Abstract

The Microsoft Kinect seems inherently useful for the field of sports biomechanics- as a cheap alternative to expensive equipment, a limit on sensors and cameras and the potential for mass production. In addition there is a limitation for the accessibility of information gained from biomechanical systems. In my research I hoped to improve on the limitations of biomechanical systems. I developed an application that utilizes the Microsoft Kinect and Unity 3D to give feedback to users to help improve on their martial arts techniques.

Introduction

The best professional athletes are able to process countless information each second in order to maximize their performances. There have been many studies that have evaluated the extent to which our brains are able to detect information embedded in biological motion. The field of sports analytics utilizes video analysis and motion capture systems in order to determine the most optimal techniques for sports. The main issue with these systems is that they are both expensive and require an expert to help turn the data into tangible information. For my research I will be utilizing the Microsoft Kinect and Unity 3D to create a tool for learning Tae Kwon Do techniques that addresses the issues of expense and expertise. I propose doing so by having an immersive training program with gesture recognition working in conjunction with real-time joint analysis.

Related Works

I had originally hoped to utilize machine learning and gesture recognition to discover the parameters that which we use when we categorize ourselves. The main idea being that social categorization has been so enforced and internalized that we function everyday making assumptions according to these labels. The idea of exploring the parameterization stemmed from point light display research which showcased the ways in which our brains have been trained to interpret information even when limited information. The experiment consisted of putting lights on key joints and body parts and having people guess the actions and gender of the participants. The results were that people were able to determine the participant's gender at high rates, suggesting that theories of gender embodiment and social conditioning could be observed at a biological level. The work ended up leaving more questions, the main one being how it is possible for our brains to evaluate this and what is it that is important in evaluating biological motion.

This same theory of identity resulted in a group of researchers creating a gesture library in the hopes of analysing identity embedded in the motion. They used a motion capture system and saved the raw data such that it could be implemented in 3ds Max and Matlab. The group got a group of people of varying identities to do the same set of actions. Each person was tasked with performing an emotion by interpreting a script. The hope was that each motion capture file would have an emotion and gender tied to the action. They had their work available for download for the hopes that it could one day be utilized for examining the data. I downloaded the database in hopes to be able to utilize the data, but the current limitations of machine learning technology made me go a different direction.

I required a more concrete problem solution framework for which to analysis motion, and this lead me to the field of sports biomechanics. In the case of sports the brain analyzes countless information each second in order to maximize performance in sports and the field of biomechanics is interested in discovering and analyzing the techniques that that the brain looks at in order to maximizes performance. One such research was done on evaluating the performance of athletes through immersing them in virtual environments. The experiment was broken down into two sports, rugby and handball. For the rugby portion participants wore a headset and were given two buttons correlating to right and left. The participants were tasked with determining what direction an oncoming player would be going towards and they were evaluated by their correctness and time taken. The second experiment utilized curved display screens, motion capture system and a physical goal to evaluate what aspects are important in determining which direction to block a shot on goal. Another example of sports analytics is a system made by the American Sports Medicine Institute for evaluating pitching mechanics. It requires a motion capture system for recording and a team of researchers for evaluating the recordings. After a thorough analysis the pitcher is given a 15 page analysis. The main problems with these systems is that they do not provide easy feedback for participants to get better or feedback at all.

The Microsoft Kinect has always been seen as a possible alternative to using a complex motion capture systems. There have been multiple experiments that evaluated the validity of this assumption. They have found that the Kinect works best for recognizing discrete movements and ones that are not extremely complicated. The main issue with utilizing the Microsoft Kinect only for gesture recognition is that at most it can be used to learn binary yes/no completed/unrecognized gesture data. Additional information is needed to be calculated in order to provide more feedback.

Methodology

The main goal of my research was to create a program that utilized sports biomechanics and intuitively taught users how to perform techniques better. This project utilized a combination of the Microsoft Kinect V2 Sensor, Kinect Studio, Visual Gesture Builder and Unity 3D. Kinect Studio and Visual Gesture Builder were utilized for recording and training gestures that users would learn. The Microsoft Kinect and Unity 3D were utilized in conjunction to both recognize these gestures and provide an immersive learning environment with helpful feedback for improvement.

The first step required training the gestures to be utilized in the application. In an effort to provide more feedback I expanded the range of positive examples of the gesture, in doing so gestures that are slightly off would still be recognized as an attempt. For example, I had examples of punches with perfect form and punches that ended outside, were too slow or too high just to name a few. These were done for the following gestures: CrossPunch, Jap, BackFist, DownBlock, OutsideBlock and HighBlock. Each gesture was trained as a negative example for each other. The recording was done in Kinect Studio and the training was done in Visual Gesture Builder. Once the database was built it could then be imported into Unity to be referenced through code.

Once the database was copied into the Unity Resource folder each of the Gestures could be referenced by name. Utilizing the Kinect Asset in Unity 3D each of the skeletal Joint Points that the Kinect Sensor provides could be have a GameObject attached to it. These GameObjects were utilized for processing information in real-time such as keeping track of sequences of angle changes and velocity by using the x,y,z coordinates of the GameObject. An example of an angle used is that of the right hand to the shoulder in order to determine how height and width of a punch. Once a gesture was recognized it would make a call to the tracked angle function in order to get the info on what type of punch was made, a similar call is made to determine the velocity of the punch.

The next step required determining the best methods for creating an immersive experience. The joint model used was made to face away from the camera such that when a user stepped forward to the right the model moved in the same direction rather than towards the user.

In this way the user can be more immersed into the program by seeing a visual representation of themselves that they can associate their movements to. Each gesture was also given a recorded tutorial, when a user asked for help the skeletal model will demonstrate the way that the gesture should be done. Then there was the development of destroyable objects in order to replicate real reaction time situations similar to training with a sparring partner. An instructor side of the application was developed, but has been limited to creating a lesson plan composed of a sequence of techniques.

Two versions of the program were created in order to showcase the improvements that visual feedback can provide. The first of which only prompted users with a sequence of text prompts of the gestures they were required to do. The second of which meet users with sequences of destroyable objects that required a specific gesture to be destroyed. In the first version users were merely reacting to reading a prompt and not practicing the correlation of a technique and its potential application. Both of these feedback such as “too high” or “too left” along with an angle and speed.

Conclusion

Creating a system that utilizes sports biomechanics information in order to improve sports performance is extremely viable, especially when it comes to techniques that are less complicated. A more immersive environment does well to replicate actual training situations. The continued implementations of visual cues increases both the ease of accepting feedback and the sense of immersion with the system. Future development could be in allowing for instructors to provide more than just a lesson plan but also have access to adding new techniques and defining the parameters of angles that are important for the techniques. The ideal format for this program is a web server where users can log in online and only need to have access to their own Kinect Sensors in order to use the program. Additional visual keys such as graphical representations of important angles and visual tracking of joints would provide more intuitive feedback for users.

As it stands the creation of all-encompassing sports biomechanics training systems are getting closer to being realized the more technology advances. The inherent limitations of the Kinect may not lend it be used for all complex sports techniques but it is exciting to imagine what technology is along the way.