Judging qualification, gender, and age of the observer based on gaze patterns when looking at faces[⋆]

Pawel Kasprowski[1][0000−0002−2090−335X], Katarzyna Harezlak[1][0000−0003−3573−9772], Piotr Fudalej[2][0000−0001−8777−7434], and Pawel Fudalej[3]

[1] Silesian University of Technology, Akademicka 16, Gliwice, Poland
pawel.kasprowski@polsl.pl
[2] Department of Orthodontics and Dentofacial Orthopedics, School of Dental Medicine, University of Bern, Bern, Switzerland
[3] Private Practice, Warsaw, Poland

Abstract. The research aimed to compare eye movement patterns of people looking at faces with different but subtle teeth imperfections. Both non-specialists and dental experts took part in the experiment. The research outcome includes the analysis of eye movement patterns depending on the specialization, gender, age, face gender, and level of teeth deformation. The study was performed using a novel, not widely explored features of eye movements, derived from recurrence plots and Gaze Self Similarity Plots. It occurred that most features are significantly different for laypeople and specialists. Significant differences were also found for gender and age among the observers. There were no differences found when comparing the gender of the face being observed and levels of imperfection. Interestingly, it was possible to define which features are sensitive to gender and which to qualification.

Keywords: Eye tracking · Gaze patterns · Faces observation · Age estimation.

1 INTRODUCTION

By registering eye movements of people observing images, it is possible to obtain much interesting information. On the one hand, it is possible to learn something about the observers - their knowledge, intentions, anxiety, etc. On the other side, by registering eye movements of several observers looking at the same image, it is also possible to learn something about the picture - if the content is interesting, shocking, which parts of the image are the most interesting, etc.
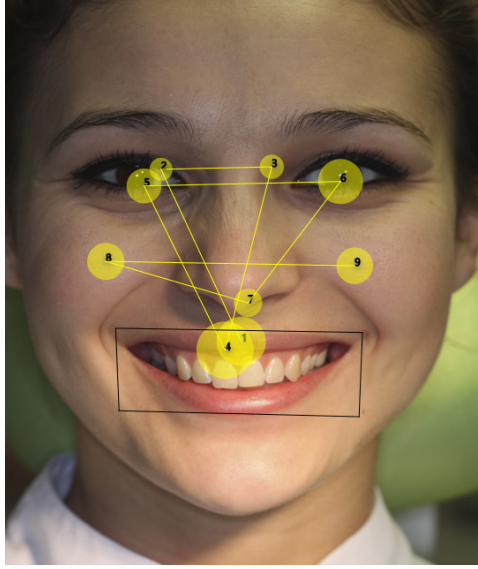
Fig. 1. A scan-path of a typical face observation

Registering eye movements while observing other people's faces is one of the most popular experiments because face observing patterns are very specific for humans. The ability to follow and analyze somebody's face is one of the first skills achieved by a newborn baby. During our whole life, we train observation of faces to recognize people's intentions, moods, etc. Lack of this ability may be considered a severe disadvantage. The analysis of eye movement patterns during face observation has been used to diagnose autism [4] or Alzheimer's disease [12]. The pattern is also supposed to be idiosyncratic [3, 8], different for races [2] and age [11], and changes if the observer knows the face being observed [7].

It is well known that people typically start by looking at the other people's eyes and then look at the mouth and nose area. It is called a 'face triangle'. However, subtle differences in this pattern may reveal information about the observer and the face being observed. It is the central assumption of the research described in this paper. We presented faces of people with and without various teeth imperfections to two groups of participants: dentists and non-specialists without any connection with dentistry. Contrary to the previous research that showed that people tend to look longer at teeth when imperfections are visible [9], this time, the differences were subtle and barely distinguishable even for specialists.

Our first assumption was that eye movement patterns of dentists and laypeople will be distinguishable. We also expected that eye movement patterns would depend on the level of teeth imperfections - especially for the dentists' group. At the same time, we wanted to explore if the eye movement patterns depend on the gender and age of both people being observed and observers.

Many papers show differences in eye movement patterns between novices and specialists, e.g., radiographers [10], or chess experts [13]. Some works were also devoted to the dentists' gaze patterns - e.g., when they looked at CT images [14]. However, to the best of our knowledge, there was no published research concerning differences between non-specialists and dentists while looking at face images. In the case of gender, some evidence of different women's and men's gaze patterns has been shown in [5] - which revealed that women are better at recognizing familiar faces.

However, most papers cited above compared gaze patterns using only traditional criteria such as the number and duration of fixations and area of interest (AOI) analysis. Our paper analyses also additional, relatively new, and unexplored measures based on recurrence plots [1] and Gaze Self-Similarity Plots [6]. Therefore, this paper's contribution is two-fold: (1) it explores the previously unchecked problem of comparing dentists with laypeople when looking at faces, and (2) it introduces novel and yet unexplored techniques to perform this comparison.

2 METHODS AND MATERIAL

The experimental design consisted of a presentation of 48 human faces with three levels of dental imperfections (16 faces for each level). Every face was presented for 6 seconds. The participants' task was to just freely observe the pictures and try to judge the attractiveness of the person being observed.

The same presentation was shown to 53 participants, including 25 non-specialists and 28 dental specialists. Participants' eye movements were recorded during each trial utilizing The Eye Tribe eye tracker with a 60Hz sampling rate. It gave about 360 subsequent gaze points for each 6-second recording - called an *observation* in the subsequent text.

The data registered during the experiment were analyzed to remove observations with improper gaze points (e.g., when a person closed eyes for some time or turned away her eyes). Observations for which there were less than 250 gaze points were removed from the dataset. Finally, 1883 observations were available for further analyses, including 929 observations of specialists and 954 observations of non-specialists. The dataset consisted of 654 observations from women and 1229 from males, and 1188 observations of female faces and 695 observations of male faces.

Every image of the face used during the experiment was processed to define two areas of interest (AOI): mouth and incisor teeth. The time to gaze at these AOIs and the overall time of staring at them were calculated for each observation. Additionally, typical eye movement signal properties such as fixation number and average fixation duration were measured.

The next step was the calculation of four measures used to quantify eye movement recordings based on the recurrence plot defined in [1]. Given a sequence of N fixations:

– **_Recurrence_** represents the percentage of recurrent fixations (i.e., how often the observer refixates previously fixated image positions).
– **_Determinism_** represents repeating fixation patterns.
– **_Laminarity_** measures how often the same areas of a scene are repeatedly fixated.
– **_Center of recurrence mass_** (CORM) measures how close the recurrent fixations points are situated.

Due to space limits, we do not present details about the way these four measures are calculated, which may be found in [1].

The next group of measures included the features derived from Gaze Self Similarity Plots (GSSP) [6]. At first, the co-occurrence matrix was calculated for every GSSP, and then the contrast, homogeneity, and uniformity were calculated.

– **_Homogeneity_** gives information to what extend nearby gazes are placed in similar locations.
– **_Contrast_** reveals long jumps (saccades) from one gaze point to another.
– **_Uniformity_** measures gaze pairs repetitions. It is high when the GSSP contains similar areas, which means that the same paired values with the same arrangement repeatedly appear in the image.

All three features were calculated for $[1, 1]$ offset. Please refer to [6] for calculation details.

The last group of features included the coverage - the whole image was divided into $N \times N$ areas and the **_coverageN_** was calculated using the Equation 1.

$$COV_N = \frac{\sum_i^N \sum_j^N (\exists Fix \in area(i,j))?1:0}{N \times N} \qquad (1)$$

All examined features are summarized in Table 1.

3   RESULTS

The research aimed to check if some particular observer's features influence the way the image is being observed. The second aim was to check if the content of the image influences gaze patterns. Therefore, we defined three independent features describing the observer: qualification, gender, and age and two independent features describing the image: the level of imperfection and *imgGender* - gender of a person on the image (see Table 1). It was first checked if there are statistically significant differences in these groups for dependent variables (all gaze-related features). According to the Kolmogorov-Smirnov test, all variables were normally distributed. Therefore, for variables with only two possible values (*qualification*, *gender* and *imgGender*) the t-Student test was used, and for variables with more than two possible values (age and level) the ANOVA test was used with Bonferroni correction. For the sake of readability, values of the homogeneity and uniformity were multiplied by $10e3$ and $10e5$, respectively.

Table 1. The features taken into account.

Feature	Description
	independent features
gender	gender of the observer
age	age of the observer
qual	qualification of the observer (layman, specialist)
level	level of imperfection (1-3)
imgGender	gender of a person on the image
	gaze related features
fixNum	number of fixations
fixAvgDur	average fixation duration
sacAvgLen	average saccade length
sacLen	summary of saccades length
mouthTimeto	time before focusing the mouth AOI
mouthTime	the overall time of gazing at mouth AOI
incisTimeTo	time before focusing the incisors AOI
incisTime	the overall time of gazing at incisors AOI
rec	reccurence
det	determination
lam	laminarity
corm	center of recurrence mass
coverage5	coverage of the image in 5x5 grid
coverage7	coverage of the image in 7x7 grid
coverage10	coverage of the image in 10x10 grid
contrast	GSSP contrast
homog	GSSP homogeneity
uniformity	GSSP uniformity

3.1 Observer related variables

The results analysis starts with discussing features related to the observer, namely: qualification, gender, and age.

Qualification We verified the hypothesis that eye movement patterns are different for specialists and laypeople. Average values of features and standard deviations were calculated, and then the t-Student test was used to decide if the differences are significant. As it is visible in Table 2, this hypothesis has been confirmed as nearly all analyzed features (apart from *fixAvgDur*) occurred to be significantly different between groups.

It is not surprising that for the chosen AOIs (mouth and teeth), the observation is longer, and *timeto* values are shorter for specialists. However, it is interesting to notice that other features exhibit significant differences as well. The only exception is *uniformity*, which is similar for both groups.

Table 2. Features values - differences for laypeople and specialists.

feature	laymen	specialists	t(1883)	p-value	
fixNum	14,6	15,4	-3,289	0,001	**
fixAvgDur	345,9	359,4	-0,704	0,482	
sacAvgLen	6,9	4,7	20,024	0	***
sacLen	100,2	70,6	14,976	0	***
mouthTimeto	2253,2	1054	14,12	0	***
mouthTime	986	1977,4	-14,818	0	***
incisTimeto	3991,3	2574,8	13,573	0	***
incistime	428,5	962,7	-10,353	0	***
rec	13,1	21,3	-12,178	0	***
det	30,8	46,1	-9,167	0	***
lam	51,6	80,4	-14,702	0	***
corm	4,8	8,1	-8,822	0	***
coverage5	0,21	0,17	13,204	0	***
coverage7	0,14	0,11	11,043	0	***
coverage10	0,08	0,07	10,17	0	***
contrast	102,4	186,6	7,23	0	***
homog	332,3	290,6	13,25	0	***
uniformity	202,4	220,3	-2,18	0,03	*

Gender The next research question was if the eye movement patterns depend on the gender of the observer. As was already stated in the Introduction, it has been confirmed previously for other datasets. As there were significant differences between specialists and laypeople, we decided to perform the t-Student test independently for both groups. The results are presented in Table 3 and Table 4. There are significant differences in most features depend on the observer's gender in both groups.

Interestingly this time, contrary to the previous case, the values of *recurrence* and *contrast* are not significantly different, while the *uniformity* may be considered as a feature that can reliably distinguish both genders.

Age of the observer Observers were divided into four groups differing with age: (20-30,30-40,40-50,50+). It occurred that for most features the result of ANOVA test revealed significant differences: $F(3,1883) > 4.8, p < 0.01$. Only for *corm* feature the effect was less significant with $F(3,1883) = 3.268, p = 0.021$.

3.2   Image properties

The second group of properties included properties related to the observed image - the different level of imperfection (*level*) and the gender of a person on the image (*imgGender*). However, this time neither the ANOVA test conducted for *level* nor t-Student test performed for *imgGender* showed significant differences in the results (Table 5).

Table 3. Gender differences for non-specialists.

feature	female	male	t(954)	p-value	
fixNum	12,9	15,6	-10,36	0	***
fixAvgLen	422,7	297,5	9,85	0	***
sacAvgLen	7,4	6,6	4,32	0	***
sacLen	95,5	103,2	-2,62	0,01	**
mouthTimeTo	2051,8	2380,2	-2,35	0,02	*
mouthTime	1409,1	719,2	8,31	0	***
incisTimeTo	4076	3937,9	0,93	0,35	
incisTime	587,1	328,4	4,35	0	***
rec	13,4	13	0,57	0,57	
det	27,9	32,7	-2,17	0,03	*
lam	39	59,5	-7,73	0	***
corm	5,3	4,5	2,73	0,01	**
coverage5	0,2	0,22	-4,02	0	***
coverage7	0,13	0,14	-4,52	0	***
coverage10	0,08	0,08	-4,1	0	***
contrast	90,1	110,2	-1,76	0,08	
homog	372,6	306,8	19,62	0	***
uniformity	236,12	181,1	5,89	0	***

3.3 Classification

Disclosing the importance of differences in eye movement features for specialists and non-specialists, males and females, and the observer's age was the motivation factor to verify the possibility of identifying the observer's qualifications based on his/her visual pattern during a single observation of a face.

For this purpose, we applied a simple k-nearest neighbor algorithm with $k = 7$. The algorithm predicts the given observation's class based on the classification of k nearest training observations. The model was tested using classic 10-fold cross-validation repeated ten times. The results presented in Table 6 confirm the previous findings that the observer's qualification and gender may be estimated using the features defined in Table 1.

The last test was an attempt to predict the observer's age. The task was more difficult than the previous ones because it was 1 out of 4 classification. The confusion matrix for the dataset is presented in Table 7. It may be noticed that the best results were obtained for the participants belonging to the group 40-50. However, the overall accuracy was about 58%.

4 SUMMARY

The research presented in this paper aimed to analyze if and how eye movement features depend on the properties of a person being examined and the properties of the image being observed. The detailed analysis confirmed that significant

Table 4. Gender differences for specialists.

feature	female	male	t(954)	p-value	
fixNum	13	16,4	-7,69	0	***
fixavgDur	662,8	225,1	8,67	0	***
sacAvgLen	5,2	4,5	4,39	0	***
sacLen	72,1	70	0,65	0,52	
mouthTimeTo	893,3	1125,2	-2,1	0,04	*
mouthTime	2997	1526,1	11,58	0	***
incisTimeTo	2359,8	2669,9	-1,87	0,06	
incisTime	1452,6	745,8	6,56	0	***
rec	21,9	20,9	0,76	0,45	
det	34,7	51,1	-6,47	0	***
lam	66,6	86,5	-6,66	0	***
corm	10,3	7,2	3,27	0	**
coverage5	0,16	0,17	-1,77	0,08	
coverage7	0,11	0,11	-2,49	0,01	*
coverage10	0,06	0,07	-3,45	0	***
contrast	167,6	194,9	-1,25	0,21	
homog	334,2	271,3	13,36	0	***
uniformity	295,8	186,8	7,07	0	***

differences in eye movement patterns might be found for qualification and gender. Probably the most interesting finding was that *contrast* and *homogeneity* proved to be useful in differentiating the level of expertise, while *uniformity* depended on the gender of a person.

It was impossible to find any significant dependencies regarding the type of an image - probably because the images represented the same kind of stimulus (face) and were characterized by similar features.

The main limitation of the presented results is the vague meaning of the novel features being analyzed. The semantics of the presented REC and GSSP features is for now not clear. More studies of these features are necessary to define what they represent precisely. The research presented in this paper shows the meaningfulness of these features; however, their values are still not as easy to interpret as average fixation duration or AOI gazing time.

Future work will include tests using other datasets and more in-depth analysis of the novel features. It is planned to check if these features could be used to distinguish a type of observed images or to identify people.

Acknowledgement

Table 5. Features values for images containing women and men (*imgGender*) - the results for specialists.

feature	female	male	t(929)	p-value	
fixNum	15,4	15,3	0,39	0,7	
fixavgDur	357,9	361,9	-0,11	0,91	
sacAvgLen	4,8	4,5	1,87	0,06	
sacLen	72,8	66,8	2,21	0,03	*
mouthTimeTo	1098,3	977,4	1,14	0,26	
mounthTme	1910,2	2093,8	-1,58	0,11	
incisTimeTo	2655,8	2434,4	1,41	0,16	
incisTime	918,6	1039	-1,28	0,2	
rec	20,9	21,8	-0,74	0,46	
det	46,3	45,8	0,18	0,86	
lam	81,4	78,8	0,86	0,39	
corm	8	8,4	-0,63	0,53	
coverage5	0,17	0,16	1,81	0,07	
coverage7	0,11	0,11	-0,19	0,85	
coverage10	0,07	0,07	0,84	0,4	
contrast	177,8	201,8	-1,08	0,28	
homog	289,1	293,2	-0,84	0,4	
uniformity	219,2	221,4	-0,17	0,86	

References

1. Anderson, N.C., Bischof, W.F., Laidlaw, K.E., Risko, E.F., Kingstone, A.: Recurrence quantification analysis of eye movements. Behavior research methods **45**(3), 842–856 (2013)
2. Blais, C., Jack, R.E., Scheepers, C., Fiset, D., Caldara, R.: Culture shapes how we look at faces. PloS one **3**(8), e3022 (2008)
3. Cantoni, V., Galdi, C., Nappi, M., Porta, M., Riccio, D.: Gant: Gaze analysis technique for human identification. Pattern Recognition **48**(4), 1027–1038 (2015)
4. Dalton, K.M., Nacewicz, B.M., Johnstone, T., Schaefer, H.S., Gernsbacher, M.A., Goldsmith, H., Alexander, A.L., Davidson, R.J.: Gaze fixation and the neural circuitry of face processing in autism. Nature neuroscience **8**(4), 519–526 (2005)
5. Hall, J.K., Hutton, S.B., Morgan, M.J.: Sex differences in scanning faces: Does attention to the eyes explain female superiority in facial expression recognition? Cognition & Emotion **24**(4), 629–637 (2010)
6. Kasprowski, P., Harezlak, K.: Gaze self-similarity plot-a new. Journal of Eye Movement Research **10**(3), 1–14 (2017)
7. Kasprowski, P.: Mining of eye movement data to discover people intentions. In: International Conference: Beyond Databases, Architectures and Structures. pp. 355–363. Springer (2014)
8. Kasprowski, P., Harezlak, K.: The second Eye Movements Verification and Identification Competition. In: 2014 IEEE International Joint Conference on Biometrics (IJCB). pp. 1–6. IEEE (2014)
9. Kasprowski, P., Harezlak, K., Fudalej, P., Fudalej, P.: Examining the impact of dental imperfections on scan-path patterns. In: International Conference on Intelligent Decision Technologies. pp. 278–286. Springer (2017)

Table 6. Results of classification for all two-class features.

feature	accuracy	precision	recall	F1-score
qual	79,4%	80,8%	78,1%	79,4%
qual/fem	86,4%	85,9%	83,4%	84,7%
qual/male	78,3%	79,8%	78,9%	79,4%
gender	77,6%	86,8%	80,4%	83,5%
gender/lay	80,6%	84,6%	83,9%	84,2%
gender/spec	78,4%	90,6%	80,7%	85,4%
imGen	57,0%	21,2%	35,9%	26,7%
imGen/spec	59,2%	27,1%	41,2%	32,7%
imGen/lay	57,5%	22,2%	37,9%	28,0%

Table 7. Confusion matrix for age classification for non-specialists.

	20-30	30-40	40-50	50+
20-30	108	50	21	9
30-40	88	211	34	27
40-50	19	18	220	4
50+	2	8	2	35

10. Kasprowski, P., Harezlak, K., Kasprowska, S.: Development of diagnostic performance & visual processing in different types of radiological expertise. In: Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications. p. 40. ACM (2018)
11. Liu, S., Quinn, P.C., Wheeler, A., Xiao, N., Ge, L., Lee, K.: Similarity and difference in the processing of same-and other-race faces as revealed by eye tracking in 4-to 9-month-olds. Journal of Experimental Child Psychology **108**(1), 180–189 (2011)
12. Mendez, M.F., Mendez, M., Martin, R., Smyth, K.A., Whitehouse, P.J.: Complex visual disturbances in Alzheimer's disease. Neurology **40**(3 Part 1), 439–439 (1990)
13. Reingold, E.M., Charness, N., Pomplun, M., Stampe, D.M.: Visual span in expert chess players: Evidence from eye movements. Psychological Science **12**(1), 48–55 (2001)
14. Suwa, K., Furukawa, A., Matsumoto, T., Yosue, T.: Analyzing the eye movement of dentists during their reading of ct images. Odontology **89**(1), 0054–0061 (2001)