Mobile HCI Coursework - A look into Head Tracking for selection interactions

NIK HARITH SHARIFUDDIN BIN MOHD SUHAIMI, 2673209B

1 INTRODUCTION

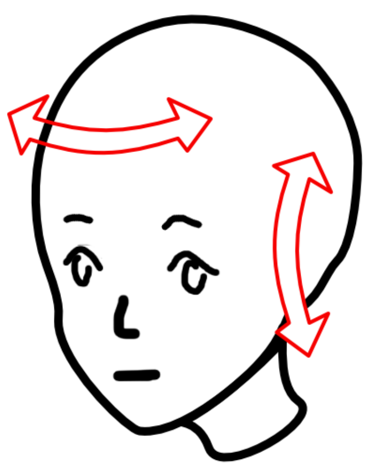
In recent years, we have seen the rise of Social AR, which are small augmented reality experiences that are available on social media platform such as Instagram, Snapchat and TikTok. These mobile-based AR experiences, or more fondly known as "filters" or "effects" by social media users introduced many novel ways of interaction enabled by the multiple ways of tracking that are available on those platforms, namely plane tracking, head tracking, hand tracking and even eye tracking. In this coursework however, I am choosing to study the interaction that utilizes the head tracking.

2 DESIGN

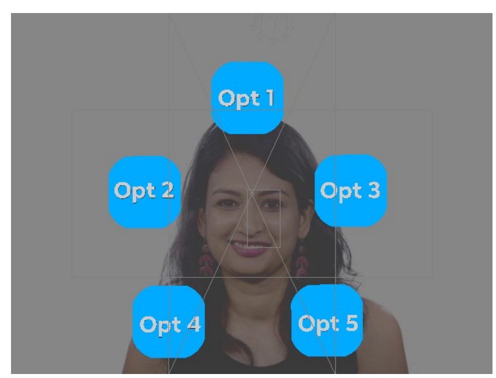
2.1 Main design

The head tracker controls for this interaction allows for 3-DoF (degrees of freedom) controls as seen in what has been implemented in [5], however for for this study I have opted to only utilize the rotation of the X and Y axes only where the tracking will only correspond to the user's movement of tilting their head up and down, and the action of "facing" left and right.

The interaction is designed so that the user will be able to choose and activate one out of five buttons, one button at a time. The buttons are placed encircling the user's face, forming a 5-point circle around the user's face. The buttons are placed and is 'parented' relative to the head tracker, so when the buttons will always follow the user's head.



(a) Illustration of how the head rotation used.



(b) Proposed arrangement of the buttons layout

2.2 Rationale of the chosen interaction and design

This type of interaction was chosen because apart from the head tracking being one of the easiest interaction to be implemented, it is also a 'hands-free' interaction where the users are not required to use their hands to operate the controls and choose the buttons. Traditional 2D flat screen interactions has long been a norm since the debut of modern smartphones, however as more and more Augmented and Mixed Reality technologies are introduced, interactions that wouldn't require the typical flat screen device should be explored as a step to further improve these kind of interactions. Wang et al. made a comprehensive research [4] on how head tracking can be used in video games.

In the context of social AR, it has become commonplace to use head tracking controls for these AR "effects" to control and use the effects. Among popular use cases are games, drawing, and as an anchor for a makeup filter or attaching objects to the head. While these use cases are mostly "for fun", it is undoubted that there were considerations and decisions for the interaction design when making these effects. It has also been seen to be used on exhibitions powered with augmented reality where there will be a big screen projecting the experience while the tracking is being done with a mobile device such as a smartphone or tablet.

2.3 Recognized weaknesses

However it is also recognized that there are certain weaknesses of this interaction as we will discuss it further in the evaluation section. Among the weaknesses that has been revealed is that sometimes it's not clear and obvious how to use this interaction, the accuracy of the head tracking can be swayed by some factors and determining the threshold for the controls can be tricky.

3 IMPLEMENTATION

3.1 Software & tools used

The prototype of the implementation was built within Meta's Spark AR software [1], which is a software to make and publish AR "effects" on Meta's platforms (Facebook, Instagram and Messenger). The logic of the project was easily configured by utilizing the software's easy to use "patches", which is a way to code using visual scripting.

The prototype was built step-by-step from a simple implementation where only 2 buttons were used to the latest prototype which implements the choosing for one of 5 buttons. This was done to incrementally test the logic for the head tracking (from simple to more complex) and spot errors early in those implementation.

3.2 Interaction flow and the use of thresholds

Upon opening the prototype, the button will automatically arrange themselves around the user's face, forming a star-like pattern for the 5 buttons. The user can then freely move their head around and select any buttons they wish. When a certain threshold is reached, it will trigger the button selection where the button will be slightly enlarged and changes it colour to be slightly darker as a form of feedback to signify that the user has successfully selected the button. A short sound effect will also be played upon the trigger as another form of feedback.

For the logic of selecting the buttons, I used the predefined triggers of the head rotation provided within the software itself. These triggers consisted of 6 different predefined threshold that sends a 'True' boolean when the user rotates their head and consisted of:

(1) Left Turn State

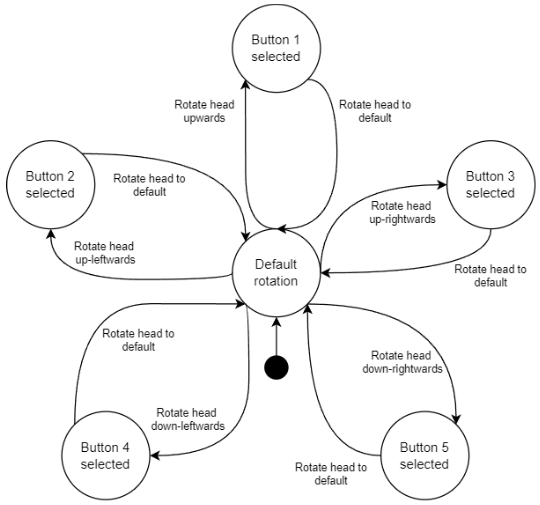


Fig. 2. A finite-state machine diagram representing the flow of the interaction.

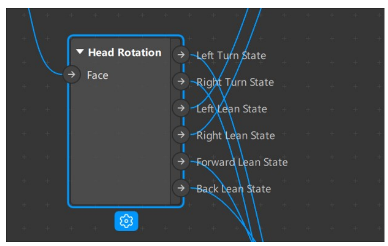


Fig. 3. The predefined threshold triggers within Spark AR.

(2) Right Turn State
(3) Left Lean State
(4) Right Lean State
(5) Forward Lean State
(6) Back Lean State

Using these triggers, I created a logic and mapped the triggers according to the position of the buttons around the user's head. I had to combine two states for some of the buttons that are positioned diagonally as they require the combination for the thresholds to be triggered. This was simply done by using the 'AND' logic to combine the two states.

3.3 Challenges and issues

One issue when implementing the prototype was the requirement to use 5 buttons. While it is a good amount to be evaluated for novel interactions, it affected how I needed to design the layout and placement of the 5 buttons. It also complicated the structuring of the code logic as the threshold detection can be a little bit finicky when two states are required. It would be easier to design if there were only 4 buttons as it would require only the four primary directions of rotating the head (Up, Down, Left, Right) and consequently would make the interaction easier to be figured out and used.

While it is possible to include more than 5 buttons, the placement of the buttons is limited due to the constraint of display size, which this prototype is implemented on mobile devices where they have vertical screens as opposed to a wider screen available on other devices such as the normal desktop computer. On top of that, an increase of the number of options also increases the user's mental load to process on how to use the prototype.

3.4 Testing and deployment

The prototype was tested live within the software itself by connecting to a webcam, or it can be tested directly on a device that has Facebook or Instagram installed, often providing better performance due to it being optimized to run on a mobile device.

After being satisfied with the tests, the prototype is then published to Instagram and Facebook platforms through a link-only access to the "effect/filter". Then, the prototype is ready to be used by sharing the generated link and it will automatically open either the Instagram or Facebook application depending on the link shared.

4 EVALUATION

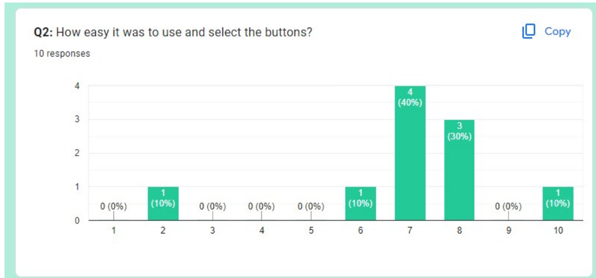
4.1 Methodology

To evaluate my implementation of the head tracking controls, I opted for user evaluation to gather feedback data from real users directly. This is done through user testing where the user will try out and experiment with the prototype and then followed by a short survey asking their thoughts about the prototype that they have just used. This method is the easiest way to gather insights on how users think about this interaction method as well as their perceived decisions when they tried it out. It is also good to gather data for rapid changes for an interface as found in [3].

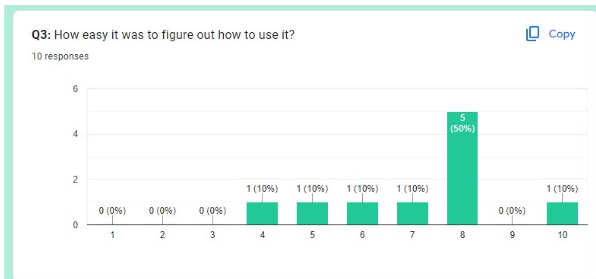
The user evaluation starts by introducing the participants to the prototype. The participants were given a task to activate the buttons using the head tracking controls. Then, the participants were instructed to play around and explore the prototype on their own without any guidance. This is to measure how easy it was to figure out how to use the head tracking controls by themselves. After completing the task, the participants were asked to fill in a short survey consisting of 10 questions related to the head tracking controls prototype that they have just used. They were asked about the ease of use, the level of difficulty on figuring out the controls, the effectiveness of the feedback forms upon activating the buttons, and their comfort when using the head tracking controls.

4.2 Findings from evaluation

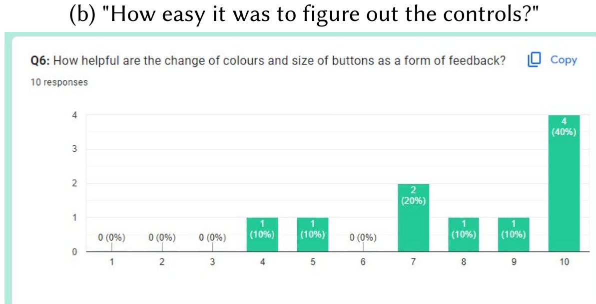
Based on the evaluations with a sample of 10 participants, 80% of them reported that the head tracking controls were easy to use with scores over 7 out of 10 were given. It is also reported that the controls were easy to be figured out, with 50% of them scoring it 8 out of 10 in terms of easiness to figure it out. The participants' early impressions on the novel interaction method was quite positive, with mostly saying it is 'unique', 'innovative' and 'interesting'. When asked why was it easy to figure out the controls, some participants reported that the interface is intuitive enough to figure out the controls while another participant reported that they have previous experience by using Instagram filters. On the contrary, some participants reported difficulties using the head tracking controls as they were struggling to make their head rotations activate the buttons.



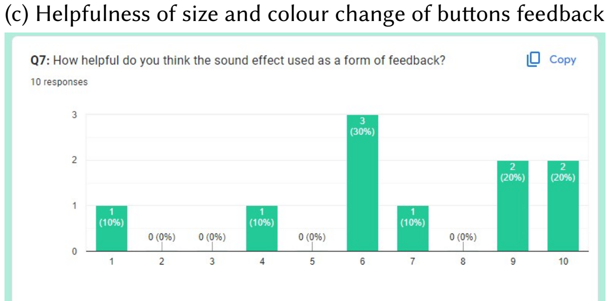
(a) "How easy it was to select the buttons?"



(b) "How easy it was to figure out the controls?"



(c) Helpfulness of size and colour change of buttons feedback



(d) Helpfulness of audio feedback

Fig. 4. Results from the survey

When asked about how the feedback of the button changing colours and size when activated, it scored an average of 8 out of 10 for being helpful in the prototype, however the audio feedback only scores an average of 6.8 out of 10 for being helpful. It is observed that the changing of colours and size of the buttons are more helpful presumably due to it being obviously displayed when a button is activated. Meanwhile, the audio feedback is less obvious and some participants also reported that they did not notice there were any sounds when the buttons are activated.

Last but not least is how comfortable the participants were when using the head tracking controls. Results showed that the comfort level only scored at an average of 5.8 points out of 10. Feedback from participants finds that there were some confusions on how to rotate their head to select the buttons. Some interpreted the rotation as 'tilting' their head, which correspond to rotate their head on the Z-axis, which is not used in the implementation of this prototype. It was also found that with longer or excessive use of the head tracking controls, it can introduce tiredness and strain to the participants. A number of participants also reported that the threshold for the head tracking seemed too high for them to activate the buttons

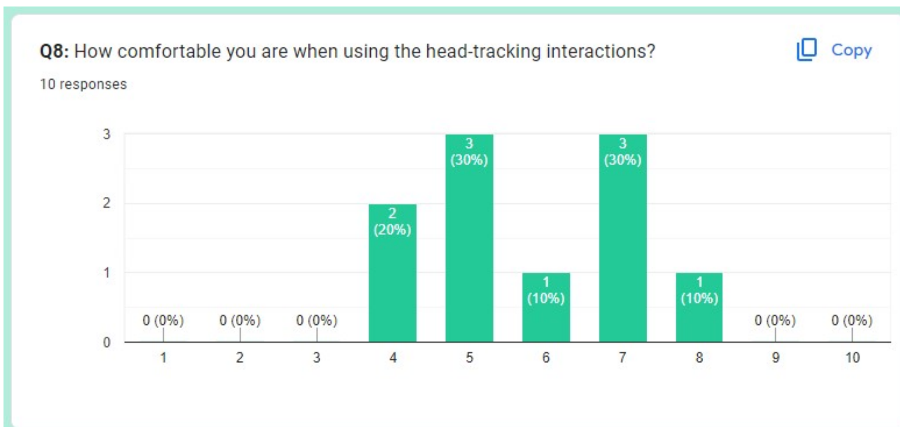


Fig. 5. Findings shows that comfort levels are average when using the head tracking controls.

4.3 Future refinements and alternative designs

As observed in the findings above, we have seen that one of the common issues that users encounter when using the prototype was the confusion on figuring out the controls upon launching it. In this experiment, this was done by design and is intentional in order to study how intuitive the interaction method is without any sort of guide and whether the users had any previous exposure to the head tracking controls.

In an ideal implementation, the prototype would be refined and added a 'hint' where the user will be shown a quick animation showing how to move their head. Another improvement that can be made is to add a pointer or cursor to show the user where their head is currently pointing. Both of these improvements can reduce the time the user need to figure out the controls and use it right away.

In this implementation, the head tracking is using the predefined thresholds that has been provided by the software. While this has proven to be easier to set it up, the more complex the layout and interactions are, the less flexible it gets to utilize those predefined thresholds. Alternatively, we can instead use the raw head rotation data and specify the thresholds ourselves and this can give us more flexibility in terms of the thresholds and the ability to specify custom points where the user need to rotate their head compared to the 6 predefined triggers.

Continuing on the same track of idea, instead of designing the implementation mainly relying on head rotation thresholds to select the button, an alternative design would be to use a combination of the raw head rotation data and use a pointer/cursor where it will act more likely like how a mouse/cursor do in a traditional desktop computer. This would result in more fluid interaction as the cursor will follow the head rotation in real time, providing instant visual feedback to the users, thus avoiding some confusion. This has been proven in [2] where they have evaluated and found that by using the 'pointing by head' (which is the same as the proposed alternative), it scored second only to traditional mouse pointer controls when compared by activation time.

5 CONCLUSION

In this study, we introduce a novel interaction method that utilizes head tracking through a mobile device's camera and explores the challenges to implement it as well as findings from users who tried out the implementation.

We observed that head tracking controls are used in only specific areas at the moment due to the way it is being widely implemented, which are mostly on social medias and sometimes augmented reality exhibitions. Even though it is not widely used, through this study it has been proven that the interaction method is still intuitive and easy to use. When used with appropriate guides and feedback, it can create a good user experience when using it.

While the findings of this study revealed helpful data to evaluate the head tracking interactions, the sample size of the user evaluation was only 10 people. It is not significant enough conclude the findings of the evaluations. This study also lack some literature reviews that would have given more insights and correlation with previous studies that may have explored this scenario.

This study has found that this novel head tracking controls interaction is viable for appropriate interactions in certain cases. Further studies could discover around the areas of best practices and general guidelines on how to effectively implement this type of interaction. It could also be supplemented with proven use cases to further support the viability of using this interaction method.

In summary, this novel interaction method of head tracking has its use cases and could be utilized with a proper design and implementation of it.

REFERENCES

[1] Meta, 2024. Spark AR Studio - Platform to build AR contents by Meta.

[2] Minakata, K., Hansen, J. P., MacKenzie, I. S., Bækgaard, P., and Rajanna, V. Pointing by gaze, head, and foot in a head-mounted display. In *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications* (New York, NY, USA, 2019), ETRA '19, Association for Computing Machinery.

[3] Savage, P. User interface evaluation in an iterative design process: a comparison of three techniques. In *Conference Companion on Human Factors in Computing Systems* (New York, NY, USA, 1996), CHI '96, Association for Computing Machinery, p. 307–308.

[4] Shuo Wang, Xiaocao Xiong, Y. X. C. W. W. Z. X. D. D. Z. Face tracking as an augmented input in video games: Enhancing presence, role-playing and control. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2006).

[5] Terziman, L., Marchal, M., Emily, M., Multon, F., Arnaldi, B., and Lécuyer, A. Shake-your-head: revisiting walking-in-place for desktop virtual reality. In *Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology* (New York, NY, USA, 2010), VRST '10, Association for Computing Machinery, p. 27–34.