Eye-tracking metrics in perception and visual attention research

Magdalena Borys, Małgorzata Plechawska-Wójcik

Institute of Computer Science, Lublin University of Technology, Poland

European Journal of Medical Technologies
2017; 3(16): 11-23

www. medical-technologies.eu
Published online 31.10.2017

Corresponding address:
Institute of Computer Science, Lublin University of Technology, Poland
ul. Nadbystrzycka 38A,
20-618 Lublin
tel. +48 81 538 43 49
e-mail: m.borys@pollub.pl

Abstract

Eye-tracking is a research method in which point of gaze is measured using a device called an eye-tracker. In medicine eye-tracking have been used for many years to evaluate the visual behaviour of experts, but also for medical diagnosis, tracking progress or regression of disease over time. In this publication, authors discussed the most often used eye-tracking metrics as well as presented usage of those metrics in two experiments. The first experiment involved the analysis of perception in the diagnostic process using medical images. The second experiment included the analysis of visual attention of participants during psychomotor tests. Applied eye-tracking metrics allowed to verify the visual patterns of the participants with the high and low cognitive performance. As a summary, the authors presented a generalized process for visual attention analysis with eye-tracking data using area of interest.

Key words:

eye movement measurements, visual perception, eye-tracking

Introduction

Perception is an active cognitive process that involves the interpretation of sensory data (e.g. visual) using contextual cues, attitudes, acquired knowledge and experience. Perception is considered as a key cognitive element characterized the highly qualified specialists in various disciplines.

In the medical fields such as radiology, dermatology, pathology, surgery or even pediatrics, perception and visual expertise have a high impact on the work efficiency and effectiveness as well as the correctness of analysis and diagnosis. Medical experts are characterized by faster decision-making and greater accuracy in performing tasks related to their specialization. Numerous studies have demonstrated a number of differences in eye movement parameters and patterns between experts and novices. Those studies are often conducted using eye-tracking techniques.

Eye-tracking (or oculography) is a research method that uses an eye-tracker device to track the point of gaze or eye movement of a person during a task execution. Currently, the most popular eye-trackers are using the video-based method, which allow to non-invasive and non-contact eye movements measurement. In medicine, eye-tracking has been used for many years to identify experts behaviours [1], but also for medical diagnosis, progression or recurrence of disease over time, and even as an objective measure of treatment's effectiveness [2].

Eye-tracking metrics

Eye-tracking metrics might be applied to different aspects of oculomotor behaviour, depending on the type of analysis [3]. The most common eye-tracking metrics are based on fixations and/or saccades. Fixations are represented as discrete samples of almost stable points where the eye is looking. Saccades are defined as eye movements between fixations.

Among the most widely applied **fixation-based metrics** one can find: number of fixations, number of fixations on each area of interest, total number of fixations, fixation duration, total fixation duration, time to first fixation on target, fixation density, repeat fixations. It was shown, that these types of metrics might be used to define observer's engagement [3].

Long fixation duration is correlated with high cognitive workload and higher cognitive effort [4]. This metric was also found as related to the difficulty of the visual content of the performed task [5]. However, fixation duration less than 300ms are believed not to be encoded in memory. The fixation duration statistic can be complemented by the total fixation time [3] and fixation density defined as the total number of gaze points divided by the minimal area to capture all the gaze points [4]. Another useful metric is inter-observer consistency applied to quantify the similarity of observer fixation patterns on an image. This metric measures differences between fixation heatmap among the group of observes. In case of expert vs. novice tasks research [6,7] show that novices made many more fixations than experienced persons. The studies showed that novices and not trained radiographers made many more fixations than experienced observers to cover the visual scene.

Eye-tracking **metrics based on saccades** are related to searching sequence of particular areas of interest. Saccades might be triggered arbitrarily or with stimulus, whereas stimulus does not to be visual [8]. In contrast to fixations, during saccades visual information is not processed, so in the context of perception and visual attention saccadic metrics are not as relevant as in the case of fixations. Among example metrics one can find number of saccades, saccadic amplitude (i.e. saccadic distance), saccade regressions, saccadic duration or saccadic velocity. In the literature one can find several saccadic-related experiments [9]. For example the experiment called pro-saccadic task is based on a single target jumping rapidly and changing its location. Saccadic latency determined by visual processing and decision making might be measured [10], increased saccadic latencies [11] might be related to Parkinson disease. Other parameters measured during this task are the accuracy of saccades or a peak and mean velocity [12]. Another experiment, anti-saccadic task, consists of the target displayed in the middle of patient's vision field. If a new point is displayed (on the right or left), the patient need to look in the opposite direction [9]. This task is unintuitive, as the natural human reaction is to follow the new point. That is why it is referred as higher mental complexity task and it might be applied in diagnosing neural disorders affecting the frontal cortex or the basal ganglia. Saccadic latency data and errors percentage might be also analysed [13,14]. In case of experts vs. novices tasks [9] the experienced users were noticed to use longer eye movements with larger saccade amplitudes, leaving larger areas of the image unfixated. Other eye-tracking metrics useful in such tasks are: time to first fixation on the condition, a number of fixations before the region of interest, total number of fixations, mean fixation time, and maximum visit (known also as dwell time) count in a region of interest. Smaller score of these metrics indicates greater expertise.

Smooth pursuit is another phenomenon detectable by eye-tracking equipment. It can occur unintentional when dynamic stimuli are presented and when an observer follow a movement in a presented stimulus [15]. Standard velocity of the eye in this move is between 10 and 30 per second. Popular experiment covering smooth pursuit consists in following with eyes an appearing object [9]. Smooth pursuit movement should be detectable when the object moves slowly (less than 100◦/s). Problems with maintaining smooth pursuit eye movement, when eyes are moving faster or slower than the moving object, might indicate schizophrenia [16]. What is more, intrusive, catch-up saccades might also occur. Problems with smooth pursuits may indicate also other diseases such as autism (Johnson et al., 2016), Alzheimer disease or Parkinson disease [17].

There are also eye-tracking metrics applying both, saccades and fixations. Any combination of **saccades and fixations** is called scanpath. An ideal scanpath is characterized with short time and straight-line saccade to the specific target [18]. Usually a deviance from this ideal straight line is interpreted as poor search [19]. Popular scanpath-based metrics are scanpath duration, scanpath length, spatial density, transition matrix, scanpath regularity (repeatability), scanpath direction (search strategies), saccade/fixation ratio [15]. Scanpath-based measures were applied for instance in comparison of gaze patterns of nurses giving medicines to patients in a simulated environment [20]. The study consisted in analysing their ability to find errors in the patient identifier. Research showed that these nurses who were able to find the mistake had more predictable eye fixation sequences in comparison to nurses who failed to detect an identification error. The successful group of nurses had their fixations in a row on the patient's chart. Other studies [21] investigates the dispersion of eye movements in the task of radiologists' lung nodules search in tomographic images. In this study two metrics were applied: the lung volume fraction that was enclosed within the gaze volume and the nearest distance (in pixels) to the search path for each embedded nodule. Results revealed that radiologists on average search only 26% of the lung parenchyma covering, however, 75% of nodules. In this study the eye movement topology was compared. The eye movement topology was considered as a scan path, where fixations represent nodes and saccades are path edges. Other study comparing experts to novices [22] shown that experts' eye movements are more systematic, but on the other hand, they covered less of the image. However further research are necessary, as the influence of the systematic viewing on improving performance has not been confirmed [9]. The scan paths analysis may cover also checking elements on which the eye fixated first and which are re-fixated [3].

Areas of interest (AOIs) or region of interest (ROIs) are high importance part of stimulus, defined usually based on the semantic information of the stimulus. They might be defined either before the experiment or after it, during the analysis process. Typical metrics related to AOIs are based on a transition, which is defined as a movement between two AOIs. Among these metrics one can find transition count defined as number of transitions between two AOIs, the visit (or dwell) time within an AOI, the AOI hit informing if a fixation is within an AOI or not. Typical AOI-based tasks cover exploration of images or multimedia divided into several AOI containing a lesion which should be noticed by an observer. Test are conducted in regard to such segments [21,23].

Other AOI-based tests apply time-based measures analysing for instance time needed to reach particular region, time necessary to gaze this region or the ratio of gaze duration at the region to the total gaze duration [9]. the percent of gaze time was applied by Almansa et al. [24] to check central vs. peripheral gaze hypothesis in the study where eight peripheral and one central segments with adenoma were presented. Results revealed a positive correlation between the detection rate and the central visual gaze pattern. On the other side, the total gaze time per area of interest was applied in the exert vs novice study [25] of the endoscopy monitor. More experience users tended to have higher mean gaze times per the AOI and higher percentage of overlap between measured gaze positions in different areas. Users with different level of experience was also analysed with the measure of time to the first fixation on AOI. For instance Cooper [26] found that experts were able to accomplish the task in the shorter time but also they spent more time looking at the AOI surrounding a challenging lesion. Experts vs novices differences were also analysed using eye-gaze patterns [23]. In the study experts and residents surgeons were observed while watching operative videos. Expert surgeons had significantly closer eye-gaze patterns.

Popular task related to AOI is the visual-paired comparison task, where an object is presented to a user and after that the same object is presented, but this time together with a new one. It is expected [27] that a user is focused on the previously seen object ("novelty preference" paradigm). This is measured with fixations duration separately for both objects treated as AOI. This task was applied for example among MCI patients [28].

Another, relatively novel and promising task is to ask users to observe the scene without any instructions. Such paradigm was applied with a set of static images [29] to diagnose autism or movie clips [30] to differentiate between ADHD, fetal alcohol spectrum disorder as well as Parkinson disease and a control group.

The case studies

Participants

The participants have normal or corrected-to-normal vision. The participation in both experiments was voluntary, no compensation was offered. The experiments were approved by the Research Ethics Committee of the Lublin University of Technology and all participants received verbal and written information about the study. All participants signed an informed consent.

Equipment and measurement

The experiment was conducted in a special testing laboratory in Laboratory of Motion Analysis and Interface Ergonomics (LARiEI) of Lublin University of Technology. The room was illuminated with standard fluorescent light and outside light was blocked to ensure stable conditions for the duration of the experiments. All participants were using the same software and hardware settings (Laptop Asus G750JX-T4191H with Intel Core i7-4700HQ and 8GB of RAM).

Eye activity were recorded using screen-based eye tracker Tobii Pro TX300 (Tobii AB, Sweden). The Tobii Pro TX300 performed binocular tracking with sampling frequency 300Hz using dark pupil and corneal reflections method to detect eye movement. The gaze was measured by the eye-tracker with accuracy of 0.5° at illumination within 600-1000 lux range.

The experiments were created using Tobii Studio 3.3 and a visual stimulus were presented on separate monitor equipped with eye tracker (23" TFT monitor at 60 Hz). The participants were seated while working with the application (fig. 1). The distance between the screen and the participant was in a range from 50 to 80 cm depending on individual participant preferences (comfortable position for working with computer).

At the beginning of each session, the eye tracker were calibrated using a 9-point built-in calibration procedure. Once calibrated, the participants were provided with the experimental instructions on the screen.

Experiment 1 – novice and expert visual attention

The goal of the experiment was to identify there are significant differences in the visual analysis of medical images with disease abnormality by experts and novices. What is more, is there a difference in the group of novices between those who managed to identify the abnormality correctly and those who did not succeed. In the paper, the partial results from the experiment will be presented.

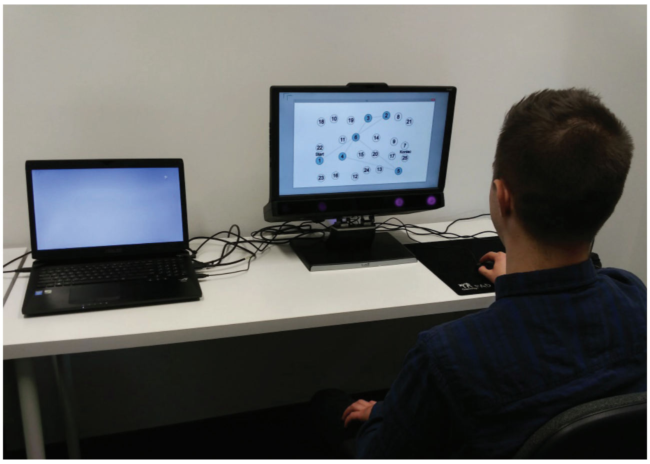

Fig. 1.
The participant working with an application in the testing room during eye-tracking recording

Participants

Eight students of the first and the second degree in Biomedical Engineering and a doctor with specialization in radiology took part in the experiment.

Experimental task

Stimuli consisted of 6 radiographs or computed tomography (CT) images, 4 images were with a specific lesion (e.g. brain/lung tumor) demanding a varying degrees of expertise to recognizing that abnormality. The task was to identify the abnormality in the presented image and indicate it using a computer mouse.

Results

The CT image with brain tumor was correctly identified in a fast manner by 6 students (fig. 2) and a radiologist. However, all students need to search a larger area to find the abnormality and make a diagnostic decision. The scanpaths presented in fig. 2 proved that visual attention of the student is distributed over whole image, while the radiologist made the decision when she spotted the tumor.

Similarly, the visualization method of eye-tracking data were used to present visual search pattern and visual attention of participants for the projectional radiograph of lung with tumour lesions. The fig. 3 presented the comparison of scanpaths between the radiologist, students who correctly identified the lesions and those who did not. Same as in the previous stimulus, the expert scanpath is the shortest with long fixation on the identified lesion. The scanpath of students who correctly identified the lesions is much longer, students give attention to much larger area but the longest fixations are also presents on area with lesion. On the other hand, the students who did not

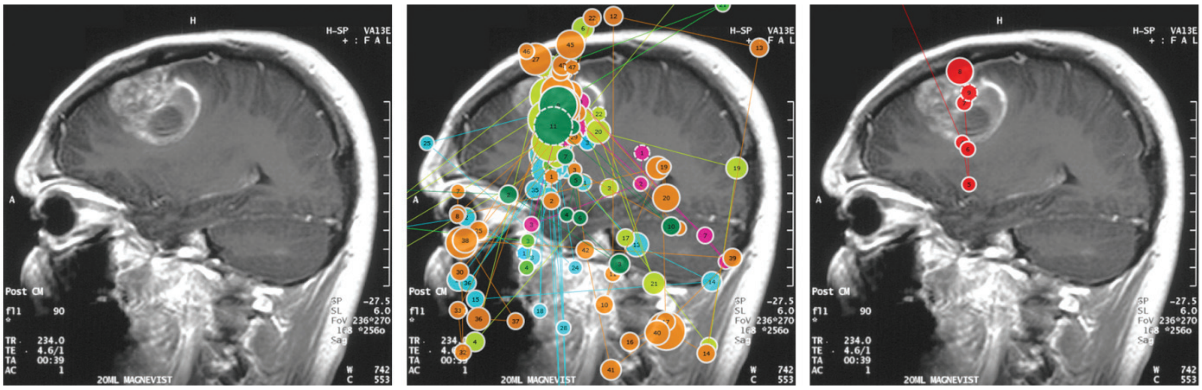

Fig. 2.
CT image of head with brain tumour, from left: original image presented to participants, the scanpaths of students who correctly identified the tumour, the scanpath of radiologist

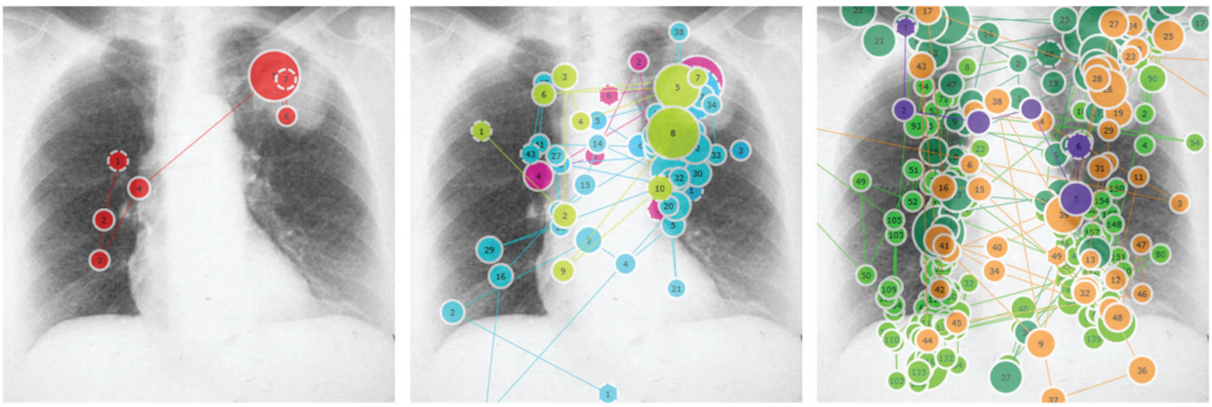

Fig. 3.
Projectional radiograph of lung with tumour lesion, from left: the scanpath of the radiologist, the scanpaths of students who correctly identified the lesions, the scanpaths of students who were not able to identified lesions

identified lesion they scanned the whole area of image not knowing where to look for potential lesions.

The visual attention can be also presented as heat map indicating where not individuals fixations points but whole groups focus their visual attention. The heatmaps of radiograph of lung with tumour lesions for the radiologist and 2 students groups are presented in fig. 4. Similarly to scanpaths, it can be notice that students, who did not identified lesion, distributed their attention on many elements of the image. Using heatmap it is possible to separate three different areas with the most visual attention given.

Moreover, the additional analysis were conducted using AOI. The tumour lesion was marked as a single elliptical AOI, while rest of image were marked as not on AOI. To compare the groups eye-tracking metrics such as time to first fixation, number of fixation, visit duration and number of visit were calculated for AOI and not on AOI (Table 1). The results confirm the previous conclusions that students who identified the lesion need less fixations on AOI and non on AOI to make decision. They also have less visits on AOI and non on AOI and duration of visit last shorter. Two of four students who identified the lesion need only

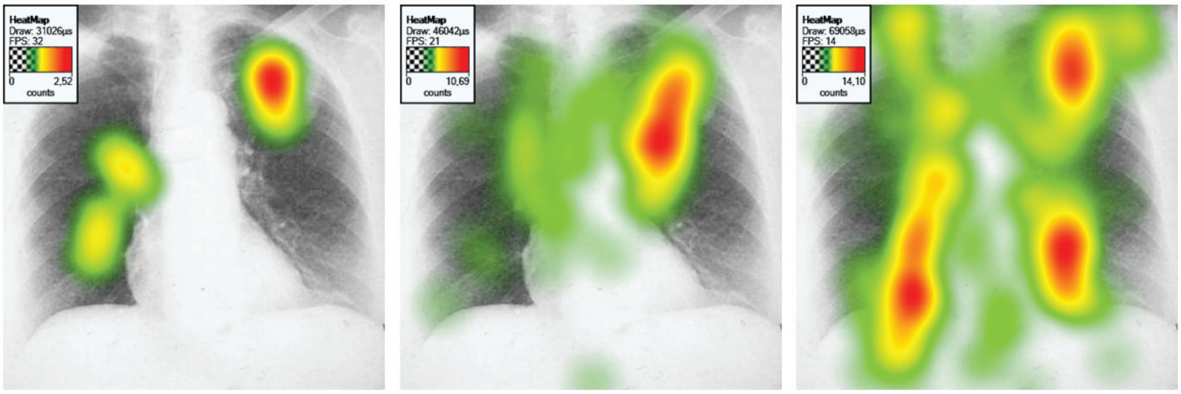

Fig. 4.
Projectional radiograph of lung with tumour lesion, from left: the heatmap of the radiologist, the heatmap of students who correctly identified the lesions, the heatmap of students who were not able to identified lesions

Table 1.
Eye-tracking metrics statistics for AOI and not on AOI

Participant	Time to first fixation (s)		Number of fixations		Visit duration (s)		Number of visits	
	AOI	Not on AOI	AOI	Not on AOI	AOI	Not on AOI	AOI	Not on AOI
Radiologist	0.45	0.00	3	4	1.40	0.45	1	1
Students who identified the lesion								
Student 1	0.86	0.00	1	5	1.37	0.86	1	1
Student 2	0.78	0.00	3	7	2.27	1.51	2	3
Student 4	0.83	0.00	1	5	1.19	0.97	1	2
Student 6	1.83	0.00	8	35	1.59	6.29	6	7
Mean (SD)	1.08 (0.50)	0.00 (0.00)	3.25 (3.30)	13.00 (14.70)	1.61 (0.47)	2.41 (2.60)	2.50 (2.38)	3.25 (2.63)
Students who did **not** identify the lesion								
Student 3	10.48	0.00	7	41	2.83	15.86	5	6
Student 5	-	0.08	-	6	-	1.52	-	1
Student 7	0.00	0.03	7	155	1.12	13.98	5	5
Student 8	1.15	0.00	8	41	2.86	10.14	5	6
Mean (SD)	3.88 (5.75)	0.03 (0.04)	7.33 (0.58)	60.75 (64.69)	2.27 (1.00)	10.38 (6.37)	5.00 (0.00)	4.50 (2.38)

one visit to recognize the lesion, while student who did not identified lesion make multiple visit on AOI and could not recognize it as tumour lesion. What is interesting one of the student did not have any visit on AOI.

Conclusions

Eye-tracking technique allow to analyses visual behaviour and reveal visual search strategies between different groups. The presented results from experiment show that both qualitative (in the form of visualizations) and quantitative data can better understand the differences between novice and more experience specialist. In the study it was not possible to compare quantitative data between experts and novice, because only one expert participated in the experiment and her data can be treated as referential.

Experiment 2 – cognitive performance

The goal of the experiment was to identify the visual attention differences between participants with high and low cognitive performance in the battery of neuropsychological tests. In the paper the results only for the first board with equally spaced dot configurations without distractors in the background from the Ruff Figural Fluency Test (RFFT) will be presented.

Participants

Seventy healthy male students at the Lublin University of Technology were recruited to the experiment. From the initial group, 11 participants were excluded due to a psychiatric medical history, low eye-tracking ratio or poor data quality. The final group consisted of 61 students of Computer Science and Biomedical Engineering ranging in age from 18 to 31, mean age was 21.4 years. All participants were right-handed.

Experimental task

The RFFT was created by Ronald M. Ruff as a nonverbal fluency measures [31]. In the experiment the computerised version of test was used. The RFFT application was developed in Java Swing and it is operated using a computer mouse. The RFFT application covers three different dot configurations and two distractors added to the first, symmetric dot configuration. Each trial lasts 60 seconds and it is preceded by a three-element testing attempt. When the participant has created single design, he had to press the "Next→" button to clear the board. The accomplished designs were presented in the form of miniatures at the left part of the screen. The computerised RFFT allows to log all user interactions, and thus to register not only the total number of designs or unique designs, but also the execution time for each individual design or part of the test.

Results

The median of the total number of unique designs created by participants on the first board was equal to 16.5. Therefore the participants were divided into two groups: high cognitive performance group where the total number of unique designs was equal or higher that 17 (30 participants), and low cognitive performance group where the total number of unique designs was lower than 17 (31 participants). The analysis for the RFFT was performed using the three AOIs marked on the interface of application (fig. 5). The AOIs covers: area for creating new design (marked with orange colour on fig. 5), area of "Next→" button to save created design and area for display of all created designs design (marked with purple colour on fig. 5).

For each AOI, the following metrics were calculated: time to the first fixation, number of fixations, mean fixation duration, max fixation duration, total fixation duration, number of visit, mean visit duration, max visit duration and total visit duration. Those features were calculated for the first trial (60s) as well as for 6 timespans lasting 10s during the first trial.

The statistical analysis was performed using Statistica 13 (Dell Inc., USA) to investigate the research question "Is there any differences in eye behaviour between participants with high and low cognitive performance (measured in the total number

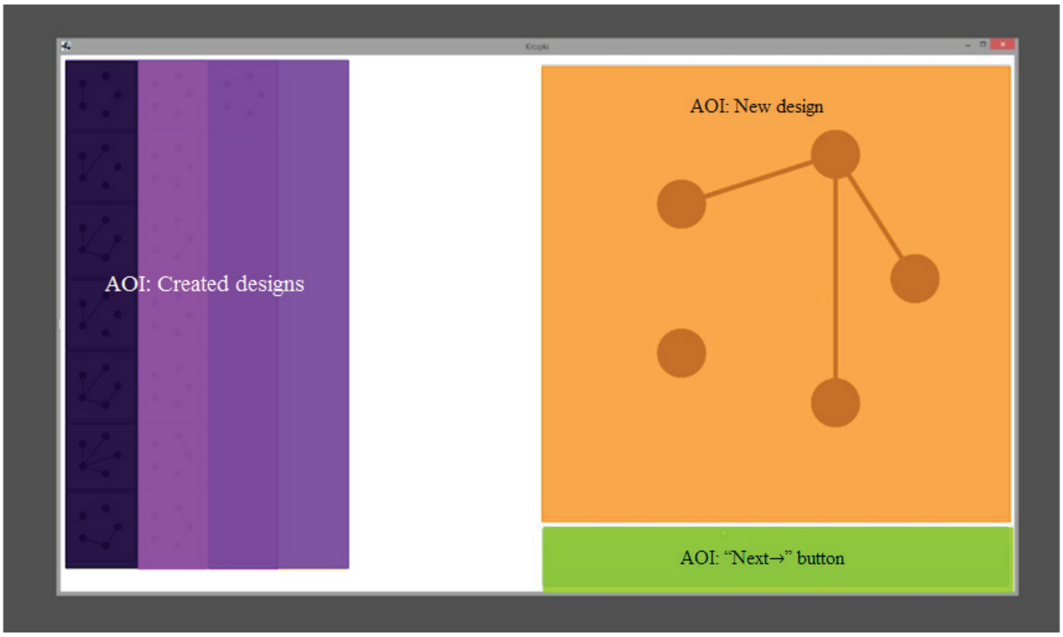

Fig. 5.
The RFFT application interface with marked AOIs

of unique designs) in defined AOIs?". The statistically significant difference (with p=0.05) between two groups using unpaired two-sample t-tests were found for mean values of 7 metrics in Created designs AOI, 4 in "Next→" button AOI and 3 in New design AOI.

The mean number of visits in each AOIs differed significantly between groups (fig. 6). The participants from the group that created less than 17 designs (low cognitive performance group) have fewer visits in New design and "Next→" button AOIs, which corresponds to the value of the total number of unique designs achieved by them on the first board in RFFT test. Moreover, statistically significant difference in the number of visits in Created designs AOI reveals that participants from low cognitive performance group twice as often checked the created designs than participants from high cognitive performance group. This may indicate that participants from low cognitive performance group did not remember the created designs therefor they looked at them more often.

Another significant difference can be observed in mean and max visit duration in Created designs AOI between groups (fig. 7). The low cognitive performance group not only made more visits but also the mean and max visit duration last longer (despite they got fewer designs to inspect). The difference in visit duration results from more number of fixation per visit. The participant from high cognitive performance group made about 5.6 fixations, the participant from low cognitive performance group made about 15 fixations. Moreover, the difference in mean fixation duration values between group is not statistically significant, what can indicate less effective searching or more complicated information to process.

The significant difference between groups is observed also in mean and max visit duration in New design AOI (fig. 8). The participants from the high cognitive performance group needed less time (about 2.95s in average) than the participants from the low cognitive performance group (about 3.97s in average).

Conclusions

The analyses of visual attention during non-verbal fluency test such as RFFT allows to discovery the behaviour patterns for users with different performance. The analysis with AOIs during the first trial of RFFT reveals that participants from low cognitive performance group twice as often checked the created designs than participants from high cognitive performance group. What is more, in the low cognitive performance group the mean and max visit duration last longer, what is cause by more fixations made by participant in average during single visit.

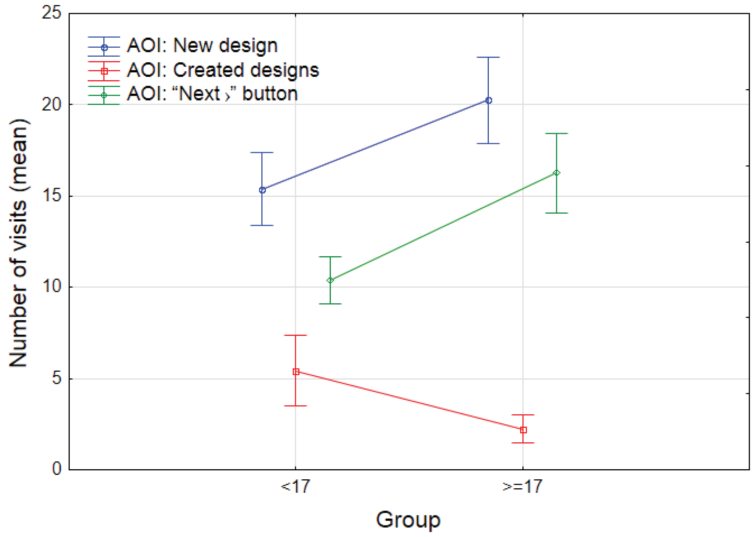

Fig. 6.
The plot of mean number of visit in each AOI for both groups (mean value is marked, vertical bars represents 0.95 confidence intervals)

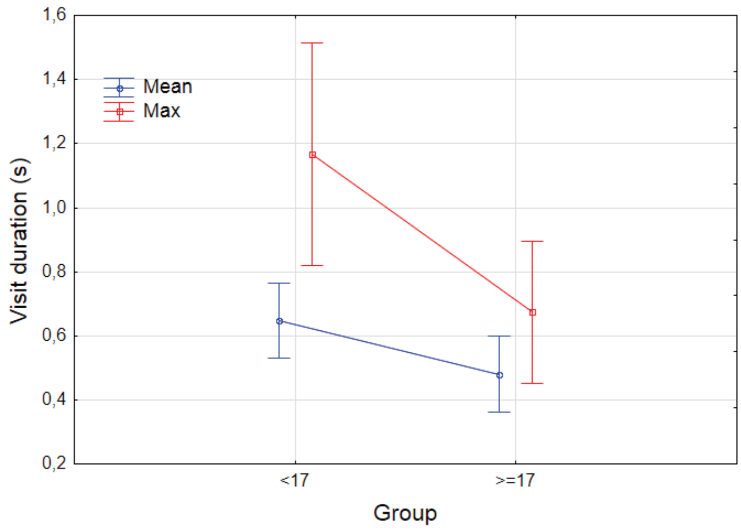

Fig. 7.
The plot of mean and max visit duration in Created designs AOI for both groups (mean value is marked, vertical bars represents 0.95 confidence intervals)

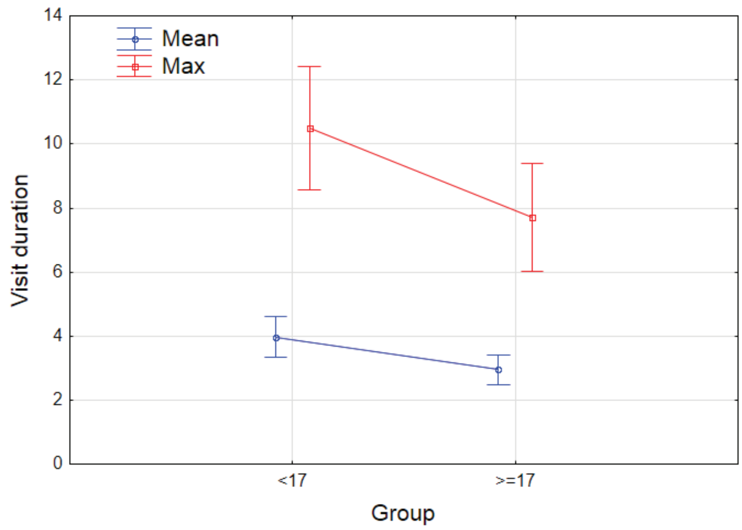

Fig. 8.
The plot of mean and max visit duration in New design AOI for both groups (mean value is marked, vertical bars represents 0.95 confidence intervals)

The generalized process of visual attention analysis using eye-tracking

The experiments conducted to measure perception and visual attention might be realized according some general standards. They might be gathered as particular steps in the analysis process (fig. 6).

Proper selection of participants to the study is the crucial issue. A researcher needs to define the expected level of participants experience and knowledge to adjust it to the complexity of the experiment. The number of participants needs to be selected depending on whether the quantitative or qualitative analysis will be carried out. This number usually needs to be higher in case of the quantitative analysis where statistical methods will be applied to ensure statistically meaningful results.

If the analysis will be based on the areas of interest, the appropriate selection of these areas needs to be specified from the graphical content before eye-tracking tests. Such areas should be selected to cover crucial content, it can take any geometric form but on the other side they cannot overlap or be too extensive. A good example of such area of interest is the specific lesion visible in the x-ray image. This area of interest meets all requirements because it covers only a particular element of the presented image and is of great importance to the analysed issue.

The next step is to determine eye-tracking metrics. In case of visual attention analysis with areas of interest examples of often applied metrics are: the number of visits, the duration of visits, the number of fixations, or the length of fixation in a particular area of interest. To properly chose metrics, a researcher must consider what differences between participants or what participants features he is looking for. Another issue related to metric selection is the data form, possibility of their further processing and ways of their presentation. Choosing the metric set one should also take into account the duration of a single observation to assume the amount of data that will be received.

Once the results are gathered, the analysis might be started. Depending on the participant number and the study characteristics, a researcher should consider grouping the participants into groups. It enables to precise the characteristics of participants. Good example of such division are experts vs. novices studies or comparison of participants with extreme differences. On the other hand, the analysis might be also conducted individually for each participant.

If the results of a study will be presented both, graphically and by statistical data, the analysis should be carried out in two steps. In the first one, the graphical representation of the study results should be emphasizes, taking into account the division of participants into groups. In the second step the gathered results should be supported with the statistical analysis results based on selected eye-tracking metrics.

Discussion

Figures and tables together with captions should appear as close as possible to the places in which they are referred to for the first time.

Eye-tracking technique might be applied in different branches such as marketing, ergonomic, medicine or psychology. Depending on the task one needs to choose proper eye-tracking metrics to reveal features related to the aims of the analysis. Metrics discussed in the paper are related to perception and visual attention. Two example case studies were presented. The first one was dedicated to identify significant differences in the analysis of medical images with disease abnormality by experts and novices. The eye-tracking analysis was based on scanpaths, heat map with fixation point and AOI. The results show that both qualitative and quantitative data might reveal the differences between novice and more experience specialist.

The second study was performed to identify differences in the visual attention of participants with high and low cognitive performance in the RFFT test. The analysis was conducted using three AOIs.

Fixation-based metrics such as scan path occurred to be useful in finding a strategy for image searching by individual participants in the experiment. What is more, heat maps provided information about attention distribution in the analysing areas. Aggregated eye-tracking data presented for example in tabular

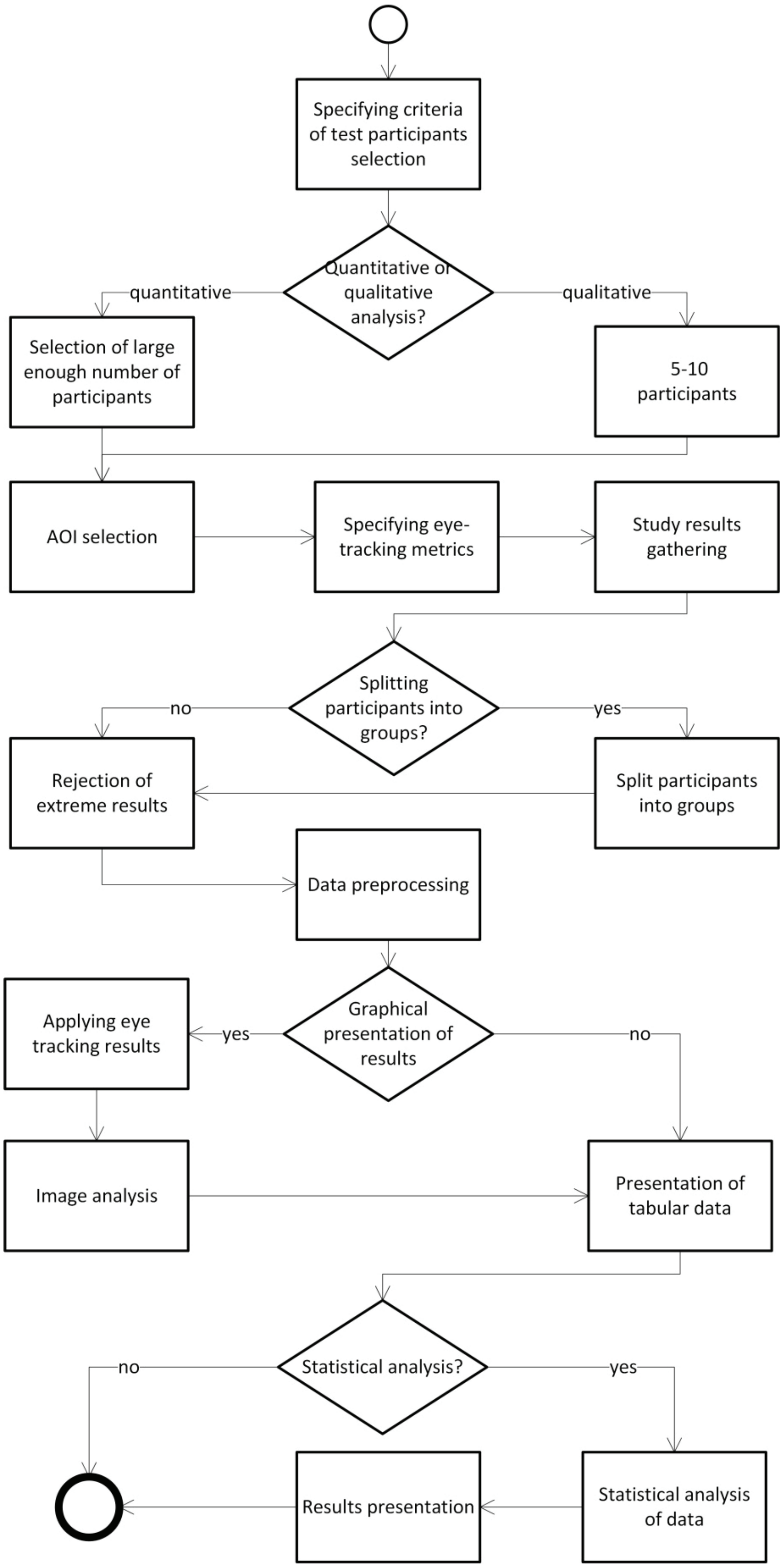

Fig. 9.
The generalized process for visual attention analysis with eye-tracking data

form also provides specific information on the differences between groups of subjects and enable to draw conclusions about participants characteristics.

Presented analysis of perception and visual attention results shows the difference between eye-tracking metrics. Proper selection of eye tracking measures is a key issue. Eye tracking metrics are not universal, many of them occurs useful only for specific purposes. This phenomena was considered in the generalized process of visual attention analysis using eye-tracking presented in the paper. This procedure is a result of the literature research and analysis of own eye-tracking studies.

Acknowledgment

Authors would like to sincerely thank the students who so graciously agreed to share their time and participate in the research.

Special thanks to Marta Walczewska, Róża Dzierżak, Weronika Borys and Grzegorz Kęsik for their contribution in conducting the experiment and analysis of data.

References

1. Kundel HL, Nodine CF, Carmody D. Visual scanning, pattern recognition and decision-making in pulmonary nodule detection. Invest Radiol 1978;13:175–81.
2. Kuechenmeister CA, Linton PH, Mueller T V, et al. Eye tracking in relation to age, sex, and illness. Arch Gen Psychiatry 1977;34:578–9.
3. Bylinskii Z, Borkin MA. Eye fixation metrics for large scale analysis of information visualizations. ETVIS Work. Eye Track. Vis., 2015.
4. Shojaeizadeh M, Djamasbi S, Trapp AC. Density of Gaze Points Within a Fixation and Information Processing Behavior. Int. Conf. Univers. Access Human-Computer Interact., 2016, p. 465–71.
5. Fitts PM, Jones RE, Milton JL. Eye movements of aircraft pilots during instrument-landing approaches. Ergon Psychol Mech Model Ergon 2005;3:56.
6. Kelly BS, Rainford LA, Darcy SP, et al. The development of expertise in radiology: in chest radiograph interpretation,"expert" search pattern may predate "expert" levels of diagnostic accuracy for pneumothorax identification. Radiology 2016;280:252–60.
7. Manning D, Ethell S, Donovan T, et al. How do radiologists do it? The influence of experience and training on searching for chest nodules. Radiography 2006;12:134–42.
8. Krauzlis RJ. Eye movements. Fundam Neurosci 2008:775–92.
9. Harezlak K, Kasprowski P. Application of eye tracking in medicine: A survey, research issues and challenges. Comput Med Imaging Graph 2017.
10. Darrien JH, Herd K, Starling L-J, et al. An analysis of the dependence of saccadic latency on target position and target characteristics in human subjects. BMC Neurosci 2001;2:13.
11. Antoniades CA, FitzGerald JJ. Using saccadometry with deep brain stimulation to study normal and pathological brain function. JoVE (Journal Vis Exp 2016:e53640--e53640.
12. Yang Q, Wang T, Su N, et al. Specific saccade deficits in patients with Alzheimer's disease at mild to moderate stage and in patients with amnestic mild cognitive impairment. Age (Omaha) 2013;35:1287–98.
13. Pereira MLGF, von Zuben AC, Aprahamian I, et al. Eye movement analysis and cognitive processing: detecting indicators of conversion to Alzheimer's disease. Neuropsychiatr Dis Treat 2014;10:1273.
14. Alichniewicz KK, Brunner F, Klünemann HH, et al. Neural correlates of saccadic inhibition in healthy elderly and patients with amnestic mild cognitive impairment. Front Psychol 2013;4.
15. Blascheck T, Kurzhals K, Raschke M, et al. State-of--the-art of visualization for eye tracking data. Proc. EuroVis, vol. 2014, 2014.
16. Diefendorf AR, Dodge R. An experimental study of the ocular reactions of the insane from photographic records. Brain 1908;31:451–89.
17. Fukushima K, Ito N, Barnes GR, et al. Impaired smooth-pursuit in Parkinson9s disease: normal cue-information memory, but dysfunction of extra-retinal mechanisms for pursuit preparation and execution. Physiol Rep 2015;3:e12361.
18. Conversy S, Hurter C, Chatty S. A descriptive model of visual scanning. Proc. 3rd BELIV'10 Work. BEyond time errors Nov. Eval. methods Inf. Vis., 2010, p. 35–42.

19. Goldberg JH, Kotval XP. Computer interface evaluation using eye movements: methods and constructs. Int J Ind Ergon 1999;24:631–45.
20. Marquard JL, Henneman PL, He Z, et al. Nurses' behaviors and visual scanning patterns may reduce patient identification errors. J Exp Psychol Appl 2011;17:247.
21. Rubin GD, Roos JE, Tall M, et al. Characterizing search, recognition, and decision in the detection of lung nodules on CT scans: elucidation with eye tracking. Radiology 2014;274:276–86.
22. Kok EM, Jarodzka H, de Bruin ABH, et al. Systematic viewing in radiology: seeing more, missing less? Adv Heal Sci Educ 2016;21:189–205.
23. Khan RSA, Tien G, Atkins MS, et al. Analysis of eye gaze: Do novice surgeons look at the same location as expert surgeons during a laparoscopic operation? Surg Endosc 2012;26:3536–40.
24. Almansa C, Shahid MW, Heckman MG, et al. Association between visual gaze patterns and adenoma detection rate during colonoscopy: a preliminary investigation. Am J Gastroenterol 2011;106:1070.
25. Dik VK, Hooge ITC, van Oijen MGH, et al. Measuring gaze patterns during colonoscopy: a useful tool to evaluate colon inspection? Eur J Gastroenterol Hepatol 2016;28:1400–6.
26. Cooper L, Gale AG, Darker IT, et al. Radiology image perception and observer performance: How does expertise and clinical information alter interpretation? Stroke detection explored through eye-tracking 2009.
27. Chau SA, Herrmann N, Eizenman M, et al. Exploring visual selective attention towards novel stimuli in alzheimer's disease patients. Dement Geriatr Cogn Dis Extra 2015;5:492–502.
28. Lagun D, Manzanares C, Zola SM, et al. Detecting cognitive impairment by eye movement analysis using automatic classification algorithms. J Neurosci Methods 2011;201:196–203.
29. Wang S, Jiang M, Duchesne XM, et al. Atypical visual saliency in autism spectrum disorder quantified through model-based eye tracking. Neuron 2015;88:604–16.
30. Tseng P-H, Cameron IGM, Pari G, et al. High--throughput classification of clinical populations from natural viewing eye movements. J Neurol 2013;260:275–84.
31. Ruff RM. Ruff Figural Fluency Test professional manual. Odessa, FL: Psychological Assessment Resources Inc.; 1988.