An approach to VR based head and eye movement correlation evaluation system

Dutko, J.[1], Vargic, R.[2]
[1,2] Affiliation (Institution) of the First Author, Postal Address, Country
xdutko@stuba.sk

Abstract - **This contribution describes a new VR based head and eye movement correlation evaluation systems. Systems allows capturing of the gaze position using eye tracker mounted in the head mounted display for the virtual reality as well as the actual head position and rotation. There values can be evaluated under number of prepared scenes and experiments. Using the environment some of the hypotheses related to the head and eye movements in the virtual reality were studied and tested. The preliminary results are shown and are promising. However, there are significant systematic errors that shall be eliminated.**

Keywords **- Component; virtual reality, eye tracking, head tracking**

I. Introduction

Nowadays is Virtual reality (VR) successfully been applied in various domain as entertainment, healthcare or education. Typical VR end device is the head mounted display (HMD). One of the directions how to enhance the capabilities of the HMD is eye tracking [1]. This allows to capture the eye gaze in real time and to provide important feedback. Depending on the usage area, the feedback can be used in many ways. The displayed content can be adapted in real time, e.g. game scenario changes and characters behave differently. Or when based on the eye movements the decrease of attention level detected or even user is falling asleep, he can we waked up by strong sound. Or the feedback can be stored and analysed offline to change the content, the layout, the presentation scenario, the interaction options. The feedback can be even used for foveated rendering [2], which is the upcoming graphics rendering technique that uses the eye tracker integrated with HMD to reduce the rendering load by significantly reducing the image quality in the marginal view (outside the zone viewed by the fovea). The applications are endless. However, there are several technical problems that shall be solved to be able to use full potential of eye tracking in VR. The scene is most cases dependent on the head position, so errors of both positions (head, eyes) contribute to the resulting error. There is eye position evaluation problem during fast head movements [3]. In some cases, the head movements and eye movements are correlated, so up to some degree the head movements (and corresponding viewport position) can be predicted from previous movements of head and eyes [4]. In this contribution we describe the system that allows the measurements, analysis and evaluation of head and eye movements and incorporation of the resulting feedback into the interactive multimedia delivery. We evaluate the preliminary results and propose further directions.

II. VR headsets and eye tracking

VR is most often used in entertainment applications such as games and 3D cinema [6]. In the field of social sciences and psychology, virtual reality offers a cost-effective tool for studying and replicating interactions in a controlled environment. The largest virtual reality launcher for regular consumers was Google Cardboard [5], which brought people into a 3D environment with the ability to interact with VR models by combining cheap cardboard and smartphone display. The basis of virtual reality is the effort to display as closely as possible the spatial models and scenes, the manipulation with them, the creation of the real world, its part with all its rules and rules, the movement in a three-dimensional space and all that in real time. Basic computer graphics techniques are used. VR headsets are supplied by a number of companies. On the [6] page there are a few VR headsets that are currently dominating the market: Sony PlayStation VR, HTC Vive, Oculus Rift, Google Daydream View, Samsung Gear VR,

Eye tracker technology is based on eye pupil position detection [7]. There exist lot of eye tracking devices on the market, among the most known are [8]: Tobi, SMI, EyeLink, ISCAN, LC Technologies, EyeTech, The Eye Tribe, Ergoneers, Smart Eye, Mirametrix, Pupil Labs, GazePoint. One of the most famous is Tobi [9]. Tobi, SMI, ISCAN, EyeTec, Pupil Labs also produce the eye tracking devices suitable for mounting into VR headset. An eye tracker built into the HTC Vive VR headset and its features were introduced at the GDC conference [1].

III. Exploitation of the eye tracker in viewport prediction process

The movement of the eyes in conditions where the head is restricted in movement is well understood, and there are a number of publications describing various types of eye movements such as saccades [10]. Saccades are characterized by [4] relationships between motion amplitude, maximum speed and duration. For example, when the amplitude and the direction of the saccade are known, it is possible to predict the progress of the saccade. The data presented in the studies [4] indicate the interaction between head and eye motor systems in which head movement commands change concurrent saccades. In contrast, the kinematics of the head movement is predictable only on the basis of the amplitude of the head movement [4]. The data provide support for the hypothesis that the eye and the head are controlled separately with the existence of interaction between these systems.

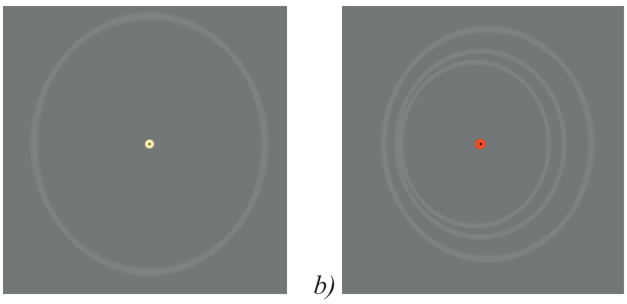
When viewing objects in conditions where the head is not restricted in motion, different models have been established considering that eye movement and head movement are driven only by the perceived velocity of the visual target, the so-called speed hypothesis [11]. Research in [11] considers the importance of the initial position of the eye, the position of the target relative to the retina and the velocity of the target for predicting the movement of the head during tracking. For example, when the eyes are already at the edges of the eyebrow, further tracking in this direction will be accompanied by a larger head movement than when the eyes are centred in the eyebrow, even if the target speed is the same. It is possible to predict [4] the contribution of the head to shift the view with reasonable accuracy with respect to the amplitude and direction of the target shift and the initial position of the eyes. When the target speed is low enough, visual tracking by combination of head movements and continuous eye movement is often referred to as eye tracking.

In [12] is described the vestibulo-oculus reflex and its contribution to coordination of the head and eyes for the acquisition of visual targets. The results of the experiment say, for example, that when the subject moved his head and eyes to gain visual objectives in the horizontal plane, the eye movements consisted of an initial saccade in the direction of the head movement followed by a slower return to the orbital centre that compensated for the remaining movement of the head.

In order to determine relationships between the head and eye movements and perform predictions, machine learning algorithms can be used. Recurring neural networks [13] [14] and its subtype LSTM have been successfully applied across different domains and are also useful for prediction. Tensorflow [15] contains an LSTM network implementation and is suitable for use also for growing popularity.

IV. Proposed Evaluation System

The system was designed to determine the relationship between eye movement and head. The system is designed to cover various aspects and hypotheses related to eye and head co-ordination. The system is fully based on nowadays available commercial core components. For eye tracking the solution from Pupil Labs [16] was chosen, for VR the HTC Vive [17] HMD. Eye tracker needs to be mounted to HTC Vive.



a) *b)*

Figure 1. Calibration screens in hmd-eyes, 2D version (a), 3D version (b)

The software part is based on Pupil Labs' open source products. Our system uses Pupil Capture, which offers real-time feedback on the level of maturity determination, and it is also possible to customize some parameters of the algorithm using. It offers a variety of plugins which can be used e.g. to determine blinks or fixations.

Pupil Labs also offers an open source plugin to the Unity called hmd-eyes [(20)]. Plugin offers the basic building blocks for building applications for VR and AR. It includes implementation of calibration for 3D and 2D eye pupil detection (see Figure1), example scenes, real time view of VR view, creation of heatmaps, basic export of data into csv format. It shows, how to implement communication to the Pupil Capture via the ZeroMQ library [19] using the publish-subscribe pattern.

The system is implemented in Unity [20] environment, where the hmd-eyes plugin is used. The system has been extended in many directions. It indicates the position of the eyes and the head, visualization of head rotation in the space. The Architecture of the proposed system is on Figure 2.

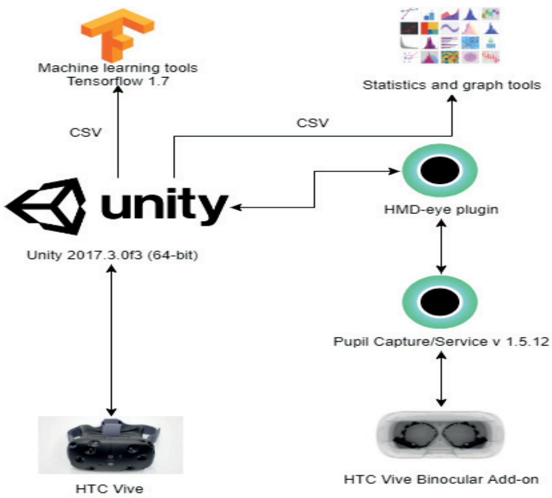


Figure 2. Architecture of the proposed evaluation system

TABLE I. Parameters of the Proposed Eye Tracking System

Parameter	Parameter value
Tracking Frequency	120Hz
Gaze Accuracy	~1.0deg
Gaze Precision	~0.08deg
Camera Latency	5.7ms
Processing latency	3-4ms on i5 CPU
Resolution	640x480

TABLE II. Parameters of the HMD and PC

Parameter	Parameter value
HMD resolution	1080 x 1200 pix. per eye (2160x1200 combined)
HMD refresh rate	90 Hz
HMD Field of view	110 degrees
OS, RAM, processor	Microsoft Windows 10 Home, 8GB, Intel(R) Core(TM) i7-6700K CPU @ 4.00GHz

The system contains 5 research scenes:

- The market scene (based on the scene from hmd-eyes)
- Cube tracking scenes
- Scenes of 360-degree videos
- Game scene

The evaluation session contains these basic steps:

- The participant sits on the seat and deploys HTC Vive.
- Pupil Settings are checked and adjusted
- Pupil Capture is started, and confidence is checked, then in HTC Vive application the focus the camera on the eye tracker checked and adjusted when needed (Figure. 2).
- Calibration is performed
- The selected scene is launched, and experiment is performed.

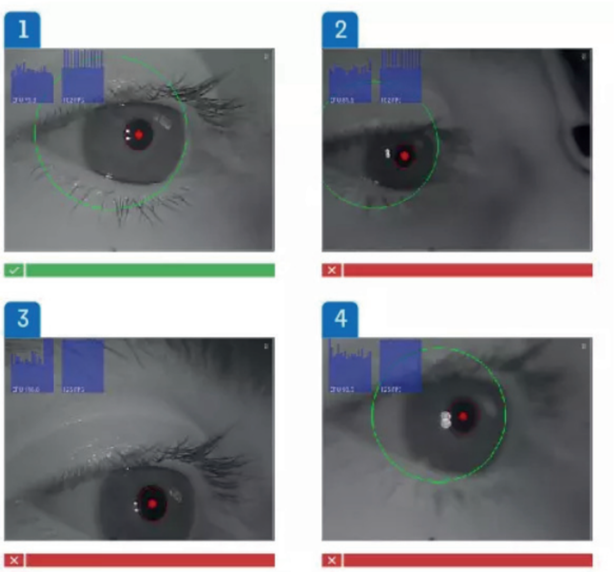


Figure 3. Good and bad eye video, the only good is (1) where the eye is sharp in focus and whole range of the eye movements are visible [2]

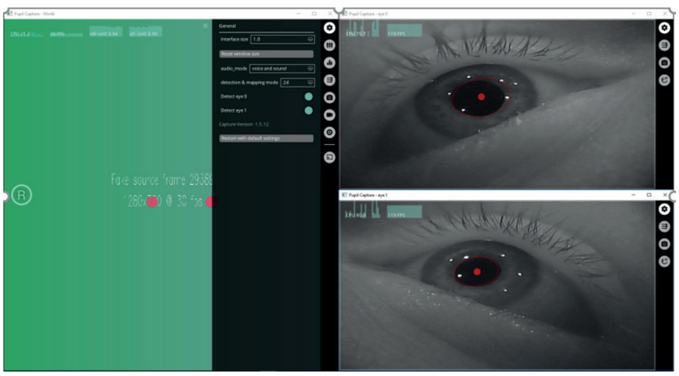


Figure 4. Pupil Capture feedback

Basic description of experiments:

- The market scene: This is the animated scene, where the participant is located on the market. The participant follows the hints from instructor where to look. These commands are forcing a participant to move his head and eyes in the space in a predictive way to the destination object. The trajectory of the head and eyes is evaluated.
- Cube tracking scenes: a) The participant is asked to track the cube, which moves horizontally or vertically. The starting positions are in the middle and on the corners. This process follows the hypothesis about the effect of the initial position on object tracking. b) The cube is also placed to arbitrary location and the participant is asked to search the cube in space. This experiment aims to check the vestibulo-oculus reflex in acquisition of visual targets. c) high speed cube movements – here the effect of the object's velocity on the prediction of the motion of the head is evaluated
- Video scenes: This experiment aims to track the movements in the natural environment. 360 degree videos with different topics and moving objects are viewed by the participant. The participant is asked to track the specified object. This allows to monitor coordination for example when riding a motorcycle or concentrating on interesting sites.
- Game Scene: In this scene enemies are randomly displayed in the different positions accompanied with the stereo sound and can indicating the arrival of enemy. There are 8 positions where enemies can be spawned. The researcher can adapt settings for enemies like speed, spawn time and attack damage or sound effects. The participant rotates his head and targets and eliminates the enemy.

Except for above mentioned dynamic scenes, also scene, suitable for static experiments based on static image observation (Figure 8) was realized.

The output from the position tracking subsystem is set of time traces containing the confidence levels, eye tracker data, and HTV Vive data on head rotation as well as its position. These data are in the evaluation systems compared against reference positions of the target objects in time. In the game and cube scenes, the reference object positions in time are generated automatically. In the video and market scenes is the reference information created manually along with the instructions that the instructor gives to the participant. The output data from the eyes and head movement monitoring system are processed in the evaluation system which is programmed in Python [21]. It uses the plotly [22] library as the chart creation tool.

Figure 5. Example screenshot of the Market scene

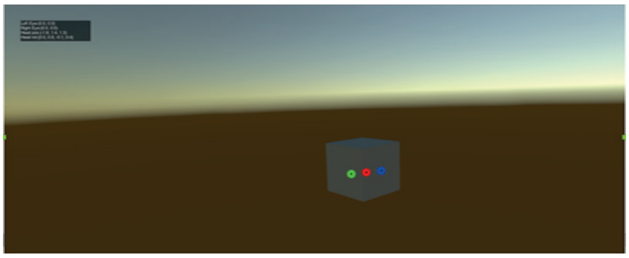


Figure 6. Example screenshot of the cube tracking scene – centered in middle of scene

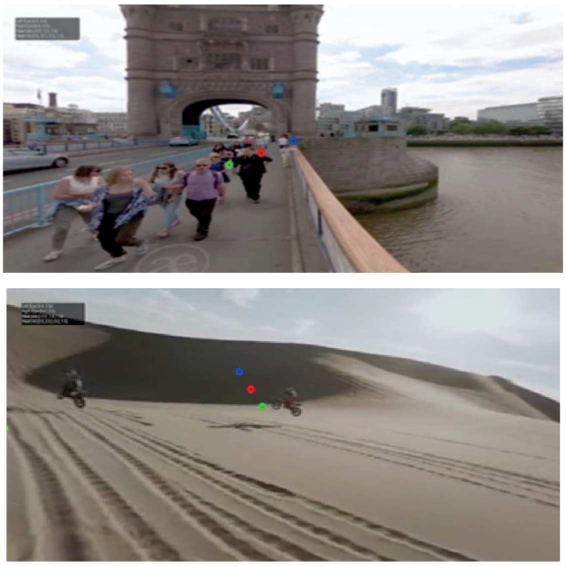


Figure 7. Example screenshots of the video scene

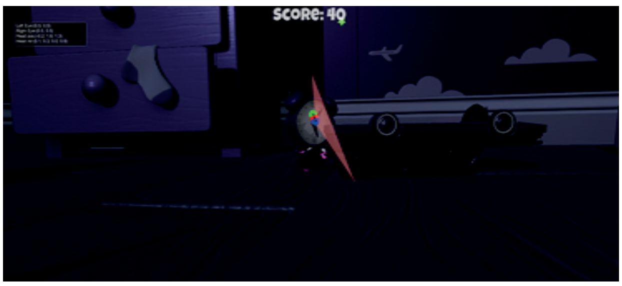


Figure 8. Example screenshot of the Game scene

Figure 9. Example screenshot of the realised variants for the experimental setup suitable for static image tracking. On the picture is depicted table from Rorschach test, which can be used for cognitive impairment tests [23]

V. SYSTEM EVALUATION AND RESULTS

The system was run in MM LAB and tested on multiple users. Calibration for the 2D detection algorithm was easy to perform against calibration for a 3D detection algorithm, whose tracking takes longer and is more difficult to follow circles. After calibration for 2D, the determination of the view was more accurate, and the deviations set during the observation were + - 2 degrees. It was sometimes necessary to perform the calibration more than once in case of large deviations. 3D calibration was sometimes performed many times, due to the large deviations in the optimal case, the results were + -3 degrees. The HMD-eyes plugin for Unity is still under development, so it's possible that the calibration will change, and the results will be more accurate.

The example results of the survey participants in graphical form in this section. Prior to plotting the data, the data was pre-processed as follows: in the first step were filtered out the data where the accuracy was less that one. In the second step, the heading position from the HMD was added to the view. In the survey, all participants noted the inaccuracy of viewing the scene. Under the initial conditions, the tracking error was of about 1 degree. However, with head movements, the position exhibited larger deviation. This may be related to insufficient calibration. HTC Vive Head Tracking showed the exact results. In addition, when tracking the cube horizontally to a greater distance, the eye tracker showed position above the tracked object. This error noticed also the participants. Therefore, the system needs to be improved either by improving the calibration or by improving the environment in which mapping is displayed. Cube tracking can also be improved with smaller distances, and we can improve the random view of the cube in space by changing positions where the cube is displayed, and thus covering a wider range of head and eye movement options. The results have shown similarity in the way, how the participants move in the same scenes. Research has shown that motion velocity in object tracking is important and creates motion variability. In addition, the initial eye position in the eyebrow also showed less variability in the result, but they could also be caused by the eye tracker. Example results are provided in the Figures 10 and 11.

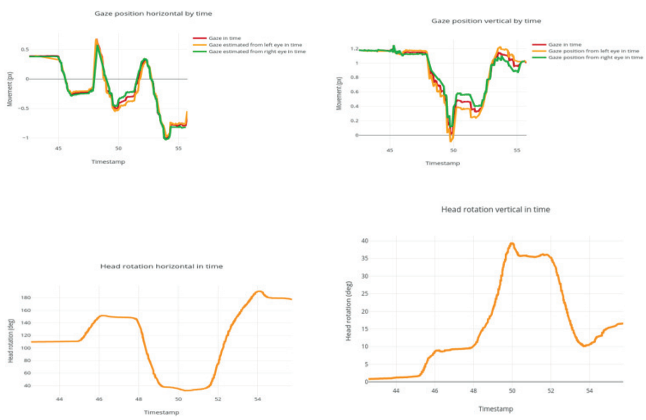


Figure 10. Examples of graphs for captured during time in experiments for gaze position in vertical and horizontal position, for head rotation is horizontal/vertical direction

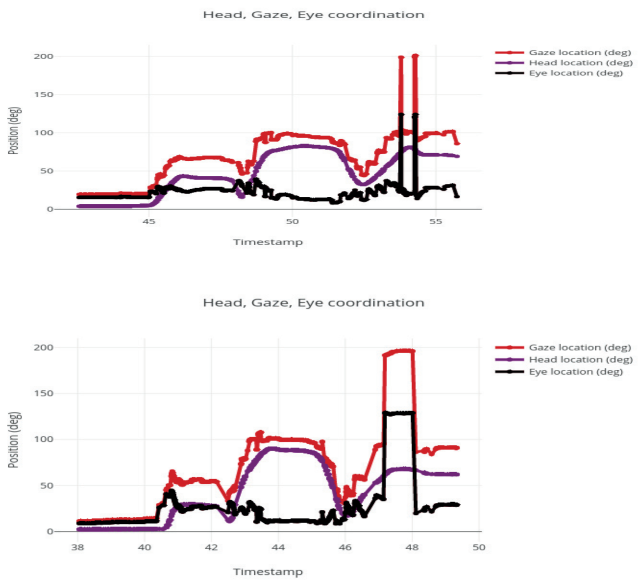


Figure 11. Examples of movement in the scene market, describing the movements to the left, right, left. The upper graphs are from the participant A, the lower ones from the participant B.

VI. Conclusion

The realised system capable to evaluate the head and eye movements based on the HTC Vive and Pupil labs eye tracker shows the great potential when coupled with the Unity platform. However as seen in the previous section, the unprecise measurements need further work to compensate the systematic errors and make the system usable for real applications.

Acknowledgment

This research was done with support of projects VEGA 1/0800/16 INOMET and H2020 NEWTON Ref. No.: 688503.

References

[1] Hardawar, Devindra. Tobii proves that eye tracking is VR's next killer feature. engadget. [Online] 01. 13 2018. https://www.engadget.com/2018/01/13/tobii-vr-eye-tracking/.

[2] Hayden, Scott. Google Shares New Research into Foveated Rendering Techniques. roadtovr. [Online] 7. Dec 2017. https://www.roadtovr.com/google-shares-new-research-foveated-rendering-techniques/

[3] Pro, Tobii. Does head movement affect eye tracking results? Tobii. [Online] https://www.tobiipro.com/learn-and-support/learn/eye-tracking-essentials/does-head-movements-affect-eye-tracking-results/.

[4] Freedman, Edward G. and Sparks, David L. Coordination of the eyes and head: movement kinematics. Experimental Brain Research. [Online] 131, 01. Mar 2000. https://doi.org/10.1007/s002219900296. 1432-1106.

[5] ČO JE VIRTUÁLNA REALITA ? virtualnarealita. [Online] https://virtualnarealita.eu/co-je-virtualna-realita/.

[6] Greenwald, Will. The Best VR (Virtual Reality) Headsets of 2018. PC Magazine. [Online] 05. 12 2017. https://www.pcmag.com/article/342537/the-best-virtual-reality-vr-headsets.

[7] Fuhl, Wolfgang and Tonsen, Marc and Bulling, Andreas and Kasneci, Enkelejda. Pupil detection for head-mounted eye tracking in the wild: an evaluation of the state of the art. Machine Vision and Applications. [Online] 27, 01. Nov 2016. https://doi.org/10.1007/s00138-016-0776-4. 1432-1769.

[8] Bryn Farnsworth, Ph.D. Top 12 Eye Tracking Hardware Companies (Ranked). imotion. [Online] 10. 01 2017. https://imotions.com/blog/top-eye-tracking-hardware-companies/.

[9] Tobii. [Online] https://www.tobii.com/.

[10] Types of eye movement. Tobii. [Online] https://www.tobiipro.com/learn-and-support/learn/eye-tracking-essentials/types-of-eye-movements/.

[11] C. Pallus, Adam and G. Freedman, Edward. Target position relative to the head is essential for predicting head movement during head-free gaze pursuit. Experimental brain research. [Online] 234, 01. Aug 2016. http://doi.org/10.1007/s00221-016-4612-x. 1432-1106.

[12] Barnes, G R. Vestibulo-ocular function during co-ordinated head and eye movements to acquire visual targets. The Journal of Physiology. [Online] 02 1979. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1281486/.

[13] Olah, Christopher. Understanding LSTM Networks. Github. [Online] 27. August 2015. https://colah.github.io/posts/2015-08-Understanding-LSTMs/.

[14] Lipton, Zachary Chase. A Critical Review of Recurrent Neural Networks for Sequence Learning. CoRR. [Online] abs/1506.00019, 2015. http://arxiv.org/abs/1506.00019.

[15] Team, Google Brain. TensorFlow. [Online] https://www.tensorflow.org/.

[16] Bulling, Moritz Kassner and William Patera and Andreas. Pupil: An Open Source Platform for Pervasive Eye Tracking and Mobile. CoRR. [Online] abs/1405.0006, 2014. http://arxiv.org/abs/1405.0006.

[17] VIVE. [Online] https://www.vive.com/eu/.

[18] 20. Building blocks for eye tracking in AR and VR. github. [Online] https://github.com/pupil-labs/hmd-eyes.

[19] Zeromq distributed messaging, [Online] http://zeromq.org/.

[20] Technologies, Unity. Unity. [Online] https://unity3d.com/.

[21] foundation, Python software. python language reference. [Online] https://www.python.org/.

[22] Plotly. [Online] https://plot.ly/feed/.

[23] Polec, J., Vargic, R., Csóka, F., Bieliková, M., Svrček, M., Móro, R., Smolejová, E., Heretik, A., Detection of schizophrenia spectrum disorders using saliency maps, ICT 2017, Moscow, Russia. September 20-22, 2017. Piscataway, IEEE, 2017, s. 398--402. ISBN 978-1-5386-0500-4