https://doi.org/10.24867/GRID-2024-p41 Original scientific paper

HEATMAP ANALYSIS OF PROFILE FACIAL IMAGES

Andrej Iskra , Helena Gabrijelčič Tomc
University of Ljubljana, Faculty of Natural Sciences and Engineering,
Chair of Information and Graphic Art Technology, Ljubljana, Slovenia

Abstract: *Facial images are a very common element on the World Wide Web and in mobile applications. With the widespread use of eye-tracking systems, research in the field of viewing and memorizing facial images has greatly expanded. The main issue in analyzing facial images is to understand the way of viewing, memorizing and recognizing. Most of the research is done for the frontal facial images, but we focused on profile facial images. One of the most widely used methods for determining memorization and recognition of facial images is the internal facial feature method. However, in our research we have focused on the new heatmap method. In this method, the results of area, perimeter and circularity of heatmaps are analyzed. For the whole research we used a standardized memory test comprising two phases: an observation test and a recognition test. To analyze the influence of time on recognition success, we set four display times of facial images in observation test (1 s, 2 s, 4 s, 8 s). The results show that the recognition success increases with increasing display time of the facial images in the observation test. The turning point occurs at a display time of 4 seconds, as the recognition performance in the 8 s test is almost the same as in the 4 s test. The heatmap analysis method shows an increase in the area and perimeter of the heatmap areas up to the 4 s test, after which both values remain practically the same. After that, both values remain practically the same. At this point, the circularity of the viewing areas also no longer decreases and is the same in the 8 s test as in the 4s test. The results of the heatmap analysis are therefore in good correlation with the results of the facial recognition and proved the usefulness of heatmap analysis for research into the observation and recognition of facial images.*

Key words: profile facial images, observation test, recognition test, eye tracking, heatmap

1. INTRODUCTION

In visual perception, we are dealing with two types of eye movements: fixations and saccades (Harrington, 1981; Robinson, 1963). Fixations are quick stops of the gaze during the process of observation (Abdallahi et al., 2007). A quick jump between one fixation and another is called a saccade. These movements are very fast and are caused by the user's desire to move the gaze (Abdallahi et al., 2007). For fixations, the duration of the fixation is important, for saccades its length.

When using eye-tracking systems, we essentially obtain two results: Gaze plots (duration of fixations and length of saccades) and heatmaps. Figure 1 shows the average of both.

Figure 1: Gazeplot (Tobiipro, 2024) and Heatmap (Ovano, 2012)

As can be seen from Figure 1, the gazeplot is a sequence of eye stops (fixations) and movements between them (saccades). The size of the fixation circle provides information about the duration of the fixation at the point where the eye stops. The heatmap provides information about the time the eyes spend in certain areas. Figure 1 shows that the eye stayed the longest on the red-coloured areas, slightly less on the yellow and least on the green.

All research in the field of facial image viewing is based on analysis, i.e. the duration of fixations and the length of saccades. These analyses usually also include the use of area of interest (AOI), with which we define areas in the image where we analyse the above results.
So far, however, we have not come across any research that includes Heatmap analysis. The first obstacle is that these images are usually in colour. In our case, we first converted the heatmaps to grayscale and then to black and white. These areas were then analysed by calculating the area, volume and circularity of these areas. The results were compared with the results of recognition success and false recognition.

2. METHODS

2.1 Participants

As mentioned above, we conducted 4 tests for each observation time. Every test was done for 6 participants, so all together we recruited 24 participants (7 male and 17 female). They were our students and had normal vision. Average age was 20,3 years (SD = 0,96).

2.2 Stimuli

We took the images from the Minear and Park Face Database (Minear & Park, 2000). The facial images were prepared according to the conditions of natural observation (Kealey, Sekuler & Bennet, 2008; Henderson et al., 2001). The conditions of natural observation mean that we look at a real face at a distance of one meter. Since the distance of the participants to the test screen was 60cm, the faces on the screen were displayed at 60% of their natural size. The average height of a face in nature is 18.5cm (Park Aging Mind Laboratory, 2024), and according to this calculation, the faces on our screen were on average 11.1cm high. Depending on the screen resolution and the size of the faces in the facial images, we prepared our facial images at a size of 640 × 480px. The set of images was equally selected by gender (10 male and 10 female). For the recognition test we added 20 new faces to these faces (also 10 male and 10 female).

2.3 Apparatus

All tests were carried out in the Laboratory of Visual Perception and Colorimetry at the Department of Textile, Graphic and Design of the Faculty of Natural Sciences and Engineering at the University of Ljubljana. Setting up the environmental and testing system was based on the standards and recommendations (Pernice & Nielsen, 2009). We performed the test with the Tobii X-120 eye tracking system. Analysis were done in Tobii Studio 3.4.5 software. The defaults setting for definition of fixation was 100 ms for 30 px area. That means if eyes stayed in the area 30 pixel for at least 100 ms it was concerned as one fixation (Tobii Connect, 2016).

2.4 Procedure

We have conducted a memory test according to the previously mentioned YES/NO principle. We performed two tests, an observation test, and a recognition test. The procedure observation test is shown in Figure 2.

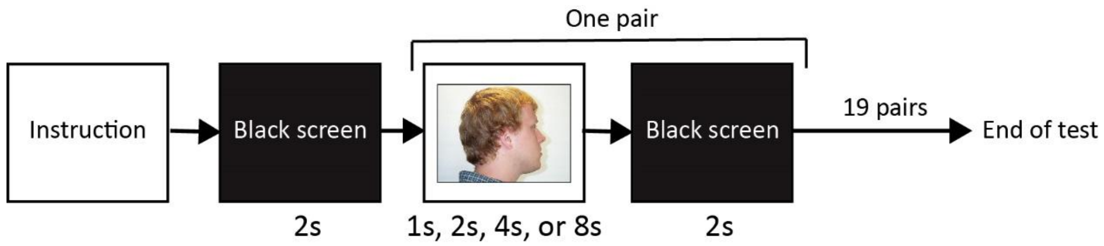

Figure 2: Procedure of the observation test

The observation test included 20 facial images (10 male and 10 female). According to the procedure in Figure 1, the tests lasted 1 minute (Test1s), 1 minute 20 seconds (Test2s), 2 minutes (Test4s), and 3 minutes 20 seconds (Test8s). The observation test was followed by the recognition test (Figure 3).

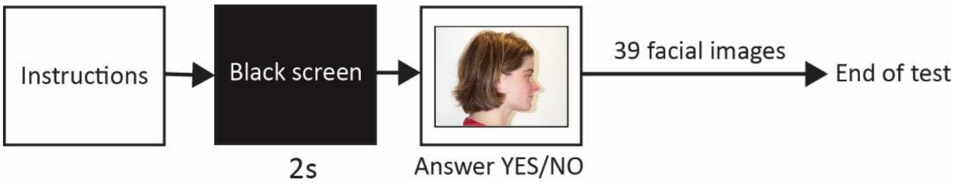

Figure 3: Procedure of the recognition test

Here we have added 20 new facial images to the original 20 facial images of the observation test. For each facial image displayed on the screen, participants were asked to indicate whether they had seen it in the observation test. We recorded their responses as correct or incorrect.

2.5 Analysis of results

The first part of the study provided us with the results of the average time of correct answers (CA) and incorrect answers (FA – False alarm) for each of the 4 tests. Those two data are then used to calculate recognition success (discrimination index A') with equation 1.

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

For the second study, we first set the eye-tracking system to create heat maps in grayscale instead of colour. So instead of red, yellow and green, we set black (R = G = B = 0), dark grey (R = G = B = 85) and light grey (R = G = B = 170). These greyscale images were then converted into black and white images using the ImageJ program. The process is shown in Figure 4.

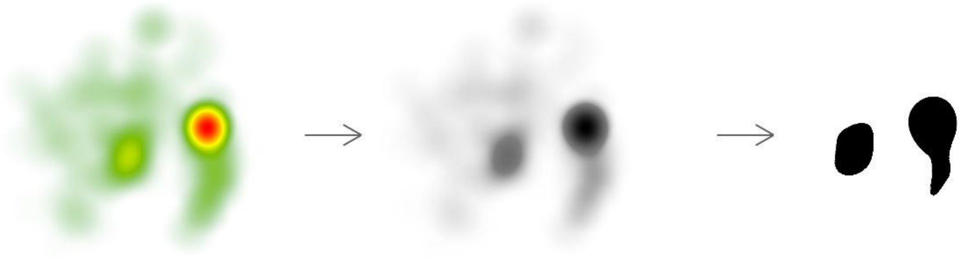

Figure 4: Converting color heatmaps to grayscale heatmaps to a black and white image

We then used the ImageJ program to calculate the area size, perimeter and circularity of the black and white heatmaps. The latter is calculated from the two previous parameters according to equation 2.

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

3. RESULTS

3.1 Recognition success

As already mentioned, we first determined the success of the recognition as a function of the observation time. The results are shown in Figure 5.

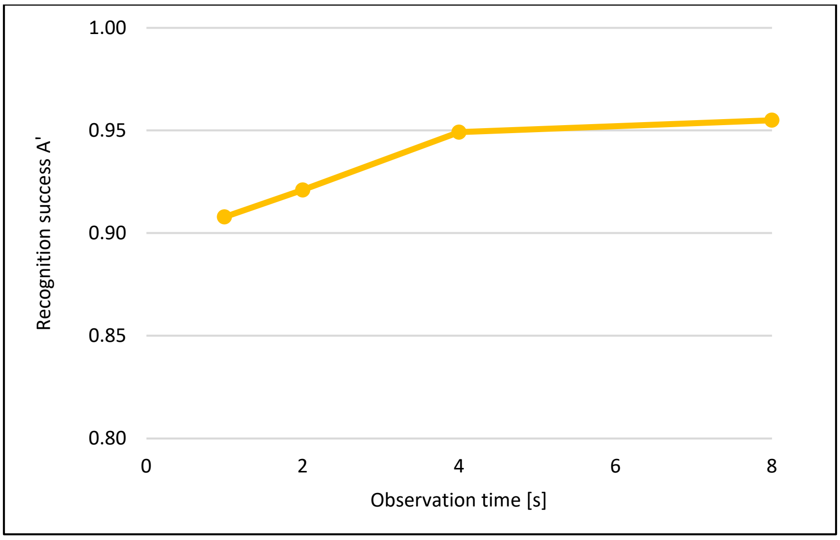

Figure 5: Results of recognition success for profile facial images

The second result of the first part of the experiment is incorrect recognition (false alarm). These results are shown in Figure 6.

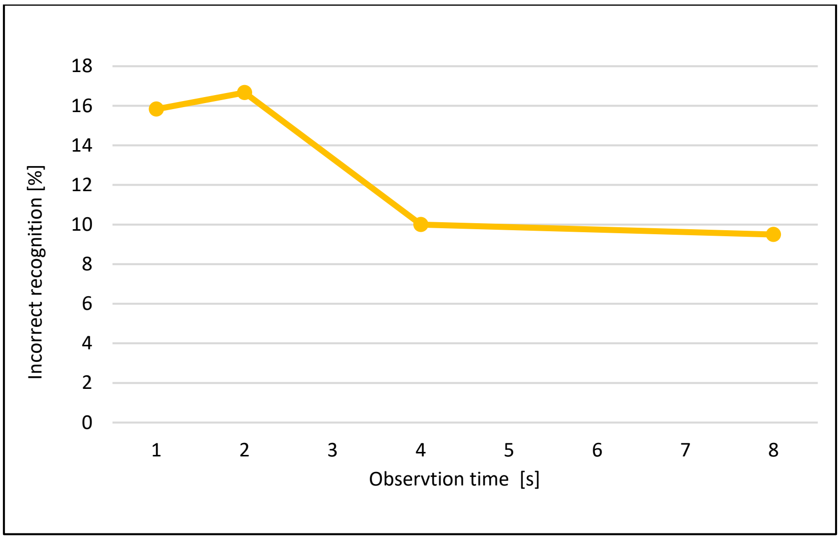

Figure 6: Results of incorrect recognition for profile facial images

3.1 Heatmaps results

As already mentioned, the second part of the experiment was the measurement of heatmap parameters. Figures 7, 8 and 9 show the results of the area size, extent and circularity of the converted heatmaps.

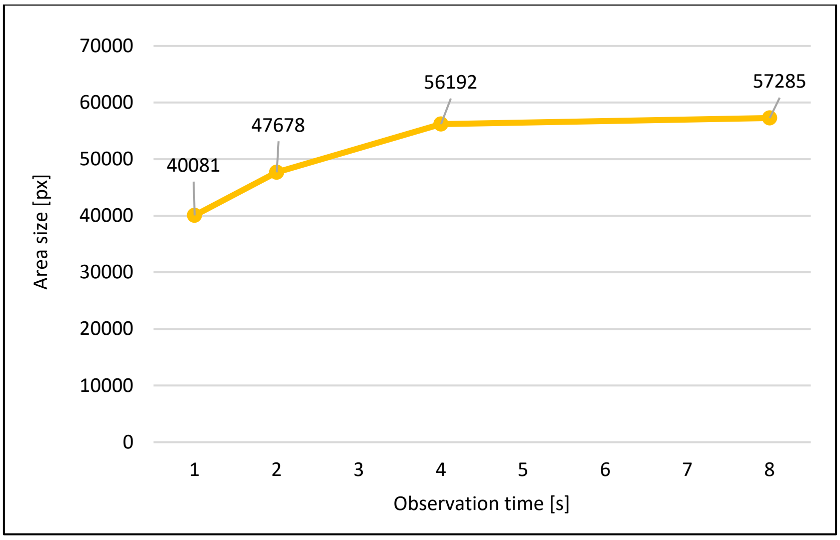

Figure 7: Results of area size of Heatmaps

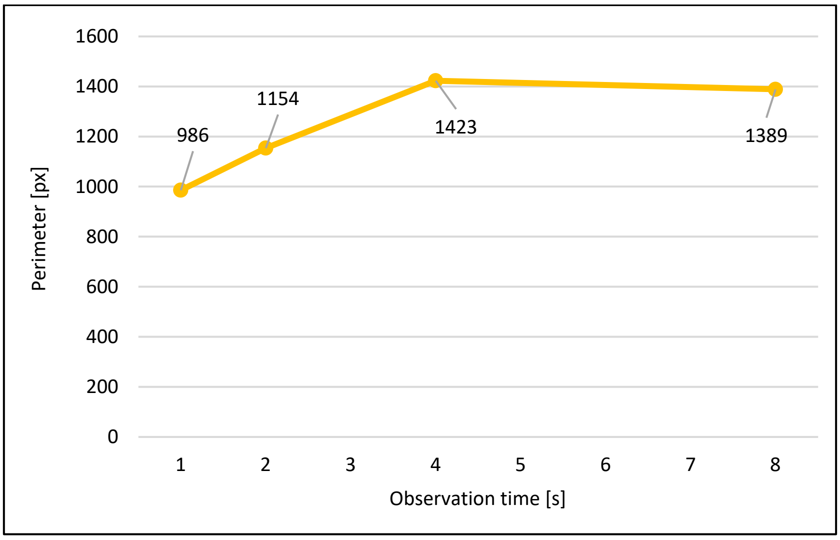

Figure 8: Results of perimeter of Heatmaps

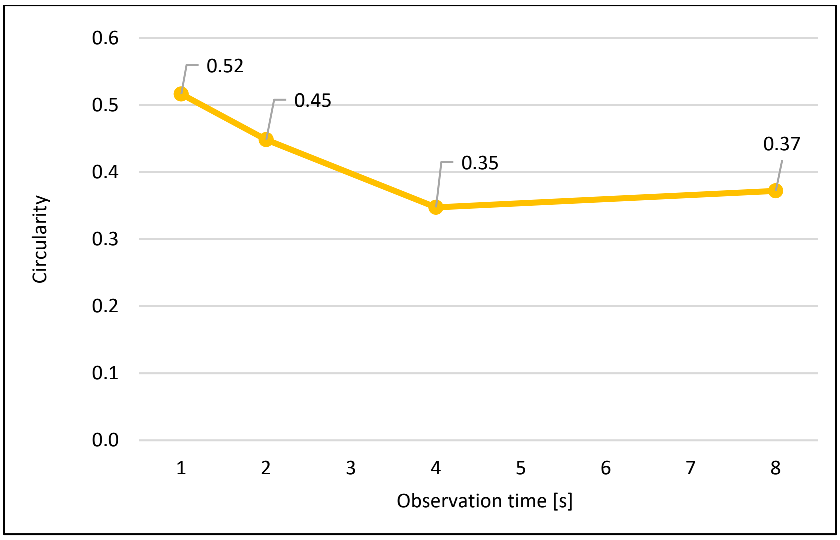

Figure 9: Results of circularity of Heatmaps

4. DISCUSSION

As we suspected, the recognition success increases with a longer observation time of the facial images in the observation test (Figure 5). This increase takes place up to an observation time of 4 seconds. In the 8-second observation test, the recognition performance is almost the same as in the 4-second test. This means that a longer observation time does not improve the memorization of facial images. For satisfactory memorization of facial images, an observation time of 4 seconds is sufficient.
The same applies to the result of incorrect recognition (Figure 6). With short observation times of facial images, the memorization ability is naturally poorer and thus the incorrect recognition is higher (15.8 % in the 1-second test and 16.7% in the 2-second test). At the same time, memorization of facial images is much better in the 4-second test and incorrect recognition falls to 10.0%, and it remains at approximately the same level in the 8-second test (9.5%).
When analysing the Heatmaps, it can be seen in Figure 7 that their area is smallest when the observation time of facial images is 1 second, which is of course to be expected. With a longer observation time of facial images, users saw a larger part of the face, which also means a larger area of the Heatmaps. This increase happens up to the 4-second test, and in the 8-second test we see that the area of the Heatmaps is almost the same as in the 4-second test. This indicates that between 4 and 8 seconds, participants mostly looked at previously observed parts of the face. Again, this does not contribute to better memorization, as we have already seen in Figures 5 and 6. The same thing happens when measuring the perimeter of the Heatmaps (Figure 8). As the observation time increases, the Heatmaps become larger and thus the perimeter also become larger. However, there is an interesting change here, as the perimeter of the Heatmaps is larger in the 4-second test than in the 8-second test. The reason for this is that there are certain cases where we obtained two separate Heatmaps in the 4-second test, for example, whereas in the 8-second test the participants observed a larger part of the facial image and these two Heatmaps merged into one. The perimeter of two Heatmaps is thus larger than the perimeter of one common heatmap. Therefore, the average perimeter of the Heatmaps for the 8-second test is even slightly smaller than for the 4-second test.
Figure 9, which shows the results of the circularity as a function of the observation time of the facial images, shows that the circularity is highest in the 1-second test and then decreases towards the 4-second test. This is because in the 4-second test the participants observe at a larger part of the facial images, the views are more scattered and the Heatmaps have more varied shapes than in the 1-second test, where the views are more concentrated on the centre of the facial image and the Heatmaps are more circular. There is practically no difference between the 4-second and 8-second test, which we have already seen from the results of the area and perimeter of the Heatmaps. During this extra time between 4 and 8 seconds, participants mostly observe parts of facial images that they have already observe before.

5. CONCLUSIONS

In our study, we attempted to compare the results of facial image recognition and the results of Heatmaps measurements as a function of the time of observing facial images in the observation test. We had 4 of these times (1, 2, 4 and 8 seconds). All results showed a similarity between the measured values. Thus, we found that a turning point occured everywhere at the observation time of 4 seconds. Up to this point, as the time observation increased, the recognition success (A') increased, the false recognition decreased, the area and perimeter of the Heatmaps increased and the circularity decreased. All these values remained almost the same for the 4-second and 8-second tests. With these results, we can explain the nature of observing facial images, which is such that the participants' gaze first land on the eyes, then in 4 seconds they observe the remaining parts of the face image, and between 4 and 8 seconds, participants observe parts of the faces in the facial images that they had already seen before.
With our study, we have demonstrated the usefulness of measuring Heatmaps. To further confirm the usefulness of this method, it should also be tested on images with other content (landscape, objects, interiors, etc.), which we plan to do in one of the future studies.

6. RERENCES

Abdallahi, O. M., Perriera Da Silva, M. & Courboulay, V. (2007) *A history of eye gaze tracking*. Hal archives–ouvertes. Report number: ffhal-00215967.

Harrington, D. O. (1981) *The Visual Fields: A Textbook and Atlas of Clinical Perimetry*. St. Louis, Mosby.

Henderson, J., Falk, R., Minut, S., Dyer, F. C., & Mahadevan, S. (2001) Gaze control for face learning and recognition by humans and machines. In: Shipley, T. F. & Kellman, P. J. (eds.) *From fragments to objects: Segmentation and grouping in vision. New York, Elsevier Science,* pp. 463–481.

Kealey, R., Sekuler, A. B. & Bennet, P. J. (2008) Effects of viewing condition and age on the functionality of eye movements for face recognition memory. *Journal of Vision*. 8 (6), 892. Available from: doi: 10.1167/8.6.892.

Minear, M. & Park, D. (2000) A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments & Computers*. 36 (4), 360−363. Available from: doi: 10.3758/bf03206543

Ovano, N. (2024) *Consumer Product Giants' Eye-Trackers Size Up Shoppers*. Available from: https://phys.org/news/2012-07-consumer-product-giants-eye-trackers-size.html [Accessed 22th September 2024].

Park Aging Mind Laboratory. (2024) *Centre for Vital Longevity*. Available from: http://agingmind.utdallas.edu/ [Accessed April 23rd September 2024].

Pernice, K. & Nielsen, J. (2009) *Eyetracking Methodology: How to Conduct and Evaluate Usability Studies Using Eyetracking*. Fremont, Nielsen Norman Group.

Robinson, D. A. A. (1963) Method of measuring eye movement using a scleral search coil in a magnetic field. *IEEE Transactions on Bio-medical Electronics*. 10 (4), 137−145. Available from: doi: 10.1109/tbmel.1963.4322822

Tobii Connect. (2016) *User's Manual Tobii Studio Version 3.4.5*. Danderyd, Tobii Technology.

Tobii Connect. (2024) *Working with Heat Maps and Gaze Plots*. Available from: https://connect.tobii.com/s/article/heat-maps-and-gaze-plots?language=ja [Accessed 2nd June 2024].