AI Automation for Sentiment Analysis in Virtual Reality Games

Erdem Murat
Department of Computer Science
George Mason University
Fairfax, VA
emurat@gmu.edu

***Abstract*—Virtual reality (VR) games are an emerging category in the consumer market and game research. Like traditional flat-screen games, there is a plethora of research opportunities that can be explored. For example, user perception is a deeply studied and well-established field in games. Likewise, virtual reality games possess equally if not more opportunities for research due to its immersive nature. As such, this research studies the user perception of events in a virtual reality game by using sentiment analysis and link perceptions to emotions. While previous studies conducted similar trials on traditional games, our work aims to conduct more large-scale and accurate sentiment analysis via artificial intelligence technologies.**

***Index Terms*—Virtual Reality (VR), Artificial Intelligence (AI), Sentiment Analysis**

I. Introduction

Video games are an effective tool for invoking human emotions through immersive experiences [2]. Past research has utilized video games and designed experiences to understand and invoke certain emotions. The methodologies for studying these emotions are vast and well-established. One popular method is the 'think-aloud' approach which has the user thinking-aloud while playing the game and voicing their inner thoughts. Typically, some form of sentiment analysis or survey is conducted to then extract emotional data. While these approaches have been researched and published for traditional flat-screen games, our implementation is the first for virtual reality.

As virtual reality games are an emerging market [4], there is a demand for new games and design improvements [6]. As such, it would be a valuable contribution for previous design and research approaches to be tested and implemented in VR. Our implementation uses the think-aloud method for players playing a virtual reality game, and uses AI to analyze gameplay recordings and sentiment throughout gameplay to generate an understanding between gameplay events and user emotion. This approach can be especially beneficial for game-testers who would like to view the change in perception / emotion of a user during gameplay.

II. Overview

A. Virtual Reality Game

The virtual reality game used for this research is called "Why did the chicken cross the road". This game was influenced by the popular mobile game "Crossy Roads" and was utilized in this research as it allowed us to have control over data collection and game parameters. The logic of the game is simple, to cross the road without getting hit by cars. There are 16 game levels which are randomly cycled for each gameplay.

Identify applicable funding agency here. If none, delete this.

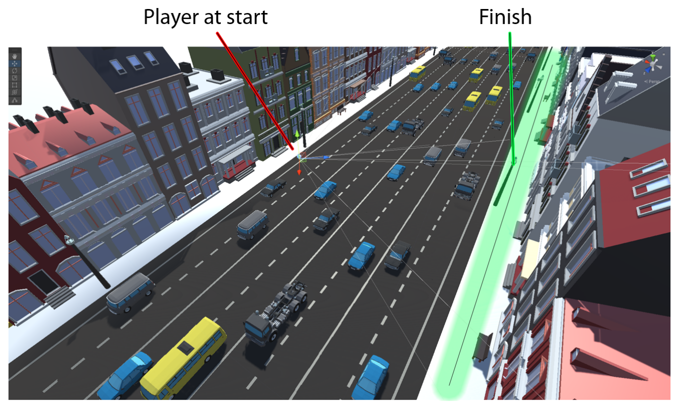

Fig. 1. A game scene from the player's point of view.

B. Data Collection

60 users played the game and had their data recorded. The coordinates and rotation of all game objects (including player) were saved locally. This resulted in enough data to replay each gameplay accurately. Along with this, the first-person view of each user playing the game was locally saved at approximately 2 frames per second. For the purposes of this project, synthetic statements were created to simulate user think-aloud responses during gameplay.

C. Replayer

A replayer was developed within Unity to access the large database of user gameplays and replay scenes. This replayer was designed to incorporate the obtained data and create meaningful visual representations of user sentiment analysis during gameplay. More on the incorporation of findings and data will be discussed in the discussions section.

D. AI Sentiment Analysis

We utilized the ChatGPT API to conduct sentiment analysis. Using this API, we were able to design a bot specifically to our needs. For the ChatGPT model, We used vision-4 as this allowed us to integrate computer vision for the analysis of game events. This meant that each sentiment could be analyzed with respect to the game events going on in the scene. Each statement analyzed is produced ratings based on emotion that was then used to evaluate the change of emotion throughout gameplay.

E. RBF and DTW

To create meaningful visual representations of the data generated, we used the Radial Basis Function (RBF). RBF is used to approximate multivariate functions [1]. In our case, it is used to approximate the intensity of emotions throughout gameplay. We did this by doing interpolation using RBF with a Gaussian Kernel.

We used DTW (Dynamic Time Warping) to compare the trajectories of the obtained interpolated data. DTW was useful in comparing gameplays with similar experiences and finding trends. It also allowed us to obtain valuable metrics such as total and average distance between warped points to aid in trajectory similarity assessment.

III. Technical Approach

A. ChatGPT API Development

To use the ChatGPT API, we created an account, uploaded funds, and followed the instructions [5] written by OpenAI to write Python code that creates a GPT client and functions similarly to the online chatbot. To develop our API to our needs, we chose the vision-4 model. We created two text files that describe what we need from the bot. All GPT bots are required to have an instruction text which describes the main functionalities of the bot. This instruction text is used every time at launch to optimize the bot to its tasks. Thus, one of text files we had contained a detailed set of instructions. This file describes the role of the bot and specifies areas in which the bot should be trained. For example, this file let the bot know that it will be expected to conduct sentiment analysis, and that it should be consistent across users. It showed the format of how the sentiment analysis is expected so it can be easily parsed and saved. The other text file is a document describing gameplay and what to expect. This file describes all of the levels of the game, the difficulty changes, and how they affect the game and users. This file also describes what to expect in the gameplay frames. For example, it mentioned which side the user would start on and where they would try to run to. It also specified how to know when a user was running, getting close to being hit by cars, winning, or when they died in the game. It also incorporated corrections about instances where the bot was tested and had a false understanding.

The ChatGPT bot was designed to conduct sentiment analysis by returning values based on emotion for each sentiment. The emotions were based on the Ekman hexagon scale [3]. This scale is established in emotion theory to classify human emotion. As such, we had our GPT model produce ratings for each emotion on a scale of 0-10. It also produced a color with

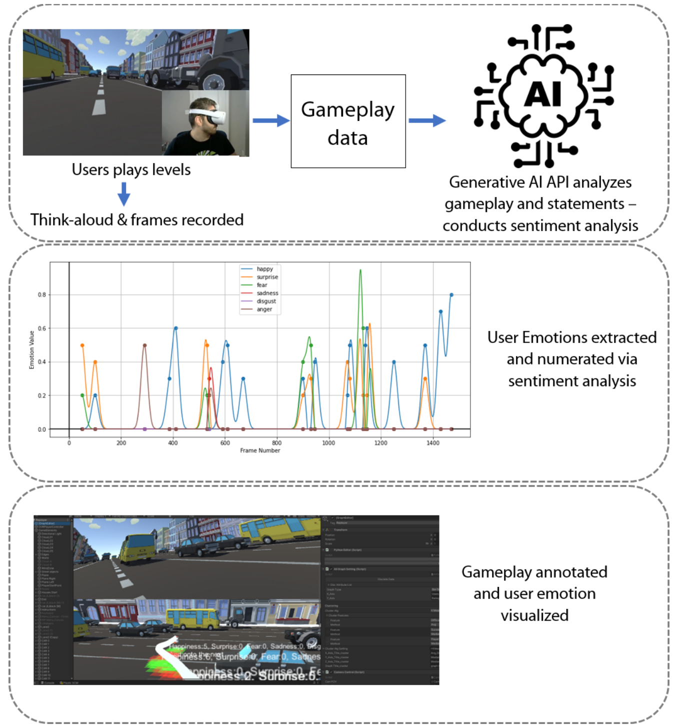

Fig. 2. Overview of Project Workflow

intensity corresponding to the most dominant emotion. The emotions and colors used for this scale are as follows:

- Happiness (red)
- Anger (magenta)
- Disgust (blue)
- Sadness (cyan)
- Fear (green)
- Surprise (yellow)

B. RBF Formula

The RBF was used to interpolate emotion ratings throughout gameplay. The formula taken as reference is

$$\sum_{i=1}^{N} w_i \varphi(||x - x_i||),$$

The Gaussian kernel was used in this formula. As this method can cause a variance between nearby points, we also incorporated regularization to reduce this.

$$\min_{w}\{||\Phi w - y||^2 + \lambda ||w||^2\}$$

C. DTW Algorithm

Dynamic Time Warping was used to calculate the similarity between trajectories. The trajectories analyzed were between different users playing the same levels. These trajectories were analyzed using the player position and emotion value to calculate costs. The formula below was used

$$DTW_q(x, x') = min_{\pi \epsilon A(x,x')} \Big(\sum_{(i,j) \epsilon \pi} d(x_i, x'_j)^q \Big)^{1/q}$$

IV. Discussion

So far in our research, we discovered that:

- Chat GPT API is excellent at conducting Sentiment analysis on a large-scale
- Chat GPT is surprisingly good at analyzing images and statements at the same time to produce consistent and realistic outcomes
- The radial basis function is an effective method for plotting values related to emotion during gameplay
- Dynamic time warping is helpful when conducting trajectory comparison on data interpolated via RBF
- The obtained results are useful in replayer visualizations

A. ChatGPT Usability

Compared to previously deployed techniques for sentiment analysis, we found that using a Chat GPT API is the most efficient and scalable solution yet. Our solution is able to view a user's entire gameplay (about 16 levels) along with statements and produce expected results in approximately 10 minutes. Although we have not deployed a technique to assess consistency or accuracy, early findings have shown that the tool's results are in line with expectations. Users typically perceived as 'happier' received higher scores in this emotion, and vice-versa for those who were perceived as less happy. These perceived early results consisted throughout the other emotion metrics as well. Thus, we deduced that ChatGPT is an ideal solution for a modern approach to gameplay testing and game design research. We also believe that this technology has the potential to be deployed in many other ways that can benefit the game design process.

B. RBF and DTW

RBF and DTW were shown to be effective techniques in visualizing the data generated by Chat GPT. Although there is still room for improvement in both, they both sufficiently created outputs that can be used to visualize data in an easy-to-understand format effectively.

RBF interpolation resulted in a meaningful connection between scattered plots. Figure 3 is an example plot showing the change in 'happiness' of a user throughout gameplay. This type of visualization makes it easy to see a continuous trend in the user's emotions.

DTW made good use of the interpolated values. Although our implementation still has a lot of room for progress in customization, it has satisfied our need for trajectory comparison so far. Figure 4 shows a useful example where DTW does trajectory comparison between two users who played the same game level. The x-axis represents the position of the user (left-to-right corresponds to start-to-finish). The y-axis represents the happiness rating interpolated value given by the RBF.

C. Data Visualization in Replayer

The replayer is a component we are looking to finish all implementations following the finalization in RBF, DTW, and data collection. However, as a proof of concept, we have deployed the following functionalities at least in full or in part.

- Augmented Annotations
- Colored emotion tiles

Figure 5 shows how these two have been implemented so far. As can be seen, this is an instance where two users had similar experiences in emotion for happiness. Especially near the end of the gameplay, there are very similar trajectories in emotion. DTW was able to present an average distance of 0.03 between warped paths, which is seemingly low in this early implementation.

So far, we found that the augmented annotations work really well in showing a numeric overview of the emotions, as well as the statements made by users to remind the tester of the ground truth. The colored emotion tiles were color-coded using the ChatGPT generated values for color. The visualization shows a good contrast and change in user emotions, one that can be helpful in assessing how game flow affects change in emotion.

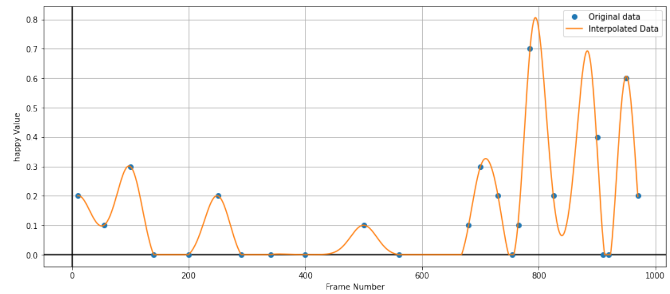

Fig. 3. Example plot for happiness value interpolated by RBF

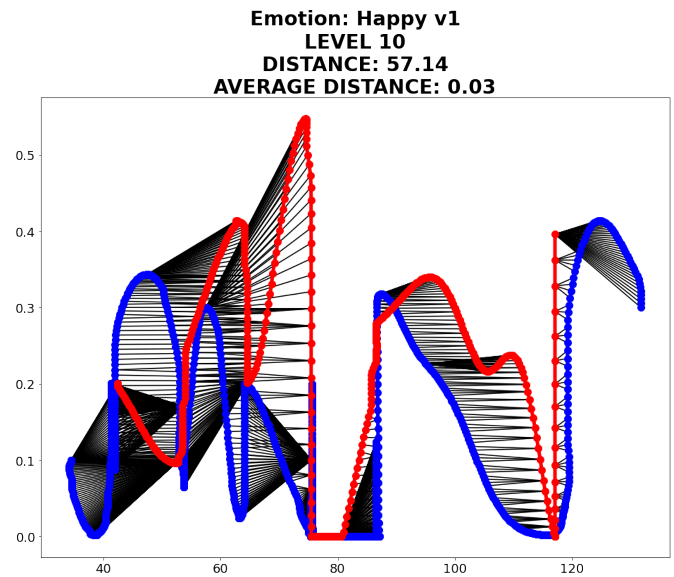

Fig. 4. Example plot for happiness value comparison between two users playing the same level

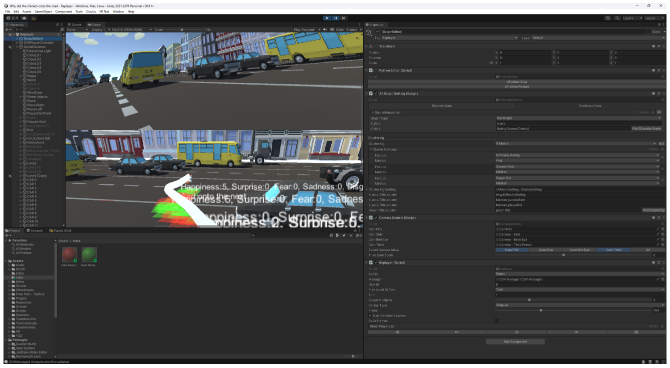

Fig. 5. Inside replayer screenshot showing augmented annotations and colored emotion tiles

V. Limitations and Future Work

Although we were able to achieve much of what we sought in AI from this project, we need to continue making considerable improvements and additions in almost all avenues. First, we aim to transform our data used to be from actual players rather than use synthetic data. A successful implementation of this will further support our AI system's robustness. Second, we hope to improve our RBF system to be adjusted for the data. Currently, there are no constraints to keep the numbers between the fixed 0-1 rating for each level. Third, we hope to improve our DTW implementation by creating more meaningful comparisons between paths. We aim to do this by adding costs into the main calculation that account more for the trajectory and positions of lanes to better align user emotions at critical points. Lastly, we hope to finalize current functionalities in the replayer as well as add a live-visual feed showing change in emotions on a graph throughout gameplay.

VI. Conclusion

To tackle the problem of understanding user emotions in virtual reality games, we devised a solution similar to previous attempts on traditional flat-screen games, transformed it to be used in VR games, and paired it with AI to create a modern and scalable sentiment analysis solution. This solution delivers a satisfactory proof of concept on the usage of AI in game testing and analysis. In future research, we hope to further uncover the potential of what more can be achieved using this technology.

Acknowledgment

I thank and acknowledge my advisor, Prof. Lap-Fai Yu, and lab-mate Yongqi Zhang, for their expertise and endless support throughout this exciting research.

References

[1] M. D. Buhmann. Radial basis functions. *Acta Numerica*, 9:1–38, 2000.

[2] Maximilian Croissant, Guy Schofield, and Cade McCall. Emotion design for video games: A framework for affective interactivity. *ACM Games*, 1(3), oct 2023.

[3] Paul. Ekman and Richard J. Davidson. *The Nature of emotion : fundamental questions / edited by Paul Ekman, Richard J. Davidson.* Oxford University Press, New York, 1994.

[4] Lik-Hang Lee, Tristan Braud, Pengyuan Zhou, Lin Wang, Dianlei Xu, Zijun Lin, Abhishek Kumar, Carlos Bermejo, and Pan Hui. All one needs to know about metaverse: A complete survey on technological singularity, virtual ecosystem, and research agenda. *CoRR*, abs/2110.05352, 2021.

[5] OpenAI. Chat GPT AI Reference. https://platform.openai.com/docs/api-reference, 2024. [Online; accessed 4-April-2024].

[6] Allied Market Research. Global virtual reality content creation market research report: By content type (videos, 360-degree photos and games), by component (software and services), by end user (real estate, travel & hospitality, media & entertainment, healthcare, gaming, automotive and others), by region (north america, europe, asia-pacific, middle east & africa, south america) - forecast till 2030. *Allied Market Research*, 2022.