EYE GAZE TRACKING SYSTEM FOR ADAPTED HUMAN-COMPUTER INTERFACE

Rafael Santos, Arnaldo Abrantes, Pedro M. Jorge
Computer, Telecommunication and Electronics Engineer Department, Instituto Superior de Engenharia de Lisboa, Portugal
a37178@alunos.isel.pt aabrantes@deetc.isel.pt pjorge@deetc.isel.pt

Abstract — **This work describes the implementation of an eye gaze tracking system for a natural user interface, based only on non-intrusive devices, as Intel RealSense camera. Through image processing, the implemented system is able to convert face elements to the corresponding focus point on the computer screen. Tests show promising results.**

Keywords: Eye gaze; image processing; human-computer interaction; 3D camera; nose gaze; mobility

I. Introduction

The advance of technology has been pushing the emerge of new human-machine interfaces (HMI) [4]. The act of looking at a screen is part of most natural interaction processes. However the information that the eye gaze can give us is still not entirely exploited today. Gathering and processing user's eye gaze to interact with the machine is a topic already studied [3][4], but mostly based on specific technologies that are not available in mass market devices, such as, laptops or tablets.
Currently most devices are equipped with webcam, which can be used to collect information and provide feedback. However this technology isn't specific for this kind of image acquisition, missing sample rate and quality. Even if those requirements were met, some conditions would necessarily be put together like, ambient light. The algorithm and techniques used for that job would have a crucial impact. To avoid some troubles an camera with infra-red light and depth information was used. So this work describes a system to detect eye gaze based on Intel RealSense F200 camera, enabling a more natural form of human-machine interaction.

The implemented system also includes other face elements, like the nose, to replace eye movements.

II. STATE OF THE ART

Eye gaze is a natural form of interaction, and identifying where a person is looking allows a machine to interact in a more human way[6][1]. However, replicating this procedure automatically in terms of human-machine interaction is not simple. Eye gaze has been the subject of several studies over the past years [1][2]. Rayner and Pollatsek [7] were the first ones using electro oculography to measure retina movements while the user reads. In 2003, Duchowski [8] proposed a method, similar to Rayner's, where he used a metallic aro on contact lens, so he can measure electromagnetic field variations created by retina movement. Morimoto and Mimica [9], in 2004, released a study where they used several techniques for eye gaze, where the main purpose was to develop interactive applications, based on eye structure and corneal reflex. Neither techniques showed a good result, though the most reliable systems that produce the best results use specific and expensive equipment. New systems are being developed that use non-intrusive and more affordable devices. Among these systems, the best performances are obtained with a source of infrared light [3]. However, distrust of infrared light exposure motivated the development of eye gaze detection and tracking systems that use current technology, such as webcams [2]. These systems still have limitations associated with head movement compensation [2][3]. It is also necessary to improve real time processing algorithms and hardware [3]. On 2014's year end, Intel presented a new camera with depth, infrared and rgb information. The Intel Real Sense is meant to work on tablets, computer and even smartphones. With this technology, the work proposed here is meant to update the old one presented in 2013 by Rafael Santos, et al [5].

III. EYE GAZE DETECTION SYSTEM

This system's objective is to study the viability to implement a new interaction method between human and machines. That method would use facial metrics to do so. The system has some requirements like the use of commercial hardware, real time execution, enough precision to use it as daily, and easy and fast calibration method. The first one is the most important as the known system are expensive to the final user.

The implemented system can be described by the block diagram of figure 1.

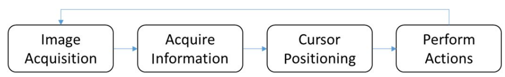

Figure 1- System block diagram

The system starts with image acquisition from Intel RealSense camera (Image Acquisition block).With the image the system go to Acquire Information block. Here the system processes the image and extracts required features from the user's face, such as, head orientation, eye, pupil and nose positions. This block is implemented in two different ways: as a common webcam, with traditional methods [5][3]; and Real Sense way.

For the common webcam method the results aren't very different of those read on bibliography. On figure 2 are presented two frames where is possible to achieve that pupil tracking with traditional methods aren't good enough.

1

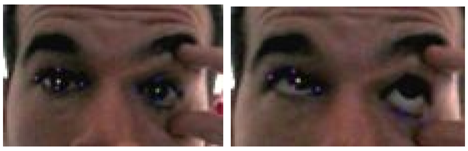

Figure 2 - Looking forward and looking up right.

With that, the time was invested on method two, Real Sense way. For that we used Real Sense SDK to acquire required features. For face pose, Euler angles were used. After face pose, the reference points were acquired. The system need points from both eyes and nose. Even with stopped eye the reference points have position oscillation. To reduce that noise a Kalman filter was used.

With all those features the system can process the information and execute actions. On Cursor Positioning block, the system uses the previous information to determine the focus of attention point and position the cursor on the screen. That could be treat as an action, but since it's the base of the system, it had a specific block.

Once the cursor is positioned small actions can be done. Human can interact with the machine (Perform Actions block). Eye blink to simulate left and right mouse click, was one of the two actions implemented.

IV. TESTING AND ANALYSIS

To evaluate system accuracy, several tests were performed. Being done some unitary testing the eye movement wasn't good, but the user could actually move the cursor horizontally as requested, but vertically, since the pupil movement is minimal, no movement was practically noticed. For eye blink action, the results were good, being the user able to perform those actions as requested(figure 3).

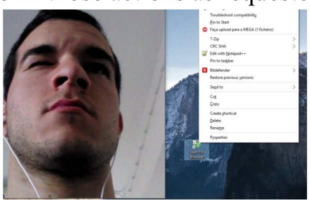

Figure 3 – Right eye blink to perform click over desktop icon.

Precision tests were done, with two screens, 24" and 15", with two resolutions, FullHD and HD. So the tests have some comparison a visual pattern is displayed on the screen, like calibration screen. The user should look to different points from this pattern and the necessary information is acquired. Tables 1 and 2 shows the results for eye and nose gaze, respectively.

Cursor positioning - Eye gaze

			Test points					
			center	bottom left	bottom right	top left	top right	Total
ERROR	FHD15	X	154	572	669	599	510	501
		Y	153	402	379	614	641	438
	FHD24	X	130	456	645	462	550	448
		Y	129	294	349	516	515	361
	HD15	X	27	479	399	307	343	311
		Y	62	262	279	344	351	260
	HD24	X	42	454	492	449	436	375
		Y	7	367	339	344	370	285
	Total	X	88	491	551	454	459	409
		Y	88	331	336	455	469	336

Table 1 - Cursor positioning error through eye gaze

From that table, eye gaze presents a mean error of 409 pixels for *X* coordinate and 336 for *Y* coordinate. The same test was performed for nose gaze and cursor positioning, the results are shown on table 2.

Cursor positioning - Nose gaze

			Test points					
			center	bottom left	bottom right	top left	top right	Total
ERROR	FHD15	X	54	366	23	3	115	112
		Y	31	21	25	0	140	43
	FHD24	X	22	5	16	0	66	22
		Y	25	16	17	9	9	15
	Total	X	25	123	13	1	60	45
		Y	19	12	14	3	50	20

Table 2 - Cursor positioning error through nose gaze

From that table, nose gaze represents a much smaller mean error. Being 45 for *X* coordinate and 20 for *Y* coordinate. Tests show a big variability in the results. Mean error is better with nose gaze than eye gaze, due to the information that can be acquired in the nose and in the eyes.

V. CONCLUSION

This project shows the possibility of a system that uses non-intrusive technology to develop a more natural human machine interface. The ambient light still being one of the main factors [5]. Other aspect is user's distance related with the camera. Since the eye has a small area in the face, the distance will influence the pupil detection as it'll have less resolution for the eye.

References

[1] Poole, A. e Ball, L. J. (2005). Eye Tracking in Human-Computer Interaction and Usability Research: Current Status and Future Prospects. In C. Ghaoui (Ed.) *Encyclopedia of Human-Computer Interaction*, Pennsylvania: Idea Group, Inc.

[2] Tunhua, B. B. W., Changle, L. S. Z. e Kunhui, L. (2010). Real-time Non-intrusive Eye Tracking for Human-Computer Interaction. *Proceedings of the 5th International Conference on Computer Science and Education (ICCSE).1092-1096.*

[3] Wild, D. J. (2012). Gaze tracking using a regular web camera.

[4] Drewes, H. (2010). Eye Gaze Tracking For Human Computer Interaction.

[5] Santos, R., Santos, N., Jorge, P. e Abrantes, A. (2013) Eye Tracking System With Common Webcam. Lisboa, ISEL. Elsevier.

[6] Jaimes, A. and Sebe, N.(2007). Multimodal human computer interaction: A survey. Computer Vision and Image Understanding, 1-2 (108):116-134.

[7] Rayner, K. and Pollatsek, A. (1989). The psychology of reading. Prentice Hall, NJ.

[8] Duchowski, A. T. (2003). Eye tracking methodology: Theory and practice. Springer-Verlag Ltd, London.

[9] Morimoto, C. H. and Mimica, M. R. (2004) Eye Gaze tracking techniques for interactive applications. Elsevier Inc.