DATE: November 6, 2020

TO: Ananth Dodabalapur
Hanan Hashem
Venkata Suresh Rayudu

FROM: Evan Canter
Tanjeem Mazid
Andrea Nguyen
Deepanshi Sharma
Alexander Voth

SUBJECT: Enhanced Augmented Reality Prior Art

INTRODUCTION

In striving to bring audio sensory data to hearing impaired individuals, this report outlines a collection of existing patents and designs the team may utilize to create an augmented reality solution. Our design seeks to address the issue of reduced sound awareness, which is a major obstacle to communication for deaf of hard of hearing individuals [1]. We seek to improve upon our design by studying prior work in the fields of augmented reality and visualization technologies. In search of patent-exclusive material, we discovered an existing holographic caption product, various sound visualization techniques, and deaf and hard of hearing user research studies [2, 3, 8, 10]. Our search for patents specifically yielded a promising design for wearable visualization devices, an optical system for implementing a near-eye display, and a thorough augmented reality implementation [7, 11, 13]. The findings from our prior art search have given us unique insight about user accessibility, acoustic interpretation, optics, and approaches to visualization that will help us develop a safer, smarter, and more useful design.

OVERVIEW OF THE DESIGN PROBLEM

Reduced sound awareness has wide-ranging impacts for deaf or hard of hearing (DHH) individuals, from missing critical notifications such as a ringing fire alarm to inconveniences like a phone ringing at an inappropriate time [1]. We seek to overcome these obstacles to communication with enhanced augmented reality technology. The purpose of our project is to merge computer data and the real world, resulting in a mixed-reality environment that will provide hearing impaired users with a graphical visualization of sound data. This will be accomplished by harnessing optical hardware to combine computer-generated images displaying sound intensities and transcriptions with what the user is experiencing in real life. The product will be a headset-mounted display consisting of an optical system and microphones that will overlay these graphical visualizations of sound onto the user's view of their environment in real time. The combination of audio data graphics and the user's perception of their real environment will constitute an augmented reality that enhances the user's ability to interact with the world around them.

REVIEW OF PRIOR ART

As augmented reality solutions for accessibility have become increasingly popular, our team found several existing research and designs relating to our goal. In several studies regarding the head-mounted displays for the hearing impaired, the effectiveness of the system was dependent on the user interface and experience. As a result, while existing HMD captioning relies either on user configuration or static placement, our design will implement facial tracking and dynamically place captions in free unobstructive space. In building this design, we will also reference Sense Intelligent's patent on devices for visualizing and localizing sound. However, our improvements to the design include more intuitive audio graphics and sentiment analysis. Additionally, we plan on modeling our optical display after David Bohn patented near-eye display with some design changes. These design changes consist of our image source being a small plane versus and moving the polarizer location to the location light exits the display. With these purposes in mind, we will draw upon existing research and implementations to fine tune our design.

Prior Art Exclusive of Patent Information

Before considering patents specifically, we began our prior art search by focusing on more general sources of information such as research papers and products. This process was initiated by brainstorming pertinent terms and embarking on an advanced Google search. Next, we honed our search terms based on the Google results and moved on to a more formal search using Web of Science, a trusted, publisher-independent research engine. Finally, we consulted popular technology blogs and news sources to survey the market for existing products related to our research. As a result, we identified many interesting research methods and applications that furthered our understanding of sound visualization, caption services, and augmented reality products.

Real-Time Captioning Autoethnographic Study

In Exploring Augmented Reality Approaches to Real-Time Captioning: A Preliminary Autoethnographic Study, Dhruv Jain (a DHH individual) tests two different caption prototypes on a head-mounted display [2]. In the first prototype, AR Windows, captions scroll up and disappear in a web browser window placed in 3D space. The user can create multiple windows and resize and move them through hand gestures. The second prototype, AR Subtitles, places one caption line at a fixed position in the user's view. Over a 45-day period, Jain documented his experience with both prototypes. He found that AR Windows was easier to use with multiple speakers. Jain was able to create caption windows for each speaker and place them out of view of the speaker's face. However, if the speakers were moving, Jain preferred AR Subtitles. While typically AR Subtitles' "captions felt disconnected from speakers," if speakers moved the captions remained in view, and Jain did not have to continuously configure the caption windows [2]. In improving DHH sound awareness, we will implement similar real-time captioning on a head-mounted display. Our design will also need to be user friendly and not obstruct the user's sight. As a result, we decided to caption similar to AR Windows and have separate captions for each speaker in 3D space. However, through the use of computer vision, we hope to track speakers' faces and have captions automatically reposition as speakers move.

Holographic Closed-Captioning Glasses

Another approach to real-time captioning was implemented in Sony's Entertainment Access Glasses, which use holographic technology to superimpose closed-caption text over a cinema screen [3]. This solution allows movie-goers to experience each film as if they were watching it with a burn-in subtitle, making Sony's Entertainment Access Glasses ideal for people with hearing impairments. While the implementation of Sony's caption service using a static holographic display is different from our dynamic augmented reality approach, we gained crucial insight on how we can improve the user accessibility of our project. Sony's product elucidates the importance of providing a comfortable viewing experience by working to reduce eye fatigue and ensuring a see-through transmittance of more than 90% [3]. Loss of light through the optical system that facilitates our near-eye display is a major obstacle in our design, so we must ensure that our display is bright enough to be comfortably perceived by the human eye [4]. In this regard, the transmittance metric published by Sony is a valuable benchmark for our design's performance. Furthermore, Sony's technology has multilingual capabilities for caption-data, allowing the user to select from 6 offered languages [3]. This feature is something that we can bring into our design, and it will greatly increase the user accessibility of our design. The Sony caption product also drew our attention to a significant pitfall of caption services: depending on the length/shape of the captions, the text can obscure the user's view (see Figure 2 below). While this would only be a minor annoyance to a customer using the product in a movie theater, this potential obstruction of vision has much stronger consequences for our design that it is intended for use in a natural environment. In order to address this concern, our software will harness computer vision to track relevant figures in the frame and automatically reposition the text in a non-distracting manner.

Figure 1: Obstructive caption text displayed by Sony Entertainment Access Glasses [3]

The Research of Visual Characteristics for the Deaf Wearing AR Glasses

A proposal by researchers at South China University of Technology explores how the visual display interface of AR glasses is suited for users with hearing loss [8]. After a survey was performed for this study, results showed that the deaf mainly place an importance of using AR for sound in a work environment, followed by home and school and other situations [8]. The common factor for all of these environments is the need for a good display interfacing. The AR glasses used in this specific study contains a control panel that the user can slide up and down, back and forth, and click to operate [8]. This study can assist us on our project by providing some insight into the display and how to experiment our prototype in regard to cognitive efficiency. For example, the study describes an experimental interface that used 10 zones according to an ergonomic optical view. The results showed that deaf people had a smaller difference in cognitive efficiency and a more sensitive peripheral vision than the control group [8]. It would be noteworthy to see if cognitive efficiency is a significant issue when using the AR glasses by performing basic tests for usability and visual processing. It is also worth considering the idea of display placement in the headset in regard to peripheral viewpoint. By combining artificial reality and efficient display interfacing, it is possible to not only collect surrounding sound information, but also to provide an interfacing that promotes cognitive efficiency.

Glass Vision 3D: Digital Discovery for the Deaf

A pilot study was done with young deaf children that allowed them to use a Google Glass app to look at a QR code through augmented reality glasses and watch a relating video in American Sign Language (ASL) [10]. The video shows a person signing the word in ASL, some clipart pictures of the item, and a video of an animated animal singing about the object in ASL. QR codes were chosen for this specific study because of their ubiquity in recent times. They provide an easy way to launch into other sources to relay information [10]. Therefore, around 25 of these videos were uploaded to YouTube and linked with a QR code for use in the study for various object (Figure 2) [10]. One of the main points of study was usability. The population was younger deaf children. If they were able to use the Glass without difficulty, then it could be useful to incorporate in schools. After a single demonstration, the fifth graders were able to wear the Glass, scan objects around the classroom with the QR flash card codes, and watch the videos. At the end of the experiment, the classroom teacher expressed that it was obvious her students were enthusiastic in using the glasses as a learning tool.

The main takeaway from this study is making sure we design a prototype that is usable for both adults and children, and especially making sure that it is also easy enough to operate for both adults and children. One concern that arose in the study and that could arise in our project is the headset overheating. In the study a message would appear warning the user that the Glass was becoming too hot, and to wait to let it cool down [10]. We would need to look into the thermal aspects of our headset to determine any possible overheating. Additionally, while our first prototype will most likely be bulkier and will attract glances, the goal would be to minimize it to where anybody, adult or child, could wear the headset walking down the street and feel normal doing so.

Figure 2: Sample flash card with QR code used in the AR education study [10]

Combining augmented reality and speech technologies

An engineering team (Mirzaei et al.) from the Sharif University of Technology in Iran has linked the use of augmented reality (AR) with the need to help people with disabilities: specifically, deaf and hard-of-hearing patients. The team utilized automatic speech recognition (ASR) and text-to-speech (TTS) software to create an integrated system named “ASRAR” [12]. The system takes a subject’s speech in through a microphone, where it is processed and converted, by the aforementioned software, to readable text that is shown on an AR display for the patient. The required hardware consisted of: a camera, a microphone, a speaker, and a display. The team developed the software to be cross-platform, meaning that it is useable on a variety of portable devices. This makes the AR system widely accessible rather than having a single platform available. The team also demonstrated use with a head mounted display and mobile phones. To show the versatility of the project, the team included hardware requirements in the report. It is shown that a system of this capacity requires very little in terms of processing power and equipment.

Device	Description
Processor	Intel Mobile or Core Due 1.0 GHz.
Memory	1 GB.
Camera	640*480 VGA sensor or more, compatible with USB 2.0 or 1.1, or built-in cameras.
Microphone	All types of built-in or external microphone.

Figure 3: Table of Hardware Requirements for the "ASRAR" system [11]

This study is useful for our project because we can see that there are indeed practical uses for AR technology in the field of disabled patient accessibility. The team explains that a survey shows that more than 90 percent of deaf people are 'very interested' in using the ASRAR system as a way of communication [12]. In addition, our project takes AR one step further by using optics to merge actual vision with a generated heads-up display as opposed to how this team just uses a camera to display an image of the real world. What's most interesting about this study, is that it shows there is an actual demand for AR hardware in the deaf community. This study gives us reason to believe that our project will have an even higher approval due to better safety and applicability.

Patent Search and Findings

We modeled our patent research procedure on the guidelines provided by the United States Patent and Trademark Office (USPTO) and Stanford University's Office of Technology Licensing [5, 6]. Once again, we began the process by brainstorming pertinent terms related to our design specifications. Next, we conducted preliminary searches using those terms on many patent databases, including Google Patents and PatentBuddy. This helped us narrow down our search terms to the most relevant keywords. Then we moved on to find associated Cooperative Patent Classification (CPC) schemes using the USPTO website search [5]. Once we had identified the relevant CPC schemes, it was much easier to find the most applicable patent publications for our purposes and sort them into functional categories. Finally, we broadened our search to include international databases such as the European Patent Office's Worldwide Espacenet patent database. Using this streamlined process, we discovered many relevant patents that improved our understanding of augmented reality in the context of accessibility, optics, and existing sound visualization approaches.

Wearable Visualization Devices

In 2018, a Toronto-based company called Sense Intelligent secured a patent on Devices and Methods for the Visualization and Localization of Sound [7]. The patent details a method for processing sound, localizing the signal from one or more sound sources, and rendering acoustic visualizations on a wearable device. Sense Intelligent's innovation touches on various aspects of our design, such as mapping frequency to color, displaying acoustic information rendered at its physical position in reality (augmented reality), and localizing sound using multiple microphones. This patented design will serve as a very useful resource for our project, especially as a reference for sound analysis calculations. For example, the document explicates well-developed methods to calculate time-shifts between microphones and compare the sound intensities from different microphones on the same frequency [7]. This will save us from having to re-invent the wheel when it comes to the sound analysis algorithms that will enable our design.

However, our project is still unique from this patented invention in the way that we will display the audio data and our use of camera tracking to create an immersive experience. Sense Intelligent’s design displays audio information at a plotted spectrograph that maps frequency against time [7]. Our project will take the stereographic data to the next level by displaying it a more digestible manner that will integrate into the user’s environment. Instead of displaying graphs in the user’s vision, our design will convey data about sound intensity as color-mapped graphics (i.e. a green box for low sounds and a red box for an intense noise). Furthermore, our design will go beyond simply processing acoustic information by using camera tracking to determine sentiment and provide orientation cognizance. Sentiment analysis be accomplished by tracking the movement of facial features to determine whether the speaker is excited or angry and display that information in an emoticon image. In this way, the speaker will get valuable information about the sentiment of the situation that he wouldn’t be able to infer simply from sound intensity data. Orientation cognizance will be implemented by using tracking algorithms to find key figures in the frame and re-positioning the displayed graphics so that they are unobstructive to the user’s view. In these key areas, we will expand on the patented design for wearable visualization devices.

Projection Optical System for Coupling Image Light to a Near-Eye Display

Inventor David Bohn patented a unique near-eye display design that focuses on light efficiency [11]. The optical system described in this patent converts a point source image (exit pupil) into an image suitable for the eye at very short distances. Since our end design will most likely use a near-eye display, this patent is useful for examining existing designs and methodologies. One of the key features of the patented display is the use of bird-bath optics, which are curved reflectors that can direct light straight into the eye. The hardware sub team has been discussing using bird-bath optics in the display, so a detailed implementation of such a display is invaluable.

While the use of bird-bath optics is a similarity between our proposed display and Bohn's, there are a lot of details that make his display much different. His image source is a point, meaning all light radiates outward from a single location. Our design will lightly use a micro display which emits light over the span of a small plane. Also, his choice of polarizer location is right before the exit pupil, whereas ours will most likely be right after the light exits the micro display [11]. With these main design differences in mind, we will merely look at Bohn's schematics and raytracing as a starting point when modeling our own optical design.

System for Fully Augmented Reality by Anchoring Virtual Objects

Inventors Grinberg et al. have invented a system that anchors virtual objects, visually, functionally, and behaviorally, to create an integrated, comprehensive, and rational augmented reality environment [13]. The system uses an input device such as a smartphone to send a virtual image input to a microprocessor, which handles all aspects of viewing such as motion processing and viewing angles. This microprocessor sends encrypted input to the wearable augmented reality headset (smart glasses).

This option is one that the hardware team has considered for our project, but the use of a separate microprocessor may be redundant when using an already powerful input device such as a PC or smartphone. It should be noted that both of these devices would also be capable of sending image input as well as processing real world stabilization of the sent image. Our design will likely use a portable computer as a power and processing source for imagery and stabilization of the imagery for output to the user headset. The software sub team will need to work with the hardware sub team to ensure that the stabilizing and functionality of anchored virtual objects is feasible and within hardware constraints.

IMPACT OF PRIOR-ART SEARCH ON DESIGN DECISION-MAKING

The information produced in our prior art research has provided valuable insight on how we can improve our design in four major areas: user accessibility, acoustic interpretation, optics, and approaches to visualization. In terms of accessibility, our overview of existing products helped us identify what areas are crucial to providing a comfortable, useful, and safe experience for users. We have concluded that obstruction of vision is a major concern, and we will address it using camera tracking to automatically re-position text based on the relevant figures in the frame [2, 3, 8]. We also gained many useful references for acoustic data analysis calculations based on patented innovations for sound localization and frequency mapping [7]. These documented algorithms will be extremely facilitative in setting the framework for our design's backend functionality. In terms of hardware, the system that will facilitate our design's optics is modeled after David Bohn's near-eye display, which will help us solve the problem of optimizing light efficiency [4]. These patented schematics will serve as a starting point for our design, but our final implementation will include some modified components to better suit our purposes. Finally, our understanding of our design was greatly impacted by exploring prior approaches to sound visualization. Based on our survey of existing implementations, this is a very strong area that we can develop to set our design apart from prior work. Many existing approaches rely on a combination of spectrographic graphs and basic sound analysis to gather and display acoustic information [1, 7]. Our design will take this a step further by using camera tracking to perform sentiment analysis. In this way, users will be able to visualize the general emotion (i.e. angry or excited) associated with the situation – something they wouldn't be able to infer simply from sound intensity. Ultimately, the insight gleaned from our prior art research about user considerations, stereo analysis, optical design, and sound visualization will help us develop a safer, smarter, and more useful design.

CONCLUSION

Our findings from this prior art search have simulated new ideas and improvements from existing designs. As our team intends to bring an immersive augmented reality solution to the reduced sound awareness of DHH individuals, we determined that our design must provide audio sensory data without obstructing the user’s other senses. Current real-time captioning technology such as AR Subtitles and AR Windows, involves static placement or manual configuration [2]. However, this is often cumbersome and splits the user’s attention. As a result, our design will combine sound localization and speaker tracking to place captions in less obstructive spaces. From research of deaf individual's experience with HMD, we factored in the importance of minimizing the size and heat of the system in order to create a comfortable and wearable product. Additionally, we looked at existing designs regarding devices localizing and visualizing sound. While Sense Intelligent’s patent is an existing solution to our design problem, we hope to seamlessly blend audio data graphics in a more useful and comprehensive manner. Another important improvement we plan on implementing is the use of sentiment analysis to observe emotion in language and facial expressions. Finally, we found David Bohn’s patented near-eye display as a good source to model our optical system around. The results from our prior art search have given us a stimulated perspective on user accessibility, acoustic interpretation, optics, and approaches to visualization that will enable us develop a safe and practical solution.

REFERENCES

[1] Tara Matthews, Janette Fong, F. Wai-Ling Ho-Ching, and Jennifer Mankoff. 2006. Evaluating non-speech sound visualizations for the deaf. Behaviour & Information Technology 25, 4 (Jul 2006), 333–351. https://doi.org/10.1080/01449290600636488

[2] Dhruv Jain, Bonnie Chinh, Leah Findlater, Raja Kushalnagar, and Jon Froehlich. 2018. Exploring Augmented Reality Approaches to Real-Time Captioning: A Preliminary Autoethnographic Study. In Proceedings of the 2018 ACM Conference Companion Publication on Designing Interactive Systems (DIS'18 Companion), 4–7.

[3] Sony. Sony Access Glasses. http://goo.gl/0DKFoQ

[4] Hong H, Dewen C., Yongtian W., Sheng L. “Near-eye displays: state-of-the-art and emerging technologies.” in Proceedings of SPIE -The International Society for Optical Engineering. 10.1117/12.852504.

[5] “Seven Step Strategy,” United States Patent and Trademark Office - An Agency of the Department of Commerce, 22-Mar-2019. [Online]. Available: https://www.uspto.gov/learning-and-resources/support-centers/patent-and-trademark-resource-centers-ptrc/resources/seven. [Accessed: 05-Nov-2020].

[6] “Performing a Basic Prior Art Search,” Performing a Basic Prior Art Search | Office of Technology Licensing. [Online]. Available: https://otl.stanford.edu/performing-basic-prior-art-search-0. [Accessed: 05-Nov-2020].

[7] H. Hu, “Devices and Methods for the Visualization and Localization of Sound,” U.S. Patent 10 111 013, 23-Oct-2018.

[8] Li Z., Mo X., Shi C., Jiang S., Jiang L. (2020) The Research of Visual Characteristics for the Deaf Wearing AR Glasses. In: Ahram T. (eds) Advances in Human Factors in Wearable Technologies and Game Design. AHFE 2019. Advances in Intelligent Systems and Computing, vol 973. Springer, Cham. https://doi.org/10.1007/978-3-030-20476-1_16

[9] “Deafness and Hearing Loss,” World Health Organization. [Online]. Available: https://www.who.int/zh/news-room/fact-sheets/detail/deafnessand-hearing-loss. [Accessed: 05-Nov-2020].

[10] Parton, B.S. Glass Vision 3D: Digital Discovery for the Deaf. TechTrends 61, 141–146 (2017). https://doi.org/10.1007/s11528-016-0090-z

[11] Bohn, D., 2017. Projection Optical System for Coupling Light to a Near-Eye Display. EP2948813B1.

[12] M. R. Mirzaei, S. Ghorshi and M. Mortazavi, "Combining Augmented Reality and Speech Technologies to Help Deaf and Hard of Hearing People," *2012 14th Symposium on Virtual and Augmented Reality*, Rio Janiero, 2012, pp. 174-181, doi: 10.1109/SVR.2012.10.

[13] Grinberg, D. et al., 2015. SYSTEM WORN BY MOVING USER FOR FULLY AUGMENTING REALITY BY ANCHORING VIRTUAL OBJECTS. US9210413B2