reflects the well-known property of second-order visual mechanisms, which ensures their scale-invariant capabilities (Sutter, Sperling and Chubb, 1995; Kingdom and Keeble, 1999; Dakin and Mareschal, 2000; Landya and Oruç , 2002).

Using the largest gradient operator, where the diameter of its central area equaled the size of the image, we were able to mark one area with the highest, lowest and intermediate modulation of the total contrast in every stimuli. Then, by repeated halving of the operator's diameter, 2, 4, 8 and 16 areas were marked for each contrast modulation amplitude (min, med and max). The total diameter of the identified at different spatial frequencies areas was equal to the diameter of the conditional circle into which the original image was inscribed. For each stimulus 3 maps of the distribution of areas with the min, med and max modulation of contrast were constructed. These maps were a superposition of Gaussians.

Statistical data analysis

The empirical maps (FDMs) were compared with calculated theoretical maps which were a result of digital processing of stimuli. To assess the similarity of the maps, two distribution-based metrics were used: Pearson's linear correlation coefficient (Cc) which shows if there is a linear relationship between two variables; EMD (Earth mover's distance or Wasserstein distance) which is a spatially robust measure that, unlike all other similar metrics, takes into account the spatial differences between theoretical and empirical

results (Bylinskii et al., 2018). To calculate the distance matrix, we used a computer implementation of the similarity metric for the Python language (Pele and Werman, 2009).

Results

First, we compared empirical maps for each of the 490 stimuli with the distribution maps of min, med, and max regions constructed from image areas identified in all five spatial frequency ranges. Due to the non-normal distribution of the data obtained and the heterogeneity of the variances, we used a non-parametric test. The medians of the correlation coefficients for min were -0.109, for med and max were 0.323 and 0.459, respectively. By comparing these scores using the Kruskal-Wallis rank sum test (df = 2, n = 1470), it was found that the similarity of theoretical and empirical maps significantly increases with an increase in the modulation amplitude of the total contrast of the selected areas (p <0.000). The median EMD scores for min, med, and max were 5.266, 3.371, and 3.266, respectively. It also should be noted that the shorter EMD indicated less the similarity between theoretical and empirical maps. The Krus-kal-Wallis rank sum test showed that this similarity significantly increases with the increase in the contrast of the selected areas (p <0.000).

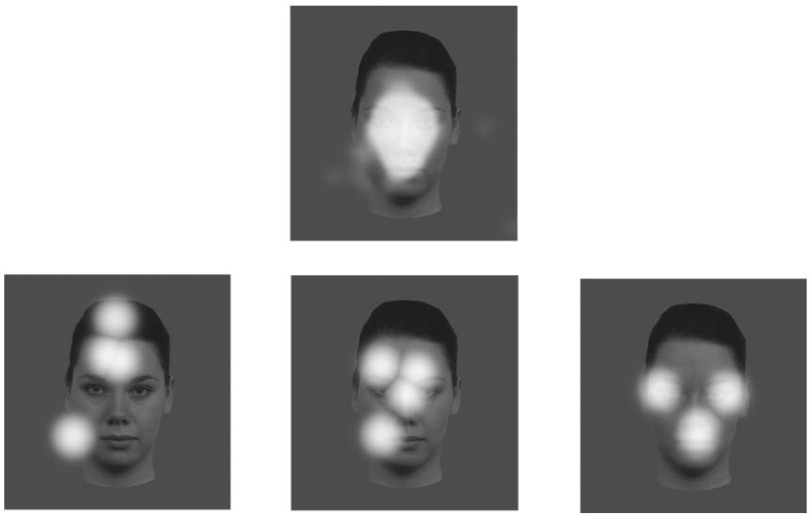

Figure 1. Examples of an empirical FDM (above) and the areas with the lowest (left), intermediate (center) and highest (right) increases in the total contrast, highlighted at a frequency of 16 cycles per image. The brightness level of the selected areas reflects the probability of gaze fixation on a given area of the image.

Then we performed a similar analysis separately for each of the spatial frequency ranges. At this stage, the empirical maps remained the same, and the calculated theoretical maps were built from the areas identified in a narrow range (1 octave) of spatial frequencies with a central frequency of 4, 8, 16, 32 and 64 cycles per image. The correlation analysis results are presented in Table 1.

Table 1.

Median scores of correlation coefficients for maps in different ranges of spatial frequencies and the effect of increasing the amplitude of contrast modulation

Cycles per image	min	med	max	Kruskal-Wallis chi-squared	p
4	-0,089	0,294	0,361	1047,0	< 0.000
8	-0,094	0,122	0,503	795,82	< 0.000
16	-0,019	0,330	0,474	963,63	< 0.000
32	-0,023	0,000	0,000	455,83	< 0.000
64	-0,023	0,066	0,137	811,64	< 0.000

The higher the Kruskal-Wallis chi-squared scores, the more pronounced the differences between the compared values (in this case, the correlation coefficients). Statistical comparison of the obtained scores using the Kruskal-Wallis rank sum test showed that the similarity of theoretical and empirical maps significantly increases with an increase in the contrast modulation amplitude of the selected areas.

Table 2.
EMD scores in different spatial frequency ranges and the effect of increasing the contrast modulation amplitude

Cycles per image	min	med	max	Kruskal-Wallis chi-squared	p
4	6,760	3,755	3,609	1009,5	< 0.000
8	6,061	3,696	2,380	776,72	< 0.000
16	3,519	1,832	1,658	843,0	< 0.000
32	7,197	4,098	4,070	77,079	< 0.000
64	5,601	2,801	2,676	328,08	< 0.000

The results of the EMD analysis (shown in Table 2) were consisted with the previous analysis. These results also support the conclusion that, the higher the increase in the total contrast of the selected areas, the more the calculated maps coincide with the empirical FDMs.

To clarify the results obtained at various spatial frequencies, we conducted a post-hoc pairwise comparison of the values obtained for min, med and max areas, using Conover test (Table 3 and 4).

Table 3.
Post-hoc analysis results for the correlation coefficients

Cycles per image	min-med	p med-max	min-max
4	< 0.0000	< 0.0000	< 0.0000
8	< 0.0000	< 0.0000	< 0.0000
16	< 0.0000	< 0.0000	< 0.0000
32	< 0.0000	= 0.14	< 0.0000
64	< 0.0000	< 0.0000	< 0.0000

Table 4.
Post-hoc analysis results for the EMD

Cycles per image	min-med	p med-max	min-max
4	< 0.0000	< 0.0000	< 0.0000
8	< 0.0000	< 0.0000	< 0.0000
16	< 0.0000	< 0.0000	< 0.0000
32	< 0.0000	= 0.14	< 0.0000
64	< 0.0000	< 0.0000	< 0.0000

The post-hoc analysis showed that the relationship between facial areas with the greatest increase in nonlocal contrast and gaze fixations is disturbed at high spatial frequencies (32 and 64 cpi). It is clear that low and medium spatial frequencies (4, 8 and 16 cpi) are more important for attention control when viewing time is limited. Higher spatial frequencies also seem to be able to direct the observer's attention, but with a longer exposure.

Discussions

The main goal of our study was to test the hypothesis that the most informative facial regions may be the regions with the greatest increase in nonlocal contrast. The results obtained definitively showed that in recognizing emotions on faces the distribution of gaze fixations significantly coincide with the layout of areas with the greatest increase in nonlocal contrast at low and medium spatial frequencies. The similarity of theoretical and empirical maps significantly decreases with a decrease in the amplitude of the contrast modulation in selected areas. This effect has been observed and confirmed comparing maps using both the correlation coefficient and the EMD. This applies to both maps that combine the selected areas from all five octaves, and maps constructed the 1-octave ranges of spatial frequencies.

Consistent with previously stated, image areas with contrast modulation activate second-order visual mechanisms in human vision. But how can the functioning of these mechanisms be related to the organization of eye movements? Based on the fact that image areas that differ from the surroundings in their physical characteristics are more informative (Itti, Koch and Niebur, 1998; Einhauser and Konig, 2003; Honey, Kirchner and VanRullen, 2008; Fuchs et al., 2011), it is logical to assume that the targets of focal attention are the areas with the greatest increase in non-local contrast. Spatially overlapping second-order visual mechanisms are able to automatically find these areas in the image at different levels of resolution. The increase in activation of this mechanisms is proportional increase of contrast modulation in the receptive field of the second order filter. We assume that the more the filter is activated, the higher its ability to draw attention to a certain part of the visual field. As a result, the most activated second-order visual mechanisms become “windows” for attention. Through these windows the higher levels of processing receive information from the preattentive stage.

We believe that the perception of a face goes through certain stages. When a new object appears in the observer’s field of view, a face in particular, the perception begins with separating this object from the background. Since second-order visual mechanisms have receptive fields of different sizes (Sutter, Beck and Graham, 1995; Kingdom and Keeble, 1999; Dakin and Mareschal, 2000; Landy and Oruç, 2002), it is always possible to find among them the one with a field that best matches the size of the appeared face. As a result, this mechanism is centered relatively towards the appeared face. It is tuned to a lower spatial frequency than other, smaller second-order visual mechanisms also involved in facial processing. Therefore, it has an advantage in initiating the saccade. This conclusion is based on the fact that ultra-rapid saccades to faces are initiated precisely by low spatial frequencies (Guyader et al., 2017). Thus, because the low-frequency second-order visual mechanism is centered relative to the face, the initial saccade with a high probability will be directed towards the center of the face. This may explain previously reported tendency of the first saccades to be directed to the geometric center of the presented image (Tatler, 2007; Bindemann, Scheepers and Burton, 2009, 2010; Atkinson and Smithson, 2020). Attention directed to the center of the face allows us to obtain general (low-frequency) information about the configuration of the appeared object and classify it as a face (Meinhardt-Injac, Persike and Meinhardt, 2010; Cauchoix et al., 2014; Comfort and Zana, 2015). As shown in Figure 1, the averaged FDM has a peak in the center of the face (between the nose bridge and the mouth). Moreover, statistical data analysis (Tables 1 and 2) confirms that the empirical map of gaze fixations most closely matches the calculated max map obtained at the lowest spatial frequency.

However, prior research, both the performance results (Leder and Bruce, 1998; Cabeza and Kato, 2000; Collishaw and Hole, 2000; Schwaninger, Lobmaier and Collishaw, 2002; Bombari, Mast and Lobmaier, 2009) and neuroimaging data (Rossion et al., 2000; Harris and Aguirre, 2008; Lobmaier et al., 2008; Betts and Wilson, 2009; Liu, Harris and Kanwisher, 2010), indicate the contribution of not only configural processing, but also feature processing to face recognition. A detailed (featural) description of faces can be performed by second-order visual mechanisms tuned to higher spatial frequencies. These filters, as the frequency setting increases, highlight smaller and smaller parts of the face. It is agreed that the most valuable frequency range for face recognition is from 8 to 32 cycles per face (Nasanen, 1999; Ruiz-Soler and Beltran, 2006; Willenbockel et al., 2010; Collin et al., 2014). As shown in Figure 1 (lower right corner), the areas with the greatest increase in contrast in frequency range from 11 to 22 cpi (the central frequency is 16 cpi) are located in the area of the eyes and mouth - areas that are most informative for the perception of faces (Butler et al., 2010; Peterson and Eckstein, 2012; Smith, Volna and Ewing, 2016; Royer et al., 2018). Therefore, the smaller image areas are highlighted by second-order visual mecha-nisms, the more detailed information is available for analysis at higher processing levels.

Conclusions

The results of the study allow us to conclude that in recognizing emotional facial expressions the higher the luminance contrast of the facial area, the higher the probability that this area will become the object of the observer's attention. It was shown that gaze fixations correlate better with the regions of maximum modulation of nonlocal contrast, containing information from the lower half of the frequency spectrum. Perhaps this can be explained with the fact that in our experiments the viewing time was limited to 700 ms per image. This amount of time is enough to make a decision about emotional expression, but during this time the observer can perform only 2-4 saccades, initiated by low-frequency information. Increasing the exposure time will allow the observer to pay attention to the details of the perceived image and can enhance the connection between gaze fixations and high-frequency information.

In our opinion, spatial modulation of contrast in an image can be extracted by the second-order visual mechanisms. The more the contrast is modulated in their receptive field, the higher their activation is. The higher the activation, the higher the probability of drawing the attention to this area of the visual field. Those mechanisms that are more activated can alternately attract visual attention and initiate saccades towards the areas with the greatest increase in nonlocal contrast, starting with lower spatial frequencies.

The results obtained set perspectives for new studies, where it could be determined the universal role of modulations of nonlocal contrast in the perception of not only faces, but also other objects, as well as examined the role of other spatial modulations of luminance gradients (modulations of orientation or spatial frequency) in bottom-up visual attention control.

The accumulation of experimental data in this field is related to the development of image segmentation algorithms and solving the problem of salience. New knowledge about the regularities and mechanisms of determining "regions of interest" will help to optimize the operations of preliminary processing of input information in artificial vision systems and can be useful in the development of image classification systems using deep learning networks.

Acknowledgements

The study was carried out with the financial support of the Russian Science Foundation (project 20-64-47057).

Conflict of interests

The authors declare no conflict of interest.

References

Açık, A., Onat, S., Schumann, F., Einhäuser, W., & König, P. (2009). Effects of luminance contrast and its modifications on fixation behavior during free viewing of images from different categories. *Vision research, 49*(12), 1541-1553. https://doi.org/10.1016/j.visres.2009.03.011

Allen, P. A., Lien, M. C., & Jardin, E. (2017). Age-related emotional bias in processing two emotionally valenced tasks. *Psychological research, 81*(1), 289-308. https://doi.org/10.1007/s00426-015-0711-8

Atkinson, A. P., & Smithson, H. E. (2020). The impact on emotion classification performance and gaze behavior of foveal versus extrafoveal processing of facial features. *Journal of experimental psychology: Human perception and performance, 46*(3), 292–312. https://doi.org/10.1037/xhp0000712

Babenko, V. V., & Ermakov, P. N. (2015). Specificity of brain reactions to second-order visual stimuli. *Visual neuroscience, 32*. https://doi.org/10.1017/S0952523815000085

Babenko, V. V., Ermakov, P. N., & Bozhinskaya, M. A. (2010). Relationship between the Spatial-Frequency Tunings of the First-and the Second-Order Visual Filters. *Psikhologicheskii Zhurnal, 31*(2), 48-57. (In Russian). https://www.elibrary.ru/download/elibrary_14280688_65866525.pdf

Babenko, V.V. (1989). A new approach to the problem of visual perception mechanisms. In *Problems of Neurocybernetics,* ed. Kogan, A. B., pp. 10–11. Rostov-on-Don, USSR: Rostov University Pub. (In Russian).

Belousova, A., & Belousova, E. (2020). Gnostic emotions of students in solving of thinking tasks. *International Journal of Cognitive Research in Science, Engineering and Education, 8*(2), 27-34. https://doi.org/10.5937/IJCRSEE2002027B

Bergen, J. R., & Julesz, B. (1983). Parallel versus serial processing in rapid pattern discrimination. *Nature, 303*(5919), 696-698. https://doi.org/10.1038/303696a0

Betts, L. R., & Wilson, H. R. (2010). Heterogeneous structure in face-selective human occipito-temporal cortex. *Journal of Cognitive Neuroscience, 22*(10), 2276-2288. https://doi.org/10.1162/jocn.2009.21346

Bindemann, M., Scheepers, C., & Burton, A. M. (2009). Viewpoint and center of gravity affect eye movements to human faces. *Journal of vision, 9*(2), 1-16. http://dx.doi.org/10.1167/9.2.7

Bindemann, M., Scheepers, C., Ferguson, H. J., & Burton, A. M. (2010). Face, body, and center of gravity mediate person detection in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance, 36*(6), 1477.

http://dx.doi.org/10.1037/a0019057
Bombari, D., Mast, F. W., & Lobmaier, J. S. (2009). Featural, configural, and holistic face-processing strategies evoke different scan patterns. *Perception, 38*(10), 1508-1521. https://doi.org/10.1068/p6117
Bruce, N. D. & Tsotsos, J. K. 2005). Saliency based on information maximization. In *Advances in neural information processing systems, 18*, 155-162. http://cs.umanitoba.ca/~bruce/NIPS2005_0081.pdf
Budanova, I. (2021). The Dark Triad of personality in psychology students and eco-friendly behavior. In *E3S Web of Conferences* (Vol. 273, p. 10048). EDP Sciences. https://doi.org/10.1051/e3sconf/202127310048
Butler, S., Blais, C., Gosselin, F., Bub, D., & Fiset, D. (2010). Recognizing famous people. *Attention, Perception, & Psychophysics, 72*(6), 1444-1449. https://doi.org/10.3758/APP.72.6.1444
Bylinskii, Z., Judd, T., Oliva, A., Torralba, A., & Durand, F. (2018). What do different evaluation metrics tell us about saliency models?. *IEEE transactions on pattern analysis and machine intelligence, 41*(3), 740-757. https://doi.org/10.1109/TPAMI.2018.2815601
Cabeza, R., & Kato, T. (2000). Features are also important: Contributions of featural and configural processing to face recognition. *Psychological science, 11*(5), 429-433. https://doi.org/10.1111/1467-9280.00283
Cauchoix, M., Barragan-Jason, G., Serre, T., & Barbeau, E. J. (2014). The neural dynamics of face detection in the wild revealed by MVPA. *Journal of Neuroscience, 34*(3), 846-854. https://doi.org/10.1523/JNEUROSCI.3030-13.2014
Chubb, C., & Sperling, G. (1989). Two motion perception mechanisms revealed through distance-driven reversal of apparent motion. *Proceedings of the National Academy of Sciences, 86*(8), 2985-2989. https://doi.org/10.1073/pnas.86.8.2985
Collin, C. A., Rainville, S., Watier, N., & Boutet, I. (2014). Configural and featural discriminations use the same spatial frequencies: A model observer versus human observer analysis. *Perception, 43*(6), 509-526. https://doi.org/10.1068/p7531
Collishaw, S. M., & Hole, G. J. (2000). Featural and configurational processes in the recognition of faces of different familiarity. *Perception, 29*(8), 893-909. https://doi.org/10.1068/p2949
Comfort, W. E., & Zana, Y. (2015). Face detection and individuation: Interactive and complementary stages of face processing. *Psychology & Neuroscience, 8*(4), 442. https://doi.org/10.1037/h0101278
Crouzet, S. M., & Thorpe, S. J. (2011). Low-level cues and ultra-fast face detection. *Frontiers in psychology, 2*, 342. https://doi.org/10.3389/fpsyg.2011.00342
Crouzet, S. M., Kirchner, H., & Thorpe, S. J. (2010). Fast saccades toward faces: face detection in just 100 ms. *Journal of vision, 10*(4), 16-16. https://doi.org/10.1167/10.4.16
Dakin, S. C., & Mareschal, I. (2000). Sensitivity to contrast modulation depends on carrier spatial frequency and orientation. *Vision research, 40*(3), 311-329. https://doi.org/10.1016/S0042-6989(99)00179-0
Einhäuser, W., & König, P. (2003). Does luminance-contrast contribute to a saliency map for overt visual attention?. *European Journal of Neuroscience, 17*(5), 1089-1097. https://doi.org/10.1046/j.1460-9568.2003.02508.x
Einhäuser, W., Rutishauser, U., Frady, E. P., Nadler, S., König, P., & Koch, C. (2006). The relation of phase noise and luminance contrast to overt attention in complex visual stimuli. *Journal of vision, 6*(11), 1-1. https://doi.org/10.1167/6.11.1
Eisenbarth, H., & Alpers, G. W. (2011). Happy mouth and sad eyes: scanning emotional facial expressions. *Emotion, 11*(4), 860-865. https://doi.org/10.1037/a0022758
Ekman, P. (1992). An argument for basic emotions. *Cognition & emotion, 6*(3-4), 169-200. https://doi.org/10.1080/02699939208411068
Fodor, J. (1983). *Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, Mass: MIT Press.
Fodor, J. A. (2000). *The mind doesn't work that way: The scope and limits of computational psychology.* MIT press. Retrieved from http://www.sscnet.ucla.edu/comm/steen/cogweb/Abstracts/Sutherland_on_Fodor_00.html
Frey, H. P., König, P., & Einhäuser, W. (2007). The role of first-and second-order stimulus features for human overt attention. *Perception & Psychophysics, 69*(2), 153-161. https://doi.org/10.3758/bf03193738
Fuchs, I., Ansorge, U., Redies, C., & Leder, H. (2011). Salience in paintings: bottom-up influences on eye fixations. *Cognitive Computation, 3*(1), 25-36. https://doi.org/10.1007/s12559-010-9062-3
Gao, D., Han, S., & Vasconcelos, N. (2009). Discriminant saliency, the detection of suspicious coincidences, and applications to visual recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 31*(6), 989-1005. https://doi.org/10.1109/TPAMI.2009.27
Gao, D., & Vasconcelos , N. (2007). Bottom-up saliency is a discriminant process . Proceedings / IEEE International Conference on Computer Vision. *IEEE International Conference on Computer Vision*. https://doi.org/10.1109/ICCV.2007. 4408851
Graham, N. V. (2011). Beyond multiple pattern analyzers modeled as linear filters (as classical V1 simple cells): Useful additions of the last 25 years. *Vision research, 51*(13), 1397-1430. https://doi.org/10.1016/j.visres.2011.02.007
Guyader, N., Chauvin, A., Boucart, M., & Peyrin, C. (2017). Do low spatial frequencies explain the extremely fast saccades towards human faces?. *Vision research, 133*, 100-111. https://doi.org/10.1016/j.visres.2016.12.019
Harris, A., & Aguirre, G. K. (2008). The representation of parts and wholes in face-selective cortex. *Journal of Cognitive Neuroscience, 20*(5), 863-878. https://doi.org/10.1162/jocn.2008.20509
Honey, C., Kirchner, H., & VanRullen, R. (2008). Faces in the cloud: Fourier power spectrum biases ultrarapid face detection. *Journal of vision, 8*(12), 9-9. https://doi.org/10.1167/8.12.9
Hou, W., Gao, X., Tao, D., & Li, X. (2013). Visual saliency detection using information divergence. *Pattern Recognition, 46*(10), 2658-2669. https://doi.org/10.1016/j.patcog.2013.03.008
Hou, X., & Zhang, L. (2007, June). Saliency detection: A spectral residual approach. In *2007 IEEE Conference on computer vision and pattern recognition* (pp. 1-8). Ieee. https://doi.org/10.1109/CVPR.2007.383267
Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature reviews neuroscience, 2*(3), 194-203. https://doi.org/10.1038/35058500
Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on pattern analysis and machine intelligence, 20*(11), 1254-1259. https://doi.org/10.1109/34.730558
Kanwisher, N. (2000). Domain specificity in face perception. *Nature neuroscience, 3*(8), 759-763. https://doi.org/10.1038/77664
Kingdom, F. A., & Keeble, D. R. (1999). On the mechanism for scale invariance in orientation-defined textures. *Vision Research, 39*(8), 1477-1489. https://doi.org/10.1016/S0042-6989(98)00217-X

Kingdom, F.A.A., Prins, N., & Hayes, A. (2003). Mechanism independence for texture-modulation detection is consistent with a filter-rectify-filter mechanism. *Vis. Neurosci., 20*, 65-76. https://doi.org/10.1017/s0952523803201073
Kosonogov , V., Vorobyeva , E., Kovsh , E., & Ermakov , P. (2019). A review of neurophysiological and genetic correlates of emotional intelligence. *International Journal of Cognitive Research in Science, Engineering and Education (IJCRSEE), 7*(1), 137–142. https://doi.org/10.5937/ijcrsee1901137K
Landy, M. S., & Oruç, I. (2002). Properties of second-order spatial frequency channels. *Vision research, 42*(19), 2311-2329. https://doi.org/10.1016/S0042-6989(02)00193-1
Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H., Hawk, S. T., & Van Knippenberg, A. D. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and emotion, 24*(8), 1377-1388. https://doi.org/10.1080/02699930903485076
Leder, H., & Bruce, V. (1998). Local and Relational Aspects of Face Distinctiveness. *The Quarterly Journal of Experimental Psychology Section A, 51*(3), 449–473. https://doi.org/10.1080/713755777
Li, G., Yao, Z., Wang, Z., Yuan, N., Talebi, V., Tan, J., ... & Baker, C. L. (2014). Form-cue invariant second-order neuronal responses to contrast modulation in primate area V2. *Journal of Neuroscience, 34*(36), 12081-12092. https://doi.org/10.1523/JNEUROSCI.0211-14.2014
Liu, J., Harris, A., & Kanwisher, N. (2002). Stages of processing in face perception: an MEG study. *Nature neuroscience, 5*(9), 910-916. https://doi.org/10.1038/nn909
Liu, J., Harris, A., & Kanwisher, N. (2010). Perception of face parts and face configurations: an fMRI study. *Journal of cognitive neuroscience, 22*(1), 203-211. https://doi.org/10.1162/jocn.2009.21203
Liu, J., Higuchi, M., Marantz, A., & Kanwisher, N. (2000). The selectivity of the occipitotemporal M170 for faces. *Neuroreport, 11*(2), 337-341. https://doi.org/0.1097/00001756-200002070-00023
Liu, L., & Ioannides, A. A. (2010). Emotion separation is completed early and it depends on visual field presentation. *PloS one, 5*(3), e9790. https://doi.org/10.1371/journal.pone.0009790
Lobmaier, J. S., Klaver, P., Loenneker, T., Martin, E., & Mast, F. W. (2008). Featural and configural face processing strategies: evidence from a functional magnetic resonance imaging study. *Neuroreport, 19*(3), 287-291. https://doi.org/10.1097/WNR.0b013e3282f556fe
Lundqvist, D., Flykt, A., & Öhman, A. (1998). The Karolinska directed emotional faces (KDEF). CD ROM from Department of Clinical Neuroscience, Psychology section, *Karolinska Institutet, 91*(630), 2-2.
Luria, S. M., & Strauss, M. S. (1978). Comparison of Eye Movements over Faces in Photographic Positives and Negatives. *Perception, 7*(3), 349–358. https://doi.org/10.1068/p070349
Marat, S., Rahman, A., Pellerin, D., Guyader, N., & Houzet, D. (2013). Improving visual saliency by adding 'face feature map'and 'center bias'. *Cognitive Computation, 5*(1), 63-75. https://hal.archives-ouvertes.fr/hal-00703762
Meinhardt-Injac, B., Persike, M., & Meinhardt, G. (2010). The time course of face matching by internal and external features: Effects of context and inversion. *Vision Research, 50*(16), 1598-1611. https://doi.org/10.1016/j.visres.2010.05.018
Mertens, I., Siegmund, H., & Grüsser, O. J. (1993). Gaze motor asymmetries in the perception of faces during a memory task. *Neuropsychologia, 31*(9), 989-998. https://doi.org/10.1016/0028-3932(93)90154-R
Näsänen, R. (1999). Spatial frequency bandwidth used in the recognition of facial images. *Vision research, 39*(23), 3824-3833. https://doi.org/10.1016/s0042-6989(99)00096-6
Olszanowski, M., Pochwatko, G., Kuklinski, K., Scibor-Rylski, M., Lewinski, P., & Ohme, R. K. (2015). Warsaw set of emotional facial expression pictures: a validation study of facial display photographs. *Frontiers in psychology, 5*, 1516. https://doi.org/10.3389/fpsyg.2014.01516
Pantic, M., Valstar, M., Rademaker, R., & Maat, L. (2005, July). Web-based database for facial expression analysis. In *2005 IEEE international conference on multimedia and Expo* (pp. 5-pp). IEEE. https://doi.org/10.1109/ICME.2005.1521424
Pele, O., & Werman, M. (2009, September). Fast and robust earth mover's distances. In *2009 IEEE 12th international conference on computer vision* (pp. 460-467). IEEE. https://doi.org/10.1109/ICCV.2009.5459199
Perazzi, F., Krähenbühl, P., Pritch, Y., & Hornung, A. (2012, June). Saliency filters: Contrast based filtering for salient region detection. In *2012 IEEE conference on computer vision and pattern recognition* (pp. 733-740). IEEE. https://doi.org/10.1109/CVPR.2012.6247743
Peterson, M. F., & Eckstein, M. P. (2012). Looking just below the eyes is optimal across face recognition tasks. *Proceedings of the National Academy of Sciences, 109*(48), E3314-E3323. https://doi.org/10.1073/pnas.1214269109
Reddy, L., Wilken, P., & Koch, C. (2004). Face-gender discrimination is possible in the near-absence of attention. *Journal of vision, 4*(2), 106-117. https://doi.org/10.1167/4.2.4
Reynaud, A., & Hess, R. F. (2012). Properties of spatial channels underlying the detection of orientation-modulations. *Experimental brain research, 220*(2), 135-145. https://doi.org/10.1007/s00221-012-3124-6
Rivolta, D. (2014). Cognitive and neural aspects of face processing. In *Prosopagnosia* (pp. 19-40). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-40784-0_2
Rossion, B., Dricot, L., Devolder, A., Bodart, J. M., Crommelinck, M., Gelder, B. D., & Zoontjes, R. (2000). Hemispheric asymmetries for whole-based and part-based face processing in the human fusiform gyrus. *Journal of cognitive neuroscience, 12*(5), 793-802. https://doi.org/10.1162/089892900562606
Royer, J., Blais, C., Charbonneau, I., Déry, K., Tardif, J., Duchaine, B., ... & Fiset, D. (2018). Greater reliance on the eye region predicts better face recognition ability. *Cognition, 181*, 12-20. https://doi.org/10.1016/j.cognition.2018.08.004
Ruiz-Soler, M., & Beltran, F. S. (2006). Face perception: An integrative review of the role of spatial frequencies. *Psychological Research, 70*(4), 273-292. https://doi.org/10.1007/s00426-005-0215-z
Schwaninger, A., Lobmaier, J. S., & Collishaw, S. M. (2002). Role of featural and configural information in familiar and unfamiliar face recognition. *Lecture Notes in Computer Science, 2525,* 643–650. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-36181-2_64
Skirtach, I.A., Klimova, N.M., Dunaev, A.G., & Korkhova, V.A. (2019). Effects of rational psychotherapy on emotional state and cognitive attitudes of patients with neurotic disorders. *Trends in the development of psycho-pedagogical education in the conditions of transitional society (ICTDPP-2019)*, 09011. https://doi.org/10.1051/SHSCONF/20197009011
Smith, M. L., Volna, B., & Ewing, L. (2016). Distinct information critically distinguishes judgments of face familiarity and identity. *Journal of Experimental Psychology: Human Perception and Performance, 42*(11), 1770. https://doi.org/10.1037/

xhp0000243
Sun, P., & Schofield, A. J. (2011). The efficacy of local luminance amplitude in disambiguating the origin of luminance signals depends on carrier frequency: Further evidence for the active role of second-order vision in layer decomposition. *Vision research, 51*(5), 496-507. https://doi.org/10.1016/j.visres.2011.01.008
Sutter, A., Beck, J., & Graham, N. (1989). Contrast and spatial variables in texture segregation: Testing a simple spatial-frequency channels model. *Perception & psychophysics, 46*(4), 312-332. https://doi.org/10.3758/BF03204985
Sutter, A., Sperling, G., & Chubb, C. (1995). Measuring the spatial frequency selectivity of second-order texture mechanisms. *Vision Research*, 35(7), 915– 924. https://doi.org/10.1016/0042-6989(94)00196-S
Tamietto, M., & De Gelder, B. (2010). Neural bases of the non-conscious perception of emotional signals. *Nature Reviews Neuroscience, 11*(10), 697-709. https://doi.org/10.1038/nrn2889
Tatler, B. W. (2007). The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions. *Journal of vision, 7*(14). http://dx.doi.org/10.1167/7.14.4
Theeuwes, J. (2010). Top–down and bottom–up control of visual selection. *Acta psychologica, 135*(2), 77-99. https://doi.org/10.1016/j.actpsy.2010.02.006
Theeuwes, J. (2014). Spatial orienting and attentional capture. *The Oxford handbook of attention*, 231-252. https://doi.org/10.1093/oxfordhb/9780199675111.013.005
Valenti, R., Sebe, N., & Gevers, T. (2009, September). Image saliency by isocentric curvedness and color. In *2009 IEEE 12th international conference on Computer vision* (pp. 2185-2192). IEEE. https://doi.org/10.1109/ICCV.2009.5459240
Vorobyeva, E., Hakunova, F., Skirtach, I., & Kovsh, E. (2019). A review of current research on genetic factors associated with the functioning of the perceptual and emotional systems of the brain. In *SHS Web of Conferences* (Vol. 70, p. 09009). EDP Sciences. https://doi.org/10.1051/SHSCONF/20197009009
Vuilleumier, P. (2002). Facial expression and selective attention. *Current Opinion in Psychiatry, 15*(3), 291-300. https://doi.org/10.1097/00001504-200205000-00011
Willenbockel, V., Fiset, D., Chauvin, A., Blais, C., Arguin, M., Tanaka, J. W., ... & Gosselin, F. (2010). Does face inversion change spatial frequency tuning?. *Journal of Experimental Psychology: Human Perception and Performance, 36*(1), 122. https://doi.org/10.1037/a0016465
Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological science, 17*(7), 592-598. https://doi.org/10.1111/j.1467-9280.2006.01750.x
Wu, J., Qi, F., Shi, G., & Lu, Y. (2012). Non-local spatial redundancy reduction for bottom-up saliency estimation. *Journal of Visual Communication and Image Representation, 23*(7), 1158-1166. https://doi.org/10.1016/j.jvcir.2012.07.010
Xia, C., Qi, F., Shi, G., & Wang, P. (2015). Nonlocal center–surround reconstruction-based bottom-up saliency estimation. *Pattern Recognition, 48*(4), 1337-1348. https://doi.org/10.1016/j.patcog.2014.10.007