การสรุปใจความสำคัญของวีดิทัศน์การฝึกซ้อมกอล์ฟโดยใช้การตรวจหาวงสวิงจริง

นายพีร โชติมนัส

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ
ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์
คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2554
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

GOLF PRACTICE VIDEO SUMMARIZATION USING REAL SWING DETECTION

Mr. Peera Chotimanus

A Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of Master of Science Program in

Computer Science and Information Technology

Department of Mathematics and Computer Science

Faculty of Science

Chulalongkorn University

Academic Year 2011

Thesis Title	GOLF PRACTICE VIDEO SUMMARIZATION USING REAL SWING DETECTION
By	Mr. Peera Chotimanus
Field of Study	Computer Science and Information Technology
Thesis Advisor	Assistant Professor Nagul Cooharojananone, Ph.D.
Thesis Co-advisor	Suphakant Phimoltares, Ph.D.

Accepted by the Faculty of Science, Chulalongkorn University in Partial Fulfillment of the Requirements for the Master's Degree

……………………………………………….. Dean of the Faculty of Science

(Professor Supot Hannongbua, Dr.rer.nat.)

THESIS COMMITTEE

……………………………………………….. Chairman

(Assistant Professor Rajalida Lipikorn, Ph.D.)

………………………………………….……. Thesis Advisor

(Assistant Professor Nagul Cooharojananone, Ph.D.)

………………………………………….……. Thesis Co-advisor

(Suphakant Phimoltares, Ph.D.)

……………………………………………….. External Examiner

(Assistant Professor Panjai Tantasanawong, Ph.D.)

พีร   โชติมนัส: การสรุปใจความสำคัญของวีดิทัศน์การฝึกซ้อมกอล์ฟโดยใช้การ
ตรวจหาวงสวิงจริง. (GOLF PRACTICE VIDEO SUMMARIZATION USING
REAL SWING DETECTION) อ. ที่ปรึกษาวิทยานิพนธ์หลัก : ผศ. ดร. นกุล คูหะโรจ
นานนท์, อ. ที่ปรึกษาวิทยานิพนธ์ร่วม : อ. ดร. ศุภกานต์ พิมลธเรศ, 60 หน้า.

กล้องวีดิทัศน์เป็นอุปกรณ์สำคัญเพื่อช่วยนักกีฬาในการพัฒนาความสามารถในหลาย
กีฬาโดยเฉพาะกีฬากอล์ฟ บ่อยครั้งที่นักกีฬากอล์ฟใช้กล้องวีดิทัศน์เพื่อบันทึกวงสวิงใน
ช่วงเวลาฝึกซ้อมเพื่อใช้ในการวิเคราะห์วงสวิงของตนเอง ในการวิเคราะห์วงสวิงกอล์ฟนักกีฬา
กอล์ฟส่วนใหญ่จะให้ความสนใจที่วงสวิงจริง เพราะว่าวงสวิงจริงสามารถบ่งชี้ถึงปัญหาของวง
สวิงของนักกีฬากอล์ฟได้อย่างแท้จริง อย่างไรก็ตามในวีดิทัศน์การฝึกซ้อมกอล์ฟนักกีฬากอล์ฟ
อาจไม่ได้บันทึกแต่วงสวิงจริงเท่านั้น ดังนั้นเพื่อที่จะย่นระยะเวลาในการวิเคราะห์วงสวิงของ
นักกอล์ฟ ระบบการสรุปใจความสำคัญของวีดิทัศน์การฝึกซ้อมกอล์ฟจึงได้ถูกพัฒนาเพื่อ
อำนวยความสะดวกให้กับนักกอล์ฟในระหว่างการวิเคราะห์วงสวิง วิทยานิพนธ์ฉบับนี้
นำเสนอวิธีการในการตรวจหาวงสวิงจริงจากวีดิทัศน์การฝึกซ้อมกอล์ฟที่บันทึกจากมุมกล้อง
ทางด้านหน้า วงสวิงจริงสามารถถูกตรวจหาได้โดยการตรวจหาการขึ้นไม้และการตีลูกของ
นักกอล์ฟในระหว่างการฝึกซ้อม วิธีการที่นำเสนอนี้ประกอบด้วยห้าขั้นตอนหลักคือ การ
ตรวจหาหัวไม้กอล์ฟ การตรวจหาลูกกอล์ฟ การตรวจหาการขึ้นไม้ การตรวจสอบของผลลัพธ์
จากการตรวจหาการขึ้นไม้ และการตรวจหาการตีลูก หลังจากทดลองวิธีนี้กับวีดิทัศน์การ
ฝึกซ้อมกอล์ฟที่ถูกบันทึกจากกล้องหลายชนิดจากมุมกล้องทางด้านหน้า ในสภาพแสงที่
หลากหลาย ผลการทดลองได้แสดงให้เห็นว่าวิธีการในการตรวจหาวงสวิงจริงที่วิทยานิพนธ์
ฉบับนี้ได้นำเสนอนั้น สามารถตรวจจับวงสวิงจริงได้อย่างมีประสิทธิภาพ ค่าความแม่นยำและ
ค่าความครบถ้วนของระบบโดยเฉลี่ยเท่ากับร้อยละ 90.78 และ 92 ตามลำดับ

ภาควิชา ..คณิตศาสตร์และวิทยาการคอมพิวเตอร์...	ลายมือชื่อนิสิต ….................................
สาขาวิชา ….................วิทยาการคอมพิวเตอร์และ	ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก
เทคโนโลยีสารสนเทศ…....….........................	…...
ปีการศึกษา …..................2554…..................	ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์ร่วม
	…..................................…..............

5273626523 : MAJOR   COMPUTER SCIENCE AND INFORMATION TECHNOLOGY

KEYWORDS : VIDEO SUMMARIZATION / VIDEO INDEXING / GOLF PRACTICE VIDEO / GOLF SWING ANALYSIS / REAL SWING DETECTION

PEERA CHOTIMANUS : GOLF PRACTICE VIDEO SUMMARIZATION USING REAL SWING DETECTION. ADVISOR : ASST. PROF. NAGUL COOHAROJANANONE, Ph.D., CO-ADVISOR : SUPHAKANT PHIMOLTARES, Ph.D., 60 pp.

Video camera is often used in many sports especially in golf to help athletes to improve their performance. Golfers often use video cameras to record their golf swing during golf practice session for golf swing analysis. In golf swing analysis, most golfers only focus on real swing because only real swings can truly show the problems of golfer's swing. However, during golf practice session golfers may not record only real golf swing. Therefore, to shorten the process of golf swing analysis, golf practice video summarization method is developed in order to help golfers locate each real golf swing in golf practice video during golf swing analysis. This thesis proposed a method to summarize golf practice video recorded from face-on camera angle. Real swing can be detected by detecting takeaway and ball hit action of the golfer during golf practice session. The purposed method consists of five main processes which are golf club head detection, golf ball detection, takeaway detection, takeaway validation and ball hit detection. The result of experiments on golf practice video recorded with different cameras from face-on camera angle within different light conditions shows that this method can efficiently extract real swings. The precision and recall rate of the system are 90.78% and 92% respectively.

Department : ..Mathematics and Computer Science..	Student's Signature …..…….…..….…..
Field of Study : …..................Computer Science and	Advisor's Signature …..……..….……..
Information Technology………………………….…..	Co-advisor's Signature …..….…....…..
Academic Year : …………….……2011………..…….	

Acknowledgements

I would like to thank to my advisor Assistant Professor Nagul Cooharojananone for all the support of my study and research at the department of Mathematics and Computer Science, Chulalongkorn University. I would also like to thank Dr. Suphakant Phimoltares and the rest of my thesis committees Assistant Professor Rajalida Lipikorn and Assistant Professor Panjai Tantasanawong for their comment and guidance.

Lastly, I would like to express my love to my family and friends who always encourage and support me.

Contents

List of Tables

List of Figures

CHAPTER I

INTRODUCTION

The advance in development of data capturing and video processing has drawn athletes' attention to sport video analysis. Nowadays, video camera is used by not only professional, but also amateur athletes in various sports as a tool to analyze their performance, because using video is an only way that athletes can truly diagnose their flaws.

Golf is one of many sports that frequently use video analysis during training session. It is a sport that required a lot of practice and effort to master like any other sports, but the key of success in golf is an achievement of a good golf swing and in order to achieve a good golf swing, golfers must know their flaws and what need to be done to fix those flaws. In golf competition the golfer who has the lowest score or number of strokes wins, which also means that a golfer who makes the least mistakes has a higher chance of winning.

Therefore practicing session is very important for every golfer to improve their performance. Video camera is an ideal tool for golf swing analysis to help improve their golf swing. It is often used to record golf swing by both professional and amateur golfers during the golf training session to analyze their golf swing. A video recorded during practice session may contain a lot of actions such as pre-shot routine, practice swing and real swing, however, when analyzing golf swing, most golfers mainly focus on real swing. Because real swing is the only swing that truly shows the problems of the golfer's swing, it is the swing that has intention to hit the golf ball. On the other hand, the purpose of practice swing is to relax and loosen up, hence practice swing cannot truly show the problems of the golfer's swing, because during practice swing golfer has no intention and effort to hit the golf ball. Therefore the concept of video summarization can be applied to extract the start of each real swing in golf practice video to allow the golfer to quickly navigate to each real swing, which facilitates and accelerates the process of swing analysis during practice session (see Fig 1.1).

Golf Practice Video
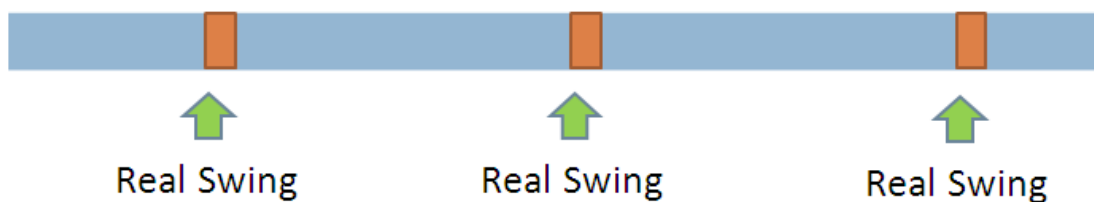
Figure 1.1: Golf practice video may contain a lot of actions, but golf swing analysis focuses on real swing. Moreover, real swing also happens in a very short period of time about 2-3 seconds which may take some times for a golfer to search for each real swing during golf swing analysis. Therefore, in order to facilitate golfers to analyze their golf swing during practice session more efficient, the video summarization technique is applied to extract only real swing.

(a)   (b)

Figure 1.2: Camera angle used in golf swing analysis. (a) Face-on camera angle. (b) Down the line camera angle.

In golf video analysis, there are two major camera angles that golfers use to record their swings which are down the line and face-on camera angle (Rice, 2009; Breed and Midland, 2008). For face-on, the camera is position in front of the golfer so that it is facing toward the golfer's shoulder (see Fig. 1.2). For down the line, the camera is position behind the golf ball so that it is looking at the right side of the golfer through the target as (see Fig. 1.2).

This thesis proposes the method to summarize golf practice video recorded in face-on camera angle in golf driving range environment by detecting every real swing in the golf practice video. Searching for a real swing shot in golf practice video might take a lot of time during swing analysis because golf practice video may contains a lot of action and real swing also happens very fast; about 2-3 seconds. Using this method, every real swing in golf practice video is extracted by detecting frame index of beginning and ending point of each real swing allowing a golfer to be able to quickly access to each real swing. This method will allow golfers to quickly review and analyze their swing efficiently. The beginning point of each real swing or address position is estimated from detected takeaway action (see Fig. 1.3) and the ending of each real swing or swing finish is estimated from detected ball hit action. Takeaway detection and ball hit detection method are based on the routine of most golfers during real swing (see Fig. 1.4). The routine of real swing action always starts with the address position followed by takeaway, hit of golf ball and ends with swing finish. At address position, the golfers place a golf club next to a golf ball preparing to swing and then they begin to takeaway. Subsequently, takeaway action begins with the golfers taking their golf club away from a golf ball to start their swing. Thus, in order to detect take away action, it is important that the system has to locate the position of golf club and golf ball.
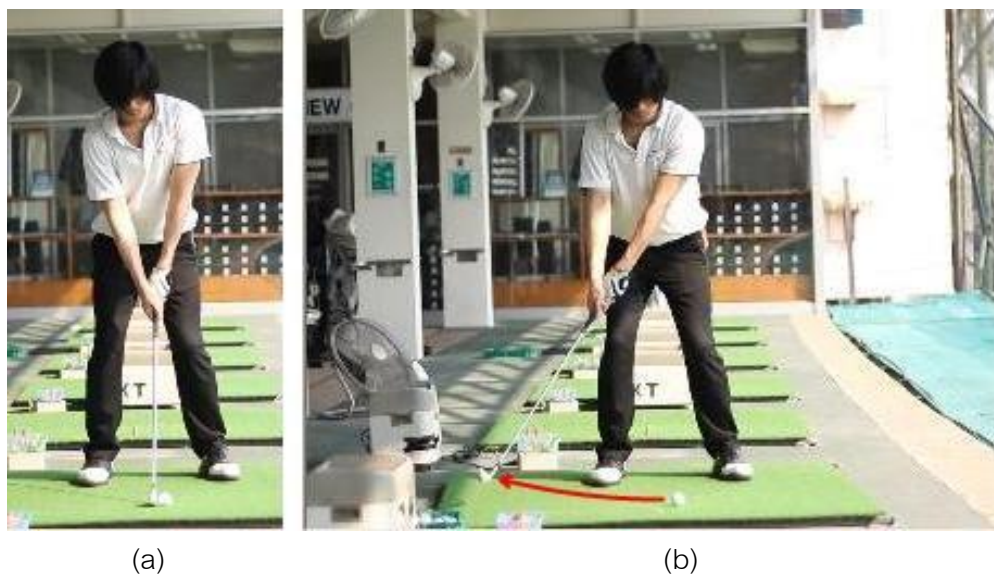
Figure 1.3: Takeaway action in Face-on camera angle. Takeaway (b) is the movement when the golfer takes the golf club away from the golf ball at the very first part of the swing starting at the address position (a).
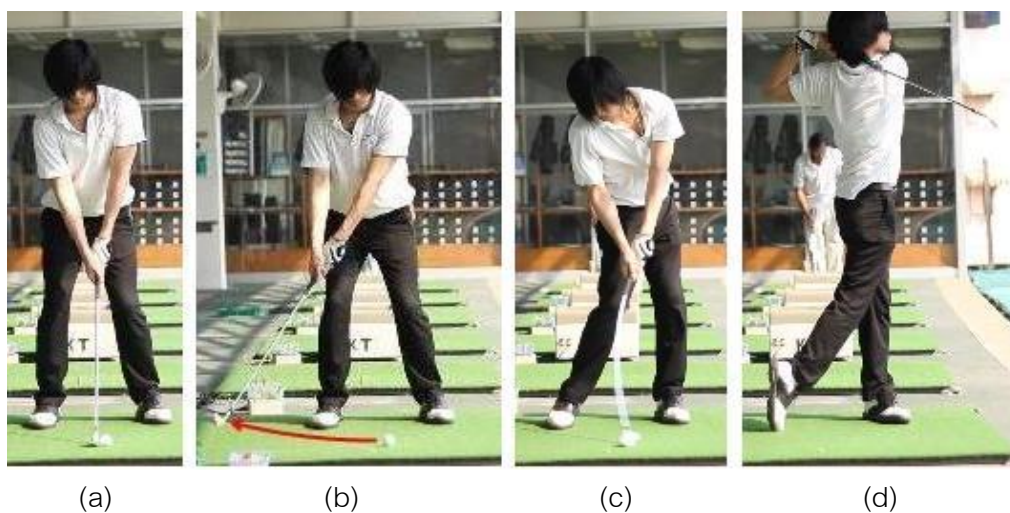
Figure 1.4: Real swing action in sequence. Real swing always starts with address position (a) then takeaway (b), hit (c) and ends with swing's finish (d) respectively.

Many researches in the area of sport video have long been active. In the area of golf, there also are many researches such as visual display of golf video providing more visual for better swing analysis (Kelly et al., 2010; Karliga and Hwang, 2007), visual display of golf ball trajectory (Zupancic and Jaklic, 2009) and impact (Kim et al., 2011) to help golfers analyzing their control of golf ball. But there are only a few researches about golf video indexing and summarization. C. Jung et al. (Jung and Kim et al., 2009) have developed a method to summarize broadcast golf videos by extracting player information from captions in broadcast golf videos. The caption in broadcast golf video is detected by using edge detection and clustering method of color patterns. D. Ring et al. (Ring and Kokaram, 2007) have researched and developed a method to summarize practice video using intrinsic motion. The practice video is summarized by detecting ready-then-action pattern in mostly static background environment. This system can be used with many kind of sport that involves repetition of ready-then-action pattern including golf, but this method cannot used to precisely extract real swings in golf practice video.

The rest of this thesis is organized as follows: Chapter 2 describes theoretical background of this research. Chapter 3 describes detail of all processes of the purposed method. Chapter 4 describes and discusses experimental result of this research. Finally conclusion and future work are described in chapter 5.

1.1 Problem Formulation

Golf swing analysis mostly focuses on real swing, but golf practice video may contain a lot of actions such as pre-shot routine, practice swing and real swing. Moreover, real golf swings scatter over golf practice video and real swings happen in a very short period of time which may take some times for a golfer to search for each real swing during golf swing analysis. Therefore, in order to facilitate golfers to analyze their golf swing during practice session more efficient, the video summarization technique is applied to extract only real swing to provide golf practice video summarization which will

allow a golfer to quickly access to each real swing and accelerate the process of golf swing analysis.

1.2 Objectives

The objectives of proposed research are the following:

1. To facilitate and accelerate the processes of golf swing analysis by allowing a golfer to quickly navigate to each real swing.

2. To propose a method used to extract the real swing in golf practice video which can be applied in embedded system such as digital camera or smart phone which the navigation control is not convenient.

3. To promote research in the area of sport especially in the area of golf.

1.3 Expected Outcomes

This proposed research is expected to implement a method which accurately extracts real swing in golf practice video, to help golfers during the process of golf swing analysis by providing video summarization of real golf swing in golf practice video.

1.4 Scope of the Work

The proposed method is limited to the following:

1. This method is designed for only golf practice video recorded in face-on view. For face-on view, the camera is position in front of the golfer so that it is facing toward the golfer's shoulder.

2. This method is designed for golf practice video recorded in sufficient lighting environment. The method can process golf practice video recorded in the morning, afternoon, evening, and at night with sufficient lighting.

3. This method is designed for golf practice video recorded with a video camera mounted on the tripod or a stabilized video camera.

1.5 The Benefits of Research

This research can be applied to implement automatic golf video summarization application to facilitate golfers by allowing them to easily locate each real swing in golf practice video which allowing their golf swing analysis during their practice session to be much more efficient. Moreover this research can be applied to use in embedded system, whose navigation control is not very convenient, to allow navigation to each real swing to be faster and more convenient.

CHAPTER II

THEORETICAL BACKGROUND

Golf practice video summarization involves with many techniques in various fields such as image processing techniques, video summarization techniques, and object detection and tracking techniques. To summarize golf practice videos video summarization techniques are applied. Video summarization techniques can be classified into three categories based on information sources which are text, image and audio. And the content of video can be roughly categorized into two categories which are event based content and informative content. In addition, this thesis also utilizes a lot of image processing techniques to process images in golf practice videos in order to identify and locate important of objects in golf practice video which are the golf club and the golf ball. These image processing techniques are frame differencing, Canny edge detection, morphological operation, Hough transform. Moreover, to locate real swings in golf practice video object tracking techniques are applied to track golf club and golf ball. Tracking of golf club is to ensure that during takeaway action the detected golf club is forming the correct trajectory. On the other hand tracking of golf ball is to ensure that during takeaway action the golf ball is not moving.

The outline of this chapter is organized as follows: Section 2.1 describes about video summarization, video summarization technique and video summary visualization Section 2.2 describes about image processing techniques, and Section 2.3 describes about object tracking techniques.

2.1 Video Summarization

The amount of digital video is growing at a very fast rate, due to the advancement of new technology in storage and communication. Nowadays, users require assistant to help them manage and access to video content. Moreover, there are limited tools and applications to help describe, summarize video data to users. Therefore, researches in video summarization are needed to develop methods or

techniques to help users access their video content faster and easier. The fundamental of video summarization is to produce the condensed version of a full length video by identifying the most important and relevant contents in the video.

2.1.1 Video Summarization Techniques

Video summarization technique can be roughly categorized into three categories based on information source used in the process of video summarization which are text, image, and audio.

2.1.1.1 Text

Text contains in the video is a valuable information. Text appeared in the video can be in the form of captions or subtitles. Text can be used in video summarization to identify important events or detail information about the contents in the video such as subtitles in the movies, captions in broadcasting of sport video, or captions in broadcasting of news. Captions in sport video contain important information such as score, player, or times. Information from the captions can be used to identify important events in sport video, for example, player captions in golf broadcasting video can be used to identify the player and score during each hole, score caption in basketball broadcasting video or football broadcasting video can be used to identify score of each team and time when each team earned score.

2.1.1.2 Image

Image is one of the main elements of video. Sequence of images in the video also can identify important events. Image features in video summarization can be used to identify object, event, and motion (Money and Agius, 2008). Change or motion of color can be used to identify shots, panning shots in the movies. Detection of objects such as left objects, people, and vehicle can be identified as important events in surveillance video summarization.

2.1.1.3 Audio

Audio is also one of the main elements of video. Audio can be used to highlight in many sports by detecting excitement sound of audiences such as after a great shot in broadcasting of golf competition, after a home run hit in broadcasting of baseball match, or after a goal is scored in broadcasting of soccer games. Moreover specific knowledge about particular video can be used to refine video summarization. For example double whistling in soccer games can be used to identify foul events (Xu et al., 2003).

However, video summarization techniques can be made up from combination of other techniques. For example audio and image features can be used in combination to detect highlights in golf, baseball, and soccer video. By using audio and image features the system is able to automatically identify type of sport from sport video whether it is excitement of audience highlights from golf, baseball or soccer sport (Xiong et al., 2005).

2.1.2 Video Summary Visualization Type

The summary of video can be display in different forms such as text, graphic or video. However, different forms of summary share the same goal which is to make redundant contents in the video more useful and convenient for users to access. The type of video summary visualization can be roughly categorized into two categories which are static visualization and dynamic visualization (Taskiran et al, 2006).

2.1.2.1 Static Visualization

Static visualization of video summary is still images, key frames, or graphics to represent the important parts in the original video which mostly automatically selected. Static visualization of video summary can be presented in the form of story board or table of contents fashion (Xiong et al., 2006) which may be associated with description text or timestamp. The table of contents of the book is a great example of static visualization. It greatly facilitates the accessibility of the book. The advantage of static visualization is that it makes the task of browsing and retrieval much more efficient. It

helps users to quickly access to parts of video of their interest. Users do not need to use fast-forward and rewind function, which is very time-consuming, to search for parts of the video of their interest.

2.1.2.2 Dynamic Visualization

Dynamic visualization of video summary is a video segment or a set of video segments to represent the most important parts in the original video which may be automatically generated. Dynamic visualization of video summary is represented in the form of a shorter version of original video which preserves both motion and audio element of the original video therefore most users may find that a dynamic visualization of video summary is more interesting and attractive. The advantage of dynamic visualization of video summary is that it can display motion which static visualization lacks.

Dynamic visualization also can be represented in the form of condensed video segment. In (Li et al., 2008) produced video summarization by condensing activities occurred separately in time close together.

2.1.3 Type of Video Content

It is important to consider about type of video contents when choosing the video summarization techniques to apply. The content in video can be roughly categorized into two categories based on type of contents in the video. These two categories are event based content and informative content.

2.1.3.1 Event based Content

Videos with event based content can be split into video segments based on identifiable important events in the videos. Examples of event based content videos are sport video, surveillance video. The important events in sports video can be goals, fouls, attacking moments in soccer broadcasting video or a great shot, hole-in-one in golf broadcasting video.

2.1.3.2 Informative Content

Videos with informative content are the video that are not easily split into video segments based on important events because the content in the video is equally important such as product review video, home video or presentation video. However, informative content video can be summarized by using the information of important parts provided from user. In order to summarize informative content video, users are required to provide annotation of important frames or video segments in the video.

2.2 Image Processing Technique

The takeaway detection system provides indexes of the beginning of real swing in golf practice video. To locate the beginning of real swing, the system has to locate the golf ball and the golf club. The system to locate golf ball and golf club consists of five main processes which are golf club head detection, golf ball detection, take away detection, take away validation and hit detection. The important image processing techniques involve in these processes are frame differencing and morphological operation.

2.2.1 Frame Differencing

Frame differencing is the very simple background subtraction method. It is one of image processing technique that is frequently applied to detect moving objects in the video. To efficiently extract motion in the video with frame differencing technique, it requires videos to be recorded from stabilized camera.

Frame differencing technique is widely used to detect moving objects in the video. In (Wei et al., 2006), frame differencing technique is used to detect motion region of table tennis players and table tennis ball. In (Kosit et al., 2010) frame differencing technique is used to detect motion region of tennis ball. Motion region can be extracted by fetching three consecutive frames, finding frame difference between first frame and second frame, finding frame difference between second frame and third frame. After
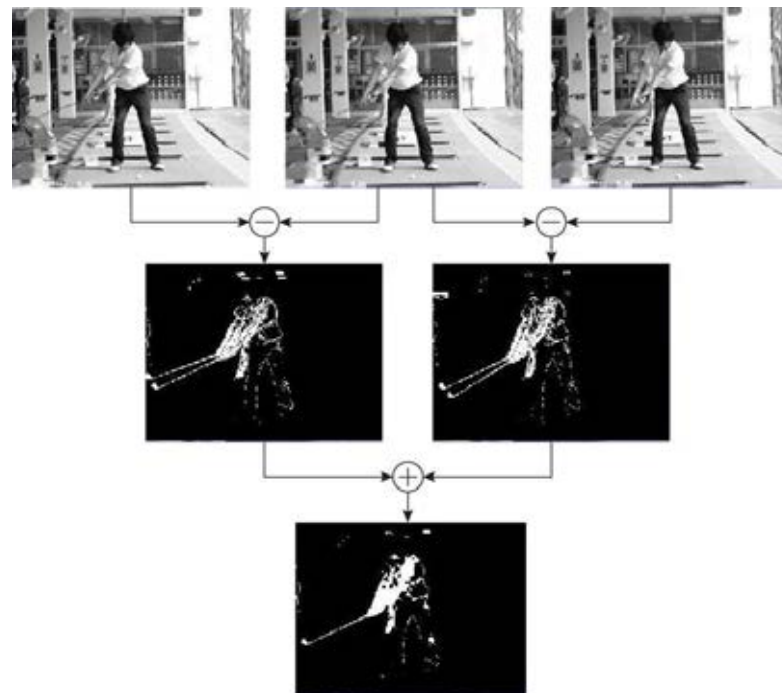
Figure 2.1: Motion region can be detected by using frame differencing between three consecutive frames (top) and applying AND operation to the result of frame differencing (middle) and finally the result of motion region is obtained (bottom).

that using AND operation between two results from frame differencing to locate motion region (see Fig. 2.1).

2.2.2 Morphological Operation

Morphological operation is a very useful operation. It is often used in the area of computer vision to removing noises, merging small elements from binary or grayscale images. The two basic operation of morphological operation are dilation and erosion. The major task of dilation is to merge small gaps, or to remove noises in the image. On the other hand, erosion is used to shrink or narrow elements in the image.

Morphological closing and opening are the application of dilation and erosion. The opening operation is used for eliminating extending elements from the object, while the closing operation is used for filling small gaps. The equations of opening and closing operation are as follows:
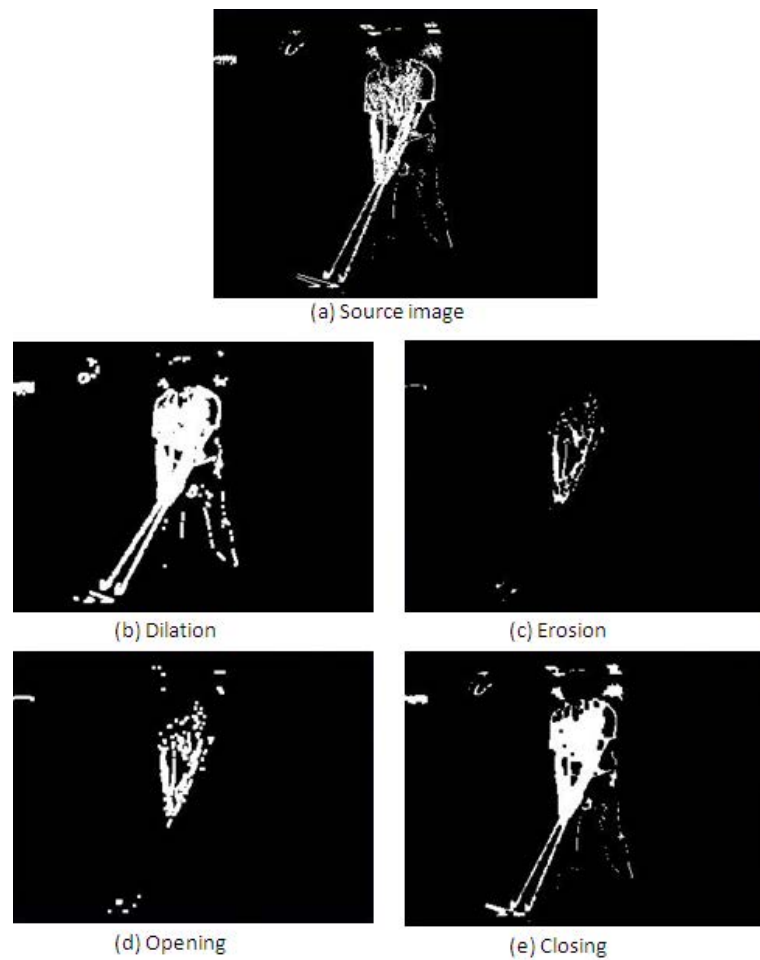
Figure 2.2: The example of morphological operation. Result of dilation operation applied to source image (b). Result of erosion operation applied to source image (c). Result of opening operation applied to source image (d). Result of closing operation applied to source image (e).

$$Opening \; = \; (A \ominus E) \oplus E$$

$$Closing \; = \; (A \oplus E) \ominus E$$

$\oplus$ is dilation operation and $\ominus$ is erosion operation respectively.

2.3 Object Tracking

Object tracking is one of the most important tasks being used in the field of computer vision. It is strongly beneficial to many areas. Object tracking is used in the area of traffic monitoring as a tool to estimate traffic flow to provide real time report of traffic condition. In the area of human and computer interaction, object tracking is often used to detect gesture as the input to interact with computer. Moreover object tracking is also used in the area of video summarization and indexing to detect important events generated from motion of objects such as detecting of suspicious activities in video surveillance system, detecting of tennis ball's trajectory to provide better visualization.

CHAPTER III

PURPOSED METHOD

In this chapter, explanation of system of purposed method is written in detail. The purposed method is designed to extract real swing actions in golf practice video by detecting takeaway actions and ball hit actions which will be used to create video summarization of golf practice session. This method facilitates golfer to quickly access all real swings in golf practice video which allows golfers to analyze their golf swing much more efficient during golf practice session; spending less time to locate each real swing. Detection of takeaway actions and ball hit actions mainly relies on the detection of two objects in golf practice video which are a golf club and a golf ball. The system of purposed method consists of five main parts which are golf club head detection, golf ball detection, take away detection, take away validation and bat hit detection as shown in Fig. 3.1.

The rest of this chapter is organized as follows: Section 3.1 explains golf club head detection process. Section 3.2 describes golf ball detection process. Section 3.3 describes takeaway detection process. Section 3.4 describes takeaway validation process. Section 3.5 describes ball hit detection process. Finally section 3.6 describes the results obtained from the purposed method for golf practice video summarization.
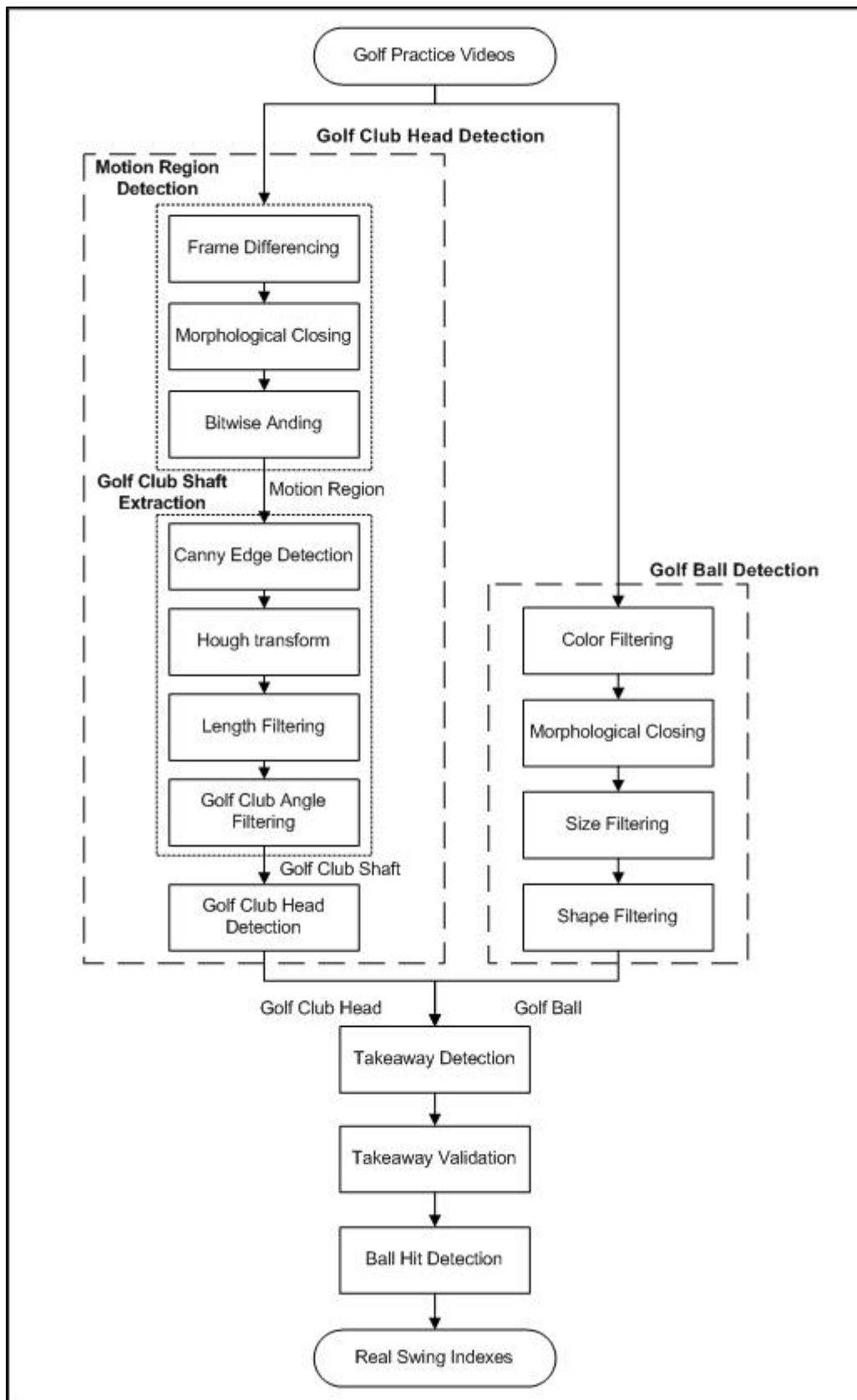
Figure 3.1: Block diagram of the proposed method

3.1 Golf Club Head Detection

The process of golf club head detection consists of three main processes which are motion region detection, golf club shaft extraction, and golf club head detection respectively. The concept of golf club head detection is to locate the motion region first. After the motion region is obtained, the golf club shaft can be located by finding straight lines and finally the golf club head can be identified.

To detect the golf club, all frames are first converted into grayscale image and the golf club is detected by locating motion or moving region which will contain the golf club. The process of golf club head detection consists of three steps which are motion region extraction, golf club shaft extraction, and golf club head detection respectively.

3.1.1 Motion Region Detection

In this sub section, the process of motion region detection is described. Motion region detection consists of three sub processes which are frame differencing, morphological closing and bitwise ANDing. The purpose of motion region detection is to detect the motion of the golfer swing in the current frame so that the system will be able to locate golf club shaft.

3.1.1.1 Frame Differencing

Frame differencing is the first process of motion region detection. It is one of image processing techniques to identify the difference between two frames in the video. The difference between two frames is used to detect the movement of the golfer swing which results as a moving region in the video frame. To identify the movement of the golfer, three consecutive frames are fetched (see Fig. 3.2). Each pair of consecutive frames in the practice video is then computed to identify the motion (see Fig. 3.3). In order to identify motion region in frame f_i, Frame differencing is applied as following: $Diff(f_{i-1}, f_i)$ and $Diff(f_i, f_{i+1})$, where $Diff(f_1, f_2)$ is frame differencing operation between frame f_1 and f_2. Then the results obtained from frame differencing are saved in binary images. The equation of frame differencing operation is as follows:

(a) frame f_{i-1}

(b) frame f_{i+1}

(c) frame f_i
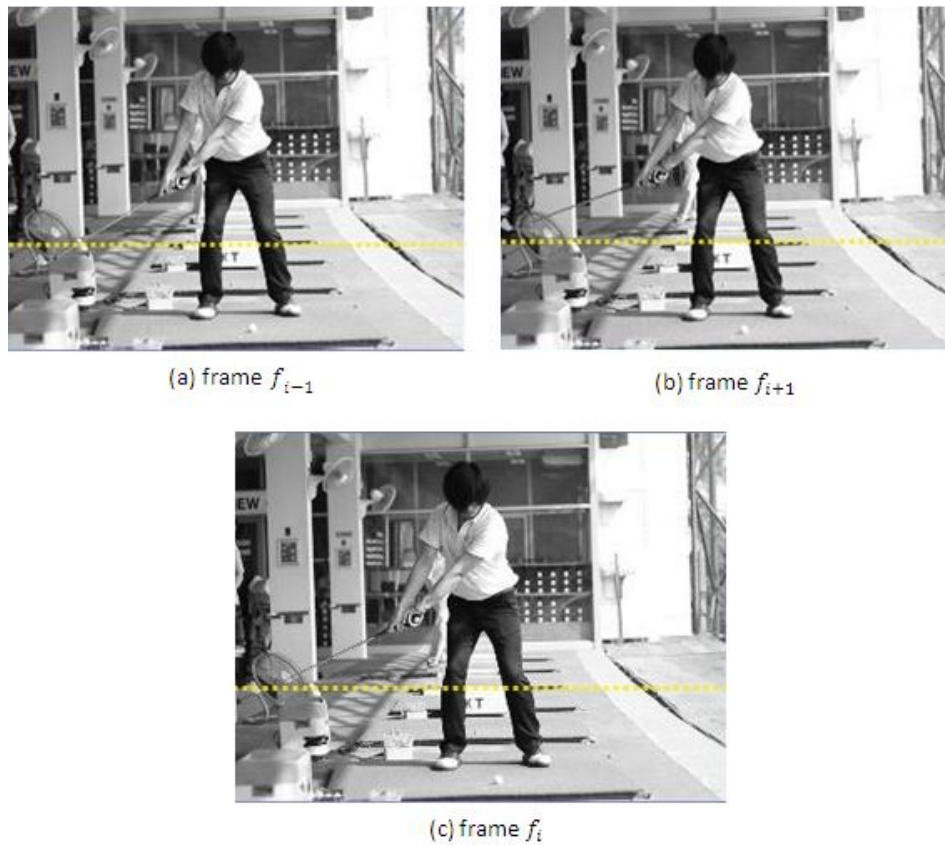
Figure 3.2: To find motion region three consecutive frames are fetched which are previous frame (a), current frame (c), and next frame (b).

$$Frame\ Difference = Diff(f_{i1}, f_{i2}) = |f_{i1} - f_{i2}| \qquad (1)$$

3.1.1.2 Morphological Closing

Morphological closing is the second process of motion region detection. After frame differencing operation is finished. The frame difference between the current frame and the previous frame is extracted. A morphological closing operation (Serra,1983) is then applied to two binary images to fuse small gaps (see Fig. 3.3). The equation of morphological closing operation is as follows:

$$Morphological\ closing = Mc(Df) = (Df \oplus E) \ominus E \quad (2)$$

where Df is binary image obtained from frame differencing process, E is 3x3 rectangular structuring element, $\oplus$ is dilation, and $\ominus$ is erosion.

$$Diff(f_{i-1}, f_i)$$

$$Diff(f_i, f_{i+1})$$

$$Mc(Diff(f_{i-1}, f_i))$$

$$Mc(Diff(f_i, f_{i+1}))$$

$$Mc(Diff(f_{i-1}, f_i)) \cap Mc(Diff(f_i, f_{i+1}))$$
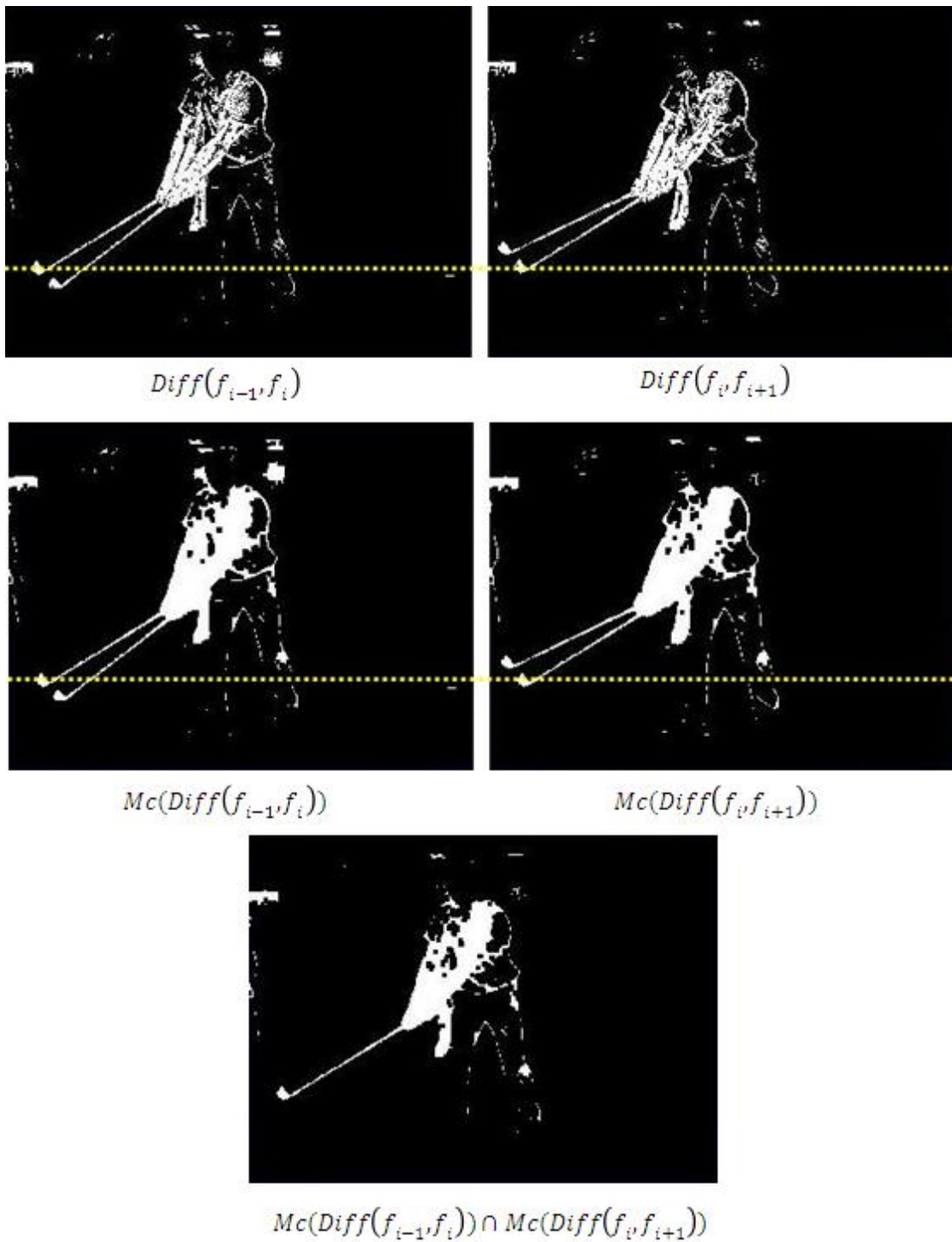
Figure 3.3: To locate motion region. (Top) Frame differencing of three consecutive frames is used to identify movement. (Middle) Morphological closing of the results from frame differencing method is used to fill small gaps. (Bottom) Bitwise AND operation is applied to obtain motion region.

3.1.1.3 Bitwise ANDing

To extract moving region in frame f_i, bitwise AND operation is applied to AND two binary images obtained from $Mc(Diff(f_{i-1}, f_i)$ and $Mc(Diff(f_i, f_{i+1})$ producing motion region which contains moving pixels in the current frame (see Fig. 3.3). Therefore, motion region can be described as the following equation:

$$Motion \ = \ Mc(Diff(f_{i-1}, f_i) \cap Mc\big(Diff(f_i, f_{i+1})\big) \quad (3)$$

where $Mc(...)$ and $Diff(...)$ is morphological closing and frame differencing respectively.

3.1.2 Golf Club Shaft Extraction

In this section, the process of golf club shaft extraction is described. Golf club shaft extraction consists of four sub processes which are Canny edge detection, Hough transform, length filtering, and golf club angle filtering. The purpose of golf club shaft extraction is to detect the golf club shaft in the motion region obtained in the first section so that the system will be able to locate golf club head in the next step.

3.1.2.1 Canny Edge Detection

Once motion region is obtained, Canny edge detection (Canny, 1986) is then used to identify the boundary of the motion region (see Fig. 3.4).

3.1.2.2 Hough Transform

Hough transform (Duda and Hart,1972) is a method used for finding straight lines, circles or simple shapes in an image. To identify line segments in the boundary of motion region obtained from canny operation, straight lines feature of Hough transform is used for to locate the golf club. After obtained motion region, Hough transform is then applied in order to identify line segments from the motion region's boundary because the shaft of the golf club looks like a straight line, therefore line segments are searched in each frame. Line segments obtained from each video frame is viewed as golf club

(a) Canny Edge detection. (b) Detected golf club shaft using Hough line transform
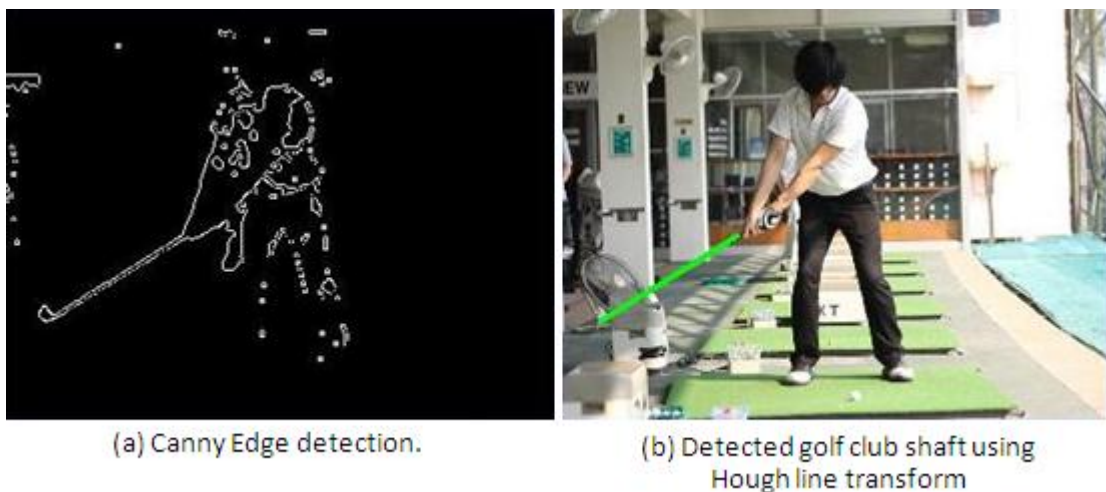
Figure 3.4: To identify golf club shaft from motion region, Canny Edge detection (a) is applied to motion region obtained from motion region detection process to identify edge of the motion region, then golf club shaft can be detected by using Hough line transform and filtering (b) to detect straight line from edge of motion region.

candidates if they meet several criteria from length filtering and golf club angle filtering (see Fig. 3.4).

3.1.2.3 Length Filtering

Length filtering is the third process of golf club shaft extraction. It is the process used to select the straight lines detected from Hough transform process whose length is close to the golf club. The video camera is often placed within a specific range from the golfer because in order to capture an analyzable golf practice video, the video usually be able to show the complete view of golf swing clearly especially all the complete path of golf club and the angle of golf club's face. In the experiments, the golf club length falls within the following range $T_{min} < Club\ Length < T_{max}$, where T_{min} is the minimum threshold of golf club length and T_{max} is the maximum threshold of golf club length. In the experiments, $T_{min} = Frame\ Height/4$ and $T_{max} = Frame\ Width/2$.
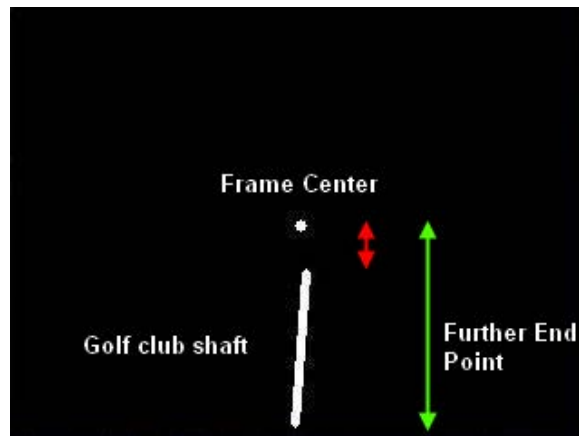
Figure 3.5: After golf club shaft is detected by Hough transform and filtering, the golf club head can be identified by using frame center as the reference point, the further end point to frame center is considered to be the golf club head.

3.1.2.4 Golf Club Angle Filtering

Golf club angle filtering is the fourth process of golf club shaft extraction. To prepare the golf clubs candidates for the takeaway detection, only the golf clubs is filtered by focusing on the golf club during takeaway action. The angle of the golf club refers to the clock position where the golf club is pointing to. During takeaway, golf club angle is always steep (Between 8 o'clock and 4 o'clock). Therefore only detected line segments with steep angle are considered to be golf clubs. On the other hand all detected line segments whose angle is flat are all discarded.

3.1.3 Golf Club Head Detection

Golf club head detection is the last step of golf club head detection process. The identification of golf club head based on the fact that golfer is always positioned at the center of the frame in most golf practice video in order to capture the complete path of golf swing. By positioning a golfer at the center of the frame, the center of golf swing is near the center of frame, so the center of the frame is used as the reference point to determine the position of the golf club head. The distance between center of the frame and two end points of line segment is computed by Euclidian distance. One end point of

(a) Wide range HSV   (b) Narrow range HSV
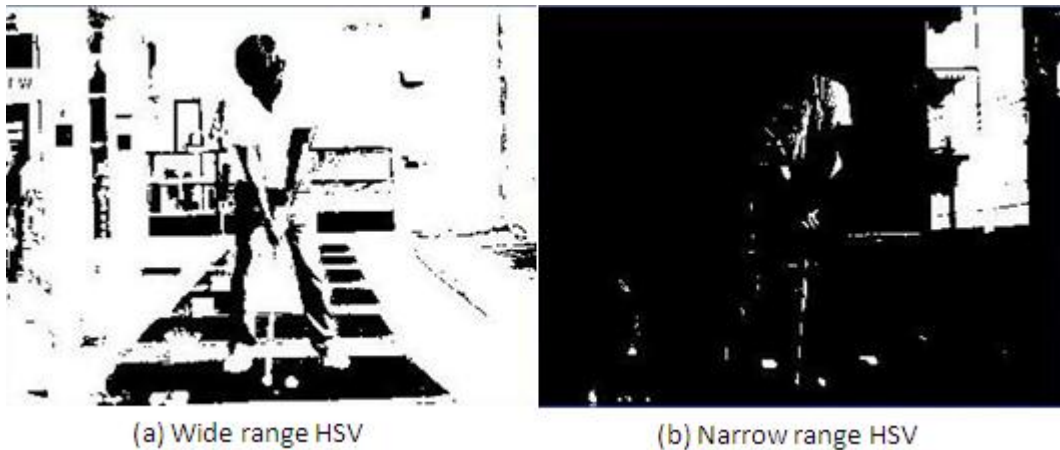
Figure 3.6: Golf ball is detected by color filtering technique. Filtering only white color in the image or using narrow range HSV may not detect the golf ball (b). Wide range HSV color filtering (a) is used extract near-white color with high mixture of whiteness which can detect the golf ball more efficient.

detected line segment which is further from the center of the frame is considered to be golf club head (see Fig. 3.5).

3.2 Golf Ball Detection

In this section the process of golf ball detection is described. The process of Golf ball detection consists of four processes which are color filtering, morphological closing, size filtering and shape filtering respectively. In order to detect takeaway actions, identification of the position of both golf club and golf ball are needed. To extract the golf ball color, size and roundness features are used for detection.

3.2.1 Color Filtering

HSV color filtering is used to extract white color which is a general color of most golf balls, and in order to detect golf balls under different light situations without often adjusting the filter's parameters, HSV color filtering has to cover wide range of HSV value to detect near white objects. To detect golf ball on green golf practice mat or grass under different light conditions, near-white color objects with high mixture of

(a) Golf practice video recorded under yellow neon light in the evening.

(b) Greyscale Image

(c) Detected near white color objects from greyscale image

(d) Golf ball detection under yellow neon light
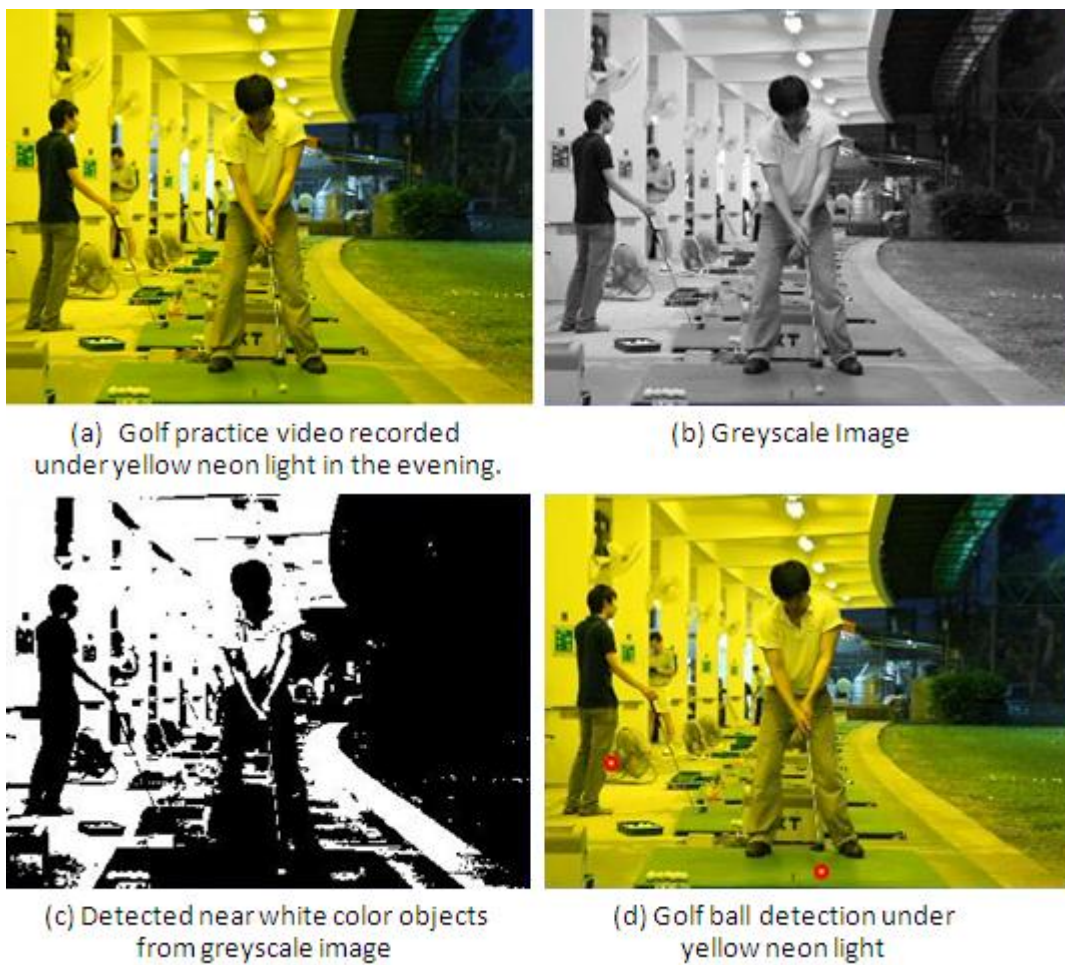
Figure 3.7: In case of detecting golf ball in golf practice video recorded under yellow fluorescent lighting in the evening, everything in the video looks more yellow under yellow fluorescent lamp (a). HSV filtering is used to extract white color from grayscale image (b) instead of color image to detect whiteness of the golf ball. After shape and size filtering the golf ball can be detected (c)(d).

whiteness are detected, except green color which is the color of golf practice mat and grass. Narrow range HSV filtering cannot be used to extract the golf ball in different light conditions (see Fig. 3.6). From the experiments the HSV value that are used for filtering in most videos is as following: $0 < H < 55 \cap 65 < H < 110, S < 27, V > 33$. The HSV value is written in the following format: H:0-255, S:0-100, V:0-100 and the white color is H:0, S:0, V:100. This HSV value can be used to detect golf ball in the morning,

Figure 3.8: After white color objects are located, golf balls can be detected by using size and shape filtering.

afternoon and evening, however, HSV value may be adjusted according to light situations to improve the accuracy of golf ball detection.

To detect golf ball at night under yellow fluorescent lighting, grayscale image is used instead of color image, because all objects in color image look yellow which causes HSV filtering from color image to be unable to accurately detect the golf ball. HSV filtering is used to detect white color objects from grayscale image, and then detected white objects are filtered by size and shape (see Fig. 3.7).

3.2.2 Morphological Closing

A morphological closing is then applied to fill small gaps. All white objects are view as golf ball candidates if they meet following criteria including size and roundness. The equation of morphological closing is the following:

$$Morphological\ closing\ =\ (Cf \oplus E) \ominus E \qquad (4)$$

where Cf is binary image obtained from color filtering process, E is 3x3 rectangular structuring element, $\oplus$ is dilation, and $\ominus$ is erosion.

3.2.3 Size Filtering

After near-white objects are detected, size filtering is applied to filter for objects whose size is near the size of the golf ball. The size of golf balls can be different in golf practice video with different camera setups. In the experiments, the resolution of the video is 320x240, the golf ball size is range from 10 to 30 pixels, depending on the camera setup.

3.2.4 Shape Filtering

After near-white color objects, whose size is near the size of the golf ball, are obtained shape filtering is used to detect the golf ball (see Fig. 3.8). The shape of the golf ball is circle, but a golf ball within different lightings is often not perfect circle. To identify roundness of the detected near white color object two thresholds are used to estimate roundness of the object. The first one is the ratio of the maximum width to the maximum height, if the ratio is close to one, the detected object is considered to be golf ball candidate. The maximum height is the distance between the topmost and the bottommost pixel of the object. On the other hand the maximum width is the distance between the rightmost and the leftmost pixel of the object. The second threshold is the result from the equation used to estimate roundness. The equation used to estimate roundness is as follows:

$$Roundness = \frac{4\pi A}{P^2} \qquad (5)$$

Where A and P are area and perimeter of a detected golf ball candidate respectively. The equation of perimeter is $2\pi r$.

In the experiments, the threshold of ratio of the maximum width to the maximum height of golf ball range from 0.6 and 1.4 and the threshold of the result from roundness equation is often higher than 0.56.

(a) Correct takeaway detection.   (b) False takeaway detection.
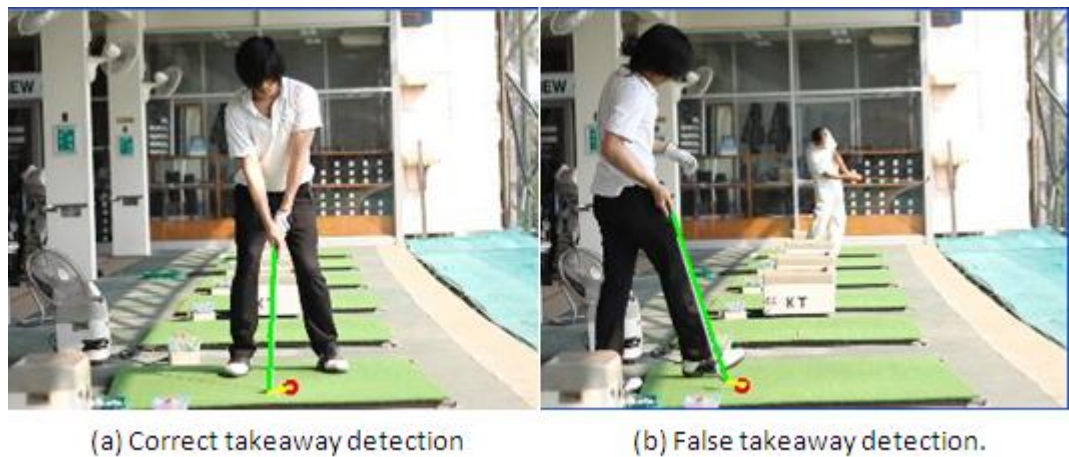
Figure 3.9: Other actions may be detected as takeaway. (a) Correct takeaway detection. (b) Tapping of golf ball action which causes false takeaway detection.

3.3 Takeaway Detection

After detection of both golf club and golf ball, the position of them are represented by points, this information is used to detect takeaway action. Takeaway is the very first step of backswing. To detect the early stage of takeaway action, the distance between golf club head during takeaway action and a golf balls is detected using Euclidean distance. If the distance between a golf club and a golf ball is less than the threshold then it is viewed as takeaway candidates. The threshold depends on the size of the video image. In the experiments, the resolution of the video frame is 320x240, the threshold is 40 pixels and the distance between Y-axis of the golf ball and the golf club head must less than 10 pixels to ensure that it is a takeaway because during the takeaway the golf club is taken away to the left or right side of the golf ball, not to the top side of the golf ball. Therefore every time the golf club head is placed next to the golf ball, it will be considered as takeaway candidate. A lot of action may be detected as a takeaway, for example, when the golfer taps a golf ball to reposition the golf ball or when the golfer walks pass a golf ball and the golfer's leg is detected as a golf club (see Fig. 3.9). Therefore, in order to avoid a lot of false detections, takeaway validation process is needed. Every time a takeaway candidate is detected, the following information will be kept for later validation.

Figure 3.10: To eliminate false detection, takeaway validation is needed. Takeaway validation keeps tracking of golf club and golf ball to ensure that the golf club is taken away forming a trajectory only in the same direction and the golf ball is not moving during takeaway action.

- Frame number: Frame number identifies the time when takeaway is first detected. Frame number will be later used to extract real swing in the golf practice video.

- Direction: The direction of takeaway is determine by the direction that the golf club is taken away from a golf ball, whether it is taken away to the left side or the right side of the frame from frontal view. Therefore, if a golfer is right-handed the take away direction is always to the left. On the other hand, if a golfer is left-handed the take away direction is always to the right. To compute the direction, the position of the golf club is considered. If the golf club is placed on the left of the golf ball from frontal view (Right-handed golfer) then the direction of takeaway is always to the left and if the golf club is placed on the right of the golf ball (Left-handed golfer) then the direction of takeaway is always to the right.

- Position of the golf ball: The position of the golf ball when takeaway candidate was detected is kept for validation.

- Position of the golf club head: The position of the golf club head when takeaway candidate was detected is kept for later validation.
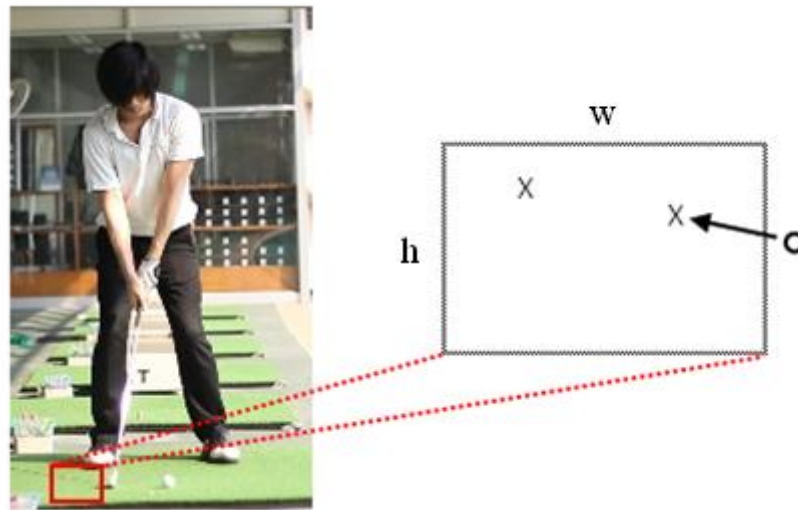
Figure 3.11: Golf club head tracking for takeaway validation process to ensure that takeaway forms a trajectory only in the same direction. The rectangular area is used to search for the position of the golf club head in the next frame.

3.4 Takeaway Validation

The main purpose of the takeaway validation is to verify that takeaway candidates are valid to avoid too much of false detection. To validate a takeaway, the position of the golf ball is tracked to ensure that during takeaway the golf ball is not moving and the movement of the golf club head forms the correct trajectory of takeaway action (see Fig.3.10). To validate detected takeaway, the position of the detected takeaway candidate's golf ball and golf club head are tracked. If a takeaway passes following tracking process then it is considered to be valid takeaway and it will be used for golf practice video summarization.

3.4.1 Golf Club Head Tracking

The golf club head will be tracked for a specific number of frames, which will depend on the FPS of the golf practice video, to ensure that the movement of that golf club head is takeaway action. If golf club head forms the trajectory heading only in the same direction then it is considered to be valid. Thus a takeaway whose trajectory is not heading in the same direction during a specific number of frames will then be rejected.
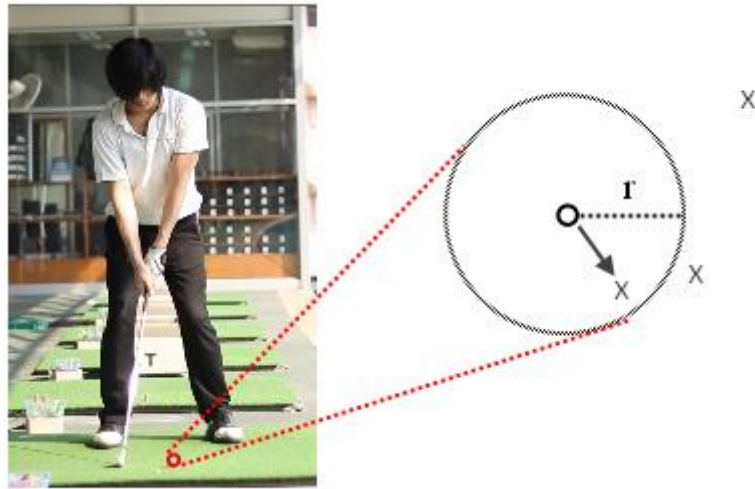
Figure 3.12: Golf ball tracking for takeaway validation and ball hit detection process to keep track of the golf ball. The circular area is used to search for the next position of golf ball in the next frame.

From the experiments, the video frame rate is 24 FPS, the number of frames used to track golf club head is 10.

To keep track of detected takeaway candidate's golf club head to ensure that the trajectory of the golf club head forming the correct trajectory of takeaway action, the rectangular area is used to search for the position of the golf club head in the next frame. If there is a golf club head detected in the next frame within the rectangular area then position of the detected golf club head within the rectangular area in the next frame is selected. In the case that there multiple of detected golf club head in the rectangular area, then the closest of detected golf club head to the current golf club head position is selected (see Fig 3.11). In the experiments, the resolution of the video is 320x240, width and height of tracking rectangular area are 25 to 16 pixels respectively.

3.4.2 Golf Ball Tracking

The golf ball will also be tracked for a specific number of frames because golf club head tracking process alone cannot validate takeaway action. The takeaway is considered to be valid, if the golf ball stays still for a specific number of frames and the

trajectory of the golf club head is only heading to the same direction. On the other hand, takeaway candidates will be rejected if the golf ball moves a lot during takeaway. In the experiments, the number of frames used to track the golf ball is also 10.

To keep track of detected takeaway candidate's golf ball to ensure that the golf ball is not moving and the golf ball is not disappeared during takeaway action, the circular area is used to search for the next position of golf ball in the next frame. If there is a golf ball detected within the circular area in the next frame then the position of the detected golf ball within the circular area in the next frame is selected. In the case that there are multiple golf balls detected within the circular area then the closet golf ball to the current golf ball position is selected (see Fig 3.12). In the experiments, the resolution of the video is 320x240, the radius of tracking circle area is 2 pixels.

3.5 Ball Hit Detection

Ball hit detection is begun after takeaway is validated to ensure that validated takeaway is real swing. The disappearance of golf ball that occur after valid takeaway is considered that the golf ball was hit (see Fig. 3.13). Ball hit detection is used to verify that the valid takeaway produces the real swing, by eliminating false detection when golfer takeaway but interrupted and does not hit the golf ball, because real swing is the swing that hit the golf ball. Every time ball hit action is detected, the following information will be kept for later validation.

- Frame number: Frame number identifies the time when ball hit action is detected. Frame number of ball hit action will be later used with frame number of takeaway action to estimate starting and ending point of real swing to extract real swing in the golf practice video.

(a) Valid takeaway   (b) Ball hit detected

Figure 3.13: Ball hit detection process by detecting disappearance of the golf ball (b), after detected valid takeaway (a).

3.6 Results

The results from the purposed method are frame numbers or indexes of takeaway actions and ball hit actions in golf practice video obtained from takeaway detection and ball hit detection respectively. To extract real swing, frame number of address position and swing finish are needed. However, frame number of address position and swing finish can be estimated from frame number of takeaway and ball hit action respectively. Because takeaway action occurred after address position, frame number of takeaway action can be used to estimate address position, which is the beginning of real swing, by subtracting frame number of takeaway action. And because ball hit action occurred before swing finish, frame number of ball hit action can be used to estimate swing finish, which is the ending of real swing, by adding frame number of ball hit action.

From the experiment frame number of address position can be estimated by subtracting frame number of takeaway action by 12 frames or about 0.5 seconds. And frame number of swing finish can be estimated by adding frame number of ball hit action by 36 frames or about 1.5 seconds.

CHAPTER IV

EXPERIMENT

In this chapter the experiment and experimental result of the purposed method is described. The performance of purposed method is measured by three features which are precision rate, recall rate, and amount of unimportant contents skipped during golf swing analysis process.

The experiment is divided into three stages: The first stage is to record golf practice session with both DSLR and point-and-shoot video cameras in different lighting; morning, afternoon, and evening in golf driving range. The second stage is to process all record video with the purposed method to locate all real swing in the video, and gather results after the completion of processing. Finally, the results from the purposed method are analyzed.

The outline of this chapter is organized as follows. The detail of golf practice video is first described in section 4.1. The process of golf video processing is then described in section 4.2, and finally the experimental results are presented in section 4.3.

4.1 Golf Practice Video Recording

To test the purposed method for precision and recall rate, golf practice session is recorded at golf driving range in face-on camera angle. The recorded camera was mounted on tripod to stabilize video images. The total number of real swing present in the all videos is about 150 real swings. The videos were recorded in different light conditions and backgrounds using different cameras, different focal length and camera setups (see Fig. 4.1). Moreover, in order to test reliability of purposed method, two golfers were recorded during their practice session. To record the video, different cameras are used which are Canon 60D with 17-55 IS lens and Panasonic DMC-FX520.
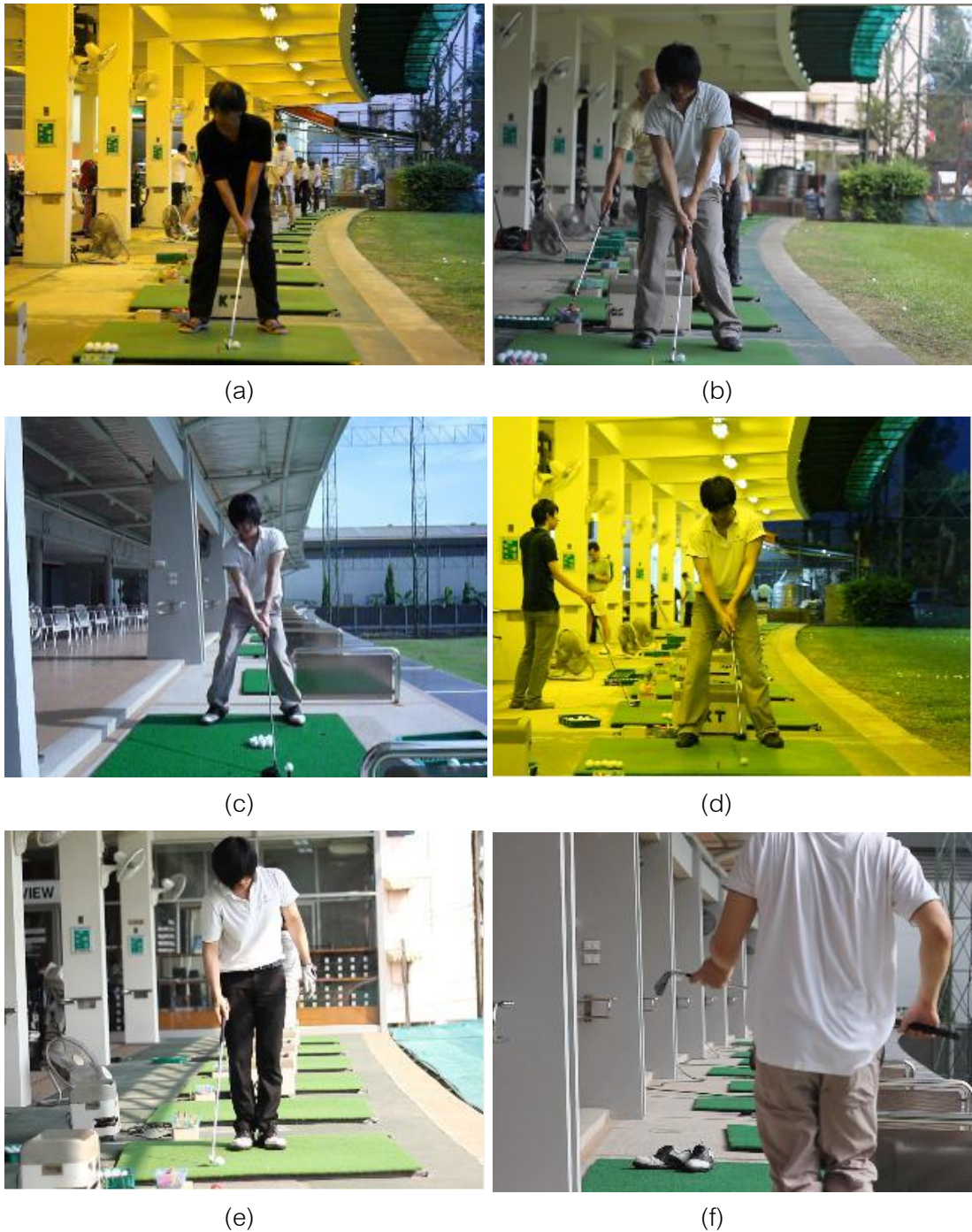
Figure 4.1: Recorded video of golf practice sessions of golfers using different golf clubs (iron, fairway woods, driver) with different camera setups and light conditions (sunlight: morning, afternoon, evening and yellow fluorescent lighting) containing variety of actions and backgrounds.

Table 4.1: The classification of test results.

Detection \ Actual	Takeaway action is actually presented	Takeaway action is not actually presented
Takeaway action is detected	TP	FP
Takeaway action is not detected	FN	TN

4.2 Golf Practice Video Processing

Each record of golf practice video is divided into 1-2 minutes video segments and resized to 320x480 pixels, so that the processing of each video does not take very long time. On the completion of processing of each golf practice video the result of each video is saved and then manually classified into three categories which are TP,FN and FP as illustrated in Table 4.1, where TP is the correct extraction of takeaway, FN is the miss of extraction of takeaway and FP is the false extraction of takeaway. The result of each video is also used to manually measure the amount of unimportant contents skipped by comparing number of frames when viewing only real swings and number of frames when viewing the entire video.

4.3 Experimental Results

Golf practice videos of golfers recorded at golf driving range and golf practice videos downloaded from the Internet in face-on camera angle were used to test the purposed method. The total real swing is 150 real swings. The videos were recorded in different light conditions and backgrounds using different cameras and camera setups. Two cameras are used in the experiments, the first one is DSLR camera and the second one is point and shoot camera. They record different qualities of video. The videos also contain actions that may cause false detection to the system during golf practice such

Table 4.2: The results of experiments.

Recorded Time	TP	FN	FP	Precision	Recall
Morning	53	6	5	91.38%	89.83%
Afternoon	62	1	7	89.86%	98.41%
Evening	23	5	2	92%	82.14%
Total	138	12	14	90.78%	92%

as tapping of a golf ball, practice of golf swing and other golfers practicing in the background. The experiments show that the purposed method can precisely extract the start of real swing in different light conditions, camera setups, and complex background. Table 4.2 shows the result of experiments on golf practice videos in different light conditions; morning, afternoon and evening. It shows that the averages of precision and recall rate are 90.78% and 92% respectively. The equations of recall and precision rate are obtained as follows:

$$Precision = \frac{TP}{TP + FP} \qquad (6)$$

$$Recall = \frac{TP}{TP + FN} \qquad (7)$$

Where TP is the correct extraction of takeaway, FN is the miss of extraction of takeaway, and FP is the false extraction of takeaway. Most of FNs (miss of extraction of takeaway) is occurred when the golf ball or the golf club cannot be detected. Most of the time that the golf ball cannot be located is because the golf ball is blocked by rubber tee or other golf balls and most of the time that the golf club cannot be detected

is because low shutter speed which is caused by the quality of the camera or lack of light. Low shutter speed causes the golf club captured in the video to be blurring. For FPs (false extraction of takeaway), FPs are often caused by false detection of the golf club or the golf ball which forms an action that is similar to takeaway and hit action. From the experiments, the quality of videos obtained from the camera directly affects the recall rate. The recall rate obtained from the videos recorded in the evening is the lowest comparing to videos recorded in the morning or afternoon, because videos recorded in evening is forced to use low shutter speed to gather more light in order to get the correct exposure. The experiments with 85,802 frames of practice videos also show that the purposed method can improve golf analysis process by extracting roughly 14,000 frames of real swing, which results in allowing the golfer to skip over approximately 71,000 frames or 83% of unimportant contents in golf practice videos.

Based on analysis, this purposed method still has some limitations which are mainly caused from golf club detection and golf ball detection. Golf club detection using frame differencing cannot be used to efficiently detect the golf club in golf practice video containing complex backgrounds because the movement in the background may overlap with the motion of the golf club which may cause the detection of the golf club to fail and takeaway action to be undetected. Moreover, golf club detection using Canny edge detection and Hough transform does not always detect the golf club, other objects that look like a straight line may be detected as well which may cause false detection of takeaway action. Golf ball detection using color, size and shape filtering cannot be used to efficiently detect the golf ball when different color of golf ball is used for practicing. It also cannot be used to detect the golf balls when golf balls are placed next to each other. These scenarios will cause detection of golf ball to fail and takeaway action to be undetected.

CHAPTER V

CONCLUSION AND FUTURE WORK

In this thesis, a novel method to extract real swing for video summarization of golf practice video is purposed. This purposed method can help increasing efficiency of swing analysis by allowing golfers to quickly access to each real swing. This purposed method can be applied to implement automatic golf practice video summarization which facilitate golfers by allowing them to quickly locate each real swing in golf practice video which improve the process of golf swing analysis during their practice session. The system of this method consists of golf club head detection, golf ball detection, takeaway detection, takeaway validation and ball hit detection. Each real swing is extracted using takeaway detection and ball hit detection. To detect takeaway and ball hit action, detection of golf club head and golf ball are needed. To detect the golf club head, the motion region is first identified by using frame differencing, morphological operation and bitwise AND operation respectively and then the motion region is identified the golf club shaft can be extracted by using Canny edge detection, Hough transform, length filtering and angle filtering respectively. After the golf club shaft is identified, the golf club head is detected by using frame center as the reference point. To detect the golf ball, color filtering is used to extract white color and then size filtering and shaped filtering is used to detect the golf ball. Takeaway action is identified by close distance between the golf club and the golf ball. Then false detections are removed using takeaway validation process by tracking of golf club and golf ball. Finally real swing is identified by valid takeaway action which followed by golf ball hit. Subsequently frame numbers of takeaway action and golf ball hit action are then used for golf practice video summarization. Although the proposed method has some limitations, the purposed method is still able to efficiently detect real swing in golf practice video. The result of experiments on golf practice video recorded in different lighting environment and different background with different camera setup shows that this method can efficiently extract real swings from golf practice videos. The precision and recall rate of are 90.78% and 92% respectively. Unfortunately, other related methods cannot be

compared with the purposed method because their methods cannot be applied to efficiently extract real swing in golf practice videos.

There are some issues remaining for further improvement of this work in the future. Currently, parameters associated with these techniques such as HSV values, parameters of filters, needed to be set up manually. In future work, these parameters can be improved to by implementing a mechanism to automatically configure all parameters for the system. To improve detection of golf club and golf ball, a more robust technique is needed to be developed which will allow the system to extract real swing in more complex environment. In addition, the concept of practice video summarization of this work can be applied to create summarization technique in practice video of other sport. Moreover, future works in the area of golf swing analysis are very interesting such as automatic golf swing analysis to automatically identify the problem of the golfer's swing.

References

[1]   Andrew Rice. It's All About Impact: The Winners of Over 100 Majors Prove It. Bally
       Down Press, 2009.

[2]   Michael Breed and Greg Midland. The Picture-Perfect Golf Swing: The Complete
       Guide to Golf Swing Video Analysis. Atria Books, 2008.

[3]   Philip Kelly, Aoife Healy, Kieran Moran and Noel E. O'Connor. A Virtual Coaching
       Environment for Improving Golf Swing. In Proceedings of ACM Workshop on
       Surreal Media and Virtual Cloning, 2010, pp. 51-56.

[4]   Ibrahim Karliga and Jenq-Neng Hwang. Extraction and Integration of Human Body
       Parts for 3-D Motion Analysis of Golf Swing from Single-Camera Video Sequences.
       In Proceeding of IEEE International Symposium on Circuits and Systems, 2007,
       pp. 3960-3963.

[5]   Tadej Zupancic and Ales Jaklic. Automatic Golf Ball Trajectory Reconstruction
       and Visualization. In Proceedings of International Conference on Computer
       Vision/Computer Graphics Collaboration Techniques, 2009, pp.150-160.

[6]   Dongwoo Kim, Youngjun Song, Jehoon Lee, Nam Kim,and Jeahyeong Ahn.
       Implementation of an Impact Position Display System of Golf Ball using an Image
       Processing in an Indoor Golf Course. In Proceeding of IEEE International
       Conference on Advanced Communication Technology, 2011, pp. 561-564.

[7]   Cheolkon Jung and Joongkyu Kim. Player Information Extraction for Semantic
       Annotation in Golf Videos. IEEE Transaction on Broadcasting. Vol. 55 , 2009, pp.
       79-83.

[8]   Dan Ring and Anil Kokaram. Online Parsing Of Sports Coaching Video Through
       Intrinsic Motion. In Proceeding of IEEE International Conference on Image
       Processing, 2007, pp. 541-544.

[9]   Jean Serra. Image Analysis and Mathematical Morphology . Academic Press,
       1983

[10] John Canny. A computational approach to edge detection. <u>IEEE Transactions on Pattern Analysis and Machine Intelligence</u>. Vol. 8, 1986, pp. 679–714.

[11] Richard Duda and Perter Hart. Use of the Hough transformation to detect lines and curves in pictures. <u>Communications of the Association for Computing Machinery</u>. Vol. 15 , 1972, pp. 11–15.

[12] Peera Chotimanus, Suphakant Phimoltares and Nagul Cooharojananone. Real swing extraction for video indexing in golf practice video. <u>In Proceeding of IEEE International Conference on Computing, Communications and Applications</u>, 2012, pp. 420 - 425.

[13] Arthur Money and Harry Agius. Video summarisation: A conceptual framework and survey of the state of the art. <u>Journal of Visual Communication and Image Representation</u>. Vol. 19, 2008, pp. 121–143.

[14] Ziyou Xiong, Regunathan Radhakrishnan, Ajay Divakaran, Yong Rui, Thomas Huang. <u>A Unified Framework for Video Summarization, Browsing & Retrieval: with Applications to Consumer and Surveillance Video</u>. Academic Press, 2005.

[15] Cuneyt Taskiran, Zygmunt Pizlo, Arnon Amir, Dulce Ponceleon, Edward Delp. Automated Video Program Summarization Using Speech Transcripts. <u>IEEE Transactions on Multimedia</u>. Vol. 8, 2006, pp. 775-791.

[16] Ziyou Xiong, Regunathan Radhakrishnan, Ajay Divakaran and Thomas Huang. Highlights extraction from sports video based on an audio-visual marker detection framework. <u>In Proceeding of IEEE International Conference on Multimedia and Expo</u>, 2005.

[17] Wei Chen and Yu-Jin Zhang. Tracking Ball and Players with Applications to Highlight Ranking of Broadcasting Table Tennis Video. <u>In Proceeding of Imacs Multiconference on Computational Engineering in Systems Applications</u>, 2006 , pp. 1896-1903.

[18] Kosit Teachabarikiti, Thanarat Chalidabhongse and Arit Thammano. Players Tracking and Ball Detection for an Automatic Tennis Video Annotation. <u>In Proceeding of IEEE International Conference on Control Automation Robotics and Vision</u>, 2010, pp. 2461 – 2494.

[19] Min Xu, Numunu Mudduge, Changsheng Xu, Mohan Kunkunhulli and Qi Tian. Creating Audio Keywords for Event Detection in Soccer Video. In Proceeding of IEEE International Conference Multimedia and Expo, 2003.

[20] Zhuang Li, Prakash Ishwar, and Janusz Konrad. Video condensation by ribbon carving. IEEE Transactions on Image Processing, 2009, pp. 2572–2583.

[21] Nevenka Dimitrova, Hong-Jiang Zhang, Behzad Shahraray, Ibrahim Sezan, Thomas Huang and Avideh Zakhor. Applications of video-content analysis and retrieval. IEEE Multimedia. Vol. 9, 2002, pp. 42-55.

[22] Kenichi Minami, Akihito Akutsu, Hiroshi Hamada, and Yoshinobu Tonomura. Video Handling with Music and Speech Detection. IEEE Multimedia. Vol. 5. 1998, pp. 17-25.

Appendices

APPENDIX A

RESULTS FROM REAL SWING DETECTION

Golf practice videos of golfers recorded at golf driving range and golf practice videos downloaded from the Internet in face-on camera angle were used to test the purposed method. The total real swing is 150 real swings. The videos were recorded in different light conditions and backgrounds using different cameras and camera setups.

Video Number	Number of Real swing	TP	FN	FP	Recorded Time
1	4	4	0	0	Morning
2	6	6	0	0	Morning
3	4	1	3	0	Morning
4	5	5	0	0	Morning
5	3	2	1	0	Morning
6	4	4	0	0	Morning
7	3	3	0	0	Morning
8	3	3	0	1	Morning
9	2	2	0	0	Morning
10	3	3	0	1	Morning
11	2	2	0	0	Morning
12	2	1	1	0	Morning
13	4	3	1	1	Morning
14	4	4	0	2	Morning
15	3	3	0	0	Morning
16	4	4	0	0	Morning
17	3	3	0	2	Morning
18	3	3	0	0	Afternoon
19	4	4	0	2	Afternoon
20	4	4	0	2	Afternoon

Video Number	Number of Real swing	TP	FN	FP	Recorded Time
21	6	6	0	0	Afternoon
22	5	5	0	0	Afternoon
24	6	6	0	1	Afternoon
25	4	4	0	2	Afternoon
26	7	7	0	0	Afternoon
27	8	8	0	0	Afternoon
28	6	6	0	0	Afternoon
29	5	5	0	0	Afternoon
30	5	4	1	0	Afternoon
31	6	5	1	1	Evening
32	5	4	1	0	Evening
33	6	5	1	0	Evening
34	7	6	1	0	Evening
35	4	3	1	1	Evening
Total	150	138	12	14	Total

Biography

Peera Chotimanus was born in Bangkok, Thailand in 1988. He received his Bachelor's Degree in Information Technology from Sirindhorn International Institute of Technology, Thammasat University in 2009. Currently, he is pursuing his master degree in image and video processing at Computer Science and Information Technology Program, Department of Mathematics and Computer Science, Faculty of Science, Chulalongkorn University.