Original Scientific Paper https://doi.org/10.24867/GRID-2026-p51

Heatmap-Based Area Analysis of Frontal Face Images Using an Eye-Tracking System

ABSTRACT

Facial images are a ubiquitous component of modern web platforms and mobile applications. The widespread use of eye-tracking technology has substantially increased research interest in facial perception and memory. A fundamental question is how facial recognition performance is influenced by observation time and by attention directed towards specific facial features. Memory experiments generally consist of two phases: an observation phase and a recognition phase. To investigate the effect of observation time on recognition accuracy, four image presentation durations were defined during the observation phase: 1 s, 2 s, 4 s, and 8 s. The results showed that correct recognition improved and incorrect recognition decreased as observation time increased. However, a plateau was observed at 4 s, as recognition accuracy in the 8 s condition was nearly identical to that in the 4 s condition, suggesting a threshold beyond which additional observation time provides limited benefit. To validate these findings, heatmap-based analysis was employed. This analysis confirmed the existence of the 4 s threshold and provided evidence for the observed trend. Both the area and perimeter of gaze regions increased with observation time up to 4 s and remained virtually unchanged thereafter. Similarly, the circularity of gaze regions decreased up to the 4 s condition and then stabilised, exhibiting comparable values in both the 4 s and 8 s conditions. These findings demonstrate the utility of heatmap analysis for investigating face observation and recognition processes and support the existence of a critical observation period of approximately 4 s for effective facial memorisation and recognition.

Andrej Iskra
Helena Gabrijelčič Tomc

University of Ljubljana, Faculty of Natural Sciences and Engineering, Department of Textiles, Graphic Arts and Design, Chair of Information and Graphic Art Technology, Ljubljana, Slovenia

KEY WORDS
Frontal facial images, eye tracking, face observation, face recognition success, heatmap analysis.

1. Introduction

Face memory and recognition under real-world conditions are influenced by a variety of factors. Among the most important are:

- face observation time,
- viewing distance,
- viewing angle,
- facial emotional expression.

Research on face memory and recognition can be conducted using facial images as experimental stimuli. Numerous facial image databases have been developed for this purpose. By combining different facial images with controlled experimental conditions, these databases make it possible to investigate the effects of the parameters listed above. Face observation time can be manipulated by varying the duration for which facial images are presented on a display device, most commonly a computer monitor. Similarly, different viewing distances encountered in real-world situations can be simulated by varying the displayed size of the facial images. Fur-

thermore, facial images can be acquired from different viewing angles and with different emotional expressions.

In the present study, we focused specifically on frontal facial images with neutral expressions, presented under conditions designed to approximate natural face observation at a real-world viewing distance of one meter. The main experimental variable was the presentation duration of the facial images, with particular emphasis on determining how presentation time affects subsequent face memory and recognition performance.

A considerable number of studies have previously investigated face memory and recognition (Goldinger et al., 2009, Leyk et al., 2008, Reynolds & Przdek, 1992). In most cases, recognition performance has been quantified using the measure of recognition success A′. Previous studies have also frequently examined recognition performance using an internal facial features approach in which the relative amount of visual attention directed toward the internal facial features is analysed. These regions primarily include the eyes, nose, and mouth, collectively referred to as internal facial features.

In the present study, we introduce a new heatmap-based area analysis method based on heatmaps generated from eye-tracking data. Relatively little research has been conducted on the quantitative analysis of heatmaps (Fux et al., 2024, Jia et al., 2024). The newly developed method was used to provide an additional means of evaluating and validating recognition-performance results. By analysing the spatial distribution of visual attention across facial regions, the proposed approach confirms conventional measures of face recognition performance and feature-based eye-tracking analysis.

Accordingly, the overall study was divided into two experiments, each designed to investigate a specific aspect of the relationship between facial image presentation time, recognition success A′ and heatmap analysis.

2. Methods

2.1. Participants

Participants were recruited from the student population of the Faculty of Natural Sciences and Engineering, University of Ljubljana (Ljubljana, Slovenia). A total of 42 participants took part in the study and were divided into three experimental groups (A, B, and C), with 14 participants in each group. All participants had normal or corrected-to-normal vision and were between 19 and 22 years of age. Participation was voluntary, and participants received a small amount of additional credit toward their laboratory course grade as compensation for their involvement.

2.2. Apparatus

Our research was performed eye tracking system. We used stand-alone eye tracking system Tobii X120. Analysis was done in Tobii Studio 3.4.5 software (User's Manual Tobii Studio, 2016). It is stand-alone eye tracking unit designed for eye tracking studies of real-world flat surfaces.

2.3. Stimuli

For the experiment, we used the publicly available Minear and Park (Minear & Park, 2000) face image database, developed in 2003 at the University of Michigan, Center for Vital Longevity Face Database. The database contains facial images depicting both neutral and emotional expressions, presented from two different viewpoints: frontal and profile views. For the purposes of the present study, only frontal-view facial images with neutral expressions were selected. The images had a resolution of 1280 × 960 pixels. Only facial images of individuals aged 18–29 years were included, ensuring that the age range of the depicted individuals was comparable to that of the study participants.

2.4. Procedure

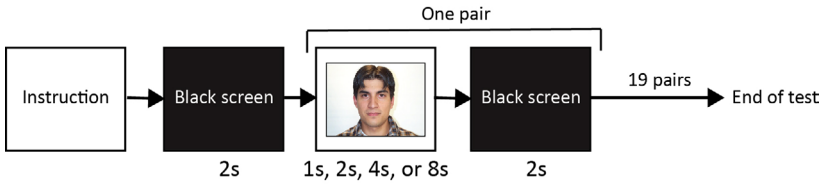
The experiment consisted of two phases: an observation phase and a recognition phase. Four stimulus presentation durations were used: 1, 2, 4, and 8 s, corresponding to four experimental conditions (Test1s, Test2s, Test4s, and Test8s). Each test comprised 20 randomly presented facial images, with an equal number of male and female faces (10 each).

The experimental procedure was based on previously well-known YES/NO recognition memory tasks. Following the initial instructions, a blank black screen was displayed for 2 s before the first facial stimulus appeared. Each subsequent facial image was preceded by a 2-s black screen, which served as a neutral interstimulus interval. The facial images were presented for the predefined duration corresponding to the respective experimental condition (1, 2, 4, or 8 s).

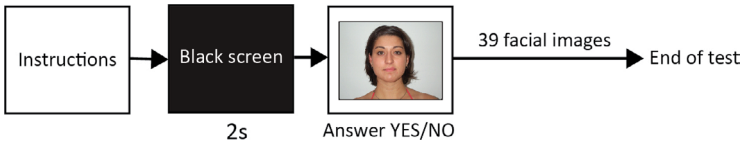
Stimulus presentation was fully automated, and participants had no control over the timing or sequence of the stimuli. The observation phase procedure is illustrated in Figure 1.

The recognition phase comprised 40 facial images: 20 previously presented during the observation phase (old stimuli) and 20 previously unseen images (new stimuli). Following the instructions and an initial 2-s black screen, facial images were presented individually.

For each image, participants indicated whether they had seen the face during the preceding observation



» **Figure 1:** Procedure of the observation phase



» **Figure 2:** Procedure of the recognition phase

phase by providing a yes/no recognition response. After responding, participants advanced to the next image by clicking the mouse button. Unlike the observation phase, the recognition phase was self-paced and had no imposed time limit. The recognition phase procedure is illustrated in Figure 2.

Participants' responses were recorded manually by the experimenter and subsequently transferred to an electronic spreadsheet for further analysis.

2.5. Experiment 1

As previously described, the first experiment examined face recognition performance as a function of stimulus presentation duration. Recognition performance was quantified using the discrimination index (Stanislaw & Todorov, 1999), calculated according to the following equation 1:

$$A' = 0{,}5 + \frac{(CR-FA)(1+CR-FA)}{4CR(1-FA)} \quad (1)$$

where:

- CR represents the proportion of correctly recognized previously presented faces, that is, faces presented during the observation phase and subsequently correctly identified as previously seen during the recognition phase.
- FA represents the proportion of false recognitions, that is, faces not presented during the observation phase but incorrectly identified as previously seen during the recognition phase.

We hypothesized that recognition performance would increase with stimulus presentation duration following a logarithmic relationship, with larger improvements in recognition performance occurring at shorter presentation durations and progressively smaller improvements as presentation duration increased.

2.6. Experiment 2

To further evaluate the results obtained in Experiments 1, a new heatmap-based area analysis was performed using eye-tracking heatmaps. Heatmaps were generated in Tobii Studio based on the recorded gaze data. The principal parameter affecting heatmap generation was the size of the gaze representation, which could be adjusted incrementally. A value of 105 pixels was selected as the optimal setting as it provided sufficiently localized gaze regions while preserving differences associated with the different facial stimulus presentation durations.

Although heatmaps are conventionally represented using a colour scale ranging from red through yellow to green, they were converted to grayscale directly in Tobii Studio. Consequently, gaze regions were represented by different grayscale intensity levels, ranging from black and dark grey to light grey.

The resulting grayscale heatmap images were subsequently imported into ImageJ for quantitative image analysis. The primary parameters of interest were the area, perimeter, and circularity of the gaze regions. Prior to measurement, the grayscale heatmaps were converted into binary (black-and-white) images, allowing the gaze regions to be segmented from the background and analysed as discrete regions. An example of the original colour heatmap, the corresponding grayscale image, and the resulting binary image is shown in Figure 3.



» **Figure 3:** Colour, grayscale, and binary heatmap image

The area, perimeter, and circularity of the resulting binary gaze regions were subsequently analysed. Area and

perimeter were measured in pixels. Circularity, a dimensionless parameter, was used to characterize the spatial dispersion of gaze across the facial image. Circularity was calculated according to Equation 2, where A represents the area and P the perimeter of the gaze region:

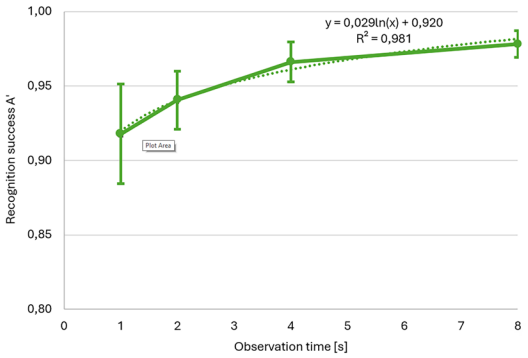
$$Circularity = \frac{4\pi A}{P^2} \quad (2)$$

Higher circularity values indicate more compact and circular gaze regions, whereas lower values indicate more irregular or spatially dispersed gaze patterns.

3. Results

3.1. Experiment 1

Figure 4 shows recognition performance for four different image presentation durations in the observation tests.



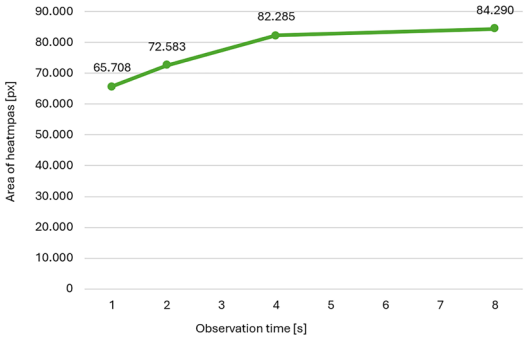
» **Figure 4:** *Results of recognition success*

It can be seen from the figure that the measured results fit the logarithmic curve very well, as indicated by the high R^2 value of 0.981.

The standard deviation decreases with increasing facial image observation time during the observation test, suggesting that longer viewing durations lead to greater similarity in the observation patterns adopted by individual participants. Conversely, shorter facial image presentation times appear to promote a wider range of observation strategies, resulting in greater inter-individual variability in recognition performance. Consequently, recognition performance shows greater dispersion, as indicated by a higher standard deviation (SD).

3.2. Experiment 2

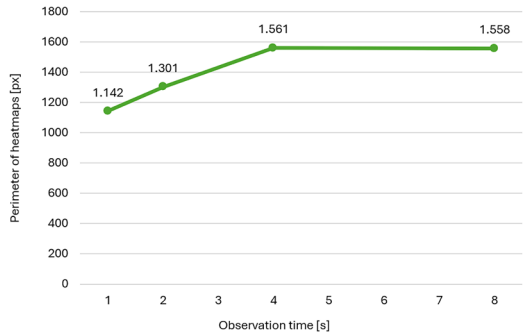
In this experiment, we focused on the analysis of heatmaps and their three parameters: area, perimeter, and circularity. The results for the area of heatmaps as a function of the observation time of facial images are presented in Figure 5.



» **Figure 5:** *Results of area of heatmaps*

The results indicate that the area of heatmaps increases as the observation time of facial images increases, up to approximately 4 seconds. Beyond this point, further increases in observation time produce little additional increase the area of heatmaps.

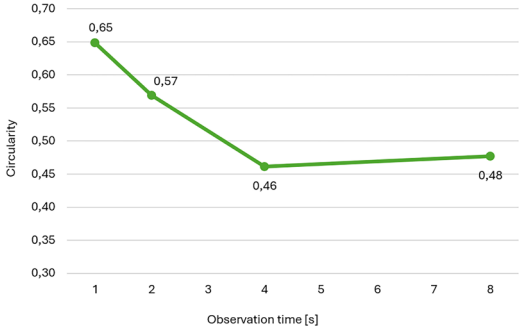
The relationship of perimeter of heatmaps and the observation time of facial images is shown in figure 6.



» **Figure 6:** *Results of perimeter of heatmaps*

Similar to the area, the perimeter of the heatmaps reaches a plateau at an observation time of 4 seconds. Interestingly, at an observation time of 8 seconds, the perimeter is slightly smaller than that measured in the 4-second test.

Figure 7 shows the circularity of the viewing regions as a function of the observation time of facial images.



» **Figure 7:** *Results of circularity of heatmaps*

A similar pattern can be observed in the circularity of heatmaps, with a distinct transition occurring at an observation time of 4 seconds.

4. Discussion

4.1. Experiment 1

The results of the first experiment show a logarithmic relationship between recognition success and the presentation time of facial images in the observation test. The recognition success graph also reveals the most pronounced change and the largest deviation from the fitted logarithmic curve at an observation time of 4 seconds.

These results suggest that an observation time of approximately 4 seconds represents a potential threshold or transition point.

4.2. Experiment 2

Using the newly introduced method for analysing heatmaps, the results for the area and perimeter of these regions showed that both measures increased with increasing observation time of the facial images. This is an expected result, as longer observation times allowed the participants to visually explore a larger area of the face.

However, this increase was observed only up to an observation time of 4 s. Beyond this point, the area of the observed regions remained practically unchanged despite the longer presentation time of the facial images. In our case, area for the 8 s test were almost identical to those obtained for the 4 s test.

The perimeter, however, was slightly smaller at 8 s. In some cases, regions that had previously been separated into distinct areas associated with the eyes and the nose and mouth merged into a single region during the 8 s condition. As a result, the overall area increased slightly, whereas the perimeter of the resulting merged region was smaller than the combined perimeter of the two previously separate regions. An example of this phenomenon is shown in Figure 8.

A similar pattern was observed for circularity, which decreased with increasing observation time up to 4 s and then showed a slight increase at 8 s. This behaviour can also be explained based on the circularity equation presented in the previous section.



» **Figure 8:** *Heatmaps of 4 seconds and 8 seconds test*

5. Conclusions

In this study, we investigated the effect of facial image presentation time on face memorization. The results showed that memorization improved by increasing presentation time, reaching a turning point at approximately 4 seconds. Under the conditions of the present experiment, this duration appeared to be sufficient for participants to memorize the presented faces and achieve a reliable level of recognition performance. Extending the presentation time beyond 4 seconds did not result in further improvements in memorization or recognition performance.

The findings were also supported by the newly developed heatmap-based area method used to analyse heatmaps of viewing areas. All three parameters, area, perimeter, and circularity, showed a clear turning point at 4 seconds. Up to this presentation time, area and perimeter increased, while circularity decreased. However, the values obtained at 8 seconds presentation time were very similar to those observed at 4 seconds, indicating that the spatial characteristics of the viewing areas remained largely unchanged with additional viewing time. This pattern is consistent with the recognition success results, further suggesting that extending viewing time beyond 4 seconds provides little additional benefit for face recognition.

In future studies, this heatmap-based area method could also be applied to investigate the recognition of facial expressions of emotion. Although different theoretical frameworks propose somewhat different classifications of basic emotions, commonly investigated facial expressions include happiness, anger, sadness, fear, and surprise. Future research will therefore focus on how different facial expressions influence users' visual attention and face-recognition performance.

6. References

Fux, M., Monshi, J., Jang, H., Lahey, C. H., Arslan, S. S., Dixon III, W. V., Groth, M. and Sinha, P. (2024) 'Comparing Human Eye-Tracking Heatmaps with DNN Saliency Maps for Faces at Different Spatial Frequencies', *Journal of Vision*, 24(10), 1329. DOI: https://doi.org/10.1167/jov.24.10.1329

Goldinger, S. D., He, Y. and Papesh, M. H. (2009) 'Deficits in cross-race face learning: insights from eye movements and pupillometry', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(5), pp. 1105–1122. DOI: https://doi.org/10.1037/a0016548

Jia, J., Zhao, T., Yang, J. and Wang, Q. (2024) 'Intelligent Evaluation Method for Design Education and Comparison Research between Visualizing Heat-Maps of Class Activation and Eye-Movement', *Journal of Eye Movement Research*, 17(2). DOI: https://doi.org/10.16910/jemr.17.2.1

Leyk, D., Sievert, A., Heiss, A., Gorges, W., Ridder, D., Alexander, T., Wunderlich, M. and Ruter, T. (2008) 'Validation of a short-term memory test for the recognition of people and faces', *Ergonomics*, 51(8), pp. 1125–1136. DOI: 10.1080/00140130802094371

Minear, M. and Park, D. (2004) 'A lifespan database of adult facial stimuli', *Behavior Research Methods, Instruments & Computers*, 36(4), pp. 630–633. DOI: 10.3758/bf03206543

Reynolds, J. K. and Przdek, K. (1992) 'Face recognition memory: the effects of exposure duration and encoding instruction', *Applied Cognitive Psychology*, 6, pp. 279–292. DOI: https://doi.org/10.1002/acp.2350060402

Stanislaw, H. and Todorov, N. (1999) 'Calculation of signal detection theory measures', *Behavior Research Methods, Instruments, & Computers*, 31, pp. 137–149. DOI: https://doi.org/10.3758/BF03207704

Tobii Connect (2016) *User's Manual Tobii Studio version 3.4.5*. Danderyd: Tobii Technology.

© 2026 Authors. Published by the University of Novi Sad, Faculty of Technical Sciences, Department of Graphic Engineering and Design. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license 4.0 Serbia (https://creativecommons.org/licenses/by/4.0/deed.en).