ANALYSIS OF OBSERVING AND RECOGNITION PROFILE FACIAL IMAGES USING EYE TRACKING SYSTEM

Andrej Iskra, Helena Gabrijelčič Tomc 匝

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

Abstract: Facial images have been the subject of research for many years, using the eye-tracking system. However, most researchers concentrate on the frontal view of facial images. Much less research has been done on faces shown at different angles or profile views of faces in facial images. However, as we know, in reality we often view faces from different angles and not just from a frontal view. In our research we used a profile presentation of facial images and analyzed memory and recognition depending on the display time and dimensions of the facial images. Two tests were performed, i.e. the observation and the recognition test, and we used the well-known yes/no detection theory. We used four different display times in the observation test (1, 2, 4 and 8 seconds) and two different dimensions of facial images 640 × 480 and 1280 imes 960). All facial images were taken from the standardized face database Minear&Park. We measured the recognition success which is mostly presented as a discrimination index A', incorrect recognition (FA - false alarm) and time-spatial method based on fixation duration and saccade length. In this case, eye tracking provides us with objective results when viewing facial images. In the results it was found that extending the display time of facial images improves recognition performance and that the dependence is logarithmic. At the same time, wrong recognition decreased. Both parameters are independent of the dimensions of the facial images. This fact has been proven by some other researchers also for frontal facial images. It was also discovered that with an increase of the display time of facial images an increase of the fixation duration and saccade lengths occurred. In all results we detected major changes at the display time of four seconds, which we consider as a time, where the subjects looked at the whole face and their gaze returned to the center of the face (in our case eye and mouth).

Key words: eye tracking, profile face images, recognition success, fixation duration, saccade length

1. INTRODUCTION

Images are one of the five basic elements of online information (Kyrnin, 2020). These include facial images, which are now an integral part of the way people present themselves on their websites, be it professional, personal or social networking sites. When publishing facial images, the dimensions and resolution of these facial images (Du and Martinez, 2011), the position of the face (frontal, mid-profile or in profile) (Liu and Chaudhiri, 2002) and the expressed emotion (Tarnowki et al, 2017; Green and Guo, 2018) are of great importance. From the memory aspect, the presentation facial image is main importance. This has a decisive influence on memory and later recognition. Much research has been done on frontal facial images, but very little research has been done on profile facial images. Above all, there is a lack of research on the influence of presentation time of facial images on the memory of faces. In our research we have demonstrated this influence of time in the results of recognition success A', incorrect recognition and two parameters of the eye tracking system, i.e. fixations duration and saccade lengths.

With the rapid technological development, very accurate, affordable and, above all, user-friendly eye tracking systems have recently emerged. We too have used such a system in our research, especially for the fact that it is an objective method for obtaining results and thus also allows for in-depth research in the behavioural and cognitive area.

2. METHODS

2.1 Participants

As mentioned above, we had eight different tests (two dimensions of facial images and four different times of facial image display in the observation test. Each test was performed by six participants, so that we had a total of 48 participants (15 male and 33 female) with an average age of 20.3 (SD = 0.96). These were students from our faculty who volunteered for the test and had normal vision.

2.2 Stimuli

The facial images were taken from the Minear & Park database (Minear and Park, 2000). In the observation test, we had 20 images of faces of the Caucasian race between 18 and 29 years of age with neutral expressions. 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. When setting up the environmental and testing system we followed the standards and recommendations (Pernice and Nielsen, 2009).

We performed the test with the Tobii X-120 eye tracking system. The distance between the test subjects and the screen with the facial images was 60 cm. The setting of the test environment and the test subject is shown in Figure 1.



Figure 1: Testing setup

Analysis were done in Tobii Studio 3.4.8 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, 2016).

2.4 Procedure

As already mentioned, we have conducted a memory test according to the well-known old/new 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 instructions were followed by a short pause of black screen, which lasted for two seconds, then the first facial image appeared. The display time of the facial image was determined according to the test (one, two, four or eight seconds). This was followed by a two-second pause with black screen and a new facial image. Like other researchers, we randomly determined the order of the images with respect to the representation of male and female faces. The black screen between the facial image representations was designed to neutralize participant's gaze before next facial image appear. The test contained 20 facial images and was performed automatically (the participants had no influence on the facial image display). Recognition test contained 40 facial images (240 YES / NO responses for each participants), which, as in the observation test, were equally distributed according to gender and existing and non-existing facial images in the observation test. The test was controlled by the participants themselves. The instructions were followed for two seconds on a black screen, then the facial images followed. The participants responded YES / NO by saying whether or not they saw the face in the observation test. After the answer, the participants clicked the mouse button to display a new face. The results were recorded manually. Procedure of recognition test in shown in Figure 3.



Figure 3: Procedure of the recognition test

2.5 Analysis of results

From the answers of the test persons we received two results. These are the percentage of correct recognition CR (the face was in the observation test and the participants confirmed this) and the percentage of incorrect recognition (FA – False alarm) (the face was not in the observation test and the participants answered that it was in the observation test). Those two data are then used to calculate recognition success (discrimination index A') with equation 1.

$$A' = 0.5 + \frac{(CR - FA)(1 + CR - FA)}{4CR(1 - FA)}$$
(1)

In addition, we presented the results as a time-spatial analysis based on the fixations duration and the saccade length. The fixations duration was obtained directly from Tobii Studio. The saccade length was calculated from the position of two consecutive fixations (F_1 and F_2) according to the following equation 2.

$$F_1F_2 = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$
(2)

3. RESULTS

3.1 Recognition success

As mentioned above, we conducted tests for two different dimensions of facial images and for four different times of facial image display. Figure 4 shows the recognition performance as a function of the facial image display time. The figure also shows the logarithmic trendline and R^2 (fitting coefficient).



Figure 4: Results of recognition success

For all eight tests we also obtained the results of incorrect recognition (FA), which are shown in Figure 5, for both dimensions of facial images depending on the time of display of facial images in the observation test.



Figure 5: Results of incorrect recognition (FA)

The main results of the eye tracking system were fixations duration and saccade length, which were obtained as described in the Method section. Both results are shown in Figure 7 and Figure 8.



Figure 6: Results of fixation duration



Figure 7: Results of fixation duration

4. DISCUSSION

The results of the recognition success show that it increases with the longer time of displaying facial images in the observation test. This is to be expected because the test subjects memorized their faces better over a longer period of display of the images and recognized them better in the recognition test. The increase in recognition success occurred up to the observation time of four seconds. In the 8-second observation test, the recognition success is almost the same (dimensions 640×480) or only slightly increased (dimensions 1280×960). However, the results are very similar for the two different dimensions of facial images, so we see that the dimensions of facial images have no influence on memory and recognition. Even before the tests, we assumed that the recognition curve would have a logarithmic shape as a function of time. This means that in the short time of displaying facial images the recognition success improves rapidly, and in the longer time it slows down or remains almost the same (we reach saturation). The very good matching of the result curves with the logarithmic curve is also shown by the determination coefficient R2 (for the dimensions 640×480 it is 0.936, for the dimensions 1280×960 0.973). These logarithmic curves can be used to predict the recognition success for other facial image display times (e.g. three seconds or six seconds).

In the case of incorrect recognition, we see that it decreases with increasing display time of the facial images. This can be explained by the fact that the test persons remembered facial images better with longer display time and therefore made fewer errors in recognition process. The results show us an interesting turning point. With short display times of facial images (one and two seconds) the error detection is high (approx. 15%), with 4s and 8s tests it is significantly lower (approx. 10%). Here, too, we see a big change in the results in the 4s test.

As for the fixation duration, we obtained the results of an increasing fixations duration with longer display time of the facial images. This is due to the fact that the eyes calmed down more with longer display times of the facial images and the facial images were observed in a more relaxed manner. This leads to a longer fixation duration. With fast change of the facial images (display time of one second) the eye was more disturbed and moved faster. As a result, gaze remained at a certain location for a shorter time and the fixations duration was shorter. Here, too, we see a tendency towards a smaller increase in the results of the fixation duration in the 4-second test. If we compare the dimensions of the facial images, we see that fixations were shorter at larger dimensions. We see the reason for this in the fact that the facial parts are larger in larger facial images. Depending on the setting of the fixation threshold, the eye tracking system detected more fixations on a larger area of the facial parts and these are consequently shorter in time. In smaller facial images, however, the facial parts were closer together and some small eye movements were not even perceived by the eye tracking system as individual fixations, but as a single fixation. As a result, we had fewer fixations that last longer.

In the results of the saccade lengths we can see that they lengthened with increasing display time of the facial images. During the short display time of the facial images, the test subjects had very little time, so their gaze was mainly directed to the central part of the face (eyes, nose and mouth). When they had more time available, they also looked at other parts of the face and thus saccades were longer. Again, we see a smaller increase in saccade lengths between the 4s and 8s tests than in the tests with shorter display time. When we compare the lengths of saccades in facial images with different dimensions, it is obvious that they were longer in images with larger dimensions. There is a factor of 2 between the two dimensions of the facial images is smaller (1.53 to 1.63). The reason for the smaller factor of the ratio of the average saccade lengths compared to the ratio of the dimensions of the facial images is that in larger images, the gaze was directed more towards the face itself and left the facial area less frequently than in smaller images. So in smaller facial images we got several individual long saccades, which increased the average saccade length.

5. CONCLUSIONS

In our research, which focused on profile facial images, we found that extending the display time of facial images in the observation test improves memory and thus the recognition of facial images in the recognition process. This is reflected in recognition success and incorrect recognition. Similar results have been demonstrated several times for frontal facial images, but there are far fewer studies for profile facial images. The main focus of these results was mainly on the acquisition of logarithmic curves of the dependence of recognition performance from the time of display of facial images in the observation process. These logarithmic curves can be used to predict recognition performance for other facial image display times (e.g. three seconds or six seconds). Both the results of recognition success and incorrect recognition show us a turning point in the test of four seconds display time of facial images. Longer display time of facial images does not improve memory and thus recognition. This was also shown with frontal facial images (Iskra and Gabrijelič, 2019).

The turning point in the observation of facial images at a display time of four seconds was also evident in the results of the fixation duration and the saccade length, where this change (the increase in both values) was not so obvious. Similar results were also obtained for frontal facial images (Iskra and Gabrijelič, 2019). In the future, we will use the area method and the internal facial features method to further confirm the results in detail.

6. RERENCES

- Du, S., Martinez, A. M.: "The resolution of facial expressions of emotion", Journal of Vision 11 (13), 1–13, 2011. doi: 10.1167/11.13.24.
- [2] Green, C., Guo, K.: "Factors contributing to individual differences in facija expression categorisation", Cognition and Emotion 32 (3), 37–48, 2018. doi: 10.1080/02699931.2016.1273200.
- [3] Iskra, A., Gabrijelič, H. T.: "Time and Spatial Eye-Tracking Analysis of Face Observing and Recognition", Technical Gazette 26 (4), 977–984, 2019. doi: 10.17559/TV-20180309142158.
- [4] Kyrnin, J.: "What is Web Content? ", URL: http://webdesign.about.com/od/content/qt/what-is-webcontent.htm (last request: 2020-04-02).
- [5] Liu, C.H., Chaudhiri, A.: "Reassesing the 3/4 view effect in face recognition", Cognition 83, 31–48, 2002. doi: 10.1016/S0010-0277(01)00164-0.
- [6] Minear, M., Park, D.: "A lifespan database of adult facial stimuli", Behavior Research Methods, Instruments & Computers 36 (4), 360–363, 2000. doi: 10.3758/bf03206543.
- [7] Pernice, K., Nielsen, J.: "Eyetracking methodology: how to conduct and evaluate usability studies using eyetracking", (Fremont, USA, Nielsen Norman Group 2009.), page 164.
- [8] Tarnowki, P., Kolidziej, M., Majkovski, A., Rak, R. J.: "Emotion recognition using facial expressions", Precedia Computer Science 108, 1175–1184, 2017. doi: 10.1016/S0010-0277(01)00164-0.
- [9] Tobii Studio, User's Manual version 3.4.5, Tobii Pro, URL: https://www.tobiipro.com/siteassets/tobiipro/user-manuals/tobii-pro-studio-user-manual.pdf (last request: 2020-06-12).



© 2020 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 3.0 Serbia (http://creativecommons.org/licenses/by/3.0/rs/).