Using a Multi-Spectral Imaging Device to Monitor the Distractive Use of Mobile Devices by Students in a Classroom Setting

John W Campbell

North Carolina A&T State University

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Department: Electrical and Computer Engineering

Major: Computer Engineering

Major Professor: Dr. Corey A Graves

Greensboro, North Carolina

2015

The Graduate School
North Carolina Agricultural and Technical State University

This is to certify that the Master's Thesis of

John W Campbell

has met the thesis requirements of
North Carolina Agricultural and Technical State University

Greensboro, North Carolina
2015

Approved by:

Dr. Corey A. Graves Major Professor	Dr. Gary L. Lebby Committee Member
Dr. Christopher C. Doss Committee Member	Dr. John C. Kelly Department Chair

Dr. Sanjiv Sarin
Dean, The Graduate School

Biographical Sketch

John W Campbell is an aspiring engineer who always has had a passion for electronics. As a child, the youngest of five, he was curious about how things worked, while following his father and watching as he worked and on numerous electrical devices, mostly stereos, amps and speakers. Fascinated with how life was given back to these devices, this encouraged him to want to learn to do the same thing. So his father taught him everything he needed to know to diagnose and resolve an issue with an electrical device. In elementary school he already understood what resistors and capacitors were, and how to use a solder iron. As he got a little older he tried to take his knowledge and talent to a new height by attempting to invent things that would make things easier around the house. One particular invention dealt with creating a helicopter, using parts from other toys and an old RC race car.

Once high school came along his older brother had graduated college with a bachelor's of science in Electrical Engineering at North Carolina Agricultural and Technical State University. While deciding a career path his freshman year in high school John also began to make plans to become an electrical engineer and made a promise that he would get a Master's degree. He made a promise to his parents that he would make it, and make a difference. His mother was supportive no matter the choices or path he would take. Her main concern was education. Now while working for an Aerospace Company his dreams are unfolding as he finishes up his journey of getting his Master's degree.

Dedication

I dedicate this thesis to my parents who pushed and made sure all of this could happen.

Acknowledgments

I would like to acknowledge and thank God for giving me the knowledge and strength needed over the years to make it this far in my education career as well as my parents, Lynwood and Rachelle Campbell, who paved the way for me to be the best I can be.

I would like to acknowledge my committee members, Dr. Gary L. Lebby, Dr. Christopher C. Doss, and the chair of my committee Dr. Corey A. Graves, for being very supportive and available to me when I needed them.

I would like to give a special thanks to Mr. Howard V. Hardiman, Mr. Malcolm Bethea, Dr. Clay Gloster, Dr. Ibraheem Kateeb, Dr. Corey A. Graves, and Dr. John C. Kelly. They all saw the potential in me, they pushed and helped me get to the next level; without them this big dream of mine, wouldn't be a reality.

To my siblings, Ralanda Waller, Shawn Campbell, Lynwood Campbell Jr., Ralyn McDougald, Keith McDougald, Alicia Campbell, who were there throughout my journey to help, provide, and give words of wisdom.

To my nieces and nephews, Daisha Campbell, Peoria McDougald, Keith McDougald Jr., Iyanla A. Campbell, Alana Waller, Jaylen Waller, Lynwood Campbell III, Amariane Campbell, Illyana Campbell, Arielle Z. Campbell, who always help lighten my journey with laughter and good times.

To my fiancée, Sasha N. Munn, who encouraged and kept me uplifted throughout my journey. She was always by my side, doing everything she could to make sure I was not distracted from my studies, her main phrase was, "do not worry about the house; do what you have to do for school". She packed lunches, made snacks and checked on me regularly on long nights. She made sure I was well taken care of and focused.

Last but not least, to the luxurious North Carolina Agricultural and Technical State University, for giving me the opportunity to receive a bachelor's degree and now a master's degree. NC A&T as a whole has had a huge impact on my life in the most positive way, and I will never forget those who help push and mold me into who I am today.

Table of Contents

List of Figures

List of Tables

Abstract

In order to understand and do well on a particular subject while in class, it is important to be completely focused on the subject that is being presented. One cause of poor student performance is a lack of focus while in class. Most students today have various forms of mobile technology that has greatly evolved. Students bring mobile devices such as smart phones and tablets to class. Even though mobile devices can be beneficial for a student, they can be very distractive.

The main objective of this research is to create a pervasive system that will autonomously monitor the distractive use of mobile devices in a classroom setting. The main features used to determine whether a student is distracted by a mobile device are the eyes and the state of the hands. Being able to track where the student is looking is very important in determining distraction.

The system revolves around the use of a Microsoft Kinect v2 and uses gesture recognition, hand tracking, elbow tracking, machine learning, face tracking, and eye gaze. Each feature is monitored by its own classifier to help the system determine whether a student is distracted by a mobile device. Eye gaze will determine whether the student is looking down or not, hand tracking will be determining whether a student is holding a device or not, and machine learning will aid in determining whether a student is holding a device by comparing the students posture to popular mobile device use posture.

Results show that the eye tracking mechanism works very well. With the hardware setup, the system was able to get 100% accuracy in determining whether a student was looking down or not. Hand tracking accuracy was 80% and trained gesture tracking, developed by collecting

posture data from human subjects, was accurate 80% of the time. Overall, the system was correct 80% of the time.

CHAPTER 1

Introduction

"Are you distracted by a mobile device in class?" is a very difficult question to answer for students, because that is not something they want to admit to a teacher especially if their grades are not exactly reflecting attentiveness. Today mobile devices are more tolerated in classrooms. They can be beneficial and can be used to follow lecture notes. However, even though mobile devices can be useful educationally, there are many other uses for a mobile device.

Today, social media can be taken everywhere there is a mobile device also and where the internet is accessible. Wifi is available free to students for educational purposes, but the access will be abused if it is not monitored, which is mostly the case. The biggest trending social media sites today are Twitter, Instagram, Facebook, Vine, and Snapchat. There are also popular video apps such as, YouTube and Worldstar. All of these apps are for entertainment purposes. Once a student is disengaged in a lecture, they can simply turn to a mobile device to be entertained.

Using social media on mobile devices does not always mean that a lecture is boring or the student is disengaged. While using the device for educational purposes the student can receive a notification from one of the listed sites and immediately feel eager to check the notification. This incident can cause a chain of interruption. The student could respond or comment to someone on a social network site which could cause a mobile conversation to spark. This can lead to text messages and emails Depending on whether the device is on silent or not it can be distracting for surrounding students The vibrate setting can be distracting also, especially while the device is sitting on a desk. A recent student has shown that students perform worse on assignments due to mobile device distraction. The test was conducted during a surprise quiz with a ringtone sounding periodically. [13]

This research is not just for catching students that are distracted by mobile devices. This research will aid in remedying problems with relaying material to students. Some students are easily distracted because of the style of teaching, the student could be disengaged, or just not understand the material relayed, and give up on what the teacher is attempting teach to the student. In order for a student to be successful

Eye gaze is big indicator for determining whether a student is distracted. The Student's eyes may tell whether they are paying attention or not. During a lecture it is natural for students to look down for a few seconds to take notes, however while the teacher is presenting material they often point out things of importance or work through examples, which require the students to keep their eyes on the teacher for a majority of the lecture.

Hand tracking will work in parallel with eye gaze to determine if the student is holding a mobile device and a trained gesture will help determine the use of a mobile device. Hand tracking will be used to track hands on the X-axis, measuring the distance between both hands to determine if a student is holding a device. Trained gestures are postures used to compare to the student's posture, to determine the distractive use of mobile devices. Eye tracking will be used to determine whether a student is looking down at a mobile device.

This system autonomously monitors the student to determine whether the student is distracted or not. It provides an answer to the following question; "how can a pervasive computing system be developed to autonomously monitor the distractive use of mobile devices?"

A multispectral device can be used to develop an autonomous system to monitor the distractive use of mobile devices in a classroom. A Microsoft Kinect can be used, due to its multiple sensors such as, an RGB camera, infrared camera, and microphone. The RGB camera and infrared camera will be the only sensors needed for this research.

CHAPTER 2

Background

2.1 Gesture Recognition

"Gesture recognition pertains to recognizing meaningful expressions of motion by a human, involving the hands, arms, face, head, and/or body." [9]

Gesture Recognition is a powerful tool that can be used to determine someone's body posture. It is crucial to be able to create unique postures to fit certain scenarios. In Korea there was a study done on smart houses, the smart house was design for disabled individuals. The smart house was able to communicate with the user through specific body gestures different events were triggered in the house, manually and autonomously, depending on the gesture. If the individual was in danger it would react autonomously. If the individual wanted to move to another room the individual would manually give a gesture to be moved to another room. [12]

2.2 Skeleton Tracking

A robotic smart house was developed to assist elderly people with everyday needs. "A Hidden Markov Model (HMM) based human behavior recognition using lower limb joint angles and body angle".[1] That's how a healthcare robot would monitor an individual's motor functions. The robotic smart house development results displayed that joint trajectory could be analyzed with high accuracy compared to a motion tracking system, and human behavior could be recognized from the joint trajectory.

The Kinect uses structured light and machine learning to track gestures. Structured light is used to create a depth map of the human figure, and then tries to determine the body location using data collected during machine learning. The Depth map is constructed by using a speckled pattern of infrared light like shown in Figure 5. Structured light is used to display known patterns over objects and display the contorted results, as shown in Figure 6. When deciding where and what the figure is doing, a choice is made using a decision tree, like the one showed in Figure 7.

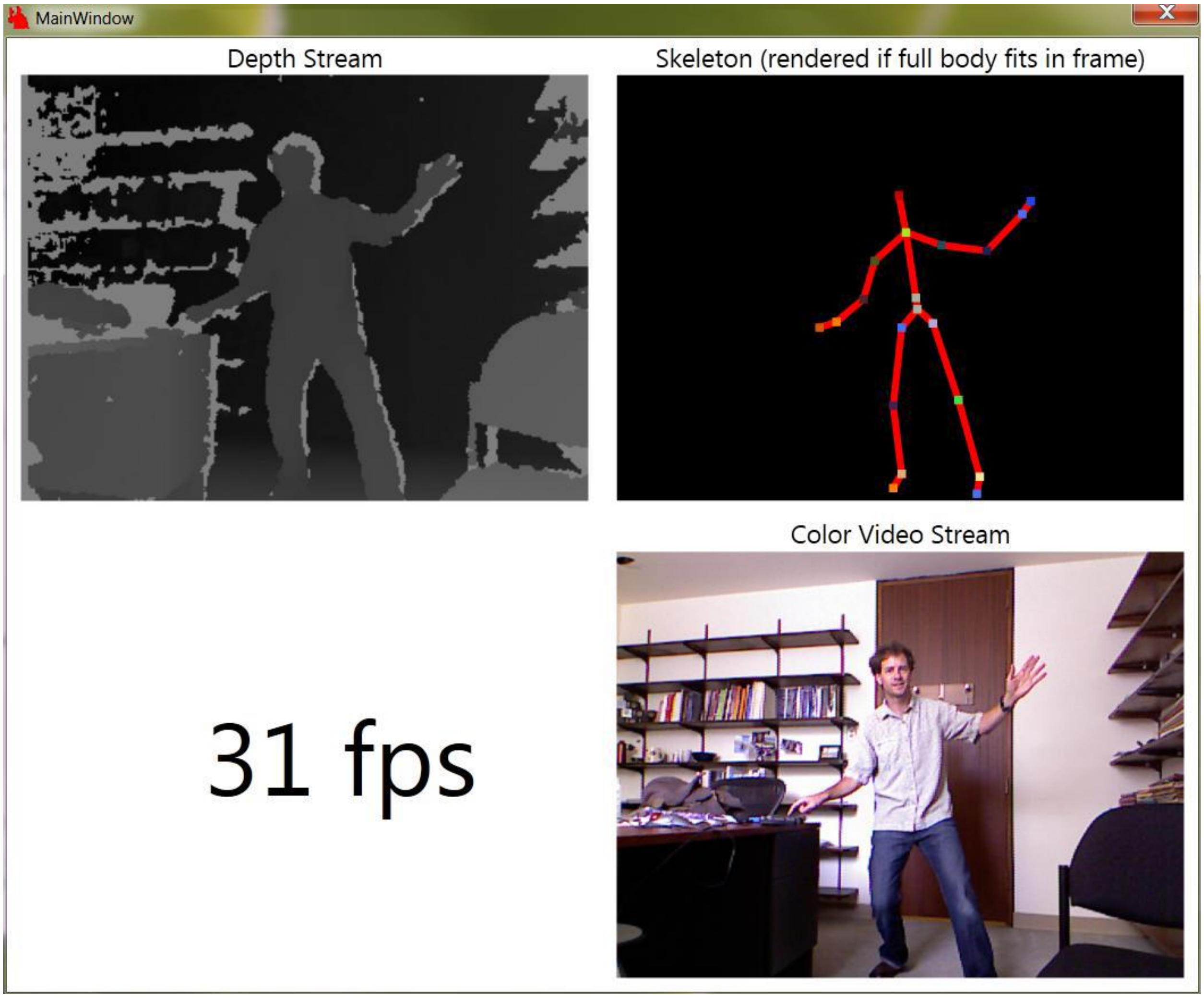


Figure 1. Depth map and skeleton. [3]

Figure 2. Example of structured light. [3]

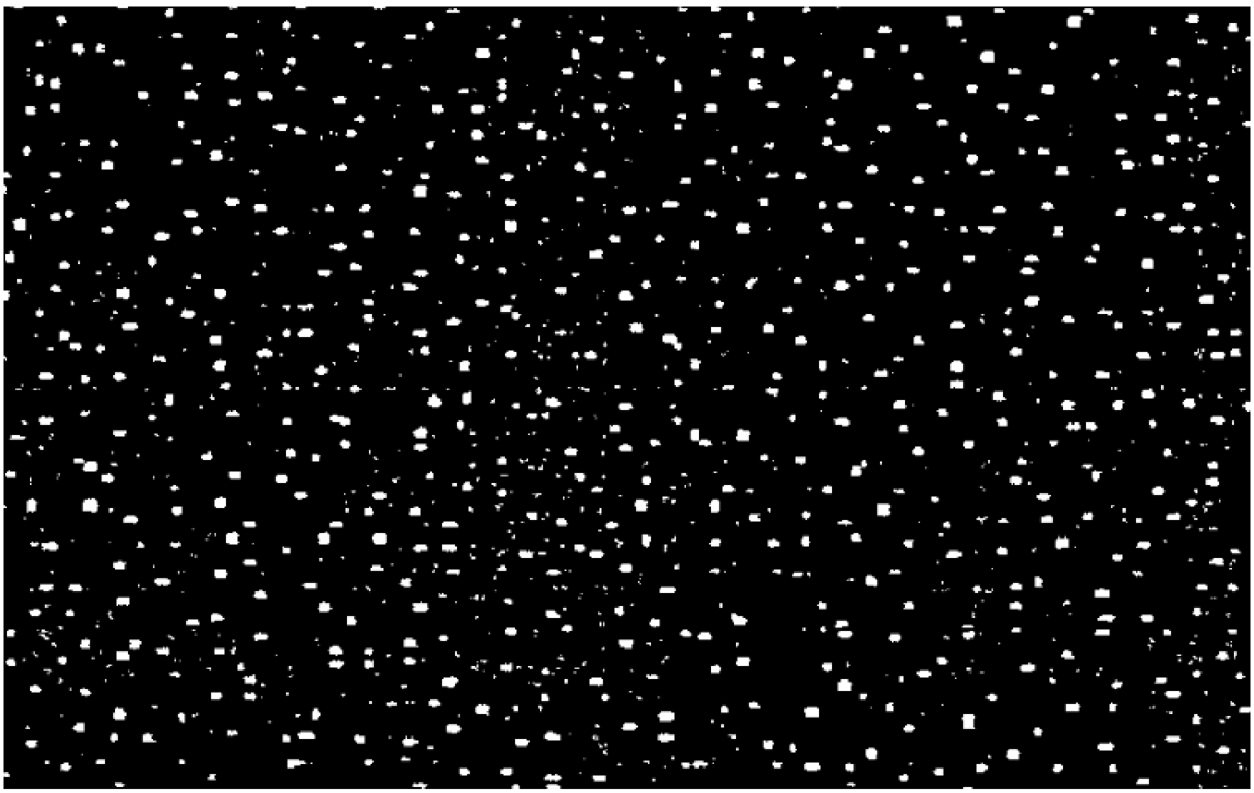


Figure 3. Speckled light used for Tracking human figures. [3]

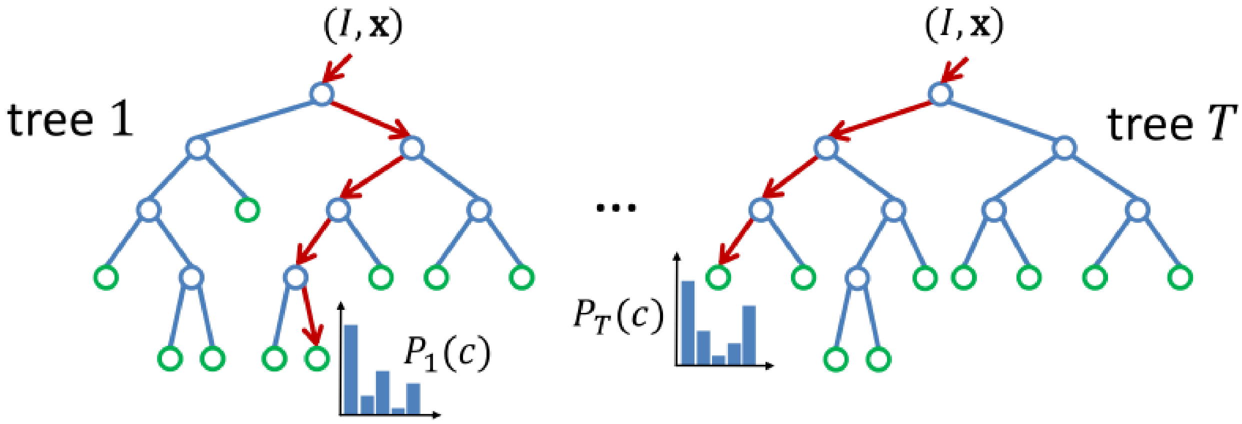


Figure 4. Decision tree used for determining joint positions. [3]

2.3 Eye Gaze

Eye gaze tracking is the process of trying to estimate the approximate direction that an eye is looking. The center of the eye can be tracked independently of one another according to research; eye gaze can be estimated using one eye. The iris is located by locating the corners of the eye then finding the center. Using an algorithm to locate the outer sides of the iris a circle can be created that tracks the iris with in the eye. [13]

2.4 Multi-Spectral Imaging Device

The Kinect is a device created by Microsoft used for tracking human skeletal structures, originally intended for game play, games that included Dance Dance Revolution, Zumba and many more. All of which requires the user to stand and use their entire body to perform certain task and dance moves. This all revolves around light cardio and fitness.

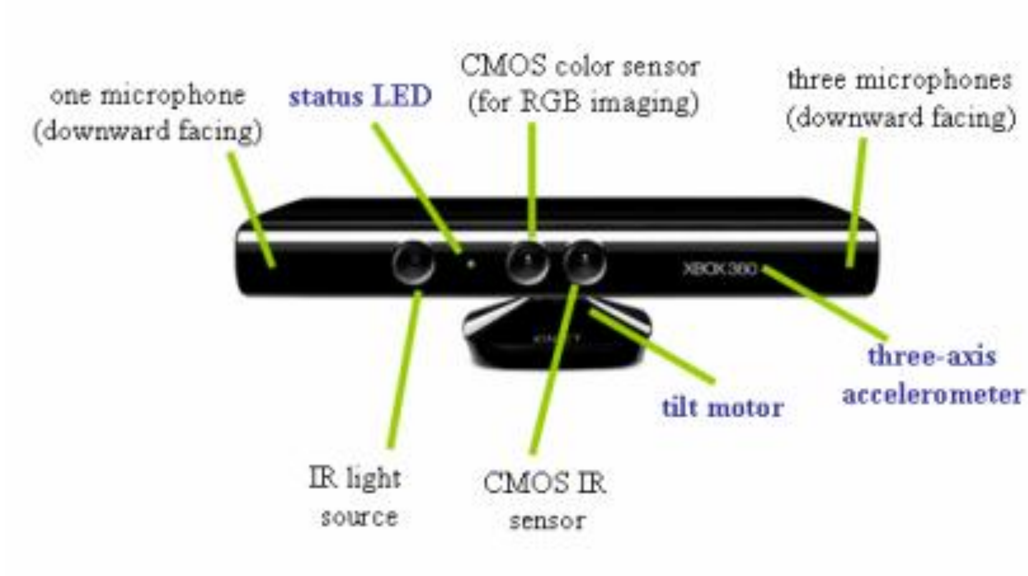


Figure 5. Kinect for Xbox360.

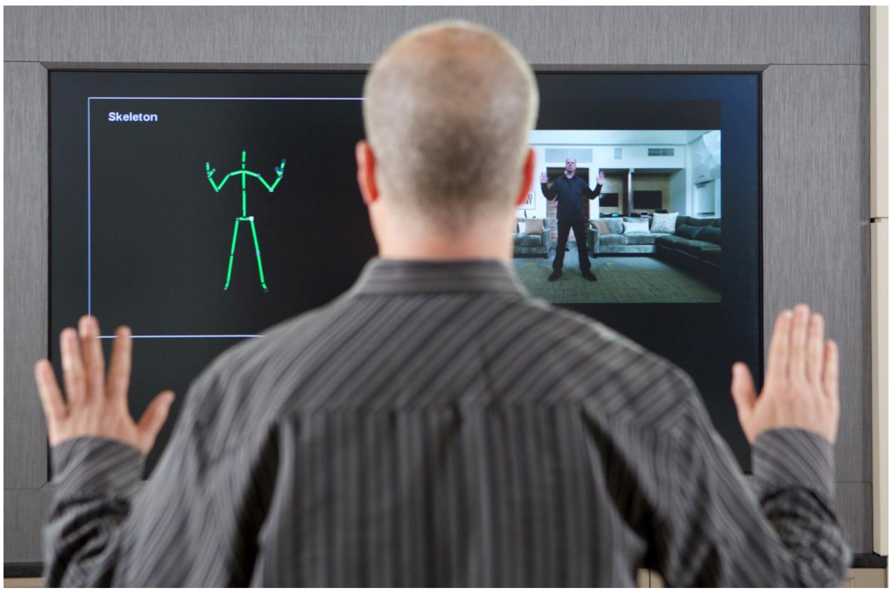


Figure 6. Skeletal capture by Kinect. [10]

Although the Kinect was innovative after its release 2009 and later releasing a software development kit (SDK) for commercial development. In 2014 a new Kinect was born. An SDK was created for version 2 of the Kinect to finally jump start some fresh innovation with the Kinect.

Figure 7. This figure displays Kinect v2.

2.4.1 Kinect New Features and Comparison. Facial recognition was a huge improvement. The Kinect v2 has the ability to capture one's own face and import it onto the main character a game. The Kinect v2 has upgraded its original features such as the RGB camera, the Infrared sensors and even the microphone array was implemented differently. The ability to track more people with one Kinect was also implemented. The Original Kinect was developed to track two full skeletons at a time. Version 2 of the Kinect can track up to 6 full skeletons at once. One nice feature from the original Kinect was removed from version 2 of the Kinect, and that was the tilt motor feature. The original Kinect had the capability to allow programmable control of the tilt of the sensor to adjust to the skeleton if needed. Version 2 of the Kinect has to be adjusted by hand. However, due to its wide angled view its tilt wouldn't need to be changed after being setup once in its targeted area. Some of the major differences between the two Kinects are listed in Table 1.

Table 1. Differences between the Kinects.

Feature	**Kinect Version 1**	**Kinect Version 2**
Color Camera	640x480 @ 30fps	1920x1080 @30fps
Depth Camera	320x240	512x424
Max depth Distance	~4.5M	~4.5M
Min Depth Distance	40 cm in near mode	50 cm
Horizontal Field of View	57 degrees	70 degrees
Vertical Field of View	43 degrees	60 degrees
Tilt Motor	Yes	No
Skeleton Joints Defined	20 joints	26 joints
Full Skeletons Tracked	2	6
USB Standard	2.0	3.0
Supported OS	Win 7/8	Win 8

2.4.2 Microsoft Kinect SDK. The Kinect 2.0 SDK has many new features that aid the user in creating new and improved apps. Body, hand and joint detection have been greatly improved this SDK is capable of detecting thumbs on hands to more accurately detect a closed, open and thumbs up hand orientations. A huge added feature is the ability to train your own gestures into the system using the machine learning software feature. The SDK comes with an application called *Gesture Builder*, which will record data and store and build code around the recorded gesture, giving the user the ability to create and explore gestures outside of the presets without in-depth manual coding. According to the manufacturer, the resolution and number of points plotted while tracking faces has improved giving more accurate results. This enables the ability to track more significant facial expressions and more areas of the face. This version of the SDK is capable of allowing multiple applications to share the same sensors. The original Kinect was not able to do this without having some sort of service in between applications that distributed the sensor data accordingly. That, however, made things confusing and complicated.

CHAPTER 3

Methodology

3.1 Research Idea

3.1.1 Hypothesis. This research seeks to the following question: "How can a pervasive system be developed to autonomously monitor the distractive use of mobile devices in a classroom setting?" In response to this question, it is hypothesized that a multi-spectral image capturing device can be used to autonomously monitor the distractive use of mobile devices in a classroom setting by tracking body movement and posture.

3.1.2 Procedure. This research assumes that the class allows mobile device usage during class lectures. A Microsoft Kinect v2 is used for this research. The Kinect is used to track hands, create posture gestures, and track eye gaze. The System detects whether a student is distracted by a mobile device. A posture gesture is created to determine if a student is using a mobile device. A custom gesture is shown in Figure 8. Figure 9 displays the gesture within the system application. The gesture has multiple actions that consist of popular mobile use postures.

The system also contains eye tracking, which utilizes the students' eye lids to determine whether they are looking down. The eye tracking has a timing system that initiates only when the student is looking down. After observing distracted students in a classroom the results concluded that students who are attentive tend to look up at the professor frequently. This find led to believe a student who is looking down for more than 10 seconds while in a distractive use posture may be using a device for distractive use. The timer is set to 10 seconds, so if the student is looking the down for more than 10 seconds they are classified as being distracted. Typically a student looks down on an average of 5 seconds to take notes then looks back at the board or professor indicating attentiveness. If the timer is triggered and it reaches the 10 seconds the system will assume the student is using a device. The hands will determine whether a student is holding a mobile device. It tracks the distance between the hands on the x-axis. If the hands are touching according the plotted skeleton hands, the system assumes the student is holding a mobile device.

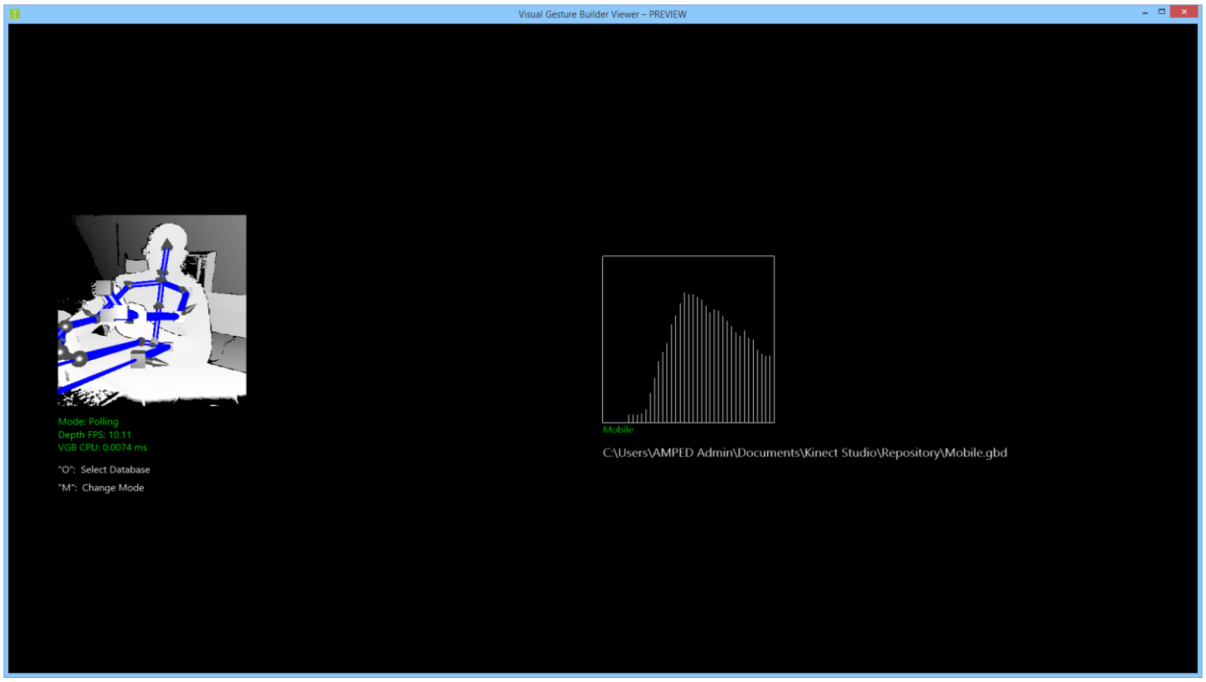


Figure 8. A custom gesture being tested.

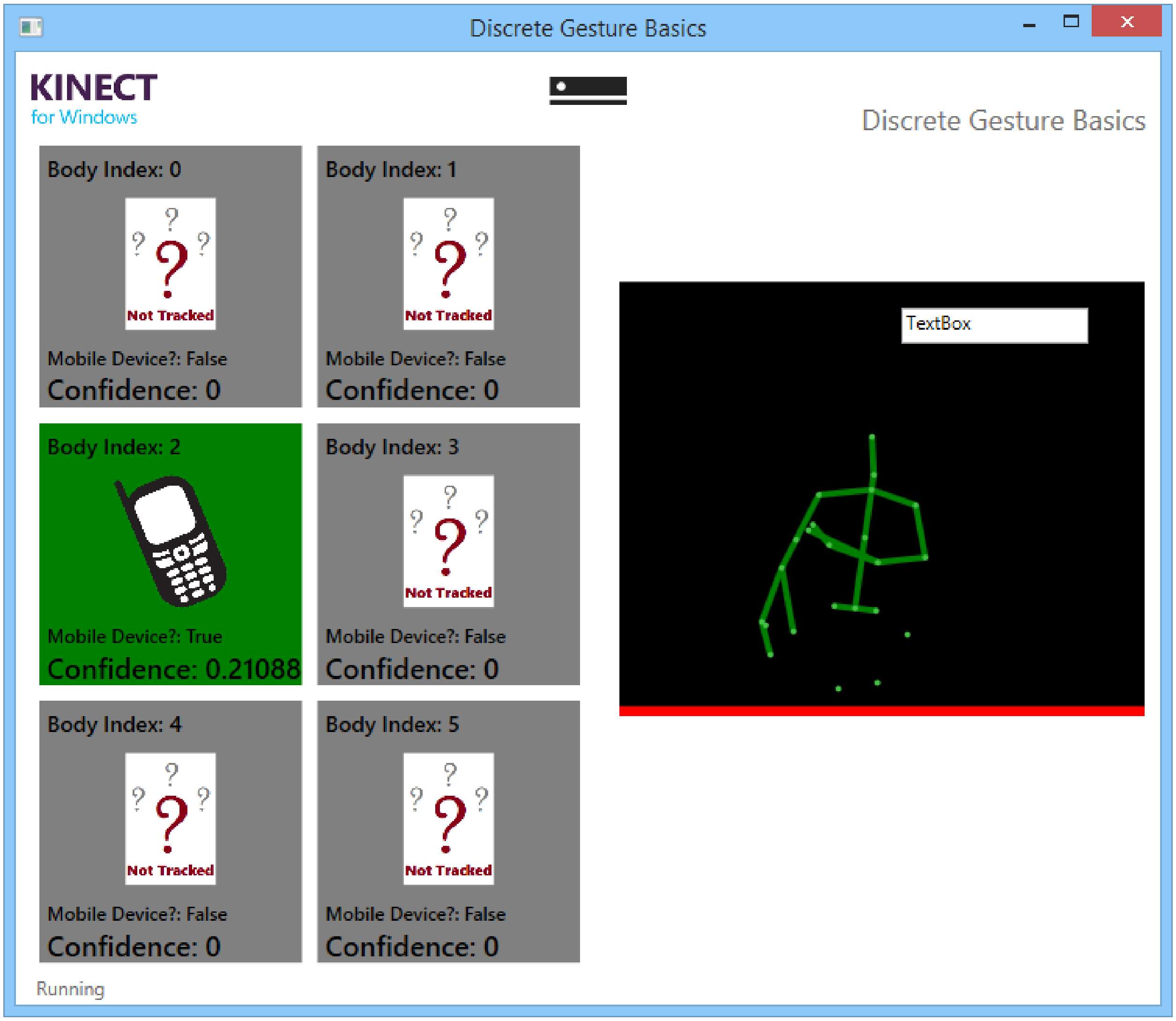


Figure 9. An in app test for the custom gesture.

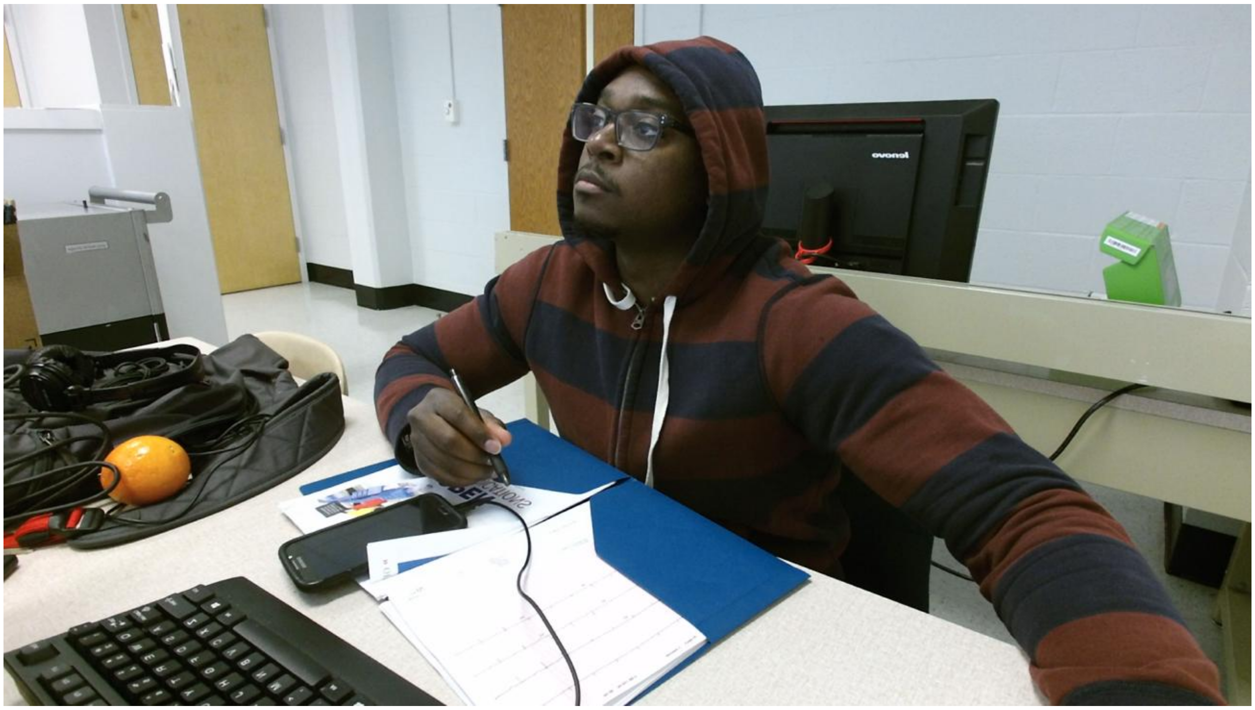


Figure 10. Attentive student.

Figure 11. Distracted student.

3.2 Prototype Development

During the prototype development the system was setup to focus on one posture group at a time to ensure visibility, accuracy and a proper position for the Kinect. Before testing the whole system with everything tied together, there had to be certainty that the Kinect could function in the environment. This included issues with lighting as well as the amount of space that is available around the workstation/desk. Most classrooms with workstations have an ample amount of desk space, to compensate for the computer and for space for the students belongings, including space to work, as shown in Figure12.

Figure 12. Typical workstation space.

3.2.1 Visibility. Version 1 of the Kinect required having lots of light, but natural sunlight works best. It's hard to know whether there will be lots of natural sunlight in a classroom setting if any at all. Version 2 of the Kinect has an auto light enhancing feature that adapts to the amount and type of available light autonomously depending on its environment.

3.2.2 Accuracy. The accuracy of tracking the student will make a huge difference in determining what she is actually doing. If the joints that are being plotted are too jumpy, this will mean the results will hardly ever be accurate. The Kinect Version 1 isn't perfect but at longer distances it had more accurate joint readings. The Kinect Version 2 has shown great improvement with locating and following plotted joint movement. Version 1 of the Kinect needed an attachment for smaller rooms or if the Kinect needed to be closer to the user, 2ft away. However, Version 1 of the Kinect did not perform much better with this attachment, as it wasn't designed to be used that close to its users. Version 2 picked up the slack however. It can view and recognize the user at much closer ranges. Version 2 of the Kinect also plots more skeleton points than Version 1, resulting in it giving more crisp skeleton and joint plots at close and long ranges.

3.2.3 Design Specifications. The placement of the Kinect is very crucial. Testing the use of the Kinect from directly under the monitor led to negative results as shown in Figure 13. For version 1 of the Kinect, the lighting is just too poor to function correctly while underneath the monitor. Version 2 of the Kinect takes a little longer to recognize the user due to trying to automatically adapt to the lighting in the environment. When the Kinect attempts to correct the lighting in dark areas, it can potentially over light the area, resulting in the features getting distorted or washed out, becoming un-viewable to the sensors.

3.2.4 Current Working Locations. Version 2 of the Kinect seems to function a little better while elevated and off to the side, as shown in Figure 12. This is due to being exposed more to the lighting in the environment. That angled elevated view gives about 2ft of distance from the user to compute more accurate results. The location of the Kinect in Figure 14 was used to capture the plotted joints, including the hand and elbow indicators shown in Figure 16. However, the best location for the Kinect is displayed in Figure 15, it is elevated, exposed to light and is able to capture all user data; such as, looking down at a mobile device and body posture. This position also gives the necessary distance the Kinect needs to function correctly while being centered to gain better eye data.

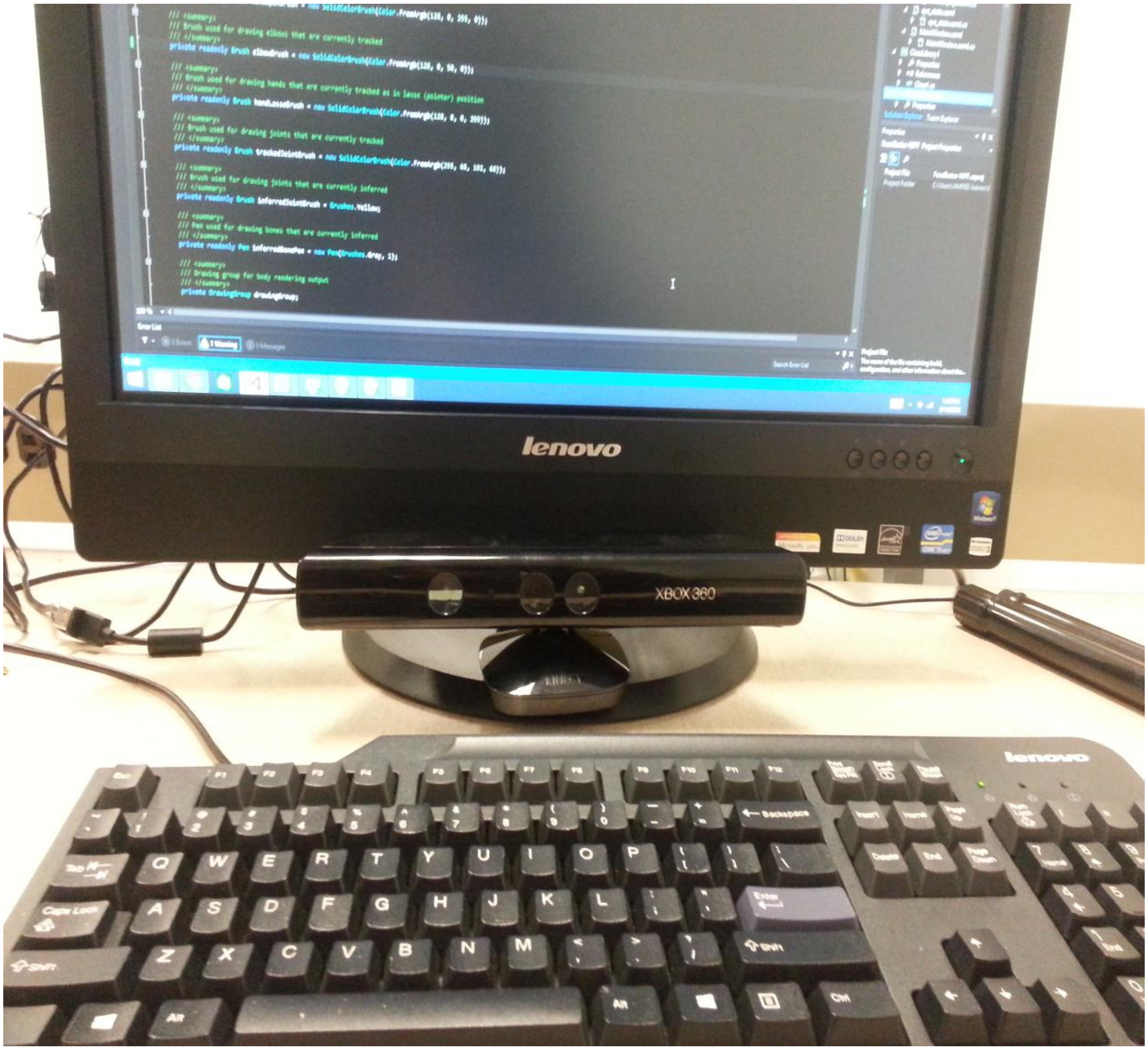


Figure 13. Kinect underneath the monitor.

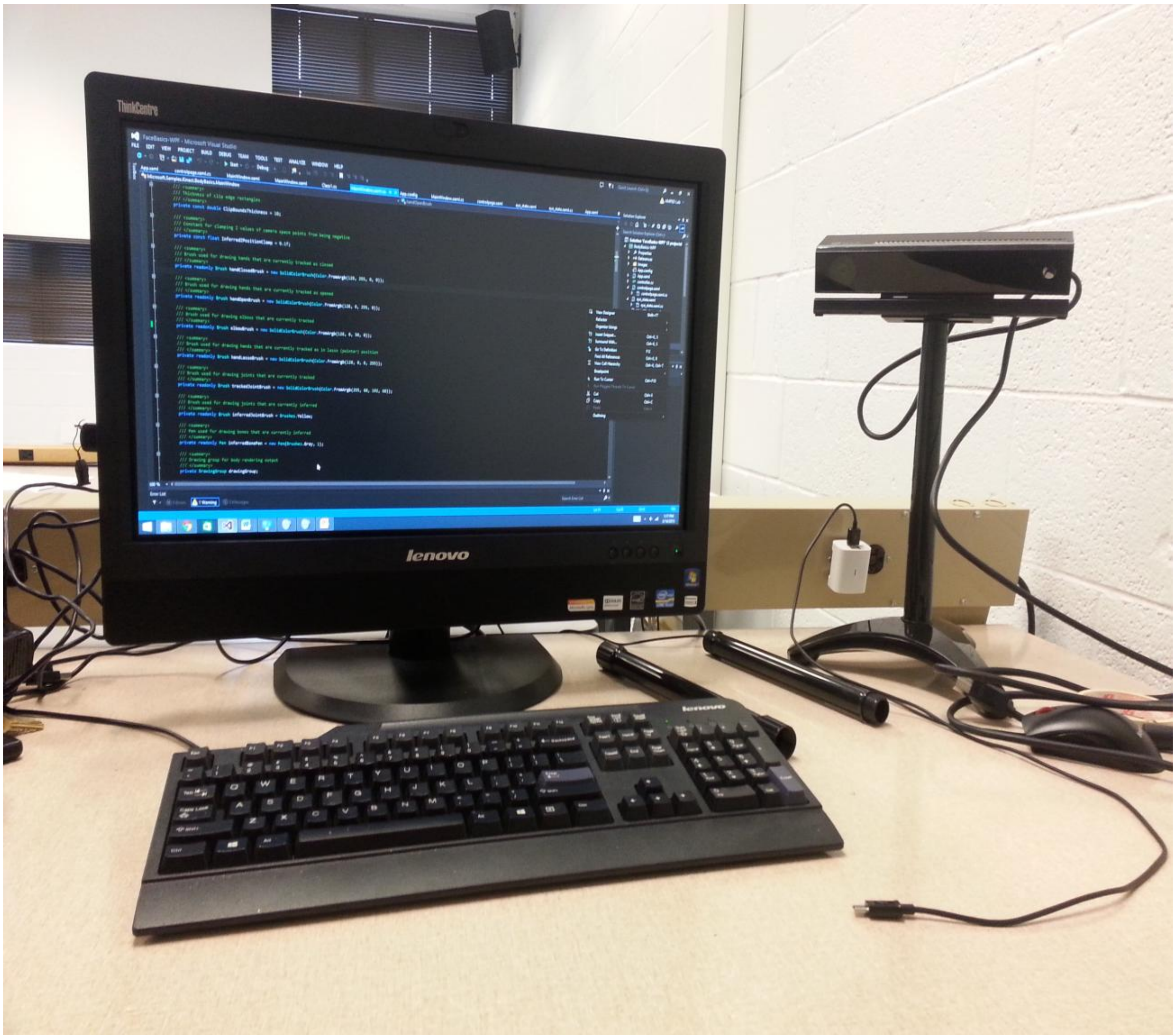


Figure 14. Closer view Kinect location.

Figure 15. Best Kinect location.

3.3 The System Process.

The system looks for data from the student that consists of tracking the eyes, hands and elbows. Since Version 2 of the Kinect cannot be programmatically tilted it will be difficult to use the tilt of the head data. However, the tracking of the eyes works well enough to determine distraction independently from the pitch of the head. Figure 16 displays a level 0 block diagram of the overall flow of data that will take place. Figure 17 displays in more depth the way the system will receive and handle the data collected from the student.

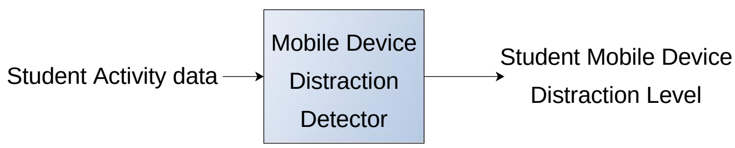


Figure 16. Level 0 block diagram.

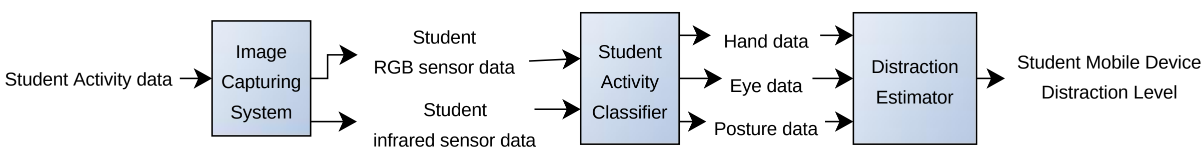


Figure 17. Level 1 block diagram.

3.3.1 What the user does. The main user of the system will be the professor. However they do nothing to or with the system. Once setup, the system will autonomously monitor the student, collecting and analyzing the data. The student will act naturally in her environment, and the system will do the rest.

3.3.2 Determining Attentiveness. The system has classifiers that deal with each bundle of data received from the student. The Bundle of data is collected from the hand classifier, eye classifier and posture classifier. The classifiers work together to determine Attentiveness.

3.3.2.1 *Hand Classifier.* The hand classifier analyzes the X-axis position of the hand and the distance between the hands. This classifier determines whether a student is holding a mobile device. The main window will update indicating whether the student is assumed to be holding a mobile device or not.

3.3.2.2 *Eye Classifier.* The eye classifier tracks whether the eyes are closed are not, which can be viewed in Figure 19 and 20. The eye lids are used to determine whether a student is looking down. When the student looks up or down the upper eye lid naturally follows the eye direction. So, while your eye is assumed to be closed or open, it also indicates the user is looking up or down.

3.3.2.3 *Posture Classifier.* The posture classifier will be monitoring the posture of the student in real time, comparing it to trained postures stored in its database. This data base was created by monitoring real subjects as they used mobile devices for distractive use.

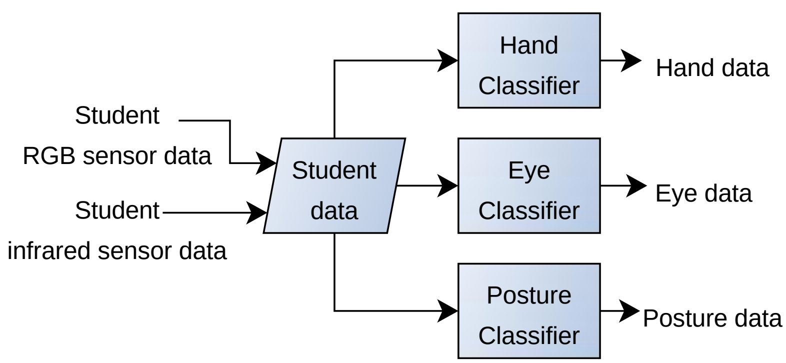


Figure 18. Student activity classifiers.

3.4 Overall System.

With all of the classifiers working together the system attempts to determine the attentiveness of a student. The hand, elbow and eye classifiers work simultaneously and continuously.

3.5 Cases.

Each classifier has cases that help determine the state of the tracked data. Hands, eyes and posture individually determine the state of data it tracks. After the cases have determined its outcome the system determines whether a student is distracted by a mobile device or not.

3.5.1 Hand Case. If the hands are a certain distance apart then the system will automatically assume the student is using a mobile device. However the hand classifier does not look for devices being held by one hand. The posture classifier will look for cases that involve the student holding a device with one hand and also using a device with one hand.

3.5.2 Eye Case. If the hands classifier gets triggered to believe the student is holding a mobile device, the eye tracking classifier will monitor the eyes on a timed basis. If the student is considered to be holding a device and the eyes are looking down, then the system will determine that student is distracted by a mobile device.

3.5.3 Posture Case. The posture classifier constantly looks for the postures that it has been trained to identify. This classifier makes its decision on whether the student is in a distractive use posture based on data collected from human subjects. Human subject posture data was collected and stored in a database for the posture classifier to refer to while monitoring a student.

The distractive posture gestures created in this research were made using the Kinect Gesture Viewer. This software allows the user to create a gesture of any sort. The user records a video of action using the software then imports that video to mark the desired area of the video that indicates the target action to be recognized. After the gesture is marked and built, it can be placed in the application for use.

3.6 Group Testing

The Distraction Measurement Application was developed to gather distractive posture behavior, while at the same time determining the level of distraction.

A publicly available comprehensive instrument was chosen, the level of randomness played a major role in the decision for the choosing the instrument used. In this case, the instrument used is about animals using their spit for survival (refer to appendix C). The instrument consists of a reading passage followed by a set of questions intended to assess the understanding of the passage. The application starts by welcoming the user and then asks the user to choose a study group, as shown in Figure 19. There are 3 study groups. Each group will sit through a computer generated slideshow that displays all the information from the reading passage portion of the instrument as shown in Figure 20.

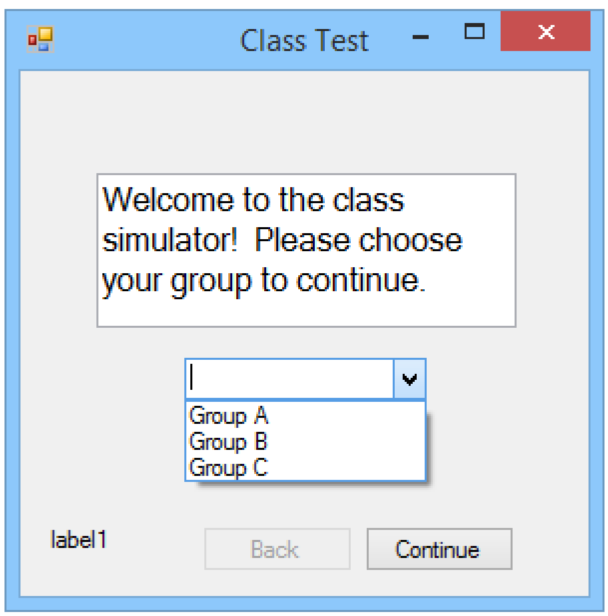


Figure 19. User choosing a group in the Distraction Tracker App.

Group A sits through the slide show without any distraction or the use of their phone. Group B sits through the slide show with the choice of using their mobile device to take notes on the presented material. Group C sits through the slideshow while receiving three distracting text messages from the application. Once received, they respond to the messages, being placed in a distractive state/posture.

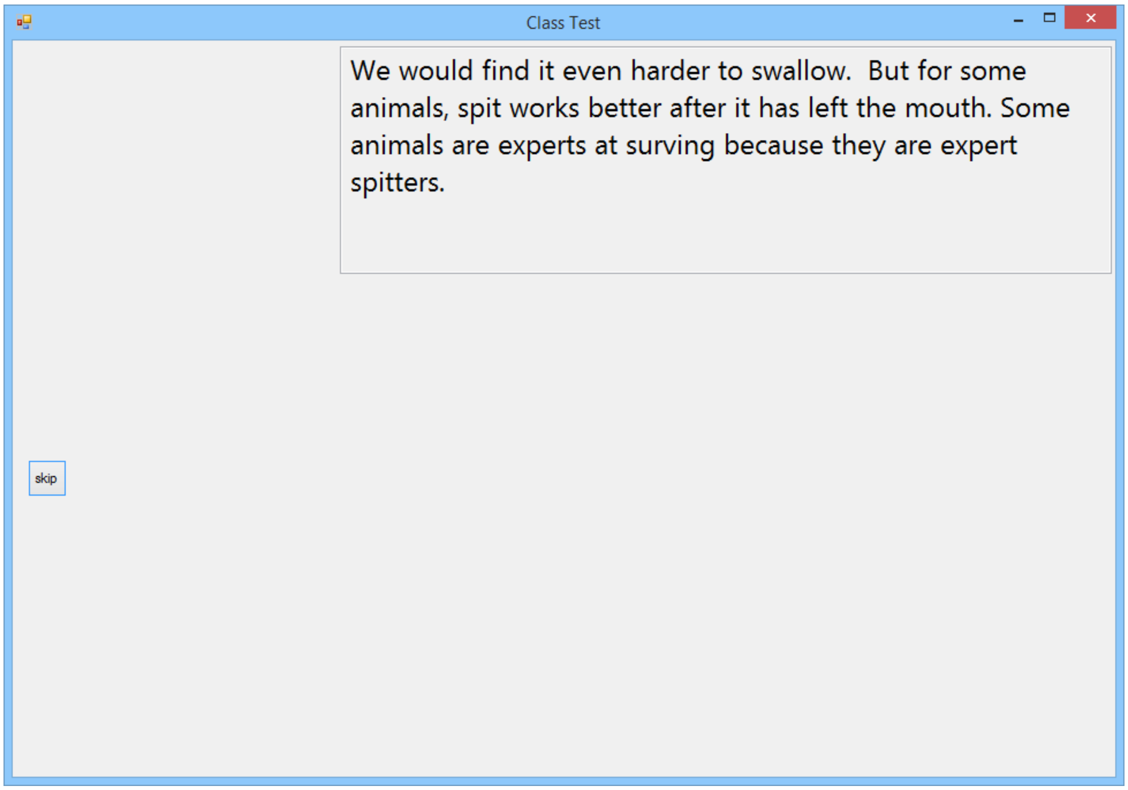


Figure 20. Slide from the instrument-based slideshow presented to the subjects.

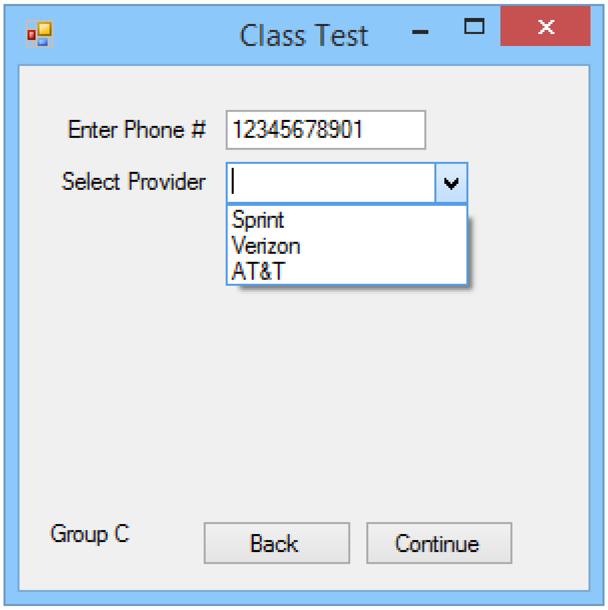


Figure 21. A user in group C inputting their mobile device information.

After the subjects complete the slideshow, a test is given on the material presented. The test is the question portion of the instrument (see appendix C). The results of the test are used to determine the level of distraction.

While the subjects are viewing the reading passage, their body posture is being video recorded by the Kinect. The video from the Kinect is transferred directly to the Kinect Studio to be used later in the Kinect gesture builder application. Use of the gesture builder is shown in Figure 22.

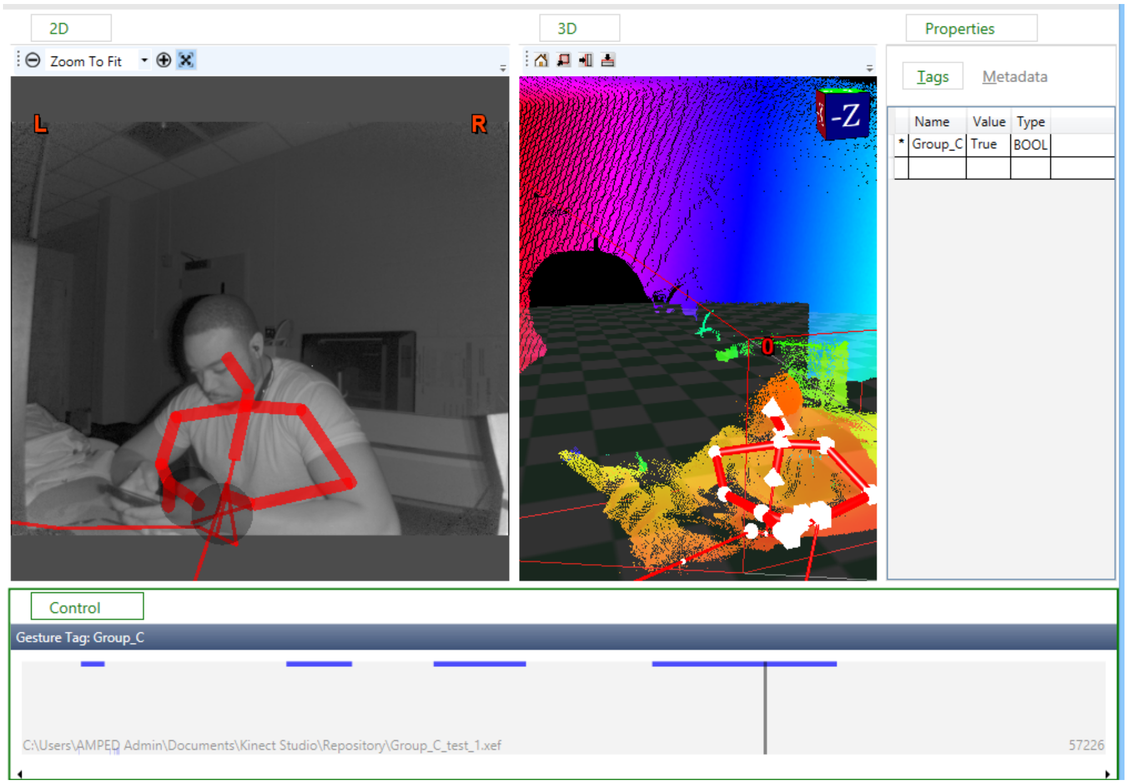


Figure 22. Use of the gesture builder.

CHAPTER 4

Results

4.1 Using the Kinect Version 1

The first stages of testing involved using version 1 of the Kinect. The Kinect gained a great reputation for tracking skeletons and being used for human/computer interactions (HCI). So naturally the Kinect is a great choice regarding using a multi-spectral device.

4.1.1 Testing with Kinect Version 1. This version was able to track the face, hands, and elbows independently. When trying to run applications simultaneously, the SDK designed for version 1 of the Kinect is not designed to share its sensors with other applications developed within the SDK. Tracking the hands, elbows, and head worked just fine, however, when adding a feature to create a mesh over the face to locate the points around the eye, the Kinect v1 was not able to handle it all at the same time. This was not the only issue regarding the use of version 1 of the Kinect. The distance the user needs to be from the Kinect plays a huge role. Version 1 of the Kinect required a range of 3 to 5 feet of distance to track the user's skeleton. However, the student will be seated at about 2ft away from the Kinect.

4.1.2 Eye Tracking with the Kinect Version 1. Eye tracking with the Kinect version 1 has been done and resolved through research. However, it's not a key feature, which means there is a lot of room for development and improvement in this area.

4.2 Using the Kinect Version 2

Version 2 of the Kinect had many upgrades from Version 1, from the resolution of the RGB camera to the overall design and style, and most importantly the SDK. The Kinect SDK 2.0 allows the user to share the Kinect's sensors with other applications. This allows more opportunities for more complex applications to be built, and for the ease of managing the overall operation of a project.

4.2.1 Testing with the Kinect Version 2. Using the same testing procedure as was used with Version 1 of the Kinect. The code was tested piece by piece before being put together. Once the Kinect was able to recognize skeletons at a reasonable distance, the process for reading the joints for that skeleton took place. Once that was accurate, indicators were made to visually show the tracked points on the skeleton, as shown in Figure 23. Textbox's were placed on the screen to show the X/Y coordinates of the hands and elbows, knowing the distance between the hands can help determine if the student is holding a mobile device. The Y axis for the elbows will help determine if the students' hands are above or below the desk.

4.2.2 Eye Tracking with the Kinect Version 2. Version 2 of the Kinect offers the advanced features that compensate for the shortcomings in Version 1. For example, Microsoft made facial tracking a main feature for the Kinect v2 in gaming and for the SDK.

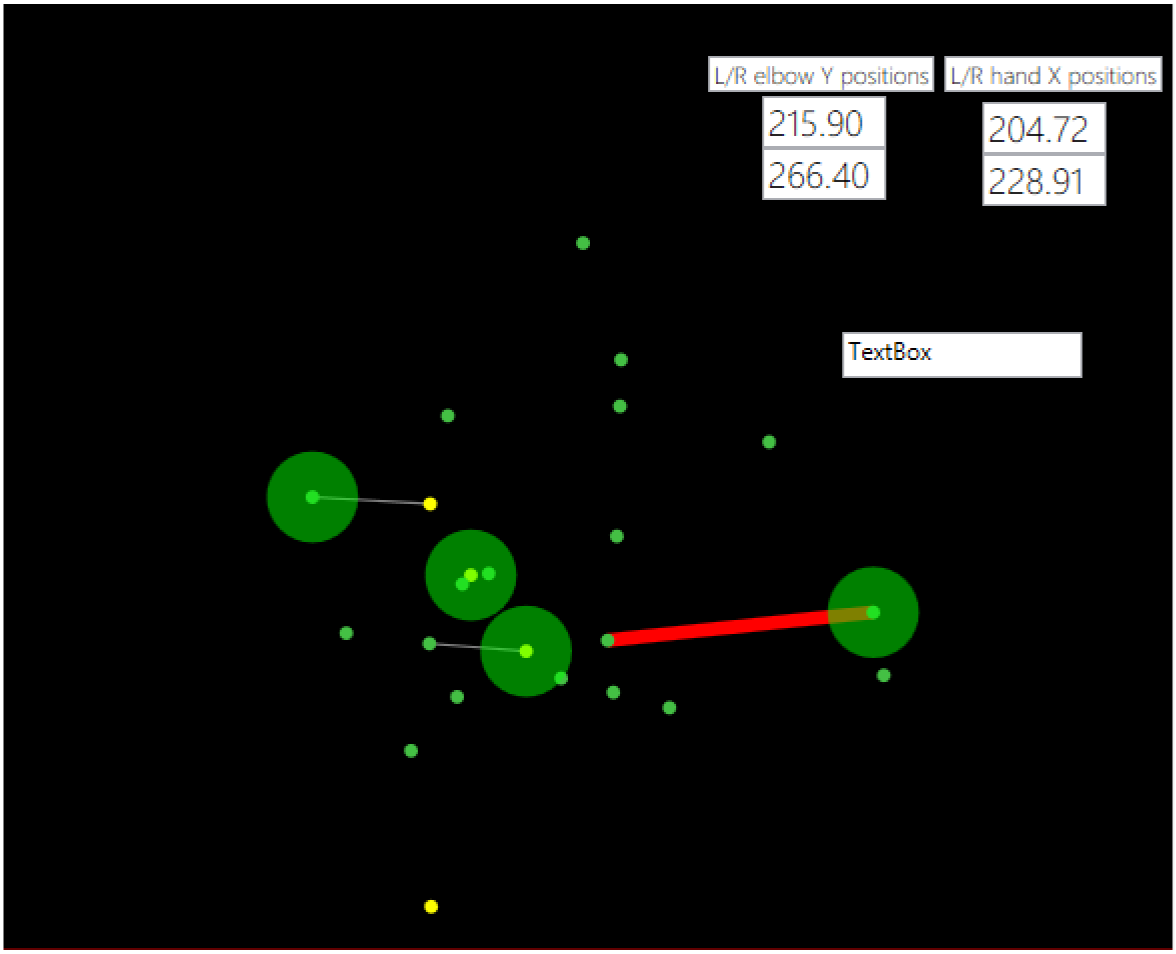


Figure 23. A student using a mobile device.

The eyes are tracked more easily and accurately with Version 2 of the Kinect and the SDK 2.0. In this research, instead of tracking the pupil for eye gaze, the user's eyelids are tracked to determine downward eye gaze. Since this research is focused on attentiveness, it isn't necessary to monitor the exact movement of the eyeballs. The system will determine if someone is looking down at their phone by looking for the movement of the top eyelid. When looking down at a mobile device (or down at anything else) the top lids tend to follow the pupil downward, as shown in Figure 24. A normal forward gaze is shown in Figure 25.

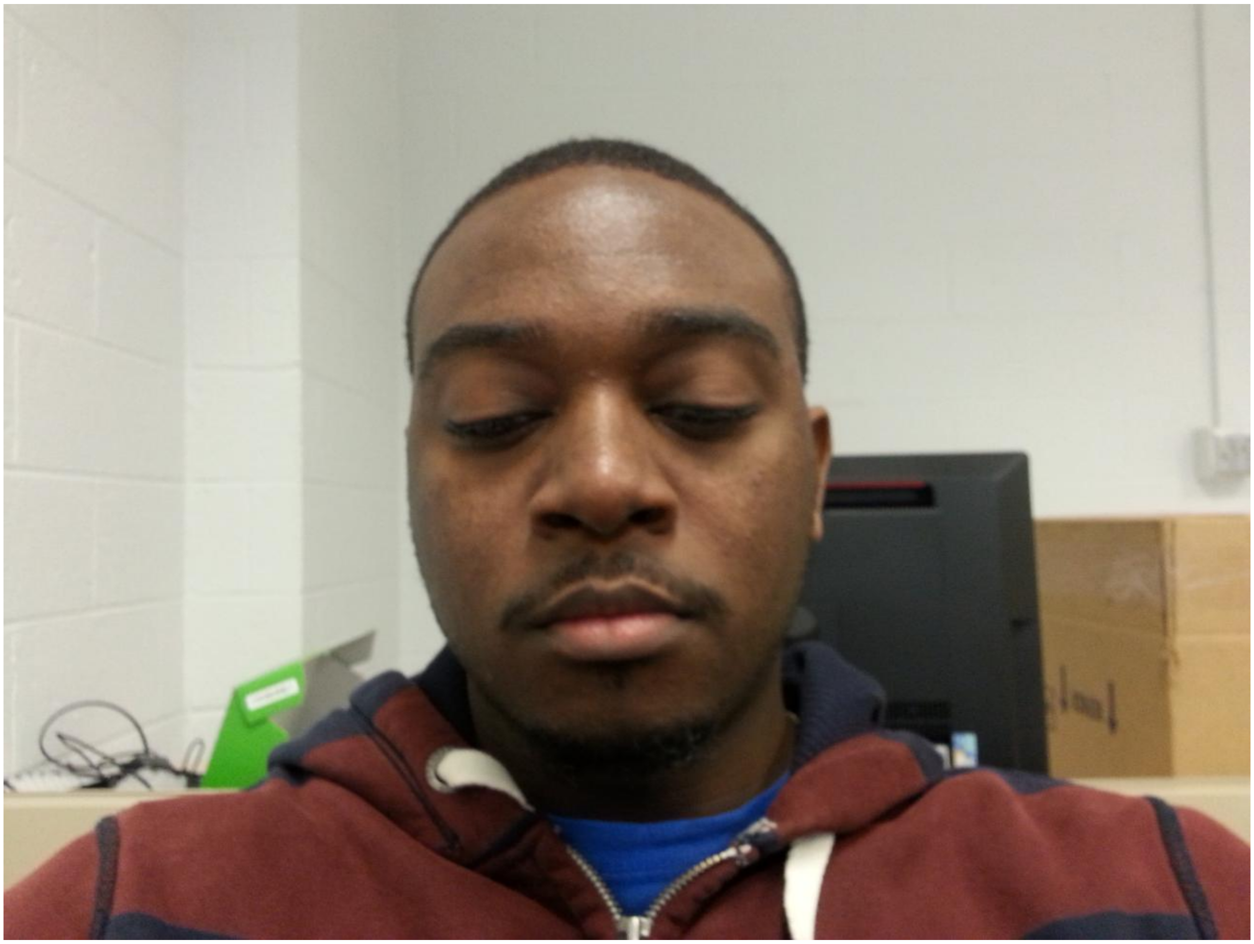


Figure 24. A student looking at their mobile device.

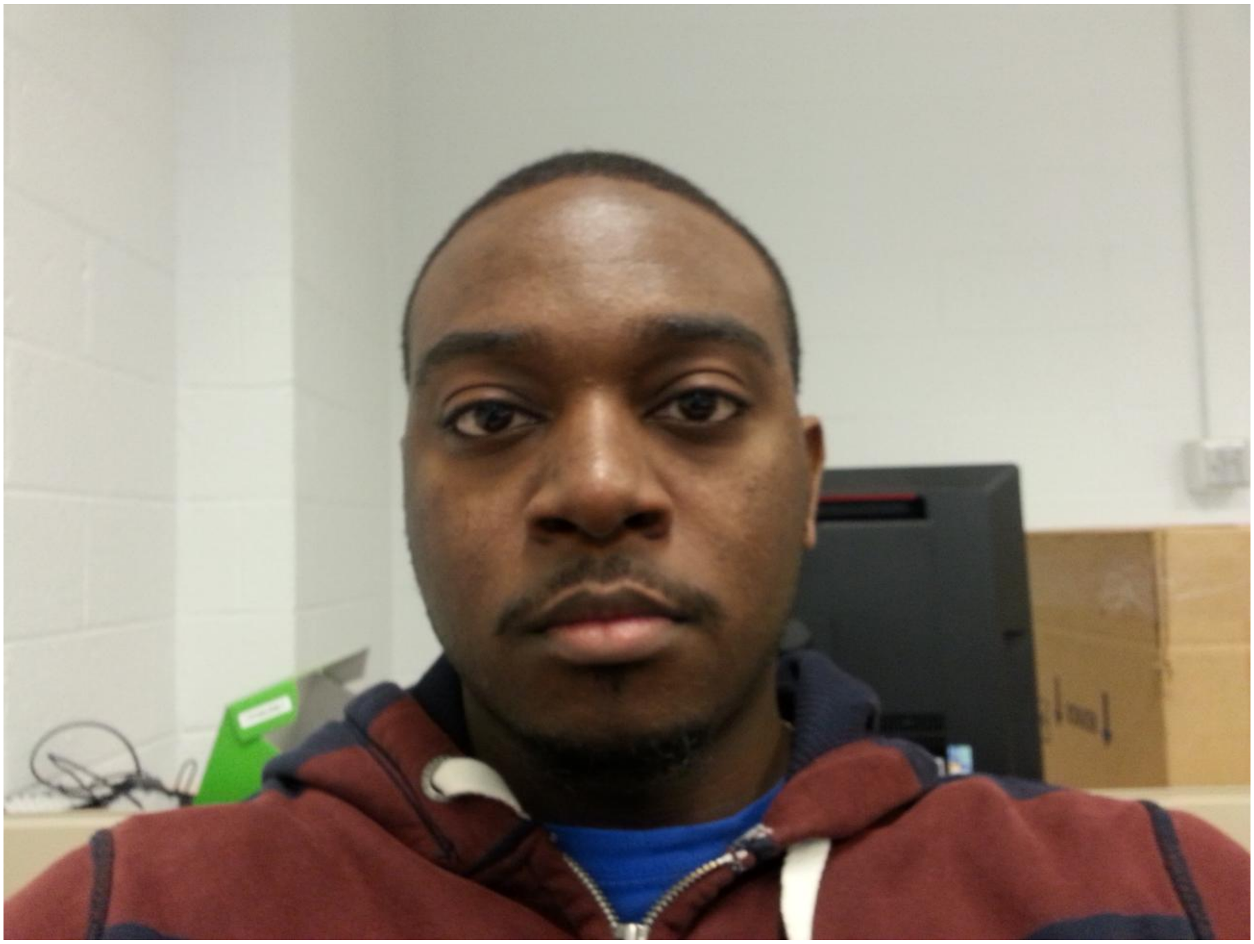


Figure 25. A student looking attentive.

If the distances between the plotted eyelids come within a certain distance the system will indicate that eyes are closed. Figure 26 displays real time behavior of a student looking down. The screenshot was taken while looking down, emulating the same effect of looking down at a mobile device.

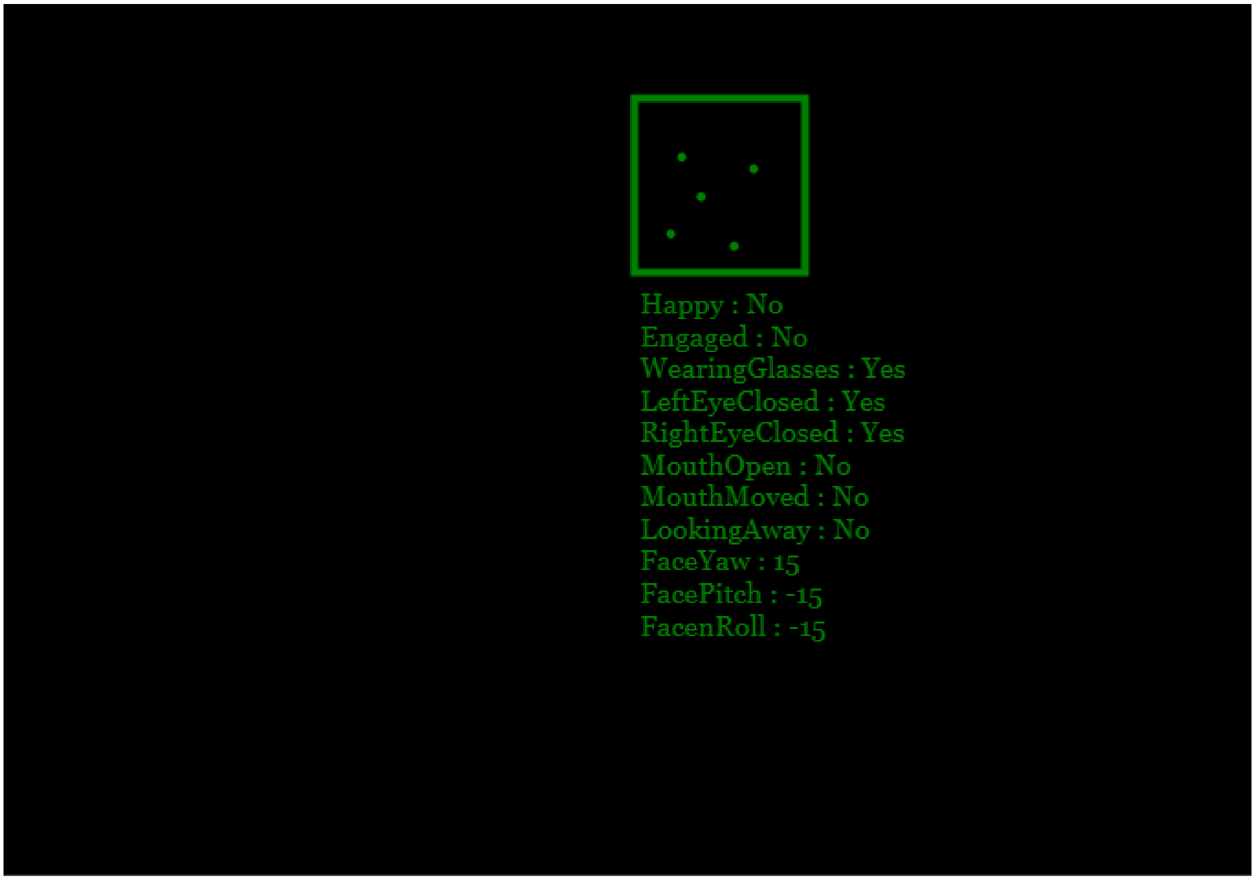


Figure 26. Real time eye tracking.

4.3 Overall System.

Testing the system with all the classifiers running not only proves that the Kinect can handle sharing sensors, but it also proves that the system can indeed monitor the attentiveness of a student. However, the system is not perfect. The system can detect the eye gaze by tracking eyelids seamlessly as long as the Kinect is in view of the eyes. The plotting of the joints for the hands can be a little jumpy depending on the distance from the Kinect.

The System is not perfect. Table 3 displays the table completed for testing of single subject (the author). Each data set was tested10 times to approximate the accuracy. The systems hand tracking is the least accurate getting 40% accuracy; the eye gaze was the most accurate at 80%. The overall system performed at 55% accuracy as shown in Table 2. The hand joint plotted by the Kinect tend to pick up the waist joints when they cross paths causing jitter and jumpy activity. The Kinect v1 had an option to use the Kinect in a seated mode which would only plot 10 major joints from the waist up. However, the Kinect v2 does not have this feature, so its tracking software attempts to plot the legs even though they aren't visible. The Kinect v2 assumes the legs are there, the same goes for the elbows. The *Custom Gesture* feature somewhat helped the issue because the gesture does not have to be 100% correct to be recognized. As long as the student's posture is close to the gesture, within 50% accuracy, it will assume the gesture matches. The percentage of accuracy can be changed in the code for the gestures.

Table 2. Accuracy with 10 attempts.

Accuracy	
Item	Percentage
Eye Track	80%
Hand Track	40%
Posture	60%
Overall	55%

4.4 Using the Distraction Tracker Application for Posture Data

To get accurate data for distractive use of mobile devices in a classroom setting, the gesture data collected from human subjects while interacting with the Distraction (or Group) Tracker Application was used to training distractive postures for the system.

Table 3 displays the results from the Distraction Tracker application test for each group. The test consisted of 10 questions and the percentage is calculated based on the max score being 100%. Group C results illustrate obvious distraction, performing the worse on the test after the slideshow, averaging 64%. Group B proved to have the best test scores having the ability to take notes as they sit through the slide show, averaging 90%. Group A scores were not too far behind group B with an average of 82% while not having the use of a mobile device at all.

Table 3. Grade results of the "Distraction Tracker application".

Participants	Group A	Group B	Group C
1	80%	80%	70%
2	100%	90%	70%
3	80%	90%	90%
4	70%	100%	40%
5	80%	90%	50%
Score Average	82%	90%	64%

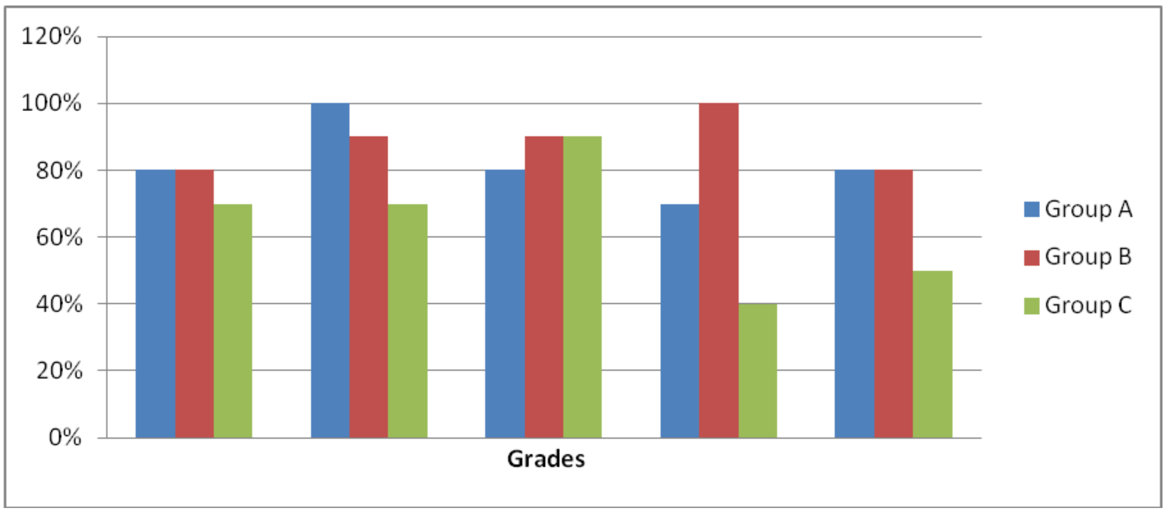


Figure 27. Grades of all 3 groups.

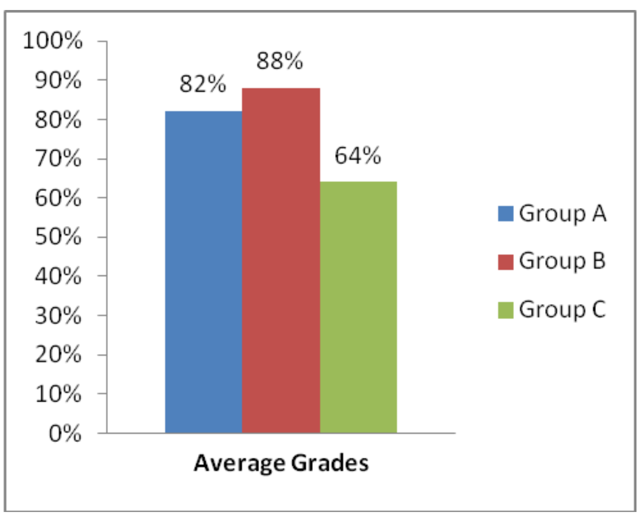


Figure 28. Average grades of the "Group Test".

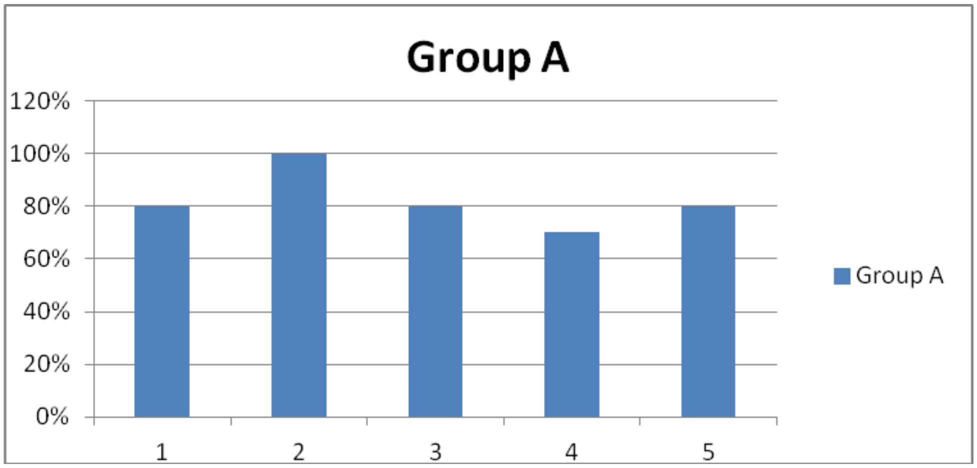


Figure 29. Test scores for group A.

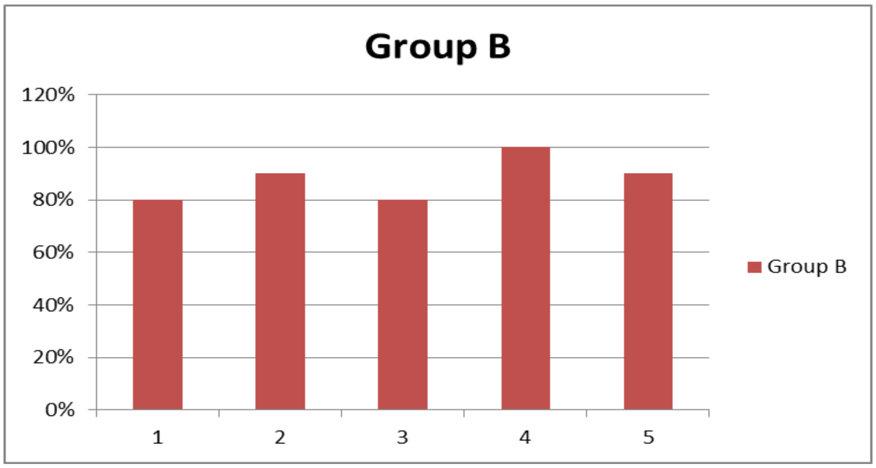


Figure 30. Test scores for group B.

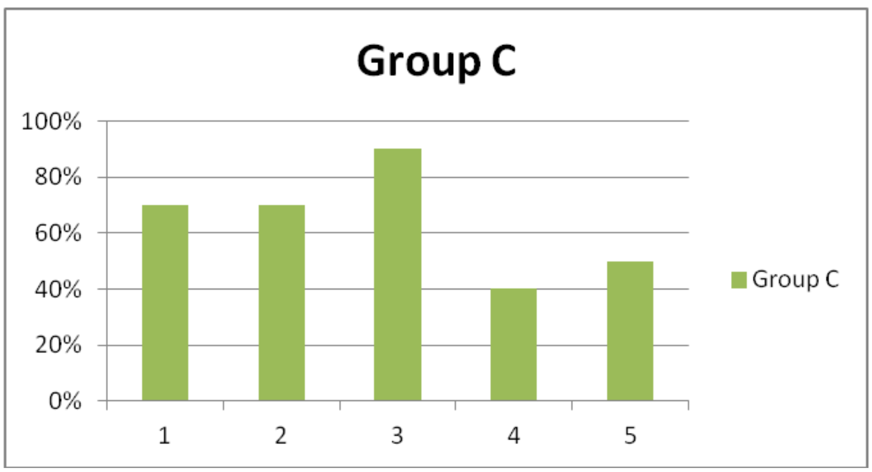


Figure 31. Test scores for group C.

The testing was done on the PC at the workstation with the Kinect. The posture data collected enhanced the accuracy and effectiveness of the system. This data was used to help create the head tilt and hand action gesture. Use of the new trained postures along with eye gaze makes this system more accurate.

4.5 Results after adding Class Data

After adding the posture data collected from the human subject testing, the accuracy of determining whether a student is distracted by mobile device increased. With the gathered data, the system has a library of gestures that helps determine distractive use of a mobile device. Without the human subject data, the system only looked for users holding a device with two hands. The collected data allows the system to also recognize a user holding or interacting with a device with one hand, as shown in Figure 32.

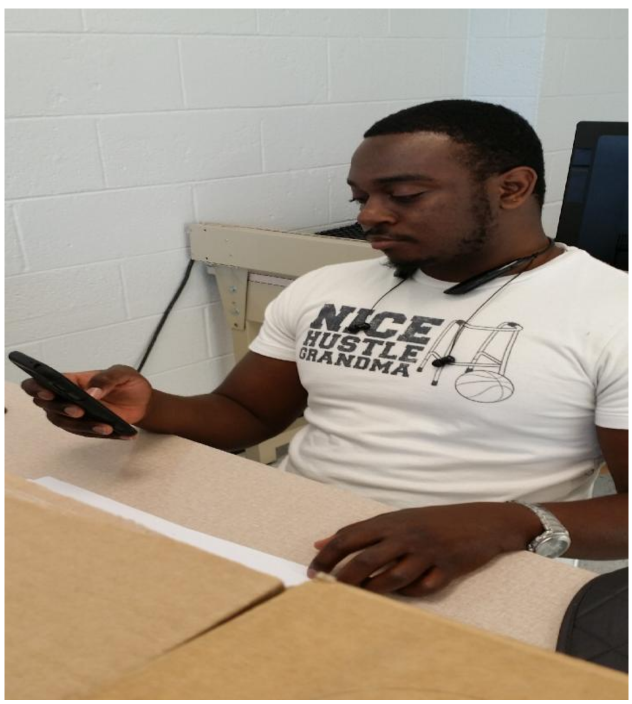


Figure 32. Using a mobile device with one hand.

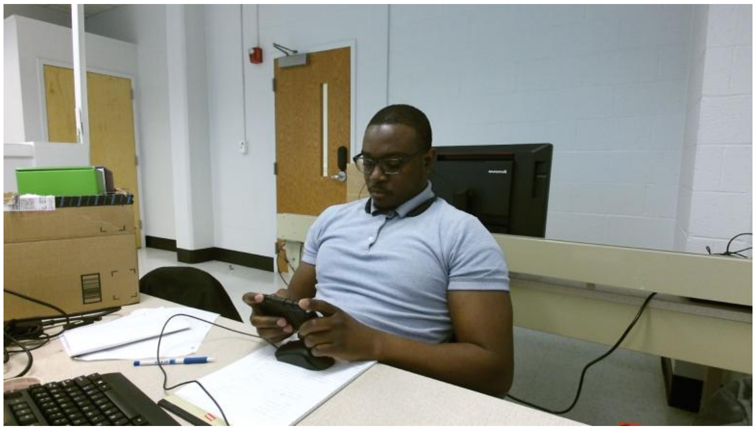


Figure 33. Using a mobile device with two hands.

The Success rate of determining the use of a mobile device however comes with its own flaws. As mention previously, version 2 of the Kinect no longer has a seated version which would only track a human skeleton from waist up while seated. As a result, Version 2 of the Kinect will estimate the position of the missing joints that are not clearly view to the Kinect's sensors. While sitting at a workstation the Kinect cannot see the skeletal joints from the waist down. While monitoring the user's posture, the skeleton may read incorrectly due to crossing the paths of missing joints, as shown in Figure 34.

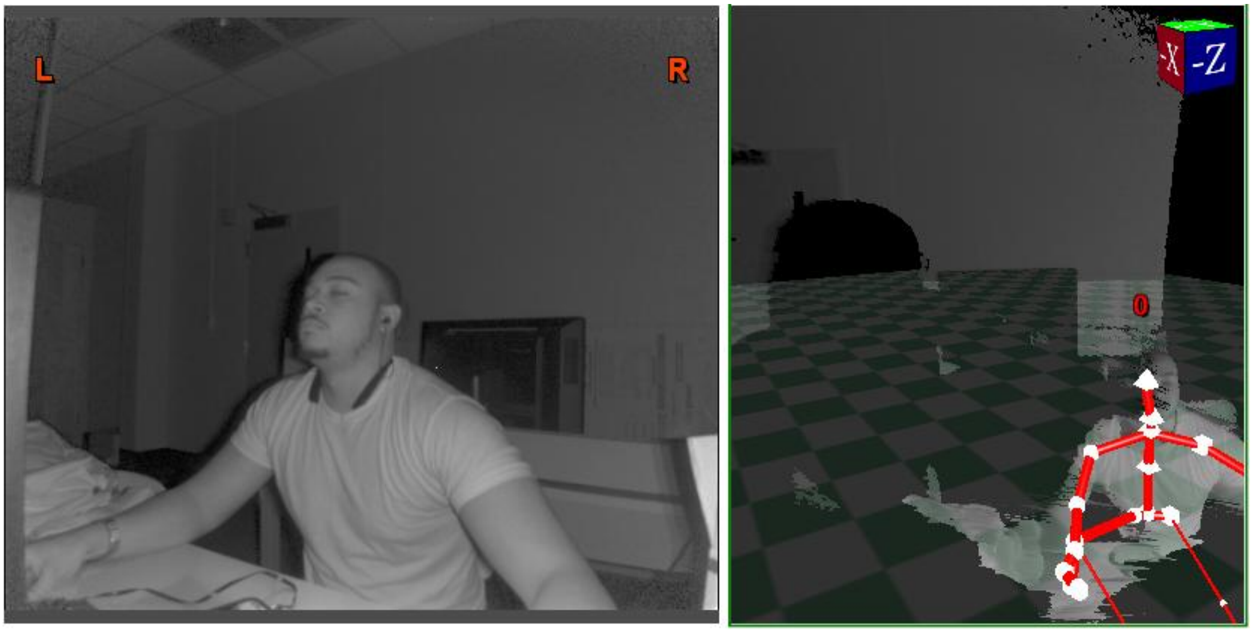


Figure 34. Displaying joint miscalculation.

After adding the collected data from the group testing, 10 trails were accessed on the system. These trails consisted of various distractive use postures. As listed in Table 4, the success rate of detecting whether a student is looking down is 100%, the rate for detecting distractive behavior posture is 80% based on data collected from five subjects, and the rate of success for determining the distractive use of mobile devices is 60%. Adding the data collected from the subjects increased the accuracy of the overall system by 25%.

The application was tested on five subjects in both locations. The new location improved eye tracking dramatically. Gesture tracking also improved to 80% opposed to 60% calculated with the Kinect positioned to the side of the monitor. Hand tracking also improved to 80% and the overall detection of distractive use increased 25% to 80%.

Table 4. System Accuracy.

Accuracy	
Posture detection	80%
looking down detection	100%
Hands	80%
Overall system success	80%

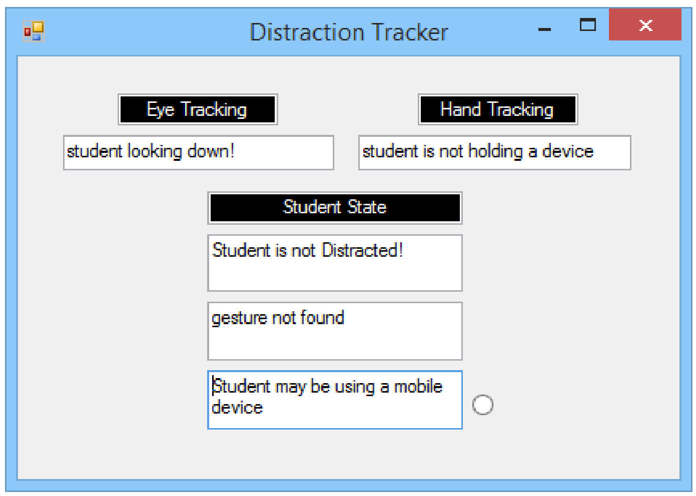


Figure 35. Distraction Tracker main window.

\

CHAPTER 5

Discussion and Future Research

The goal of this research was to develop a pervasive system to autonomously monitor the distractive use of mobile devices in a classroom setting. The system achieved this by tracking the student's hands and elbows, eye gaze and posture. The system was able to predict the eye gaze with reasonable accuracy, however the hand tracking is an issue now caused by the fact that the Kinect v2 SDK does not support a seated mode at the moment. In the future, if the SDK 2.0 supports a seated skeleton the performance of the current system would be expected to improve greatly. The collection of posture data from human subjects greatly compensated for inaccurate hand joint detection. The data also allowed additional posture and gestures to be detected while determining whether a student was using a mobile device n a distractive way.

The system should be mostly based off of custom gestures as they have a consistent success rate. However, the eye gaze estimation should be still used, as it outperformed the other tracked features. Measuring eye gaze along the y-axis using eye lids proved to be very successful. Also the ability to identify and track the student's mobile device would be very convenient. Allowing the system to scan the mobile device to get the size and shape of the device, would possibly help to determine whether the student is actually holding a mobile device. Currently, the system will assume the student is holding a device. The ability to identify and track the device itself will greatly improve the system. Currently this is possible while pairing the Kinect with OpenCV.

References

1. Park, K., Bien, Z., Lee, J., Lim, B., Kim, J., Lee, H., ... Lee, W. (2007). Robotic smart house to assist people with movement disabilities. *Autonomous Robots, 22*(2), 183-198.

2. Su, C., Rakheja, S., & Liu, H. (Eds.). (2012). Human Behavior Recognition by a Bio-monitoring Mobile Robot. *Intelligent Robotics and Applications*, 21-30.

3. MacCormick, J. (Director) (2011, January 1). How does the Kinect work? *How does the Kinect work?* Available at: http://pages.cs.wisc.edu/~ahmad/kinect.pdf.

4. Marchall, M., Harrington, A., & Steele, J. (1995). Effect of work station design on sitting posture in young children. Ergonomics, 38(9), 1932-1940.

5. Lee, H., Choi, Y., & Lee, S. (2012). Mobile posture monitoring system to prevent physical health risk of smartphone users. *UbiComp '12,* 592-593.

6. Kinect for Windows features. (2014, January 1). Retrieved January 1, 2015, from http://www.microsoft.com/en-us/kinectforwindows/meetkinect/features.aspx

7. Szymczyk, M. (2014, December 19). How Does The Kinect 2 Compare To The Kinect 1? Retrieved January 1, 2015, from http://zugara.com/how-does-the-kinect-2-compare-to-the-kinect-1

8. Pterneas, V. (2014, December 14). How to use Kinect v2 Face Basics. Retrieved January 1, 2015, from http://pterneas.com/2014/12/21/kinect-2-face-basics/

9. Mitra, S., & Acharya, T. (2007). Gesture Recognition: A Survey. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, 37*(3), 311-324. doi: 10.1109/TSMCC.2007.893280

10. Mogg, T. (2013). Apple just bought the company that invented Microsoft's Kinect Read more: http://www.digitaltrends.com/home-theater/apple-primesense/#ixzz3UkAiFlCa

11. Nergui, M., Yoshida, Y., Imamoglu, N., Gonzalez, J., & Yu, W. (2012). Human Behavior Recognition by a Bio-monitoring Mobile Robot. In C.-Y. Su, S. Rakheja & H. Liu (Eds.), *Intelligent Robotics and Applications* (Vol. 7507, pp. 21-30): Springer Berlin Heidelberg.

12. Park, K.-H., Bien, Z., Lee, J.-J., Kim, B., Lim, J.-T., Kim, J.-O., . . . Lee, W.-J. (2007). Robotic smart house to assist people with movement disabilities. *Autonomous Robots, 22*(2), 183-198. doi: 10.1007/s10514-006-9012-9

13. Wang, J.-G., Sung, E., & Venkateswarlu, R. (2005). Estimating the eye gaze from one eye. *Computer Vision and Image Understanding, 98*(1), 83-103. doi: http://dx.doi.org/10.1016/j.cviu.2004.07.008

Appendix A

Test Writer: John Campbell						
Test Case Name:	Monitor Distractive use of a Mobile Device.	Test ID#	Attentive-01			
Description:	This is a test that varifies the system can detect a distractive student.	Type:	White Box X Black Box			
Test Information						
Name of Tester	John Campbell	Date:	3/18/2015			
Software version	1.0	Time:	2:33am			
setup:	For each test confirm if the user is label distracted for using a mobile device. Check for eye tracking and check if system detects both hands holding a device.					
Test		expected outcome	Pass	Fail	N/A	Comments
1	look down	system displays "student not looking down", "student is not holding a device"	x			System worked as expected, the test passed.
2	look down for 10secs	system displays "student not looking down", "student is not holding a device", "student is distracted"	x			
3	hold phone with two hands	system displays "student is holding a mobile device", "student may be using a mobile device"				
4	hold phone with two hands and look down	system displays "student is holding a mobile device", "student may be using a mobile device"				
5	hold phone with two hands and look down for 10 secs	system displays "student is holding a mobile device", "student is using a mobile device"				
5	spread hands apart	system displays "not holding a device"				
6	lookdown while holding device	system displays "holding a device", "student distracted", "student using a device"				
7	look attentive	system displays "student not looking down", "student is not holding a device", "student is attentive"				

Figure A.1. Test document for the system.

Appendix B

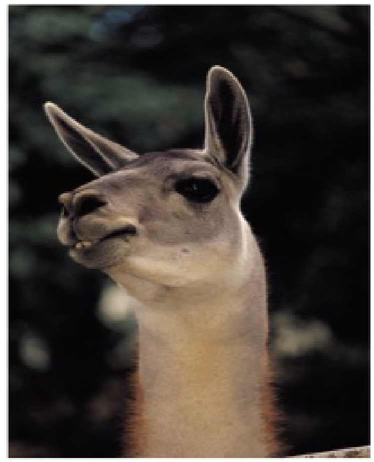
Spitting to Survive

by Liana Mahoney

Spit keeps our mouths moist and softens our food when we chew. Without spit in our mouths, we would have a hard time talking. We would find it even harder to swallow. But for some animals, spit works better after it has left the mouth. Some animals are experts at surviving because they are expert spitters.

Llamas are animals often found in petting zoos and farms. These animals seem to like their personal space. A llama that feels threatened or annoyed will spit slimy gobs at you to get you to leave it alone. Sometimes llamas even spit on each other to steal food! This trick usually works, because llama spit includes food from the llama's stomach, and it can be quite smelly. When a llama spits on another animal, the animal usually loses its appetite and walks away, leaving its food behind.

The archer fish is a very skilled spitter. This fish is like a submarine with a loaded weapon. It takes aim and spits jets of water at insects and other small creatures to knock them into the water. Then it gulps them down quickly. To create such a forceful stream of water, an archer fish closes its gills, and uses its tongue to form a tube in its mouth. Then the fish sticks its snout out of the water and aims. Aim! Launch! Lunch!



Spitting cobras are also known for their expert aim. These snakes spray poisonous venom from their fangs to protect themselves. Scientists believe that these snakes actually aim for the eyes! When the cobra's venom gets into the eyes of an animal, the venom causes terrible pain, and even blindness. This gives the snake plenty of time to get away.

Spitting is considered to be rude behavior in people. But for some animals, spitting can be a smart way to get lunch –or a clever way to avoid becoming lunch!

Figure B.1. Page 1 of the comprehensive material.

Name: ____________________

Spitting to Survive

by Liana Mahoney

1. List the three ways spit helps humans.

____

____

2. Which animal creates a forceful stream of water to capture insects?
   - **a.** humans
   - **b.** archer fish
   - **c.** spitting cobras
   - **d.** llamas

3. Name two reasons a llama might choose to spit.

____

____

4. How does a spitting cobra use its spit to protect itself?

____

____

5. What is the author's purpose for writing this passage?
   - **a.** to tell funny stories about animals
   - **b.** to teach the reader how animals survive
   - **c.** to express opinions about animals
   - **d.** to show how animals are different

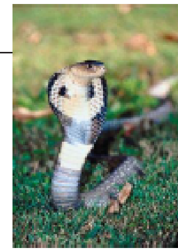
Super Teacher Worksheets - www.superteacherworksheets.com

Figure B.2. Page 2 of the comprehensive material.

Name: ____________________

Spitting to Survive

by Liana Mahoney



1. List the three ways spit helps humans.

   Spit helps humans chew by softening our food. It also helps us talk and swallow.

2. Which animal creates a forceful stream of water to capture insects? B

   a. humans b. archer fish
   c. spitting cobras d. llamas

3. Name two reasons a llama might choose to spit.

   A llama spits when it feels threatened or when it is trying to steal food.

4. How does a spitting cobra use its spit to protect itself?

   A spitting cobra spits poisonous venom in the eyes of its predators so it has time to escape.

5. What is the author's purpose for writing this passage? B

   a. to tell funny stories about animals b. to teach the reader how animals survive
   c. to express opinions about animals d. to show how animals are different

Super Teacher Worksheets - www.superteacherworksheets.com

Figure B.2. Page 3 of the comprehensive material.