ISSN 1330-3651 (Print), ISSN 1848-6339 (Online)
https://doi.org/10.17559/TV-20180309142158
Original scientific paper

Time and Spatial Eye-Tracking Analysis of Face Observing and Recognition

Andrej ISKRA, Helena GABRIJELČIČ TOMC

Abstract: Facial images are one of the most common images in the online and offline media, which raises the question as to how people look at and remember faces. The aim of our research was to determine relation between fixation duration and saccade length in the observation process, recognition performance and furthermore to find possible relation between recognition performance and portion of observation time for internal facial features. The observation test was designed to perform time (fixations) and spatial (saccades) eye-tracking analysis when observing the facial images for three different dimensions and four different presentation times (1 second, 2 seconds, 4 seconds and 8 seconds). Recognition test showed how different presentation time and different dimensions of face image influenced recognition performance. For the recognition performance presentation time is much more important than image dimensions. Furthermore, the analysis of observing the internal facial features (eyes, nose and mouth) was also carried out. All results revealed that 4 seconds of observation time is turning point in face observation. For the 4 seconds observation test fixation duration and saccade length reached constant value, wrong recognition dropped significantly and portion of observation internal facial features reached lowest point and then increased for longer observation times.

Keywords: eye-tracking; face image; face observation; face recognition; facial features; fixation; saccades

1 INTRODUCTION

The use of facial images is nowadays very widespread. It is included in many different fields of our lives, e.g. web presentations, criminalistics, security, psychology, neuroscience, advertising, marketing, etc. [1]. In the online presentations of companies, institutions, associations, etc., it is important that the employees be also represented by the image. Face images are highly informative, so they should be prepared with high accuracy of colour, contrast and details. How the observers look at and remember faces has been part of researches for many decades. Humans have the ability to memorize and identify thousands of different faces with very similar patterns [2]. Researches in the early 90's have shown that the face recognition consists of visual perception and cognition [3]. In the area of visual perception, movement of the eyes (eye-tracking) had been used already at the end of the 19th century. Analysis of facial images really boosted with the wide use of eye-tracking systems. First eye-trackings were built in the beginning of 20th century and first experiments were done for the investigation of reading process. Eye-tracking is a process of precise measurement of ocular activity and is considered as an objective method [4]. The movement of the eye consists of two parts: fixation and saccades. Fixation means that visual gaze stays for a short time on a particular location. Saccade is a quick, simultaneous movement of both eyes in the same direction between two fixations.

There are two main directions in face image research. The first research subfield has a focus on emotional facial expression [5–8], while the second exploratory area of face image research studies the observations at different angles of presented faces [9–12].

In the process of face observation and recognition the most important parameter is time of face presentation during observation phase. Longer observation time results in better recognition performance. That was proved for correct recognition [13, 14] and false alarm [15].

Many analyses have been done in the field of fixation duration, but mostly in observing scene images [16–20]. The review of the references revealed that the area of saccade length is less explored [21–23]. But no research was performed to explore possible relation between recognition success and fixation duration and saccade length.

Great number of researches was also carried out in the field of facial features. They mainly focused on internal facial features (eyes, mouth and nose) and their role in face recognition process [24–27].

There were three main goals of our research. We focused on variations of observation time and different dimensions of face images in the process of face image observation. First investigation was how different observation time and face image dimensions influence different eye tracking data: fixation duration and saccade length. Next question was how these changing parameters in observation test further impact the recognition performance. As mentioned before we also focused on three main facial features and tried to find out the pattern of how humans look at face images in the observation process. We examined portions of observation time for each of internal facial features at different observation times. Previous researches [27–30] have revealed that internal facial features are much more important in the face recognition process than the external one. This is related with observation time of face image. In a short observation time participants spend most time looking at internal facial features. By extending observation time these portions drop in favour of external facial features. Our assumption is that at one point when observation time is long enough, participants' gaze returns back to observe internal facial features. Consequently, we were interested in discovering at what observation time this turn happens and whether it has any relation with eye tracking data (fixation duration and saccade length) and recognition performance (correct and incorrect recognition). At the end, all the data were used to examine possible relation between eye tracking data (fixation duration and saccade length), recognition performance (correct and wrong recognition) and observation portions of internal facial features in the processes of face observation and face recognition.

2 METHODS

Cases mentioned in Introduction (angle of face orientation, emotion expression) contain many parameters that influence the way participants observe and recognize faces. In our investigation, we focused only on frontal facial images with a neutral facial expression. We used the faces of Caucasian race and a similar age to our participants. We consider them to be typical faces.

2.1 Participants

The test participants were students, all of whom were aged between 19 and 23 with normal vision. The participants were divided into 12 groups according to 4 different observation times and 3 different facial image dimensions (12 tests in total). There were 93 participants. Since Tobii eye tracking detection rate of 90% was required [31], we performed our tests in such a way as to ensure that there was an equal number of participants in each test group that passed that criterion. Six participants were allocated to each group, which meant that the results for 72 participants were analysed. All of the participants were volunteers and received a bonus during the study course.

2.2 Stimuli

The facial images were obtained from the Minear and Park facial database [32] which was created by photographing volunteers in controlled conditions. The two main categories for organizing this database were age and emotion expression. The facial images were divided into four age groups:

- 18–29 years old,
- 30–49 years old,
- 50–69 years old,
- 70–94 years old.

Figure 1 Male facial image from Minear and Park database

Database contains a total of eight basic facial expressions: neutral, happy, sad, angry, annoyed, surprised, disgusted and grumpy. There is approximately equal number of male and female facial images and all four main races were included (Caucasian, Black, Asian and Indian). We took the natural expression facial images of 20 male and 20 female Caucasian people aged 18 to 29 years old. This was done in order to have similar aged stimuli and test participants. Fig. 1 shows samples of the male facial images from the Minear and Park database.

2.3 Apparatus

All of the tests were performed using the Tobii X-120 eye tracker with the sample rate of 120 Hz for eye detection. The monitor was 24" with a resolution of 1440 × 900. Although a higher resolution could be set on the monitor, we were precluded from taking advantage of this owing to the requirements of the Tobii studio 3.4.4 software which was used to collect and analyse the testing data. The fixation definition in the Tobii software was set to 100 ms and a 30 px area, which was also used by other researchers [33, 34]. The distance between the participants and the monitor was approximately 60 ± 3 cm [7, 35]. For our monitor resolution settings, faces of the medium dimension images were approximately 13 cm in height, which corresponds to the conditions of natural observation (20 cm face height at distance of 1 m) [36]. We also had smaller and larger facial images.

2.4 Procedure

We designed 12 tests according to the different observation times and dimensions of the facial images presented. All of the facial images in different tests were the same (20 male and 20 female facial images). Each participant performed one of the 12 tests. As mentioned above we investigated relation between observation time and dimension of facial image and recognition success. Observation times were set at 1 second, 2 seconds, 4 seconds and 8 seconds. The second parameter was the dimensions of the facial images. We prepared three groups of different dimension images: "small" (320 px × 240 px), "medium" (640 px × 480 px) and "large" (1280 px × 960 px). All of these combinations yielded 12 different tests. The test for the participants was divided into two parts. The first part was observation test and the second part was recognition test. Calibration of the eye tracking system was carried out for each participant at five control points and was performed at the beginning of the test.

The observation part consisted of 20 images (10 male and 10 female images). This testing procedure was automatic and is shown in Fig. 2.

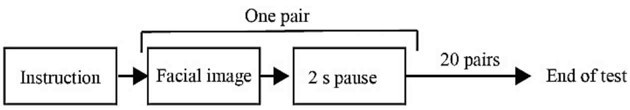

Figure 2 Procedure of the observation test

After the initial instructions users clicked mouse button for the first facial image. After the observation time of facial image (depending on the test), there was a 2 seconds pause with dark screen. The purpose was to neutralize the position of the eyes [7, 37]. In that case last fixation of the previous facial image had no influence on the first fixation on the next facial image. By inserting a dark screen pause, the eyes position at the appearance of new facial image was placed in a neutral position (somewhere near the centre of the screen).

The second part of our testing was recognition test, which was controlled by the participants themselves. The test comprised 40 facial images, 20 from observation test and 20 new faces from the same group as before (Caucasian race, neutral expression, aged 18–29). These 20 new faces were also equally divided into 10 male and 10 female facial images. After the instructions were provided, the first facial image was displayed. The participants were required to answer YES or NO if they saw that facial image in the observation test. The answers were recorded manually. After providing each answer, the participants clicked the mouse and the next image appeared. Since there was no time limit, the participants had as much time as they needed to think about each image; however, the answers were usually provided very soon after the facial image appeared. The recognition test procedure is presented in Fig. 3.

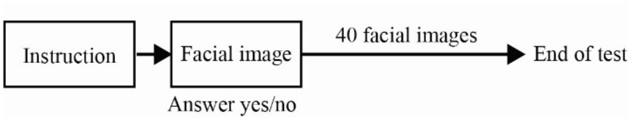

Figure 3 Recognition test procedure

3 EXPERIMENT

As described in the introduction, our research was divided into three parts.

3.1 Fixation Duration and Saccade Length

In the observation test, we focused on fixation duration and saccade length and their dependence on different observation times and different dimensions of facial images. The fixation duration data were obtained directly from Tobii Studio by setting Area of Interest (AOI) for the whole facial image. All of the participants' fixations and duration times were measured.

The saccade length was calculated from the fixation position. We exported the horizontal (x) and vertical (y) positions for all the fixations of each facial image. From this data, we used Eq. (1) to calculate the distance between two consecutive fixations:

$$F_1F_2 = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \qquad (1)$$

where F_1 and F_2 represent two consecutive fixations and x_1, x_2, y_1, and y_2 represent the horizontal and vertical position on the screen. These data from Tobii Studio were in pixels and we calculated saccade length from pixels into degrees of visual angle (°). The analysis included the facial images of the same dimensions as we were interested in ascertaining the relationship between observation time and saccade length for facial images of the same dimensions.

3.2 Recognition Performance

In the recognition we measured the effectiveness of recognition in relation to the observation time and to the dimensions of the facial image. We obtained recognition success for 12 different parameters in the observation tests (4 observation times and 3 facial image dimension parameters) and we defined two terms:

- CR (correct recognition). The facial image was in the observation test and the participants confirmed by answering YES.
- IR-FA (incorrect recognition – false alarm). The facial image was not included in the observation test, but the participant stated that he/she saw it.

3.3 Internal Facial Features

The focus of the last investigation centred on internal facial features. This investigation was done only for the largest facial images (1280 × 960 px). The reason for this was the unreliable levels of precision of the eye tracking system. The facial features on small images are closer together than for large images and consequently it can occur that while the participant looked at the eyes the eye tracking recorded a fixation on the nose. These errors happen very rarely with large images where facial features are far apart. We analysed the observation made on the internal facial features (eyes, nose and mouth) and drew an AOI for each manually. An example of an AOI is shown in Fig. 4.

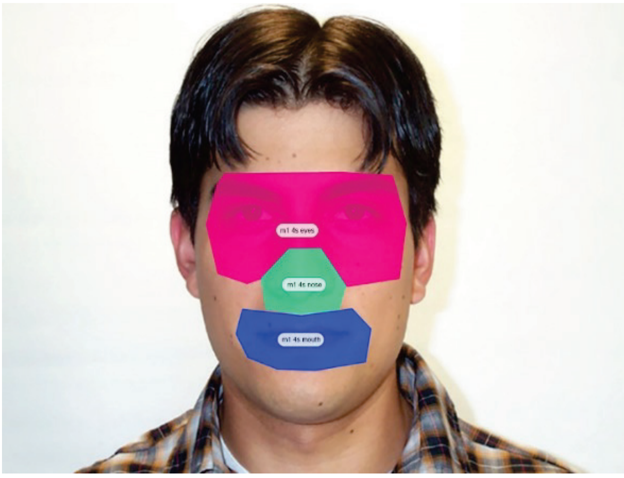
Figure 4 Setting of AOI on facial image.

Defining AOI meant that we got number of fixation, average fixation duration and total observation time for all AOI's of all facial images in the observation tests. Average observation time for all participant and each facial feature for different time tests (1 s, 2 s, 4 s and 8 s) was calculated. The final results were presented as a percentage of the total observation time for each facial feature.

Another way to investigate how participants observe facial images is to calculate return level for the eyes area [14]. Eye → mouth → nose is the most common sequence for observing facial images. After observing internal facial features, people look at external facial features (chin, cheeks, forehead, ears and hair) and when the observation of all face is completed, peoples' gaze mostly goes back to the eyes. We calculated return level for all participants for different observation time and the result was the percentage of participants that return to the eyes area during observation test. Our presumption was that the percentage of observation time for internal facial features is in relation with return rate.

4 RESULTS
4.1 Fixation Duration and Saccade Length

Average value of fixation duration for every individual facial image was measured from the results on the whole facial image and obtained directly from Tobii Studio. For example, six participants performed 1 second test on small facial images. Each participant observed 20 facial images, which meant that the results for the average fixation duration were based on 120 observed images. The results of the fixation duration are shown in Fig. 5.

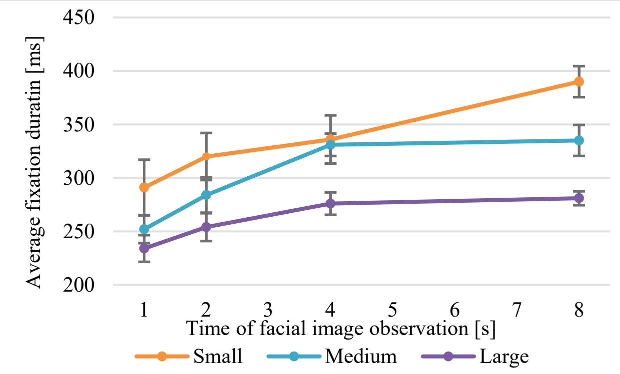

Figure 5 Average fixation duration and standard deviation bars for different observation times and different dimensions of facial images

Average fixation duration was the longest for small dimension images (M = 334 ms, SD = 36.1). Middle dimension facial images have average fixation duration M = 304 ms, (SD = 34.4) and the shortest average fixation time was for large dimension facial images (M = 260 ms, SD = 18.2). Our results showed that observation time has the lowest influence at large facial images.

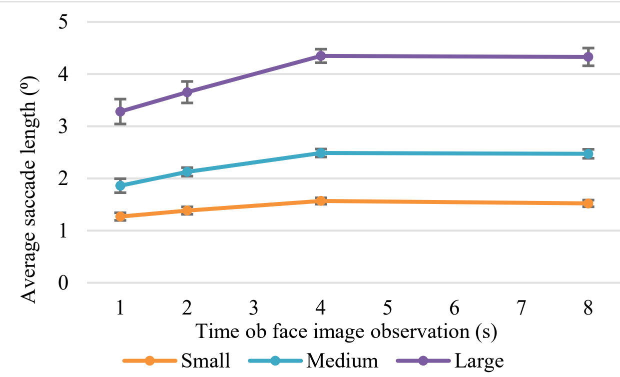

Figure 6 Average saccade length and standard deviation bars for different observation times and different dimensions of facial images

For saccade length, we exported the horizontal and vertical position data for all the fixations during the observation tests. Using Eq. (1) (described above), we calculated the distances between two consecutive fixations and calculated them into degrees of visual angle (°). Again, the results were taken for groups of 120 observed facial images for each of observation tests. Fig. 6 shows the results for saccade length.

Average saccade length was shortest for small dimension facial images (M = 1.44°, SD = 0.12), for middle dimension facial images was M = 2.24° (SD = 0.24) and the longest was for large dimension facial images (M = 3.9º, SD = 0.45).

4.2 Recognition Performance

Recognition test gave us results for correct and incorrect recognition for all 12 observation tests and can be seen in Fig. 7.

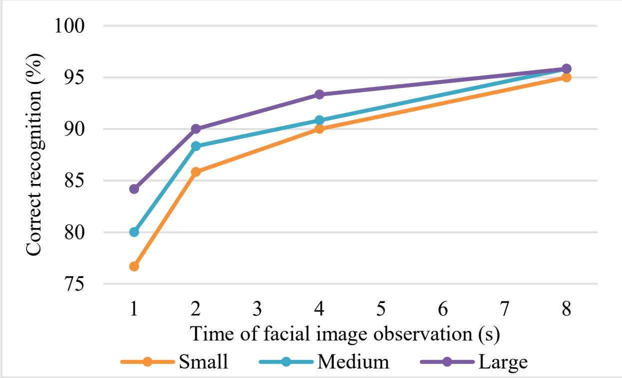

Figure 7 Recognition success (%) for different observation times and different dimensions of facial images.

Fig. 7 shows that the recognition had the highest percentage (independent from facial image dimension) at observation time 8 seconds (M = 95.6%, SD = 0.39). For 4 seconds observation test correct recognition was 91.4% (SD = 1.42), for 2 seconds observation test correct recognition was 88.1% (SD = 1.71) and, as expected, the results show that the recognition success was the worst at 1 second observation test (M = 80.3%, SD = 3.07). Results of SD showed that, at longer observation time, the dimensions of facial images had less influence on correct recognition.

We also analysed incorrect recognition, which is commonly also referred to as False Alarm. In this case, the participant responded that the face had featured in the observation test, despite this not being the case (the participant made a mistake) (Fig. 8).

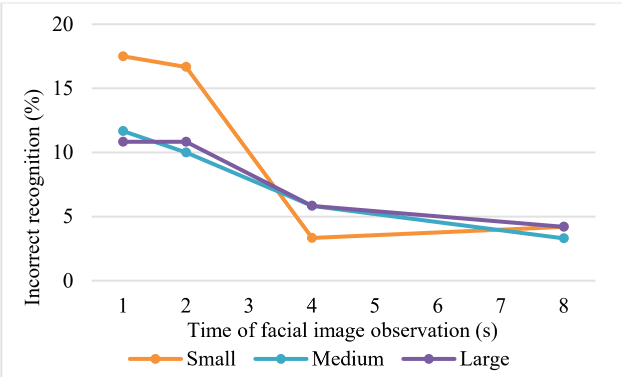

Figure 8 Incorrect recognition (%) for different observation times and different dimensions of facial images

The analysis for average incorrect recognition for different observation test gave the following results: 1 second observation test (M = 13.3%, SD = 2.97), 2 seconds observation test (M = 12.5%, SD = 2.97), 4 seconds observation test (M = 5.0%, SD = 1.18) and 8 seconds observation test (M = 3.9%, SD = 0.42). Again, SD for different observation times showed for longer observation times dimensions of facial image had less impact on incorrect recognition.

Fig. 9 and Fig. 10 show the relation between Incorrect recognition and fixation duration and saccade length, respectively.

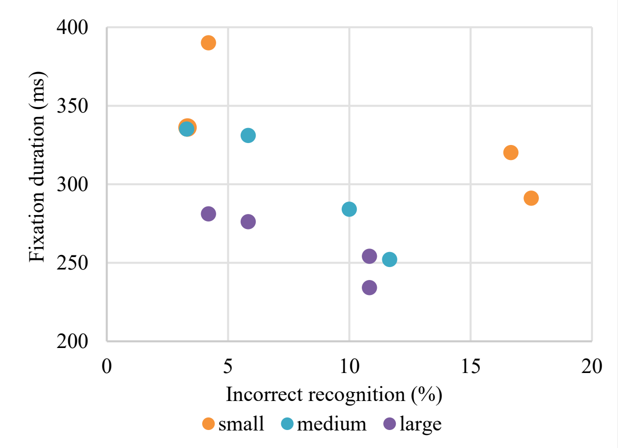

Figure 9 Relation between Incorrect recognition (%) and fixation duration (ms)

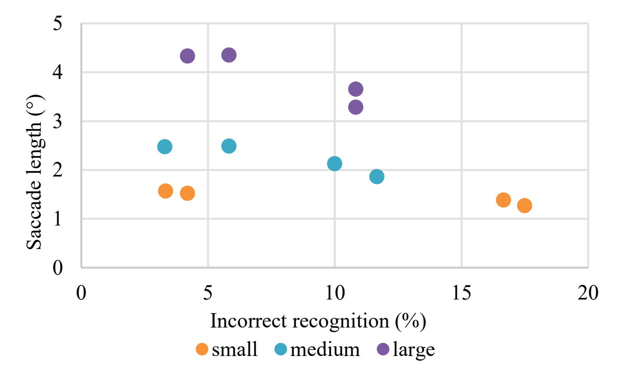

Figure 10 Relation between Incorrect recognition (%) and saccade length (°)

Good recognition means low level of incorrect recognition. If we observe left part of graph in Fig. 9 we can see two points relatively close and also in horizontal alignment for medium and large facial images for observation times 4 and 8 seconds. It means that fixation duration became almost constant at 4 seconds observation test and incorrect recognition dropped significantly compared to short observation tests.

Fig. 10 shows similar pattern as Fig. 9. We had almost horizontal alignment (constant saccade length) at low incorrect recognition (left side of graph). This turning point happened at 4 seconds observation test.

4.3 Internal Facial Features

Results of the third experiment were to investigate portion of time spent on internal facial features (eyes, mouth, nose) in dependence on observed time. As explained before, we took only large facial images and set the AOI for every facial feature on all the facial images. Results of average time that the participants spent on AOI were acquired from the Tobii Studio followed by the calculations of the relevant portions for all three facial features. The results are presented in Fig. 11.

Average portion for observing eyes area for all different observation times was 39.18% (*SD* = 2.30), for mouth area 15.92% (*SD* =4.10) and for nose area 12.50% (*SD* = 3.31). In all tests together participants observed internal facial features for more than two thirds of total observation time.

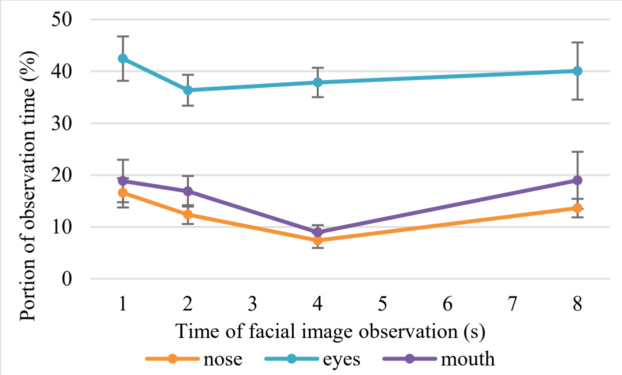

Figure 11 Portion of observation time and standard deviation bars for the main three facial features

In Fig. 12 return rate for the eyes area can be observed and it shows the percentage of returns to the eyes area for each observation test (20 images each user, 6 users per test).

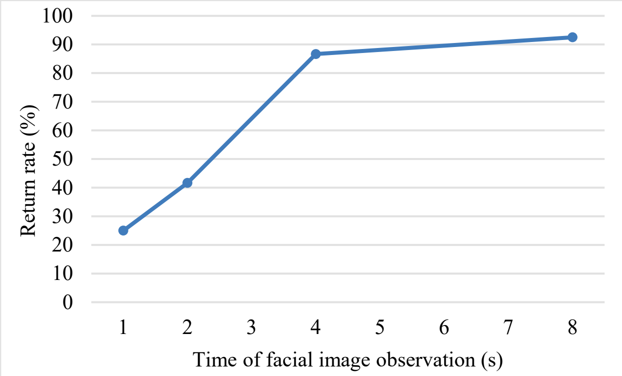

Figure 12 Return rate (%) for different observation times

1 second of observation time is so short that most users only manage to visit eyes area and continue with the quick observations of nose or mouth as other two internal facial features. Return rate can be, in some cases, problematic value. Due to the error detection of precise position of participants' gaze, some fixation can be detected outside eyes AOI, in spite of their observing the eyes. If next fixation is again inside eyes AOI it is treated as a positive value for return rate and presumably a certain portion of the results in Fig. 12 is also consequence of the mentioned mistake. Nevertheless, Fig. 12 clearly shows a huge increase of return level at 4 seconds observation time (86.7%). Longer observation time (8 seconds) did not increase that level by much (92.5%).

Fig. 13 shows gaze plot for different observation times. First face image (1 second observation time) shows participants' gaze mostly stayed at internal facial features, at 2 seconds observation time participants started to observe also other facial features. In 4 seconds observation time test participants' gaze already returned to the eyes and from that point on (8 seconds observation time) they mostly observed internal facial features. Images in Fig. 11 are from different observation tests (1 s, 2 s, 4 s, and 8 s), i. e. from different users and not from users of 8 seconds observation test for different time spans.

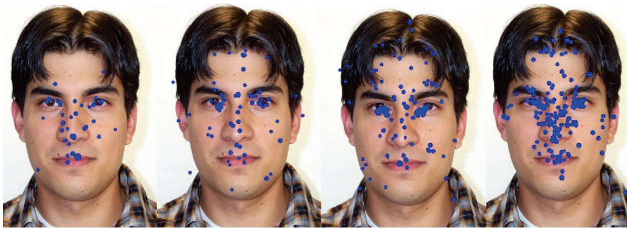

Figure 13 Gaze plots for different observation times (1 second, 2 seconds, 4 seconds and 8 seconds)

5 DISCUSSIONS

5.1 Fixation Duration and Saccade Length

The results of the first experiment showed that a longer observation time resulted in longer fixation duration (Fig. 5). This was because the participants' eyes adapted to the longer presentation of the facial image and they could observe the facial elements in greater details. We can also see the smallest influence of observation time for the large dimension facial images (SD data). There is an interesting similarity between the medium and large dimension images. There was practically no difference in fixation duration between the 4 seconds and 8 seconds tests (medium dimension facial images 0.331 ms and 0.335 ms, large dimension facial images 0.276 ms and 0.281 ms).

Shorter fixation duration was in the 1 s and 2 s tests due to the eyes being more agitated. In other words, to quickly changed images eyes were adopted with faster change of gaze. Not such pattern was found with small facial images. Small images were inconvenient for the used eye tracking experiments, because facial feature in smaller images are really close to each other and the eye tracking system cannot track eye movement so accurately. We can say that observation time of 4 seconds is kind of turning point where eye became calm and observed facial images in more relaxed way. If we compare results for medium and large dimension facial images, we can see that difference in fixation duration was present in all tests and it was slightly growing from 1 second to 4 seconds observation time. Large facial images meant larger facial features that are also farther apart in the image space. Therefore, in this case, participants made more fixations (depending on fixation definition in Tobii Studio) and more fixations in the same time meant shorter fixation duration. In small facial images, the facial features are so close together that the eye movements are short and some are not even detected as a new fixation on the eye tracker. Consequently, less number of fixations meant they were of longer duration.

When investigating saccade length, we detected the same pattern for all image dimensions (Fig. 6). The saccade length increased almost linearly from 1 second to 4 seconds test. 1 second observation time was so short that many times participants' gaze just stayed in the centre of the face (on main three facial features). Longer observation time meant participant had more time to observe more facial features, so there were more fixations and were farther apart. However, the observation times of 4 seconds and 8 seconds resulted in almost the same saccade distance. Again, this confirmed the assumption made during the fixation duration investigation that an observation period of 4 seconds or more caused more relaxed eye movement with constant saccade length.

5.2 Recognition Performance

Fig. 7 shows the predicted relation between observation time and correct recognition. Longer presentation of facial images meant better recognition performance. This was confirmed by other researchers [13−15]. The shapes of the correct recognition curves (Fig. 7) for all three different dimensions of facial images are similar. The curve is steep for a short observation time, which means that recognition success improves significantly as the observation time increases until the 4 seconds presentation time, which is the point when the curve starts to flatten out. Recognition success at 8 seconds observation time is just a little better than at 4 seconds observation test. We assume that if we had increased the observation time even more (i.e. to 10 or 15 seconds), the recognition results would not be any better than for 8 seconds presentation time. This could be an interesting goal for the future experiments, as we assume that correct recognition would get even worse.

As expected, incorrect recognition (Fig. 8) gave opposite curve shape from the correct recognition. Longer observation time for facial images meant less number of incorrect recognitions. Here, the turning point is again at 4 seconds observation time, because at this point the level of incorrect recognition dropped significantly from the level at 2 seconds test. On the other hand, 8 seconds test did not improve incorrect recognition significantly. Also 1 second test had almost the same incorrect recognition level as 2 seconds test. This showed that 4 seconds face presentation is sufficient enough that incorrect recognition level is satisfactorily small. Our results of recognition success confirmed previous researches [14, 32].

Dimensions of presented facial images (small, medium, large) had less influence than observation time. Fig. 7 shows that dimensions of facial images had more importance for correct recognition in short observation time (1 seconds) (76.7%, 80.0% and 84.2%). With increasing observation time dimensions of facial images had less influence. At the presentation time of 8 seconds, we can see that correct recognition was almost the same for all three dimensions (95.0%, 95.8% and 95.8%). Incorrect recognition (Fig. 8) was almost the same for all dimensions for long observation times (4 seconds and 8 seconds). For short observation times (1 second and 2 seconds), small difference in incorrect recognition level is shown only for small and medium dimension images. Significant drop of incorrect recognition can be seen at 4 seconds observation time (5.0%), and SD showed relatively high independence from dimensions of facial image. Low level of incorrect recognition stayed also for 8 seconds observation time. Observation times of 1 and 2 seconds had relatively high level of incorrect recognition (more than 10%), especially for small dimension facial images (17.5% and 16.7% respectively). Again, 4 second observation time appeared to be turning point for incorrect recognition. 4 seconds turning point can be seen in Fig. 9 and Fig. 10 with presentation of relation between incorrect recognition and fixation duration and saccade length, respectively.

5.2 Internal Facial Features

At investigation of the portion of observation time for internal facial features (eyes, mouth and nose) we attempted to discover the pattern of how participants look at facial images when they are presented at different observation times. For observation time of 1 second the percentage of observation of three main face elements (Fig. 11) was quite high. All three took over more than three quarters of the total observation time (77.9%). In these short observation times, the participants' gaze mostly remained directed towards the centre of the face. At 2 seconds observation time participants had enough time to start observing also other facial features (chin, cheeks, forehead and ears), so the portion of observation for main three facial features dropped (eyes from 42.5% to 36.4%, mouth from 18.9% to 16.9%, nose from 16.6% to 12.4%). These trends continued also for 4 seconds observation time (eyes 37.9%, mouth 9%, nose 7.4%) where observation percentage for mouth and nose dropped significantly. On the other hand, percentage of eyes observing already slightly increased. In absolute time this was more than double time for observing eyes between 2 seconds and 4 seconds tests (0.73 s and 1.51 s). Consequently, this meant also that 4 seconds observation time was a turning point where participants mostly looked at all face and returned back to the eyes area. For observation time 8 seconds this return effect was even more obvious, as the percentage of observing main three facial features again increased significantly (from 54.2% to 72.7%). In terms of absolute observation time, this meant an increase from 2.17 s to 5.81 s and specially mouth (from 0.36 s to 1.52 s) and nose (from 0.3 s to 1.09 s) areas were observed much longer than in the 4 seconds test.

Return rate (Fig. 12) for 1 and 2 seconds observation times was relatively small (25.0% and 41.7%), and was, partly, also the consequence of error detection of precise position. 4 seconds observation time was the turning point, as return rate increases significantly (86.7%). That indicated that most participants already observed whole face image and returned their gaze back to the eyes area.

All of these results and the scan path of the participants' gazes revealed that time period of approximately 4 seconds was sufficient to observe the whole face, before reverting back to the eyes area and then to the mouth and nose areas.

6 CONCLUSION

In the three presented experiments it was investigated how different observation times and the various dimensions of facial images refer to the results of eye tracking parameters (fixation duration and saccade length), recognition success and distribution of observation of internal facial features. An important milestone occurred at an observation time of 4 seconds. At this point, the fixation duration (for medium and large dimension facial images) and saccade length (for all dimensions) became almost constant. Also at this point, the incorrect recognition level drops to < 5% and stayed relatively constant even for longer observation time. Although correct recognition increased rapidly for short observation times, a 95% correct recognition rate (the saturation level) was recorded for longer observation times (8 seconds). Increasing the observation time would not improve recognition success and it is more likely that the results of this parameter would drop due to the total length of the test (the participants would experience problems maintaining concentration) [38].

Results of portion of observation of internal facial features and return rate showed that most participants finish observing whole face (all facial features) in less than 4 second and their gaze were already back to the eyes area.

All results have showed similarity between several parameters in our experiments: fixation duration, saccade length, incorrect recognition, observation portions of internal facial features and return rate. When fixation duration and saccade length became nearly constant the percentage of incorrect recognition dropped significantly (around 5% or less), observation portion for eyes area started increasing and return rate increased to the very high level. All these turning points happened at 4 second observation time test.

We can conclude that 4 seconds of observation time is long enough for face memorization. At that time eye movement became more stable and constant. This resulted in relatively good remembrance of facial images which resulted in low incorrect recognition. Longer observation times meant looking back to already seen facial features and not improving recognition success.

7 REFERENCES

[1] Senior, A. W. & Bolle, R. (2002). Face recognition and its applications. D. Zhang, ed. *Biometric Solutions for Authentication in an E-World.* Boston, MA: Kluwer Academic Publishers, 101-115. https://doi.org/10.1007/978-1-4615-1053-6_4

[2] Bahrick, H. P., Bachick, P. O., & Wittleinger, R. P. (1975). Fifty years of memory for names and faces: A cross-sectional approach. *Journal of Experimental Psychology*, *104*(1), 54-75. https://doi.org/10.1037/0096-3445.104.1.54

[3] Kowler, E. (1990). The role of visual and cognitive processes in the control of eye movements. Kowler, E. ed. *Eye Movements and Their role in Visual and Cognitive Processes*. Michigan: Elsevier, 1-70.

[4] Collewijn, H. (1999). Eye movement recording. Carpenter, R. H. S. & Robson, J. G. ed. *Vision research: A practical guide to laboratory methods.* Oxford: Oxford University Press, 245-285. https://doi.org/10.1093/acprof:oso/9780198523192.003.0009

[5] Malcom, G. L., Lanloy, L. J., Fugard, A. J. B., & Barton, J. J. S. (2008). Scan patterns during the processing of facial expression versus identity: An exploration of task-driven and stimulus-driven effects. *Journal of Vision*, *8*(2), 1-9. https://doi.org/10.1167/8.8.2

[6] Rachael, J. E., Blais, C., Scheepers, C., Schyns, P. G., & Caldara, R. (2009). Cultural confusions show that facial expressions are not universal. *Current Biology 19*(18), 1543-1548. https://doi.org/10.1016/j.cub.2009.07.051

[7] Cangöz, B., Altun, A., Aşkar, P., Baran, Z., & Mazman, S. G. (2013) Examining the visual screening patterns of emotional facial expressions with gender, age and lateralization. *Journal of Eye Movement Research*, *6*(4), 1-15.

[8] Schurgin, M. W., Nelson, J., Iida, S., Ohira, H., Chiao, J. Y., & Franconeri, S. L. (2014). Eye movements during emotion recognition in faces. *Journal of Vision*, *14*(13), 1-16. https://doi.org/10.1167/14.13.14

[9] Saether, L., Van Belle, W., Laeng, B., Brennen, T., & Øvervoll, M. (2009). Anchoring gaze when categorizing faces' sex: Evidence from eye-tracking data. *Vision Research*, *49*(23), 2870-2880. https://doi.org/10.1016/j.visres.2009.09.001

[10] 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. https://doi.org/10.1167/9.2.7
[11] Chelnokova, O. & Laeng, B. (2011) Three-dimensional information in face recognition: an eye-tracking study. *Journal of Vision*, 11(13), 1-15. https://doi.org/10.1167/11.13.27
[12] Brielmann, A. A., Bülthoff, I., & Armann, R. (2014). Looking at features in Asian and Caucasian faces. *Vision Research*, *100*, 105-112. https://doi.org/10.1016/j.visres.2014.04.011
[13] Leyk, D., Sievert, A., Heiss, A., Gorges, W., Ridder, D., Alexander, T., Wunderlich, M., & Rüther, T. (2008) Validation of a short-term memory test for the recognition of people and faces. *Ergonomics*, *51*(8), 1125-1136. https://doi.org/10.1080/00140130802094371
[14] Iskra, A. & Gabrijelčič Tomc, H. (2016). Eye-tracking analysis of face observing and face recognition. *Journal of Graphic Engineering and Design*, *7*(1), 5-11. https://doi.org/10.24867/JGED-2016-1-005
[15] Reynolds J. K. & Pezdek K. (1992). Face Recognition Memory: The Effects of Exposure Duration and Encoding Instruction. *Applied Cognitive Psychology*, *6*(4), 279-292. https://doi.org/10.1002/acp.2350060402
[16] Unema, P. J. A., Pannasch, S., Joos, M., & Velichkovsky, B. M. (2005). Time course of information processing during scene perception: The relationship between saccade amplitude and fixation duration. *Visual Cognition, 12*(3), 473-494. https://doi.org/10.1080/13506280444000409
[17] Henderson, J. M. & Pierce, G. L. (2008). Eye movements during scene viewing: Evidence for mixed control of fixation durations. *Psychonomic Bulletin & Review*, *15*, 566-573. https://doi.org/10.3758/PBR.15.3.566
[18] Henderson, J. M. & Smith T. J. (2009). How Are Eye Fixation Durations Controlled during Scene Viewing? Evidence from a Scene Onset Delay Paradigm. *Visual Cognition*, *17*(6/7), 1055-1082. https://doi.org/10.1080/13506280802685552
[19] Calen-Walshe, R. & Nuthmann, A. (2014). Asymmetrical control of fixation durations in scene viewing. *Vision research, 100*, 38-46. https://doi.org/10.1016/j.visres.2014.03.012
[20] Pannasch, S., Helmert, J. R., Roth, K., Herbold, A.-K., & Walter, H. (2008). Visual Fixation Durations and Saccadic Amplitudes: Shifting Relationship in a Variety of Conditions. *Journal of Eye Movement Research*, *2*(2), 1-19.
[21] Andrews, T. J. & Coppola, D. M. (1999). Idiosyncratic characteristics of saccadic eye movements when viewing different visual environments. *Vision Research*, *39*, 2947-2953. https://doi.org/10.1016/S0042-6989(99)00019-X
[22] Henderson, J. M. & Hollingworth, A. (2003). Eye movements and visual memory: Detecting changes to saccade targets in scenes. *Perception & Psychophysics*, *65*(1), 58-71. https://doi.org/10.3758/BF03194783
[23] Glen, F. C., Smith, N. D., & Crabb, D. P. (2013). Saccadic eye movements and face recognition performance in patients with central glaucomatous visual field defects. *Vision Research*, *82*, 42-51. https://doi.org/10.1016/j.visres.2013.02.010
[24] Hsiao, J. H. & Cottrell, G. W. (2008). Two fixations suffice in face recognition. *Psychological Science*, *19*(10), 998-1006. https://doi.org/10.1111/j.1467-9280.2008.02191.x
[25] Buchan, J. N., Paré M., & Munhall, K. G. (2007). Spatial statistics of gaze fixations during dynamic face processing. *Social Neuroscience*, *2*(2), 1-13. https://doi.org/10.1080/17470910601043644
[26] Heisz, J. J. & Shore, D. I. (2008). More efficient scanning for familiar faces. *Journal of Vision*, *8*(1), 1-10. https://doi.org/10.1167/8.1.9
[27] Hills, C., Romano, K., Davies-Thompson, J. &. Barton, J. J. S. (2014). An adaptation study of internal and external features in facial representations. *Vision research*, *100*, 18-28. https://doi.org/10.1016/j.visres.2014.04.002
[28] Logan, A. J., Gordon, G. E., & Loffler, G. (2017). Contributions of individual face features to face discrimination. *Vision research*, *137*, 29-39. https://doi.org/10.1016/j.visres.2017.05.011
[29] Longmore, C. A., Liu, C. H., & Young, A. W. (2015).The importance of internal facial features in learning new faces. *The Quarterly Journal of Experimental Psychology*, *68*(2), 249-260. https://doi.org/10.1080/17470218.2014.939666
[30] Kramer, R. S. S., Young, A. W., & Burton. A. M. (2018) Understanding face familiarity. *Cognition*, *172*, 46-58. https://doi.org/10.1016/j.cognition.2017.12.005
[31] (2012) User Manual Tobii Studio version 3.2. Tobii Technology, Danderyd, Sweden.
[32] Minear, M. & Park, D. C. (2004). A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments, & Computers*, *36*(4), 630-633. https://doi.org/10.3758/BF03206543
[33] Goldinger, S. D., He, Y., & 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), 1105-1122. https://doi.org/10.1037/a0016548
[34] Mäntylä, T. & Holm, L. (2006). Gaze control and recollective experience in face recognition. *Visual Cognition*, 13(3), 365-386. https://doi.org/10.1080/13506280500347992
[35] Franken, G., Podlesek, A., & Možina, K. (2015). Eye-tracking Study of Reading Speed from LCD Displays: Influence of Type Style and Type Size. *Journal of Eye Movement Research*, *8*(1), 1-8
[36] Henderson, J. M., Williams, C. C., & Falk, R. J. (2005). Eye movements are functional during face learning. *Memory & Cognition*, *33*(1), 98-106. https://doi.org/10.3758/BF03195300
[37] Peterson, M. F. & Eckstein, M. P. (2014). Learning optimal eye movements to unusual faces. *Vision Research*, 99, 57-68. https://doi.org/10.1016/j.visres.2013.11.005
[38] Read, J. D., Vokey, J. R., & Hammersley, R. (1990). Changing photos of faces: Effects of exposure duration and photo similarity on recognition and the accuracy-confidence relationship. *Journal of Experimental Psychology Learning Memory and Cognition, 16*(5), 870-882. https://doi.org/10.1037/0278-7393.16.5.870

Contact information:

Andrej ISKRA, MSc. Assistant
University of Ljubljana,
Faculty of Natural Sciences and Engineering,
Aškerčeva 12, 1000 Ljubljana, Slovenia
andrej.iskra@ntf.uni-lj.si

Helena GABRIJELČIČ TOMC PhD, Associate Professor
University of Ljubljana,
Faculty of Natural Sciences and Engineering,
Aškerčeva 12, 1000 Ljubljana, Slovenia
helena.gabrijelicic@ntf.uni-lj.si