

https://dspace.org/

Super vision: an investigation into the potential use of a player's face, using computer vision, as an alternative to traditional video game controllers

Suleman, Farouk

2021
https://hdl.handle.net/10539/32116

Downloaded from a DSpace repository. https://dspace.org/

MA Digital Arts Research Essay

Author Farouk Suleman, 765795 Supervisors Raymond Whitcher Ethics Clearance: H19/05/29

WITS UNIVERSITY



PLAGIARISM DECLARATION

1. I know that plagiarism means taking and using the ideas, writings, works or inventions of another as if they were one's own. I know that plagiarism not only includes verbatim copying, but also the extensive use of another person's ideas without proper acknowledgement (which includes the proper use of quotation marks). I know that plagiarism covers this sort of use of material found in textual sources and from the Internet.

2. I acknowledge and understand that plagiarism is wrong.

3. I understand that my research must be accurately referenced. I have followed the rules and conventions concerning referencing, citation and the use of quotations as set out in the Departmental Guide.

4. This assignment is my own work, or my group's own unique group assignment. I acknowledge that copying someone else's assignment, or part of it, is wrong, and that submitting identical work to others constitutes a form of plagiarism. 5. I have not allowed, nor will I in the future allow anyone to copy my work with the intention of passing it off as their own work.

Name……

Student #.....

Signed …….

Date …….

Table of Contents

Abstract 5

Title 6

Section 1 6

1.1 Introduction 6

1.2 Terminology and Definitions 8

1.2.1 Real-Time 8

1.2.2 Computer Vision 9

1.3 Aim 9

1.4 Rationale 10

1.5 Approach 13

1.6 Related Works 14

1.7 Risks 14

1.8 History of Computer Vision in Gaming 16

1.8.1 Nintendo Game Boy Camera (1998) 16

1.8.2 Sony EyeToy in (2003) 17

1.8.3 Xbox LIVE Vision (2006) 17

1.8.4 Sony Go!Cam (2007) 18

1.8.5 Sony PlayStation Eye (2007) 18

1.8.6 Nintendo Wii Camera (2009) 19

1.8.7 Xbox 360 Kinect (2010) 19

1.8.8 Xbox One Kinect (2013) 20

1.8.9 PlayStation 4 Camera (2013) 20

1.8.10 Smartphones 20

1.9 Types of Cameras 21

1.9.1 Monocular (The Camera I will be using for the creative Practice) 21

1.9.2 Stereo 21

1.10 Theory Part 1 21

1.10.1 Literature Review 21

1.11 Theory Part 25

1.11.1 Creating Haar Cascades 25

1.11.2 Creating Convolutional Neural Network (CNN) 27

Section 2 29
2.1 Creative Output – Application of Theory 29
2.1.1 Music 29
2.1.2 Game Feel 29
2.1.3 Ethical Considerations 33
Section 3 35
3.1 Method 35
3.2 Technology Tools 36
3.3 Creative Practice 38
Section 4 46
4.1. Playtesting 46
4.2 Feedback 46
4.3 How to Use 47
Section 5 48
5.1 Conclusion and Discussion 48
References 51
Image List 58
Appendix 1 60
Facial Feature points 60
Appendix 2 61
Some Screenshots from The First and Second Games 61
Appendix 3 62
Volunteer permission letter 62
Appendix 4 65
Ethics Form 65
Appendix 5 66
Interview Questions 66

Abstract

Video Games has created its own community and are played by millions of people every day on a global scale. Usually, these games are played through mainstream peripherals, relegated to standard controllers, or mouse and keyboard. However, newer forms of control are coming to light especially with the massive proliferation of smartphones becoming more powerful and widely consumed.

The camera, in my opinion, can be one of the best peripherals used to control a video game as there is no external device between the player and the game. The player would simply complete an action, and a reaction in the game will occur. This effect may be even greater if the player could use even less of their body to interact with the game. Furthermore, in my opinion, using the camera and facial gestures as input to games, could have a wider appeal to both mainstream functioning and motor impaired players. Unfortunately, many people with motor impairments are often excluded from this community due to the nature of the control mechanisms used to play these games. (Scardovelli, Terigi & Frère, A. (2014) [43])

Thus, my research aims to test if the use of a camera and facial gestures, can be utilised as a viable input mechanism to the gaming community and the impact on their playing experience. In addition, my research highlights the benefits of camera and facial gestures when used by people with motor impairments. This paper serves as an investigation into the feasibility of using camera peripherals as an alternative means to traditional video game controllers.

My research method was based on a qualitative study of volunteers testing the game by using their face as an input device. I chose a qualitative study due to the uniqueness of each person's face. Furthermore, I tested users from every demographic background to determine the usability and experience of the game play. My sample of volunteers included a person with a physical impairment. The result of this exercise showed the effectiveness of facial gestures with the use of a camera in playing simple games with a strong bias towards mobile games.

In my conclusion, it was evident from the study that this form of input mechanism creates a unique gaming experience for the player, one that he/she would not have experienced using an external controller. However, this form of input would require more development and possibly used in conjunction with other input methods for competitive gaming.

Title

Super Vision: An investigation into the potential use of a player's face, using computer vision, as an alternative to traditional video game controllers.

Section 1

1.1 Introduction

Current traditional video game peripherals (mouse, keyboard, controllers, etc.) require many hand and wrist movements, and though it is now a norm and players can do it without thinking, I feel that there are other tools available to enhance the gaming experience. For movements to become automatic while playing a game, new players need to practice for hours (Razbuten (2019) [44]). This is further exacerbated by the different game design control conventions. For example, if a game player is a *Tekken (Bandai Namco (1994) [50])* player the game control is different to that of *Mortal Kombat (NetherRealm Studios (1992) [51])*. Each game has different buttons and movements making it difficult for players to master every form of game. As mentioned above, the use of traditional periphery devices excludes many players with motor impairments from this form of sport. The use of facial gestures will allow a far more accessible means of control. The ability to interact with a computer game without the requirement of a traditional game controller device, or using full body gestures, such as with the Xbox Kinect, provides a player with the unique ability to use minimum effort facial gestures to directly interact with the game they are playing.

This research is concerned with real-time human-computer interaction for computer games using a single camera webcam system to allow a player to interact directly with a video game in the same manner as they would with body gestures, however only utilizing the players face. This will be expanded upon later in the paper.

Video Games are one of the many means of entertainment consumed by people daily. This has been moved from solely entertainment to enhancing education, for example, the use of children's games to help with reading and comprehension, mathematics, etc. As such video games can be found everywhere from small flash games[1] to massive AAA Games[2] being played by an immense amount of people all

[1] Small usually browser-based games that do not require many resources to play. Installation of the game is not needed.

[2] Pronounced Triple-A is a term used for games with a publisher usually having high development and marketing budgets.

around the world. Some of these games are geared to be simple and whimsical, while others are designed to be more serious, violent, or even controversial. No matter the type of game, it usually attempts to invoke feelings in its players and viewers.

In my opinion, video games are a very emotional art form. In an article on *Psychology Today (Video Games Impact Your Feelings in Real Life,* 2015[37]) the author states where immoral actions done in video games causes a feeling of guilt, players become more sensitive to similar situations in real life. Thus, making the use of video games in therapy for potential future psychological interventions. In my opinion, video games generally try to trigger emotions that can range from comedic to the moribund, depending on what the game designers and developers want the player to experience[3]. (Oswald, Christopher & Prorock, Chris & Murphy, Shane. (2014) [45])

Some games would allow users to bring their own meaning and understanding to the game based on the player's own game experiences such as *Journey* (Thatgamecompany, 2012 [21]). This game has no storyline and allows the player to interpret the game in their own way. Creating a sense of profound depth of meaning that is uniquely personal. Other games would much rather guide a player along with a linear story taking the player on a journey of discovery. Games have a unique ability to invoke imagination through immersing the player in a fictional realm. Each of these games provides players opportunities to learn and participate in a community.

In my opinion, experiences are what is most important in games, thus, how a player encounters the game is also exceptionally important. On the same note, the method of how a player interacts with a game will also affect the experience in the game. As mentioned earlier, most games use the standard control methods (Mouse and Keyboard or controller). However, in certain game types, other methods of control can enhance how the player interacts with the game. The game *RainBrow* (Nathan, 2017[22]) on the iPhone X device has pioneered the way forward using facial gestures to differentiate the player experience. The game uses the players' eyebrows to jump to different colours of the rainbow. *Nose Zone* (Hatchlings, Inc [48]) & *Rainbrow* are early games that utilize the iPhone X's TrueDepth sensor for controls, (Chance Miller, Dec. 18th, 2017, [8]) the whole appeal of the game is that you are playing it

[3] Experiences - an encounter, observation which leaves and impression on someone.

with your eyebrows. What makes it appealing is the fact that it is new and different and offers players a different way to play a game than simply using their fingers.



Figure 1: The RainBrow Interface (Nathan, 2017[22])

1.2 Terminology and Definitions

1.2.1 Real-Time

The definition of real-time as defined in The Oxford Dictionary is, real-time relates to a system in which input data is processed within milliseconds so that it is available virtually immediately as feedback to the process from which it is coming, e.g., in a missile guidance system. (Oxford,2019 [23]) Real-Time is a term used to describe and compare the performance and practicality of using different algorithms, in this case for a computer vision-based system. It is describing processing at a fast enough speed that allows the user a sufficient level of interaction for the interface.

In the gaming world, this is of utmost importance as lack of processing speed may make the game feel unresponsive, and can lead to a frustrating game experience, which is unacceptable as ultimately, in my opinion, the purpose of a game is to bring entertainment to the players. Frustration impedes on this effect. In Game Feel, this is describing response. According *to Game Feel: The Secret Ingredient* (Steve Swink , [39]) response is how the system processes, modifies, and responds to player input in real time. This affects the games feel. Bad response will negatively impact a games enjoyment. As stated in the paper *Playing Video Games as Entertainment*. [Vorderer, P., Bryant, J., Pieper, K. M., & Weber, R. (2006), 38] Video games have surpassed the designation of "fad" or "new technology" to become a staple of contemporary entertainment. Respondents overwhelmingly indicated that when forced to choose between video games and other media, they would rather give up television (73.6%) or movies (69.3%).

1.2.2 Computer Vision

Computer vision is a form of artificial intelligence where computers can "see" the world, analyze visual data, and then gain an understanding about the environment and situation. (Calo, 2019 [24]) Computer Vision applies to pictures as well as videos. It is useful in a variety of fields, from the medical examination of images such as an x-ray, to customer service, i.e., by determining customer satisfaction via their facial expressions. Computer vision systems can detect if a customer is happy with the service they are receiving.

In the gaming industry, though, it has mostly been used as full-body real-time videos, i.e., the Microsoft Xbox Kinect peripheral or the Sony PlayStation Eyetoy with later adaptations over newer consoles including the PS Move for the PlayStation 4.

In my research I will focus solely on the face as an input device, to test the viability of using the face and facial gestures to control the game. The viability of the technology will be based on the following key criteria.

1. Responsiveness[4] - The level of responsiveness required depends on the system in mind. This was tested through playtesting and improved upon incrementally leveraging a few different technologies.
2. Player Experience[5] - The player tested the system and feedback was gained through using the survey asking players if they enjoyed the experience.
3. Practicality – Type of games and multiple devices. - Through testing and observation, I found that many players could not utilize their face to make gestures as easily as others. This was tested by observing how players interacted with the games. As such, for this input method to be practical for more complex games, it would need supplementary input devices to keep the gestures simple.

Once people grow accustomed to the idea of controlling systems with facial gestures, it may even become more convenient. For more complex games, it could supplement the overall experience by engaging the player more, thus inducing more immersion.

1.3 Aim

The aim of this paper is to explore the notion of Computer Vision in video games, and how it can be used to analyze a player's facial features to control actions based on their facial expressions, or movements of

[4] Quick and efficient reaction to a scenario. In my case player facial gestures.
[5] View of how the player interacts and experiences the system.

the player's in-game avatar. As things stand, there are apps that allow for user interaction via these methods, for example, Snapchat filters, Facebook picture filters, etc. However, although these allow interactions via facial tracking, these apps do not take it one step further, and implement these features as controls into a game (except for the new iPhone X games and Snapchat games). My intention is to demonstrate through creative practice, a fun, engaging game using a human player's face as the primary controller, and use a test group of players to test the system to determine if this is a viable and exciting method of controlling a video game, or if it is just a novelty system.

This is important as previously the use of this technology has always been to augment the user experience, and not the primary mode of input or interaction. As such I am leveraging this technology to create a game that uses facial gestures as the sole means of interaction. In doing this I would like to test how players interact with the system.

I plan to achieve this by focusing on designing a simple game, that apart from the navigation system, uses a player's face as the primary input device to interact with the game world. This game will then be play-tested to investigate how players react to this method of interacting with video games. Though the focus will be more on the technology than the actual game.

1.4 Rationale

As a semi-professional gamer, I find that game controllers, periphery and input devices are key to being competitive and enhancing the game experience. It becomes an extension of the gamers' personality. Gamers, game development companies and technology companies are continually pushing the limits to immerse the gamer further into the game. For example, game controllers have evolved from the first-generation input devices i.e., the original keyboards, ball mouses, arcade sticks, etc., to the controllers we have today with vibration and actual trigger mechanisms based on the type of weapon you use such as the PlayStation 5 controller.

This is further enhanced by other periphery devices such as screens with higher frame refreshes, fighting sticks, chairs, and gaming rigs. Each of these devices hope to provide a competitive edge and enhance the gamers experience as mentioned above. This can be seen in the paper written by Birk and Mandryk (*Usability Testing of Video Game Controllers: A Case Study (2013)* [46]) which states "the performance of a user while playing a game can also be strongly influenced by the type of controller used".na

My rationale for undertaking this project was to see if the technology could remove a traditional input device, replacing it with the gamer directly as an input mechanism, specifically the face. I chose the face to see whether the performance of the technology would be able to be used effectively in the gaming experience. From an incredibly young age, technology always fascinated me. This passion was ignited when I received my first game console at eight years old. When I found out that games could be programmed, I knew from that age that this was what I wanted to do. Later during my undergraduate years, I learnt about Computer Vision, as defined earlier in the terminology section 1.2.2. Again, I was really taken using technology to solve real world problems. I chose to use my two areas of passion which is Computer Vision and gaming, to develop this research paper. In my opinion, the face is not just a mechanism for identity, but a key communicator of expression as it forms part of nonverbal cues. With that said, I think that this would be an ideal mechanism to unlock fun and entertainment by using your face and facial gestures. This led to me to investigate the viability of this technology in the gaming field. An additional motivation is having a brother that is Autistic, who struggles to read facial cues and struggles with playing games. His passion to play games and to be part of the gaming community is visibly noticeable. If this technology becomes mainstream, he will have more access and opportunity to learn about the world of gaming and integrate himself within this unique community.

In testing this technology, my primary motivation was to test the performance in a game setting, through analysing the responsiveness of facial gestures during the game play. The performance included lag times, player experience of the technology and practicality of this input mechanism for gaming. An additional motivator would be using this means of technology if viable, for motor impaired players to compete competitively against abled bodied players. In this way it would become evident whether the technology would be capable of processing complex instructions quickly and efficiently and if there would be any system lag (delays in processing). It would also assist in foreseeing any possible challenges using this technology that may be encountered by game designers and developers, as people have unique facial features, skin tones, textures, and expressions.

This paper explores how Computer Vision would fare in video games when not focused on big body movement and hand gestures, but solely on a player's facial gestures. This stems from my interest in exploring alternate methods of player control in the gaming environment, beyond the conventional hand-held controller or even augmented body mapping (The way the body is tracked). This requires a system to be far more sensitive, as the over exaggerated motions (air punching, jumping, etc.) will not be present in place of subtler facial movements.

As stated by the aim, the system designed for the Creative Practice will be used to analyze a player's facial features to control actions based on the facial expressions or movements of a player. The system will be play tested to observe how players interact with the game and how the system is able to react to different player's facial features and movements, which would differ far more than if general body gestures where used.

In the thesis *Computer Vision Based Interfaces for Computer Games* it is stated that there is a lot of interest in using this technology as an alternative to traditional game controller devices (Jonathan Rihan, 2010 [4]). The vast majority is still currently focused on body movement. In contrast, this paper and creative practice aims to show the potential of facial based computer vision in the gaming industry.

Personally, I feel like there is a vast amount of unlocked potential other than simply using facial gestures on a character to mimic the player. Aspects where facial gestures can actually be used as input or affect the virtual world for example, think of a horror game that detects facial expressions to determine when a next scare should happen or in a narrative style role playing game(RPG)[6], in-game characters behaving differently towards a player based on the players expressions, for instance, should the player have a scowl, NPCs will avoid them. Thus, controlling specific experiences in the game based off a player facial gesture. In the creative practice the whole game is controlled using only facial gestures. Even though the creative practice uses multiple simple mini games, there is potential to do the same in larger scaled games.

Facial gestures have applications in systems in many industries not solely the Gaming Industry. What the Creative Practice accompanying this research paper will attempt to prove, is that the use of facial gestures can be a fun, enjoyable experience allowing players to control a game using only their face. This will hopefully display some of the potentials of using facial gestures as input.

In the future, this could be used in conjunction with traditional controls to create more immersive experiences by using more subtle, natural facial gestures. This could also be used in the education field in the form of educational games for example, especially for special needs children who cannot learn using traditional learning methods and need other forms of input and interactions in the classroom. It could also be useful for people with Autism, teaching and enabling them to understand facial cues and to express themselves. This can also be used for people who are physically disabled and cannot use all

[6] Rpg-Defined as a role-playing game where the player assumes roles of characters in a fictional setting e.g. becoming *dovahkiin* in *Skyrim* (Bethesda, 2011 [53])

the parts of their bodies for traditional or full body controls. As this technology advances it could perhaps even allow disabled and able-bodied people to compete in esports[7] tournaments allowing both traditional and nontraditional methods of control.

Disabled people do often have means of control using an analysis method that uses their eyes as input. However, due to the number of gestures that can be performed, if a person can move their face these technologies could be combined to give an extra level of control. (F. Iacopetti, L. Fanucci, R. Roncella, D. Giusti and A. Scebba, 2008 pp. 757-762 [47]). Another area where facial gesture detection can be used is within hospitals. Should a patient be in pain and require assistance, a nurse could be alerted and thus the patient can receive help more efficiently.

1.5 Approach

The research in this essay considers two approaches to the problem of detecting the facial gestures of a player. Both approaches utilize feature-based methods of detection as will be explained in greater detail later in the paper.

The first approach presents an algorithm that uses Haar Cascades[8] for the detection method since its speed and efficiency is far beyond what other methods can achieve. However, it is susceptible to loss of a target if the target being tracked moves too quickly. It is also more sensitive to distance and lighting than the second approach being used. This approach uses a face cascade to detect and track a player's face and a smile cascade to detect the smile gesture.

The second approach presents an algorithm that uses Convolutional Neural Network (CNN)[9] and Dlib[10] for the detection method because it's reliable and that the Dlib library can handle facial tracking in a more superior way than the Haar Cascade method. A CNN is then created using the facial gestures required. This method is slower than the first approach, however, better for detecting at further distances as well as better for adding more gestures that are needed for detection, as fewer samples are needed to create a detection system. Dlib is more reliable that Haar Cascades because it maps out points

[7] Esports is a form of competition involving video games.
[8] As defined by Will Berger (Berge, [9]) Haar Cascades are machine learning object detection algorithms used to identify objects in an image or video and based on the concept of features proposed by Paul Viola and Michael Jones.
[9] A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. (Sumit Sasha, 2019 [25])
[10] Dlib is a modern C++ toolkit containing machine learning algorithms and tools for creating complex software in C++ to solve real world problems. (http://dlib.net/ , 2019 [26])

on the face and uses the different starting and ending points on the face to determine the facial gestures. Haar Cascading only matches the faces to trained cascades, like a template, but with hundreds of faces trained into.

For the creative practice only one method is used in the game. While initially finding the CNN method far too slow to be used practically on my system; I have dramatically improved the accuracy and processing speed of the CNN by combining both Haar Cascading and CNN. The Haar cascade quickly detects faces and the CNN then analyzes the gestures made by the face using an open-source detection model trained on fer2013 [11]dataset.

1.6 Related Works

Advertising Liking Detection System – A system that plays a video advertisement and detects facial gestures of a viewer for the duration of the video. After the video has been watched the system uses a ratio depending on the length of the video and the number of smiles a user made during the duration of the video, to determine if the advertisement was liked or not by the user. A lot of this work is relevant to facial gestures for video games as there are certain similarities, for instance, the real – time facial gesture detection. However, for games, the gestures need to be read even faster depending on the type of game. This was part of my motivation when doing the Advertising liking detection system as I could identify its potential in certain aspects of the business and gaming world alike.

Other Computer Vision based systems I have previously created are:

- ID number Detection – Detects and old ID Book and get the ID number from the book.
- HSV Tracking System (Hue Saturation Value) – Tracking an object from a user selected ROI (Region of Interest) using and objects HSV.

1.7 Risks

These were some of the risks that have hindered the initial progress of the project. Some of these risks have materialized, which were resolved, and others would require further analysis as they could not be determined in time for the analysis and customer tests.

[11] fer2013 is an open-source dataset which was created by Pierre-Luc Carrier and Aaron Courville. This dataset consists of 35.887 grayscale, 48x48 sized face images.

• Inability to get required data - Not being able to get access to the facial gesture database showing user emotions. I tried to get the data through usage of the GENKI[12] database. This risk materialised, to mitigate the risk I manually collected images to train the system from image search engines and image libraries. This was not enough though, thus making the system slow so I utilized and open-source CNN with the fer2013 dataset.

• Technical Problems Risk - This entails Data Loss, Corruption, PC problems, etc. To counter act these, the only course of action would be to make backups of the project. There was no data loss encountered during the trials. This risk did not materialise.

• Health Problems Risk- Unforeseen Problems such as illness, or injury. There is no way to counteract this risk. We did not experience any problems caused using this technology. It must be noted that the study was over a truly short time and research needs to be undertaken as to the long-term potential impacts of using this technology.

• Failure to produce fully operational system – This risk materialised due to the complexity of integrating the computer vision technology to the gaming software. I had to make several adjustments in the computer vision technology to improve system performance. This took longer than expected and delayed my system tests with the users.

• Run Times – This risk materialised. The system gathers data from viewers in real time, it is continuously polling data inputs from the users face. This led to the system hanging, crashing and errors etc. This also led to extremely long processing times. To mitigate this risk, I used a hybrid of a CNN and Haar-Cascade to improve overall performance.

• Falling Behind on schedule – This risk materialised because I needed to get ethics clearance which took some time as well as the technical challenges from the technology integration of the computer vision and gaming code. This risk resulted in additional time being allocated to compete the user testing and had an impact on the overall time to complete this project.

[12] An expanding database of images containing faces spanning a wide range of illumination conditions, geographical locations, personal identity, and ethnicity.

• Inability to find willing players to test system – This risk did not materialise. The venue that I had access to had sufficient test subjects willing to participate in the test. However, due to Covid-19 the tests had to be undertaken at the players homes as the building was closed.

• Some users may be affected with epilepsy or vertigo – I have advised people who are affected by these not to participate in the project. If a participant does experience any discomfort the system can be stopped at any time.

1.8 History of Computer Vision in Gaming

This section talks about some of the evolutions of the technology made by some of the pioneering entertainment companies in the gaming industry such as Sony, Microsoft, and Nintendo. Though not all the systems use facial gesture detection, many of them still use CV which is preceded by gesture detection systems. As such I feel it is still important to mention these systems.

1.8.1 Nintendo Game Boy Camera (1998)

The earliest Camera used in gaming that I have found was back in 1996 by Nintendo. This camera was used with Nintendo's Game Boy by using low-level vision image processing algorithms implemented on a hardware chip to try and achieve Real-Time Human–computer interaction. I have found that no games were made to use the camera and it instead created means to take photographs and apply filter effects. Due to the fact that filters could be applied, I have added this camera as it was one of the first systems to be able to utilize this functionality, and served as the pioneer for CV in gaming and other mediums, leading to where it is now. Sega also released a camera known as the Dreamcast Dreameye Camera in 2000 in Japan only. This camera was used mainly for video recording and editing images and not primarily for gaming. (Nintendo.fandom [27])



Figure 2: Nintendo Gameboy Camera

(https://www.dpreview.com/articles/1838149063/throwback-thursday-gameboy-camera, [61])

1.8.2 Sony EyeToy in (2003)

Sony Eye toy was the first suitable camera used as an input device in gaming. This camera allowed players to interact with games using motion, color detection, and also sound, through its built-in microphone. This camera had a wide variety of games that both used the camera as a sole means of input as well as simply enhanced other games such as EyeToy: Operation Spy, EyeToy: Monkey Mania, etc. (Bernadette Johnson, 2014 [28]). The Eye toy did allow for some gesture detection motions such as waving however this system was quite primitive and often blocking the camera would pick up as player input.



Figure 3: Sony EyeToy Camera (2003)

(https://www.ebay.com/itm/PS2-Playstation-2-official-Sony-EyeToy-USB-Camera-Eye-Toy-Camera-silver-/203005105108,[62])

1.8.3 Xbox LIVE Vision (2006)

Microsoft Released its first video camera for the Xbox360. Though it was mostly for live video feeds of players in online games there were a handful of games that used it for some input and some other games used it for facial mapping purposes known as Digimask (Tom Clancy's Rainbow Six: Vegas1 and 2, FIFA 10, etc.). (Wikmili, 2019 [29])



Figure 4: Xbox Live Vision Camera

(https://www.walmart.com/ip/xbox-360-live-vision-camera/5065209,[63])

1.8.4 Sony Go!Cam (2007)

A Sony PSP Camera that attached to the back of the camera. This camera used augmented reality and allowed players to scan their surroundings using the camera while the game placed items in various locations. These game items were then used by the player, for instance, an enemy could be placed in an area and the player would have to fight the creature. InviZimals is an example of a game that used this system. Players place a card known as a trap over a location where a creature is hiding to capture them. (Bernadette Johnson, 2014 [30])



Figure 5: Sony Go!Cam

(https://www.amazon.com/Official-SONY-PSP-Cam-450x-Camera/dp/B0047ERZ7Q,[64])

1.8.5 Sony PlayStation Eye (2007)

Released on the PlayStation 3 a year after the console was released. The camera was much more sensitive than the original EyeToy and was able to cope with lower light conditions. Due to the increased performance resulting from more processing power, more complex algorithms could be developed for both gesture recognition and augmented reality. This was generally used with a peripheral called the PlayStation Move. (Bernadette Johnson, 2014 [30]). Gesture detection became a console mainstay from the PlayStation 3 and Xbox 360 era.



Figure 6: Sony PlayStation EyeToy Camera (2007)

(https://www.amazon.com/Sony-Playstation-Move-Motion-Camera/dp/B0107WL3BA,[65])

1.8.6 Nintendo Wii Camera (2009)

This camera device was released for the console with a game called Your Shape a fitness game that utilized gesture recognition to establish if a player did the correct exercise. (Petrović, Ivetić 2013 [31])



Figure 7: Nintendo Wii Camera

(https://www.neogaf.com/threads/ubisofts-wii-camera-anyone-use-it-and-how-does-it-compare-to-kinect-and-such.1415935/, [66])

1.8.7 Xbox 360 Kinect (2010)

The Xbox 360 Kinect improved on its competitor Sony's PS3 Eyetoy by implementing multiple cameras allowing for depth perception, being the first console-based camera to utilize this. It used a system known as Light Coding which generated a grid from a 3D object's location using a variant of image-based 3D reconstruction. (Cong, Winters [32])



Figure 8: Xbox 360 Kinect Camera (2010)

(https://www.techradar.com/reviews/gaming/games-consoles/controllers/microsoft-kinect-for-xbox-360-905010/review,[67])

1.8.8 Xbox One Kinect (2013)

Originally designed to be an integral part of the Xbox One Experience thus was sold together with the Xbox One. This was changed later due it being poorly received. Utilized upgraded technology from its predecessor the Xbox 360 Kinect. (Bernadette Johnson, 2014 [30])



Figure 9: Xbox One Kinect (2013)

(https://www.ebay.com/p/9032778187, [68])

1.8.9 PlayStation 4 Camera (2013)

This was Sony's move to implement depth perception into their camera-based games thus allowing better body tracking than its predecessor. (Cooper, 2013 [33])



Figure 10: PlayStation 4 Camera (2013)

(https://www.bestbuy.com/site/sony-playstation-camera-for-playstation-4 new/5622702.p?skuId=5622702, [69])

1.8.10 Smartphones

Smartphone cameras and processing speeds have improved vastly, and filters have become a norm in the selfie taking world. Snapchat and Facebook have implemented Augmented Reality filters using face detection and gesture detection algorithms to determine where a user's face is and thus apply the filter

at specific locations on the users face. Recently there has also been the implementation of games, for instance, the snapchat games.



Figure 11: Smartphone Camera

(https://www.dxomark.com/the-best-smartphone-cameras/, [70])

1.9 Types of Cameras

1.9.1 Monocular (The Camera I will be using for the creative Practice)

A standard single lens webcam used commonly in most laptops though they also do come as standalone USB cameras. These cameras generally return the image or video data in the form of either Red Blue Green (RGB) colour, Greyscale or Bayer filter mosaic. The Bayer Filter mosaic is a color filter array for arranging RGB colour filters on a square grid of photo sensors. As mentioned, this is the camera I will be using for the creative practice as it is the camera available on my laptop. (Wikipedia [34])

1.9.2 Stereo

A stereo camera is a type of camera with multiple lenses with a separate image or video frame feed for each lens. This is done in to try and simulate the human binocular vision thus allowing image and video feed to be three dimensional. These cameras can retrieve depth as well as colour image and video data. An example would be the Xbox Kinect that utilizes multiple cameras for depth perception. (Wikipedia [35])

1.10 Theory Part 1

1.10.1 Literature Review

1.10.1.1 Facial Action Coding System

The Facial Action Coding System or FACS is a catalogue of unique action units relating to each independent movement of the face's muscles. Using this system, a computer vision system can code many of these actions automatically as described by Daniel McDuff, Rana El Kaliouby, Thibaud Senechal, David Demirdjian, and Rosalind Picard in the paper *Automatic measurement of ad preferences from facial responses gathered over the internet* (McDuff et al 2014, [3]).

1.10.1.2 Methods of detecting expressions

There are two major ways of detecting facial expressions each with its own pros and cons:

- The feature-based method
- The appearance-based method

1.10.1.2.1 The feature-based method

The feature-based method requires to find facial parts i.e., eyes, mouth, etc. Thus, allowing variations of the subject face positions and angles. However, because this method requires facial parts, if those parts are not found it cannot provide results as mentioned in the paper *Appearance-based smile intensity estimation by cascaded support vector machines (*Shimada et al (2010, pages 277–286,[6])).

1.10.1.2.2 The appearance-based method

The appearance-based method provides results without having to find facial parts but therefore is sensitive to the positioning and angles of the subjects face as mentioned in the paper *Appearance-based smile intensity estimation by cascaded support vector machines (*Shimada et al (2010, pages 277–286,[6])).

1.10.1.3 Intensity differences between pixels

One less common way of detecting an expression is detecting a facial movement by utilizing the intensity differences between pixels in grayscale images which are used as features as described in *Smile detection by boosting pixel differences* (Shan, 2012, page 431-436, [5]).



Figure 12: The top two rows depict smile and bottom two rows depict non-smile as depicted by (Shan, 2012, page 431-436, [5])

1.10.1.4 Data Collection

Real time data collection will be required as the purpose is to monitor players and utilize their facial movements and expressions as player input. In other words, data will be collected spontaneously from players utilizing a camera so that the system can decide in real time what facial expression a player is making. I considered using the GENKI database to train the system should there not be premade Haar Cascades of what I needed. This also eliminates the process of me having to get people in person to allow me to take a series of expression pictures. The GENKI contains pictures photographed by subjects from various locations in real-world conditions. Alternatively, I could always just manually collect a database of images to train my system on via image searching for expressions.

1.10.1.5 Data preparation

Due to the nature of the data, lots of the data may yield inconclusive results. As a result, I will most likely have to implement Adaboost which is a machine learning algorithm utilized to find the best classifiers with minimum error as described by Yu-Hao Huang and Chiou-Shann Fuh in *Face detection and smile detection* (Yu-Hao Huang and Chiou-Shann, 2009, [1]). Adaboost is one of the first boosting algorithms to be adapted in solving practices. Adaboost helps you combine multiple "weak classifiers" into a single "strong classifier**"**. AdaBoost works by putting more weight on difficult to classify instances and less on those already handled well. (Desarda , 2019 [54])

However, since I chose to use OpenCV, this problem was taken care of by its built-in functions. OpenCV, is an opensource library of Computer Vision modules that can read and write images from inception using computer code. This technology can capture, save, process and perform facial feature detection to determine various changes in facial gestures.

This technology uses a multitude of highly optimized computer vision algorithms that are developed in an open-source environment. The libraries allow algorithms that can be used to detect and recognize faces and facial gestures. I used OpenCV libraries with Haar Cascades and Convolutional neural network, defined in the footnotes 8 and 9 to optimize the performance of system. (opencv.org [55])

In essence, I leveraged the OpenCV libraries to improve the efficiency and effectiveness of the combined technical solution resulting in a higher rate or recognition and performance in terms of response times.

1.10.1.6 Feature Extraction

Face detectors do not ensure the accuracy of the subject's position and thus is susceptible to angles and position of the subject's face. As such the facial parts are not always detected properly. A face image can be divided into a grid and cells to extract local features for each cell. The final feature is built then concatenates all the local features. A local Intensity Histogram (LIH) and Center-Symmetric Local Binary Pattern (CS-LBP) can be used as local features as described by *(*Shimada et al (2010, pages 277–286,[6])).

Steps Taken for Local Intensity Histogram

- Divide face image into a matrix of M x N
- Build LIH with L bins for each cell
- Normalize LIH
- Concatenate Normalized LIH of all cells to form M x N x L dimensional vector

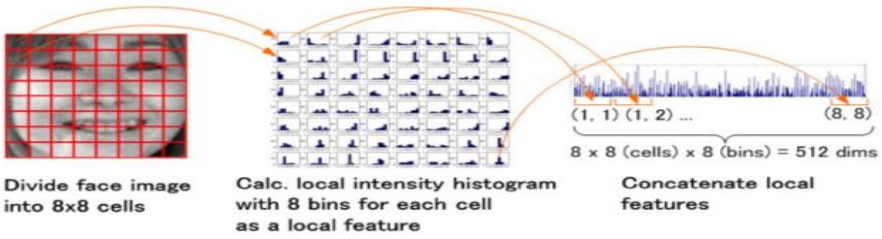


Figure 13: *(*Shimada et al (2010, pages 277–286,[6])). *Appearance-based smile intensity estimation by cascaded support vector machines*.

1.10.1.7 Center-Symmetric Local Binary Pattern

The CS-LBP can also be used to extract features and can represent texture information in more compact binary patterns.

The steps to achieve this are as follows:

- Divide face image into a matrix of M x N
- Calculate and build a CS-LSP for each cell
- Concatenate the CS-LBP of all calls to form M x N x 16 dimensional vector

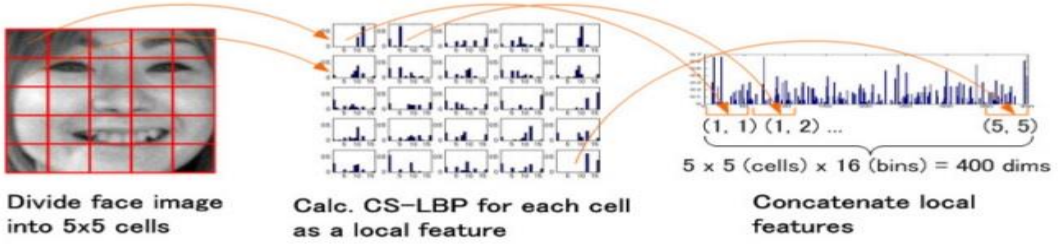


Figure 14: (Shimada et al (2010, pages 277–286,[6])). Appearance-based smile intensity estimation by cascaded support vector machines.

1.11 Theory Part

1.11.1 Creating Haar Cascades

Haar Cascades is an xml file that classifies positive and negative images to detect specific objects. Haar cascades uses machine learning to train the system using positive image samples of the object and negative image samples not containing the object, to identify the subject or Region of Interest (ROI) [13]. Positive image samples consist of the object you wish to detect, in my case, people with their facial expression neutral, smiling and lastly pulling tongue (This gesture was phased out so that more natural expressions such as surprised, and angry expressions could be used). (Patil, 2014 [56])

Negative Samples consist of images that does not contain the object you wish to detect. In my case anything other than a human face.

The Haar Cascade analyzes changes in pixel intensity on the image. Once trained the system will detect objects in the region of interest. For example, I trained the Haar Cascade to identify the various facial gestures by setting the region of interest, and through numerous examples of facial gesturing. The Haar Cascade then could use its Computer Vision to identify a facial gesture like a smile, frown, open mouth or closed mouth. In training the positive and negative images, I trained the system to identify what a face is and what a face is not. I did this through manually selecting images of faces and non-face objects to train the system. I used approximately 2000 images of faces and random household furniture and other appliances to train the system.

[13] ROI – A subset of samples within a dataset in this case a subset of an image of the camera video identifying a players face and mouth.

The system, then identifies the face as a region of interest. Once done there are sub regions of interest in the face. In essence it plots the face on a grid from the edge of the region of interest. Identifying specific areas on the face is done through identifying light and dark areas from the edge of the region of interest. It does this through identifying pixel intensity. To explain this through an example, a picture of the face will have dark areas and light areas. Areas which have hair are generally darker and areas which have skin are generally lighter. So, if we take a greyscale picture, the eyebrows are generally darker from the forehead and the area below the eyes. Similarly, the centre of the nose is lighter than the sides of the nose in a picture. This is how the Haar Cascade determines the sub regions of interest in the face. (Patil, 2014 [56])

Once this was done, I trained the system to identify facial gestures using the same method above.

1.11.1.1 Creating samples

A list is required for the positive and negative images to allow the software to determine where the images are, and where the object is in relation to the picture. OpenCV has built in functions (example opencv_createsamples) to assist with this task, but it is optimized for Linux, so it is recommended to run a Linux server if using windows or mac. The tool opencv_createsamples allows generation of samples from input images and returns *.vec files which is used to train the classifier. This tool distorts and transforms images to create even more positive images. Another tool by Naotoshi Seo called createsamples.pl combines positive images with random negative images to gain an even greater set of positive images to help with the classification.

1.11.1.2 Training the classifier

The Two main OpenCV tools used are opencv_haartraining and opencv_traincascade these are similar and accomplish the same task while the latter allows for multithreading[14]. Also, the formats generated differ depending on what is required. Once the tool is done it will return an .xml [15]file which will be used as the classifier in the system as shown in the article by (Thorsten Ball, 2013 [10]).

[14] The ability of a program to run multiple tasks simultaneously by generating multiple threads thus increasing speed and efficiency.
[15] Extensible Markup Language (XML) is a markup language used to store and transport data

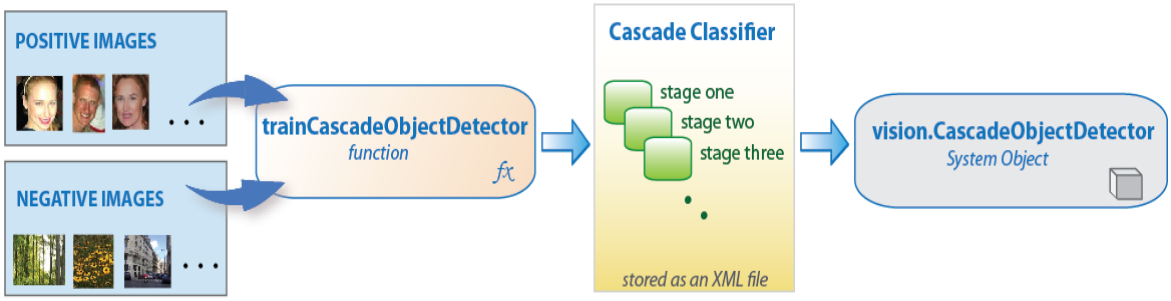


Figure 15: *Deep learning Haar Cascade explained* (Will Berger, [9])

1.11.2 Creating Convolutional Neural Network (CNN)

A CNN is a specialised Artificial Neural Network used to detect and identify patterns. It can interpret and understand the patterns it identifies. Therefore, this technology is ideal in image recognition. With regards to facial recognition, this technology is ideal because it has convolutional layers in the artificial neural network. These layers take in input from one layer and passes that to the convolutional layer of the network. The convolutional layer transforms the input and then passes this to the next layer of the network. This result is a feature map of the image. (Brownlee, 2019[57])

1.11.2.1 Data Collection

As mentioned previously to train a classifier, a wide variety of samples are needed however no negative samples are needed for the CNN method. This was achieved using Image Searches to get as many pictures as possible of people with neutral expressions, smiling and pulling tongue.

1.11.2.2 Data preparation

Once all the images are collected, they are first made to all be one specific size to keep everything constant and help with processing speeds so that the algorithm does not have to manually change all the sizes of the pictures at run time. The next step is to crop out all the images so that only the mouth is visible. This is again done for speeding up processing time as well as allowing the system to create a smaller Region of Interest thus easing the load on the classifier. The images are then sorted into folders and named. This helps with the labeling process as the algorithm sees the separate folders, and is able to determine which images are from which folder thus speeding up the training aspect and also saving the programmer from

having to manually label each item. Finally, the images are changed to Grey Scale (Black and White) as this helps with processing speeds as RGB is no longer factors in the classification.

1.11.2.3 Facial Feature Extraction

The algorithm uses the Dlib Library, that finds and tracks the face and mouth region. The Region of Interest (ROI) is then created around the user's mouth. The ROI is then cropped out in order to allow the CNN to categorize what the player's mouth is doing whether it be neutral, smiling or pulling tongue. An alternative to find the ROI is using Haar cascades instead. I found that this method was much faster and easier on my system. As such this was the method I used in the creative practice.

1.11.2.4 Categorization

The system determines the most likely gesture the player is making thus forcing an input allowing the game to be controlled.

A huge factor in this would probably be the lighting as this is a problem in all Computer Vision systems. Regardless, provided the object is not way too far away from the camera, the system should be able to track the face much better than Haar Cascading can, due to the Dlib Library facial tracking, at the cost of processing speed.

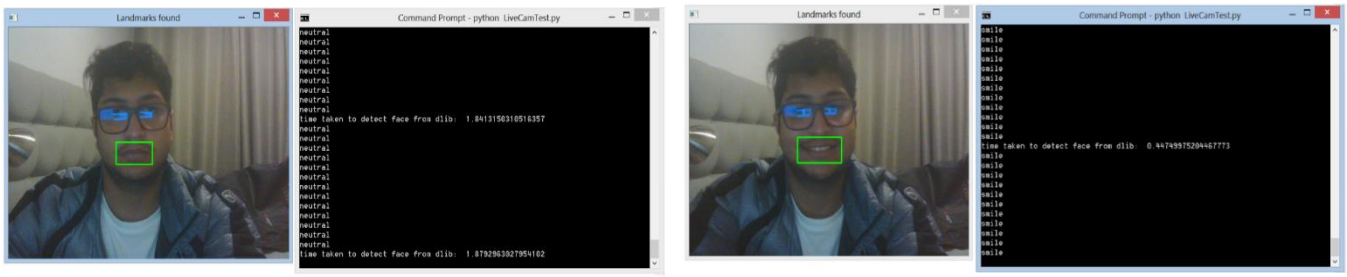


Figure 16 and 17: Depicting system detecting neutral and smiling gestures

Section 2

2.1 Creative Output – Application of Theory

2.1.1 Music

After doing some research into music for games I came across the term earworms also known as “brain worms”. This is described as sticky, catchy music that is designed to continuously run through the hearer’s mind. Generally, earworms have simple, often familiar sounding beats. (Pillay, 2017 [58])

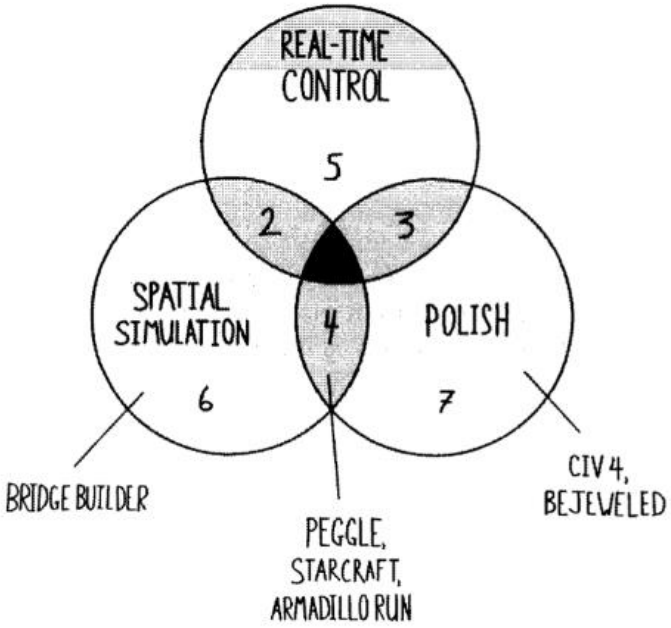
The simplicity helps the beat melody to get stored into a person’s memory and causes the continuous repetition to occur. The problem with earworms is that they often cause annoyance due to it constantly running through a person’s mind. In gaming, this is important as it “hooks” players. This is especially good to keep people coming back to play the game more.

In my opinion, an amazing example of an earworm used in a video game would be the Star Road theme song from Super Mario World. I decided I would like to use this tactic in my game however since I have no music experience I decided to turn to non-copyright music and came across a song called Feelin' Good. This song is a simple upbeat song that I feel meets the criteria to be considered an earworm. In the game currently in development, this song plays throughout the game continuously looping in order to help it get stuck into the players’ minds.

2.1.2 Game Feel

According to the book *Game Feel: A Game Designer’s Guide to Virtual Sensation Game Feel* Game Feel [16] the experience by players is built from real-time control, simulated space and polish.

[16] Game Feel or Game Juice is difficult to describe. It was popularized by a book called *Game Feel: A Game Designer’s Guide to Virtual Sensation* by Steve Swink. It is the intangible, tactile sensation experienced when interacting with video games.



Game Feel: A Game Designer's Guide to Virtual Sensation (Steve Swink, 201, [40]) Figure 18

In the book Swink talks about Super Mario and 64 and how players interact with the game and the world. This gave way to a methodology known as The Garden's Ecosystem. This methodology is used to separate the aspects of game feel in a more manageable form. *Game Feel: The Secret Ingredient* (Steve Swink,2020,[39])

- Input -- How the player can express their intent to the system.
- Response -- How the system processes, modifies, and responds to player input in real time.
- Context -- How constraints give spatial meaning to motion.
- Polish -- The interactive impression of physicality created by the harmony of animation, sounds, and effects with input-driven motion.
- Metaphor -- The ingredient that lends emotional meaning to motion and provides familiarity to mitigate learning frustration.
- Rules -- Application and tweaking of arbitrary variables that give additional challenge and higher-level meaning to motion and control.

2.1.2.1 Input

Input is the means by which a player controls and interacts with the game. This is the physical input device used by the player, usually traditional peripherals such as game controllers, mouse and keyboard etc., however in the case of this study the input is a camera, which uses facial gesture detection as input. One advantage of using traditional peripherals is the ability to naturally map game movement to controls that have been embedded into players over the years. In an fps[17] gamers know the default is usually right trigger to shoot left trigger to aim down sights.

With facial gestures there is no previous knowledge. This can lead to some awkward initial control. Over time, players can adjust, and should facial gesture control become a mainstay in the video game community, general "rules" will eventually be established to this form of control mapping.

In the context of the creative practice the facial gestures used are happy, angry and surprised. The first level contains players having to open their mouth signalling a surprise gesture causing the on-screen character to open their mouth at the same time. The player then needs to time the mosquito and eat it. When the player smiles indicating their mouth is now closed, the on-screen characters mouth closes with an expressive Nom-nom sound. Figure 19, 20, 21, 22 on page 40 depicts the actions and resulting outcomes in the game.

The second level uses a joke to hint the player to smile. Should the player smile within the time, the player will pass this stage. If a player does not smile in time, then the Mafia man will shoot the player.

In the third level the player needs to crush a bug. The player needs to time the bar in order to gain enough power to crush the bug. The input mapped to hitting the bug is smiling, thus a player needs to smile at the right moment to kill the bug otherwise the bug will catch the swatter and laugh at the player.

The fourth level has a player playing as a goblin trying to escape the dungeon. The player simply has to avoid the boulders by climbing and descending from ladders. This action is mapped to smiling once again as it is the fastest most natural expression players make.

The last level requires players to use all three expressions, as the mimic forces a player to copy them. The mimic will do a random expression and the player will have to emulate the expression namely smiling,

[17] First Person Shooter – A type of shooting game where the camera is in First person thus simulating real life scenarios such as the Call of Duty series.

angry and surprised expressions. If the expression is not done in time the player will die. If the player is successful the game is over, and the player will win.

2.1.2.2 Response

Response is how the game handles the player's inputs. With the nature of gesture detection there often can be false-positive inputs. This is where the system detects a player doing an expression the player is not actually doing. In order to counteract this, I used expressions that were not used in the game due to them not being easily replicated for everyone. I used fear and sad. Sad looks similar to angry thus when the system detects a sad expression, the system interprets an angry expression. With regards to fear it looks similar to surprise. Thus, when the system detects fear it registers the player as surprised. This tremendously improved game performance however the system is still not perfect and still receives false positives when lighting, distance and angles are bad.

2.1.2.3 Context

Context refers to the environment of the game. In essence the game world. In a platformer a player has the ability to move and jump and this should be reflected by the environment giving players reasons to utilize these mechanics. In the context of the creative practice, the worlds of the mini-games are very simplistic as this is not a traditional medium. Players would tire extremely quickly if the game incorporated more intricate control. The player does not really interact much with the environment except for the enemies and in the fourth level, the ladders.

2.1.2.4 Polish

Polish refers to the sound and graphics aspect of the game. This incorporates things like muzzle flash in an fps, dust particles in a platformer, screen shake, etc.

In the creative practice all the stages use screen shake. In the first level when the player munches, the screen shakes, and the character says "Nom-Nom" giving a more tactile feel to the stage.

If a player is killed on any of the levels there is also screen shake. On the second level should a player die, there will be an audio cue of a gunshot and a muzzle flash followed by a red flash in the screen which fades away and reveals blood splatter on the screen. The effect is amplified as when the gun is shot there is screen shake. On the other hand, if a player smiles there is a yellow flash with screen shake followed by a sizzling sound that fades away to reveal the mafia man burnt to a skeleton.

On the third level, should the player not gather enough power to kill the bug, there will be screen shake and a flash fading away to reveal the bug holding the swatter as he lets out a laugh audio clip. Should the player kill the bug there is a flash with screen shake followed by a squish sound which then reveals the bug to be dead.

The fourth level on death, there is screen shake with an “Oof” audio cue. The last level on death, has screen shake and a flash of red with blood splatter similar to the mafia stage. When the mimic makes a new expression, it lets out an audio cue depending on the expression.

2.1.2.5 Metaphor

Metaphor in a game refers to the theme of a game based on which players bring into the game, preconceived notions of how things should be and what should be done. In the context of the game though, it has some scenarios where the players will know what to do. For instance, dodge boulders and squash the bug, however it also contains scenarios that are not typical for a player such as eating a mosquito.

2.1.2.6 Rules

As stated earlier rules are the application and tweaking of arbitrary variables that give additional challenge and higher-level meaning to motion and control. These can be aspects such as score, time limit, etc. which are goals for the player to achieve.

The first level requires a player to eat 10 mosquitos. The second level has a timer. If the player does not smile in time, they lose. In the fourth level, the goblin has to survive for a specific duration, and in the last level, the player has to copy the mimic in a specific time or die. The player also has to copy the mimic a specific number of times.

2.1.3 Ethical Considerations

Ethical approval is required whenever we conduct any form of research with people be it playtesting or undertaking a survey on the topic. All ethical considerations were appraised regarding this project. The ethics process started by me applying for approval to the university Ethic Committee. This process involved me having to complete an application form to undertake my playtesting. In this form the initial project proposal was submitted with the relevant documentation such as the volunteer permission letter (Appendix 3) and the interview questions (Appendix 5). Once I received approval, I then undertook the playtesting process.

Before any of the testing took place, I undertook a process of explaining the reason for the research. The participants were invited to participate in the testing and all potential risks were communicated. Before testing began, I ensured that all participants understood how the testing will take place and informed of the risks. Participants were informed that if they have vertigo or suffer from epilepsy, they would not be able to participate in the research.

The participants were informed about how the interview would be evaluated, and how the data would be stored and processed to retain confidentiality. Each participant was informed that their participation in this process was entirely voluntary and that they would experience no benefits or disadvantages should they participate or withdraw at any stage. As the project was playtesting the game, should they experience any form of distress or discomfort they could stop immediately. In addition, if the participants had any questions post the test, they could contact me for further clarification.

Participants were carefully chosen from various ethnic groups with varying educational backgrounds as well as familiarity with the gaming industry. In addition, one participant had a visible motor impairment.

Finally, all participants were guaranteed full anonymity, privacy and confidentiality.

Section 3

3.1 Method

Since this research project is Creative Practice, my first priority was to get the base game done without any Computer Vision making sure to get rid of as many bugs as possible. Initially, I had planned to make a platformer[18] however when talking to peers and friends I realized that quite a few of them cannot actually move specific parts of their faces, for instance only one eyebrow. This prompted me to move away from the platformer towards an endless runner *Geometry Dash (RobTop Games (2013) [52])* style game, since only one movement would be required.

Due to the fact that a colleague stated players faces would tire, I decided to take a more narrative approach. This led me to the third option. A narrative game which is driven by choice. I later decided to steer away from the narrative game in favour of a more casual experience as the simplistic narrative style did not suit the purpose of my research i.e., demonstrating how games can be played using only a player's face. As such, I developed a game that used whimsical scenarios. This would allow a player to play out these scenarios using only their face.

Once the game was completed, playtesting commenced without CV in order to test the game and discover any bugs. Thereafter, I focused solely on the Computer Vision Aspect of the game. Following the completion of the CV system, playtesting with CV began.

In order to validate the results of my research, I undertook to use a qualitative study. This took place through various users interacting with the game and completing a structured questionnaire thereafter. During all of the tests I observed the user interaction with the game through how they undertook the various stages of the game. I observed facial expression, body language as well as conversations with the users whilst playing the game. During this time, I managed to ascertain additional data and information on the player group I tested. Data and information regarding the players experience and the ethnicity of the players.

I chose a qualitative study due to the fact that I needed to take various different people through the experience, and see the impact of different facial features and gestures on the performance of the technology, as well as question the players on their experience.

[18] A platform game where player runs continuously the stage while the level gets progressively getting more difficult

During this process I took a cross section of the population, using various different ethnic groups with varying academic backgrounds, as well as people who were familiar and unfamiliar to the gaming community. My sample size consisted of 15 people that undertook the test. Out of the study one tester had visible motor impairments.

Table 1 below depicts the breakdown of the sample size by ethnic group.

Indian	Coloured	Black	White	Other
4	6	4	1	0

Table 2 depicts the breakdown of the group by familiarity in the gaming industry.

Low	Medium	High
4	6	5

Questions were designed to gain insight into the overall experience of the technology. On observation, I could see the response times of the technology, and during conversation post the testing, I got a view into the applicability of this technology on a variety of game types and complexities. To elaborate further, the game test generated some interesting discussions post the game play, whereby users indicated that the technology is great for simplistic games and mobile games but would be difficult to use in multi-button complex games please see Appendix 5.

3.2 Technology Tools

As mentioned previously I chose to steer away from using Haar Cascades due to the fact that I could not gather enough images to make an accurate enough Cascade for the pulling tongue Gesture. The databases I looked at e.g., the GENKI database, for instance, was in video format and didn't have the required gestures. I managed to create my own Haar Cascade, however it was not detecting as expected and was plagued with false positives[19].

In addition, the Haar Cascades would lose player's faces easily if a player moved around too quickly. As such, I started afresh and designed a CNN (convolutional neural network, a network of layered

[19] A term when a prediction is predicted as positive but in actuality negative.

independent filters that allow for image and video processing). This allowed me to bypass having to create Haar Cascades and also implement a better Facial Landmark Detection System Library known as Dlib.

The problem then arose that my system was not fast enough to run this library efficiently, however out of all the methods I have attempted aside from Haar Cascades, Dlib is by far the fastest.

This led me to combine Haar Cascading with the CNN. This however did not initially produce accurate results. I was advised to create a system that could detect different emotional ranges. However, due to my personal technological constraints with the equipment available to me, it was not practical for me to use a system that could do this, even though I had an algorithm that could accomplish it. It was way too slow to even attempt to speed up the algorithm.

Eventually I chose to return to Haar Cascades and change the game structure from a narrative to a casual scenario-based game. This was done with the view to simplify the instruction sets to allow for a more suited gameplay style.

As mentioned above, I decided to create a third game aimed at more casual players. This game was to be lighthearted and funny. The CNN worked with it but the longer the CNN ran, the slower it became. This was because the game consisted of different little mini games as levels, it was exceptionally inefficient to have to restart the CNN at every stage.

Haar Cascades granted me access to the faster more efficient speed and by weighing up the pros and cons of each method, I decided it would be more beneficial for me to use the Haar Cascade method for the third game. Through continuous refinement and research into more efficient CNN methods, I have managed to successfully combine Haar Cascades and CNNs together. This is by far the most efficient system I have come up with. The CNN I developed and trained albeit efficient, resulted in slow responses that would have impacted the gameplay experience i.e., creating intermittent system hanging in the gesture detection system.

As such I have decided to use an open-source CNN trained on the fer2013 dataset to increase performance. This system uses the speed of Haar Cascades to quickly find and extract ROIs, then uses the CNN to analyze the facial gestures. It allows for a much smoother experience as well as allows for an increase in the number of usable gestures. Once the game was completed, I had it playtested to give me a view on how the game performed with multiple players, and to gauge the overall game play experience.

3.3 Creative Practice

Before the first step in facial detection and gesture recognition can take place, the required data must first be collected. As previously mentioned, I had planned on using premade Haar Cascades, however upon testing I learnt that I required more gestures. I then planned to use the GENKI database to create more Haar Cascades. This database did not contain the gestures I required so I decided to use Google Images and other Image based libraries to find pictures I could use for alternate gestures. This was due to the difficulty in finding willing candidates to make facial gestures to train the CV system.

Through this method, I was not able to collect enough images to create Haar Cascades that detected well enough for the narrative game. This resulted in me using CNN as a reduced number of images are required in the CNN algorithm as opposed to the Haar Cascades method.

Once the images were collected preparation of the data was required. Creating a new Haar Cascade entailed creating folders for both positive and negative images. Positive Images consists of the object you wish to detect, in my case people with their facial expression neutral, smiling and lastly, pulling tongue. Negative Images consist of images where the object you wish to detect is not represented in the picture. This is then used to create a list of where the positive and negative images are, using an OpenCV function called opencv_createsamples.

This is a long tedious process where the object has to be manually detected in each individual picture. This is then turned into samples using a function called opencv_createsamples which allows the generation of samples from input images and returns *.vec file. This file is then used to create the classifier using the opencv_traincascade function which creates the xml file used to classify gestures.

Unfortunately, this did not grant me sufficiently accurate readings to allow this to be used in a game. Sometimes it would detect the pulling tongue gesture but most of the time it did not. For the CNN method Data preparation required making all the images one size after which they are then cropped so that only the mouth region is visible. This is then sorted into folders and labeled appropriately depending on the gesture. This allows the learning algorithm to differentiate the pictures from the labels as it is still learning and cannot determine which gesture is which. Finally, the images are changed into grayscale which speeds up processing time.

The very first step in coding a facial gesture detection system is to detect a player's face. Without being able to determine where the Region of interest is, the system will not be able to determine any gestures made. The method used varies from what the system requires. For instance, in a game that required more

processing speed, I opted to use Haar Cascading to make the game feel responsive. As mentioned previously, I found it much faster than using Dlib. Thus, in the latest game I developed for the creative practice aspect of this project, I used this method.

Dlib though more reliable, uses numerous points mapped onto a player's face once detected. As such the longer the system runs the slower it becomes. The points mark different features on the face such as different parts of each eyebrow, parts of each eye, parts of the mouth, etc. As such for the latest game it was not efficient for me to continuously have the Computer Vision system restart for every level especially since I planned to make quite a few levels.

The next step in a facial gesture detection system is to detect the actual gesture the player is making. As aforementioned, there are three ways of detecting expressions. The feature-based method and the appearance-based method being the most prominent ways. Both the Haar Cascading method and the CNN method are part of the feature-based method. This method first requires the face to be detected before gestures can be recognized. If the system cannot find the face and determine the ROI (Region of Interest), in this case the mouth, a gesture cannot be determined.

By using the Haar Cascading method to detect the ROI, I have managed to speed up the CNN method significantly. This allows me to now use multiple gestures effectively. These include the neutral, smiling, angry and surprised gestures. By default, any action other than these are set to neutral. Hence if a player is, for instance, pulling their tongue, it will not bug out the game by reading in unknown inputs, it will instead just pick up as neutral.

The third step is to develop the actual game to be experienced by the player. It is important to determine the type of game you want to build beforehand, as this has an impact on the gesture detection system to be used. As mentioned before, for the fast-paced endless runner type games, I decided to use Haar Cascades as only one input was required, and the speed required to process inputs were fast.

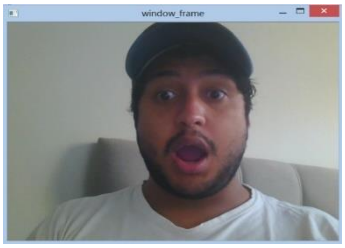
When I was advised to change to a narrative, due to my lack of good quality images to create an efficient Haar Cascade method, I decided to create a CNN gesture detection system.

Once advised to make a more "tongue-in-cheek" game and have fun with the technology, I decided to use the Haar Cascade method to develop a one button style game that uses absurd little scenarios or jokes as levels for players to experience. This would hopefully create a few laughs along the way.

As the game is a “One Button” game or in this case, a one gesture game, it is extremely simple with a similar gameplay style to the game *Dumb Ways to Die (Metrp Trains (2012))*. *Dumb Ways to Die* consists of a whole lot of mini-games and the purpose is to keep your character alive.

The game developed currently uses many different and strange scenarios. In the final build, I planned to have the levels randomized with lives and scores, but currently for demonstration purposes the game’s levels are set in an order.

1. The first of the levels is a scenario where there is a new virus carried by mosquitos. If a mosquito were to bite the character, it would mean the end of that level for the player as the character has died. The game hints with a text box that says “Say ahhh!” and an emoticon that makes a gesture of an open mouth closing with teeth showing. This is a prompt for the players to emulate this action. If done, the characters’ mouth will close when the players’ mouth closes thus eating the mosquito. In this scenario, should the mosquito bite you the virus will be transmitted into your blood stream leading to death; however, our saliva enzymes are enough to detain the virus until it reaches the stomach where the acid destroys it. The mouth action is opened by making a surprised face and closed once a player smiles.



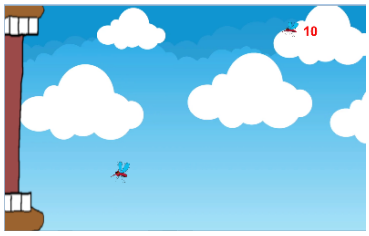


Figure 19 and 20: Image of a surprise gesture and its depiction in the game



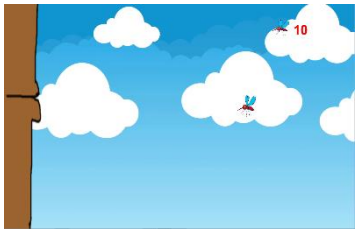


Figure 21 and 22: Image of a smile and its depiction in the game

The images above depict the action used to play this level of the game. Figure 19, 20, 21, 22 Depicting surprise and smile gestures with the outcomes in game.

2. The second level sees a man with a gun who is about to shoot you. The game uses a child's playground insult, to hint a player to smile at him. The insult being "Your teeth are so yellow you make the sun jealous!". During the playtesting process this almost always brought laughs to younger audiences when they saw it. Some of the testers could relate to the insult above when they were younger.

   This level has a timer where if a player does not smile within a specific time, the man shoots and kills the player. If the player smiles, then the man is burnt to a crisp with nothing but his skeleton remaining.

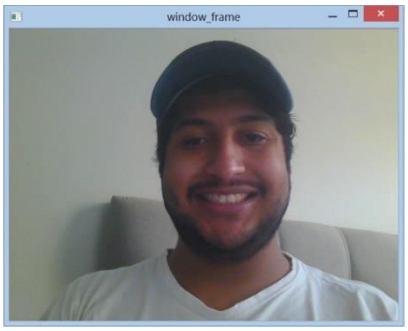


   Figure 23: The image above depicts the smile action required for this level.

3. The third level sees the player holding a fly swatter. On the side is a power gauge continuously filling and emptying again. The level takes place in a kitchen where on the table is a menacing looking bug. The player attempts to squash the bug with the fly swatter which is active using the smile gesture.

   If the power gauge has enough power the bug will be squashed, however, should there be too little power the bug will effortlessly catch and hold the fly swatter as he gives an evil smile back at the player. Figure 23 above indicates the motion required to play this level.

4. The next level seen in the game sees a goblin trying to escape a dungeon riddled with boulder traps. The only means of escape is to dodge the boulders by utilizing the ladders. When a player nears a ladder, should they smile, the player will ascend or descend the ladder depending on if they are above or below the ladder. Due to players struggling in playtesting I got advised to leave in an indicator of where the boulders would go thus helping players navigate using the ladders. Figure 23 above depicts the gesture.

5. The final level of the game leverages the CNN algorithm and Haar Cascade methods to allow for multiple gestures to be detected and used. This performance of the CNN algorithm was enhanced through using the open-source CNN dataset resulting in real time gesture detection.

   The scenario in this level creates a copycat type of scenario where the player has to copy the gestures of the mimic in order to survive. The gestures include smiling, surprised and angry. If the player does not copy the mimic, the player will die. However, if they copy the mimic successfully, they will have completed the game. I also decided to change the mechanics of the first level due to the new gestures available. Initially the level would only use the smile gesture. It now uses the surprise gesture to detect an open mouth as well as the smile gesture to detect a closed mouth simulating an eating motion.

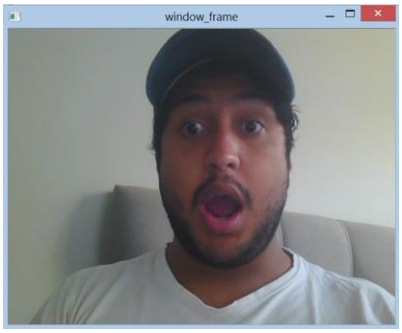




Figure 24 and 25: Images depicting a surprise gesture and how it looks in the game

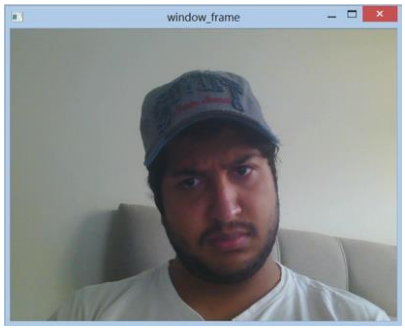




Figure 26 and 27: Images depicting an angry gesture and how it looks in the game

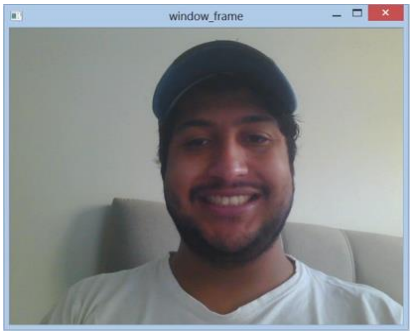




Figure 28 and 29: Images depicting an angry gesture and how it looks in the game

Some other mini-game ideas for levels I was thinking of implementing in the creative practice, contain scenarios where players would have to smile at specific intervals enabling them to dodge something, almost like a two-lane endless runner type game. An Endless Runner is characterized as an action game, linear in design with no end, without pauses or breaks for rest and no stages or changing levels. It has one continuous level. The game difficulty starts slow and easy and becomes more difficult over time. (Momoda, [59]) e.g., *Subway Surfer* (Kiloo, SYBO GAMES (2012) [60])

This will use the endless runner format, however, it will have a defined end in the game. As the game uses your face it could be extremely tiresome to use facial gestures for a long period of time. The game would be kept short in multiple levels, each one with a themed scenario.

Another idea is a meme[20] mimic. This little creature takes the form of many things. In this case, it takes the form of different memes. If a player laughs thus producing a smile, the mimic kills the player. Should the player last to the end without laughing or smiling, the mimic allows them to pass.

The sprite design for the current game was created in Photoshop. After the core mechanics of a level was created, an image search of the kind of characters required would be done. This would give me some ideas as to what could be used. Once I found a character design that I liked, the image was used as a guide to develop my character. I used Photoshop to develop a character that looked like the guide.

[20] an image, video, piece of text, etc., typically humorous in nature, that is copied and spread rapidly by internet users, often with slight variations.

The images below depict the characters used in the game.

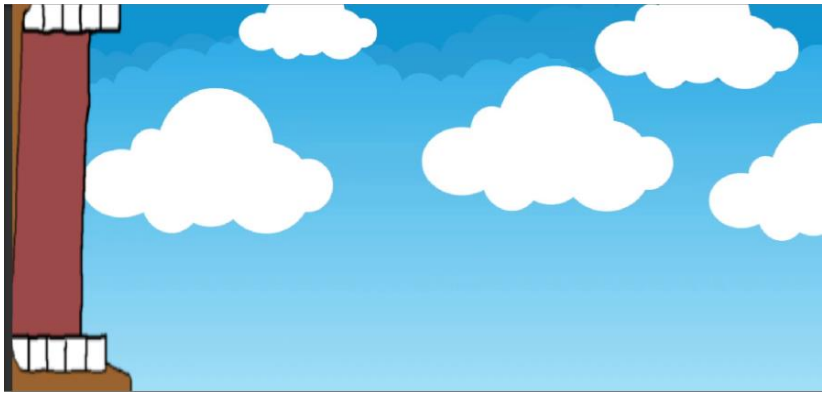


Figure 30: The image above is the design of the first level



Figure 31: The image above is in the second level of the game



Figure 32: The image above is in the third level of the game

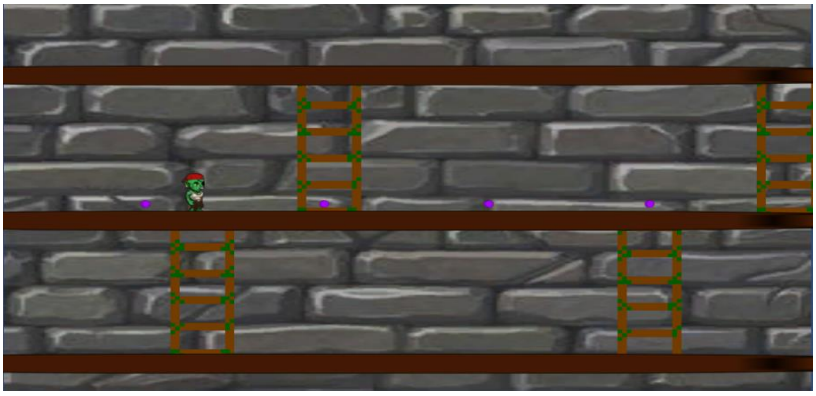


Figure 33: The image above is in the fourth level of the game

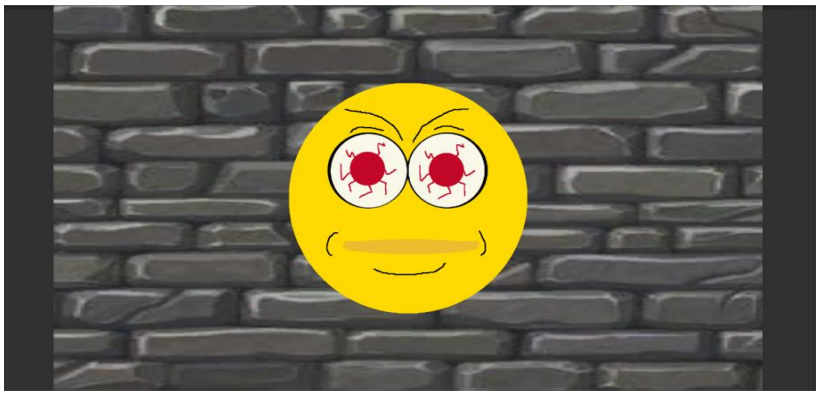


Figure 34: The image above is in the final level of the game

Section 4

4.1. Playtesting

My initial strategy to test the game and the viability of the technology, was to undertake a qualitative study of how the user interacts with the game. This study would have been undertaken by me observing the user during game play and a discussion on the experience post gameplay. For me to gain useful insight, I decided to have a more structured approach to the qualitative study and developed a questionnaire regarding their experiences.

The questions I asked after the play testers finished playing the game were:

1. Did they enjoy the experience?

2. Would they consider playing other games using only their face? and

3. If they thought it was a viable means to control even more complex games in the future?

I initially organized a location to test before gaining ethics clearance. However, because the process to get ethics clearance had taken a long time, and due to the COVID-19 pandemic, I had lost the use of the location to test my game. I was then forced to conduct the testing and analysis in a more informal setting. Players who tested the game were required to sign the information sheet thus giving consent to the testing process. Under normal circumstances a third-party tester would be used in my stead to test the system with the volunteers. However due to the pandemic, in order to minimize the number of human interactions, I dealt with the participants personally.

4.2 Feedback

I received mixed feedback from players. Some players really enjoyed the experience while others found it an ineffective means of control. A common complaint I received was that faces would tire especially on the level with the mosquitos. I was advised to place that level last as a “Boss-Level” as it was by far the most difficult out of the mini-games.

Due to the location not being controlled, lighting and distance were also factors of concern as the system is extremely susceptible to bugs in bad light conditions. This caused some frustration amongst players when false positives occurred causing a player to die.

On the other hand, some players had enjoyable experiences. Many seemed to enjoy the concept and said they would try other games that used only their face.

Out of the 15 players, 10 stated they would consider more complex games that used only facial expressions as a means of control. I had hoped the number would have been greater, but as the conditions were not controlled, the false positives had caused a lot of issues. This hindered the experience greatly. A solution to this would be to use a well-lit room or a higher quality camera to detect the players face. In addition, the game can have specific on-screen indicators of where the player needs to position themselves for optimal performance of the game.

4.3 How to Use

Currently, as the Gesture Detection system and the game are two individual systems, both would have to be run simultaneously to use the systems as intended. In the future, I am considering opening the python script that controls the Gesture Detection system via C# script. This would allow players to just run the game causing the python script to automatically open and run. This makes the gameplay seamless for the user.

The required dependencies for the Gesture Detection Haar Cascade method are python 3, OpenCV, NumPy and Pynput as well as the Haar Cascades. The CNN method requires the same dependencies as well as TensorFlow, Dlib and the trained emotion model.

Section 5

5.1 Conclusion and Discussion

The topic of the proposed research focuses on designing and implementing a game that is controlled using a player's facial gestures. The game utilizes the facial gestures to allow players to play the game using CNN and Haar Cascades to detect the face and facial gestures. I believe that this research can serve as a foundation for other researchers and organisation's hoping to leverage this technology. Many companies today have already begun leveraging this technology in the social media gaming world such as *SnapChat* and Google. Examples of using this technology are *Mimics- THE party game* (Everyone, 2014 [71]) and *the Face Dance Challenge!* (Everyone, [72]). These games are developed to create fun and entertainment socially through playing the game with friends and family. There are currently very few games with this feature. The research also aims to decipher whether the technology is viable as an input mechanism to enhance and improve the overall user experience. I categorized the viability and user experience in three areas. These are Responsiveness, Customer Experience and Practicality.

The following observations were made based on the viability of the technology.

1. Responsiveness

   The overall responsiveness of the technology was fast enough and comparable to other traditional input devices, e.g., game controllers. During this project, I had to utilize several different technologies such as OpenCV libraries, Haar Cascades and CNN to improve the speed and accuracy of the technology. Most of the players did not find any issue with the responsiveness of the technology.

   Out of the user group across all 15 players, one of the players had issues with the responsiveness of the technology. All the other participants found the responsiveness comparable to traditional input devices i.e., controllers or a keyboard and mouse combinations.

2. Customer Experience

   The overall customer experience was good, and all the users had fun with the technology. Both user groups of gamers and non-gamers alike, thought the concept was fun and enjoyable. During the testing process some of the players experienced false positives. For example, the system picked up surprise as frowning which resulted in player frustration. For a few users, the system struggled to pick up the player. This was remedied by moving to an area that was well lit and moving the player closer to the camera. A possible solution is to improve the lighting, or a better-quality camera, and to provide some on screen

instructions of where the player needs to position themselves. When navigating through the various levels, many of the users indicated that the technology was “gimmicky” but also fun.

In this area we had a few of the players that experienced false positives. This ranged across the cohorts of gamers and non-gamers. These were experienced by people with darker skin tones as the technology struggled to pick up facial gestures accurately. Although the sample size was small, this impacted 46% of the user group. All players in the Black ethnic group, two in the Indian, and one in the Coloured group experienced false positives. This needs to be tested across a wider sample group to unpack in more detail, the impact of false positives.

3. Practicality

On the practicality area, as mentioned above, many of the players indicated that the technology was “gimmicky” and would be ideal for mobile games. The avid gaming cohort of users indicated that this technology would be difficult to play on more complex multi-button games, games with quick movements, or games that would take long time to play. They indicated that facial fatigue would ruin the game experience. With that said, most of the players across the sample size indicated that they would be willing to try this technology in the future and showed a willingness to try and use it on more complex games. The user with the motor impairments really found this technology enabling and indicated that they would love to see more games being developed using this technology.

In this element we had a definitive split in the user population. Most of the gamer cohort thought that the technology was “cool” but “gimmicky” and impractical on more complex multi-buttoned games. They indicated that this would be good for mobile games with a simple play style. All five high familiarity gamers indicated the above, while one of the medium familiarity gamers offered the same sentiment.

Most of the low and medium familiarity gamers, as well as the motor impaired gamer, thought that the technology was great and useful.

To improve the overall viability of this technology, the research could be extended further by utilizing facial gestures in conjunction with a traditional means of control. There could also be space for improvement in the algorithm that automatically refines for real time usage as more efficient methods of this technology are evolving daily.

Once the technology is more refined, it could afford physically disabled people the opportunity to compete against fully able-bodied people by choosing the means of control they would like to use, whether it be

gestures or traditional controls. This would make the gaming industry more inclusive. This system could also be beneficial in creating educational games and be used in physiotherapy settings.

In conclusion, the research proves that the technology can work in a game setting, however with the various issues experienced during the testing and the feedback from the user community, I do not think that this form of control for a mainstream peripheral device would be practical for competitive gaming in its current form. The technology would require a lot more development to provide a consistent and stable player experience in a competitive scene. On the other hand, as a casual gaming experience, it can provide a fun level of party games with friends who are not overly competitive and do not mind missing inputs now and then. Ten of the fifteen players from the user group indicated that they believe that the technology can be a viable option in the future.

References

[1] Yu-Hao Huang and Chiou-Shann Fuh. Face detection and smile detection. In Proceedings of IPPR Conference on Computer Vision, Graphics and Image Processing, Shitou, Taiwan, A5-6, page 108, 2009.

[2] Daniel McDuff, Rana El Kaliouby, Jeffrey F Cohn, and Rosalind W Picard. Predicting ad liking and purchase intent: Large-scale analysis of facial responses to ads. Affective Computing, IEEE Transactions on, 6(3):223–235, 2015.

[3] Daniel McDuff, Rana El Kaliouby, Thibaud Senechal, David Demirdjian, and Rosalind Picard. Automatic measurement of ad preferences from facial responses gathered over the internet. Image and Vision Computing, 32(10):630–640, 2014.

[4] Jonathan Rihan. Computer Vision Based Interfaces for Computer Games. PhD thesis, Citeseer, 2010.

[5] Caifeng Shan. Smile detection by boosting pixel differences. Image Processing, IEEE Transactions on, 21(1):431–436, 2012.

[6] Keiji Shimada, Tetsu Matsukawa, Yoshihiro Noguchi, and Takio Kurita. Appearance-based smile intensity estimation by cascaded support vector machines. In Computer Vision–ACCV 2010 Workshops, pages 277–286. Springer, 2010.

[7] Jacob Whitehill, Gwen Littlewort, Ian Fasel, Marian Bartlett, and Javier Movellan. Toward practical smile detection. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 31(11):2106– 2111, 2009.

[8] Chance Miller Nose Zone & Rainbrow are early games that utilize the iPhone X's TrueDepth sensor for controls - Dec. 18th, 2017 6:48 pm PT https://9to5mac.com/2017/12/18/iphone-x-truedepth-face-control-games/

[9] Will Berger DEEP LEARNING HAAR CASCADE EXPLAINED www.willberger.org/cascade-haar-explained/

[10] Thorsten Ball TRAIN YOUR OWN OPENCV HAAR CLASSIFIER 22 July 2013 https://coding-robin.de/2013/07/22/train-your-own-opencv-haar-classifier.html

[11] Sentdex Haar Cascade Object Detection Face & Eye - OpenCV with Python for Image and Video Analysis 16 https://www.youtube.com/watch?v=88HdqNDQsEk 10 Jan 2016

[12] Dasaradh Makes Make your own Haar Cascade on Windows | Quick & Simple https://www.youtube.com/watch?v=Dg-4MoABv4I 30 Sep 2016

[13] Dileep Kumar Creating a haar cascade classifier aka haar training.

http://opencvuser.blogspot.com/2011/08/creating-haar-cascade-classifier-aka.html Thursday, August 18, 2011

[14] Sefik Ilkin Serengil Real Time Facial Expression Recognition on Streaming Data

https://sefiks.com/2018/01/10/real-time-facial-expression-recognition-on-streaming-data/ January 10, 2018

[15] https://docs.opencv.org/3.4/db/d28/tutorial_cascade_classifier.html

[16] Jason Brownlee Multi-Class Classification Tutorial with the Keras Deep Learning Library https://machinelearningmastery.com/multi-class-classification-tutorial-keras-deep-learning-library/

[17] Mathworks, Mathworks. “Train a Cascade Object Detector” Mathworks, 2017, www.mathworks.com/help/vision/ug/train-a-cascade-object-detector.html.

[18] Rainer Lienhart and Jochen Maydt. An extended set of haar-like features for rapid object detection. In Image Processing. 2002. Proceedings. 2002 International Conference on, volume 1, pages I–900. IEEE, 2002.

[19] Bay, H., Tuytelaars, T. and Van Gool, L. [2006], ‘Surf: Speeded up robust features’, Computer Vision–ECCV 2006.

[20] Freeman, W., Tanaka, K., Ohta, J. and Kyuma, K. [1996], Computer vision for computer games.

[21] Thatgamecompany, 2012. Journey. [Art] (Sony)

[22] Gitter, N., 2017. Rainbrow. [Art] (Washington University).

[23] Oxford, 2019. Lexico. [Online]
Available at: https://www.lexico.com/en/definition/real_time
[Accessed 22 07 2019].

[24] Calo, C., 2019. Towards Data Science. [Online]
Available at: https://towardsdatascience.com/microsoft-introduction-to-ai-part-3-cb21d7a5e119
[Accessed 06 June 2019].

[25] Saha,S., 2018 Towards Data Science. [Online]
Available at: https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53
[Accessed 06 June 2019].

[26] Dlib user documentation http://dlib.net/

[27] Nintendo.Fandom https://nintendo.fandom.com/wiki/Game_Boy_Camera[Accessed 06 June 2019].

[28] Bernadette Johnson "How the PlayStation Camera Works" 16 January 2014. HowStuffWorks.com. <https://electronics.howstuffworks.com/playstation-camera.htm> 20 August 2019

[29] Wikmili, 2019 https://wikimili.com/en/Xbox_Live_Vision

[30] Bernadette Johnson "How the PlayStation Camera Works" 16 January 2014. HowStuffWorks.com. <https://electronics.howstuffworks.com/playstation-camera.htm> 20 August 2019

[31] Veljko B. Petrović, Dragan Ivetić Active Learning Attributes Of Wii Peripherals In CS Education, Faculty of Technical Sciences-University of Novi Sad, Trg Dositeja Obradovića 6, Novi Sad 21000, Serbia 2013

[32] Robert Cong & Ryan Winters" How does the Xbox Kinect work" https://www.jameco.com/jameco/workshop/howitworks/xboxkinect.html [Accessed 04 July 2019]

[33]Daniel Cooper"Sony Reveals how the PlayStation 4 Eye works"(2013) https://www.engadget.com/2013/02/21/sony=playstation-4-eye-works/?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAAKKRCnxKt-qpozTunzPHz8yM58l2BNxVnZTX5jTesAmHzZbYQ9DuJvL3SHlyFGurRGvWVQM4kDiB7ALrcuVQZloNXX368UrnjryfeFeWWfOV-GIcg4A1X1oow5SKMHpWkWvw_USGjprGVYNOP7Q28AuqNeHuFb0A1JpuWcIXF1a4 [Accessed 04 July 2019]

[34]Wikipedia https://en.wikipedia.org/wiki/Monocular

[35] Wikipedia https://en.wikipedia.org/wiki/Stereo_camera

[36] Rihan, J., 2010. Computer Vision Based Interfaces for computer games. Citeseer.

[37] Dr Guy Winch Video Games Impact Your Feelings in Real Life Posted Sep 21, 2015 [Accessed 12 February 2020] https://www.psychologytoday.com/us/blog/the-squeaky-wheel/201509/video-games-impact-your-feelings-in-real-life

[38] Vorderer, P., Bryant, J., Pieper, K. M., & Weber, R. (2006). Playing Video Games as Entertainment. In P. Vorderer & J. Bryant (Eds.), Playing video games: Motives, responses, and consequences (p. 1–7). Lawrence Erlbaum Associates Publishers. [Accessed 12 February 2020]

[39] Game Feel: The Secret Ingredient by Steve Swink https://www.gamasutra.com/view/feature/130734/game_feel_the_secret_ingredient.php November 23, 2007 [Accessed 20 February 2020]

[40] Game Feel: A Game Designer's Guide to Virtual Sensation by Steve Swink https://books.google.co.za/books/about/Game_Feel.html?id=3oFjDwAAQBAJ&printsec=frontcover&source=kp_read_button&redir_esc=y#v=snippet&q=garden&f=false 2017 [Accessed 1st March 2020]

[41] The Art and Science of Game Feel | How Game Designers Juice Games with Mechanics,Pacing and Effects by The Game Overanalyser https://www.youtube.com/watch?v=EDElfx2qo_M June 2019 [Accessed 1st March 2020]

[42] Secrets of Game Feel and Juice | Game Design Basics by Game Maker's Toolkit https://www.youtube.com/watch?v=216_5nu4aVQ February 2015 [Accessed 3rd March 2020]

[43]Scardovelli, Terigi & Frère, A. (2014). The design and evaluation of a peripheral device for use with a computer game intended for children with motor disabilities. Computer Methods and Programs in Biomedicine. 118. 10.1016/j.cmpb.2014.10.002. [Accessed 30 October 2020]

[44] Razbuten https://www.youtube.com/watch?v=ax7f3JZJHSw What Games Are Like For Someone Who Doesn't Play Games] September 28, 2019[Accessed 2 December 2020]

[45] Oswald, Christopher & Prorock, Chris & Murphy, Shane. (2014). The perceived meaning of the video game experience: An exploratory study. Psychology of Popular Media Culture. 3. 110. 10.1037/a0033828. [Accessed 2 December 2020]

[46] Birk and Mandryk 2013 and A Case Study Approach (pp.145-188) Usability Testing of Video Game Controllers: A Case Study May 2016 [Accessed 4 December 2020]

[47] F. Iacopetti, L. Fanucci, R. Roncella, D. Giusti and A. Scebba, "Game Console Controller Interface for People with Disability," *2008 International Conference on Complex, Intelligent and Software Intensive Systems*, Barcelona, 2008, pp. 757-762, doi: 10.1109/CISIS.2008.77. [Accessed 4December 2020]

[48] Hatchlings, Inc. *Nose Zone*. [Art]

[49] *Metrp Trains (2012) Dumb Ways to Die. [*Art*]*

[50] *Bandai Namco (1994) Tekken [Art]*

[51] NetherRealm Studios (1992) [Art]

[52] *RobTop Games (2013)[Art]*

[53] *Skyrim*(Bethesda, 2011)[Art]

[54]Akash Desarda Jan 17, 2019 https://towardsdatascience.com/understanding-adaboost-2f94f22d5bfe [Online] [Accessed 21 January 2021]

[55] https://opencv.org/about/ [Online] [Accessed 21 January 2021]

[56] https://www.youtube.com/watch?v=_QZLbR67fUU&t=278s Viola Jones Face Detection Explained Rahul Patil 2014 [Video] [Accessed 21 January 2021]

[57] Jason BrownLee How Do Convolutional Layers Work In Deep Learning Neural Netwoks https://machinelearningmastery.com/convolutional-layers-for-deep-learning-neural-networks/ 2019 [Online] [Accessed 21 January 2021]

[58] Srini Pillay Why you can't get a song out of your head and what to do about it https://www.health.harvard.edu/blog/why-you-cant-get-a-song-out-of-your-head-and-what-to-do-about-it-2017100412490 2017 [Online] [Accessed 21 January 2021]

[59] Jerry Momoda Endless Runner Games: Evolution And Future http://jerrymomoda.com/analysis-endless-runners/ [Online] [Accessed 21 January 2021]

[60] *Subway Surfer* Kiloo,SYBO GAMES (2012) [Art]

[61]https://www.dpreview.com/articles/1838149063/throwback-thursday-gameboy-camera [Image] [Accessed 21 January 2021]

[62]https://www.ebay.com/itm/PS2-Playstation-2-official-Sony-EyeToy-USB-Camera-Eye-Toy-Camera-silver-/203005105108[Image] [Accessed 21 January 2021]

[63] https://www.walmart.com/ip/xbox-360-live-vision-camera/5065209[Image] [Accessed 21 January 2021]

[64]https://www.amazon.com/Official-SONY-PSP-Cam-450x-Camera/dp/B0047ERZ7Q[Image] [Accessed 21 January 2021]

[65]https://www.amazon.com/Sony-Playstation-Move-Motion-Camera/dp/B0107WL3BA[Image] [Accessed 21 January 2021]

[66]https://www.neogaf.com/threads/ubisofts-wii-camera-anyone-use-it-and-how-does-it-compare-to-kinect-and-such.1415935/[Image] [Accessed 21 January 2021]

[67]https://www.techradar.com/reviews/gaming/games-consoles/controllers/microsoft-kinect-for-xbox-360-905010/review[Image] [Accessed 21 January 2021]

[68]https://www.ebay.com/p/9032778187[Image] [Accessed 21 January 2021]

[69]https://www.bestbuy.com/site/sony-playstation-camera-for-playstation-4-new/5622702.p?skuId=5622702[Image] [Accessed 21 January 2021]

[70]https://www.dxomark.com/the-best-smartphone-cameras/[Image] [Accessed 21 January 2021]

[71] *Mimics- THE party game* Everyone, 2014 [Art]

[72] *Face Dance Challenge!* Everyone [Art]

Image List

Figure 1 – RainBrow (https://www.dxomark.com/the-best-smartphone-cameras/, [70])

Figure 2 – Nintendo Game Boy Camera (https://www.dpreview.com/articles/1838149063/throwback-thursday-gameboy-camera, [61])

Figure 3 – Sony EyeToy (https://www.ebay.com/itm/PS2-Playstation-2-official-Sony-EyeToy-USB-Camera-Eye-Toy-Camera-silver-/203005105108,[62])

Figure 4 – Xbox Live Vision (https://www.walmart.com/ip/xbox-360-live-vision-camera/5065209,[63])

Figure 5 – Sony Go! Cam (https://www.amazon.com/Official-SONY-PSP-Cam-450x-Camera/dp/B0047ERZ7Q,[64])

Figure 6 – Sony PlayStation Eye (https://www.amazon.com/Sony-Playstation-Move-Motion-Camera/dp/B0107WL3BA,[65])

Figure 7 – Nintendo Wii Camera (https://www.neogaf.com/threads/ubisofts-wii-camera-anyone-use-it-and-how-does-it-compare-to-kinect-and-such.1415935/, [66])

Figure 8 – Xbox 360 Kinect (https://www.techradar.com/reviews/gaming/games-consoles/controllers/microsoft-kinect-for-xbox-360-905010/review,[67])

Figure 9 – Xbox one Kinect (https://www.ebay.com/p/9032778187, [68])

Figure 10 – PlayStation 4 Camera (https://www.bestbuy.com/site/sony-playstation-camera-for-playstation-4 new/5622702.p?skuId=5622702, [69])

Figure 11 – Smartphones (https://www.dxomark.com/the-best-smartphone-cameras/, [70])

Figure 12 – The Top Two rows depict some and bottom two rows depict non-smile as depicted by (Shan,2012 page 431-436[5])

Figure 13 – (Shimada et al(2010, pages 277-286)) Appearance – based smile intensity estimation by cascaded support vector machines

Figure 14 – (Shimada et al(2010, pages 277-286)) Appearance – based smile intensity estimation by cascaded support vector machines

Figure 15 – Deep Learning Haar Cascade Explained (Will Berg , [9])

Figure 16 and 17 - Depicting system detecting neutral and smiling gestures

Figure 18 – Game Feel: A Game Designer’s Guide to Virtual Sensation(Steve Swing, 201, [40])

Table 1 – Ethnic Groups

Table 2 – Familiarity in gaming industry

Figure 19 ,20 ,21,22 – Surprise and smiling gesture with the outcomes in game.

Figure 23 – Smiling gesture

Figure 24,25,26,27,28,29 – Surprise , angry and smiling gestures with the outcomes in game.

Figure 30 – First Level of game

Figure 31 – Second Level of game

Figure 32 – Third Level of game

Figure 33 – Fourth Level of game

Figure 34 – Fifth Level of game

Appendix 1

Facial Feature points

There are many features on the face that can be used as features for detecting facial expressions. According to Huang, Yu-Hao and Fuh, Chiou-Shann the following features are used:

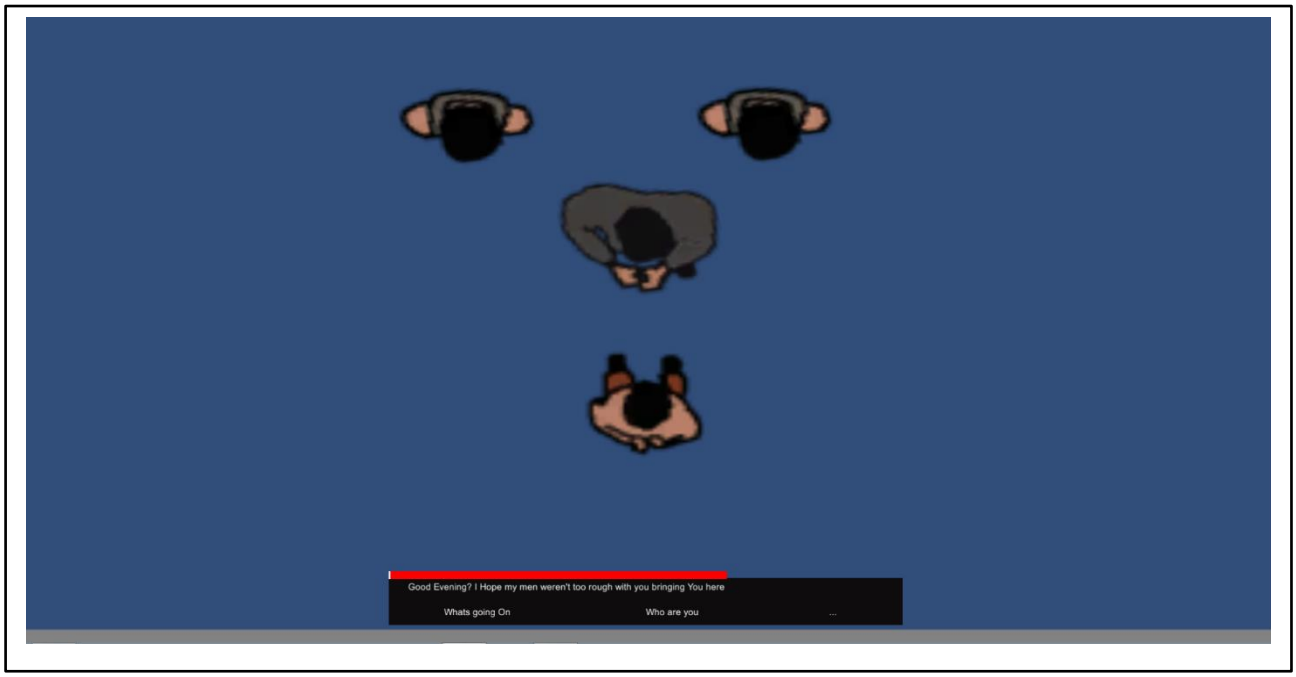
- right eye pupil
- left eye pupil
- right mouth corner
- left mouth corner
- outer end of right eyebrow
- inner end of right eyebrow
- inner end of left eyebrow
- outer corner of right eye
- inner corner of right eye
- inner corner of left eye
- outer corner of left eye
- left temple
- tip of nose
- right nostril
- left nostril
- center point on outer edge of upper lip
- center point on outer edge of lower lip
- tip of chin
- outer end of left eyebrow
- right temple

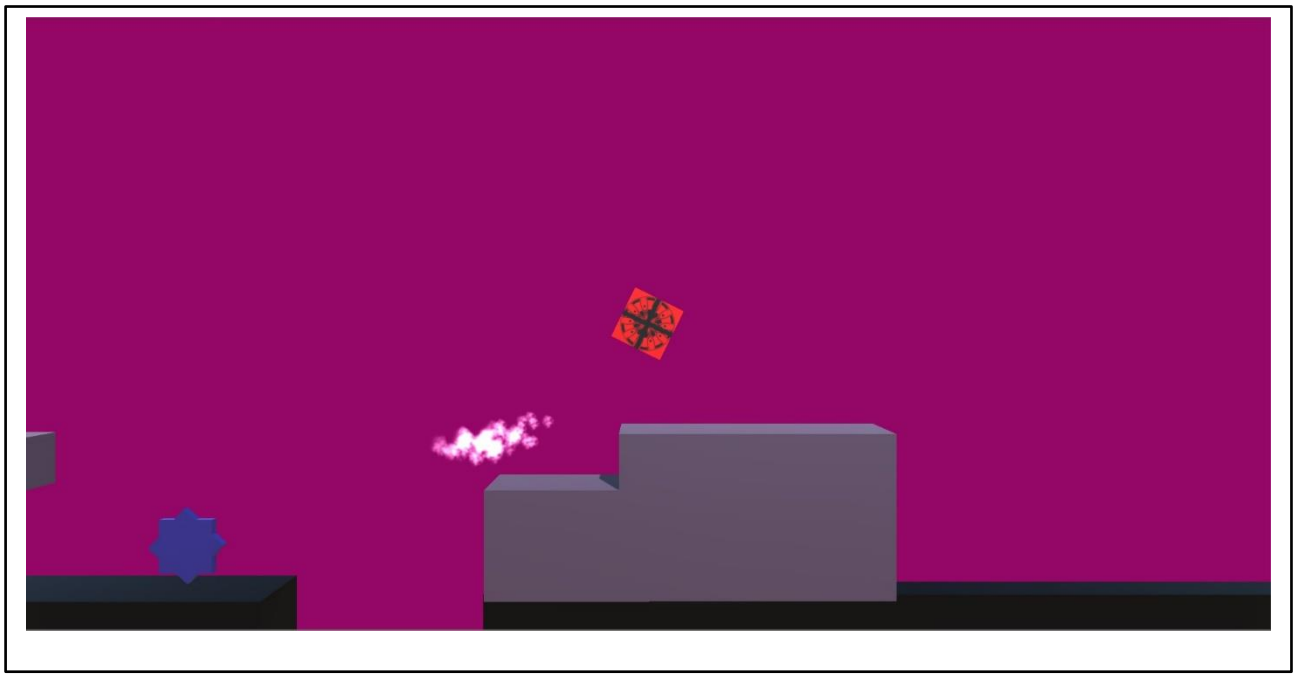


Face detection and smile detection (Yu-Hao Huang and Chiou-Shann, 2009, [1])

Appendix 2

Some Screenshots from The First and Second Games





Appendix 3

Volunteer permission letter

Dear Sir/Madam

My name is Farouk Suleman, and I am a Masters student in Digital Arts Game Design at Wits University in Johannesburg. As part of my studies, I have to undertake a research project, and I am investigating the potential of using a player's face via Computer Vision as an alternative to traditional video game controllers. Computer Vision is a subset field of Artificial Intelligence that aims at giving computers a visual understanding of the world.

The aim is to explore the notion of Computer Vision in video games and how it can be used to analyze a player's facial features in order to control actions based on the facial expressions or movements of the player's in-game avatar. As things stand there are apps that allow for user interaction via these methods, for example, Snapchat filters, Facebook picture filters, etc. However, although these allow interactions via facial tracking, these apps do not take it one step further and implement these features as controls into a game (with the exception of the new iPhone X games and Snapchat games). As such the intention is to demonstrate through Creative Practice a fun, engaging game using a human player's face as the primary controller and playtest in order to determine if this is a viable and exciting method of controlling a video game, or if it is just a gimmicky system. Due to the nature of the system, there will be a camera capturing the face of users however the system does not save any of faces. The face is only analyzed and then forgotten.

As part of this project, I would like to invite you to take part in experiencing the game and sharing your thoughts and experiences with me so that I may report on how players are able to interact with the game after which, with your permission, I will ask verbally and record what has been said via a note pad app. This will require you to sit down and play the game. Should the user be affected with epilepsy or vertigo, I would strongly advise against participating in this project. There should be no other risks in relation to playing the game.

The process will involve you playing the game and answering those three questions with as much or as little detail as you deem fit and should take around 5-10 minutes depending on how quickly you progress through the game and how much detail you would provide to the answering of the questions.

The game consists of various different mini game levels mostly consisting of “one button” style games however, the last few levels will require multiple facial gestures.

There will be no direct benefits for participating in this research study, and no disadvantages for not participating or withdrawing from the study. You may withdraw at any time or not answer any question should you not want to. The interview will be completely confidential and anonymous as I will not be asking for your name or any identifying information, and the information you give to me will be held securely and only be used for my final research report.

There should not be any discomfort during the process however, if you do experience any distress or discomfort, we will stop the experience or resume another time and if you have any questions afterwards about this research, feel free to contact me on the details listed below. This study will be written up as a research report which will be available online through the university library website. If you have any queries, concerns or complaints regarding the ethical procedures of this study, you are welcome to contact the University Human Research Ethics Committee(non-medical), telephone + 27(0)11 717 1408, email Shaun.Schoeman@wits.ac.za

Yours sincerely,

Farouk Suleman

Farouk.suleman3@gmail.com, 0715410566 Supervisor: Raymond Whitcher, raymond.whitcher@wits.ac.za , 083 466 7854

Farouk Suleman,

I ……Hereby agree to participate in this research project. The research has been explained to me and I understand what my participation will involve.

(Please Circle Yes or No)

I agree that my participation will remain anonymous ……..Yes / No

I agree that the researcher may use anonymous quotes in his research report……..Yes / No

I agree that the information I provide may be used anonymously by

other researchers following this study……..Yes / No

…… (signature)

…… (name of participant)

…… (date)

Appendix 4

Ethics Form

UNIVERSITY OF THE WITWATERSRAND JOHANNESBURG

Research Office

HUMAN RESEARCH ETHICS COMMITTEE (NON-MEDICAL)
R14/49 Suleman

CLEARANCE CERTIFICATE	**PROTOCOL NUMBER: H19/05/29**
PROJECT TITLE	Supervision: An investigation into the potential use of a player's face, using computer vision, as an alternative to traditional video game controllers
INVESTIGATOR(S)	Mr F Suleman
SCHOOL/DEPARTMENT	Arts/
DATE CONSIDERED	24 May 2019
DECISION OF THE COMMITTEE	Approved
EXPIRY DATE	13 November 2022

DATE 14 November 2019 **CHAIRPERSON** (Professor J Knight)

cc: Supervisor : Mr R Whicher

DECLARATION OF INVESTIGATOR(S)

To be completed in duplicate and **ONE COPY** returned to the Secretary at Room 10004, 10th Floor, Senate House, University. Unreported changes to the application may invalidate the clearance given by the HREC (Non-Medical)

I/We fully understand the conditions under which I am/we are authorized to carry out the abovementioned research and I/we guarantee to ensure compliance with these conditions. Should any departure to be contemplated from the research procedure as approved I/we undertake to resubmit the protocol to the Committee. **I agree to completion of a yearly progress report.**

Signature

13 / 05 / 2020
Date

PLEASE QUOTE THE PROTOCOL NUMBER ON ALL ENQUIRIES

Appendix 5

Interview Questions

Interview

Please circle Yes or No.

Did you enjoy the experience? (Yes) / No

Would you consider playing other games using only your face? Yes / (No)

Do you think it was a viable means to control even more complex games in the future? Yes / (No)

Comments

My ~~face~~ Jaw gets tired

Thank you for choosing to participate in this study. If you have any queries, concerns or complaints regarding the ethical procedures of this study, you are welcome to contact the University Human Research Ethics Committee(non-medical), telephone + 27(0)11 717 1408, email Shaun.Schoeman@wits.ac.za

Yours sincerely,

Farouk Suleman

Farouk.suleman3@gmail.com, 0715410566

Supervisor: Raymond Whitcher, raymond.whitcher@wits.ac.za , 083 466 7854

Interview

Please circle Yes or No.

Did you enjoy the experience? Yes /(No)

Would you consider playing other games using only your face? Yes /(No)

Do you think it was a viable means to control even more complex games in the future? Yes /(No)

Comments

Glitchy, Doesn't work properly

Thank you for choosing to participate in this study. If you have any queries, concerns or complaints regarding the ethical procedures of this study, you are welcome to contact the University Human Research Ethics Committee(non-medical), telephone + 27(0)11 717 1408, email Shaun.Schoeman@wits.ac.za

Yours sincerely,

Farouk Suleman

Farouk.suleman3@gmail.com, 0715410566

Supervisor: Raymond Whitcher, raymond.whitcher@wits.ac.za , 083 466 7854

Interview

Please circle Yes or No.

Did you enjoy the experience? (Yes)/ No

Would you consider playing other games using only your face? (Yes)/ No

Do you think it was a viable means to control even more complex games in the future? (Yes)/ No

Comments

Glitchy

Thank you for choosing to participate in this study. If you have any queries, concerns or complaints regarding the ethical procedures of this study, you are welcome to contact the University Human Research Ethics Committee(non-medical), telephone + 27(0)11 717 1408, email Shaun.Schoeman@wits.ac.za

Yours sincerely,

Farouk Suleman

Farouk.suleman3@gmail.com, 0715410566

Supervisor: Raymond Whitcher, raymond.whitcher@wits.ac.za , 083 466 7854

Interview

Please circle Yes or No.

Did you enjoy the experience? (Yes) / No

Would you consider playing other games using only your face? (Yes) / No

Do you think it was a viable means to control even more complex games in the future? (Yes) / No

Comments

The first stage is the hardest not maybe. make it last

Thank you for choosing to participate in this study. If you have any queries, concerns or complaints regarding the ethical procedures of this study, you are welcome to contact the University Human Research Ethics Committee(non-medical), telephone + 27(0)11 717 1408, email Shaun.Schoeman@wits.ac.za

Yours sincerely,

Farouk Suleman

Farouk.suleman3@gmail.com, 0715410566

Supervisor: Raymond Whitcher, raymond.whitcher@wits.ac.za , 083 466 7854

Interview

Please circle Yes or No.

Did you enjoy the experience? (Yes) / No

Would you consider playing other games using only your face? (Yes) / No

Do you think it was a viable means to control even more complex games in the future? Yes / (No)

Comments

_____

_____

_____

Thank you for choosing to participate in this study. If you have any queries, concerns or complaints regarding the ethical procedures of this study, you are welcome to contact the University Human Research Ethics Committee(non-medical), telephone + 27(0)11 717 1408, email Shaun.Schoeman@wits.ac.za

Yours sincerely,

Farouk Suleman

Farouk.suleman3@gmail.com, 0715410566

Supervisor: Raymond Whitcher, raymond.whitcher@wits.ac.za , 083 466 7854

Interview

Please circle Yes or No.

Did you enjoy the experience? Yes / (No)

Would you consider playing other games using only your face? (Yes) / No

Do you think it was a viable means to control even more complex games in the future? (Yes) / No

Comments

Couldn't pick me up nicely

Thank you for choosing to participate in this study. If you have any queries, concerns or complaints regarding the ethical procedures of this study, you are welcome to contact the University Human Research Ethics Committee(non-medical), telephone + 27(0)11 717 1408, email Shaun.Schoeman@wits.ac.za

Yours sincerely,

Farouk Suleman

Farouk.suleman3@gmail.com, 0715410566

Supervisor: Raymond Whitcher, raymond.whitcher@wits.ac.za , 083 466 7854

Interview

Please circle Yes or No.

Did you enjoy the experience? (Yes) / No

Would you consider playing other games using only your face? (Yes) / No

Do you think it was a viable means to control even more complex games in the future? (Yes) / No

Comments

_____

_____

_____

Thank you for choosing to participate in this study. If you have any queries, concerns or complaints regarding the ethical procedures of this study, you are welcome to contact the University Human Research Ethics Committee(non-medical), telephone + 27(0)11 717 1408, email Shaun.Schoeman@wits.ac.za

Yours sincerely,

Farouk Suleman

Farouk.suleman3@gmail.com, 0715410566

Supervisor: Raymond Whitcher, raymond.whitcher@wits.ac.za , 083 466 7854

Interview

Please circle Yes or No.

Did you enjoy the experience? (Yes) / No

Would you consider playing other games using only your face? Yes / (No)

Do you think it was a viable means to control even more complex games in the future? Yes / (No)

Comments

_____

_____

_____

Thank you for choosing to participate in this study. If you have any queries, concerns or complaints regarding the ethical procedures of this study, you are welcome to contact the University Human Research Ethics Committee(non-medical), telephone + 27(0)11 717 1408, email Shaun.Schoeman@wits.ac.za

Yours sincerely,

Farouk Suleman

Farouk.suleman3@gmail.com, 0715410566

Supervisor: Raymond Whitcher, raymond.whitcher@wits.ac.za , 083 466 7854

Interview

Please circle Yes or No.

Did you enjoy the experience? (Yes) / No

Would you consider playing other games using only your face? (Yes) / No

Do you think it was a viable means to control even more complex games in the future? (Yes) / No

Comments

Thank you for choosing to participate in this study. If you have any queries, concerns or complaints regarding the ethical procedures of this study, you are welcome to contact the University Human Research Ethics Committee(non-medical), telephone + 27(0)11 717 1408, email Shaun.Schoeman@wits.ac.za

Yours sincerely,

Farouk Suleman

Farouk.suleman3@gmail.com, 0715410566

Supervisor: Raymond Whitcher, raymond.whitcher@wits.ac.za , 083 466 7854

Interview

Please circle Yes or No.

Did you enjoy the experience? (Yes) / No

Would you consider playing other games using only your face? (Yes) / No

Do you think it was a viable means to control even more complex games in the future? Yes / (No)

Comments

_____

_____

_____

Thank you for choosing to participate in this study. If you have any queries, concerns or complaints regarding the ethical procedures of this study, you are welcome to contact the University Human Research Ethics Committee(non-medical), telephone + 27(0)11 717 1408, email Shaun.Schoeman@wits.ac.za

Yours sincerely,

Farouk Suleman

Farouk.suleman3@gmail.com, 0715410566

Supervisor: Raymond Whitcher, raymond.whitcher@wits.ac.za , 083 466 7854

Interview

Please circle Yes or No.

Did you enjoy the experience? (Yes) / No

Would you consider playing other games using only your face? (Yes) / No

Do you think it was a viable means to control even more complex games in the future? (Yes) / No

Comments

_____

_____

_____

Thank you for choosing to participate in this study. If you have any queries, concerns or complaints regarding the ethical procedures of this study, you are welcome to contact the University Human Research Ethics Committee(non-medical), telephone + 27(0)11 717 1408, email Shaun.Schoeman@wits.ac.za

Yours sincerely,

Farouk Suleman

Farouk.suleman3@gmail.com, 0715410566

Supervisor: Raymond Whitcher, raymond.whitcher@wits.ac.za , 083 466 7854

Interview

Please circle Yes or No.

Did you enjoy the experience? (Yes) / No

Would you consider playing other games using only your face? (Yes) / No

Do you think it was a viable means to control even more complex games in the future? (Yes) / No

Comments

_____

_____

_____

Thank you for choosing to participate in this study. If you have any queries, concerns or complaints regarding the ethical procedures of this study, you are welcome to contact the University Human Research Ethics Committee(non-medical), telephone + 27(0)11 717 1408, email Shaun.Schoeman@wits.ac.za

Yours sincerely,

Farouk Suleman

Farouk.suleman3@gmail.com, 0715410566

Supervisor: Raymond Whitcher, raymond.whitcher@wits.ac.za , 083 466 7854

Interview

Please circle Yes or No.

Did you enjoy the experience? Yes / No

Would you consider playing other games using only your face? Yes / No

Do you think it was a viable means to control even more complex games in the future? Yes / No

Comments

_____

_____

_____

Thank you for choosing to participate in this study. If you have any queries, concerns or complaints regarding the ethical procedures of this study, you are welcome to contact the University Human Research Ethics Committee(non-medical), telephone + 27(0)11 717 1408, email Shaun.Schoeman@wits.ac.za

Yours sincerely,

Farouk Suleman

Farouk.suleman3@gmail.com, 0715410566

Supervisor: Raymond Whitcher, raymond.whitcher@wits.ac.za , 083 466 7854

Interview

Please circle Yes or No.

Did you enjoy the experience? (Yes) / No

Would you consider playing other games using only your face? (Yes) / No

Do you think it was a viable means to control even more complex games in the future? (Yes) / No

Comments

_____

_____

_____

Thank you for choosing to participate in this study. If you have any queries, concerns or complaints regarding the ethical procedures of this study, you are welcome to contact the University Human Research Ethics Committee(non-medical), telephone + 27(0)11 717 1408, email Shaun.Schoeman@wits.ac.za

Yours sincerely,

Farouk Suleman

Farouk.suleman3@gmail.com, 0715410566

Supervisor: Raymond Whitcher, raymond.whitcher@wits.ac.za , 083 466 7854

Interview

Please circle Yes or No.

Did you enjoy the experience? (Yes) / No

Would you consider playing other games using only your face? (Yes) / No

Do you think it was a viable means to control even more complex games in the future? (Yes) / No

Comments

_____

_____

_____

Thank you for choosing to participate in this study. If you have any queries, concerns or complaints regarding the ethical procedures of this study, you are welcome to contact the University Human Research Ethics Committee(non-medical), telephone + 27(0)11 717 1408, email Shaun.Schoeman@wits.ac.za

Yours sincerely,

Farouk Suleman

Farouk.suleman3@gmail.com, 0715410566

Supervisor: Raymond Whitcher, raymond.whitcher@wits.ac.za , 083 466 7854