# ZerotoHero

*Subtitle*

By

LIAM O'SHEA

Department of Computer Science
UNIVERSITY OF BRISTOL

A dissertation submitted to the University of Bristol in accordance with the requirements of the degree of DOCTOR OF PHILOSOPHY in the Faculty of Engineering.

MAY 2014

Word count: UNKNOWN-REPLACE

## Abstract

Here goes the abstract

# Dedication and acknowledgements

Here goes the dedication.

# Author's declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: .................................................... DATE: ..........................................

# Table of Contents

# List of Tables

**Table** **Page**

# List of Figures

**Figure** **Page**

# Chapter 1

## Introduction

Computer aided coaching has revolutionised training approaches and been a key driver in improving sports performance at the highest level. Professional athletes can now use advanced augmented coaching to accurately and reliably measure a range of metrics that are then used as indicators of performance.[9] This gives the coaching teams unique insight which they use to tailor training programs to give the athlete a goal to focus on and improve.

Huge injection of research funding in preparation for the London Olympics such as The Elite Sport Performance Research in Training (ESPRIT) Programme, meant that UK Sport was allocated £2,000,000 while the Engineering and Physical Sciences Research Council were allocated £6,100,000.[14] This money was spent on enhancing elite performance training with technology [4] [13] which allowed British Olympians to take advantage of these training methods and return with a medal record making it the most successful team since 1908.[15] These hi-tech training methods were used by Team GB Boxing [2] and have started to become incorporated in the USA Boxing team.[8]

However despite the recent success, sports technology is still in its infancy, with the majority of devices being expensive, specialist and proprietary[6]. There is an array of consumer devices such as the Fitbit, Nike+ Fuelband, Garmin Vivofit and Jawbone that function as general activity trackers but there is currently nothing that offers specialist training or coaching advice. The purpose of ZeroToHero was to use an already widespread, low cost consumer device to bring specialist boxing coaching to everyone. The low-cost is especially crucial since most boxing clubs struggle with funding, rely on volunteers and traditionally have members that are the young and least wealthy in society.

ZeroToHero explores the ability of the Kinect to act as a electronic boxing trainer, offering advice on punch & stance performance.

## 1.1 Goals

Maybe talk about the pipeline here? Segmentation?

### 1.1.1 Limitations & Scope

Begins a subsection.

### 1.1.2 Other applications?

# Chapter 2

## Background & Research

This chapter describes and explains boxing concepts which are important to understanding the goal of the project. It gives an overview of the types of punches a boxer must execute along with common errors associated with those. It also reviews relevant literature in areas covered by this thesis and explores the current cutting edge possibilities.

## 2.1 Boxing Background

Boxing requires an incredible amount of co-ordination and timing as well as the ability to rapidly execute punches in a controlled and precise manner. Unlike professional boxing which is 12 x 3 minute rounds an amateur bout is 3 x 2 minute rounds which changes the dynamics of the contest. Amateur boxing relies on a points scoring system since there is often insufficient time for knockouts. therefore amateur boxers rely on the mastery of technique and proper form. Dropping you hand for a split second can open you up to an experienced boxer and could spell disaster.

As someone who has boxed for over 5 years and captain of the University of Bristol's Amateur Boxing Club (UOBABC) I know how difficult it is to develop good technique and how much time and experience is required from a coach to develop a new boxer. This one-on-one time is incredibly valuable but expensive and for the large majority very hard to get. ZerotoHero aim is to be able to identify different types of punches and to offer feedback on the quality of the movement. This will bring some much needed expert advice to a beginner who can practice in the comfort of their own home.

I am using my own experience and that of local professionals, coaches and local legend Denis Stinchcombe MBE the centre director of Riverside Youth Project and Registrar for the Western Counties for the Amateur Boxing Association.

### 2.1.1 Motivation

This research is borne out of a desire to improve access and cost to boxing coaching which are problems I have encountered first hand through the University Boxing Club. In the wider environment It could be used in developing countries where physical access to coaches with the required expertise may be difficult as well as local clubs in the UK.

Every year UOBABC takes in new members that are total beginners. We spend an enormous amount of time and effort helping them learn the basics and encourage people to practice at home. The problem is that from a boxers perspective it‚Äôs incredibly hard to spot your own faults. If it was possible to practice at home with the benefits of coaching it would bring massive improvements to their abilities. The current most effective way to train as an individual is to stand in front of a mirror and observe yourself while shadow boxing.

It could also be used as a way to introduce younger children to the sport since the Kinect is incredibly popular in that demographic cite and so it‚Äôs great from an inclusion perspective.

## 2.2 Boxing Technique

For the scope of this project I am going to focus on the most common orthodox stance. There are tens of slight variations on each punch but I am going to focus on the core and important principles from which these can be built.

### 2.2.1 Stance

The most fundamental building block of boxing is the stance, that is how you hold and position your body as well as the placement & orientation of your feet. A good stance is crucial since it allows the boxer to be well balanced and light on their feet, allowing fast movement in any direction as well as the ability to quickly duck, weave, slip and bob and lay back to avoid punches. It is also crucial for offence since the power from punches come from the transfer of weight from one leg to another which requires a very specific twisting hip movement. Often beginners forget this crucial step and so I'm hoping to use this unique trait to help me judge quality later on. A successful stance should have the following characteristics:

- Left foot forward, right foot back with the feet slightly wider than shoulder width with a 45 degree angle twist. - Right heel of the ground at all times with weight distribution mostly on your back leg - Slightly bent knees - Chin tucked down - Right hand on the right hand side of your chin, left hand should be a few inches in front of the left side of the face - Elbows tucked in

### 2.2.2 Punches

#### Jab

The elbow should stay tucked in while the left fist extends with palms facing inwards before twisting your wrist at the last moment. The natural thing to do is extend the punch with palms facing down, unfortunately this immediately makes the elbow stick out which allows the opponent to easily see you are about to throw a punch (telegraphing) while opening up your body for a counter attack. The punch should also finish so your arm is fully extended which helps to extend your reach and protect your chin before speedily returning it to the guard position. Characteristic to target: Elbow movement

#### Cross

The cross is designed as your heavy straight punch and as such is slower but more powerful. To get a śnappyánd powerful punch it is important to transfer your weight rapidly from your back leg to your front leg, twisting your hips. Characteristic to target: Rapid and specific hip movement

#### Hooks

Your elbow should come out so it is shoulder height and your fist and shoulder should be at 90 degrees to each other. Again a rapid transfer of weight between hips is needed.

Characteristic to target: Dropping the hand before punching Characteristic to target: Hip movement

#### Uppercuts

This required the fighter to crouch down into the squat position and throw a punch vertically upwards, with the aim of striking the opponent's chin. Characteristics: Sufficient Crouch, Characteristic: directly vertical punch, keeping guard close at all times.

## 2.3 Kinect

This section provides some background information about Microsoft Kinect that is important for understanding the features and limitations of Kinect Analysis. According to Microsoft, the Kinect has worldwide sales of approximately 28 million units and contains an RGB camera, an infrared (IR) emitter and an IR depth sensor as well as a multi-array microphone. The interaction space of the Kinect is limited by the field of view of the Kinect cameras. The Kinect has a 43°vertical by 57°horizontal field of view. The Kinect sensor can be tilted using a built-in tilt motor. Tilting the Kinect increases the interaction space by $+27$ and $-27$ degrees. The Kinect sensor provides sensor data in form of data streams. It can capture audio, color and depth data. In addition, it can process the depth data to generate skeleton data. Therefore, the Kinect offers four different data streams that can be accessed: audio stream, color stream, depth stream and skeleton stream. The streams can deliver at most 30 frames per second (FPS) using a resolution of $640 \times 480$ which drops to 12 FPS with a resolution of $1280 \times 960$.

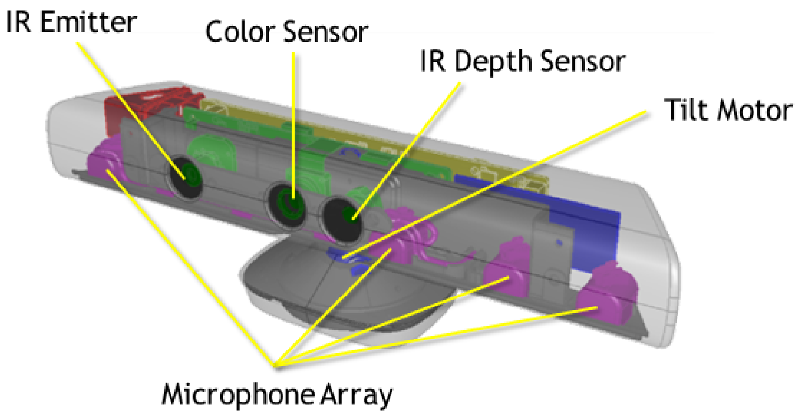


FIGURE 2.1. Kinect Device.

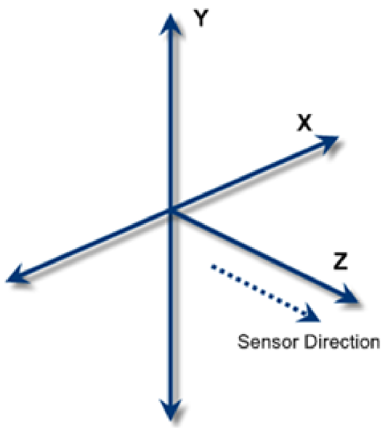


FIGURE 2.2. 3D representation

**Depth & Infrared Stream**

The depth sensor generates invisible IR light to determine an object's depth from the sensor. The primary use for the IR stream is to improve external camera calibration using a test pattern observed from both the RGB and IR camera to more accurately determine how to map coordinates from one camera space to another. [11] The NUI API uses the depth stream to detect the presence of humans in front of the sensor.[3] Skeletal tracking is optimized to recognize users facing the Kinect, so sideways poses provide some challenges because parts of the body are not visible to the sensor.

| Kinect | Array Specifications |
|---|---|
| Viewing angle | 43° vertical by 57° horizontal field of view |
| Vertical tilt range | ±27° |
| Frame rate (depth and color stream) | 30 frames per second (FPS) |
| Audio format | 16-kHz, 24-bit mono pulse code modulation (PCM) |
| Audio input characteristics | A four-microphone array with 24-bit analog-to-digital converter (ADC) and Kinect-resident signal processing including acoustic echo cancellation and noise suppression |
| Accelerometer characteristics | A 2G/4G/8G accelerometer configured for the 2G range, with a 1° accuracy upper limit. |

FIGURE 2.3. Kinect Specifications

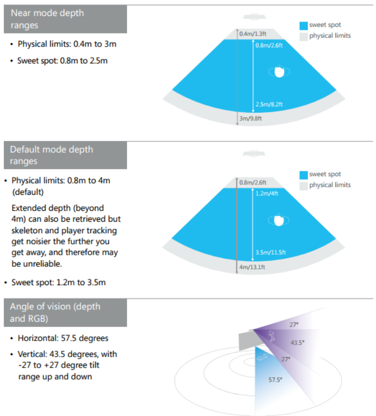


FIGURE 2.4. Kinect Specifications

### 2.3.1 Skeleton & Joint Tracking

The Kinects default tracking mode can track up to 20 joints per skeleton providing the subject is standing relatively face on and are fully visible to the sensor. Although the Kinect is capable of different modes (e.g. sitting) this is not relevant in the context of my project. Each skeleton frame contains the position of each joint as well as information about the tracking quality. Joints can have one of three different tracking states, Not Tracked (0), Inferred (1) and Tracked (2), this flag is useful as an indicator of the quality of the measurements you are receiving for a particular joint. When possible tracked joints are used to help calculate the position of those joints that cannot be directly tracked hence the inferred joints. The Kinects default tracking mode is designed to track people who are standing and fully visible to the sensor. The default range requires skeletons to be at least 80 centimetres away from the device to be tracked properly.

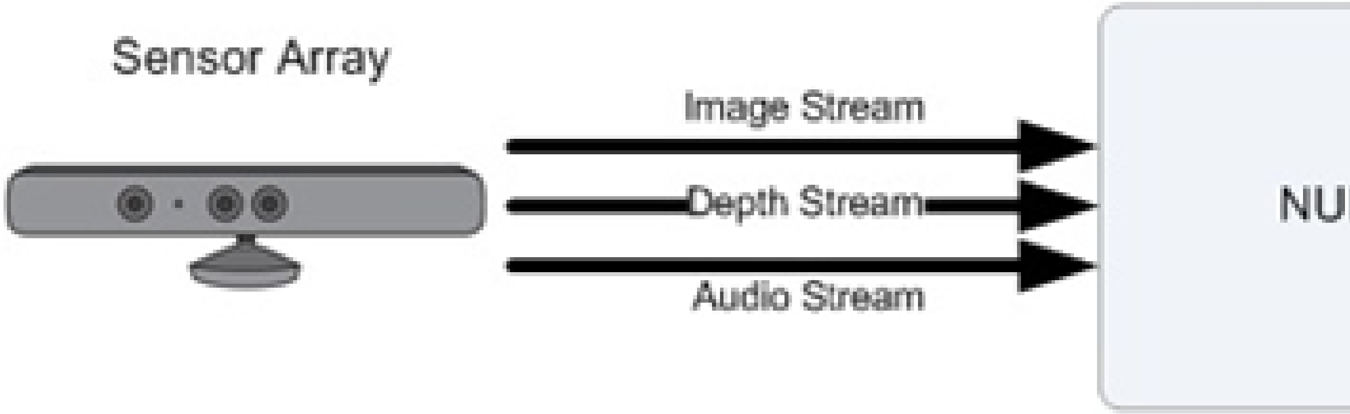


FIGURE 2.5. Kinect Specifications

## 2.4 Literature Review

The Kinect is currently a rich area of research in its own right and it straddles the fields of Image Processing and Computer Vision as well as Human computer Interaction which are exciting and popular areas of research. I wanted to see what was capable with the Kinect at the cutting edge of research so I made use of the Conference on Computer Vision and Pattern Recognition (CVPR), International Conference on Computer Vision (ICCV) & European Conference on Computer Vision (ECCV) to find relevant and useful research.
In 2013 a controlled study of 16 people who were blind or who had low vision was ran to test the usefulness of Eyes-Free Yoga: An Exergame Using Depth Cameras for Blind & Low Vision Exercise. The purpose of this was to teach the participants yoga poses using audio feedback and a positive response from participants.[12] The study took the Kinect joint positions and calculated joint angles while using heuristics for each pose. However the study only measured success using 6 static movements, which unlike boxing were considerably different from each other and so did not help me evaluate the feasibility of my goal.

The Kinect has also been used to evaluate a dancers performance with comparison to a professional dancer in real-time[1]. A score was achieved by adding three different metrics, one from the correlation coefficient of quaternions, another using joint velocities and finally a "3D flow Error," calculated from frame vectors.
The experimental results were encouraging with most of the scores consistent with real-life rankings with the exception of a few poor results due to bad skeleton calibration and tracking. This work draws strong parallels in what I am trying to achieve and demonstrated that the Kinect was a viable option for my work.

Other work such as disc throwing performance[16] also showed some promise with limited success with the lower ability groups improving their movement. This approach simply used joint

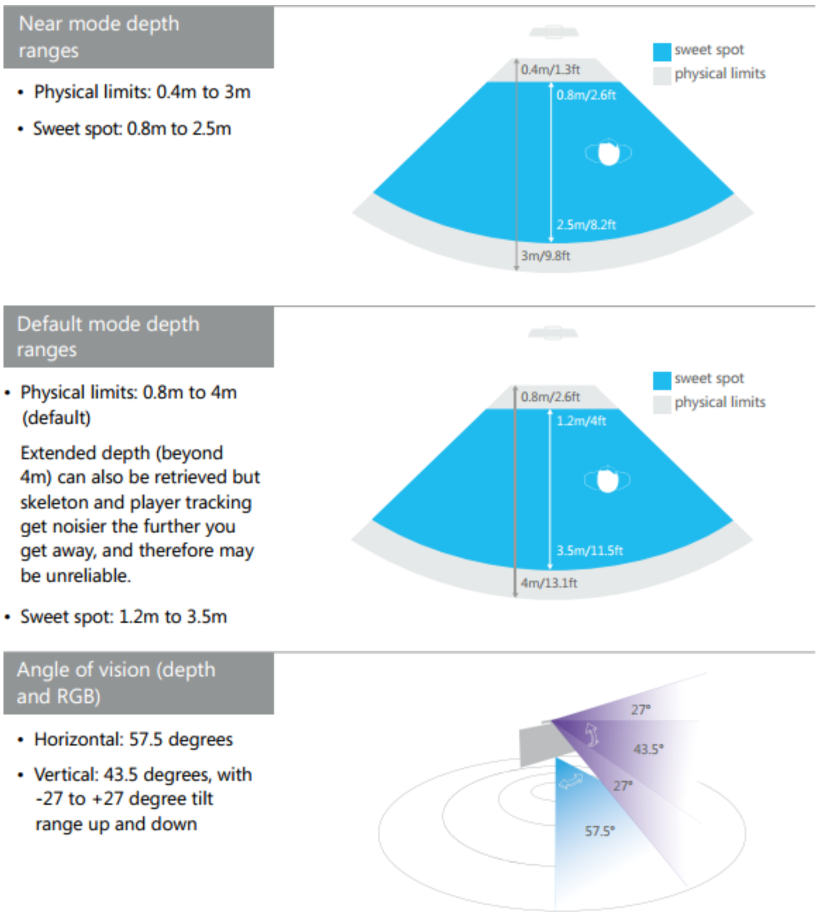


FIGURE 2.6. Kinect Depth Specifications.

angles to measure 5 different phrases of the throw and compare that to joint angle rules.

From my research I concluded that the Kinect was a viable solution to my problem and worth pursing. Crucially however there has been no research into the area of boxing which brings it's own unique challenges. Boxers are trained to be fast, well guarded and to give very little away in their movements, especially punches. Therefore many of the punches and poses are very similar since the goal is to be naturally evasive. This will make segmenting punches and giving useful quality metrics challenging in it's own way unlike say a discus throw.

Do I need to do a better evaluation of techniques used here?

cite aug7 as expensive proprietary solution It has also been used to recognise Karate movements and give a quality measurement on the movements performed to

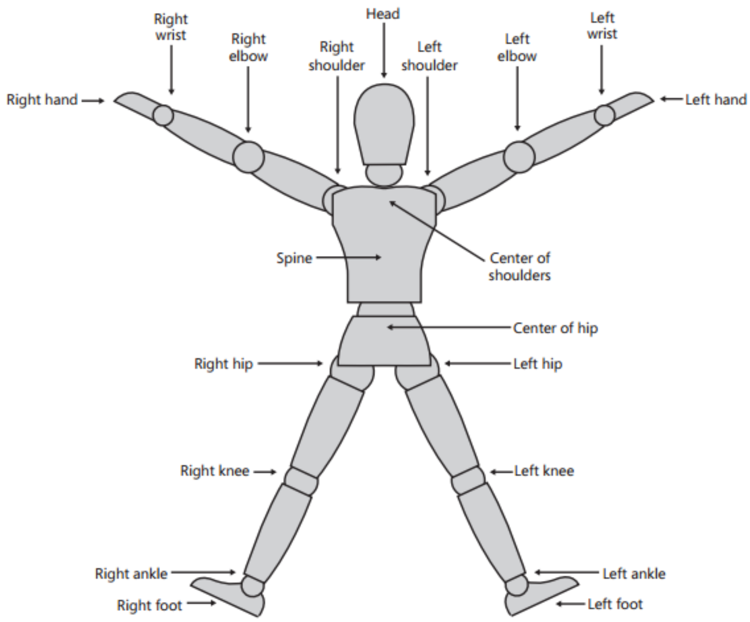


FIGURE 2.7. Kinect Joint Tracking Skeleton.

## 2.5 Existing Products

### UFC Personal Trainer

The closest commercial product is a game called ÚFC Personal Trainerẃhich is a very broad commercial product aimed at introducing people to UFC. After evaluating this product it became clear to me that it‚Äôs main focus is exercise regimes rather than technical fighting. Therefore it does not offer the preciseness or technical fighting focus that I require. In this game even technically wrong punches will still register.

### Kinect sports boxing game

This has very poor punch discrimination and does not require any sort of real boxing ability. The goal here is usually to punch towards the Kinect controller as quickly as possible. It fails to recognise properly thrown hooks and uppercuts and translate those into the game.

### Fighers Uncaged

This game has generally poor reviews from reviewers, with most complaints involving the reliability of the Kinect to accurately measure fighting moves. [7] “When the fights actually start, pulling off moves becomes a series of desperate flails, trying to get the game to recognize your actions.”[5] and “The game fails to register most of the movements”, “The idea was great for

Kinect, but something went horribly wrong along the way. Kinect is supposed to register your every kick and punch, but it only catches about the half of them.", "the Kinect control is lazily implemented."[10]

Do I need to extend the background for boxing?

# Chapter 3

# Specification & Design

Begins a chapter. Example: When the beloved cellist (Christopher Walken - outstanding) of a world-renowned string quartet receives a life-changing diagnosis, the group's future suddenly hangs in the balance: suppressed emotions, competing egos and uncontrollable passions threaten to derail years of friendship and collaboration. Featuring a brilliant ensemble cast (including Philip Seymour Hoffman, Catherine Keener and Mark Ivanir as the three other quartet members), it is a fascinating look into the world of working musicians, and an elegant homage to chamber music and the cultural world of New York. The music, of course, is ravishing (the score is the work of regular David Lynch collaborator Angelo Badalamenti): A Late Quartet hits all the right notes.

## 3.1 Scope of Project

Begins a section.

## 3.2 Punch Segementation Algorithm

## 3.3 Punch Quality Algorithm

### 3.3.1 Dimensionality Reduction

Begins a subsection. bnm m

Kinect pipeline?

# Chapter 4

# Implementation

Begins a chapter. Example: When the beloved cellist (Christopher Walken - outstanding) of a world-renowned string quartet receives a life-changing diagnosis, the group's future suddenly hangs in the balance: suppressed emotions, competing egos and uncontrollable passions threaten to derail years of friendship and collaboration. Featuring a brilliant ensemble cast (including Philip Seymour Hoffman, Catherine Keener and Mark Ivanir as the three other quartet members), it is a fascinating look into the world of working musicians, and an elegant homage to chamber music and the cultural world of New York. The music, of course, is ravishing (the score is the work of regular David Lynch collaborator Angelo Badalamenti): A Late Quartet hits all the right notes.

## 4.1 Comparison Methods

Hidden Markov models?

### 4.1.1 PCA

Principal Component Analysis is a statistical procedure that transformed a set of observations of potentially correlated variables into a set of linearly uncorrelated variables called principal components. The number of principal components should always be less than or equal to the number of original values with the first principal component having the largest possible variance. In my case I will be looking to reduce my 20 points per frame for each joint into a low dimensionality set that will help me to uniquely identify punches.

Svm?

Chapter 3: Specification & Design

Scope Algorithms
Chapter 4: Implementation
Chapter 5: Data capture??
Need to make and collect data consent forms to run a study?
Need to gather more data?

**Data Format**

I record data from the Kinect in a space separated text file with each line corresponding to one timeframe. The structure of a line is: tracking_flag x_0 y_0 z_0 tracking_flag x_1 y_1 z_1 ... tracking_flag x_19 y_19 z_19, where x_i,y_i,z_i are the x,y,z coordinates representing the position of the ith joint. Each new line is represented by a very large value that could not represent a Kinect measurement. (e.g. 2000000) The tracking_flag is an integer which describes the status of the joint: Joint not tracked = 0, Joint position inferred = 1, Join position tracked = 2. If the joint is not tracked the position is set to (-10000, -10000, -10000) and it should not be used. The position of the camera is (0,0,0).

The joints are: i=0: NUI_SKELETON_POSITION_HIP_CENTER i=1: NUI_SKELETON_POSITION_SPINE i=2: NUI_SKELETON_POSITION_SHOULDER_CENTER i=3: NUI_SKELETON_POSITION_HEAD i=4: NUI_SKELETON_POSITION_SHOULDER_LEFT i=5: NUI_SKELETON_POSITION_ELBOW_LEFT i=6: NUI_SKELETON_POSITION_WRIST_LEFT i=7: NUI_SKELETON_POSITION_HAND_LEFT i=8: NUI_SKELETON_POSITION_SHOULDER_RIGHT i=9: NUI_SKELETON_POSITION_ELBOW_RIGHT i=10: NUI_SKELETON_POSITION_WRIST_RIGHT i=11: NUI_SKELETON_POSITION_HAND_RIGHT i=12: NUI_SKELETON_POSITION_HIP_LEFT i=13: NUI_SKELETON_POSITION_KNEE_LEFT i=14: NUI_SKELETON_POSITION_ANKLE_LEFT i=15: NUI_SKELETON_POSITION_FOOT_LEFT i=16: NUI_SKELETON_POSITION_HIP_RIGHT i=17: NUI_SKELETON_POSITION_KNEE_RIGHT i=18: NUI_SKELETON_POSITION_ANKLE_RIGHT i=19: NUI_SKELETON_POSITION_FOOT_RIGHT

# Chapter 5

# Results, Conclusions, and future work

Begins a chapter. Example: When the beloved cellist (Christopher Walken - outstanding) of a world-renowned string quartet receives a life-changing diagnosis, the group's future suddenly hangs in the balance: suppressed emotions, competing egos and uncontrollable passions threaten to derail years of friendship and collaboration. Featuring a brilliant ensemble cast (including Philip Seymour Hoffman, Catherine Keener and Mark Ivanir as the three other quartet members), it is a fascinating look into the world of working musicians, and an elegant homage to chamber music and the cultural world of New York. The music, of course, is ravishing (the score is the work of regular David Lynch collaborator Angelo Badalamenti): A Late Quartet hits all the right notes.

## 5.1 Section

Begins a section.

### 5.1.1 Subsection

Begins a subsection.

APPENDIX A

# APPENDIX A

Begins an appendix

# Bibliography


[1] D. S. Alexiadis, P. Kelly, P. Daras, N. E. O'Connor, T. Boubekeur, and M. B. Moussa, *Evaluating a dancer's performance using kinect-based skeleton tracking*, Proc. 19th ACM Int. Conf. Multimed. - MM '11, (2011), p. 659.

[2] D. S. BBC, *Usa boxing goes high tech in training in colorado springs*. http://gazette.com/usa-boxing-goes-high-tech-in-training-in-colorado-springs/article/1500163, May 2013.

[3] D. Catuhe, *Programming with the kinect for windows software development kit*.

[4] E. . P. S. R. Council, *New technology will help improve athletes‚Äô performance*. http://www.epsrc.ac.uk/newsevents/news/2009/Pages/improveathletesperf.aspx, 2009.

[5] J. DeVries, *Fighters uncaged review | ign*. http://www.ign.com/articles/2010/11/09/fighters-uncaged-review, 2010.

[6] E. F. dynamics. http://elliottfightdynamics.com/, 2014.

[7] Gamesrader, *Fighters uncaged review | gamesradar*. http://www.gamesradar.com/fighters-uncaged-review/, 2011.

[8] Gazette, *Usa boxing goes high tech in training in colorado springs*. http://www.bbc.co.uk/news/technology-18735629, May 2013.

[9] T. Guardian, *London 2012 olympics: How athletes use technology to win medals*. http://www.theguardian.com/sport/2012/jul/04/london-2012-olympic-games-sport-technology, 2012.

[10] Metacritic, *Fighters uncaged critic reviews for xbox 360*. http://www.metacritic.com/game/xbox-360/fighters-uncaged/critic-reviews, Jan. 2011.

[11] Microsoft, *Infrared stream - msdn*. http://msdn.microsoft.com/en-us/library/jj663793.aspx, 2012.

[12] K. RECTOR, C. L. BENNETT, AND J. A. KIENTZ, *Eyes-Free Yoga : An Exergame Using Depth Cameras for Blind & Low Vision Exercise*, Int. ACM SIGACCESS Conf., (2013).

[13] R. C. UK, *Cutting edge 2012: The research behind sport*.
`http://www.rcuk.ac.uk/media/CuttingEdge2012/#bike`, 2012.

[14] ——, *Memorandum from research councils uk (rcuk) in response to the house of lords inquiry into sports and exercise science and medicine.*
`http://www.rcuk.ac.uk/RCUK-prod/assets/documents/submissions/SportsExercise.pdf`, 2012.

[15] WIKIPEDIA, *Great britain at the 2012 summer olympics*.
`http://en.wikipedia.org/wiki/Great_Britain_at_the_2012_Summer_Olympics`, 2014.

[16] K. YAMAOKA, M. UEHARA, T. SHIMA, AND Y. TAMURA, *Feedback of Flying Disc Throw with Kinect and its Evaluation*, Procedia Comput. Sci., 22 (2013), pp. 912–920.