Video Games For Visual Interface (Eye Tracking)

Sannara EK
UNIVERSITE GRENOBLE ALPES
GRENOBLE, FRANCE
Sannara.Ek@etu.univ-grenoble-alpes.fr

Supervised by: Didier SCHWAB.

I understand what plagiarism entails and I declare that this report is my own, original work.
Name, date and signature:

Ek Sannara , 12 / 06 / 2019

Abstract

Children with Multiple Disabilities, limited by their movements, are deprived of many forms of activities. This study focuses on two prime aspects that of studies done on GazePlay, An open multi-gaming platform mainly to interact with the human eye (gazes), founded by Didier Schwab [**schwab:hal-01804271**]. The principal purpose is to develop a novel manner to entertain these multi-disabled children's by developing and improving specifically catered software and games. The goals of these games are to enhance the 'children's cognitive ability and allow them to familiarise interactions with computers better. The second purpose is to understanding, integrate and utilize eye tracking metrics. Notably, the Area of Interest (AOI) and its extensions, to better understand the semantics of the extracted data obtained from Human-Computer Interactions through gazes.

1 Introduction

Having multiple disabilities forbid an individual to do as one wishes and in our particular case, we will mainly dig into the topic of interaction and concentrate on a subgroup of multiple disabled people, the children (individuals age 20 or below). With the coming and the improvement of gaze technologies and its use as an augmentative and alternative communication (AAC), efficient interactivity with computers without the need of physical motions is obtainable [**inproceedings**, **Qvarfordt244543**].

The AOI is crucial in visualizing GazePlay's, or other eye tracking application, impact towards the users of the platform to profoundly investigate the user's perception and cognitive skills. Question on knowing if people are genuinely noticing contents that are intended to be and if not, why and what the disturbance is shall obtain an answer. These vital are clues to retrieve the semantics of the user's behaviour and in this study, knowledge.

This study aims to empower these children with disabilities, through the GazePlay platform, by striving to embed the technology as a natural means of communication through entertainment and to understand them better.

2 Backgrounds

In order to establish a footing in this study, it is vital to elaborate on essential meanings and phenomenons. It is essential to clarify multi-disabled individual, how games are beneficial, how gazes detection works as well as the principles of eye tracking metrics and its applications.

2.1 Multi-disabled Individuals

disability is describable as "any restriction or lack (resulting from an impairment) of ability to perform an activity in the manner or within the range considered normal for a human being" [**Jones377**]. The Rett Syndrome (RTT), a postnatal neurological disorder mainly found amongst girls that directly pose movement challenges to the individual [**Lariviere**], was particularly noted for this study. Widely proven, using modern eye-gaze equipment are of the best means of communication with those affected with the syndrome [**inbook**].

2.2 The Benefits of Games

Rather than the apparent outcome of pleasure when indulging in games, there are lists of varieties of development and stimulus to parts of human skills in Detections, Memorizations, Attentions & Cognition and Motor abilities based on the type of game [**schwab:hal-01804271**, **Feng**]. Games are a means of ways of helping the subjects of the study to integrate themselves into a healthy regular life.

2.3 Eye Tracking and Gaze as an Interface

This study revolves around the use of 2 different commercial eye tracking products, TheEyeTribe device and Tobii's 4C model. The data input of the eye's movement is obtainable from a non-intrusive eye tracking system which utilizes Pupil Centre Corneal Reflection (PCCR) and real-time video captures. Through the usage of sets of principles and methodologies, the majority of meaningful data is extracted

[**Anuradha**]. Using Gaze Point, an assistive eye tracking software provided by Tobii dynavox, allows the substitution of the mouse cursors with gazes for the GazePlay platform.

This study stresses the use of equipment that is widely available and affordable to the public. The Tobii's 4c is acquirable through any major market for an estimated price of 170 euros.

The movement of the eye is definable with several sets of behaviours, namely such as those stated below [**schwab:hal-01804271**, **Hedberg**]:

- Pursuit: defines the following behaviour of the eyes on a specific context.
- Saccades: defines the small abrupt movement in the eyes.
- Fixation: defines when a specific area observed upon over an extended time.

2.4 Eye Tracking Metrics

Utilizing obtained eye-movement data, eye-tracking tools & metrics such as AOIs, Heatmaps, Fixation Sequences, Time to First Fixations (TTFF) and Revisits can be defined. The AOI, the focus of topic, is a tool which defines an area calculated from multiple sets of neighbouring fixations which contains meaningful or attractive semantics towards the perceiver of an image or video [**Hedberg**]. This tool, when applied for novices, in any context, is well-suited to assist in identifying mistakes or inefficiencies that may occur based on their observations. While applied to experts, they are to highlight eye-movement paths of corrections that can significantly elevate learning processes for novices. A similar study on this matter, although on the field of surgical education shows excellent results in the application of the AOIs [**articleHealth**].

The establishment of these metrics allows the event to create a visual attention model of the subjects of study, allowing understanding of user's perceptions and visual behaviours [**drusch:hal-01223743**]. This implementation brings fourth methods for an in-depth investigation into the understanding the user's perception, unravelling the ambiguity of user's movement semantics.

With this information at hand, noise can be defined, sidetracking intrusive contents during the gameplay, and see what catches the attention of the users. What actions they take, their reasoning and most importantly, performance measurement which reflects their cognitive abilities and level.

3 Game

3.1 Ribbit Race

The first game of this study was Ribbit Race. The player takes control of a frog against three others to partake in a race. The player's frog moves toward the finishing line after a Biboule, Gaze 'Play's mascots that scatter over the screen, are hit by the 'user's gaze. This action is to imitate the act of feeding the frogs flies, giving it incentives to advance. The level of difficulty progresses after every round of races; more flies needed to move forward.

The implemented game, classified as a competitive shooting game, centres on the stimulation of the player's cognitive ability on three main domains:

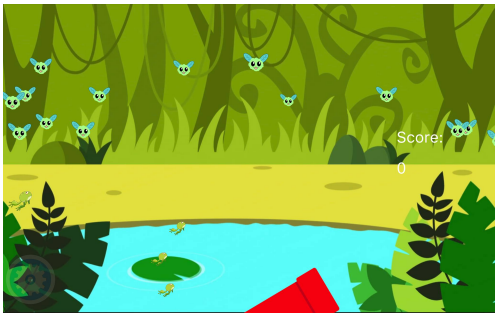

Figure 1: A Scene of Ribbit Race

- Speed (reaction time), by collecting biboules as quickly as possible to achieve first place in the race
- Judgements, due to necessary real-time decision making on which particular zone of biboules to capture, as some may be clustered together, to obtain oi better-capturing efficiency.
- Attention capacity, for the need of the user to focus on several sections of the screen in a single frame, ever analyzing for the creation of new biboules.

4 Methodology

The establishment of AOIs in real-time as the user of the platform interact with the game has posed many performance issues. Thus, in order to deal with this problem, post-processing methods after each game were employed. For a game instance, the data of eye movements are saved, and the gameplay shall be video recorded to make a playback. This playback is shown together with all the AOIs that were defined from the post calculations.

4.1 General ideas and Approaches

AOIs, generally, are applied on stationary images. This study, differently, centres on ever-changing scenes, videos. An AOI defined in one instance of time will most likely not represent the same area as it previously. Hedberg's method for defining the AOI [**Hedberg**] must be changed, adapted and improved for the GazePlay platform.

There are two main theories of how visual attention formulate [**Hedberg**, **drusch:hal-01223743**]. The Top-Down Control which bases that perceptive content mapping of attractiveness is by taking into account the contexts which hold relevant information inside an image. The Bottom-Up Control, on the other hand, does not rely on the context to establish what attracts people but rather the raw eye movement of the perceiver. The bottom-up approach to realize the AOI shall be the focus of this study in order to make a universal AOI builder that can be utilized and applied for all 48 games on GazePlay and other domains.

The development of AOIs is primarily with gaze-fixations. Establishing the fixation is the initial step and involves a two stages process. Data given by the eye-trackers API are collected labelled. The data must undergo filtering of noises, such as saccades and movements, to identify the points of

fixation. AOIs are obtainable through a series of calculations of for clusters of fixation points.

4.2 Data Input & Filtering

For this study, we refer to coordinates of the eyes on the screen provided by the eye tracker as a point.

Each point holds their x,y coordinates and the time of their occurrence. The movement from points to points is initially made as a directed graph although later on in the project, for performance gain, we have decided not to display the edges while only showing the vertices.

Establishing Fixations Points

Fixation occurs during relative stability over a small area. Fixation of a users instance obtainable by taking into account the time, and spatial properties of all the point.

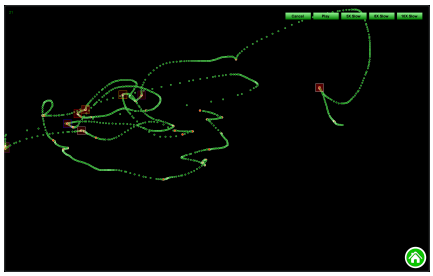

Figure 2: Visualization of eye movements with the Tobii 4c

The temporal intervals, the time between the appearance of a point and its following point, is used for separating many points clustered together. Through a series of analysis with on TheEyeTribe equipment, the minimum interval duration for fixations was at ten milliseconds before another fixation occurs was determined.

The speed is, at times taken into consideration depending on the specifications of the eye-tracking device. Particularly on TheEyeTribe eye-tracker, where eye movement data exhibits a warping effect due low-frequency rate of eye movement capturing.

The speed, or velocity, is obtained by having the Euclidean distance divided by the time interval between 2 points. Groups of points that belong to a fixation point results in relatively slow velocity.

$$\vec{v} = \frac{\sqrt[2]{(x_2-x_1)^2+(y_2-y_1)^2}}{(t_2-t_1)}$$

Findings on the behaviour and characters of many fixations have allowed us to deduct one critical property. When an individual has their fixation on an object, the eye movement will tend to linger in that particular area. Points are clustered together if their Euclidean distance falls below 10.

Approaching the phenomenon that a point tends to come to revisit the previous location, the criteria to add the next point into the fixation cluster is extended. The extension is to not only check the lastest point in the cluster but the distance between all of points and the new point.

When no more points are discoverable, which matches the criteria, the center point of all the previous points that were in a group together undergoes calculation. This center point shall be the fixation point.

4.3 Defining the Area Of Interest

The Areas Of Interest is determined primarily through clusters of fixations. Its implementations resemble the process for establishing the fixation points.

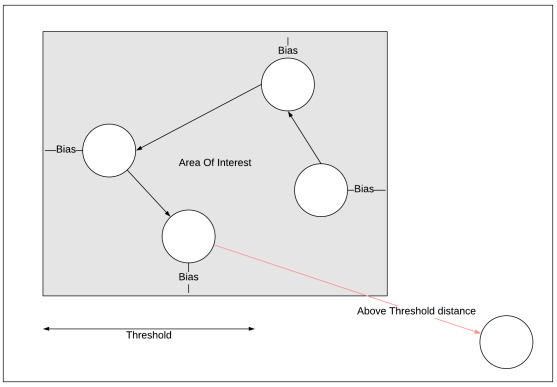

Figure 3: Visualization of AOI calculations based on fixation points, context hidden

The spatial difference between each fixation point is taken into calculation through a defined Euclidean distance. Points that are under a specific threshold undergo clusterization. Tests and experiments have found that using a threshold of 150 is highly efficient in determining the areas of representations of AOIs. Due to their usages in establishing polygons, AOIs can only be built when having three or more fixations point.

Area of representations

Most often, the AOIs are represented as a rectangle for their simplicity. Among the cluster of fixation points, identify the local maximum and minimum x and y coordinates to build the edges of the rectangles. A bias is added or subtracted to every edge based on their direction from the AOIs centre point to increase the boundary size of the rectangle to increase visibility.

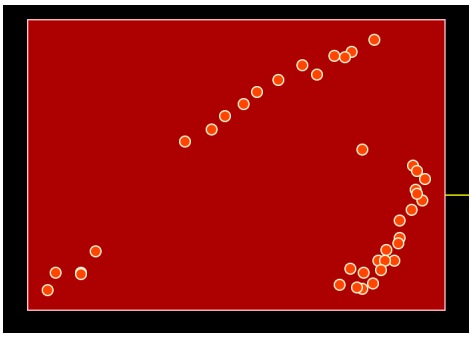

Figure 4: The area of representation as a rectangle

Convex Hulls

A more accurate way of representing the AOIs are through the use of convex hull, although they require additional computational resources. In setups where clustered fixations points

are very close to one another, there may pose visibility problems due to their smallness in size. The convex hulls are calculated with the Jarvis march algorithm which searches for the far-most fixation point of every direction from the centre of the cluster [**Jarvis**].

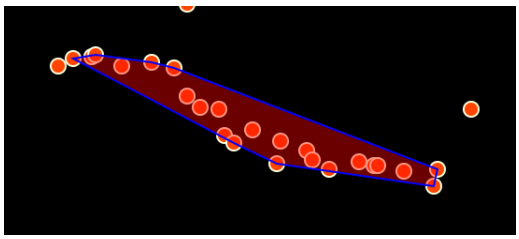

Figure 5: The area of representation as a convex hull

4.4 Concatenation of AOI

The defined AOIs only takes into account the areas in a single instance of time. It does not explicitly provide details on areas of interest that are represented throughout the entire lifetime of the game.

Thus a general summary of AOIs which represents the most attractive zones, rather than content, throughout the entire instance of a game is built.

These concatenated AOIs can further expand important details where there are little motion or changes in scenes.

Methodology

The concatenation is achievable through a recursive method where each area is checked and compared with all other areas if there is an over-lapse above a defined ratio threshold in the area of representation of another. Matched areas shall be combined and be removed from the list of areas to check while forming a new area. The newly combined area must undergo the recursion process again to check for any areas that it may over-lapse. The recursion ends when overlapping areas can no longer be detectable. The concatenation process does not necessarily limit to the combination of 2 areas but in fact, may consist of an indefinite amount so long as the criteria match. We have determined through multiple tests that a threshold of 70 percent achieves the desired result. In our studies, areas that have are the result of concatenations are represented by a yellow shade.

4.5 Parameters

The AOI on its own may not necessarily be considered as an eye metric but is essential on the extension of other metrics. Out of the total five that are extendable on the AOI, 4 of which has been implemented in this study [**imotion**]. Below states the process in establishing these parameters and their details.

- Time To First Fixation: The time from the beginning of the instance to when the area's first fixation occurs.
- Time Spent: The time regarding the time interval between the first and last fixation which belongs to the area.
- Fixations: The number of fixation points which made up the area.
- Revisit: The number of times a user comes back to the same area. This parameter is only achievable through the concatenation of AOIs. In our study, the revisit is obtainable by the number of areas that were used for the concatenations.
- Ratio: The number of times different users has fired up the same area. This parameter is yet to be ready for implementation in this study as the history of the user interactive record is not compared with current instances.

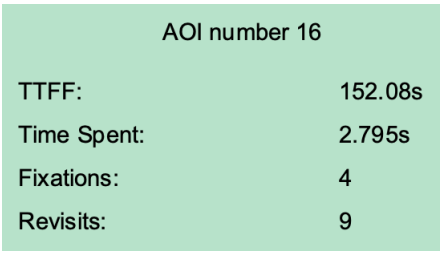

Figure 6: The parameter table of a concatenated AOI

Accessing information regarding the parameters were strived to perform as intuitive as possible. The parameter's table placements are next to their corresponding AOI during a playback instance of a game. The respective parameters of concatenated areas are accessible by hovering or gazing over them in the AOI playback screen.

4.6 Priorities of AOI

The AOIs are not equal in terms of the interest amounts it holds towards the perceiver. Some are more important than others, some less. We can determine these difference through the utilization of one of the parameters we have defined, the time spent. The notion that AOIs that has a longer fixation duration may hold higher interest contents than others is confirmed. This applies to the opposite; less fixation time infers less importance. Thus, AOI priorities are made identifiable by different shades of the colour red. A darker tone of red would symbolize greater importance while lighter tones would depict less importance.

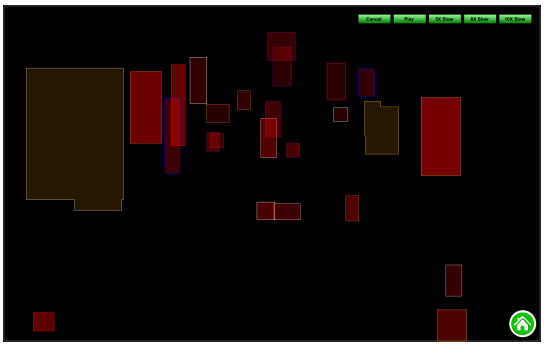

Figure 7: The general AOIs, without the context, with difference in shades of red, Yellows are concatenated AOI

4.7 Visualization of the AOI

Recordings

Video recording of the game instance is saved and shown alongside the playback of the AOI. This feature, novel to the GazePlay platform, held various constraints due to the lack of support in video recordings on JavaFX's native libraries. This problem was approached by integrating Monte media library with Techsmith video codec[**monte**].

Playbacks

Upon the exit of any game on GazePlay, users of the platform will have an option to go to a new window to review the scientific elaboration of their gaming instance. This window initially displays all the AOIs and concatenated AOIs that were built. To watch the corresponding run-time AOI playbacks, users of the platform may choose the AOIs recording in a standard speed rate or a 3, 5 or 10 time slower rate with or without the video recording of the game instance.

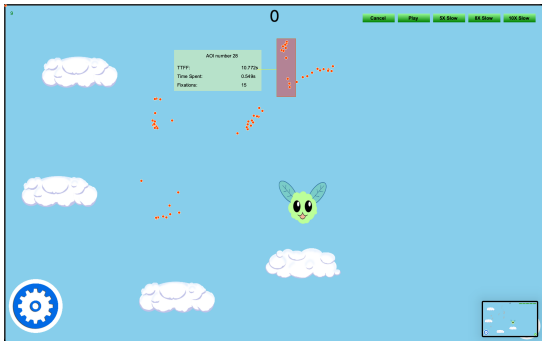

Figure 8: A scene during the playback for bibouille Jump

When played, every coordinate of the user's eye movement, treated and untreated, are shown at their respective time of occurrence for 2 seconds. They are then is removed after the stated period to produce a moving-trailing like the effect of the eyes for visual purposes and to allow a sense of order to contrast the newest to the oldest. Fixations are highlighted as orange dots while as other eye movements are smaller and green.

Performance Constraint

Challenges were encountered when it came to the playback of the eye movements due to its entanglement with the matter of short periods. The intervals of displaying coordinate points one after another were as fast as 0.1 milliseconds. The computing equipment used for these studies were unable to achieve the same time, causing a variation in the playback sequence. The eye movements were not in synchronization with the video recording playback. This problem was tackled by changing the playback rate of the video to match the playback of eye-movements. Further improvements are needed in this area.

5 Utilization of AOIs

As stated, One of the platforms and its contents objective is to improve the cognitive ability of their users. Although, the fact that there is an improvement is instead a matter of deduction that is based on by other scientific studies. Clear visualization of this improvement is not observable. It is abstract and cannot be easily inferable.

This problem is approached in the study by the extension of what we have accomplished with the tools. We defined a new type of AOI, a "Targeted AOI". This new area tool can be defined by the content developer who knows the optimal path. This area represents the optimal area of which any individual should behave to obtain the best results or efficiency.

Another method in setting the targeted AOIs is to save the general AOIs of players who obtains the highest score. The person's route and path of observation are set as the standards for target AOIs.

It is possible to conduct an assessment of performances of the perceiver through a comparison of the AOIs and Targeted AOIs. Based on the similarities of properties between the two, a score is generated. This score, hidden to typical users, may be used to represent and measure the cognitive ability level to visualize traces of changes and learning.

6 Conclusion and Perspective

What we have accomplished in this study, the addition of a new game and novel tools on Gaze play gives further facilitation to conduct qualitative and quantitative research on multiply disabled children, or general users, who enjoy the games provided. Through the elaboration of the eye-movement data, our work done is highly informative to a various degree of researches in physiological, perceptual studies or more.

The application of AOIs to establish a user performance score is a novel implication. More studies are necessary to improve the accuracy of this methodology to bring light to a different and new eye-tracking metric.

Further beyond this, the application of the Top-Down Control method which has yet to be implemented, due to its specific need of catering for all 48 game on the GazePlay platform, may result in AOIs that are possibly more accurate and yield more considerable information.

The benefits of this project studies are not limited to only multi-disabled children. It is in-fact applicable to a variety of usages such as programming problem-solving definitions and many other domains [**Obaidellah:2019:CSS:3314111.3319825**]. We have established, studied, developed and integrated the AOI to the GazePlay platform. Since what is built is not necessarily content dependent, the Bottom-Up Control approach method is chosen for the study. This choice allows the findings of this research to be extendable as an external library on its own. A standalone library, universally integrate-able on Java applications, solely to understand the ambiguity of individuals gazes through the utilization of AOIs.

Finalizing this report, It is worthy of reiterating that this study strives as an objective to help mitigate the burden of unfortunate children from their disabilities.