AN INSTRUCTIONAL SUPPORT SYSTEM FOR MUAY THAI

BY

KETCHART KAEWPLEE

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF MASTER OF
ENGINEERING (INFORMATION AND COMMUNICATION
TECHNOLOGY FOR EMBEDDED SYSTEMS)
SIRINDHORN INTERNATIONAL INSTITUTE OF TECHNOLOGY
THAMMASAT UNIVERSITY
ACADEMIC YEAR 2015

AN INSTRUCTIONAL SUPPORT SYSTEM FOR MUAY THAI

BY

KETCHART KAEWPLEE

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF ENGINEERING (INFORMATION AND COMMUNICATION TECHNOLOGY FOR EMBEDDED SYSTEMS)

SIRINDHORN INTERNATIONAL INSTITUTE OF TECHNOLOGY

THAMMASAT UNIVERSITY

ACADEMIC YEAR 2015

AN INSTRUCTIONAL SUPPORT SYSTEM FOR MUAY THAI

A Thesis Presented

By

KETCHART KAEWPLEE

Submitted to

Sirindhorn International Institute of Technology

Thammasat University

In partial fulfillment of the requirements for the degree of

MASTER OF ENGINEERING (INFORMATION AND COMMUNICATION TECHNOLOGY FOR EMBEDDED SYSTEMS)

Approved as to style and content by

Advisor and Chairperson of Thesis Committee N.K.

(Assoc. Prof. Nirattaya Khamsemanan, Ph.D.)

Co-Advisor

(Asst. Prof. Cholwich Nattee, D.Eng.)

Committee Member and
Chairperson of Examination Committee Ithipan M.

(Dr. Ithipan Methasate)

Committee Member

(Prof. Shoichi Hasegawa)

June 2016

Acknowledgements

This research is supported by Thammasat University Research Fund 2012, Thailand Advanced Institute of Science and Technology (TAIST), National Science and Technology Development Agency (NSTDA), Tokyo Institute of Technology and Sirindhorn International Institute of Techcnology (SIIT), Thammasat University (TU) and the National Research University Project, Thailand Office of Higher Education Commission.

Abstract

AN INSTRUCTIONAL SUPPORT SYSTEM FOR MUAY THAI

by

KETCHART KAEWPLEE

Bachelor of Electrical Engineering, Chulalongkorn University, 2011
M.Eng (Information and Communication Technology for Embedded Systems), Sirindhorn International Institute of Technology Thammasat University, 2015

An Instructional Support System for Muay Thai is a program that is in the form of a game. Training data is collected from Muay Thai Professional boxers. The aim of the program is to make people understand Muay Thai by self-learning practice. The user watches a maneuver animation and performs the maneuver. Then, our program will show the performance result. This program uses skeleton data contained by the Microsoft Kinect. Because there are different organ lengths from different people, the system focuses on an angle of joints of each organ. In this thesis, we also propose algorithms which are Depth Tracking and Boundary Adjustment to improve the rate of skeleton recognition. With the Kinect recognition system alone, the accuracy of skeleton recognitions of all 24 standard Muay Thai maneuvers is 51%. After applying our proposed algorithms, we are able to bring the accuracy up 26%. On average, our proposed algorithms along with the Kinect recognition system yields accuracy of 77% on average standard Muay Thai maneuvers.

Keywords: Instruction Support System, Muay Thai, Kinect

Table of Contents

List of Figures

Chapter 1
Introduction

1.1 Motivation

The appearance of the Microsoft Kinect makes us realize the potential of motion sensing input device. The Microsoft Kinect recognition in frame per second is very high. So, the Microsoft Kinect can be used to control a character in a fighting game better than a game controller which has about 10-20 buttons. There are a lot of martial arts fighting games. Muay Thai is a combat that use the full function of human body. It uses a hand, elbow, knee, and foot. Consequently, we are interested in making an instructional program to teach Muay Thai using the Microsoft Kinect.

1.2 Objective

It is our goal to make an Instruction Support System for Muay Thai that,

• The program must be able to capture player's movements correctly.

• The program can calculate an accuracy of a movement compared to a given Muay Thai maneuver.

• The training set of this program is collected from professional Muay Thai masters.

1.3 Expected Outcome

The expected outcome of this thesis is an Instruction Support System for Muay Thai that is easy to use and understand. The player can use it individually. We will make an Instruction Support System for Muay Thai as a self-learning program.

1.4 Scope and Limitation

The Microsoft Kinect can not recognize a hidden parts of an object. The out of the line of sight may cause a problem in the Microsoft Kinect skeletal tracking. We will make an algorithm to improve skeletal tracking with only one Microsoft Kinect used in our system. The algorithm will be able to predict the position of a hidden organ when the out of the line of sight problem appears. The algorithm will use the

1.5 Structure of the thesis

This thesis contains 6 chapters.

Chapter 1: Introduction

This chapter serves as an introduction to our project.

Chapter 2: Literature review

This chapter reviews previous works and techniques used.

Chapter 3: A Rule-based Approach for Improving Kinect Skeletal Tracking

This chapter introduces various skeletal tracking problems and how we improve them.

Chapter 4: Maneuvers Recognition

This chapter will show all result of maneuvers recognition by our proposed algorithms.

Chapter 5: AN INSTRUCTIONAL SUPPORT SYSTEM FOR MUAY THAI

This chapter will describe our program and it performance.

Chapter 6: Conclusion

This chapter will summarize the entire of this project and direction.

Chapter 2
Literature Review

2.1 Muay Thai

Muay Thai [5] is a type of martial arts used in Thailand. It is known also as "the art of eight limbs" because both sides of hands, elbows, knees, and legs can be used to attack an opponent. Muay Thai has more than two hundred fighting techniques. Its maneuvers are a combination of fists, elbows, knees, shins and feet, which makes a deadly combat.

Muay Thai has more than 200 maneuvers, but there are only a few standard maneuvers. In this thesis, we present 24 standard maneuvers which refer to "Mae Mai Muay Thai The Art of Self-Defense"[6]. Standard Muay Thai maneuvers have the lead and rear maneuvers as shown in figure 2.1. The lead maneuvers use the body parts which are close to the opponent. The rear maneuvers use the body parts which are far away from the opponent.

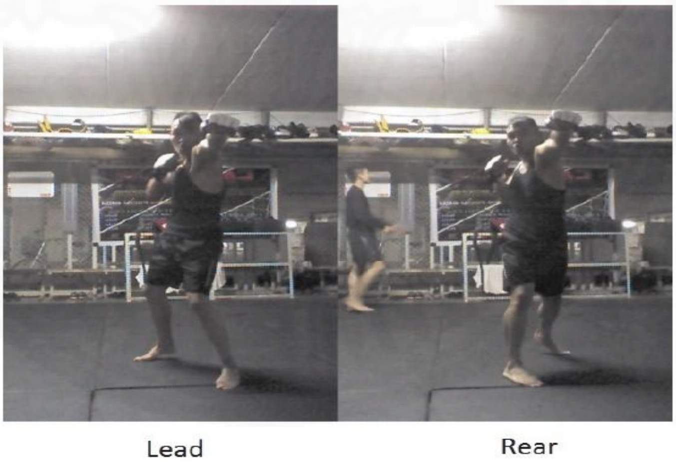

Figure 2.1: Lead Straight Punch and Rear Straight Punch

2.1.1 Lead Straight Punch (LSP)

The Lead Straight Punch maneuver is the striking punch with the fist from the shoulder level straight to the target. LSP is the punch of the leading fist throwing directly to the target by forces from the shoulder of the upright body while the feet are pushing against the floor. The feet do not move, but the weight is always transferred to the leading foot after throwing the punch.

2.1.2 Rear Straight Punch (RSP)

The Rear Straight Punch maneuver is the punch of rear fist of the fighting posture throwing directly to the target by forces from the shoulder, trunk, and feet. When launching the punch, the trunk, waist and hip will rotate forward with forces driven by the rear foot. The body weight is on the leading foot.

2.1.3 Lead Upper Punch (LUP)

The Lead Upper Punch maneuver is a swinging punch vertically directed upward with a bent, flexed elbow and a face-up fist. The Lead Upper Punch maneuver is best for striking at close quarters when the opponent is lowering his head. The landing targets are the opponent's chin, abdomen, chest, and face. The Lead Upper Punch maneuver is executed with the leading hand after engaging at close quarters. The boxer slightly bends his knees down with feet apart, lower his face-up lead fist and twist his body before launching the punch straight upward vertically. The rear arm is pulled by the rotation force. The lead foot is pushing the floor while stretching the body with the launching punch. The force of the punch depends on the flexing arm, the twist of the body, the twitch of the left arm and most importantly the stretching leg of the pushing foot against the floor.

2.1.4 Rear Upper Punch (RUP)

The Rear Upper Punch maneuver is executed with the rear hand. The boxer bends his rear knee down, twists his body, hip and shoulder to the other side before twitchingly launching the face-up fist of the flexing arm. The lead shoulder and lead fist must be punching quickly at the same time as the rear fist is launched.

2.1.5 Lead Hook Punch (LHP)

The Lead Hook Punch maneuver is the strike of the lead fist in the fighting posture. The boxer slightly bends his knees and twists his body to the other side, pushing the foot against the floor, dropping his lead arm at full stretch with the face-down fist before turning the rear arm and rear shoulder and swinging horizontally. The landing target is hit with the knuckle which is reinforced by the waist, trunk, shoulder, arm and lead fist while the body weight transferred from the rear foot to the lead foot.

2.1.6 Rear Hook Punch (RHP)

The Rear Hook Punch maneuver is executed with the rear hand. The boxer bends both knees and twists his hip, body and shoulder to the other side, stretches his rear arm away with the face-down fist and turns his lead arm before swinging the rear fist horizontally while the body weight is transferred from rear foot to the lead foot. The magnitude of force from the twisting of the body, hip and shoulder depends on the pushing feet against the floor, the body weight transferred the twitch of lead arm and the swing of the flexed rear arm.

2.1.7 Side Foot-Thrust (SFT)

The Side Foot-Thrust maneuver can be used as an ofensive tactic by thrusting with a foot targeting the upper part of the opponent's body such as chin, chest or abdomen, leaving him no chance to fight back after being pushed far behind. This can also be used as a defensive tactic. The Side Foot-Thrust maneuver requires a sideways twist of the body before delivery. The bent leg suddenly flicks forward to hit the target with its heel while the body weight is transferred to the standing leg.

2.1.8 Lead Foot-Thrust (LFT)

The Lead Foot-Thrust maneuver is the forward foot-thrust with toe curling to strike the opponent with the lead foot knuckles and heels. The landing target can be above or below the opponent's thigh or waist. Before launching the strike, the boxer stands upright to bluf the opponent, and moves around slightly to be ready on both feet consciously. The Lead Foot-Thrust can be used as a tactical ofense to put the opponent of-balance before attacking with another maneuver, or as a counterattack when the opponent makes a rush move and opens gaps. For example, if the opponent delivers the low round kick, the boxer can use the Lead

Foot-Thrust to counterattack at his kicking thigh, groin, or chest. Even when the opponent is well guarded, The Lead Foot-Thrust can be used to poke at the opponent's abdomen or groin, making him of-guard to parry it and opening the gap around the upper part of his body.

2.1.9 Rear Foot-Thrust (RFT)

The Rear Foot-Thrust maneuver is the forward foot-thrust with toe curling to strike the opponent with the rear foot knuckles and heels which similar to the Lead Foot-Thrust maneuver.

2.1.10 Lead Front Kick (LFK)

The Lead Front Kick maneuver is executed by lifting the knee straight forward, while keeping the foot and shin either hanging freely or pulled to the hip, and then straightening the leg in front of the opponent and striking the target area. It is desirable to retract the leg immediately after delivering the kick, to avoid the opponent trying to grapple the leg and to return to stable fighting stance.

2.1.11 Rear Front Kick (RFK)

The Rear Front Kick maneuver is the movement of the rear leg from its knee down to the end of its foot to strike the target which is similar to the Lead Front Kick maneuver.

2.1.12 Lead Low Round Kick (LLRK)

The Lead Low Round Kick maneuver is a kick in which the boxer swings his or her leg around in a semicircular motion, striking with the front of the leg or foot. This type of kick is utilized in many diferent martial arts and is popular in both non-contact and full-contact martial arts competitions.

2.1.13 Rear Low Round Kick (RLRK)

The Rear Low Round Kick maneuver is the maneuver that the rear leg movement is curving up and striking horizontally which is similar to the Lead Low Round Kick maneuver.

2.1.14 Lead High Round Kick (LHRK)

The Lead High Round Kick maneuver is executed at the head height requires the boxer to torque around the hip as a fulcrum. Traditional stylists would prefer to keep their weights firmly under their hips and do short range kicks and retain their abilities to counter a strike. The Lead High Round Kick maneuver requires the boxer to raise the center of gravity somewhat, tilt the hip, lean back and swing the lead leg forward.

2.1.15 Rear High Round Kick (RHRK)

The Rear High Round Kick maneuver is the maneuver that the rear leg movement is curving up and striking horizontally which is similar to the Lead High Round Kick maneuver.

2.1.16 Swing Knee (SK)

In the Swing Knee maneuver, the knee is executed by twisting the hip forward to swing the knee strike down to the target. The foot tip, the knee, and the thigh of the standing leg must align in a straight line.

2.1.17 Low Straight Knee (LSK)

To execute the Low Straight Knee maneuver, the body weight is mainly on the rear foot. The boxer transfers his weight to the lead foot. When the foot firmly touches the ground, the boxer raises the lead knee up directly while transferring the weight back to the rear foot and augmenting force to the striking knee. The sudden tightening of abdominal muscle that pulls the knee up will drive the strike straight ahead, this is good for engaging at close quarters, clinching or yanking the opponent toward the strike.

2.1.18 High Straight Knee (HSK)

The High Straight Knee maneuver is similar to Low Straight Knee maneuver, but High Straight Knee maneuver raises the lead knee up much higher.

2.1.19 Lead Shooting Elbow (LSE)

The Lead Shooting Elbow maneuver is the strike where the bending lead elbow points straight to the opponent or in the direction of incoming attack. The lead arm of the striking elbow must be flexed and kept close to the chin, shoulder and head to reinforce an impact on the opponent's forehead, eyebrows, eyes, nose, mouth, chin, chest, shoulder and abdomen.

2.1.20 Rear Shooting Elbow (RSE)

The Rear Shooting Elbow maneuver is similar to Lead Shooting Elbow maneuver, but Rear Shooting Elbow maneuver uses the rear elbow maneuver.

2.1.21 Lead Downward Elbow (LDE)

The Lead Downward Elbow maneuver is executed with the lead elbow. The fighter raises his lead elbow up high and twists his body and his lead shoulder to the other side before throwing the elbow down to the target while transferring the body weight from the rear foot to the lead foot. The violence of the strike depends on the force from pushing the feet against the floor, the twist of body, the twitch of arm and shoulder, the raising of the elbow, and the quickness of strike.

2.1.22 Rear Downward Elbow (RDE)

The Rear Downward Elbow maneuver is similar to the Lead Downward Elbow maneuver, but the Rear Downward Elbow maneuver uses the rear elbow.

2.1.23 Lead Round Elbow (LRE)

To execute the Lead Round Elbow maneuver, the boxer bends his lead knee down and throws the lead elbow horizontally to the other side while stretching the knee to transfer the weight to the lead foot and the lead elbow. The body and rear shoulder are twisted to the other side. These movements must be done in a twitch all at once to reinforce the strike. The violence of the strike depends on the force from pushing the feet against floor, the transfer weight to the lead foot and lead elbow, and the twist of body and shoulder, the swing of the elbow, the stretch of shoulder, abdomen and rear arm. The landing targets of Lead Round Elbow are the opponent's chin, lip, nose, eyes, eyebrows, forehead, and temples.

2.1.24 Rear Round Elbow (RRE)

The Rear Round Elbow maneuver is similar to the Lead Round Elbow maneuver, but the Rear Round Elbow maneuver uses the rear elbow.

Figure 2.2: Muay Thai

2.2 Related Work

2.2.1 Martial Arts in Artificial Reality [4]

This paper presents an artificial reality martial arts game installation as shown in Figure 2.3, together with the results and observations from testing the game. With real-time image processing and computer vision, the video image of the user is embedded inside 3D graphics on a virtual playfield facing virtual opponents. Collisions are detected between enemies and the outline pixels of the user. The velocity of the outline pixels is estimated using the OpenCV implementation of pyramidal Lucas-Kanade optical flow. This technique focus on the player hitting the target. It was not designed to recognize the entire body movement. Therefore, it is not suitable for Muay Thai maneuvers which are consider many body parts.

Figure 2.3: Martial Arts in Artificial Reality

2.2.2 A Kinect based Golf Swing Score and Grade System Using GMM

In this work, Kinect is used to capture skeleton data from a user. A vector quantization is used to transform the data to symbol sequence. The Serial correlation GMM model and GMM-KL divergence kernel are used to evaluate and recognize the grade of Golf Swing as shown in Figure 2.4 These methods separate the extracted data from skeleton data into states of movements and evaluate them separated by Kullback-Leibler divergence and grade them by Support Vector Machine. These methods only recognize velocity and angle on single joint. This technique can not applied to Muay Thai maneuvers because, in a standard maneuver several joints must be considered.

Figure 2.4: A Kinect based Golf Swing Score and Grade System Using GMM and SVM

2.2.3 Game Based Approach to Learn Martial Arts for Beginners [2]

In this paper, the composers introduce an interactive digital system that promote informal learning of martial arts, Karate, with the aid of Kinect as shown in figure 2.5. They explain that Kinect is a good device for camera motion tracking because Kinect price is relatively cheap compared to other available 3D sensors in the market, Kinect does not require to put the sensor on player's body and Kinect provides 3D data modeling which reduces their computation work. The composers propose a system with four steps of the method of Kinect interaction with the user. Firstly, Kinect with NITE middleware captures player position. Secondly, The system identifies player's joint positions. Thirdly, the system calculates the accuracy of player maneuver by measuring the diferent angle between player position and correct position from the definition of karate. Finally, The system instructs the player to improve maneuver accuracy by conveying the tips. The score of maneuver accuracy is announced in this step. In this work, the ma- neuver accuracy was calculated using the diferent angle between player position and correct position from the definition of karate. However, the out of line sight problems can occur during the accuracy calculation. This calculation will be less efective when it is used with Muay Thai maneuvers which have a lot of rotation movement that causes the out of line of sight problems.

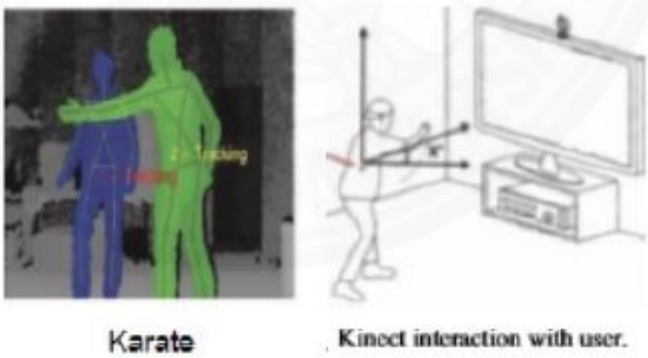

Figure 2.5: Game Based Approach to Learn Martial Arts for Beginners

2.2.4 Edutainment - Thai Art of Self-Defense and Boxing by Motion Capture Technique [3]

This work uses 9 optical motion capture cameras and 42 sensor spot in both 2 actors to capture data as shown in Figure 2.6. Though the movement data is more precise, this technique is expensive and not practical. The 3D animations show highly detail about Muay Thai movement. This software is not an interactive program. Users only can watch the animations.

Figure 2.6: Edutainment - Thai Art of Self-Defense and Boxing by Motion Capture Technique

2.3 Kinect

Microsoft Kinect [8] is a line of motion sensing input devices created by Microsoft for Xbox 360 and Xbox One video game consoles and Windows PCs as shown in Figure 2.7 Based around a webcam-style add-on peripheral, it enables users to control and interact with their console/computer without the need for a game controller, through a natural user interface using gestures and spoken commands..

Figure 2.7: Kinect

This picture is from https://msdn.microsoft.com/en-us/library/hh855355.aspx

2.3.1 Kinect Data

2.3.1.1 Joints

Data collected from Microsoft Kinect consist of 20 joints as show in Figure 2.8.

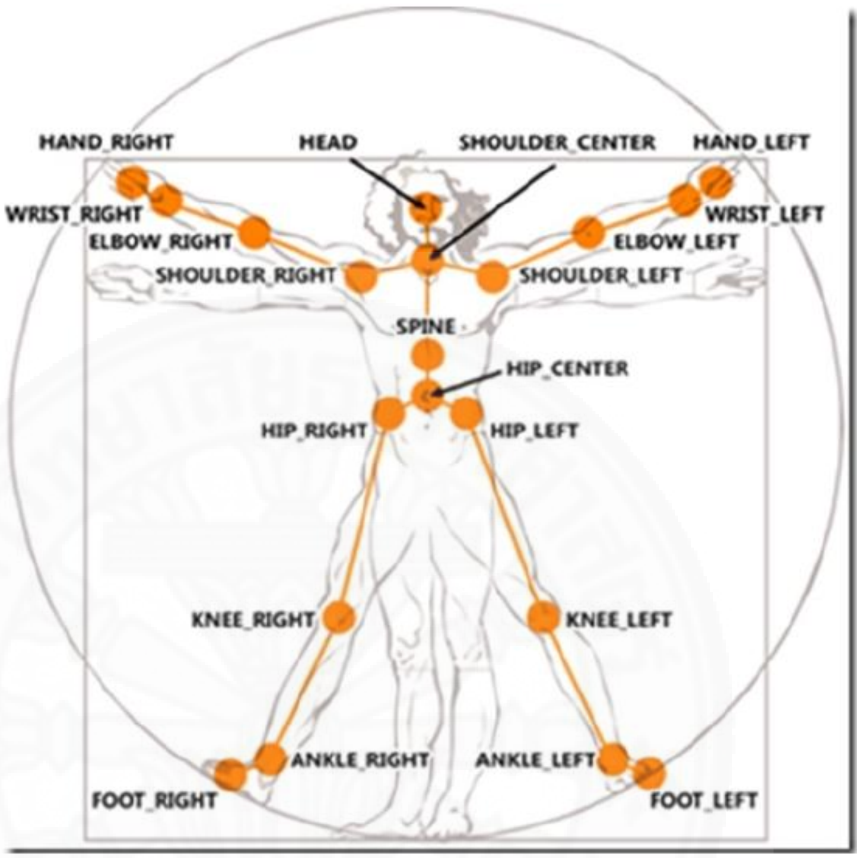

Figure 2.8: Joints

This picture is from

https://msdn.microsoft.com/enus/library/jj131025.aspx

2.3.1.2 Posture

A posture S_i , is represented by a sequence of frames $\{f_{ij}\}_{j=1}^{N_i}$ where f_{ij} is a vector containing 20 coordinates of body joints of the frame j th.

$$f_{ij} = \begin{bmatrix} P_{hip\ center}(j) \\ P_{spine}(j) \\ P_{shoulder\ center}(j) \\ \vdots \\ P_{foot\ right}(j) \end{bmatrix} \tag{2.1}$$

Where $P_{joint}(j)$ is the coordinate of a joint in the frame j^{th} .

2.3.1.3 Bone

Each bone can be represented by 3 bone vectors. as Figure 9. For example, the left hip bone is represented as a vector below;

$$\vec{B}_{Hip\ left} = \begin{bmatrix} \vec{X}_{Hip\ left} \\ \vec{Y}_{Hip\ left} \\ \vec{Z}_{Hip\ left} \end{bmatrix} \tag{2.2}$$

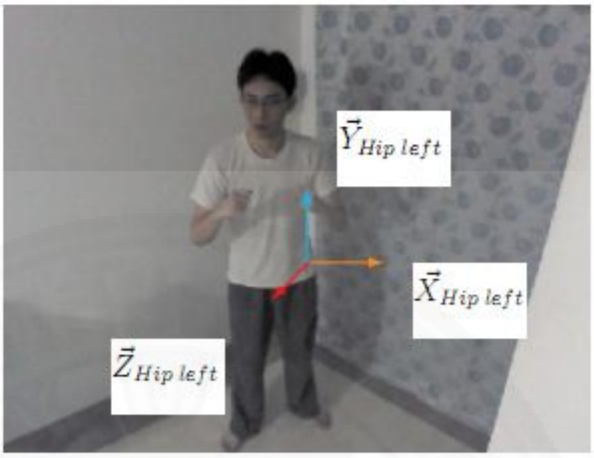

Figure 2.9: Bone

2.3.1.4 An Absolute Bone Rotation

Every bone has an Absolute Bone Rotation. An Absolute Bone Rotation of any joint is a 3×3 matrix that informs the direction between Kinect and joint.

For example, if the direction of Kinect's camera is $\begin{bmatrix} \vec{X}_{kinect} \\ \vec{Y}_{kinect} \\ \vec{Z}_{kinect} \end{bmatrix}$ and the direction of hip center joint is $\begin{bmatrix} \vec{X}_{Hip\ center} \\ \vec{Y}_{Hip\ center} \\ \vec{Z}_{Hip\ center} \end{bmatrix}$ as shown in Figure 10, then an absolute bone rotation $R^{(a)}_{Hip\ center}$ of a hip center is defined by.

$$\begin{bmatrix} \vec{X}_{Hip\ center} \\ \vec{Y}_{Hip\ center} \\ \vec{Z}_{Hip\ center} \end{bmatrix} = R^{(a)}_{Hip\ center} \times \begin{bmatrix} \vec{X}_{kinect} \\ \vec{Y}_{kinect} \\ \vec{Z}_{kinect} \end{bmatrix} \tag{2.3}$$

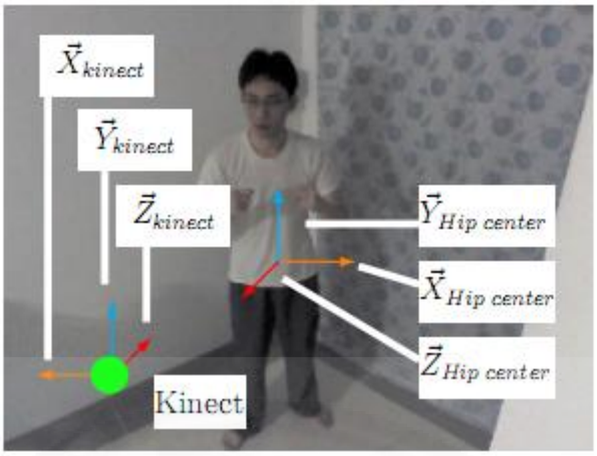

Figure 10: Absolute Bone Rotation

2.3.1.5 A Hierarchical Bone Rotation

Every joint has a Hierarchical Bone Rotation which is a matrix that informs the direction between joints that stay next to each other.

For example, if the direction of left shoulder is $\begin{bmatrix} \vec{X}_{Shoulderleft} \\ \vec{Y}_{Shoulderleft} \\ \vec{Z}_{Shoulderleft} \end{bmatrix}$ and the direction of left elbow is $\begin{bmatrix} \vec{X}_{Elbowleft} \\ \vec{Y}_{Elbowleft} \\ \vec{Z}_{Elbowleft} \end{bmatrix}$ as shown in Figure 11, then a hierarchical bone rotation of left elbow $R^{(a)}_{Elbowleft}$ of left elbow is defined by.

$$\begin{bmatrix} \vec{X}_{Elbow\ left} \\ \vec{Y}_{Elbow\ left} \\ \vec{Z}_{Elbow\ left} \end{bmatrix} = R^{(h)}_{Elbow\ left} \times \begin{bmatrix} \vec{X}_{Shoulder\ left} \\ \vec{Y}_{Shoulder\ left} \\ \vec{Z}_{Shoulder\ left} \end{bmatrix} \tag{2.4}$$

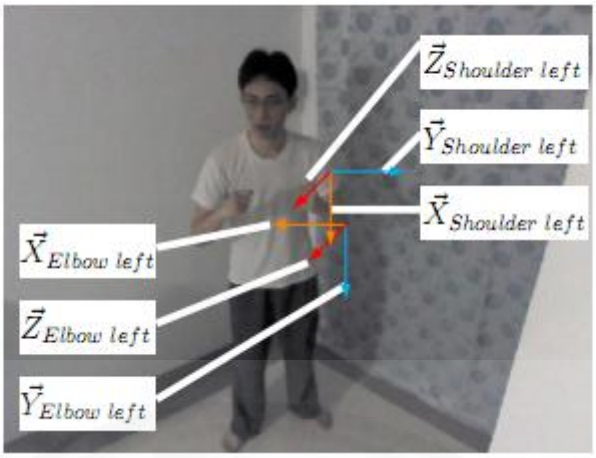

Figure 11: Hierarchical Bone Rotation

2.4 Techniques used in this system

2.4.1 Dynamic Time Warping (DTW)

In time series analysis, the dynamic time warping (DTW) is an algorithm for measuring a similarity between two temporal sequences which may vary in time or speed. For instance, similarities in walking patterns could be detected using DTW, even if one person was walking faster than the other, or if there were accel- erations and decelerations during the course of an observation. Any data which can be turned into a linear sequence can be analyzed with DTW.

Let

$$X = (x_1, x_2, \ \dots \ , x_M) \tag{2.5}$$

$$Y = (y_1, y_2, \ \dots \ , y_N) \tag{2.6}$$

be two sequences. We define a cost $M \times N$ Matrix, C where

$$C_{i,j} = d(x_i, y_j) \tag{2.7}$$

and d is the euclidean distance between x_i and y_j.

If,

$$\gamma_{i,j} = C_{i,j} + min\{\gamma_{i-1,j}, \gamma_{i,j-1}, \gamma_{i-1,j-1}\} \tag{2.8}$$

$$\gamma_{1,1} = 0 \tag{2.9}$$

$$\gamma_{i,0} = \infty \ (i = 1, 2, ..., M) \tag{2.10}$$

$$\gamma_{0,j} = \infty \ (j = 1, 2, ..., N) \tag{2.11}$$

then, the Dynamic Time Warping distance of X and Y is defined as

$$DTW(X, Y) = \gamma_{M,N} \tag{2.12}$$

2.4.2 Sliding Window Dynamic Time Warping (SW DTW)

The Sliding Window Dynamic Time Warping (SW DTW) is another form of DTW that find the least euclidean distance between two data in different sliding. Let

$$X = (x_1, x_2, \dots, x_M) \tag{2.13}$$

$$Y = (y_1, y_2, \dots, y_N) \tag{2.14}$$

be two sequences. If $M < N$,

$$Y_1 = (y_1, y_2, \dots, y_M) \tag{2.15}$$

$$Y_2 = (y_2, y_3, \dots, y_{M+1}) \tag{2.16}$$

$$\vdots \tag{2.17}$$

$$Y_{N-M+1} = (y_{N-M+1}, y_{N-M+2}, \dots, y_N) \tag{2.18}$$

We define a cost $M \times M$ Matrix, C, where $k = 1,2,\dots,N-M+1$, as

$$C_{k,i,j} = d_k(x_i, y_j) \tag{2.19}$$

and d_k is euclidean distance between x_i and y_j where $y_j \in Y_k$.

If,

$$\gamma_{k,i,j} = C_{k,i,j} + min\{\gamma_{k,i-1,j}, \gamma_{k,i,j-1}, \gamma_{k,i-1,j-1}\} \tag{2.20}$$

$$\gamma_{k,1,1} = 0 \tag{2.21}$$

$$\gamma_{k,i,0} = \infty \ (i = 1, 2, \dots, M) \tag{2.22}$$

$$\gamma_{k,0,j} = \infty \ (j = 1, 2, \dots, M) \tag{2.23}$$

then, the Sliding Window Dynamic Time Warping distance of X and Y is defined as

$$SWDTW(X,Y) = min\{\gamma_{1,M,M}, \dots, \gamma_{N-M+1,M,M}\} \tag{2.24}$$

2.4.3 Hidden Markov Model (HMM)

The hidden Markov model (HMM) is a statistical Markov model in which the system being modeled is assumed to be a Markov process with unobserved (hidden) states. An HMM can be presented as the simplest dynamic Bayesian network.

Suppose we have a sequential data

$$u = \{u_1, u_2, \dots, u_t, \dots, u_T\}, u_t \in R^a \tag{2.25}$$

Every $u_t, t = 1,\dots,T$ is generated by a hidden state, s_t.

Given the present, the future is independent of the past.

$$P(s_{t+1}|s_t, s_{t-1}, \ldots, s_0) = P(s_{t+1}|s_t) \tag{2.26}$$

We therefore define a transition probability from state k to state l, $a_{k,l}$ as

$$a_{k,l} = P(s_{t+1} = l | s_t = k), \tag{2.27}$$

where k and l are state numbers $\in \{1,2,\ldots,M\}$. M is the total number of states. Let π_k be an initial probability of a state k. We have

$$\sum^{M} a_{k,l} = 1 \text{ for any } k \text{ and } \sum^{M} \pi_k = 1 \tag{2.28}$$

Consequently,

$$P(s_1, s_2, \ldots, s_T) = P(s_1)P(s_2|s_1)P(s_3|s_2)\cdots P(s_T|s_{T-1}) = \pi_{s_1} a_{s_1,s_2} a_{s_2,s_3} \cdots a_{s_{T-1},s_T} \tag{2.29}$$

2.4.4 Baum-Welch Algorithm

Baum-Welch algorithm is used to find the unknown parameters of a hidden Markov model (HMM). It makes use of the forward-backward algorithm.

Let Xt be a discrete hidden random variable with N possible values. We assume that $P(X_t/X_{t-1})$ is independent of time t, which leads to the definition of the time independent stochastic transition matrix

$$A = \{a_{ij}\} = P(X_t = j | X_{t-1} = i). \tag{2.30}$$

The initial state distribution (i.e. when $t = 1$) is given by

$$\pi_i = P(X_1 = i). \tag{2.31}$$

The observation variables Y_t can take one of K possible values. The probability of a certain observation at time t for state j is given by

$$b_j(y_t) = P(Y_t = y_t | X_t = j). \tag{2.32}$$

Forward procedure

Let $\alpha_i(t) = P(Y_1 = y_1,..., Y_t = y_t, X_t = i \mid \theta)$ be the probability of the $y_1, y_2,..., y_t$ and being in state i at time t. This is defined recursively as

$$\alpha_i(1) = \pi_i b_i(y_1) \tag{2.33}$$

$$\alpha_j(t+1) = b_j(y_{t+1}) \sum_{i=1}^{N} \alpha_i(t) a_{ij} \tag{2.34}$$

where π_i is initial probabilities of states in formula 2.31, a is the transition probabilities in formula 2.27 and b is the probability of a certain observation in formula 2.32.

Backward procedure

Let $\beta_i(t) = P(Y_{t+1} = y_{t+1},..., Y_T = y_T \mid X_t = i, \theta)$ be the probability of the ending partial sequence $y_{t+1},..., y_T$ given starting state i at time t. We calculate $\beta_i(t)$ as,

$$\beta_i(T) = 1 \tag{2.35}$$

$$\beta_i(t) = \sum_{j=1}^{N} \beta_j(t+1) a_{ij} b_j(y_{t+1}) \tag{2.36}$$

Recursive Part

We can calculate the temporary variables, according to Bayes' theorem: The probability of being in state i at time t, given the observed sequence Y and the parameters θ is defined as

$$\gamma_i(t) = P(X_t = i \mid Y, \theta) = \frac{\alpha_i(t)\beta_i(t)}{\sum_{j=1}^{N} \alpha_j(T)}, \tag{2.37}$$

The probability of being in state i and j at times t and $t+1$ respectively, given the observed sequence Y and parameters θ , $\xi_{ij}(t)$ is defined as

$$\xi_{ij}(t) = P(X_t = i, X_{t+1} = j \mid Y, \theta) = \frac{\alpha_i(t) a_i j \beta_j(t+1) b_j(y_{t+1})}{\sum_{k=1}^{N} \alpha_k(T)} \tag{2.38}$$

The expected frequency spent in state i at time 1, π_i^* is defined as

$$\pi_i^* = \gamma_i(1) \tag{2.39}$$

The expected number of transitions from state i to state j compared to the expected total number of transitions away from state i, a_{ij}^* is defined as

$$a_{ij}^* = \frac{\sum_{t=1}^{T-1} \xi_{ij}(t)}{\sum_{t=1}^{T-1} \gamma_i(t)} \tag{2.40}$$

This is equivalent to the number of times state i is observed in the sequence from $t=1$ to $t=T-1$ (T is total sequence). The expected number of times the output observations have been equal to v_k while in state i over the expected total number of times in state i, $b_i^*(v_k)$ is defined as

$$b_i^*(v_k) = \frac{\sum_{t=1}^{T} I(y_t, v_k)\gamma_i(t)}{\sum_{t=1}^{T} \gamma_i(t)} \tag{2.41}$$

where

$$I(y_t, v_k) = \begin{cases} 1, \text{if } y_t = v_k \\ 0, \text{otherwise.} \end{cases} \tag{2.42}$$

which is an indicator function. These steps are now repeated iteratively until a desired level of convergence. It is possible to over-fit a particular data set. That is $P(Y \mid \theta_{final}) > P(Y \mid \theta_{true})$. The algorithm also does not guarantee a global maximum.

2.4.5 K-Nearest Neighbors Algorithm(k-NN)

The K-Nearest Neighbors algorithm (k-NN) is a distance-based method used for classifications and regressions. The input consists of the k closest training examples in the feature space. The output depends on whether k-NN is used for classifications or regressions.

In k-NN classifications, the output is a class of membership. An object is classified by a majority vote of its neighbors. The object being assigned to the class most common among its k nearest neighbors (k is a positive integer, typically small). If k = 1, then the object is simply assigned to the class of that single nearest neighbor.

In k-NN regressions, the output is the property value for the object. This value is the average of the values of its k nearest neighbors.

2.4.6 Hierarchical Clustering

Hierarchical Clustering (also called hierarchical cluster analysis or HCA) is a method of cluster analysis which seeks to build a hierarchy of clusters. Strategies for hierarchical clustering generally fall into two types:

• Agglomerative: This is a "bottom up" approach. Each observation starts in its own cluster, and pairs of clusters are merged as one moves up the hierarchy.

• Divisive: This is a "top down" approach. All observations start in one cluster, and splits are performed recursively as one moves down the hierarchy.

Let xi be an input sample with $(x_1, x_2,$ to be clustered, where n is the total number of input samples $(i = 1,2,$. In order to decide which clusters should be combined (for agglomerative), or where a cluster should be split (for divisive), a measure of dissimilarity between sets of observations is required. In most methods of hierarchical clustering, this is achieved by a use of an appropriate metric (a measure of distance between pairs of observations, mostly euclidean distance), and a linkage criterion which specifies the dissimilarity of sets as a function of the pairwise distances of observations in the sets.

Linkage criteria

The linkage criterion determines the distance between sets of observations as a function of the pairwise distances between observations. Some commonly used linkage criteria between two sets of observations A and B are

Maximum or complete linkage clustering is defined by

$$\max \{d(x,y) : x \in A, y \in B\}. \tag{2.43}$$

Minimum or single-linkage clustering is defined by

$$\min \{d(x,y) : x \in A, y \in B\}. \tag{2.44}$$

Mean or average linkage clustering or UPGMA is defined by

$$\frac{1}{|A||B|}\sum_{x\in A}\sum_{y\in B} d(x,y) \tag{2.45}$$

Chapter 3

A Rule-based Approach for Improving Kinect Skeletal Tracking

3.1 False Skeleton Problem

False skeleton recognitions by Microsoft Kinect happen when there are out of the line of sight, unnatural human posture or rapid movement. In working with Muay Thai standard maneuvers, we can classify false skeleton recognition into three categories.

3.1.1 Dead Limb

Dead Limb is a false skeleton problem when the sequence of postures is too short (very fast movement). Kinect is unable to recognize the quick change of limbs especially a rapid movement of a foot. This makes a false skeleton recognition where quick moving limb appears unmoved.

Figure 3.1 demonstrates an example of a dead limb scenario in Rear High Round Kick posture. Note that the skeleton for the right leg is not moved. However, the skeleton of the upper body is false recognition.

Figure 3.2. is another demonstration of a dead limb scenario. The Kinect Skeletal Tracking system is not able to realize that the right leg has already moved.

Standard Muay Thai maneuvers that frequently contain dead limb scenario are Lead-Rear Front Kick, Low-High Round Kick, Foot- Thrust, Low-High Straight Knee and Swing Knee.

3.1.2 Swap

Swap is the situation when the Kinect recognition system recognize wrong limb positions, e.g. a leg is interpreted as an arm and vice versa.

An example of a swap scenario is shown in Figure 3.3. Notice the Kinect recognition system can correctly recognize right arm and right leg in the first two frames. However, after the quick limb position change in the third frame where the right arm is unseen behind the right leg, Kinect recognition system interprets the right arm as the right leg and vice versa since generally that an arm is always above a leg in a human skeleton.

Standard Muay Thai maneuvers that frequently contain the swap scenarios are Lead-Rear Front Kick and Low-High Round Kick.

3.1.3 Short-Coming

Short-Coming is a false skeleton detection scenario when a skeleton of a limb is moved in the same direction as the limb itself but the skeleton is not embedded inside the silhouette of the body. This scenario includes quick moving sequences of postures and unnatural skeleton forms.

An example of a short-coming scenario is shown in Figure 3.4. Note that the boxer spreads out left leg straightly to the left side. However, the left leg in Kinect skeleton data is shorter than real left leg and it is outside the silhouette of the body. The foot, ankle and knee joint positions stay in wrong positions.

The short-coming scenario can be found in all Muay Thai maneuvers.

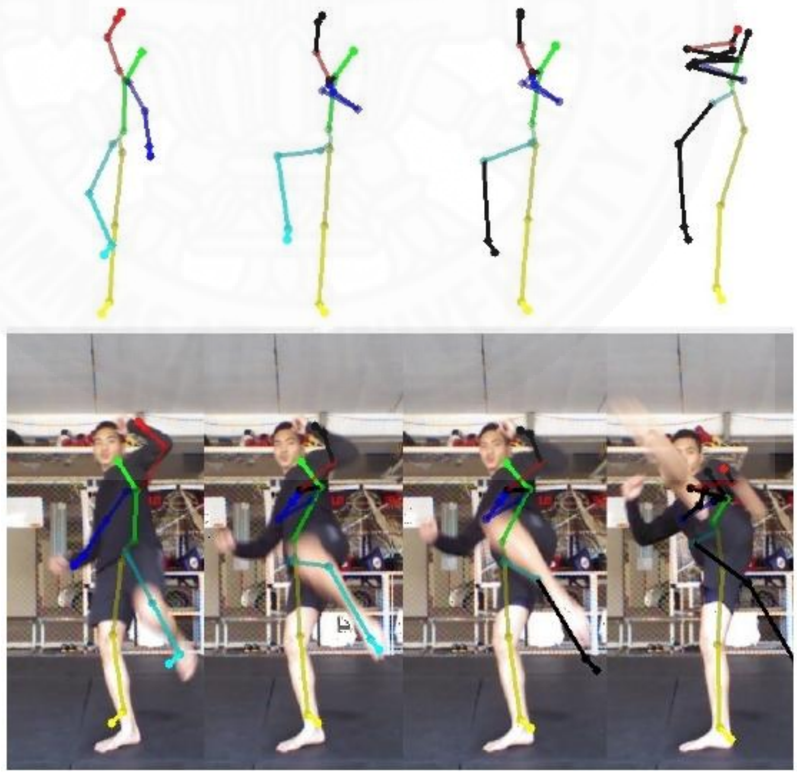

Figure 3.1: Skeleton images of a Rear High Round Kick maneuver by Kinect recognition system. The top image displays a side-view of the bottom image.

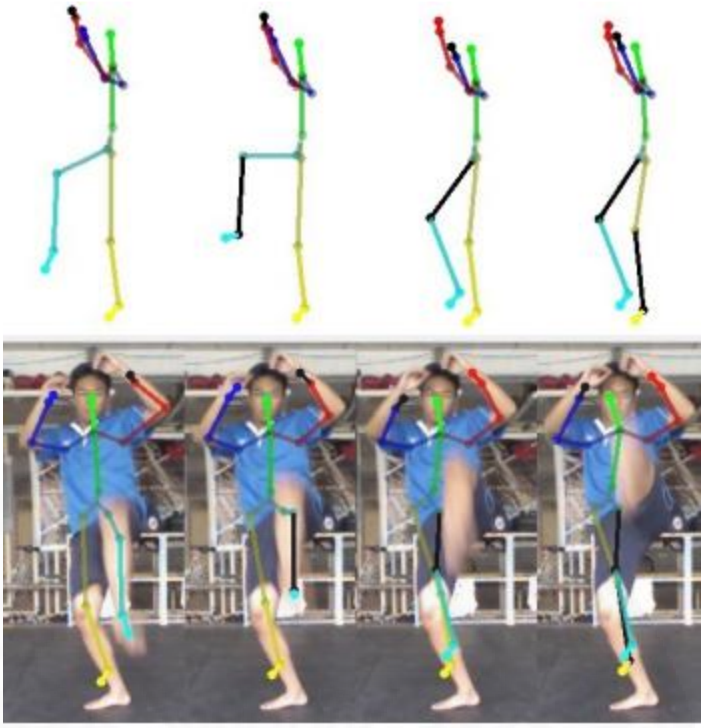

Figure 3.2: Skeleton images of a Lead Foot-Thrust maneuver by Kinect recognition system. The top image displays a side-view of the bottom image.

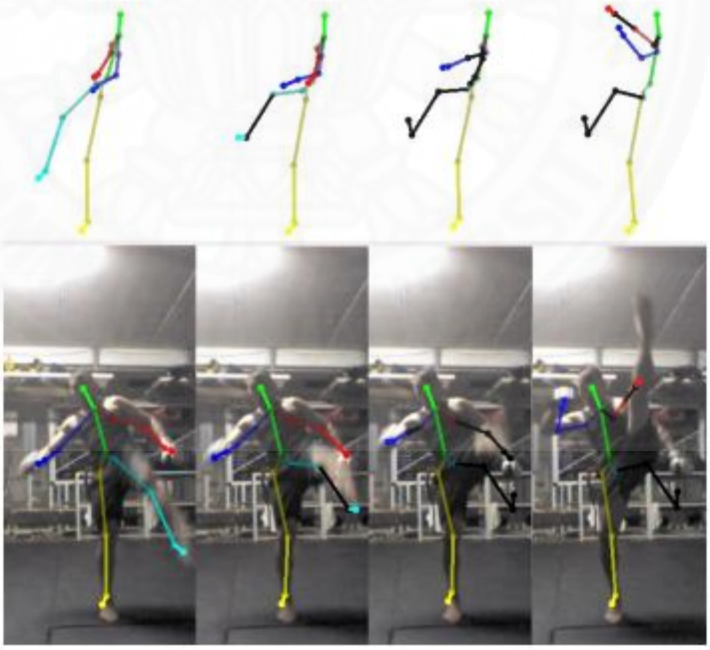

Figure 3.3: Skeleton images of a Rear Front Kick maneuver by Kinect recognition system. The top image displays a side-view of the bottom image.

Figure 3.4: Skeleton images of a Side Foot-Thrust recognition maneuver by Kinect Skeleton Tracking system.

3.2 Methodology

For a frame t, we can obtain a 640x480 matrix of skeleton data, denoted by p(t), from Kinect SDK. Each entry of the matrix, $p(t)_{ij} \in \mathbb{R}^3$, represents a 3-dimensional coordinate in the skeleton space. Let $u(t) \in \mathbb{R}^3$ be 3-dimensional coordinate of a skeleton joint u in the frame t. Thus, u(t) is an entry of p(t) identified by Kinect Runtime as a joint coordinate.

From our observation, if the angle between u(t) and u(t+1), denoted by $\Theta(u(t), u(t+1))$, is greater than or equal to 45^o, then the false skeleton detection is very likely to happen. This drastic change in the angle is resulted from a rapidly moving sequence of postures. In this case, we use "Depth Tracking" (as described in algorithm 1) to adjust u(t + 1). In case that $\Theta(u(t), u(t+1))$ is less than 45^o, we check if u(t + 1) is an "inferred" joint. If so, we also apply the Depth Tracking in this case as well. Then, we apply the "Boundary Adjustment" (as described in algorithm 2) to the output skeleton of Depth Tracking.

Algorithm 1 shows the pseudo-code of the Depth Tracking. The algorithm checks the condition of the given joint coordinate, u(t). If it is an inferred joint, or its angle of rotation is greater than or equal to 45 degrees, the coordinate of the joint in frame t + 1, u(t + 1) is corrected to the closest coordinate in the skeleton space. Here, d(p, q) is the Euclidean distance between coordinates p and q.

Algorithm 2 shows the pseudo-code of the Boundary Adjustment. This algorithm adjusts positions of joints at the boundary of the silhouette of each frame. The algorithm finds boundary points which are left most point u_{xmin}, right most point u_{xmax} and the closest point to Kinect u_{zmin}. Then it searches through all 20 joints to find joints that nearest to these boundary points. The positions of these joints, then, are replaced by the nearest boundary positions. Due to the out of a line of sight problem, it is not useful to take the furthest point to Kinect into our consideration.

Algorithm 1: Depth Tracking

Input : $u(t)$, $u(t+1)$, and $p(t+1)$
Output: $u'(t+1)$

if $\Theta(u(t), u(t+1)) \geq 45^o$ ***or*** type$(u(t+1)) = \textit{inferred}$ **then**
  $u'(t+1) \leftarrow \arg\min_{c \in p(t+1)} d(u(t), c)$
  BoundaryAdjustment$(u'(t+1), p(t+1))$
else
  $u'(t+1) \leftarrow u(t+1)$

Algorithm 2: Boundary Adjustment

Input : $u(t)$, $p(t)$
Output: $u'(t)$

$\Pi \leftarrow \textit{player}(p(t))$
$\Omega \leftarrow \textit{joint}(p(t))$
$u_{xmin} \leftarrow \arg\min_{\xi=(x,y,z)\in\Pi} x$
$\arg\min_{u\in\Omega} d(u_{xmin}, u) \leftarrow u_{xmin}$
$u_{xmax} \leftarrow \arg\max_{\xi=(x,y,z)\in\Pi} x$
$\arg\min_{u\in\Omega} d(u_{xmax}, u) \leftarrow u_{xmax}$
$u_{zmin} \leftarrow \arg\min_{\xi=(x,y,z)\in\Pi} z$
$\arg\min_{u\in\Omega} d(u_{zmin}, u) \leftarrow u_{zmin}$

3.3 Experiment

We captured all 24 standard Muay Thai maneuvers from 4 professional Muay Thai boxers. We asked each boxer to perform 5 times for each maneuver. We have a total of 480 videos. Each video has a sequence of skeleton frames. We process each video two diferent ways. First we run each video through the Kinect Skeleton Tracking system alone. Then we also run the same video through our proposed algorithms with Kinect skeleton system. The result of each way of processing is a video with skeleton data. For each output video, we consider it as an accurate result only when the skeleton data is accurate on all the frames in a video. If not, the result is considered as a false skeleton detection video.

3.4 Results & Discussion

Figure 3.5, 3.6, 3.7, 3.8 show the comparison of skeleton data result of the Kinect Skeleton Tracking system alone and the result of the proposed algorithm. Note that both depth tracking and boundary adjustment are needed to receive the correct skeleton data in all three types of problems.

Table 3.1 shows the experimental results after running all 24 maneuvers through our proposed algorithm. Notice that the proposed algorithm performs well especially maneuvers that involve with leg's movements such as Rear Front Kick, and Lead Foot- Thrust maneuvers. Over all maneuvers, the accuracy is up 26% on average.

Table 3.1: Experimental Result

Maneuver	Kinect skeleton detection	Proposed Algorithm
Lead Straight Punch	80%	95%
Rear Straight Punch	75%	100%
Lead Upper Punch	90%	90%
Rear Upper Punch	90%	95%
Lead Hook Punch	60%	85%
Rear Hook Punch	90%	90%
Side Foot-Thrust	75%	80%
Lead Front Kick	5%	60%
Rear Front Kick	0%	90%
Lead Low Round Kick	40%	90%
Rear Low Round Kick	45%	70%
Lead High Round Kick	0%	40%
Rear High Round Kick	0%	40%
Lead Foot-Thrust	0%	85%
Rear Foot-Thrust	0%	85%
Low Straight Knee	95%	95%
High Straight Knee	10%	40%
Swing Knee	20%	45%
Lead Shooting Elbow	90%	90%
Rear Shooting Elbow	95%	95%
Lead Downward Elbow	80%	80%
Rear Downward Elbow	35%	45%
Lead Round Elbow	95%	95%
Rear Round Elbow	65%	65%
Average	51%	77%

3.5 Summary

In this chapter, we propose an algorithm to improve Kinect skeleton detection ability. Our algorithm is designed to eliminate the false skeleton detection in situations that involved rapidly moving sequences of postures, out of a line of sight of the camera and inhuman postures. We used an angle of rotations, depth information and boundary of a silhouette to help adjust the position of joints. We apply our algorithm to standard Muay Thai maneuvers and able to bring the skeleton detection accuracy up from 51% to 77%.

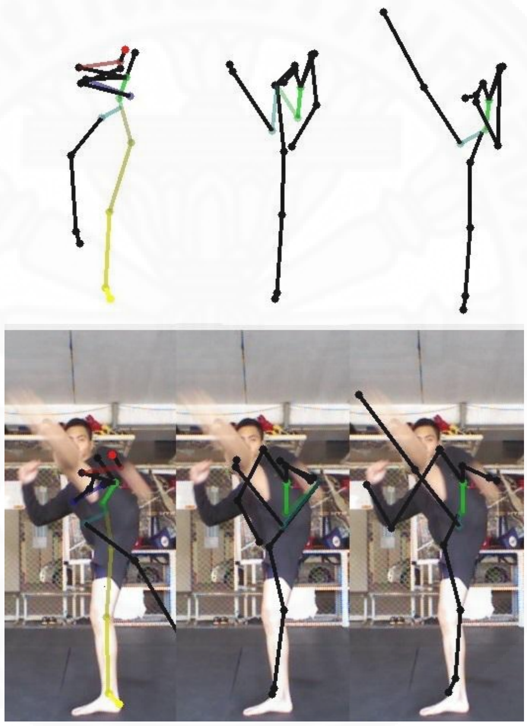

Figure 3.5: Skeleton data the video in Fig. 1. The first frame is the result from the Kinect Skeleton Tracking system alone. The second frame is the result after depth tracking. The third frame is the result from the depth tracking and boundary adjustment. The top frames are the side view of the bottom frames.

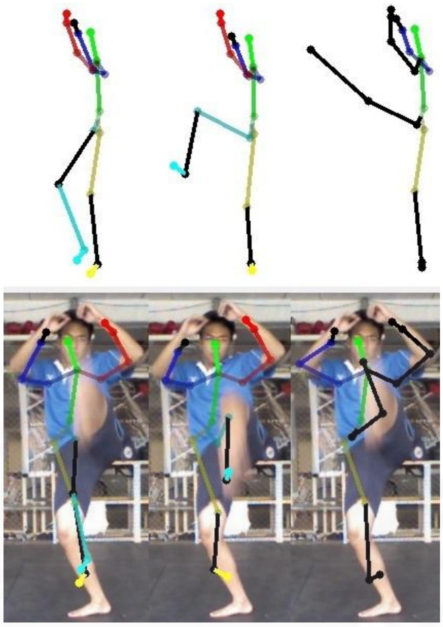

Figure 3.6: Skeleton data the video in Fig. 2. The first frame is the result from the Kinect Skeleton Tracking system alone. The second frame is the result after depth tracking. The third frame is the result from the depth tracking and boundary adjustment. The top frames are the side view of the bottom frames.

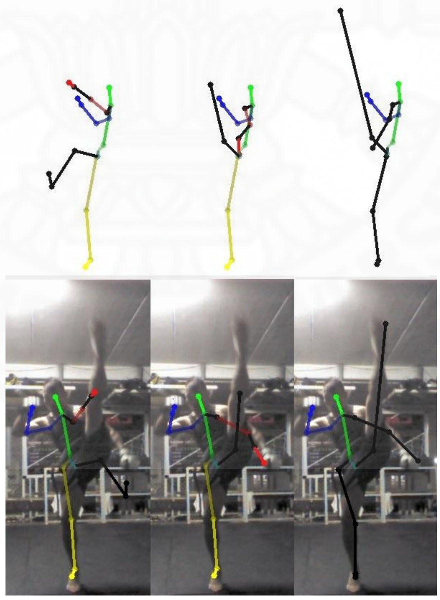

Figure 3.7: Skeleton data the video in Fig. 3. The first frame is the result from the Kinect Skeleton Tracking system alone. The second frame is the result after depth tracking. The third frame is the result from the depth tracking and boundary adjustment. The top frames are the side view of the bottom frames.

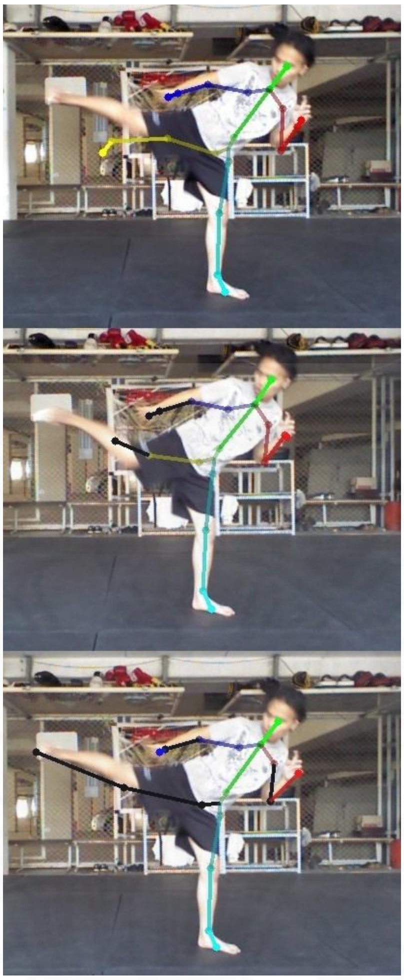

Figure 3.8: Skeleton data the video in Fig. 4. The top frame is the result from the Kinect Skeleton Tracking system alone. The middle frame is the result after depth tracking. The bottom frame is the result from the depth tracking and boundary adjustment.

Chapter 4
Maneuver Similarity Measurement

We propose the maneuver similarity measurement techniques that is a part of our *Instruction Support System f or M uay T hai*. The maneuver similarity measurement is a process that computes the closeness between new learner's maneuvers and the professional boxer's maneuvers in our database. This helps the new learner understand the correct movement of maneuvers and makes the new learner able to practice by himself or herself. We propose a number of techniques to measure similarities between maneuvers. The similarity measurement techniques are evaluated with the *k*-Nearest Neighbors algorithm. This chapter also shows the results and the analysis of the accuracy of the proposed similarity measurement techniques.

4.1 The Processes for Maneuver Similarity Measurement

We first explain the overall processes for measuring the similarity between two maneuvers. Figure 4.1 shows how two maneuvers captured by Microsoft Kinect are compared to yield a similarity value. First, The Kinect is used to capture *Sequences of F rames* from the professional boxer and new learner. Each frame is a collection of skeletal data captured by the Kinect. Second, we extract the necessary data by using *F eature Extraction* process that is explained in details in Section 4.2. Third, *Sequence of Postures* is obtained as the result of *Sequence of Frames* by the *F eature Extraction* process. Finally, we compute the similarity value between 2 *Sequences of Postures*. This will be explained the detail in Section 4.3.

4.2 Feature Extraction

A maneuver S is represented as a set of maneuvers of 20 joints S_{joint},

$$S = \begin{bmatrix} S_{Spine} \\ S_{Shoulder\ Center} \\ S_{Head} \\ \vdots \\ S_{Foot\ Right} \end{bmatrix} \tag{4.1}$$

Each S_{joint} is a sequence of hierarchical bone rotation matrices of a frame $R_{joint,i}$, where $i = 1, \ldots, N$ and N is the number of frames.

$$S_{joint} = (R^{(h)}_{joint,1}, R^{(h)}_{joint,2}, R^{(h)}_{joint,3}, \ \ldots \ , R^{(h)}_{joint,N}) \tag{4.2}$$

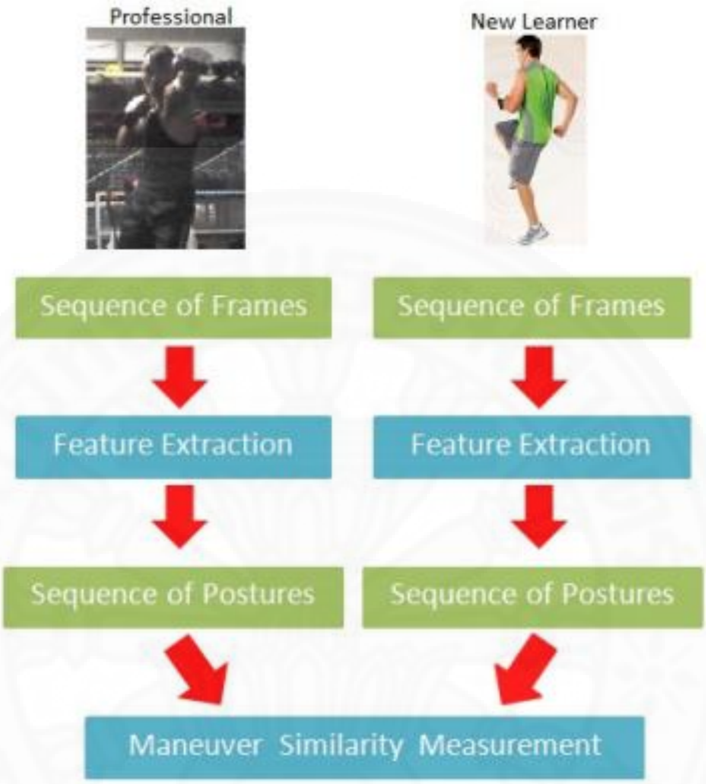

Figure 4.1: Maneuver Similarity Measurement Diagram

Each hierarchical bone rotation matrix is a 3x3 matrix;

$$R^{(h)}_{joint} = \begin{bmatrix} m_{11,joint} & m_{12,joint} & m_{13,joint} \\ m_{21,joint} & m_{22,joint} & m_{23,joint} \\ m_{31,joint} & m_{32,joint} & m_{33,joint} \end{bmatrix} \tag{4.3}$$

4.3 Maneuver Similarity Measurement

We propose techniques to measure similarities between two maneuvers based on three approaches i.e. Dynamic Time Warping (DTW), Sliding Window (SW), and Hidden Markov Models (HMM). Both DTW and SW are functions designed to measure similarities between two sequences of data. We can use them to directly measure similarities between a learner's maneuver and the professional boxer's maneuver. The similarity values can then be displayed to the learner as a feedback.

For HMM, we train an HMM for each type of maneuvers performed by the professional boxers. Then, a similarity of a learner's maneuver is computed based on the probability of the learner's maneuver given the HMM.

We propose the new modified of similarity measurement based on the DTW, SW and HMM. We want this new maneuver similarity measurement to be used in our *Instruction Support System for Muay Thai* which shows the maneuver similarity measurement in percentage.

4.3.1 Dynamic Time Warping(DTW)

Calculating a DTWbased similarity between 2 maneuvers requires calculating Euclidean distances between all combinations of postures from the 2 maneuvers. Since the Euclidean distance between 2 hierarchical bone rotation matrices which shown in Equation 4.3 is not in the range of [0,1]. Thus, it is difficult for the new learner to comprehend the similarity values. We then normalize the values so that they are in the range of [0, 1] to make it easier for the learner to use the similarity values at the feedback. In order to normalize the values, we know that the maximum value of Euclidean distance between 2 hierarchical bone rotationmatrices that is $\sqrt{12}$. This comes from the Euclidean distance between 2 opposite hierarchical bone rotation matrices. When the Euclidean distance between 2 hierarchical bone rotation matrices divided by $\sqrt{12}$, this number is in the range of [0,1]. For example, let A and B be the hierarchical bone rotation matrices that are in the opposite direction

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$B = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

Therefore, we can calculate the Euclidean distance between two hierarchical bone rotation matrices as shown in the following equation:

$$d(A, B) = \sqrt{(1-(-1))^2 + (1-(-1))^2 + (1-(-1))^2}$$
$$d(A, B) = \sqrt{12}$$

Therefore, the formula that converts the Euclidean distance to [0,1] is

$$d'(A,B) = \frac{d(A,B)}{\sqrt{12}} \tag{4.4}$$

Since DTW distance which in the form of Equation 2.12 increases when the length of data is longer, we average the value of distance by the number of steps in the path m. Therefore, we can define the maneuver similarity based on DTW as:

$$DTW'(S_{i,joint1}, S_{j,joint1}) = \frac{\gamma_{N_{i,joint1}, N_{j,joint1}}}{m} \tag{4.5}$$

Then, we define maneuver similarity measurement for a joint as

$$SIM(S_{i,joint1}, S_{j,joint1}) = (1 - DTW'(S_{i,joint1}, S_{j,joint1})) \tag{4.6}$$

Finally, we define maneuver similarity measurement for entire body as

$$SIM(S_i, S_j) = \frac{SIM(S_{i,joint1}, S_{j,joint1}) + \ldots + SIM(S_{i,joint20}, S_{j,joint20})}{20} \tag{4.7}$$

4.3.2 Sliding Window (SW)

The calculation in the Euclidean distance between 2 hierarchical bone rotation matrices is the similar manner as DTW, but the calculation in the path which minimizes the total distance is diferent due to Equation 2.20. Consequently, we can normalize Equation 2.20 as Equation 4.5 by the number of steps in the path m.

$$SW'(S_{i,joint1}, S_{j,joint1}) = \frac{min\{\gamma_{1,M,M}, \cdots, \gamma_{N-M+1,M,M}\}}{m} \tag{4.8}$$

Then, the maneuver similarity measurement for a joint is similar to Equation 4.6, when we define

$$SIM(S_{i,joint1}, S_{j,joint1}) = (1 - SW'(S_{i,joint1}, S_{j,joint1})) \tag{4.9}$$

Finally, we define maneuver similarity measurement for entire body as

$$SIM(S_i, S_j) = \frac{SIM(S_{i,joint1}, S_{j,joint1}) + \ldots + SIM(S_{i,joint20}, S_{j,joint20})}{20} \tag{4.10}$$

4.3.3 Hidden Markov Model (HMM)

In DTW and SW, the maneuver similarity measurement is measured directly from two sequences of data. But the HMM maneuver similarity measurement is measured by using training data created by many data. For example, we have 5 data of Lead Straight Punch maneuver. We want to know how similar between the unknown data and Lead Straight Punch maneuver. Then, 5 data of Lead Straight Punch maneuver will become the training data. After that, we find the similarity of unknown data by the training data. We use publicly available C# libraries for Hierarchical Clustering [10] and HMM [9]. In our program, Accord.net library is used to make training data and find the probability of unknown data by the training data. The probability means the similarity in our program. The training data must be in the form of an array in the function of Accord.net library. The length of the array is equal to the length of the total of frames of maneuver as in Equation 4.2. Each data of array is the hierarchical bone rotation as in Equation 4.2 which is clustered by hierarchical clustering. For example, let S1,joint1, S2,joint1, S3,joint1 be 3 maneuver data for a maneuver m in a joint joint1.

$$
\begin{aligned}
S_{1,joint1} &= (R^{(h)}_{1,joint1,1}, R^{(h)}_{1,joint1,2}, R^{(h)}_{1,joint1,3}, R^{(h)}_{1,joint1,4}) \\
S_{2,joint1} &= (R^{(h)}_{2,joint1,1}, R^{(h)}_{2,joint1,2}, R^{(h)}_{2,joint1,3}) \\
S_{3,joint1} &= (R^{(h)}_{3,joint1,1}, R^{(h)}_{3,joint1,2}, R^{(h)}_{3,joint1,3}, R^{(h)}_{3,joint1,4}, R^{(h)}_{3,joint1,5})
\end{aligned}
$$

If the training data has 4 clusters or observations, the 3 data will be clustered as

$$
\begin{aligned}
S'_{1,joint1} &= (1, 2, 3, 4) \\
S'_{2,joint1} &= (1, 2, 4) \\
S'_{3,joint1} &= (1, 2, 2, 3, 4)
\end{aligned}
$$

Then, they are ready to be in HMM process.

4.3.4 Weighted Separated Body Part (WSBP)

The maneuver similarity measurement based on DTW in Equation 4.7 or by SW in Equation 4.10 or by HMM are the recognition as the entire body. However, most of Muay Thai maneuvers focus only on some parts of the body. For example, Lead Straight Punch mainly requires only the arm movement. We can use maneuver similarity measurement by DTW, SW or HMM in only some parts of body. The body parts are Body, Left Arm, Right Arm, Left Leg and Right Leg.

- Body: Head, Shoulder Center, Spine, Hip Center
- Left Arm: Shoulder Left, Elbow Left, Wrist Left, Hand Left
- Right Arm: Shoulder Right, Elbow Right, Wrist Right, Hand Right
- Left Leg: Hip Left, Knee Left, Ankle Left, Foot Left
- Right Leg: Hip Right, Knee Right, Ankle Right, Foot Right

For example,

$$SIM(S_{i,Body}, S_{j,Body}) = \frac{SIM(S_{i,Head}, S_{j,Head})}{4} + \frac{SIM(S_{i,Shoulder\ Center}, S_{j,Shoulder\ Center})}{4} + \frac{SIM(S_{i,Spine}, S_{j,Spine})}{4} + \frac{SIM(S_{i,Hip\ Center}, S_{j,Hip\ Center})}{4}$$

After we have the result of maneuver similarity measurement by DTW, SW or

$$SIM(S_i, S_j) = \frac{Weight \times SIM(S_{i,D}, S_{j,D}) + \sum SIM(S_{i,O}, S_{j,O})}{Weight + 4} \qquad (4.11)$$

where $S_{i,D}$ represents dominant part and $S_{i,O}$ represents other part

$$S'_{1,joint1} = (1,2,3,4)$$
$$S'_{2,joint1} = (1,2,4)$$
$$S'_{3,joint1} = (1,2,2,3,4)$$

Example:

If the maneuver is Lead Straight Punch and Weight = 4, the result depends mostly in Right Arm.

$$SIM(S_i, S_j) = \frac{4 \times SIM(S_{i,Right\ Arm}, S_{j,Right\ Arm})}{8} + \frac{SIM(S_{i,Body}, S_{j,Body})}{8} + \frac{SIM(S_{i,Left\ Arm}, S_{j,Left\ Arm})}{8} + \frac{SIM(S_{i,Left\ Leg}, S_{j,Left\ Leg})}{8} + \frac{SIM(S_{i,Right\ Leg}, S_{j,Right\ Leg})}{8}$$

4.4 Experiment Setup

We conduct a number of experiments to evaluate the performance of the proposed three similarity measures. We measure the performance based on the predictive per- formance of the similarity measures using the k-Nearest Neighbors algorithm. We also compare the performance of the similarity measures using both the entire body of the players and the weighted separated body parts. The experiments are conducted using the 5-fold cross validation technique explained in Section 4.3.1, 4.3.2 and 4.3.3. First, we set up 5 folds. We randomly separate data in 5 folds as shown in Figure 4.2. Each fold contains all Maneuver. We have to find every fold accuracy.

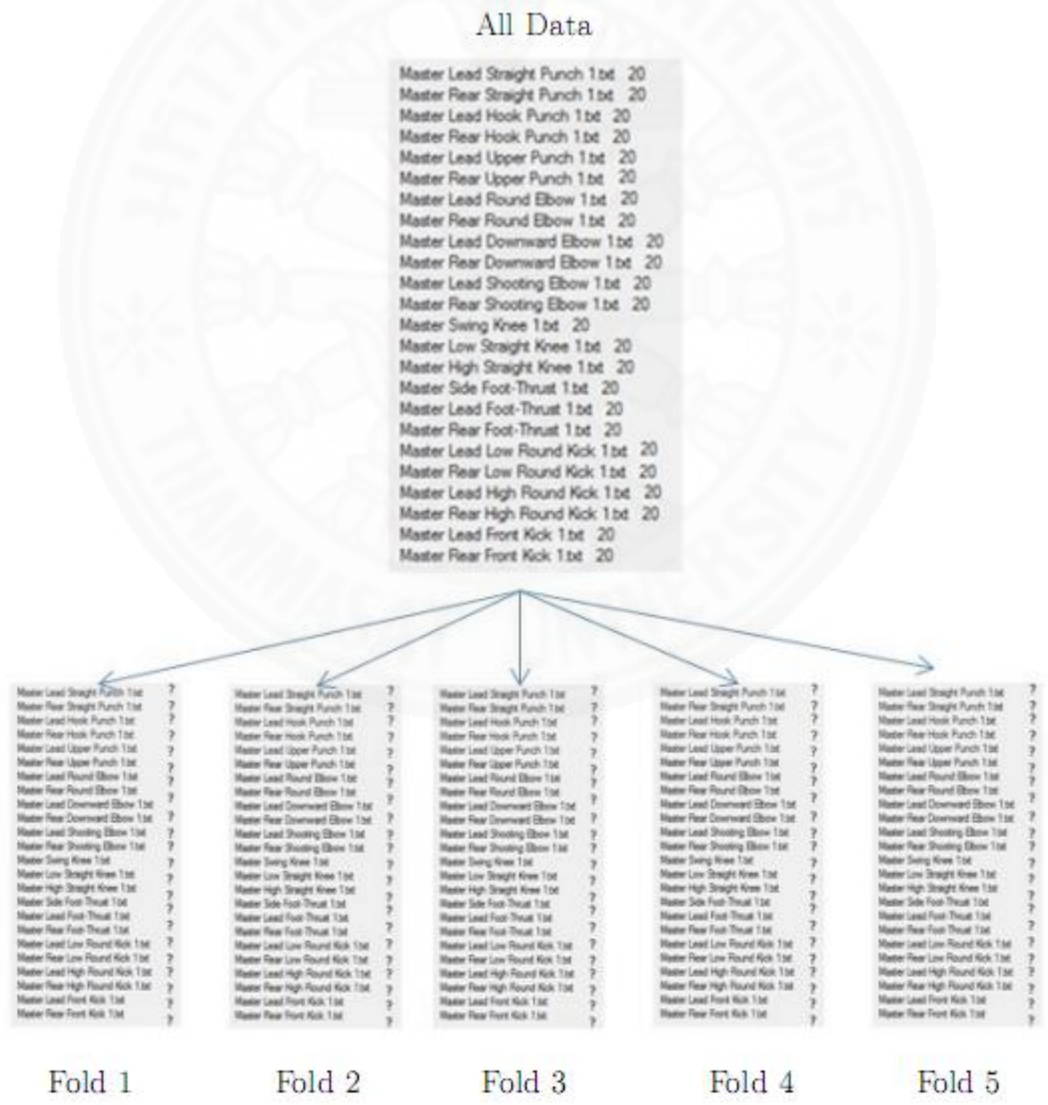

Figure 4.2: Experiment Setup Step 1

To find the accuracy in each fold, we have to find maneuver similarity measurement (*SIM*) of every data accuracy in the fold as shown in Figure 4.3. In DTW or SW, we find *SIM* by DTW explained in Section 4.3.1 or SW as Section 4.3.2. and *k*-Nearest Neighbors to identify the accuracy of each data in the fold. In each data, we have to find top 1, 3, 5, 7 and 9 for 1-NN, 3-NN, 5-NN, 7-NN and 9-NN. If top 1, 3, 5, 7, 9 (1-NN, 3-NN, …, 9-NN) of maneuver similarity measurement is(are) the same maneuver more than 1, 2, 3, 4, 5, we set it to 1. Otherwise, we set it 0. In HMM, training data of each maneuver finds the accuracy of each data as explained in Section 4.3.3. If the highest accuray comes from the same maneuver, we set it 1. Otherwise, we set it 0.

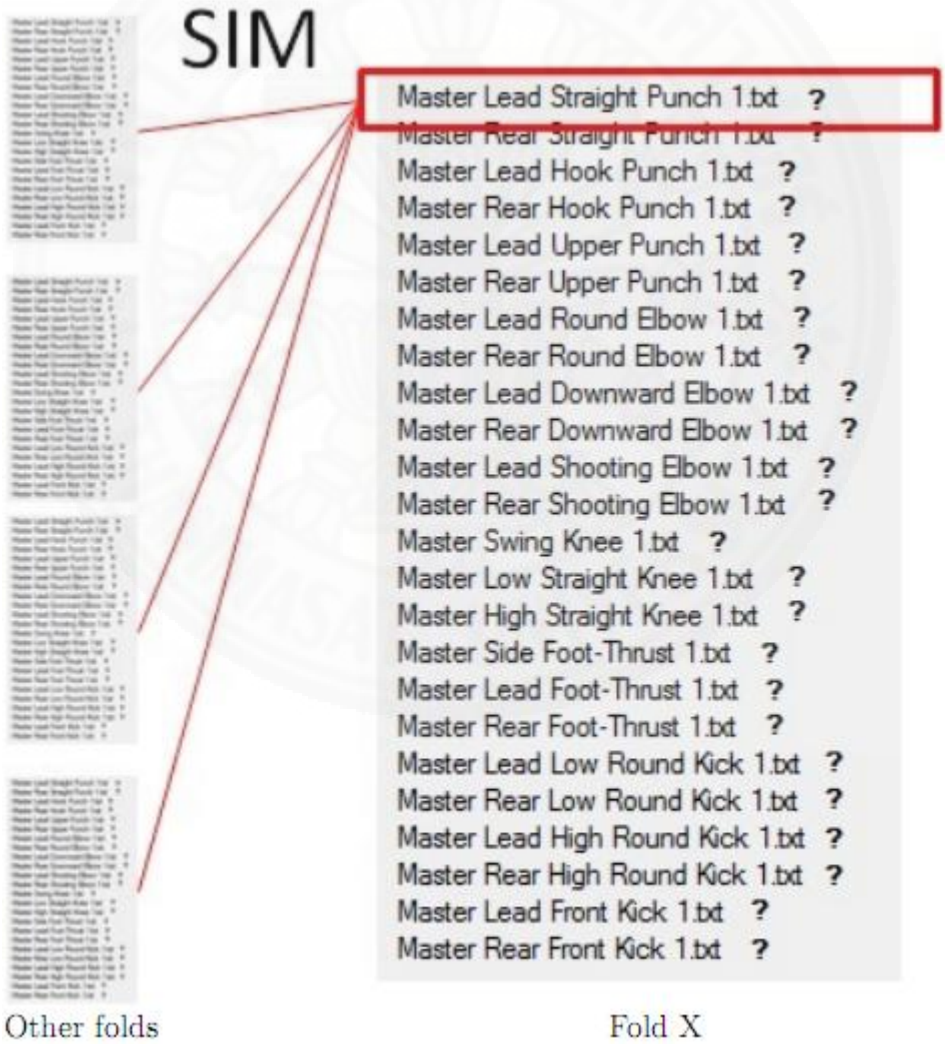

Figure 4.3: Experiment Setup Step 2

We average percentage of 1 in each maneuver to make accuracy in each maneuver as shown in Figure 4.4.

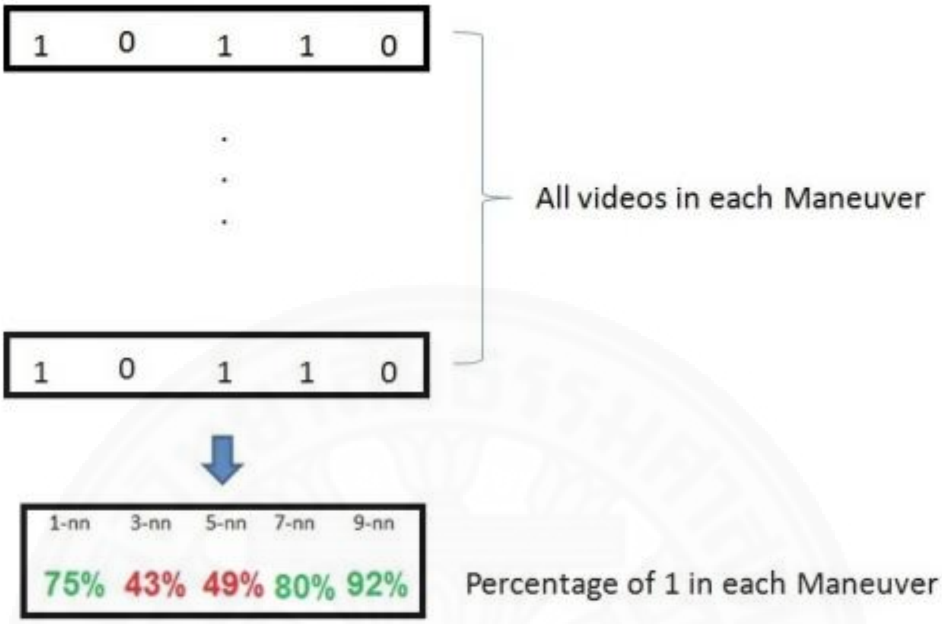

Figure 4.4: Experiment Setup Step 3

Then, we average percentage of all maneuvers in each fold as shown in Figure 4.5.

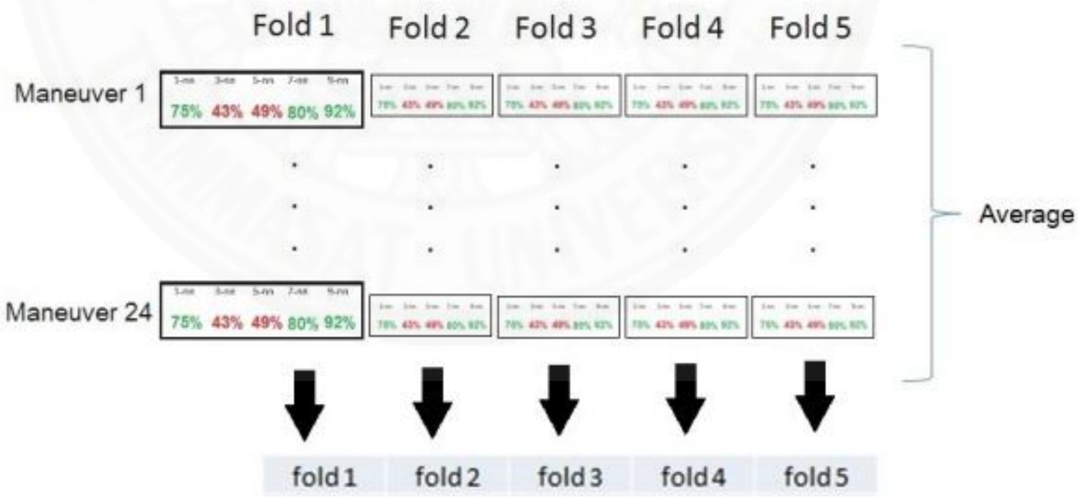

Figure 4.5: Experiment Setup Step 4

Finally, we average all accuracy values in each fold as shown in Figure 4.6.

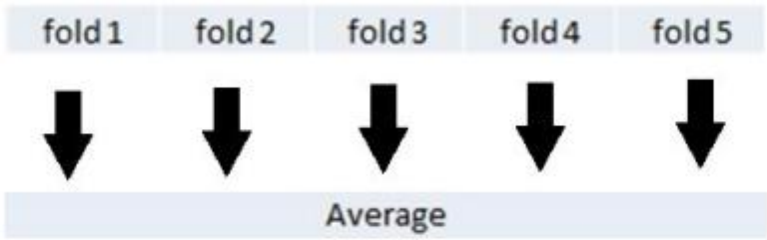

Figure 4.6: Experiment Setup Step 5

4.5 HMM Fine-tuning

In HMM, we need to find the best observations and number of states that make the highest percentage of maneuvers recognition. These graph (Figure 4.7, 4.9, 4.8) and 4.10) run with 100 states in each observation. We run a total of 1000 observations. In HMM Kinect, best parameters are observation at 569, states at 8 and it make 78.67%. In HMM Kinect+, best parameters are observation at 580, states at 10 and it make 88.95% In WSBP HMM Kinect, best parameters are observation at 575, states at 5, weight = 5 and it makes 69.04%. In WSBP HMM Kinect+, best parameters are observation at 560, states at 4, weight = 4 and it makes 78.06%.

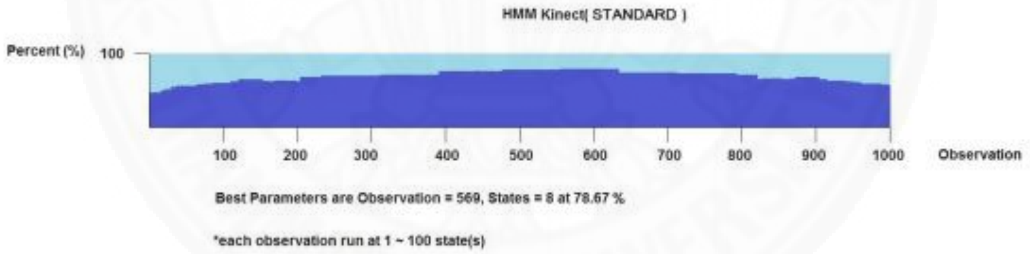

Figure 4.7: HMM Kinect Fine-tuning

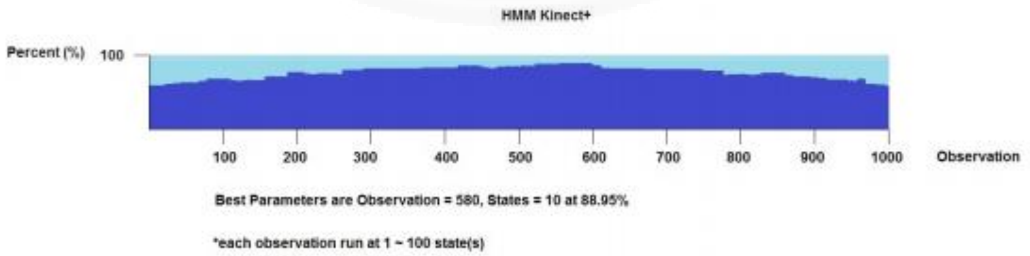

Figure 4.8: HMM Kinect+ Fine-tuning

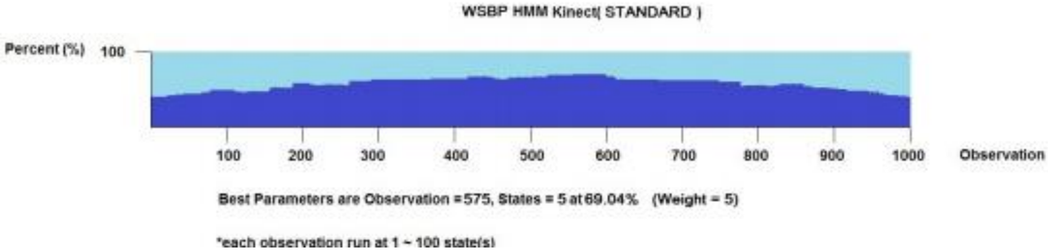

Figure 4.9: WSBP HMM Kinect Fine-tuning

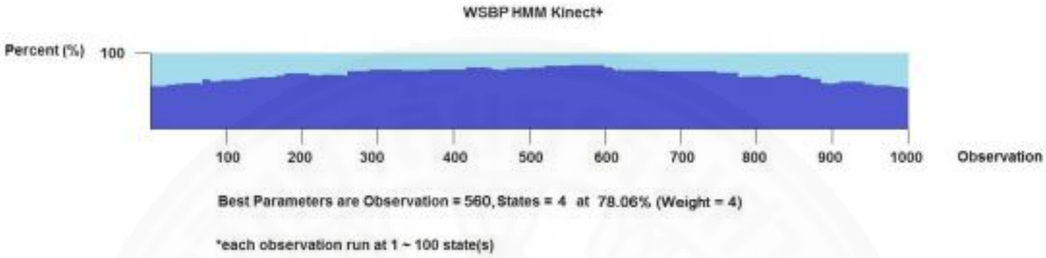

Figure 4.10: WSBP HMM Kinect+ Fine-tuning

4.6 Experiment Results

Table 4.1 is the results of our maneuver prediction accuracy. Standard Kinect is the original Kinect recognition system. Kinect+ is the original Kinect recognition system with our proposed algorithms in Chapter 3. We apply *k*-Nearest Neighbors only to DTW and SW because *k*-Nearest Neighbors is unable to apply to HMM. DTW performs slightly better than SW. However, HMM with Kinect+ yields the best result which is 88.95%. Moreover, the result shows that the percentage in maneuvers recognition improve dramatically by the algorithm in Chapter 3. When we apply WSBP, the result accuracy decreases.

Table 4.1: Best average accuracy of maneuver predictions based on different techniques with 5-fold cross validation. The dataset consists of 24 maneuvers from 480 videos.

Algorithm	Average
DTW: Standard Kinect (1-NN)	65.86
DTW: Kinect+ (1-NN)	70.62
SW: Standard Kinect(1-NN)	57.51
SW: Kinect+ (1-NN)	64.69
HMM: Standard Kinect	78.67
HMM: Kinect+	**88.95**
WSBP DTW: Standard Kinect (1-NN, Weight = 4)	57.80
WSBP DTW: Kinect+ (1-NN, Weight = 3)	61.97
WSBP SW: Standard Kinect(1-NN, Weight = 3)	50.47
WSBP SW: Kinect+ (1-NN, Weight = 5)	56.77
WSBP HMM: Standard Kinect (Weight = 5)	69.04
WSBP HMM: Kinect+ (Weight = 4)	78.06

Table 4.2: Accuracy of maneuver predictions based on DTWand k-NN techniques with 5-fold cross validation. The dataset consists of 24 maneuvers from 480 videos. The features are extracted by Standard Kinect feature extraction.

	fold 1	fold 2	fold 3	fold 4	fold 5	Average
1-NN	69.41	68.97	63.10	66.67	61.18	65.86
3-NN	59.77	60.47	63.10	55.95	59.52	59.76
5-NN	55.95	51.19	58.43	50.60	47.06	52.65
7-NN	42.86	38.10	32.94	41.18	37.93	38.60
9-NN	17.86	9.52	8.33	16.09	10.47	17.16

Table 4.3: Accuracy of maneuver predictions based on DTWand k-NN techniques with 5-fold cross validation. The dataset consists of 24 maneuvers from 480 videos. The features are extracted by Kinect+ feature extraction.

	fold 1	fold 2	fold 3	fold 4	fold 5	Average
1-NN	70.97	69.15	74.73	66.67	71.58	70.62
3-NN	62.77	67.39	69.23	63.16	65.96	65.70
5-NN	48.39	59.14	54.84	47.31	57.45	53.42
7-NN	37.63	38.04	36.17	43.62	45.16	40.13
9-NN	20.65	18.28	18.09	18.09	16.13	18.25

Table 4.4: Accuracy of maneuver predictions based on SW and k-NN techniques with 5-fold cross validation. The dataset consists of 24 maneuvers from 480 videos. The features are extracted by Standard Kinect feature extraction.

	fold 1	fold 2	fold 3	fold 4	fold 5	Average
1-NN	58.06	57.45	57.45	58.70	55.91	57.51
3-NN	53.19	52.75	53.76	52.63	50.54	52.57
5-NN	41.94	41.49	43.01	44.09	43.01	42.71
7-NN	27.96	27.17	34.41	24.47	27.66	28.33
9-NN	12.90	17.39	12.77	12.90	11.70	13.53

Table 4.5: Accuracy of maneuver predictions based on SW and k-NN techniques with 5-fold cross validation. The dataset consists of 24 maneuvers from 480 videos. The features are extracted by Kinect+ feature extraction.

	fold 1	fold 2	fold 3	fold 4	fold 5	Average
1-NN	68.97	65.85	63.10	63.33	62.20	64.69
3-NN	58.14	50.00	56.47	54.12	55.29	54.80
5-NN	48.81	45.88	44.83	39.02	42.53	44.21
7-NN	32.14	30.00	26.83	29.89	37.80	31.33
9-NN	12.05	11.63	10.84	8.14	10.34	10.60

Table 4.6: Accuracy of maneuver predictions based on HMM with 5-fold cross validation. The dataset consists of 24 maneuvers from 480 videos. The features are extracted by Standard Kinect feature extraction.

fold 1	fold 2	fold 3	fold 4	fold 5	Average
79.58	78.48	77.34	78.82	78.37	78.67

Table 4.7: Accuracy of maneuver predictions based on HMM with 5-fold cross validation. The dataset consists of 24 maneuvers from 480 videos. The features are extracted by Kinect+ feature extraction.

fold 1	fold 2	fold 3	fold 4	fold 5	Average
89.48	88.48	86.34	87.82	88.37	88.95

Table 4.8: Accuracy of maneuver predictions based on DTW and k-NN techniques with 5-fold cross validation and WSBP. The dataset consists of 24 maneuvers from 480 videos. The features are extracted by Standard Kinect feature extraction.

		fold 1	fold 2	fold 3	fold 4	fold 5	Average
Weight = 1	1-NN	58.81	58.43	53.46	56.48	51.83	55.80
	3-NN	50.64	51.23	53.46	47.41	50.43	50.63
	5-NN	47.41	43.37	49.50	42.87	39.87	44.60
	7-NN	36.31	32.28	27.91	34.89	32.14	32.70
	9-NN	15.13	8.07	7.06	13.63	8.87	14.54
Weight = 2	1-NN	59.89	59.50	54.44	57.52	52.78	56.83
	3-NN	51.57	52.17	54.44	48.28	51.36	51.56
	5-NN	48.28	44.17	50.41	43.66	40.60	45.42
	7-NN	36.98	32.87	28.42	35.53	32.73	33.30
	9-NN	15.41	8.22	7.19	13.88	9.03	14.81
Weight = 3	1-NN	60.58	60.19	55.07	58.18	53.39	57.48
	3-NN	52.16	52.77	55.07	48.83	51.95	52.16
	5-NN	48.83	44.68	50.99	44.16	41.07	45.95
	7-NN	37.40	33.25	28.75	35.94	33.10	33.69
	9-NN	15.58	8.31	7.27	14.04	9.13	14.98
Weight = 4	1-NN	60.91	60.52	55.37	58.51	53.69	57.80
	3-NN	52.45	53.06	55.37	49.10	52.24	52.45
	5-NN	49.10	44.92	51.27	44.41	41.30	46.20
	7-NN	37.61	33.43	28.91	36.14	33.29	33.87
	9-NN	15.67	8.36	7.31	14.12	9.18	15.06
Weight = 5	1-NN	60.24	59.85	54.76	57.85	53.09	57.16
	3-NN	51.87	52.47	54.76	48.56	51.66	51.86
	5-NN	48.56	44.42	50.70	43.91	40.84	45.69
	7-NN	37.19	33.06	28.59	35.73	32.92	33.50
	9-NN	15.50	8.26	7.23	13.96	9.08	14.89
Weight = 6	1-NN	59.54	59.15	54.12	57.18	52.47	56.49
	3-NN	51.27	51.86	54.12	47.99	51.05	51.26
	5-NN	47.99	43.91	50.11	43.40	40.36	45.16
	7-NN	36.76	32.67	28.25	35.32	32.53	33.11
	9-NN	15.32	8.17	7.15	13.80	8.98	14.72
Weight = 7	1-NN	59.18	58.79	53.79	56.83	52.15	56.15
	3-NN	50.96	51.55	53.79	47.70	50.75	50.95
	5-NN	47.70	43.64	49.81	43.14	40.12	44.88
	7-NN	36.54	32.48	28.08	35.10	32.34	32.91
	9-NN	15.22	8.12	7.10	13.72	8.92	14.63
Weight = 8	1-NN	58.44	58.06	53.12	56.13	51.50	55.45
	3-NN	50.32	50.91	53.12	47.11	50.11	50.31
	5-NN	47.11	43.10	49.19	42.60	39.62	44.32
	7-NN	36.08	32.07	27.73	34.67	31.93	32.50
	9-NN	15.03	8.02	7.02	13.55	8.81	14.45
Weight = 9	1-NN	58.06	57.69	52.78	55.76	51.17	55.09
	3-NN	50.00	50.58	52.78	46.80	49.79	49.99
	5-NN	46.80	42.82	48.87	42.33	39.36	44.04
	7-NN	35.85	31.87	27.55	34.44	31.73	32.29
	9-NN	14.94	7.97	6.97	13.46	8.75	14.36
Weight = 10	1-NN	57.68	57.31	52.43	55.40	50.83	54.73
	3-NN	49.67	50.24	52.43	46.49	49.46	49.66
	5-NN	46.49	42.54	48.55	42.05	39.10	43.75
	7-NN	35.61	31.65	27.37	34.22	31.52	32.07
	9-NN	14.84	7.91	6.92	13.37	8.70	14.26

Table 4.9: Accuracy of maneuver predictions based on DTW and k-NN techniques with 5-fold cross validation and WSBP. The dataset consists of 24 maneuvers from 480 videos. The features are extracted by Kinect+ feature extraction.

		fold 1	fold 2	fold 3	fold 4	fold 5	Average
Weight = 1	1-NN	61.23	59.66	64.47	57.52	61.76	60.93
	3-NN	54.15	58.15	59.73	54.49	56.91	56.69
	5-NN	41.75	51.03	47.32	40.82	49.57	46.10
	7-NN	32.47	32.82	31.21	37.63	38.97	34.62
	9-NN	17.82	15.77	15.60	15.60	13.92	15.74
Weight = 2	1-NN	61.59	60.01	64.85	57.85	62.12	61.28
	3-NN	54.47	58.48	60.08	54.81	57.24	57.02
	5-NN	41.99	51.32	47.59	41.06	49.85	46.36
	7-NN	32.66	33.01	31.39	37.85	39.19	34.82
	9-NN	17.92	15.86	15.69	15.69	14.00	15.83
Weight = 3	1-NN	62.28	60.68	65.58	58.51	62.82	61.97
	3-NN	55.08	59.14	60.76	55.43	57.88	57.66
	5-NN	42.46	51.90	48.13	41.52	50.41	46.88
	7-NN	33.03	33.39	31.74	38.28	39.63	35.21
	9-NN	18.12	16.04	15.87	15.87	14.15	16.01
Weight = 4	1-NN	61.94	60.35	65.22	58.18	62.47	61.63
	3-NN	54.78	58.82	60.42	55.12	57.56	57.34
	5-NN	42.23	51.61	47.86	41.29	50.14	46.63
	7-NN	32.85	33.20	31.57	38.07	39.41	35.02
	9-NN	18.02	15.95	15.78	15.78	14.08	15.92
Weight = 5	1-NN	60.87	59.31	64.09	57.18	61.39	60.57
	3-NN	53.83	57.80	59.38	54.17	56.57	56.35
	5-NN	41.50	50.72	47.04	40.58	49.27	45.82
	7-NN	32.28	32.63	31.02	37.41	38.74	34.42
	9-NN	17.71	15.68	15.51	15.51	13.83	15.65
Weight = 6	1-NN	60.50	58.95	63.71	56.83	61.02	60.20
	3-NN	53.51	57.45	59.02	53.84	56.23	56.01
	5-NN	41.25	50.42	46.75	40.33	48.97	45.55
	7-NN	32.08	32.43	30.84	37.18	38.50	34.21
	9-NN	17.61	15.58	15.42	15.42	13.75	15.56
Weight = 7	1-NN	60.13	58.59	63.31	56.48	60.65	59.83
	3-NN	53.18	57.10	58.66	53.51	55.88	55.67
	5-NN	41.00	50.11	46.46	40.09	48.67	45.26
	7-NN	31.89	32.23	30.65	36.95	38.26	34.00
	9-NN	17.50	15.49	15.32	15.32	13.67	15.46
Weight = 8	1-NN	59.75	58.22	62.91	56.13	60.26	59.45
	3-NN	52.84	56.74	58.29	53.17	55.53	55.31
	5-NN	40.74	49.79	46.17	39.83	48.36	44.98
	7-NN	31.68	32.03	30.45	36.72	38.02	33.78
	9-NN	17.39	15.39	15.23	15.23	13.58	15.36
Weight = 9	1-NN	59.36	57.84	62.50	55.76	59.87	59.07
	3-NN	52.50	56.37	57.91	52.83	55.17	54.96
	5-NN	40.47	49.47	45.87	39.57	48.05	44.69
	7-NN	31.48	31.82	30.26	36.48	37.78	33.56
	9-NN	17.27	15.29	15.13	15.13	13.49	15.26
Weight = 10	1-NN	58.97	57.46	62.09	55.40	59.48	58.68
	3-NN	52.15	56.00	57.53	52.48	54.81	54.59
	5-NN	40.21	49.14	45.57	39.31	47.73	44.39
	7-NN	31.27	31.61	30.06	36.24	37.53	33.34
	9-NN	17.16	15.19	15.03	15.03	13.40	15.16

Table 4.10: Accuracy of maneuver predictions based on SW and k-NN techniques with 5-fold cross validation and WSBP. The dataset consists of 24 maneuvers from 480 videos. The features are extracted by Standard Kinect feature extraction.

		fold 1	fold 2	fold 3	fold 4	fold 5	Average
Weight = 1	1-NN	50.10	49.57	49.57	50.64	48.24	49.62
	3-NN	45.89	45.51	46.39	45.41	43.60	45.36
	5-NN	36.18	35.80	37.11	38.04	37.11	36.85
	7-NN	24.12	23.45	29.69	21.11	23.86	24.45
	9-NN	11.13	15.01	11.01	11.13	10.10	11.68
Weight = 2	1-NN	50.39	49.85	49.85	50.94	48.52	49.91
	3-NN	46.16	45.78	46.66	45.67	43.86	45.62
	5-NN	36.39	36.01	37.33	38.26	37.33	37.06
	7-NN	24.26	23.58	29.86	21.23	24.00	24.59
	9-NN	11.20	15.09	11.08	11.20	10.16	11.74
Weight = 3	1-NN	50.96	50.41	50.41	51.51	49.07	50.47
	3-NN	46.68	46.29	47.18	46.19	44.35	46.14
	5-NN	36.80	36.41	37.75	38.69	37.75	37.48
	7-NN	24.53	23.85	30.20	21.47	24.27	24.86
	9-NN	11.32	15.26	11.20	11.32	10.27	11.88
Weight = 4	1-NN	50.68	50.14	50.14	51.23	48.80	50.19
	3-NN	46.42	46.03	46.92	45.93	44.11	45.88
	5-NN	36.60	36.21	37.54	38.48	37.54	37.27
	7-NN	24.40	23.72	30.03	21.35	24.14	24.73
	9-NN	11.26	15.18	11.14	11.26	10.21	11.81
Weight = 5	1-NN	49.80	49.27	49.27	50.34	47.96	49.33
	3-NN	45.62	45.24	46.11	45.14	43.35	45.09
	5-NN	35.97	35.59	36.89	37.81	36.89	36.63
	7-NN	23.98	23.31	29.51	20.99	23.72	24.30
	9-NN	11.07	14.92	10.95	11.07	10.04	11.61
Weight = 6	1-NN	49.50	48.97	48.97	50.04	47.67	49.03
	3-NN	45.35	44.97	45.83	44.87	43.08	44.82
	5-NN	35.75	35.37	36.67	37.58	36.67	36.41
	7-NN	23.83	23.17	29.33	20.86	23.58	24.15
	9-NN	11.00	14.83	10.88	11.00	9.98	11.54
Weight = 7	1-NN	49.20	48.67	48.67	49.73	47.37	48.73
	3-NN	45.07	44.69	45.55	44.59	42.82	44.54
	5-NN	35.53	35.15	36.44	37.35	36.44	36.18
	7-NN	23.69	23.02	29.15	20.73	23.43	24.01
	9-NN	10.93	14.73	10.82	10.93	9.91	11.47
Weight = 8	1-NN	48.88	48.36	48.36	49.42	47.07	48.42
	3-NN	44.78	44.41	45.26	44.31	42.55	44.26
	5-NN	35.31	34.93	36.21	37.12	36.21	35.95
	7-NN	23.54	22.88	28.97	20.60	23.29	23.85
	9-NN	10.86	14.64	10.75	10.86	9.85	11.39
Weight = 9	1-NN	48.57	48.05	48.05	49.10	46.77	48.11
	3-NN	44.49	44.12	44.97	44.02	42.27	43.98
	5-NN	35.08	34.70	35.98	36.88	35.98	35.72
	7-NN	23.38	22.73	28.78	20.47	23.14	23.70
	9-NN	10.79	14.55	10.68	10.79	9.79	11.32
Weight = 10	1-NN	48.25	47.73	47.73	48.77	46.46	47.79
	3-NN	44.20	43.83	44.67	43.73	41.99	43.69
	5-NN	34.85	34.48	35.74	36.63	35.74	35.49
	7-NN	23.23	22.58	28.59	20.33	22.98	23.54
	9-NN	10.72	14.45	10.61	10.72	9.72	11.25

Table 4.11: Accuracy of maneuver predictions based on SW and k-NN techniques with 5-fold cross validation and WSBP. The dataset consists of 24 maneuvers from 480 videos. The features are extracted by Kinect+ feature extraction.

		fold 1	fold 2	fold 3	fold 4	fold 5	Average
Weight = 1	1-NN	58.06	55.44	53.12	53.32	52.36	54.46
	3-NN	48.95	42.10	47.54	45.56	46.55	46.14
	5-NN	41.09	38.63	37.74	32.85	35.80	37.22
	7-NN	27.06	25.26	22.59	25.16	31.83	26.38
	9-NN	10.14	9.79	9.13	6.85	8.71	8.92
Weight = 2	1-NN	58.79	56.14	53.79	53.99	53.02	55.15
	3-NN	49.57	42.63	48.14	46.14	47.14	46.72
	5-NN	41.61	39.12	38.22	33.27	36.26	37.69
	7-NN	27.40	25.58	22.87	25.48	32.23	26.71
	9-NN	10.27	9.91	9.24	6.94	8.82	9.04
Weight = 3	1-NN	59.50	56.82	54.44	54.64	53.66	55.81
	3-NN	50.16	43.14	48.72	46.69	47.71	47.29
	5-NN	42.11	39.59	38.68	33.67	36.69	38.15
	7-NN	27.73	25.88	23.15	25.79	32.62	27.03
	9-NN	10.40	10.03	9.36	7.02	8.93	9.15
Weight = 4	1-NN	59.85	57.15	54.76	54.96	53.97	56.14
	3-NN	50.45	43.39	49.01	46.96	47.99	47.56
	5-NN	42.36	39.82	38.90	33.87	36.91	38.37
	7-NN	27.89	26.03	23.28	25.93	32.81	27.19
	9-NN	10.46	10.09	9.41	7.06	8.98	9.20
Weight = 5	1-NN	60.52	57.79	55.37	55.58	54.58	56.77
	3-NN	51.02	43.88	49.56	47.49	48.53	48.10
	5-NN	42.83	40.27	39.34	34.25	37.32	38.80
	7-NN	28.21	26.33	23.54	26.23	33.18	27.50
	9-NN	10.57	10.20	9.52	7.14	9.08	9.30
Weight = 6	1-NN	60.19	57.47	55.07	55.27	54.28	56.46
	3-NN	50.74	43.64	49.28	47.23	48.26	47.83
	5-NN	42.60	40.04	39.12	34.06	37.12	38.59
	7-NN	28.05	26.18	23.42	26.08	32.99	27.35
	9-NN	10.51	10.15	9.46	7.10	9.03	9.25
Weight = 7	1-NN	59.15	56.48	54.12	54.32	53.35	55.48
	3-NN	49.87	42.89	48.44	46.42	47.43	47.01
	5-NN	41.86	39.35	38.45	33.47	36.48	37.92
	7-NN	27.57	25.73	23.01	25.63	32.43	26.87
	9-NN	10.33	9.97	9.30	6.98	8.87	9.09
Weight = 8	1-NN	58.43	55.79	53.46	53.66	52.70	54.81
	3-NN	49.26	42.36	47.84	45.85	46.85	46.43
	5-NN	41.35	38.87	37.98	33.06	36.03	37.46
	7-NN	27.23	25.42	22.73	25.32	32.03	26.55
	9-NN	10.21	9.85	9.19	6.90	8.76	8.98
Weight = 9	1-NN	57.69	55.08	52.78	52.98	52.02	54.11
	3-NN	48.63	41.82	47.24	45.27	46.25	45.84
	5-NN	40.83	38.38	37.50	32.64	35.57	36.98
	7-NN	26.89	25.09	22.44	25.00	31.62	26.21
	9-NN	10.08	9.73	9.07	6.81	8.65	8.87
Weight = 10	1-NN	57.31	54.72	52.43	52.63	51.68	53.75
	3-NN	48.31	41.55	46.92	44.97	45.95	45.54
	5-NN	40.56	38.13	37.25	32.43	35.34	36.74
	7-NN	26.71	24.93	22.29	24.83	31.41	26.04
	9-NN	10.01	9.66	9.01	6.76	8.60	8.81

Table 4.12: Accuracy of maneuver predictions based on HMM with 5-fold cross validation and WSBP. The dataset consists of 24 maneuvers from 480 videos. The features are extracted by Standard Kinect feature extraction.

	fold 1	fold 2	fold 3	fold 4	fold 5	Average
Weight = 1	67.00	66.07	65.11	66.36	65.98	66.23
Weight = 2	67.42	66.49	65.53	66.78	66.40	66.66
Weight = 3	68.66	67.71	66.73	68.00	67.62	67.88
Weight = 4	69.45	68.49	67.50	68.79	68.40	68.66
Weight = 5	69.84	68.87	67.87	69.17	68.77	69.04
Weight = 6	69.06	68.10	67.12	68.40	68.01	68.27
Weight = 7	68.25	67.31	66.34	67.60	67.22	67.48
Weight = 8	67.84	66.90	65.93	67.19	66.81	67.07
Weight = 9	66.56	65.64	64.69	65.93	65.55	65.81
Weight = 10	66.12	65.21	64.27	65.49	65.12	65.37

Table 4.13: Accuracy of maneuver predictions based on HMM with 5-fold cross validation and WSBP. The dataset consists of 24 maneuvers from 480 videos. The features are extracted by Kinect+ feature extraction.

	fold 1	fold 2	fold 3	fold 4	fold 5	Average
Weight = 1	75.81	74.96	73.15	74.40	74.87	75.36
Weight = 2	76.75	75.89	74.05	75.32	75.79	76.29
Weight = 3	77.65	76.78	74.93	76.21	76.69	77.19
Weight = 4	78.52	77.65	75.77	77.07	77.55	78.06
Weight = 5	78.09	77.22	75.35	76.64	77.12	77.63
Weight = 6	77.20	76.34	74.50	75.77	76.24	76.75
Weight = 7	76.28	75.43	73.61	74.87	75.34	75.83
Weight = 8	75.33	74.49	72.69	73.93	74.40	74.89
Weight = 9	74.85	74.01	72.22	73.46	73.92	74.40
Weight = 10	74.35	73.52	71.74	72.97	73.43	73.91

4.7 Result Analysis

Figure 4.14 - 4.37, we sum all result of the database maneuvers accuracy in 1, 3, 5, 7 and 9-NN nearest neighbors in DTW and SW. The HMM uses training data to recognize maneuver. Table 4.14, 4.26, 4.16, 4.28, 4.18, 4.30, 4.20, 4.32, 4.22, 4.34 , 4.24 and 4.36 show the result of database maneuvers accuracy. The number in a diagonal line is the number of correct maneuver recognitions. Table 4.15, 4.27, 4.17, 4.29, 4.19, 4.31, 4.21, 4.33, 4.23, 4.35, 4.25 and 4.37 show the percentage of correct manuever recogntion of Table 4.14, 4.26, 4.16, 4.28, 4.18, 4.30, 4.20, 4.32, 4.22, 4.34 , 4.24 and 4.36 in upper and lower body part.

The data show that our proposed algorithms in Chapter 3 makes maneuver recognition improvement significantly in leg maneuvers.

4.7.1 DTW: Standard Kinect

Table 4.15 shows that the percent accuracy of upper body maneuvers (71%) is higher than the percent accuracy of lower body maneuvers (60%). As shown in Table 4.14, Rear Hook Punch (RHP), Lead Upper Punch (LUP), Rear Upper Punch (RUP), Rear Round Elbow (RRE), Lead Downward Elbow (LDE), Rear Downward Elbow (RDE), Lead Shooting Elbow (LSE) and Rear Shooting Elbow (RSE) are maneuvers with similar arm movement. Lead Foot-Thrust (LFT) and Lead Front Kick (LFK) are maneuvers with similar leg movement. Thus, they are mistakenly classified by the technique.

4.7.2 DTW:Kinect+

Table 4.17 shows that the percent accuracy of upper body maneuvers (76%) is higher than the percent accuracy of lower body maneuvers (67%). As shown in Ta- ble 4.16, Lead Hook Punch (LHP), Rear Hook Punch (RHP), Lead Upper Punch (LUP), Rear Upper Punch (RUP), Lead Round Elbow (LRE), Rear Round El- bow (RRE), Lead Downward Elbow (LDE), Rear Downward Elbow (RDE), Lead Shooting Elbow (LSE) and Rear Shooting Elbow (RSE) are maneuvers with simi- lar arm movement. Lead Foot-Thrust (LFT), Lead Low Round Kick (LLRK), Rear Low Round Kick (RLRK), Lead High Round Kick

(LHRK), Rear High Round Kick (RHRK) and Lead Front Kick (LFK) are maneuvers with similar leg movement. Thus, they are mistakenly classified by the technique.

4.7.3 SW: Standard Kinect

Table 4.19 shows that the percent accuracy of upper body maneuvers (63%) is higher than the percent accuracy of lower body maneuvers (51%). As shown in Ta- ble 4.18, Lead Straight Punch (LSP), Lead Hook Punch (LHP), Rear Hook Punch (RHP), Rear Upper Punch (RUP), Lead Round Elbow (LRE), Rear Round El- bow (RRE), Lead Downward Elbow (LDE), Rear Downward Elbow (RDE), Lead Shooting Elbow (LSE) and Rear Shooting Elbow (RSE) are maneuvers with sim- ilar arm movement. Low Straight Knee (LSK), High Straight Knee (HSK), Lead Foot-Thrust (LFT), Rear Foot-Thrust (RFT), Lead Low Round Kick (LLRK), Rear Low Round Kick (RLRK), Lead High Round Kick (LHRK), Rear High Round Kick (RHRK), Lead Front Kick (LFK) and Rear Front Kick (RFK) are maneuvers with similar leg movement. Thus, they are mistakenly classified by the technique.

4.7.4 SW: Kinect+

Table 4.21 shows that the percent accuracy of upper body maneuvers (67%) is higher than the percent accuracy of lower body maneuvers (61%). As shown in Table 4.20, Lead Straight Punch (LSP), Lead Hook Punch (LHP), Rear Hook Punch (RHP), Lead Upper Punch (LUP), Rear Upper Punch (RUP), Lead Round Elbow (LRE), Rear Round Elbow (RRE), Lead Downward Elbow (LDE), Rear Downward Elbow (RDE), Lead Shooting Elbow (LSE) and Rear Shooting Elbow (RSE) are maneuvers with similar arm movement. Lead Foot-Thrust (LFT), Rear Foot-Thrust (RFT), Lead Low Round Kick (LLRK), Rear Low Round Kick (RLRK), Lead High Round Kick (LHRK), Rear High Round Kick (RHRK), Lead Front Kick (LFK) and Rear Front Kick (RFK) are maneuvers with similar leg movement. Thus, they are mistakenly classified by the technique.

4.7.5 HMM: Standard Kinect

Table 4.23 shows that the percent accuracy of upper body maneuvers (80%) is higher than the percent accuracy of lower body maneuvers (76%). As shown in Table 4.22, most of the data are recognized as a correct maneuver.

4.7.6HMM:Kinect+

Table 4.25 shows that the percent accuracy of upper body maneuvers (89%) is higher than the percent accuracy of lower body maneuvers (87%). As shown in Table 4.24, most of the data are recognized as a correct maneuver.

4.7.7 WSBP DTW: Standard Kinect (Weight = 4)

Table 4.27 shows that the percent accuracy of upper body maneuvers (60%) is higher than the percent accuracy of lower body maneuvers (51%). As shown in Table 4.26, Rear Hook Punch (RHP), Lead Upper Punch (LUP), Rear Upper Punch (RUP), Rear Round Elbow (RRE), Lead Downward Elbow (LDE), Rear Downward Elbow (RDE), Lead Shooting Elbow (LSE), and Rear Shooting Elbow (RSE) are maneuvers with similar arm movement. Lead Foot-Thrust (LFT) and Lead Front Kick (LFK) are maneuvers with similar leg movement. Thus, they are mistakenly classified by the technique.

4.7.8 WSBP DTW: Kinect+ (Weight = 3)

Table 4.29 shows that the percent accuracy of upper body maneuvers (64%) is higher than the percent accuracy of lower body maneuvers (57%). As shown in Table 4.28, Lead Hook Punch (LHP), Rear Hook Punch (RHP), Lead Upper Punch (LUP), Rear Upper Punch (RUP), Lead Round Elbow (LRE), Rear Round El- bow (RRE), Lead Downward Elbow (LDE), Rear Downward Elbow (RDE), Lead Shooting Elbow (LSE) and Rear Shooting Elbow (RSE) are maneuvers with simi- lar arm movement. Lead Foot-Thrust (LFT), Lead Low Round Kick (LLRK), Rear Low Round Kick (RLRK), Lead High Round Kick (LHRK), Rear High Round Kick (RHRK) and Lead Front Kick (LFK) are maneuvers with similar leg movement. Thus, they are mistakenly classified by the technique.

4.7.9 WSBP SW: Standard Kinect (Weight = 3)

Table 4.31 shows that the percent accuracy of upper body maneuvers (54%) is higher than the percent accuracy of lower body maneuvers (43%). As shown in Table 4.30, Lead Straight Punch (LSP), Lead Hook Punch (LHP), Rear Hook Punch (RHP), Rear Upper Punch (RUP), Lead Round Elbow (LRE), Rear Round El- bow (RRE), Lead Downward Elbow (LDE), Rear Downward Elbow (RDE), Lead Shooting Elbow (LSE) and Rear Shooting Elbow (RSE) are maneuvers with sim- ilar arm movement. Low Straight Knee (LSK), High Straight Knee (HSK), Lead Foot-Thrust (LFT), Rear Foot-Thrust (RFT), Lead Low Round Kick (LLRK), Rear Low Round Kick (RLRK), Lead High Round Kick (LHRK), Rear High Round Kick (RHRK), Lead Front Kick (LFK) and Rear Front Kick (RFK) are maneuvers with similar leg movement. Thus, they are mistakenly classified by the technique.

4.7.10 WSBP SW: Kinect+ (Weight = 5)

Table 4.33 shows that the percent accuracy of upper body maneuvers (57%) is higher than the percent accuracy of lower body maneuvers (49%). As shown in Table 4.32, Lead Straight Punch (LSP), Lead Hook Punch (LHP), Rear Hook Punch (RHP), Lead Upper Punch (LUP), Rear Upper Punch (RUP), Lead Round Elbow (LRE), Rear Round Elbow (RRE), Lead Downward Elbow (LDE), Rear Downward Elbow (RDE), Lead Shooting Elbow (LSE) and Rear Shooting Elbow (RSE) are maneuvers with similar arm movement. Lead Foot-Thrust (LFT), Rear Foot-Thrust (RFT), Lead Low Round Kick (LLRK), Rear Low Round Kick (RLRK), Lead High Round Kick (LHRK), Rear High Round Kick (RHRK), Lead Front Kick (LFK) and Rear Front Kick (RFK) are maneuvers with similar leg movement. Thus, they are mistakenly classified by the technique.

4.7.11 WSBP HMM Standard Kinect (Weight = 5)

Table 4.35 shows that the percent accuracy of upper body maneuvers (70%) is higher than the percent accuracy of lower body maneuvers (65%). As shown in Table 4.34, most of the data are recognized as a correct maneuver.

4.7.12 WSBP HMM Kinect+ (Weight = 4)

Table 4.37 shows that the percent accuracy of upper body maneuvers (76%) is higher than the percent accuracy of lower body maneuvers (73%). As shown in Table 4.36, most of the data are recognized as a correct maneuver.

Table 4.14: Maneuver predictions based on DTW and k-NN techniques with 5-fold cross validation. The dataset consists of 24 maneuvers from 480 videos. The features are extracted by Standard Kinect feature extraction.

DTW: Standard Kinect	LSP	RSP	LHP	RHP	LUP	RUP	LRE	RRE	LDE	RDE	LSE	RSE	SK	LSK	HSK	SFT	LFT	RFT	LLRK	RLRK	LHRK	RHRK	LFK	RFK
LSP	68	0	1	0	11	0	16	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
RSP	0	60	0	4	0	0	3	3	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0
LHP	2	0	91	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RHP	0	0	0	96	0	0	0	3	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
LUP	0	0	1	1	64	0	29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RUP	1	5	0	16	6	57	1	3	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
LRE	4	0	1	1	11	0	64	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	2	0
RRE	0	2	0	26	0	2	1	67	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
LDE	4	0	0	1	2	0	17	0	61	0	3	0	0	0	0	0	1	0	0	0	0	0	1	0
RDE	1	0	0	10	0	0	14	3	2	68	0	2	0	0	0	0	0	0	0	0	0	0	0	0
LSE	0	0	0	1	19	0	18	0	2	0	59	0	0	0	0	0	2	0	0	0	0	0	0	0
RSE	0	2	0	11	1	0	17	0	0	1	0	58	0	0	0	0	0	0	0	0	0	0	0	0
SK	0	0	0	0	0	0	3	3	0	0	0	0	58	3	1	0	0	1	0	4	0	2	0	0
LSK	0	0	0	0	0	0	1	0	0	0	0	0	4	61	3	0	0	7	0	0	0	0	0	4
HSK	0	0	0	1	0	0	1	0	0	0	0	1	0	15	55	0	0	1	0	0	0	0	0	1
SFT	0	0	1	2	0	0	7	0	0	0	0	0	0	0	0	59	1	0	0	0	0	0	0	0
LFT	0	0	0	0	3	0	3	0	0	0	2	0	0	0	0	0	56	18	0	0	0	0	8	0
RFT	0	0	0	5	0	0	1	0	0	0	0	0	1	5	10	0	0	76	0	0	0	0	0	2
LLRK	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	10	0	61	0	7	0	0	0
RLRK	0	0	0	4	0	0	0	4	0	1	0	0	1	0	2	0	0	1	10	69	0	3	0	1
LHRK	0	0	0	1	4	0	11	0	0	0	0	0	1	1	0	0	7	0	36	0	29	0	1	0
RHRK	0	0	0	10	0	3	9	0	0	0	0	0	8	3	2	0	0	5	1	19	0	20	0	5
LFK	0	0	0	1	1	0	15	0	1	0	1	0	0	0	0	0	33	0	0	0	0	0	39	0
RFK	0	0	0	1	0	0	0	0	0	0	0	1	0	6	5	0	0	11	0	0	0	1	0	60

Table 4.15: Upper and lower body percentage of maneuver predictions based on DTW and k-NN techniques with 5-fold cross validation. The dataset consists of 24 maneuvers from 480 videos. The features are extracted by Standard Kinect feature extraction.

DTW: Standard Kinect		
Lead Straight Punch	68%	71%
Rear Straight Punch	67%	
Lead Hook Punch	91%	
Rear Hook Punch	96%	
Lead Upper Punch	67%	
Rear Upper Punch	63%	
Lead Round Elbow	75%	
Rear Round Elbow	77%	
Lead Downward Elbow	61%	
Rear Downward Elbow	68%	
Lead Shooting Elbow	59%	
Rear Shooting Elbow	58%	
Swing Knee	77%	60%
Low Straight Knee	76%	
High Straight Knee	73%	
Side Foot-Thrust	84%	
Lead Foot-Thrust	62%	
Rear Foot-Thrust	66%	
Lead Low Round Kick	76%	
Rear Low Round Kick	72%	
Lead High Round Kick	21%	
Rear High Round Kick	24%	
Lead Front Kick	42%	
Rear Front Kick	41%	

Table 4.16: Maneuver predictions based on DTW and k-NN techniques with 5-fold cross validation. The dataset consists of 24 maneuvers from 480 videos. The features are extracted by Kinect+ feature extraction.

DTW: Kinect+	LSP	RSP	LHP	RHP	LUP	RUP	LRE	RRE	LDE	RDE	LSE	RSE	SK	LSK	HSK	SFT	LFT	RFT	LLRK	RLRK	LHRK	RHRK	LFK	RFK
LSP	77	0	11	4	6	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
RSP	0	61	0	13	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LHP	2	0	79	0	3	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RHP	1	2	0	95	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LUP	1	0	1	1	91	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RUP	1	5	0	9	8	67	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LRE	3	0	23	1	3	0	63	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0
RRE	3	7	1	25	1	1	0	62	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LDE	6	0	7	1	25	0	2	0	51	1	7	0	0	0	0	0	0	0	0	0	0	0	0	0
RDE	1	6	0	18	1	0	0	4	1	69	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LSE	9	0	4	1	32	0	6	0	0	0	45	3	0	0	0	0	0	0	0	0	0	0	0	0
RSE	1	1	2	19	9	9	0	3	1	0	4	50	0	0	0	0	0	0	0	0	0	1	0	0
SK	0	0	0	1	4	4	0	3	0	0	1	1	52	11	2	4	1	8	0	2	1	1	1	3
LSK	0	0	0	2	3	1	0	1	0	0	0	1	3	77	10	0	0	1	0	0	0	0	0	1
HSK	0	1	1	3	0	7	0	0	0	3	0	2	7	32	42	0	0	2	0	0	0	0	0	0
SFT	0	0	0	0	0	1	0	0	0	0	0	0	0	3	0	95	1	0	0	0	0	0	0	0
LFT	5	0	3	1	9	0	1	0	0	0	0	0	0	0	0	3	74	1	0	0	0	0	3	0
RFT	0	0	2	2	2	6	0	0	0	0	0	2	0	2	0	0	2	79	0	0	0	1	0	2
LLRK	1	0	2	0	5	0	0	0	0	0	1	0	0	0	0	0	9	0	71	1	4	0	1	0
RLRK	0	4	0	6	2	14	0	0	0	0	0	1	2	1	0	0	1	5	4	58	0	0	0	2
LHRK	2	0	2	3	10	5	1	0	1	0	1	0	1	0	1	0	15	0	18	0	30	0	5	0
RHRK	1	7	1	3	2	7	0	3	1	1	0	9	3	4	1	0	0	6	0	16	0	26	0	9
LFK	2	0	1	5	18	0	3	0	0	0	6	0	0	0	0	3	30	0	0	0	0	0	32	0
RFK	0	1	0	6	5	2	0	2	0	1	0	1	1	12	1	0	0	5	0	3	0	0	1	59

Table 4.17: Upper and lower body percentage of maneuver predictions based on DTW and k-NN techniques with 5-fold cross validation. The dataset consists of 24 maneuvers from 480 videos. The features are extracted by Kinect+ feature extraction.

DTW: Kinect+		
Lead Straight Punch	87%	76%
Rear Straight Punch	81%	
Lead Hook Punch	93%	
Rear Hook Punch	95%	
Lead Upper Punch	96%	
Rear Upper Punch	94%	
Lead Round Elbow	63%	
Rear Round Elbow	62%	
Lead Downward Elbow	61%	
Rear Downward Elbow	69%	
Lead Shooting Elbow	55%	
Rear Shooting Elbow	50%	
Swing Knee	72%	67%
Low Straight Knee	77%	
High Straight Knee	72%	
Side Foot-Thrust	95%	
Lead Foot-Thrust	74%	
Rear Foot-Thrust	79%	
Lead Low Round Kick	75%	
Rear Low Round Kick	78%	
Lead High Round Kick	32%	
Rear High Round Kick	36%	
Lead Front Kick	52%	
Rear Front Kick	59%	

Table 4.18: Maneuver predictions based on SW and k-NN techniques with 5-fold cross validation. The dataset consists of 24 maneuvers from 480 videos. The features are extracted by Standard Kinect feature extraction.

SW: Standard Kinect	LSP	RSP	LHP	RHP	LUP	RUP	LRE	RRE	LDE	RDE	LSE	RSE	SK	LSK	HSK	SFT	LFT	RFT	LLRK	RLRK	LHRK	RHRK	LFK	RFK
LSP	75	0	3	0	5	0	4	0	5	0	1	0	0	0	0	0	2	0	3	0	0	0	2	0
RSP	0	59	0	10	0	0	0	2	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0
LHP	1	0	69	0	3	0	10	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0
RHP	1	4	1	83	1	5	0	2	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0
LUP	4	0	8	0	73	2	4	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0
RUP	1	4	1	11	2	64	0	3	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0
LRE	22	0	10	4	7	0	34	0	0	0	5	0	0	0	0	0	3	0	0	0	0	0	0	0
RRE	2	2	3	31	5	4	0	47	0	0	0	5	0	0	0	0	0	1	0	0	0	0	0	0
LDE	11	0	5	8	25	0	4	0	28	4	4	0	0	0	0	0	1	0	0	0	0	0	0	0
RDE	1	1	0	25	2	0	3	1	0	64	0	3	0	0	0	0	0	0	0	0	0	0	0	0
LSE	10	0	0	2	25	0	11	1	2	0	49	0	0	0	0	0	0	0	0	0	0	0	0	0
RSE	0	1	0	22	0	6	3	1	0	0	6	51	0	0	0	0	0	0	0	0	0	0	0	0
SK	3	0	0	5	0	1	0	0	0	0	0	1	40	7	2	0	0	3	0	4	0	0	0	0
LSK	3	0	0	4	4	0	2	0	0	0	0	1	0	49	9	0	1	7	0	0	0	0	0	0
HSK	1	0	0	6	0	2	0	0	0	0	0	1	2	22	40	0	0	1	0	0	0	0	0	0
SFT	2	1	2	1	2	1	0	0	0	0	0	0	0	0	0	46	12	3	0	0	0	0	0	0
LFT	1	0	0	1	5	1	4	0	0	0	0	0	0	0	0	4	61	1	0	0	0	0	12	0
RFT	0	2	0	3	1	1	0	0	0	0	0	3	1	2	0	0	11	74	0	1	0	0	0	1
LLRK	5	1	3	2	12	0	0	0	0	0	3	0	1	0	0	0	13	0	30	0	8	1	1	0
RLRK	4	5	0	5	0	1	0	0	0	0	0	3	7	0	0	0	0	4	18	47	0	1	0	0
LHRK	7	0	3	6	18	0	0	0	0	0	4	0	4	0	0	0	23	0	6	0	16	0	3	0
RHRK	3	5	0	11	0	4	0	0	0	0	0	3	9	4	2	0	0	6	0	18	0	20	0	0
LFK	4	0	0	0	7	0	1	0	0	0	2	0	0	4	0	0	32	0	5	0	0	0	35	0
RFK	0	3	0	9	1	3	0	0	0	0	0	1	2	2	2	0	1	13	0	0	0	1	0	47

Table 4.19: Upper and lower body percentage of maneuver predictions based on SW and k-NN techniques with 5-fold cross validation. The dataset consists of 24 maneuvers from 480 videos. The features are extracted by Standard Kinect feature extraction.

SW: Standard Kinect		
Lead Straight Punch	75%	63%
Rear Straight Punch	79%	
Lead Hook Punch	81%	
Rear Hook Punch	83%	
Lead Upper Punch	77%	
Rear Upper Punch	71%	
Lead Round Elbow	40%	
Rear Round Elbow	47%	
Lead Downward Elbow	31%	
Rear Downward Elbow	64%	
Lead Shooting Elbow	49%	
Rear Shooting Elbow	57%	
Swing Knee	65%	51%
Low Straight Knee	61%	
High Straight Knee	53%	
Side Foot-Thrust	66%	
Lead Foot-Thrust	68%	
Rear Foot-Thrust	74%	
Lead Low Round Kick	38%	
Rear Low Round Kick	49%	
Lead High Round Kick	17%	
Rear High Round Kick	24%	
Lead Front Kick	39%	
Rear Front Kick	55%	

Table 4.20: Maneuver predictions based on SW and k-NN techniques with 5-fold cross validation. The dataset consists of 24 maneuvers from 480 videos. The features are extracted by Kinect+ feature extraction.

SW: Kinect+	LSP	RSP	LHP	RHP	LUP	RUP	LRE	RRE	LDE	RDE	LSE	RSE	SK	LSK	HSK	SFT	LFT	RFT	LLRK	RLRK	LHRK	RHRK	LFK	RFK
LSP	83	0	6	0	9	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0
RSP	3	59	0	11	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LHP	4	0	68	0	9	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RHP	0	4	1	87	2	4	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LUP	5	0	5	1	84	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RUP	1	3	0	19	1	66	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LRE	9	0	24	5	13	0	40	0	0	0	3	0	0	0	0	0	1	0	0	0	0	0	0	0
RRE	1	5	0	25	2	3	5	59	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LDE	12	0	7	2	25	0	19	0	33	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
RDE	2	11	1	16	0	4	3	4	1	56	0	2	0	0	0	0	0	0	0	0	0	0	0	0
LSE	9	1	8	3	33	0	12	0	0	0	26	8	0	0	0	0	0	0	0	0	0	0	0	0
RSE	4	11	1	19	12	5	0	3	1	3	1	39	0	0	0	0	1	0	0	0	0	0	0	0
SK	10	10	1	7	0	0	1	0	0	0	0	0	49	6	1	2	0	9	1	2	1	0	0	0
LSK	0	5	0	1	2	2	1	0	0	0	0	0	3	66	15	0	0	4	0	0	0	0	0	1
HSK	5	3	0	4	2	0	2	0	0	0	0	0	3	45	31	0	0	5	0	0	0	0	0	0
SFT	13	0	2	4	2	1	0	0	0	0	0	0	0	0	0	74	4	0	0	0	0	0	0	0
LFT	3	0	4	0	2	0	2	0	0	0	0	0	0	0	0	8	55	20	0	0	0	0	5	1
RFT	5	3	1	6	2	6	0	0	1	0	0	1	0	3	0	2	4	59	0	0	0	1	0	6
LLRK	2	0	11	0	1	0	6	0	1	0	0	0	0	1	0	0	22	0	30	0	1	0	0	0
RLRK	8	18	0	15	0	4	0	0	0	0	0	1	2	0	0	0	0	0	6	46	0	0	0	0
LHRK	2	0	9	2	6	0	0	0	0	0	1	0	1	0	0	0	4	0	22	0	26	0	0	2
RHRK	8	12	0	6	6	0	4	0	1	0	0	9	3	2	0	2	0	4	2	18	0	23	0	0
LFK	0	0	3	4	15	0	9	0	0	0	0	0	0	0	0	0	22	2	0	0	0	0	45	0
RFK	0	4	0	14	0	3	0	1	0	0	0	1	1	7	0	2	0	20	0	1	0	0	0	46

Table 4.21: Upper and lower body percentage of maneuver predictions based on SW and k-NN techniques with 5-fold cross validation. The dataset consists of 24 maneuvers from 480 videos. The features are extracted by Kinect+ feature extraction.

SW: Kinect+		
Lead Straight Punch	83%	67%
Rear Straight Punch	89%	
Lead Hook Punch	80%	
Rear Hook Punch	87%	
Lead Upper Punch	88%	
Rear Upper Punch	83%	
Lead Round Elbow	52%	
Rear Round Elbow	59%	
Lead Downward Elbow	53%	
Rear Downward Elbow	56%	
Lead Shooting Elbow	36%	
Rear Shooting Elbow	39%	
Swing Knee	69%	61%
Low Straight Knee	66%	
High Straight Knee	65%	
Side Foot-Thrust	74%	
Lead Foot-Thrust	75%	
Rear Foot-Thrust	79%	
Lead Low Round Kick	42%	
Rear Low Round Kick	46%	
Lead High Round Kick	28%	
Rear High Round Kick	28%	
Lead Front Kick	56%	
Rear Front Kick	56%	

Table 4.22: Maneuver predictions based on HMM with 5-fold cross validation. The dataset consists of 24 maneuvers from 480 videos. The features are extracted by Standard Kinect feature extraction.

HMM: Standard Kinect	LSP	RSP	LHP	RHP	LUP	RUP	LRE	RRE	LDE	RDE	LSE	RSE	SK	LSK	HSK	SFT	LFT	RFT	LLRK	RLRK	LHRK	RHRK	LFK	RFK
LSP	17	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RSP	0	13	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LHP	1	0	13	0	2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RHP	0	1	0	19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LUP	0	0	1	0	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RUP	0	2	0	1	0	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LRE	4	0	0	1	0	0	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RRE	1	0	0	5	0	0	0	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LDE	2	0	0	0	0	0	6	0	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RDE	0	3	0	3	0	0	0	0	0	13	0	1	0	0	0	0	0	0	0	0	0	0	0	0
LSE	1	1	1	1	0	0	3	0	0	0	12	1	0	0	0	0	0	0	0	0	0	0	0	0
RSE	1	3	0	0	1	0	0	1	1	0	0	13	0	0	0	0	0	0	0	0	0	0	0	0
SK	3	2	0	2	0	0	0	0	0	0	0	0	11	0	0	0	0	2	0	0	0	0	0	0
LSK	0	2	0	0	0	1	1	0	0	0	0	0	0	13	3	0	0	0	0	0	0	0	0	0
HSK	1	0	0	2	0	0	1	0	0	0	0	0	1	1	13	0	0	1	0	0	0	0	0	0
SFT	3	0	0	1	0	0	0	0	0	0	0	0	0	0	0	16	0	0	0	0	0	0	0	0
LFT	1	0	2	0	4	0	0	0	0	0	0	0	0	0	0	1	11	0	0	0	0	0	1	0
RFT	1	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	1	15	0	0	0	0	0	0
LLRK	0	0	2	0	0	0	2	0	0	0	0	0	0	1	0	0	0	0	14	0	0	0	0	0
RLRK	2	4	0	2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	10	0	0	0	0
LHRK	0	0	1	1	0	0	3	0	0	0	1	0	0	0	0	0	1	0	0	0	12	0	0	0
RHRK	0	0	0	2	1	0	1	0	1	0	0	1	0	0	0	1	0	1	0	1	0	11	0	0
LFK	0	0	1	0	0	0	3	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	14	0
RFK	0	1	0	3	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	13

Table 4.23: Upper and lower body percentage of maneuver predictions based on HMM with 5-fold cross validation. The dataset consists of 24 maneuvers from 480 videos. The features are extracted by Standard Kinect feature extraction.

HMM: Standard Kinect		
Lead Straight Punch	97%	80%
Rear Straight Punch	91%	
Lead Hook Punch	93%	
Rear Hook Punch	95%	
Lead Upper Punch	96%	
Rear Upper Punch	94%	
Lead Round Elbow	73%	
Rear Round Elbow	72%	
Lead Downward Elbow	71%	
Rear Downward Elbow	79%	
Lead Shooting Elbow	65%	
Rear Shooting Elbow	60%	
Swing Knee	82%	76%
Low Straight Knee	87%	
High Straight Knee	82%	
Side Foot-Thrust	95%	
Lead Foot-Thrust	84%	
Rear Foot-Thrust	89%	
Lead Low Round Kick	75%	
Rear Low Round Kick	78%	
Lead High Round Kick	52%	
Rear High Round Kick	56%	
Lead Front Kick	62%	
Rear Front Kick	69%	

Table 4.24: Maneuver predictions based on HMM with 5-fold cross validation. The dataset consists of 24 maneuvers from 480 videos. The features are extracted by Kinect+ feature extraction.

HMM: Kinect+	LSP	RSP	LHP	RHP	LUP	RUP	LRE	RRE	LDE	RDE	LSE	RSE	SK	LSK	HSK	SFT	LFT	RFT	LLRK	RLRK	LHRK	RHRK	LFK	RFK
LSP	18	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
RSP	0	13	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LHP	1	0	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RHP	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LUP	0	0	0	0	19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RUP	0	1	0	2	0	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LRE	1	0	2	5	1	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RRE	1	1	0	4	0	0	0	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LDE	0	0	2	1	3	0	0	0	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RDE	0	1	0	4	0	0	0	0	0	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LSE	0	0	0	1	4	0	1	0	0	0	14	0	0	0	0	0	0	0	0	0	0	0	0	0
RSE	1	1	0	3	1	0	0	0	0	0	0	14	0	0	0	0	0	0	0	0	0	0	0	0
SK	0	0	0	0	0	0	0	0	0	0	1	0	14	1	1	0	1	2	0	0	0	0	0	0
LSK	0	0	0	1	0	0	0	0	0	0	0	0	1	15	2	0	0	1	0	0	0	0	0	0
HSK	0	0	0	0	0	2	0	0	0	1	0	1	1	6	9	0	0	0	0	0	0	0	0	0
SFT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0
LFT	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	16	0	0	0	0	0	1	0
RFT	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	19	0	0	0	0	0	0
LLRK	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	2	0	15	0	0	0	0	0
RLRK	0	1	0	1	0	3	0	0	0	0	0	0	0	0	0	0	0	0	1	14	0	0	0	0
LHRK	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	0	3	0	13	0	0	0
RHRK	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	2	0	15	0	1
LFK	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	2	0	0	0	0	0	0	16	0
RFK	0	0	0	2	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	16

Table 4.25: Upper and lower body percentage of maneuver predictions based on HMM with 5-fold cross validation. The dataset consists of 24 maneuvers from 480 videos. The features are extracted by Kinect+ feature extraction.

HMM: Kinect+		
Lead Straight Punch	97%	89%
Rear Straight Punch	91%	
Lead Hook Punch	93%	
Rear Hook Punch	95%	
Lead Upper Punch	96%	
Rear Upper Punch	94%	
Lead Round Elbow	83%	
Rear Round Elbow	82%	
Lead Downward Elbow	81%	
Rear Downward Elbow	89%	
Lead Shooting Elbow	85%	
Rear Shooting Elbow	80%	
Swing Knee	82%	87%
Low Straight Knee	87%	
High Straight Knee	82%	
Side Foot-Thrust	95%	
Lead Foot-Thrust	84%	
Rear Foot-Thrust	89%	
Lead Low Round Kick	85%	
Rear Low Round Kick	88%	
Lead High Round Kick	82%	
Rear High Round Kick	86%	
Lead Front Kick	82%	
Rear Front Kick	89%	

Table 4.26: Maneuver predictions based on DTW and k-NN techniques with 5-fold cross validation and WSBP (Weight = 4). The dataset consists of 24 maneuvers from 480 videos. The features are extracted by Standard Kinect feature extraction.

WSBP DTW: Standard Kinect	LSP	RSP	LHP	RHP	LUP	RUP	LRE	RRE	LDE	RDE	LSE	RSE	SK	LSK	HSK	SFT	LFT	RFT	LLRK	RLRK	LHRK	RHRK	LFK	RFK
LSP	58	0	1	0	13	0	19	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
RSP	0	51	0	5	0	0	4	4	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0
LHP	2	0	69	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RHP	0	0	0	82	0	0	0	4	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
LUP	0	0	1	1	55	0	34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RUP	1	6	0	19	7	49	1	4	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
LRE	5	0	1	1	13	0	55	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	2	0
RRE	0	2	0	30	0	2	1	57	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
LDE	5	0	0	1	2	0	20	0	52	0	4	0	0	0	0	0	1	0	0	0	0	0	1	0
RDE	1	0	0	12	0	0	16	4	2	58	0	2	0	0	0	0	0	0	0	0	0	0	0	0
LSE	0	0	0	1	21	0	21	0	2	0	50	0	0	0	0	0	2	0	0	0	0	0	0	0
RSE	0	2	0	13	1	0	20	0	0	1	0	49	0	0	0	0	0	0	0	0	0	0	0	0
SK	0	0	0	0	0	0	4	4	0	0	0	0	49	4	1	0	0	1	0	5	0	2	0	0
LSK	0	0	0	0	0	0	1	0	0	0	0	0	5	52	4	0	0	8	0	0	0	0	0	5
HSK	0	0	0	1	0	0	1	0	0	0	0	1	0	18	47	0	0	1	0	0	0	0	0	1
SFT	0	0	1	2	0	0	8	0	0	0	0	0	0	0	0	50	1	0	0	0	0	0	0	0
LFT	0	0	0	0	4	0	4	0	0	0	2	0	0	0	0	0	48	21	0	0	0	0	9	0
RFT	0	0	0	6	0	0	1	0	0	0	0	0	1	6	12	0	0	65	0	0	0	0	0	2
LLRK	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	12	0	52	0	8	0	0	0
RLRK	0	0	0	5	0	0	0	5	0	1	0	0	1	0	2	0	0	1	12	58	0	4	0	1
LHRK	0	0	0	1	5	0	13	0	0	0	0	0	1	1	0	0	8	0	42	0	24	0	1	0
RHRK	0	0	0	12	0	4	11	0	0	0	0	0	9	4	2	0	0	6	1	22	0	17	0	6
LFK	0	0	0	1	1	0	18	0	1	0	1	0	0	0	0	0	39	0	0	0	0	0	32	0
RFK	0	0	0	1	0	0	0	0	0	0	0	1	0	7	6	0	0	13	0	0	0	1	0	51

Table 4.27: Upper and lower body percentage of maneuver predictions based on DTW and k-NN techniques with 5-fold cross validation andWSBP (Weight = 4). The dataset consists of 24 maneuvers from 480 videos. The features are extracted by Standard Kinect feature extraction.

WSBP DTW: Standard Kinect		
Lead Straight Punch	58%	60%
Rear Straight Punch	57%	
Lead Hook Punch	78%	
Rear Hook Punch	82%	
Lead Upper Punch	57%	
Rear Upper Punch	54%	
Lead Round Elbow	64%	
Rear Round Elbow	66%	
Lead Downward Elbow	52%	
Rear Downward Elbow	58%	
Lead Shooting Elbow	50%	
Rear Shooting Elbow	49%	
Swing Knee	66%	51%
Low Straight Knee	65%	
High Straight Knee	62%	
Side Foot-Thrust	72%	
Lead Foot-Thrust	53%	
Rear Foot-Thrust	56%	
Lead Low Round Kick	65%	
Rear Low Round Kick	61%	
Lead High Round Kick	18%	
Rear High Round Kick	20%	
Lead Front Kick	36%	
Rear Front Kick	35%	

Table 4.28: Maneuver predictions based on DTW and k-NN techniques with 5-fold cross validation and WSBP (Weight = 3). The dataset consists of 24 maneuvers from 480 videos. The features are extracted by Kinect+ feature extraction.

WSBP DTW: Kinect+	LSP	RSP	LHP	RHP	LUP	RUP	LRE	RRE	LDE	RDE	LSE	RSE	SK	LSK	HSK	SFT	LFT	RFT	LLRK	RLRK	LHRK	RHRK	LFK	RFK
LSP	66	0	13	5	7	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
RSP	0	52	0	15	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LHP	2	0	67	0	4	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RHP	1	2	0	81	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LUP	1	0	1	1	78	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RUP	1	6	0	11	9	57	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LRE	4	0	27	1	4	0	54	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0
RRE	4	8	1	29	1	1	0	53	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LDE	7	0	8	1	29	0	2	0	43	1	8	0	0	0	0	0	0	0	0	0	0	0	0	0
RDE	1	7	0	21	1	0	0	5	1	59	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LSE	11	0	5	1	38	0	7	0	0	0	38	4	0	0	0	0	0	0	0	0	0	0	0	0
RSE	1	1	2	22	11	11	0	4	1	0	5	43	0	0	0	0	0	0	0	0	0	1	0	0
SK	0	0	0	1	5	5	0	4	0	0	1	1	44	13	2	5	1	9	0	2	1	1	1	4
LSK	0	0	0	2	4	1	0	1	0	0	0	1	4	66	12	0	0	1	0	0	0	0	0	1
HSK	0	1	1	4	0	8	0	0	0	4	0	2	8	38	36	0	0	2	0	0	0	0	0	0
SFT	0	0	0	0	0	1	0	0	0	0	0	0	0	4	0	81	1	0	0	0	0	0	0	0
LFT	6	0	4	1	11	0	1	0	0	0	0	0	0	0	0	4	63	1	0	0	0	0	4	0
RFT	0	0	2	2	2	7	0	0	0	0	0	2	0	2	0	0	2	67	0	0	0	1	0	2
LLRK	1	0	2	0	6	0	0	0	0	0	1	0	0	0	0	0	11	0	61	1	5	0	1	0
RLRK	0	5	0	7	2	16	0	0	0	0	0	1	2	1	0	0	1	6	5	49	0	0	0	2
LHRK	2	0	2	4	12	6	1	0	1	0	1	0	1	0	1	0	18	0	21	0	26	0	6	0
RHRK	1	8	1	4	2	8	0	4	1	1	0	11	4	5	1	0	0	7	0	19	0	22	0	11
LFK	2	0	1	6	21	0	4	0	0	0	7	0	0	0	0	4	35	0	0	0	0	0	27	0
RFK	0	1	0	7	6	2	0	2	0	1	0	1	1	14	1	0	0	6	0	4	0	0	1	50

Table 4.29: Upper and lower body percentage of maneuver predictions based on DTW and k-NN techniques with 5-fold cross validation andWSBP (Weight = 3). The dataset consists of 24 maneuvers from 480 videos. The features are extracted by Kinect+ feature extraction.

WSBP DTW: Kinect+		
Lead Straight Punch	74%	64%
Rear Straight Punch	69%	
Lead Hook Punch	79%	
Rear Hook Punch	81%	
Lead Upper Punch	82%	
Rear Upper Punch	80%	
Lead Round Elbow	54%	
Rear Round Elbow	53%	
Lead Downward Elbow	52%	
Rear Downward Elbow	59%	
Lead Shooting Elbow	47%	
Rear Shooting Elbow	43%	
Swing Knee	61%	57%
Low Straight Knee	66%	
High Straight Knee	61%	
Side Foot-Thrust	81%	
Lead Foot-Thrust	63%	
Rear Foot-Thrust	67%	
Lead Low Round Kick	64%	
Rear Low Round Kick	66%	
Lead High Round Kick	27%	
Rear High Round Kick	31%	
Lead Front Kick	44%	
Rear Front Kick	50%	

Table 4.30: Maneuver predictions based on SW and k-NN techniques with 5-fold cross validation and WSBP (Weight = 3). The dataset consists of 24 maneuvers from 480 videos. The features are extracted by Standard Kinect feature extraction.

WSBP SW: Standard Kinect	LSP	RSP	LHP	RHP	LUP	RUP	LRE	RRE	LDE	RDE	LSE	RSE	SK	LSK	HSK	SFT	LFT	RFT	LLRK	RLRK	LHRK	RHRK	LFK	RFK
LSP	64	0	4	0	6	0	5	0	6	0	1	0	0	0	0	0	2	0	4	0	0	0	2	0
RSP	0	50	0	12	0	0	0	2	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0
LHP	1	0	59	0	4	0	12	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0
RHP	1	5	1	71	1	6	0	2	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0
LUP	5	0	9	0	62	2	5	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0
RUP	1	5	1	13	2	55	0	4	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0
LRE	26	0	12	5	8	0	29	0	0	0	6	0	0	0	0	0	4	0	0	0	0	0	0	0
RRE	2	2	4	36	6	5	0	40	0	0	0	6	0	0	0	0	0	1	0	0	0	0	0	0
LDE	13	0	6	9	29	0	5	0	24	5	5	0	0	0	0	0	1	0	0	0	0	0	0	0
RDE	1	1	0	29	2	0	4	1	0	55	0	4	0	0	0	0	0	0	0	0	0	0	0	0
LSE	12	0	0	2	29	0	13	1	2	0	42	0	0	0	0	0	0	0	0	0	0	0	0	0
RSE	0	1	0	26	0	7	4	1	0	0	7	43	0	0	0	0	0	0	0	0	0	0	0	0
SK	4	0	0	6	0	1	0	0	0	0	0	1	42	8	2	0	0	4	0	5	0	0	0	0
LSK	4	0	0	5	5	0	2	0	0	0	0	1	0	42	11	0	1	8	0	0	0	0	0	0
HSK	1	0	0	7	0	2	0	0	0	0	0	1	2	26	34	0	0	1	0	0	0	0	0	0
SFT	2	1	2	1	2	1	0	0	0	0	0	0	0	0	0	39	14	4	0	0	0	0	0	0
LFT	1	0	0	1	6	1	5	0	0	0	0	0	0	0	0	5	52	1	0	0	0	0	14	0
RFT	0	2	0	4	1	1	0	0	0	0	0	4	1	2	0	0	13	63	0	1	0	0	0	1
LLRK	6	1	4	2	14	0	0	0	0	0	4	0	1	0	0	0	15	0	26	0	9	1	1	0
RLRK	5	6	0	6	0	1	0	0	0	0	0	4	8	0	0	0	0	5	21	40	0	1	0	0
LHRK	8	0	4	7	21	0	0	0	0	0	5	0	5	0	0	0	27	0	7	0	14	0	4	0
RHRK	4	6	0	13	0	5	0	0	0	0	0	4	11	5	2	0	0	7	0	21	0	17	0	0
LFK	5	0	0	0	8	0	1	0	0	0	2	0	0	5	0	0	38	0	6	0	0	0	30	0
RFK	0	4	0	11	1	4	0	0	0	0	0	1	2	2	2	0	1	15	0	0	0	1	0	40

Table 4.31: Upper and lower body percentage of maneuver predictions based on SW and k-NN techniques with 5-fold cross validation and WSBP (Weight = 3). The dataset consists of 24 maneuvers from 480 videos. The features are extracted by Standard Kinect feature extraction.

WSBP SW: Standard Kinect		
Lead Straight Punch	64%	54%
Rear Straight Punch	67%	
Lead Hook Punch	69%	
Rear Hook Punch	71%	
Lead Upper Punch	66%	
Rear Upper Punch	61%	
Lead Round Elbow	34%	
Rear Round Elbow	40%	
Lead Downward Elbow	26%	
Rear Downward Elbow	55%	
Lead Shooting Elbow	42%	
Rear Shooting Elbow	49%	
Swing Knee	55%	43%
Low Straight Knee	52%	
High Straight Knee	45%	
Side Foot-Thrust	56%	
Lead Foot-Thrust	58%	
Rear Foot-Thrust	63%	
Lead Low Round Kick	32%	
Rear Low Round Kick	42%	
Lead High Round Kick	14%	
Rear High Round Kick	20%	
Lead Front Kick	33%	
Rear Front Kick	47%	

Table 4.32: Maneuver predictions based on SW and k-NN techniques with 5-fold cross validation and WSBP (Weight = 5). The dataset consists of 24 maneuvers from 480 videos. The features are extracted by Kinect+ feature extraction.

WSBP SW: Kinect+	LSP	RSP	LHP	RHP	LUP	RUP	LRE	RRE	LDE	RDE	LSE	RSE	SK	LSK	HSK	SFT	LFT	RFT	LLRK	RLRK	LHRK	RHRK	LFK	RFK
LSP	71	0	7	0	11	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0
RSP	4	50	0	13	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LHP	5	0	58	0	11	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RHP	0	5	1	74	2	5	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LUP	6	0	6	1	72	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RUP	1	4	0	22	1	56	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LRE	11	0	28	6	15	0	34	0	0	0	4	0	0	0	0	0	1	0	0	0	0	0	0	0
RRE	1	6	0	29	2	4	6	50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LDE	14	0	8	2	29	0	22	0	28	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
RDE	2	13	1	19	0	5	4	5	1	48	0	2	0	0	0	0	0	0	0	0	0	0	0	0
LSE	11	1	9	4	39	0	14	0	0	0	22	9	0	0	0	0	0	0	0	0	0	0	0	0
RSE	5	13	1	22	14	6	0	4	1	4	1	33	0	0	0	0	1	0	0	0	0	0	0	0
SK	12	12	1	8	0	0	1	0	0	0	0	0	42	7	1	2	0	11	1	2	1	0	0	0
LSK	0	6	0	1	2	2	1	0	0	0	0	0	4	56	18	0	0	5	0	0	0	0	0	1
HSK	6	4	0	5	2	0	2	0	0	0	0	0	4	53	26	0	0	6	0	0	0	0	0	0
SFT	15	0	2	5	2	1	0	0	0	0	0	0	0	0	0	63	5	0	0	0	0	0	0	0
LFT	4	0	5	0	2	0	2	0	0	0	0	0	0	0	0	9	47	23	0	0	0	0	6	1
RFT	6	4	1	7	2	7	0	0	1	0	0	1	0	4	0	2	5	50	0	0	0	1	0	7
LLRK	2	0	13	0	1	0	7	0	1	0	0	0	0	1	0	0	26	0	26	0	1	0	0	0
RLRK	9	21	0	18	0	5	0	0	0	0	0	1	2	0	0	0	0	0	7	39	0	0	0	0
LHRK	2	0	11	2	7	0	0	0	0	0	1	0	1	0	0	0	5	0	26	0	22	0	0	2
RHRK	9	14	0	7	7	0	5	0	1	0	0	11	4	2	0	2	0	5	2	21	0	20	0	0
LFK	0	0	4	5	18	0	11	0	0	0	0	0	0	0	0	0	26	2	0	0	0	0	39	0
RFK	0	5	0	16	0	4	0	1	0	0	0	1	1	8	0	2	0	23	0	1	0	0	0	39

Table 4.33: Upper and lower body percentage of maneuver predictions based on SW and k-NN techniques with 5-fold cross validation and WSBP (Weight = 5). The dataset consists of 24 maneuvers from 480 videos. The features are extracted by Kinect+ feature extraction.

WSBP SW: Kinect+		
Lead Straight Punch	71%	57%
Rear Straight Punch	76%	
Lead Hook Punch	68%	
Rear Hook Punch	74%	
Lead Upper Punch	75%	
Rear Upper Punch	71%	
Lead Round Elbow	44%	
Rear Round Elbow	50%	
Lead Downward Elbow	45%	
Rear Downward Elbow	48%	
Lead Shooting Elbow	31%	
Rear Shooting Elbow	33%	
Swing Knee	59%	49%
Low Straight Knee	56%	
High Straight Knee	55%	
Side Foot-Thrust	63%	
Lead Foot-Thrust	64%	
Rear Foot-Thrust	67%	
Lead Low Round Kick	36%	
Rear Low Round Kick	39%	
Lead High Round Kick	24%	
Rear High Round Kick	24%	
Lead Front Kick	48%	
Rear Front Kick	48%	

Table 4.34: Maneuver predictions based on HMM with 5-fold cross validation and WSBP (Weight = 5). The dataset consists of 24 maneuvers from 480 videos. The features are extracted by Standard Kinect feature extraction.

WSBP HMM: Standard Kinect	LSP	RSP	LHP	RHP	LUP	RUP	LRE	RRE	LDE	RDE	LSE	RSE	SK	LSK	HSK	SFT	LFT	RFT	LLRK	RLRK	LHRK	RHRK	LFK	RFK
LSP	14	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RSP	0	11	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LHP	1	0	11	0	2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RHP	0	1	0	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LUP	0	0	1	0	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RUP	0	2	0	1	0	13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LRE	5	0	0	1	0	0	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RRE	1	0	0	6	0	0	0	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LDE	2	0	0	0	0	0	7	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RDE	0	4	0	4	0	0	0	0	0	11	0	1	0	0	0	0	0	0	0	0	0	0	0	0
LSE	1	1	1	1	0	0	4	0	0	0	10	1	0	0	0	0	0	0	0	0	0	0	0	0
RSE	1	4	0	0	1	0	0	1	1	0	0	11	0	0	0	0	0	0	0	0	0	0	0	0
SK	4	2	0	2	0	0	0	0	0	0	0	0	9	0	0	0	0	2	0	0	0	0	0	0
LSK	0	2	0	0	0	1	1	0	0	0	0	0	0	11	4	0	0	0	0	0	0	0	0	0
HSK	1	0	0	2	0	0	1	0	0	0	0	0	1	1	11	0	0	1	0	0	0	0	0	0
SFT	4	0	0	1	0	0	0	0	0	0	0	0	0	0	0	14	0	0	0	0	0	0	0	0
LFT	1	0	2	0	5	0	0	0	0	0	0	0	0	0	0	1	9	0	0	0	0	0	1	0
RFT	1	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	1	13	0	0	0	0	0	0
LLRK	0	0	2	0	0	0	2	0	0	0	0	0	0	1	0	0	0	0	12	0	0	0	0	0
RLRK	2	5	0	2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	9	0	0	0	0
LHRK	0	0	1	1	0	0	4	0	0	0	1	0	0	0	0	0	1	0	0	0	10	0	0	0
RHRK	0	0	0	2	1	0	1	0	1	0	0	1	0	0	0	1	0	1	0	1	0	9	0	0
LFK	0	0	1	0	0	0	4	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	12	0
RFK	0	1	0	4	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	11

Table 4.35: Upper and lower body percentage of maneuver predictions based on HMM with 5-fold cross validation and WSBP (Weight = 5). The dataset consists of 24 maneuvers from 480 videos. The features are extracted by Standard Kinect feature extraction.

WSBP HMM: Standard Kinect		
Lead Straight Punch	83%	70%
Rear Straight Punch	78%	
Lead Hook Punch	79%	
Rear Hook Punch	81%	
Lead Upper Punch	82%	
Rear Upper Punch	80%	
Lead Round Elbow	62%	
Rear Round Elbow	61%	
Lead Downward Elbow	61%	
Rear Downward Elbow	67%	
Lead Shooting Elbow	55%	
Rear Shooting Elbow	51%	
Swing Knee	70%	65%
Low Straight Knee	74%	
High Straight Knee	70%	
Side Foot-Thrust	81%	
Lead Foot-Thrust	72%	
Rear Foot-Thrust	76%	
Lead Low Round Kick	64%	
Rear Low Round Kick	66%	
Lead High Round Kick	44%	
Rear High Round Kick	48%	
Lead Front Kick	53%	
Rear Front Kick	59%	

Table 4.36: Maneuver predictions based on HMM with 5-fold cross validation and WSBP (Weight = 4). The dataset consists of 24 maneuvers from 480 videos. The features are extracted by Kinect+ feature extraction.

WSBP HMM: Kinect+	LSP	RSP	LHP	RHP	LUP	RUP	LRE	RRE	LDE	RDE	LSE	RSE	SK	LSK	HSK	SFT	LFT	RFT	LLRK	RLRK	LHRK	RHRK	LFK	RFK
LSP	15	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
RSP	0	11	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LHP	1	0	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RHP	0	0	0	17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LUP	0	0	0	0	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RUP	0	1	0	2	0	13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LRE	1	0	2	6	1	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RRE	1	1	0	5	0	0	0	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LDE	0	0	2	1	4	0	0	0	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RDE	0	1	0	5	0	0	0	0	0	13	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LSE	0	0	0	1	5	0	1	0	0	0	12	0	0	0	0	0	0	0	0	0	0	0	0	0
RSE	1	1	0	4	1	0	0	0	0	0	0	12	0	0	0	0	0	0	0	0	0	0	0	0
SK	0	0	0	0	0	0	0	0	0	0	1	0	12	1	1	0	1	2	0	0	0	0	0	0
LSK	0	0	0	1	0	0	0	0	0	0	0	0	1	13	2	0	0	1	0	0	0	0	0	0
HSK	0	0	0	0	0	2	0	0	0	1	0	1	1	7	8	0	0	0	0	0	0	0	0	0
SFT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	17	0	0	0	0	0	0	0	0
LFT	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	14	0	0	0	0	0	1	0
RFT	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	16	0	0	0	0	0	0
LLRK	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	2	0	13	0	0	0	0	0
RLRK	0	1	0	1	0	4	0	0	0	0	0	0	0	0	0	0	0	0	1	12	0	0	0	0
LHRK	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	0	4	0	11	0	0	0
RHRK	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	2	0	13	0	1
LFK	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	2	0	0	0	0	0	0	14	0
RFK	0	0	0	2	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	14

Table 4.37: Upper and lower body percentage of maneuver predictions based on HMM with 5-fold cross validation and WSBP (Weight = 4). The dataset consists of 24 maneuvers from 480 videos. The features are extracted by Kinect+ feature extraction.

WSBP HMM: Kinect+		
Lead Straight Punch	83%	76%
Rear Straight Punch	78%	
Lead Hook Punch	79%	
Rear Hook Punch	81%	
Lead Upper Punch	82%	
Rear Upper Punch	80%	
Lead Round Elbow	71%	
Rear Round Elbow	70%	
Lead Downward Elbow	69%	
Rear Downward Elbow	76%	
Lead Shooting Elbow	72%	
Rear Shooting Elbow	68%	
Swing Knee	70%	73%
Low Straight Knee	74%	
High Straight Knee	70%	
Side Foot-Thrust	81%	
Lead Foot-Thrust	72%	
Rear Foot-Thrust	76%	
Lead Low Round Kick	72%	
Rear Low Round Kick	75%	
Lead High Round Kick	70%	
Rear High Round Kick	73%	
Lead Front Kick	70%	
Rear Front Kick	76%	

4.8 Summary

HMM Kinect+ yields the best results with 88.95% accuracy of maneuver prediction among the proposed techniques as shown in Table 4.1. By HMM fine-tuning, HMM Kinect+'s best parameters are observation at 580 and states at 10 as shown in the Graph 4.8. Table 4.1 shows that our *Rule−based Approach for Improving Kinect Skeletal Tracking* in Chapter 3 makes the better results of maneuver predictions. Maneuvers with similar arm or leg movement are mistakenly classified by the DTW and SW technique as the results in Section 4.7. Weighted Separated Body Part (WSBP) does not returns the good results of maneuver predictions as shown in Table 4.1.

Chapter 5
INSTRUCTIONAL SUPPORT SYSTEM FOR MUAY THAI

Our objective is to make a full experience of Muay Thai maneuvers training. The instructional support system uses HMM with Kinect+ to find the maneuver similarity between a player maneuver data and our database maneuvers as explained in Chapter 4. Figure 5.1 shows the flowchart of the system. When the player starts the system, the player must do the calibration first as shown in Figure 5.2. Then, the system will show the introduction screen as shown in Figure 5.4. When the player understands the body- based control in Section 5.1, the player will go to a menu as shown in Figure 5.3. There are 2 main training mode which are single maneuver training and multiple maneuvers training. In single maneuver training, the player selects a maneuver and sees a sample maneuver animation of the professional Muay Thai master. The player can perform the maneuver and see a feedback result from the system. The player can also see the animations between the player and a professional Muay Thai master. This is the long- term training which the player spends a lot of time to practice and improves the player maneuvers. The player must recognize what is the diference between the professional animation and the player animation. The player must perform the maneuver as same as the professional performance. The more similar movement, the better feedback result. In multiple maneuvers training, the player does not select a maneuver and sees any animation, but the system will random a sequence of diferent maneuvers. This is the training for a real situation which the player has to apply the maneuvers practiced in the single maneuver training.

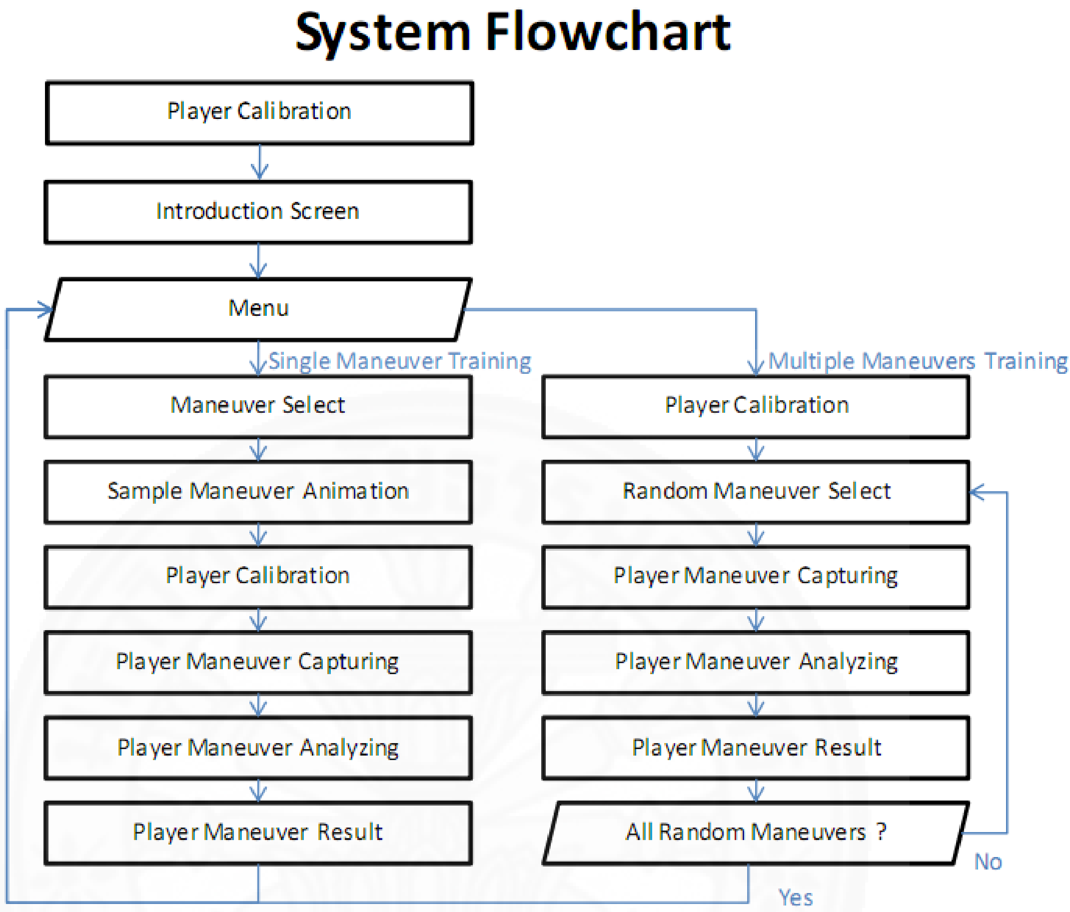

Figure 5.1: System Flowchart

Figure 5.2: Player Calibration

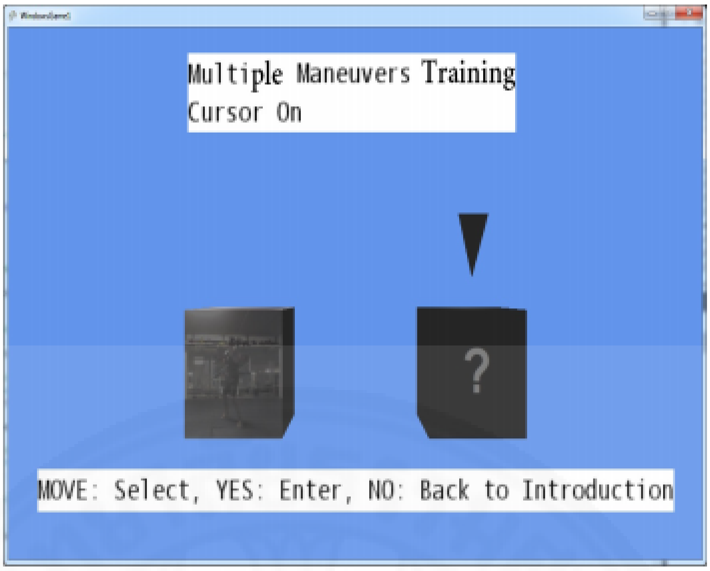

Figure 5.3: Menu

5.1 Body-based Control

There are 3 main controls by the player body as shown in Figure 5.4.

- "Yes"

  If the player raises the right hand up above the head position, the system recognizes it as "Yes".
- "No"

  If the player crosses the hands or the left hand position is on the right more than the right hand, the system recognizes it as "No".
- "Move"

  If the player lowers the right hand to the hip position in the 10 centimeters range, the system will trigger cursor system to "On". The cursor system is similar to a mouse of a computer. The direction of hand is the direction of the cursor.

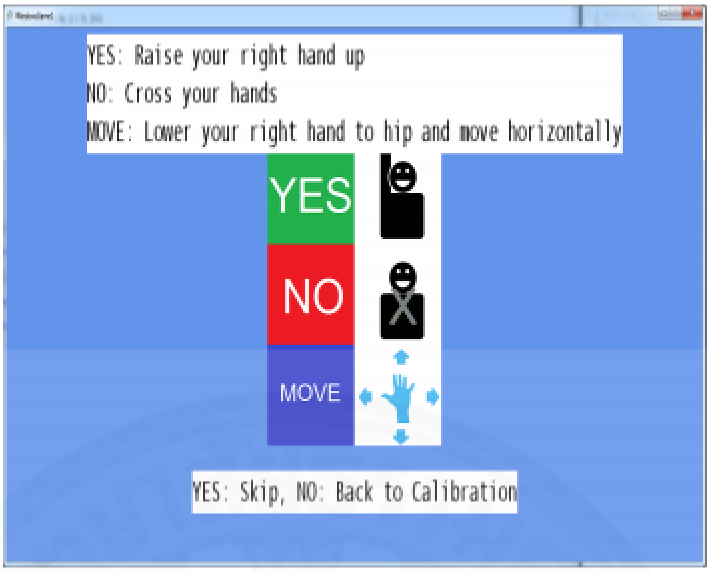

Figure 5.4: Introduction Screen

5.2 Single Maneuver Training

The single maneuver training focuses on each maneuver separately. To make the player understand the maneuver, the system first demonstrates the maneuver performed by the professional boxer as shown in Figure 5.6. The player can rotate the animation to see the maneuver in diferent angles. After studying the maneuver, the player performs the maneuver on his/her own. The system records the maneuver and replay it so that the player can check his/her form. Moreover, the system also computes the similarity between the maneuver performed by the player and the stored maneuver performed by the professional boxer. The result will show the percentage of similarity between the player maneuver and the correct maneuver. The result will also show animations between the player and the demonstration animation which can be rotated in diferent angles as shown in Figure 5.11. Next, we explain the details of this function.

First, the player selects a maneuver from 24 maneuvers as shown in Figure 5.5. There are 12 upper part body maneuvers and 12 lower part body maneuvers. The player can select a maneuver by doing "Move" and start training by doing "Yes".

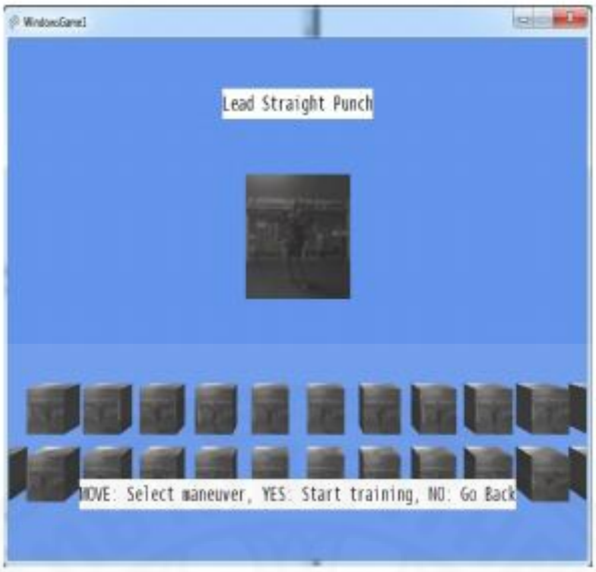

Figure 5.5: Maneuver Select

When the player starts training, the system will show the animation of selected maneuver as in Figure 5.6. It appears repeatedly. The player can rotate the camera view by doing "Move". If the player understands the movement of all postures, the player can start training by doing "Yes".

Figure 5.6: Sample Maneuver Animation

Then, the system will calibrate the player as shown in Figure 5.7. The player must find the location that the whole body of the player can be seen. When the player finds the perfect location for training, the player must do "Yes" to begin the calibration.

Figure 5.7: Stand by!!

The system calibration will make a new position for the player body. It will bring to the center position as Figure 5.8. If the calibration is done, there will be a message "Ready" as in Figure 5.8.

Figure 5.8: Ready!!

After that, an object will appear as in Figure 5.9. The object means to be the correct position of main joint movement in each maneuver. The player must remember the animation and repeat the performance in this step.

Figure 5.9: Go!!

After the player break the object as Figure 5.10 and the system will analyze the data. The system will recognize the player performance.

Figure 5.10: Analyzing

Finally, the result screen appears as in Figure 5.11. It shows both average percentages of similarity and percentage of similarity at all moment of the player data. The player can rotate camera view by doing "Move". If the player wants to skip and practice another maneuver, the player can do "Yes".

Figure 5.11: Result Screen

5.3 Multiple Maneuvers Training

Multiple maneuvers training mode is designed to allow the player to practice a sequence of diferent maneuvers in one setting. This will make the player understand how to apply the maneuvers practiced in the single maneuver training mode in a real situation. In this mode, the system will randomly select a series of maneuvers. There is no animation. The multiple maneuvers training focuses on combining diferent maneuvers in a performance. This training simulates the situation when the player is in the real combat and the player has to use diferent maneuvers due to multiple targets which appear in diferent positions during the training.

When the player starts the training, There will be a message of maneuver as in Figure 5.12. The player has to be ready at this moment.

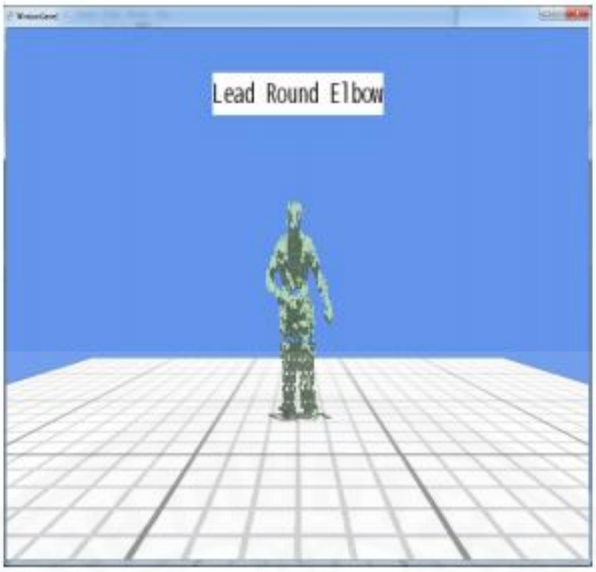

Figure 5.12: Multiple Maneuvers Training Ready Screen

The player has to perform continuously. The result of previous maneuvers will show on the bottom as in Figure 5.13.

When the player completes all random maneuvers, the system will show all percentage and average percentage as in Figure 5.14.

5.4 Summary

The instructional support system for Muay Thai is the tool to understand the basic Muay Thai maneuvers. The player practices the maneuver by following the animation in single maneuver training mode. The animation is captured from the professional Muay Thai master. The object that the player punches or kicks is the position of opponent where the player should hit, it would be opponent's neck, chest, arm or leg in real combat. When the player clearly understands all of the maneuvers in Single Maneuver Training mode, the player should try multiple maneuvers training mode which makes the player react to random maneuver as fast as in real combat. If the player can make a high score in this mode, that player has already understood the Muay Thai maneuvers.

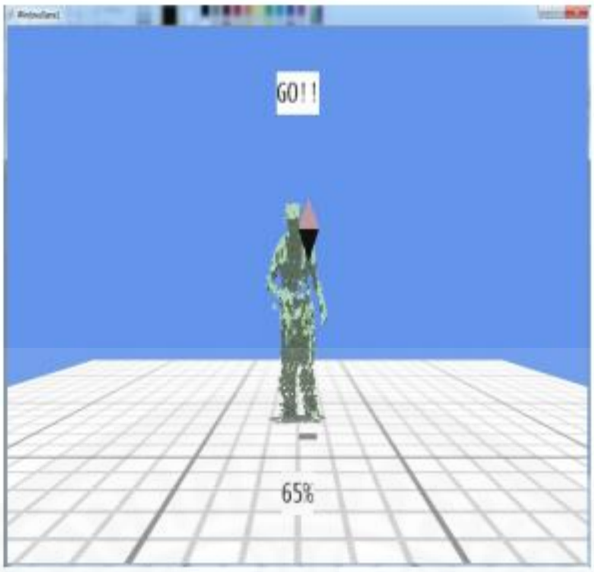

Figure 5.13: Multiple Maneuvers Training Perform Screen

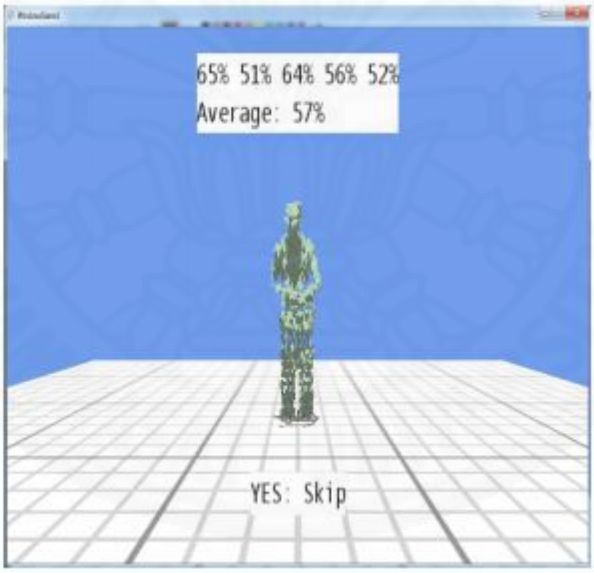

Figure 5.14: Multiple Maneuvers Training Result Screen

Chapter 6
Conclusion

When we start testing the Kinect, we realize that we can not make full experience of Muay Thai maneuvers which have a lot of rotation movement because the out of the line of sight problem. The out of the line of sight makes a big problem in Kinect skeletal tracking. It makes difculty detect the position of a body part. Kinect uses infrared to target all object into depth data. Then, there is skeletal recognition to detect player skeletal from depth data. When there is a body part that stay behind another body part, it will not appear in depth data. Consequently, this issue makes a false skeletal tracking. Therefore, we have created our proposed algorithms to solve this problem as shown in Chapter 3. We used an angle of rotations, depth information and boundary of a silhouette to help adjust the position of joints. We apply our algorithm to standard Muay Thai maneuvers and able to bring the skeleton detection accuracy
up from 51% to 77%

We have made totally 12 types of maneuver similarity measurements as explained in Chapter 4. In Table 4.1, HMM with Kinect+ is the best maneuver similarity mea- surement which makes 88.95% of database maneuvers accuracy. HMM with Kinect+ is the maneuver similarity measurement that we use in our *Instruction Support System f or M uay T hai*. We also analyze the database maneuvers accuracy of all maneuver similarity measurements. The data proof that our proposed algorithms in chapter 3 makes maneuver recognition improvement significantly in leg maneuvers.

We have successfully made *Instruction Support System f or M uay T hai* as shown in Chapter 5. There are 2 main training mode which are single and multiple maneuver training. In single maneuver training, the player selects a maneuver and practice with the professional Muay Thai maneuver animation. In multiple maneuvers training, the program will random maneuvers and player practice continuously without the anima- tion. There are totally 24 maneuvers that we have captured since in Chapter 3. The player can improve Muay Thai maneuvers by performing the maneuver to be similar to the animation as shown in Figure 5.6. The player can also see the similarity between player maneuver and the animation as shown in Figure 5.11.

References

[1] Lichao Zhang, A Kinect based Golf Swing Score and Grade System Using GMM and SVM, 2012 5th International Congress on Image and Signal Processing (CISP 2012)

[2] Connsynn Chye, Tatsuo Nakajima, Game based approach to learn Martial Arts for beginners, 2012 IEEE International Conference on Embedded and Real-Time Computing Systems and Applications

[3] Suwichai Phunsa, Nawuttagorn Potisarn, Suwich Tirakoat, Edutainment - Thai Art of Self-Defense and Boxing by Motion Capture Technique, International Conference on Computer Modeling and Simulation

[4] Perttu Hämäläinen,Tommi Ilmonen,Johanna Höysniemi, Mikko Lindholm, Ari Nykänen, Martial Arts in Artificial Reality, CHI 2005 PAPERS: Enhancing Virtual Spaces and Large Displays, April 2-7 Portland, Oregon, USA

[5] "Muay Thai". https://en.wikipedia.org/wiki/Muay_Thai

[6] "Mae Mai Muay Thai The Art of Self-Defense", Kukdej Kantanara

[7] https://www.facebook.com/MuchzhimaMMA

[8] Microsoft Corp. Kinect for XBOX 360. Redmond WA.

[9] http://accord-framework.net/

[10] http://www.extremeoptimization.com/QuickStart/CSharp/ClusterAnalysis.aspx

Appendix

Appendix A
List of Publications

A.1 International Conference

1. Ketchart Kaewplee, Nirattaya Khamsemanan, and Cholwich Nattee, "Muay Thai Posture Classication using Skeletal Data from Kinect and k-Nearest Neighbors," *In proceedings of the International Conference on Information and Communica- tion Technology for Embedded Systems (ICICTES 2014)*, Ayutthaya, Thailand, 23-25 January 2014

2. Ketchart Kaewplee, Nirattaya Khamsemanan, and Cholwich Nattee,"A Rule-based Approach for Improving Kinect Skeletal Tracking System with an Application on Standard Muay Thai Maneuvers," *In proceedings of the International Conference on Soft Computing and Intelligent Systems (SCIS 2014) and 15th International Symposium on Advanced Intelligent Systems (ISIS 2014)* , Kita-Kyushu, Japan. 3-6 December 2014