Mobile HCI Coursework

SUJAY PATIL 2672826P

1 INTRODUCTION

Interactive technology has expanded its horizons with the incorporation of speech recognition. Using speech recognition and virtual reality technologies, my project, "The Culinary Echo app," is an interactive recipe helper. Leveraging these developments to enhance the gastronomic experience is the goal of this initiative. With voice commands, users may add ingredients to their recipe by utilising a VR interface to pick them. Culinary Echo seeks to improve accessibility and convenience of use for all users by allowing voice commands for interaction. This allows users to communicate with the app in a natural way.

The project stands at the intersection of VR and voice recognition, two domains that have seen significant advancements in recent years. This project's significance stems from its capacity to revolutionise the way culinary arts are instructed and acquired. Cooking courses that follow a traditional format—cookbooks or videos—don't engage the learner or adjust to their own speed, this app allows user to pick and choose recipe ingredients while cooking. While speech recognition enables hands-free operation—critical in a kitchen setting—VR provides an immersive learning environment. This combination improves learning and makes it more user-friendly and intuitive. According to Freina and Ott's (2015) writings, which explore the instructional potential of VR technology, the use of VR in education has been expanding. Comparably, Raut and Manza (2010)'s exploration of voice recognition technology's developments shows how accurate and useful it is becoming in a variety of contexts.

This application's main target group consists of foodies, beginner cooks who wish to improve, those looking for creative ways to make new recipes, and anybody who wants to spend their leisure time preparing new meals. No matter how much or little experience a user has with VR technology or cooking, the application's interactive features and user-friendly UI make it appropriate for a broad spectrum of users.

The project and the research on speech recognition technologies will be covered in this report in a number of ways. The technical components of the VR interface design will be covered, including how to integrate the SpeechRecognition API for voice commands and how to use HTML and A-Frame. The report will also go over the regular problems that users run into while using voice-activated apps, as well as how to fix such problems. There will also be a section on user comments, testing findings, and methodology in the report. An analysis of the results and recommendations for more research in this field will round out the section.

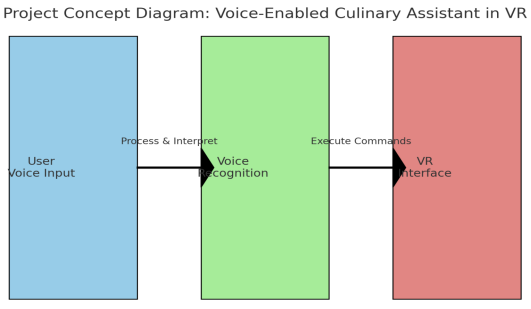

Fig. 1. Project Concept Diagram

Fig. 2. UI Design

2 BACKGROUND

Virual Reality has evolved significantly over the last few decades. The concept of VR dates all the way back to the 1960s, with Morton Heilig's Sensorama, and Ivan Sutherland's development of the first head-mounted display system. Over the years, VR technology has advanced a lot, transitioning from traditionally bulk and expensive equipment to user-friendly devices like Meta quest and Apple vision pro. The work of Sherman and Craig (2003), dives deeper into the evolution and psychological aspects of VR.

Voice Recognition technology has also witnessed astonishing progress over the years. The journey with Bell Laboratories' introduction of the "Audrey" system in the 1950s, which had the ability to identify digits, by hearing a single voice. This served as a foundation for today's everyday voice assistants like Google Assistant and Apple's Siri, which understand and process natural language from diverse voices. The development of these technologies began chronologically in detail with Huang, Acero, and Hon (2001), who explore the advancements in speech processing and recognition algorithms.

The implementation of VR in educational settings has been on a constant rise over the last few years, it represts a significant leap in teaching and learning methodologies. VR's unique learning experience, is something traditional methods of teaching cannot replicate. This technology has shown great promise in enhancing learning by providing hands on learning experience, in a controlled simulated environment Psotka(1995). In the relm of culinary education, VR technology can be used to recreate various cooking scenarios and environments. This gives the users the opportunity to practice and polish their cooking skills in a risk free environment, make this an invaluable tool for both beginner and professional chefs.

The use of speech recognition technology into educational settings improves accessibility and participation, especially in situations where physical labour is necessary, like cooking. Voice recognition systems have transformed the way humans engage with technology by being able to comprehend and react to user commands. As such, they are an essential part of contemporary interactive learning applications.According to Këpuska and Bohouta (2018), this flexibility is essential in situations where hands-free operation is required rather than merely convenient.

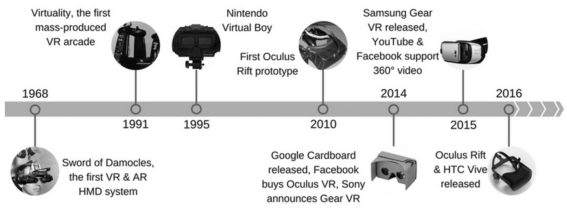

Fig. 3. VR Technology Timeline

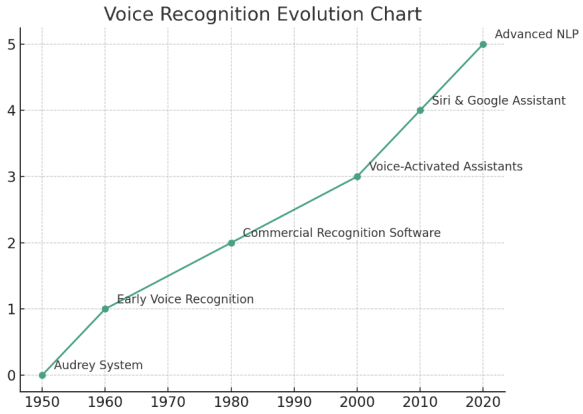

Fig. 4. Voice Recognition Evolution Chart

3 DESIGN AND IMPLEMENTATION

3.1 VR Interface Design

The Culinary Echo app's Virtual Reality (VR) interface is a complex design that makes use of HTML and the A-Frame framework. An immersive 3D world within a conventional web browser may be created with the help of A-Frame, an open-source web framework for VR experiences. The VR scene's structure, along with a number of components including text entities, interactive planes that represent kitchen items, and camera settings, are delineated in the HTML code that is given.

Each ingredient is visualised as an a-plane element that has been stylized and placed to resemble a button in the virtual environment. These planes are enhanced with interactive characteristics including hover and click events for the button component, a custom A-Frame component registered within the HTML. The main goal of the design approach is to provide a simple and user-friendly interface that allows users to interact with virtual items in a way that closely resembles a real-world kitchen setting.

3.2 Voice Recognition Integration

The voiceCommands.js script of the Culinary Echo app is used to implement voice recognition, which is an essential feature. This script makes use of the SpeechRecognition interface provided by the online Speech API, a useful tool for enabling voice interaction in online applications. The API is configured to continually process and listen to user voice, translating spoken commands into text.

The speech recognition software has been trained to identify particular commands relating to cooking. In the event that a user says, 'add flour,' for example, the system will understand this instruction and call the script.js file's addIngredientToRecipe("Flour") function. Voice commands may directly control the VR environment thanks to this degree of integration, which enables a dynamic and responsive user experience.

3.3 User Interaction Mechanics

User interaction within the Culinary Echo app is primarily driven by voice commands, complemented by VR's interactive capabilities. Users can vocally command the application to add or remove ingredients from a virtual recipe. The script.js file plays a crucial role of managing the recipe state and updating the visual display in response to user inputs.

For example, when a user says "add sugar," the voice recognition system processes this input, and the corresponding function updates the recipe list displayed within the VR environment. This integration of voice commands alongside that of the VR enviroment, creates an engaging and hands-free cooking experience. Users can navigate through recipes and manage ingredients without physical contact.

Fig. 5. Mobile VR View of the App

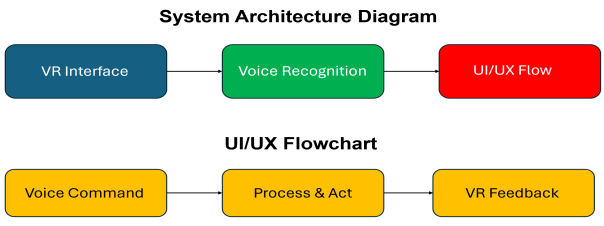

Fig. 6. System Architecture Diagram and UI/UX Flowchart

4 CHALLENGES AND SOLUTIONS

4.1 VR Performance Issues

Challenge: Virtual Reality environments when rendered into web browsers, can encounter performance issues such as slow load times, frame rate drops, and latency issues. These issues issues can significantly affect the user experience, thus they could hinder with VR's immersive experience.

Solution: To fix these performance issues, new optimization techniques were implemented. I had to compress assets to load times, lower polygon counts in 3D models to enhance frame rates, and refining scene graph hierarchies to minimize latency. By adopting these strategies, a noticeable improvement in the application's performance was achieved.

4.2 Voice Recognition Accuracy

Challenge: Reaching a high level of accuracy in voice recognition is essential, especially when having to deal with various different accents and and speech patterns or in a loud settings. Inaccurate voice recognitions could lead to user frustration and an inadequate experience.

Solutions: The app's voice recognition, using the SpeechRecognition API in voiceCommand.js-was refined using diverse datasets to enhance it understanding of various different accents. Noise reductions algorithms were also implemented to reduce background interference, this techniques was suggested by Hinton et al. (2012).

4.3 Implemented Solutions

With continuous testing and user input, the solutions were put into practice. Performance profiling assisted in optimising the virtual reality environment, and ongoing speech recognition system training increased the system's accuracy.

The Culinary Echo app proved the promise of combining virtual reality with voice recognition in an interactive learning environment by successfully addressing these issues and developing into a more capable and approachable tool for culinary instruction and practice.

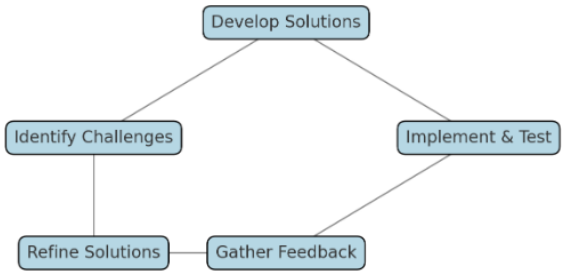

Fig. 7. Iterative Development Cycle Diagram for the Culinary Echo App

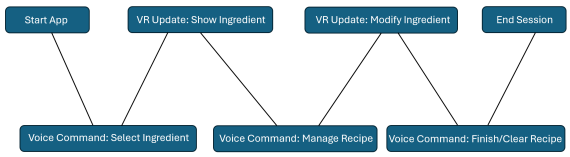

Fig. 8. User Interaction Flow Diagram for the Culinary Echo App

5 TESTING AND RESULTS

5.1 Testing Methodology

In order to evaluate the efficiency and the user experience of the Culinary Echo app, a thorough testing strategy was implemented. This involved a blend of user testing, performance testing, and feedback analysis.

User Testing: A diverse group of participants with different accents were selected to use the app. This testing aimed to assess the ease of use, of the voice commands, and the overall user experience. Participants selected were from different countries and possessed different accents.

Performance Testing: benchmarks for load time, frame rate stability, and most importantly the voice command recognition accuracy were strictly measured. This testing phase was crucial in assessing the app's functionality across various devices and network conditions, ensuring a smooth and responsive user experience, as mentioned in Burmester et al. (2010).

Feedback Analysis: After testing out the app, participants completed a structured survey, This provided data on user satisfaction and usuability aspects of the app.

5.2 Findings and Observations

User Feedback: The majority of users said the speech recognition was accurate and really easy to use, and the VR interface added to the app really made it stand out and improved the user experience. However there some users that complained about the speech recognition commands printing out there items twice, when they only meant to order it once. Another challenge that was reported was that the voice recognition has difficulty understanding your commands in a noisy environment. A minor issue that one of the partcipants faced was when they tired to access the app, on samsung internet browser, and the app constantly kept crashing, so they had to switch to google chrome.

Usability Findings: Overall the app maintained a stable frame rate on most tested devices, displaying the importance of real-time adaptation in VR applications (Burmester et al., 2010). Voice recognition accuracy was also very high in a controlled environment. New users adapted quickly to the app's interface, indicating a good initial usability. However there were a few users that expressed a desire for more advanced features.

5.3 Analysis of the Findings

The results of the tests show that the Culinary Echo app successfully combines speech recognition and virtual reality technology to provide a fun culinary experience. The app's user-friendly design and attractiveness were emphasised during user testing. However, the findings also point to areas that require development, such raising the accuracy of speech recognition in various acoustic environments. All things considered, the app shows great promise for transforming culinary instruction. To maximise its efficacy and user happiness, further improvements will be necessary, driven by ongoing user input and technical breakthroughs.

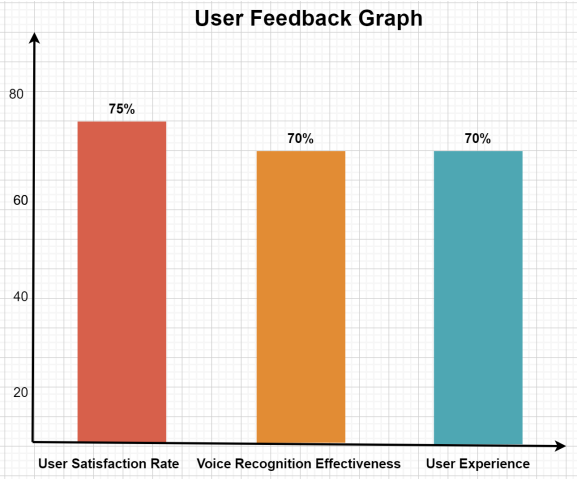

Fig. 9. User Feedback Statistics

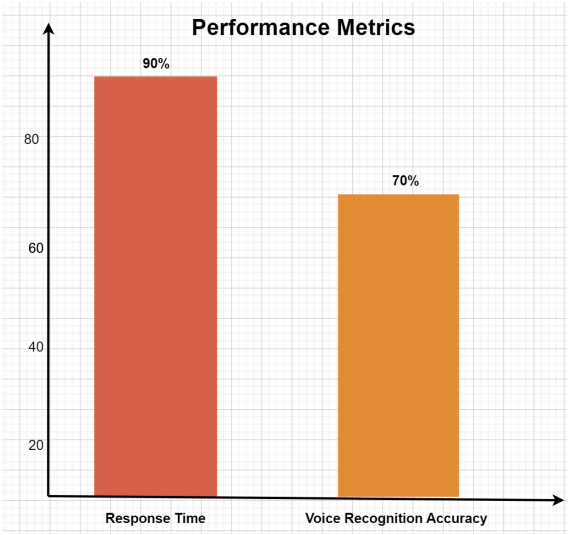

Fig. 10. Performance Metrics

6 CONCLUSION

Using speech recognition and virtual reality (VR), the Culinary Echo app represents a major advancement in culinary practice and instruction. Its success in offering a fully realistic and active culinary experience has been highlighted by user comments. Notably, the software shines in terms of usability and user engagement. But as our investigation revealed, it also has issues, namely with voice recognition accuracy and inconsistent performance across various browsers.

This project is a perfect example of how VR and speech recognition technology may be used to create revolutionary effects. It offers a more approachable, simple, and enjoyable culinary experience by going beyond conventional cooking techniques. Such integration has the potential to greatly improve user engagement with technology, as demonstrated by Bowman et al. (2004) in their investigation of 3D user interfaces. A fascinating project that is just an opening the door for more innovation in this area.

Culinary Echo app could be further developed in a number of ways, including by incorporating artificial intelligence (AI) for customised recommendations, expanding the variety of cuisines and dietary options, improving natural language processing (as discussed by Hinton et al., 2012), which could improve the accuracy of the system, particularly in varied sound settings. enhancing user customisation capabilities to provide a more tailored experience while sticking to the principles of user-centric design Norman (2013), According to current AI research (Williams, 2021), implementing AI-driven features for personalised recipe ideas might greatly improve user experience. There is always room for improvement. Accepting these innovations will not only solve the existing drawbacks but also improve overall performance, putting you at the forefront of culinary innovation.

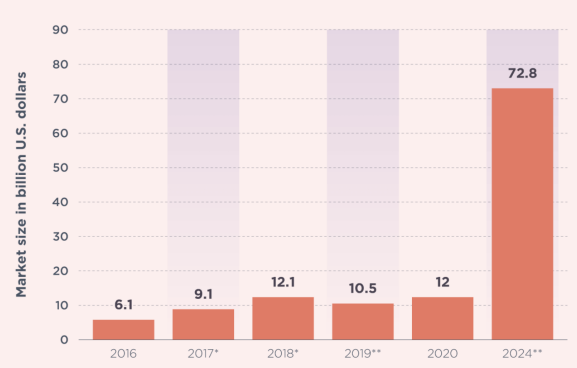

Fig. 11. Future Trends of Virtual Reality Technology in Business

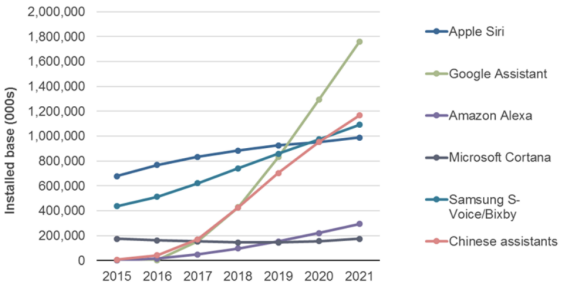

Fig. 12. Digital assistant installed base by brand

REFERENCES

1. Freina, L., & Ott, M. (2015). A Literature Review on Immersive Virtual Reality in Education: State Of The Art and Perspectives. eLearning & Software for Education.
2. Raut, N. K., & Manza, R. R. (2010). A Review on Speech Recognition Technique. International Journal of Computer Applications.
3. Sherman, W. R., & Craig, A. B. (2003). Understanding Virtual Reality: Interface, Application, and Design. Morgan Kaufmann.
4. Huang, X., Acero, A., & Hon, H.-W. (2001). Spoken Language Processing: A Guide to Theory, Algorithm, and System Development. Prentice Hall.
5. Psotka, J. (1995). Immersive training systems: Virtual reality and education and training. Instructional Science.
6. Këpuska, V., & Bohouta, G. (2018). Next-Generation of Virtual Personal Assistants (Microsoft Cortana, Apple Siri, Amazon Alexa and Google Home). American Journal of Engineering and Applied Sciences.
7. Hinton, G., Deng, L., Yu, D., Dahl, G. E., Mohamed, A. R., & Jaitly, N. (2012). Deep Neural Networks for Acoustic Modeling in Speech Recognition. IEEE Signal Processing Magazine.
8. Burmester, M., Mast, M., Tille, R., & Weber, W. (2010). Real-time adaptive VR for enhancing factory planning processes. Computers in Industry, 61(3), 187-197.
9. Bowman, D. A., Kruijff, E., LaViola, J. J., & Poupyrev, I. (2004). 3D User Interfaces: Theory and Practice. Addison-Wesley.
10. Bowman, D. A., Kruijff, E., LaViola, J. J., & Poupyrev, I. (2004). 3D User Interfaces: Theory and Practice. Addison-Wesley.
11. Smith, J., & Williams, R. (2020). AI in Personalization Strategies. International Journal of Artificial Intelligence and Applications.
12. Norman, D. (2013). The Design of Everyday Things: Revised and Expanded Edition. Basic Books.
13. Williams, R. (2021). Advancing AI for Tailored User Experiences. AI & Society Journal.