Technology Review

Carl Benson

Group 9

No more touch. No more Keyboard. Bring it All Together. Using Technology to Teach Humans.

Instructors: Kirsten Winters & Kevin McGrath

November 20, 2017

Abstract

Technology is increasing at an amazing rate, and with it comes many opportunities. Headmounted displays can completely immerse a user in a virtual world. Computers can share the contents of nearly all human knowledge in the blink of an eye. They can even identify and react to occurences in ways similar to that of humans. This paper will cover thre options each for virtual reality headsets, data hosting, and machine learning frameworks as well as compare and contrast them.

CONTENTS

I. ROLE

I am one of two people in the group with a computer suitable for running VR applications, so I will be doing part of the work for the VR develepment portion of this project. Along with this, I will also be working on the serverside coding for responses and data management.

II. VR HEADSET

A. Description

The VR headset is one of the critical systems for our project. Without it, there would be nowhere to display the information. There are several options for headsets that have emerged in recent years. These displays have evolved to include similar features and specifications but each still has differences when compared to each other.

B. Criteria

1) A high quality display is a must. Early models featured displays with low resolution and refresh rate which could cause nausea after even a short session.
2) An interaction method is very nice to have. The minimum viable product at the end of this project does not require interaction with the display data, however it would be required for certain stretch goals.
3) The headset should be price conscientious. These systems may be deployed in multiple locations, so the lower the cost is the better.
4) Development shouldn't be overly complicated or full of hoops to jump through.

C. Hololens

The Hololens is the headset created by Microsoft to provide a mixed reality environment. Unlike the Vive and Rift, which completely encompass the view of the user, the Hololens overlays objects onto the current environment. Surroundings are still available, but through the display, floating items are overlaid onto it. Interaction is handled without the use of controllers. Cameras integrated into the headset track hand movements and interpret gestures. Items are selected by simply pointing at them. These technologies come at a premium however, which is the primary drawback to the Hololens: pricing begins at $3,000. [1]

D. HTC Vive

The HTC Vive provides a full virtual reality experience. The head mounted display fully obscures the user's vision, allowing them to only see what is displayed on screen. There are two displays, one for each eye. Each display features a resolution of 1080x1200 at a refresh rate of 90hz. Two external devices set with a high point of view track the headset to allow for head tracking and positioning. This tracking allows for a user to look and move around the environment. Interaction is handled through the use of two handheld controllers that feature buttons as well as pointing ability. The Vive costs $600 and all the features are included at that price. [2]

E. Oculus Rift

The Oculus Rift is another full VR headset. Like the Vive, the user's vision is fully obscured by the display. The twin displays in the Rift feature the same resolution and refresh rate as the Vive as well. Head tracking is handled through integrated sensors, with optional external devices for more accurate tracking as well as tracking position within the room. The Rift costs $500 for the headset and controllers, with another $135 for position tracking. [3]

F. Comparison

The Rift and Vive both feature similar features to the Hololens. While the extra functionality of mixed reality and controller-less interaction would be nice, they don't justify a over a $2,000 cost increase, so the Hololens isn't a good option. The Rift and Vive both feature very close features and specifications. Both have the same quality of screen, both come with controllers, allow for head tracking, have the option of positional tracking, and are supported by the same game engines.

III. DATA HOSTING

A. Description

Hosting the data somewhere is a requirement for it to be retrieved. Even if it is saved on the computer running the display, it still is then being hosted by that computer. There are a number of requirements for the storage location of the data.

B. Criteria

1) The server must be reliable. If the server goes down, the data is inaccessible.
2) Hosting the server isn't free, however the cost shouldn't be a significant cost in the project.
3) The server must be fast enough to load the data and send it to the headset within the timelimit.
4) The data stored may contain sensitive information, so it must be stored in a secure manner.

C. Amazon Web Services

Amazon Web Services (AWS) is a wide range of services offered by Amazon. Among these are the AWS Lambda and S3 services. AWS Lambda provides computational services in response to events. This would allow for the virtual assistant to send an event to the server which would then process the request and return the resulting data to the headset. If we choose to use the Alexa, the Alexa skills tie directly into the Lambda service. AWS S3 is a storage and content delivery system. It stores given data in an accessible form to then deliver on request. There are currently a total of 15 data centers in the US offering these services. Having multiple data centers allows for better response time by offering a server physically close by. It also increases reliability, as if something goes wrong in one area there is a fallback already in place. AWS offers a service known as Macie for their storages. This service utilizes machine learning to ensure that any data is stored in such a way to keep even personally identifiable information secure. Their pricing starts at $0.023/GB per month. Additionally, the client already has systems in place utilizing AWS services. [4]

D. Google Cloud Platform

Google has their own Google Cloud Platform for the services they offer. They offer data storage, processing, and delivery services. They have 13 data centers in the US. Their pricing starts at $0.026/GB per month for storage. Google's services go through routine audits to ensure that they are compliant with a number of security standards, such as for storage of personally identifiable information. [5]

E. Local Hosting

Local hosting is another option for data storage. The only cost is the upfront equipment cost and the cost to operate the system. However, this requires all operation and setup to be done in-house. The software to enable data request, keep it safe, and process it need to be configured and kept up to date. If anything were to happen to the local server, the data would be gone. Access to the data outside of the location needs to be set up separately and protected from attacks.

F. Comparison

Local hosting is far too large of a hassle to be reasonable. Any possible cost saving benefits are outweighed by the risk of data loss, security breach, and maintenance. The services offered by Google and Amazon both offer very high reliability and availability at a low cost. Google and Amazon both offer a similar set of services, at similar cost, and boast multiple data centers. The largest deciding factor is that CDK already utilizes and has data stored on AWS, which is why we have elected to go with it.

IV. MACHINE LEARNING FRAMEWORK

A. Description

Machine learning introduces functionality to a computer system to identify and react to events. In the context of this project, machine learning will be utilized to identify data trends that cause stress in the user. Once identified, the system can learn to warn users of similar data trends in the future.

In 2005, a paper was published in the *Engineering in Medicine and Biology Society* using heart rate variability and skin response with a neural network to identify human emotional responses with a finaly result of 80.2% accuracy. In the paper, a Multilayer Perceptron (MLP) neural network structure was used. [6]

B. Criteria

1) Speed is important. If the system runs too slowly, it may not be able to keep up with the provided data or take far too long to train the system. If it doesn't track trends quickly enough, it will not be able to provide forewarning.
2) Accuracy is a must. If the system tracks information incorrectly, it may give warnings without warrant.
3) Machine learning has the capability to do amazing things such as driving or optical image recognition. The usage initially in our project will be nowhere near that level of complexity, so the framework needs to be easy enough to use without being bogged down by unneeded functionality.

C. MXNet

MXNet is a deep learning library designed by the Apache Software Foundation. It has APIs for Python, Scala, R, Julia, C++, and Perl. It offers installers for the typical computer operating systems as well as the Raspberry Pi and NVIDIA Jetson TX2. [7]

D. TensorFlow

TensorFlow was originally designed for internal use at Google to detect patterns in a similar way to the human brain. It was released under an open source license in 2015. [**?**] Installers are available for all flavors of computer operating systems as well as Android and iOS. TensorFlow scales across multiple processing units. There are official APIs for Python, C++, Java, and Go. [8] [9]

E. Caffe2

Caffe was originally developed in the Berkeley Vision and Learning center. Facebook developed and released Caffe2 in April of 2017. It has installers for all the standard computer and mobile device operating systems, as well as Raspberry Pi and Tegra devices. There are APIs for C++ and Python.

F. Comparison

All three frameworks support the three primary operating systems, Windows, Mac OS, and Linux. TensorFlow and Caffe2 both support running on mobile devices as well, which while not currently something necessary, having the support could potentially create oportunities later down the road. All three of the frameworks support functionality for creating and utilizing an MLP, which has been shown to be effective for identifying emotional responses. MXNet and TensorFlow feature open source code for using it, which would decrease the learning curve in setting it up. All three frameworks feature extensive documention, examples, and community tutorials. TensorFlow and MXNet both have full implementation examples using MLP. While Caffe2 supports it, there is not a ready made implementation of MLP with it. Testing by Jasmeet Bhatia, a Data and AI Solution Architect for Microsoft, showed MXNet performing MLP operations in 56% of the time compared to TensorFlow. [10] MLP examples and documentation give TensorFlow and MXNet a small edge compared to Caffe2. The efficiency testing by Jasmeet adds to that of MXNet due to the speed over TensorFlow. For these reasons, we will be utilizing MXNet for the machine learning framework.

REFERENCES

[1] "Microsoft hololens — the leader in mixed reality technology." https://www.microsoft.com/en-us/hololens.
[2] "Vive — discover virtual reality beyond imagination." https://www.vive.com/us/.
[3] "Oculus rift." https://www.oculus.com/rift/.
[4] "Amazon web services (aws) - cloud computing services." https://aws.amazon.com.
[5] "Google cloud platform." https://cloud.google.com/.
[6] C. Lee, S. Yoo, Y. Park, N. Kim, K. Jeong, and B. Lee, "Using neural network to recognize human emotions from heart rate variability and skin resistance," vol. 5, pp. 5523–5, 02 2005.
[7] "Mxnet: A scalable deep learning framework." https://mxnet.incubator.apache.org.
[8] "Api documentation — tensorflow." https://www.tensorflow.org/api_docs/.
[9] E. Friedman, "Tensorflow, mxnet, caffe, h2o - which machine learning tool is best?," Aug 2017.
[10] J. Bhatia, "Search for the fastest deep learning framework supported by keras," Sep 2017.