Željko Nenadić

Capstone Mentor Michael Ang

Final Capstone Paper

12.04.2022

Capstone Project: Machine Vision

Introduction

Inspired by the rapid technological development occurring in the world and the fantastic experience and knowledge that I received throughout the past four years within the Interactive Media department at NYU Abu Dhabi, I have been creating and working on my capstone project for the past two semesters. I designed and developed a fully functioning wearable headset that utilizes LiDAR technology to scan its immediate environment and display it in first-person mode to its user. The goal of my project was to create an interactive educational tool that would help people learn and better understand LiDAR as a novel technology and motivate curiosity for further research. The question that I attempted to answer through this project is whether such an interactive yet puzzling device would create a sense of interest in its users and guide them to learn more about the technology without much external intervention. Although the road to completion was a difficult one, filled with obstacles that inevitably shaped my capstone project, and I would say for the better, it was a truly amazing learning experience that I genuinely hope and believe has given me the essential expertise in coding, design, and production which will be implemented in my future projects.

Initially, throughout my late sophomore and early junior years, when I began brainstorming rough ideas regarding my capstone project, like many of my peers, I was not wholly certain regarding what I wanted to work on. However, during my first January term in Madrid, I was enrolled in a course on surveillance. The course introduced me to the many forms of surveillance methods used in history and the present, from espionage and the panopticon prison design to modern surveillance technologies such as cameras, radar, audio recording devices, etc. Due to the exciting nature of the topic, and my personal interest in technology, for some period after the class had ended, I was particularly interested in researching modern surveillance technologies. During that period, I came across a technology that I was completely unfamiliar with, unlike the previously mentioned ones.

This particular technology was Light Detection and Ranging (LiDAR). Although its primary use is not surveillance, I was introduced to it by researching surveillance technologies. Therefore, I believed that it would be best used in that field for some time. During this period, it was the first time that I thought of pursuing something related to these technologies as a capstone project topic. As the time to begin working on my capstone project came closer, I began to become increasingly interested in LiDAR technology. The first obstacle, however, was the fact that LiDAR technology, among other factors, is relatively new and extremely expensive. Although I will explain in detail how this technology works and what its uses are further throughout this paper, it is essential to note that it is divided into two primary groups: two-dimensional and three-dimensional laser scanners. The two-dimensional scanners are relatively lower priced compared to three-dimensional scanners. However, for my capstone project, I was intending to use the three-dimensional versions.

During the initial stages of brainstorming ideas for my capstone project, I came to the conclusion that I would like to create and work on a device that would utilize LiDAR technology, be relatively straightforward to use, enhance human sensory perception, and serve as an educational tool. These were my primary criteria that built the foundation for my capstone project, and although I was fully aware that throughout the completion of this project, many aspects of it would evolve, I was determined that these four criteria would constantly follow it and be fulfilled when the final product was complete.

Pre-production and Research

Although I had a rough idea as to what my project would be, during the initial stages of working on the project, I conducted research as to how LiDAR and similar technologies were used in a combination of engineering, science, and art. Throughout this research period, I discovered four particular artists and projects that served as a significant inspiration for my own work. One of the fascinating combinations of technology and art that I found during my research was cyborg art. It is a form of artistic expression where the artist creates bodily modifications to enhance the biological and organic limitations of the human body. Most often, these modifications are permanent or permanently change the body of the person.

One of the most outstanding works of cyborg art, which alters a human's sensory perception, meaning it alters the way in which one senses their environment, is Neil Harbisson's Cyborg Antenna. Neil Harbisson, who proclaimed himself as a cyborg artist, is suffering from achromatopsia, which is a disease in the spectrum of color blindness that makes it impossible for him to perceive color. Harbisson has been suffering from this condition his entire life and, through the use of bodily modifications, has created an art piece that allows him to perceive color in a certain way. The device called Cyborg Antenna is an antenna that is implanted from one end into his skull and from the other end stems above his head. The end that is implanted into his skull contains a chip that is connected to Harbisson's brain, directly interacting with his perception, and the other end of the antenna includes a camera that senses color (Donahue). Through this contraption, Harbisson gathers the data of colors with light and converts it into electrical signals that are processed in his brain into sound, allowing him to practically hear color. The current version of his modification will enable him to hear colors that are outside of the visible spectrum, as well as hear the color of images that can be wirelessly transmitted into the antenna (Donahue). This artwork serves as a perfect example and my personal inspiration for the use of technology to modify and enhance the human body and its capabilities. Although for my capstone project, I did not intend to create a technology that would permanently modify the human body; therefore not, particularly within the field of cyborg art, this piece served as a major inspiration for creating a device that would enhance the human perception of the environment in some form.

Another artwork that served as an inspiration for my project is Osmose by Canadian artist Char Davies. The artist intended to create an artwork in the form of a medium in which an observer would be able to be immersed in a world that would react in real-time to their biological responses such as breathing and movement. Initially, Davies used canvas and oil paint to create immersive environments; however, after discovering virtual reality technology, she perceived it as the ideal medium for her work. Therefore, she created Osmose, an immersive virtual reality world where the observer could interact with their surroundings through their own biological signals (Grau 197). What I find particularly inspiring about this project is that Davies utilized a novel technology for the time, the 1990s, and although this particular technology had been primarily used in the military and high-end computer companies, she successfully used it in an artwork (Malloy and Davies 324). This directly connects to my idea to use a relatively novel technology such as LiDAR to display a different perspective of the environment to the observer and make them become more interested in interacting and learning about the tech.

The third major inspiration for my artwork, which is also the first artwork that used LiDAR technology, is Antoine Delach's Ghost Cell. Delach used LiDAR as well as other photogrammetry tools to create a detailed scan of Paris and compiled the footage into a documentary of the city. What I found particularly interesting about this project is the visual aesthetic and how it obviously portrays objects and environments that are highly recognizable; however, it does that through the lens of a machine (Failes). When I discovered this project, I was inspired to create something that would allow people to see through the eyes of a machine, even if the things they are seeing are objects and environments from their everyday life. I imagined giving people the opportunity to see the contrast between seeing them through their own eyes and through the eyes of a machine.

It is important to note that not everyone shares the same enthusiasm for LiDAR technology. Elon Musk, CEO of Tesla Motors is one of the most prominent critics of this technology. On multiple occasions, he has stated that this technology is unreliable, expensive, and that among its competitors, Tesla Motors is unique in the way that they are not using LiDAR for their autonomous vehicles, and plan to, even in the future, focus on cameras and radar for their self-driving vehicles (Templeton). However, some hold the opinion that LiDAR's most vital characteristic is its precision. Research conducted on robot localization has shown that LiDAR is a promising tool for high precision localization of robots and autonomous machines, especially when combined with other localization technologies such as GPS (Ma, et al.).

Description and Production

For my capstone project, Machine Vision, I created a wearable technology, a headset that would allow its users to see through the eyes of a machine with the use of light detection and ranging (LiDAR). The main purpose of my project is to make an educational tool that would allow people to experience how most modern machines that use three-dimensional vision perceive their environment, as well as make the users curious as to what LiDAR is and get more familiar with this relatively new, and to most, alien technology. The headset, similar to existing virtual reality headsets, displays in real-time and in a first-person perspective the wearer's environment. The image is generated with the use of a LiDAR scanner mounted on the front side of the headset, giving the wearer the perception that they are observing the environment through their own eyes. In order to achieve this goal, the headset needs to be capable of performing three fundamental actions. The first action that the headset needs to perform is to collect data from its environment. This is done with the use of a Cygbot CygLiDAR D1 scanner. The scanner, which is practically a distance measurement tool creates a three-dimensional matrix of points in space.

This is done through a time of flight (ToF) method. The time of flight method is a method used for measuring distance and is used by technologies such as radar, sonar, and LiDAR. The sensor transmits a wave, in the case of LiDAR, a light beam, that comes in contact with a surface. The beam is reflected by the surface and absorbed by a camera in the sensor. The sensor measures the time (t) between transmitting the beam and receiving it from the reflected surface. Since the beam travels at a constant speed, in the case of LiDAR the speed of light (c), the sensor then determines the distance between itself and the surface from which the beam was reflected. This calculation is done through a simple formula, in which the sensor, having both the time and speed, calculates the distance, and assigns an x, y, and z coordinate in a three-dimensional space. The CygLiDAR D1 scanner that I am using for this project has a resolution of 160x60 pixels, each pixel being a single beam of light that is transmitted and reflected by a surface, meaning the matrix consists of 9600 individual beams of light which are constantly being transmitted, reflected, and absorbed by the sensor at a rate of 15 times per second. The matrix covers an area of 120 degrees horizontally and 60 degrees vertically. This closely simulates human vision, closely mimicking how the user would see the environment under default conditions.

After the data has been collected by the sensor, the second fundamental action the headset must perform is to process and render the data in order to visualize it. There are two particular ways in which my project visualizes the data. The first one is creating a two-dimensional image in which every pixel in the 160x60 matrix is assigned a color value based on the distance between the sensor and the surface that the respective pixel is scanning. This visualization and data processing is done by utilizing a stock software - CygLiDAR viewer, developed by the manufacturers of the sensor. The software runs on a Windows machine that is connected to the headset's sensor.

The second method of data visualization is done with the use of the Robot Operating System (ROS) in a CatKin workspace running Rviz software, all of which is hosted on a Linux Ubuntu 18.04 operating system. Rviz collects the data from a sensor and displays it in a three-dimensional environment in which the camera perspective can move freely. This gives the headset the capability to collect and store point cloud data that could later be used to analyze the environment which the user was observing from a third-person perspective as well as create three-dimensional models of the scanned surfaces. The third and final fundamental action that needs to be performed by the headset is displaying the visualized data to the user. This is done through a screen in front of the user's eyes which wirelessly mirrors the screen of the machine that is used to process and visualize the data. The machine and display are connected via a WiFi connection.

It is important to note that the display is placed directly in front of the user’s eyes completely within their field of view, and is absolutely unobstructed, meaning there are no optical components such as lenses to focus the user's view upon the screen. The construction of the project was finalized, when it was determined that these three fundamental actions could all be performed separately. The headset was created using a lightweight helmet which was modified to contain the scanner and the display chamber, and maximize the user’s comfort. In regards to comfort, the helmet was padded from the inside using thin layers of foam which also served as protectors for the cables and electrical components. A slot for the sensor was mechanically extruded in the front side of the helmet, making the sensor easily detachable for troubleshooting purposes.

The display chamber that contains the screen was modeled using Solidworks and 3D printed in PLA material. To protect the chamber from being damaged by frequent use, a protective layer of thin cardboard was attached to the outer walls. The chamber is permanently attached to the helmet with the use of two machine screws and a strong adhesive. The display, however, is completely detachable from the chamber, since it is attached by weak adhesives and holding slots. Throughout the project presentation, the headset was fully functional and the audience was able to interact with it. For presentation purposes, a large screen was displaying what the headset was displaying, making it possible for the audience to see what the user wearing it is experiencing in first-person view.

Observation and results

Throughout the three-day Interactive Media Capstone Project Exhibition, my project was displayed and available for audience interaction. Each day I dedicated inconsecutive four hours to being present and assisting the users in interacting with the project. While I was present, the audience was able to interact with and test the headset, and ask any questions regarding the project, technology, and hardware, or software being used. Throughout the interactions with the users and the audience, I observed several systematic ways in which they interacted with my project. After approaching, and asking me if they could interact with the headset, the audience was initially curious about the video projected on the large screen.

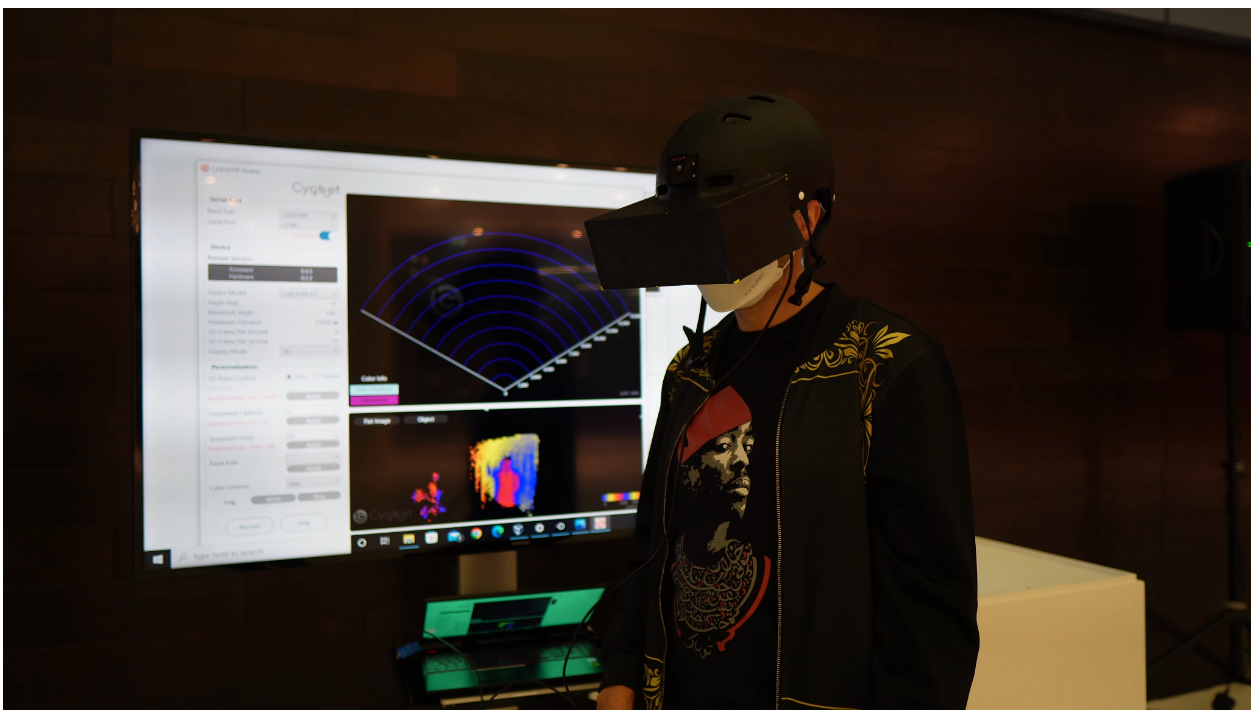

Image 1. A user wearing the headset with their view visible on the background screen.

I offered each audience member to pick up the headset, without instructing them to put it on, or in the cases in which they were hesitant to touch it at first, I would pick it up and hand it to them, again without any further instructions. The user, noticing that the device was a helmet, instinctively but slowly placed it on his head, and after noticing the first-person screen, carefully adjusted it to their preference. After adjusting it, the users would gradually turn their heads in any of the four directions until they recognized either the rectangular pedestal on their right side, a wall on their left side, or a shape of a person standing in front of them.

Immediately after realizing that they are viewing their real-time surroundings, the users would lift their hand up, rotating and observing it for some period of time. Then they would proceed to observe other physical objects around them, such as other members of the audience, etc. Many users initially assumed that they were viewing the objects in thermal vision, and asked if that was correct. However, as I explained that it is not thermal vision, but LiDAR technology, they would continue to interact with their environment, remove the headset and ask about LiDAR technology. All participants became interested to learn what LiDAR is, how it works, and how it can be used.

Image 2. A user lifts and observes their hand while wearing the headset.

As I was explaining, they would continue to interact with the headset and the scanner. For example, as I elaborated that the scanner uses ToF to measure distance, and that the color of the objects depends on the distance from the sensor, many participants would attempt to move an object such as their hand closer to and away from the sensor and observe how the colors on the screen would change. This was followed by both philosophical and technical discussions on how this form of technology could be used in the future, what that means for humanity, and many asked if I could show them the three-dimensional storing of point cloud systems. Although I did not have the answers to some of the more philosophical questions, I was extremely pleased to see that my project sparked such deep and interesting discussions, similar to the ones found in the interactive media classroom.

The main factor that led to the design choice to remove optical components such as lenses was the user feedback when the project was tested prior to the showcase. To simulate a virtual reality environment using a display in a Google Cardboard VR box. The users often stated that they felt nausea and fatigue after some time of wearing the headset and that the environment felt too overwhelming. There are two reasons behind this issue. The first reason is that the framerate in which the data is visualized comes at a maximum of 18 frames per second due to the technological limitations of the scanner and the processing power of the computer. Second, because of the wireless connection between the display and the computer, the rate at which the image was being mirrored from one to another depended on the connection strength. This caused some levels of lag between the user's movements and the visual feedback they received, which may cause fatigue if the user is immersed in the environment.

The issue, however, was fixed when the users observed the screen directly, giving them the sense that they were only semi immersed in the virtual environment. By being constantly aware that there is a screen in front of their eyes, all of the problems with lag, low FPS, and screen delay were eliminated. Therefore, for the sake of the user's comfort, I decided that the only reasonable solution would be to altogether remove any optical devices that I initially thought would enhance the experience and allow the users to be only semi-immersed in the virtual environment.

Conclusion

After a year of research, design, and testing of my project and after a very successful exhibition, I have concluded that my project has been very successful in serving as a tool for education on LiDAR technology. Although initially puzzled, the users were eager to explore and discover the capabilities of the device, and as they kept updating their knowledge, the interaction with the headset and the environment displayed in it became increasingly more intense. Beginning with observing their own hands, making sure that they themselves are physically present in the new virtual environment, the users continued exploring and interacting with their environment through LiDAR sight, attempting to discover how their vision works. As they became more familiar with the technology, their curiosity about its functionality, future, and obstacles grew. Many technical and philosophical questions emerged entering the fields of robotics, industry, military use, etc. Especially since many of the users were either completely unfamiliar with or had only limited knowledge of this technology, the curiosity regarding the depths of the back-end development of the project and the deep topics of conversation which emerged led me to conclude that the project was a success.

Works cited:

Davis, Erik. “Osmose.” Wired, Conde Nast, 1 Aug. 1996, https://www.wired.com/1996/08/osmose/.

Donahue, Michelle Z. “How a Color-Blind Artist Became the World's First Cyborg.” Science, National Geographic, 3 May 2021, https://www.nationalgeographic.com/science/article/worlds-first-cyborg-human-evolution-science.

Failes, Ian. “'Ghost Cell' Uses Cutting-Edge CG Tech to Reveal a New Side of Paris.” Cartoon Brew, 12 May 2016, https://www.cartoonbrew.com/cgi/ghost-cell-uses-cutting-edge-cg-tech-reveal-new-side-paris-138954.HTML.

Grau, Oliver. “Virtual Art—Digital! The Natural.” Virtual Art: From Illusion to Immersion, MIT Press, Cambridge, MA, 2007, pp. 193–211.

Malloy Judy, and Davies Charlotte. “Women, Art, and Technology - Landscape, Earth, Body, Being, Space, and Time in the Immersive Virtual Environments Osmose and Ephémère.” Women, Art, and Technology, MIT Press, 2003, pp. 322–337.

Templeton, Brad. “Elon Musk's War on Lidar: Who Is Right and Why Do They Think That?” Forbes, Forbes Magazine, 28 June 2021, https://www.forbes.com/sites/bradtempleton/2019/05/06/elon-musks-war-on-lidar-who-is-right-and-why-do-they-think-that/?sh=72671a2b2a3b.

Ma, Wanfeng, et al. “Two Particle Filter-Based INS/LiDAR-Integrated Mobile Robot Localization.” Multimedia Technology and Enhanced Learning Second EAI International Conference, ICMTEL 2020, Leicester, UK, April 10-11, 2020, Proceedings, Part I, edited by Zhang Yu-Dong and Wang Shui-Hua, Springer International Publishing, Cham, 2020, pp. 349–357.