Personalised annotations for augmented video coaching

A thesis submitted to Auckland University of Technology

in partial fulfilment of the requirements for the degree of

Master of Computer and Information Sciences (MCIS)

Feb 2018

By

Kam Yuk (Leona) Chan

School of Engineering, Computer and Mathematical Sciences

Supervisor

Dr Boris Bačić

Abstract

Modern mobile video and camera technologies can provide high frame rates and high-quality videos. Such video technologies can advance augmented coaching and support qualitative movement diagnostics to help sports enthusiasts to improve their golf swing. This thesis presents novel approaches and technology for privacy-preserving personalised annotations for augmented video coaching using golf as a case study.

Preliminary experiments indicated that the commonly-used foreground-background separation algorithms would not perform well in golf-specific contexts. Initial evaluations included benchmarking and combining commonly used surveillance algorithms (using Matlab) that could provide a silhouette of a golfer and a club. Evaluated solutions combined frame difference, erosion/dilatation, blob detection and Gaussian mixture models from the captured video at two diverse-characteristic golf driving ranges. The adaptive multi-layered solution for privacy preservation of golfing activity can provide pseudo-3D binary silhouette transformation of video/image that can be used for augmented coaching and providing anonymised visual annotation feedback while preserving players' privacy. Producing a video or generating a report with annotated angles, golf club head trajectory and other elements of swing performance are important coaching tools to facilitate golf learning from novice to intermediate skill level players. In addition, future work is aimed at further advancements of silhouette-based video streaming solutions and technology transfer to advance the diversity of sports disciplines, and general sports science.

Contents

Preamble

Before enrolling to the video and image processing paper at my university, I had only rudimentary knowledge of this subject. In the past, I though this subject was only useful for editing and composing photos. At the end of my Master-level journey, I had learned to appreciate that image processing techniques can also be applied for extracting various and useful information from media files including both image and video. In my view, this branch of computer science has important role in many application areas including sports coaching, robotics, medicine, surveillance, transportation, satellite imaging and, in the near future, self-driving cars.

When I had canoe training, I was curious to find out how effective would be the placement of the paddle blade into the water and what other elements of performance I could improve that would collectively contribute to making the boat go faster while conserving my energy. Similarly, in the diving competition in the Olympic games, commentators would explain the points awarded that relate to the amount of splash on entry, which is directly linked to the angle of entry in the water and a consequence of the stylistic execution of a diver. In both cases there are stylistic elements of performance that are hard to see and judge. My motivation is to combine my interests to use video and image technology to benefit people who respect the physical movement.

After moving to New Zealand, I had a chance to develop my interest in golf. Prior to commencing this thesis, I haven’t had any formal golf training. The experimental part of golf data collection and analysis provided me with my first-time golf experience. I believe that the literature review, collected video analysis and discussions with my supervisor, had a positive influence on my golf learning. In support of the use of augmented coaching systems and technology, I was pleased to see from collected data that both my golf swing and posture had improved during the three practise/data collection sessions. I hope that the developed artefacts for privacy-preservation-driven augmented coaching systems and technology will help other golfers to improve their game and awareness regarding the elements of swing performance.

List of tables

List of figures

X

Attestation of Authorship

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person (except where explicitly defined in the acknowledgements), nor material which to a substantial extent has been submitted for the award of any other degree or diploma of a university or other institution of higher learning.

Signature:

Date: 01 Mar 2018

Acknowledgments

First of all, I would like to thank God for sustaining me in the best and worst times of my life.

I want to express my deepest gratitude to my supervisor Dr Boris Bačić for the patient guidance, encouragement, advice and support in my study. I was very lucky to have him care so much about my work, and respond to questions. I also wish to thank Qi Meng (Matthew), Assoc. Prof. Wei Qi Yan, Dr Kirsten Spencer, Terry Brydon and Bumjun Kim for their help and support.

I must thank my husband Gavin, for his continued support and my two sons, Winton and Ethan for their daily co-operation. They allowed me to have more time to concentrate on my study.

I wish to thank Auckland University of Technology (AUT) for the paid conference registration fee for the SPRINZ conference 2017 and providing the help of use of the library and golf driving range. In addition, I also wish thank Takapuna golf course for allowing customers to record their own golf swing videos.

Ethics disclaimer

All recorded motion data and reported author's self-activities are subject to an exception from the AUT ethics committee approval ("Guidelines and Procedures", http://www.aut.ac.nz/researchethics/guidelines-and-procedures, retrieved 16 Feb. 2018), under Section 6 ("6.4. Research and teaching in which a single investigator is the subject of his/her own research and where no physically or psychologically hazardous procedure is involved.", http://www.aut.ac.nz/researchethics/guidelines-and-procedures/exceptions-to-activities-requiring-autec-approval-6 , retrieved 16 Feb. 2018).

List of terms

The list of terms provides terminology used in this study for multidisciplinary audiences and to prevent possible disambiguation.

Term	Description
Address	Golfer's stance in preparing for the swing
Annotation	An abstract way of augmenting explanation by added comments to a text or abstract drawing over the picture or video.
Background	The part of a picture that which is behind the main object and appears furthest from the viewer.
Coach	An instructor, sport expert or trainer in sports
Digitalisation	Convert (pictures or sound) into a digital form that can be processed by a computer.
Dilations	The action or condition of becoming or being made wider, larger, or more open. In image processing, dilation function expands the selected objects boundaries (Gonzalez & Woods, 2008, p. 523).
Edge	The outside limit of an object, area, or surface.
Erosions	In image processing, erosion function shrinks the selected objects boundaries (Gonzalez & Woods, 2008).
Follow-through	Continuation of the movement of a club after striking a ball
Foreground	The main object a picture that is nearest to the observer.
Impact	The moment when the golf club hits on the golf ball
Silhouette	A two-colour representation of a two- or three-dimensional object.
Sweet spot	A subarea of the clubface intended to provide optimal feel and energy transfer to the ball at impact. Sweet spot is specific to the club model and may be provided by the club manufacturer.
Takeaway	Initial part of the backswing
Trajectory	The path followed by a projectile flying or a moving object (e.g. golf ball and golf club).

Chapter 1 **Introduction**

Figure 1.1: Naturally occurring silhouette phenomenon due to lightning conditions, in which human vision and brain have can separate the object from its background (Photo taken by the author on 15 Apr 2016 at Cathedral Cove, Hahei Beach, New Zealand)

Images and videos are used for presenting our memories and precious life moments. In cinematography, people often find various information from the videos and images that are pertinent to the storyline. For instance, in Figure 1.1, people can understand the foreground object is a human jumping in a cave, while the background consists of waves lapping at the beach with plenty of other people enjoying the sun and azure skies with scattered clouds.

Modern sports gadgets, exergames, wearable devices and various sensors integrated in sports equipment are collecting various motion data including video. In general, such *augmented coaching systems and technology* (ACST) can record human activity and provide various numerical analysis associated with sports-specific activity (Bačić, 2016b; Lightman, 2016). However, teaching how to improve sports-specific technique (such as golf swing) is still left to a coach. One of the popular sports that requires complex movement is golf. For example, one of the early mobile apps relying on machine learning and computer vision is the SwingProfile mobile app (Z. Chan) that was introduced at the 2012 PGA merchandise show. This mobile application can automate video capture and swing detection, and provide rudimentary video analysis of the observed golf swing. However, the growing number of mobile applications (Ashford, 2012) and other forms of video-based coaching that facilitate video capturing, storing and sharing captured video information on-line also presents challenges such as privacy concerns and the need for data protection.

For end users who wish to record video on local devices using augmented coaching video analysis there are also available free or open source software alternatives including VirtualDub (*VirtualDub*), LongoMatch (*LongoMatch*), Kinovea (*Kinovea*) and VLC (*VideoLAN*) media player (Bačić & Hume, 2012; Bačić, Meng, & Chan, 2017).

The purpose of this research is to develop new algorithms and approaches to support biomechanics, and promote open source software initiatives and general video-based coaching with additional concerns regarding privacy preservation in video and image analysis.

1.1 Research questions

Given the increasing availability of video technology, there is a potential to advance sports science, coaching technology and practice by using a computing discipline of video and image processing. While intending to preserve privacy information in the video, there is also the requirement to annotate critical features that guide the assessment of static and dynamic movement patterns for the golf swing that could apply, to some degree, to other sports. The intended outcomes are also intended to foster advancements in the next generation of augmented coaching systems, exergames and rehabilitation devices.

Given the increase of privacy-invasive technologies and the availability and diversity of video-based technologies used for sports coaching and rehabilitation[1], the aim of this thesis is to advance augmented video coaching systems and technology while preserving the end-users' privacy.

Using a case study on golf,

(1) Can develop algorithms and other artefacts to preserve privacy from video while providing the evidence of performance and safety elements of the golf swing? In particular:

  a. Can silhouette filtering be suitable for privacy preservation?
  b. To what degree it would be possible to analyse and annotate observed diagnostic elements of golf swing performance, while at the same time suppressing personal information to preserve privacy?
  c. What performance elements of a golf swing can a coach qualitatively or quantitatively assess from the transformed video with suppressed personal information?
  d. Are there any elements of potential coaching bias using the intended solution?

(2) What computing techniques would be suitable to develop a solution that would be robust to changes in brightness, colour variety, contrast, shadows and static or dynamic elements of foreground and background regarding the captured videos?

(3) Can the developed solutions be applied to annotate hard-to-quantify common-sense elements in augmenting video coaching scenarios for a chosen sports activity?

(4) What are practical implications for amateur golf coaching?

(5) What are the benefits and limitations of monocular camera view found in commonly used mobile and pocket-sized camera technologies for augmented golf coaching?

[1] The last phase of rehabilitation typically involves return to sport or to normal daily activities and in this thesis is also seen as a coaching process.

To facilitate the coaching process of sports-specific techniques, the elements of augmented coaching include video capture, video replay and overlay annotation functionality of movements of interest.

1.2 Scope of the thesis

The purpose of this multi-disciplinary research is to combine sports science and domain expertise with video and image analysis. The objective is to develop new algorithms and approaches to enable the usage of a variety of video sources. Privacy preservation of the image is the object with the annotated critical features that guide the assessment of static and dynamic movement patterns found in a golf swing. Aligned with the design science approach, the chosen method in this study includes incremental prototyping and produced artefacts evaluated on a variety of captured golf videos recorded in common resolutions and frame rates for mobile technology.

1.3 Thesis structure

This research in this thesis is organised as follows:

Chapter one introduces an overview of the thesis contributions in the discipline of computer science and sports science. This chapter also provides the motivation, background and initial research questions that guided how ideas were logically developed, implemented, evaluated, improved and presented.

Chapter two presents a multidisciplinary literature review combining sports science, augmenting golf coaching tools with the computer science disciplines of computer vision and video and image processing. The chapter provides a summary and synthesis of a literature review that has influenced the directions of the experimental development for advancements of augmented video coaching systems and privacy preservation.

Chapter three provides the methodology and data collection, and illustrates the observations and interpretations of the experimental design, intermediate evidence and recommendations.

This chapter also compares the performance of computer vision methods such as foreground-background separation and morphological operations.

Chapter four describes the developed solution for the privacy preservation of golf videos along with annotation methods to support an augmented video golf coaching system. This chapter also presents the design and prototype of the developed system and critically reflects on the result and strengths of the proposed solution in coaching practice.

The last chapter provides conclusions, contextualises the achievements and limitations, addresses the research questions, and summarises the contribution and experimental approaches based on the presented evidence and novel approaches from the previous chapters. This chapter concludes with future work and the opportunities for multidisciplinary advancements based on this study.

1.4 Thesis contributions

This thesis makes multiple contributions to the growing area of computer vision systems and sports science technology, especially on privacy preservation augmented video coaching systems, ubiquitous computing and eSystems design. The developed framework, algorithms, software tools and techniques for augmented video replay-based coaching:

(1) provide silhouetted-based image or video that can show the modified foreground object of interest, preserving golf player’s privacy while removing the unwanted information from the background;

(2) provide digitised movement evidence that could help to facilitate communicating between coach and player or aid in self-coaching;

(3) provide the annotations and user interactive function help to calculate and analyse the angles of interest, such as parts of the body during the stance and throughout the swing motion;

(4) annotate the estimated swing plane;

(5) recognise and include in the foreground, information of the golf club in order to produce automated tracking of the golf club head trajectory and present the throughput of the selected golf swing motion;

(6) provide a visual coaching analysis report with coaches' comment and instruction.

Authors' publication related to this research thesis

Bacic, B., Meng, Q., & **Chan, K. Y.** (2017, 14-16 June 2017). Privacy preservation for eSports: A case study towards augmented video golf coaching system Symposium conducted at the meeting of the 2017 10th International Conference on Developments in eSystems Engineering (DeSE) doi:10.1109/DeSE.2017.34.

> Conference proceeding publication in the combining the fields of eSystems, ubiquitous computing, and video and image processing. As a co-author, I declare that I reported the parts related to experimental evidence of foreground-background separation algorithms for silhouette video filtering with relevant literature review in surveillance and related contexts. I also contributed to the motivation and introduction sections of this research paper.

Chan, K.Y & Bačić, B (2017). Producing silhouette-based augmented feedback: What elements of performance can a golf coach analyse? SPRINZ strength and conditioning conference – AUT SPRINZ, Auckland, New Zealand.

> Peer-reviewed conference publication in combining the fields of sports science, ubiquitous computing, and video and image processing. As a co-author, I declare that I reported and presented to the sport audience the parts related to the finding, experimental evidence of foreground-background separation algorithms for silhouette video filtering and result related contexts.

Chan, K.Y (2017). Personalised annotations for augmented video coaching, Centre for Robotics and Vision (CeRV), AUT, Auckland, New Zealand.

> I declare that I presented the key finding and main part of the developed solution in the centre for robotics and vision research seminar. https://cerv.aut.ac.nz/2017-seminars/

Chapter 2 ***Literature review***

Figure 2.1: Combined multiple exposures for stroboscopic action effect conveying the golf swing motion. Modified from SwingMoment (2016).

Computer vision and image discipline is used to provide evidence of captured movement, eliminate redundant information while preserving privacy and facilitate coach-to-player communication. This chapter covers multidisciplinary literature supporting and complementing the main work in this thesis. There are three main parts: in the first place is the review of related work around sports science, also specifying on golf. The second part is coaching knowledge and coaching systems. Finally, is the various methodology of video and image processing including the existing state of the art in foreground-background separation for the golf swing.

The purpose of this multidisciplinary literature review is to research sport science and golf with computer science, providing state-of-the-art computer vision and image analysis in the context of technology-mediated augmented coaching. For example, Figure 2.1 shows an image processing technique with a stroboscopic effect depicting the trace of the golf swing by combining multiple images on supressed background information. Such an artefact provides movement information indicating some of the elements of golf swing technique within a specific time interval while suppressing background information.

2.1 Multidisciplinary nature of human movement in sports

In general, human movement is covered by a collection of scientific disciplines with focus on sports, fitness and wellbeing activities (Stuart, 2017). The daily activity expressions such as ‘bending knees’ and ‘flexing trunk’ are well-understood, but they also might be considered ambiguous in communication between biomechanists, athletes, coaches and other sports enthusiasts. For biomechanists, and other sports scientists precise and non-ambiguous terminology is important, however such communication may also make it difficult for an athlete to improve. It is also common for a coach to interpret scientific analysis of observed human movement and convert it into an actionable intervention (Bartlett, 2007, pp. 38-39). From the point of view of sports biomechanics, when two body joints are fully extended, the angle is 180°. In the contrast, the clinical analysis may also report the fully extended joint angle as 0° (Bartlett, 2007, p. 266; Winter, 2009, p. 77).

2.1.1 Three- and two-dimensional representation of human movement

Human body movements are in three dimensions, and the joints focus on three axes of rotation that can be observed from three different views or perspectives: sagittal, frontal and horizontal planes (Bartlett, 2007). These three planes are commonly referred to when describing the motion in three dimensions (Table 1) and are important for motion data acquisition protocol design, analysis and augmented coaching feedback.

Table 1: Human movement plane definition and description (Bartlett, 2007, pp. 3-6, 40, 107, 130, 141)

Plane	Main movement	Movement definition	Description the sports and exercise movements
Frontal	Abduction	A limb away from the axis of the body	Sideways movement of the trunk
	Adduction	A limb toward the centre line of the body	
Sagittal	Flexion	The action of bending	Two-dimensional movement, analysis: Walking, running, long jumping
	Extension	Increasing the inner angle of the joint	
Horizontal	Medial rotation	Turning toward the midline of the body	Rotation of head, neck and trunk
	Lateral rotation	Turning away from the midline of the body	

2.1.2 Golf background

It is widely acknowledged that golf is one of the attractive and popular sports to a wide range of ages. In 2016 Americans had 24 million golfers (Heitner, 2016). The increasing number of the youth players can benefit from the *Professional Golf Tours* (PGA Tour) and *Ladies Professional Golf Association* (LPGA) growth for the future (Heitner, 2016).

Golf is a closed motor skill, which can be acquired and improved by repetitive practice and coaching feedback (Courtney & Thomas, 2004, pp. 147, 156; Dayan & Cohen, 2011; Jensen; Kooyman, James, & Rowlands, 2013; *Practice to improve your golf motor skills*). With repetitive practice motor skill may be acquired, and can also be retained over a long period (Dayan & Cohen, 2011).

While in general golf is considered of Scottish origin, first recorded game of golf was on 1297 in the Netherlands where people had similar sports activity (*History of golf*, 2011). Over the course of the last nine decades, golf has undergone great transformations in terms of definitions, rules, equipment and technique advancements. The classic swing includes recommendations regarding posture, rhythmic and smooth movement and finishing the swing with relaxed pelvis and thorax. On the other hand, the golf swings, aiming to deliver power at impact is also knowing to produce stress to the muscles, tendons, ligaments, joints and spine. Unfortunately, many golfers (55% of professionals and 35% of amateurs) will experience some form of back injuries as reported by Cole and Grimshaw (2016).

2.1.3 Different definitions of golf movements

Over time, there has been a different phasing analysis of the golf swing. In the past, there were three phases: preparation, execution and follow-through (Maddalozzo, 1987). More recently, golf movement is divided into four phases: address, backswing, downswing and follow-through (Cole & Grimshaw, 2016). Table 2 shows comparisons of the two different approaches in phasing analysis in golf coaching that have evolved in literature over time.

Table 2: Comparisons of two different approaches in golf swing phasing analysis (Bartlett, 2007, pp. 3-6, 40, 107, 130, 141; Cole & Grimshaw, 2016; Maddalozzo, 1987)

Phase defined by Maddalozzo (Maddalozzo, 1987)	**Component**	**Motion**	**Phase defined by Cole** (Cole & Grimshaw, 2016)	**Requirement of augmented coaching tools**
Preparation	Grip, Stand, knee, hip	Static, small movement	Address	Detect the smearing artefacts. Detect, compare and analyse the angle of body parts.
Execution	Backswing	Greatest acceleration at the top of the swing. Fast-movement golf club. Pelvic and thorax rotate Small motion in lower part of the body of the golfer.	Backswing	High frame rate camera to catch high-speed movement of golf club. Detect, compare and analyse the angle of body parts.
	Downswing, impact of the golf club and ball	Highest speed of the golf ball	Downswing	Trace the golf club head and golf ball
Recovery	Follow-through and finishing the swing	Deceleration of movement	Follow-through	Detect, compare and analyse the angle of body parts.

In spite of the different phasing analysis of the swing, the important performance elements in golf coaching include grip, posture, stance, ball position, backswing, downswing and follow-through.

Given the qualitative and personalised nature of golf coaching and in agreement with golf literature (Table 3) a coach may include in his/her feedback various elements of swing performance and their recommended values.

Table 3: Modern swing body position (Cole & Grimshaw, 2016; Hume, Keogh, & Reid, 2005)

Phase	**Body part**	**Position**
Address	Feet	About shoulder width
	Knees	Bent 20–25°
	Trunk	Flexed about 45°
	Body	Lean forward to their lead foot 20–30°
Backswing	Pelvic	40–50° rotation
Downswing	Club and lead arm	90°
Follow-Through	Pelvic	Turned through about 90° to be 40°–45°
	Upper thorax	Turned through around 105° to be 20°–25°
After impact	Thorax	Turns 120° or above
	Upper body part	Over 120° rotation

2.1.4 Elements of swing performance and safety consideration

Golf swing kinematics includes terms such as flexion, lateral bending and rotation. Understanding the swing mechanisms can improve elements of the performance (such as: consistency, accuracy, distance and ball trajectory) and prevent injuries. Many golf swing errors can be compounded and initially caused by the setup and backswing. Furthermore, the accuracy of a golf swing is also determined by the swing speed and path (Hume et al., 2005; Knudson, 2013, p. 84).

Injuries

Golf is a sport popular with seniors. Analysed information from North America and France shows they are 25% of the total golfer population (Cann, Vandervoort, & Lindsay, 2005; Rouillon & Rouch, 2015). From the research, back and wrist are the most injured body parts. Proper posture to prepare the golf swing is recommended to reduce injury (Cann et al., 2005). A lot of researchers reported that poor golf swing can cause injuries (Cole & Grimshaw, 2016; *Common golf injuries and how to avoid them*, 2017). Generated speed and force are considered as elements of swing performance aiming to maximise force at impact, however, they also produce stress to golfers' joints, in particular when ball impact is outside of the clubface *sweet spot*. A smooth golf swing is important for injury prevention and transfer of power from your body into impact (Choi, Joo, Oh, & Mun, 2014). For veteran-level golfers (Table 4), common injuries specific to body parts were reported by Cann et al. (2005).

Table 4: Most common injuries of elder people during golf swing (Cann et al., 2005)

Phase	Most commonly injured body parts	Occurrence
Takeaway	Back and wrist	<25%
Impact	Back, wrist, elbow and hand	Not known
Follow-through	Back, shoulder, ribs, knees and wrists	25%

Golf swing plane

The circular moment of the golf swing influences the ball trajectory. Nonetheless, the outcome of swing impact is greatly influenced by the swing plane (Bačić, 2016b). The effective swing plane is coordinating the movement of wrist, shoulders, hips and spine. Most golfers know those characteristics at the address position. Nevertheless, one of these characteristics changes at the impact position, which can change the ball trajectory directly (Bačić, 2016b; Stickney II, 2016).

2.2 Coaching practical and augmented coach technology

Q*ualitative movement diagnosis* (QMD) help to improve human movement as prove by many kinesiology professionals. The four circular modelling tasks of QMD are preparation, observation, evaluation and diagnosis, and intervention (Knudson, 2013, p. 218). Sports coaches contend that experience exerts a great influence on knowledge. They can help players to discover their potential and aims to achieve these goals by certain actions. Sports coaches can identify the player's target by defining and running appropriate personalised training programmes. In addition, sports coaches are the guidance; they analyse player performance and provide feedback as well as instruction on the critical elements. These professional messages help and motivate players or athletes enhance their skill (AGCAS; Bentley, 2015).

The coaching role should include (Johnson, n.d.):

- Providing direction and ensuring their learner stays focused and understands priorities
- Improving learner performance by assessing their current capability and identifying the need
- Providing feedback to the learner to prevent repeating in a wrong way
- Creating opportunities, and hence motivating learner performance

2.2.1 Existing sports coaching systems

Besides recognised golf augmented coaching equipment, there is a large amount of product on the market which can analyse and feedback on golf training. The attractiveness of the invention is the benefit to sports training, and precisely to trajectory detection and the feedback system on golf. Based on the capture and analysis of the initial trajectory and immediate reporting, it can positively improve the golfer’s performance in addition to providing entertainment. The portable analysis system can fit any area including a golf course or driving range. It can be combined with other golf equipment, Meanwhile, the system should not disrupt golf activities.

In one example, the system includes cameras; the videos are uploaded via a network to a storage or database. It can facilitate measuring numerous variables. In another example, the system can predict trajectories of shots from the captured part of the trajectory. The initial portion of a trajectory is captured at difference distances by different devices. As a result, the trajectory, ground track and location can be predicted by the system. In addition, the device can recognise the club which a player used. For example, a unique identifying sticker or attachment on the club can be identified by the vision system. The device has the ability to measure environmental conditions including wind, temperature, humidity and air density (Marty & Edwards, 2013).

It is acknowledged that there is much augmented coaching equipment in the golf market including golf balls, tees, clubs, golf bags, smartwatches and smartphones, which are embedded with different types of sensors to help golfers enhance their golf technique by, for example, analysing a golf swing (David & Bondu, 1975; Kawasaki, 1996; Rocha, 2006; Ruiter).

Currently technology with different sensors used for different sports, including punch rate, punch type and posture of the boxing and martial arts player. Speed and hitting zone of a baseball and softball can be tracked by embedded sensors in the handle of the bat. Nevertheless, not many commercial sports sensors can export the processed data for further data analysis. An existing sports coaching application – SmartSwing can export to PDF reports, data normalisation process and numeric format spreadsheet file processes by Matlab ver. 7.1 (*SmartSwing*).

2.2.2 Video-based augmented coaching systems

Currently, it is widely asserted that people could enjoy a video analysis as a part of augmented coaching. Many books provide high-quality images and videos for sports coaching as an example of movement diagnosis for guided tutorials (Knudson, 2013, p. 164; Tinning, 2009, pp. 41, 61 & 67). Early video pedagogies were developed over 30 years ago, like aerobics instruction by a Jane Fonda videotape. They benefit people learning aerobics or yoga at home without formal sessions (Tinning, 2009, p. 41). Furthermore, there are later version videos on the internet and exergames (Best, 2013; Tinning, 2009).

Two examples, *Professional Golfers Association* (PGA) and ERTHEO are international sports organisations. PGA coaches have used video for over 11 years in golf sessions to communicate and follow-up with players (Hughes, 2010). As well as ERTHEO, they have video analysis for training (*Training course for coaches - Barcelona*).

Currently, many sports coaching products on the market, including Dartfish; MAR; Quintic; SiliconCoach; Simi (Bartlett, 2007; *Dartfish*; *MAR*; *Quintic*; *SiliconCoach*; *Simi*) aim to provide vision analysed feedback to the athlete. One commercial sports coaching system Hawk-Eye, use high-quality cameras and complex vision processing technology to track balls and player position with high accuracy. Hawk-Eye combines an intelligent computer system with high-speed biomechanical analysis software. The system provides online and offline video to analyse the angle and performance of the player (*Hawk-eye*). However, that certain specialist commercial software is high cost and time consuming (Bartlett, 2007). Moreover, it is a documented fact that some open source software (OSS) including VirtualDub (*VirtualDub*), LongoMatch (*LongoMatch*), Kinovea (*Kinovea*) and VLC (*VideoLAN*) media player experiences exert an extremely large influence on augmented coaching. Evidently, some functions of VirtualDub are helpful for coaches, athletes and scientists (Bačić & Hume, 2012).

The qualitative analysis should include immediate feedback. It is most popular in using video to provide qualitative analysis. Many studies reported that the video replay should provide useful information for the coach, which can help the learner to improve performance. The video facilitates capturing a human movement hard to see with the naked eye, which detail cannot be seen in the live analysis (Knudson & Morrison, 2002).

One of the important roles of the video augmented coach is self-modelling. People can compare themselves to the other performers. Nevertheless, there is a psychological impact of negative comments when watching the video with other performers (Knudson, 2013, p. 169). It is one of the goals of this study’s solution that by using the silhouette-based human images to protect player privacy while analysing video, the critical features still can be presented.

Compare with the current commercial augmented coaching technologies, this video based augmented coaching study can benefit from shared resources and more economically by using a smartphone or digital video camera. Therefore, the user can avoid the high cost specialist commercial hardware and software (Bačić & Hume, 2012). Traditional coaching commonly communicates by verbally, such as ‘bend knees’ and ‘raise arms’, but they are ambiguous in communicating with sports practitioners. The visually evident is better explained scientifically. The visually evident also can benefit from comparison with a similar player or the same player previously recorded. For instance, comparing the angle of bending knees between previous and current swing records and visual evidence of the golf ball trajectory as well (Bačić, 2016b).

2.2.3 Benefits of high-frame-rate video analysis

Video can help to analysis sports movement patterns to observe and compare more easily. High frame rate and quality digital video cameras are necessary equipment for coaching (Bartlett, 2007).

Lower speed cameras are not enough to capture the full golf swing. Some scientific research reported the acceleration of the golf club to be more than 160km/h in 0.25 seconds (Cole & Grimshaw, 2016; Kim, Millett, Warner, & Jobe, 2004; Thériault & Lachance, 1998). Consequently, high-speed cameras are required, for instance when capturing 'impact' in accurate timing. Additionally, quality of the video pictures is another issue, so it is important to maximise the image quality (Knudson, 2013, p. 164).

Video feedback is effective for advanced beginners and intermediates and feedback focusing on the parts and practice immediately is good for beginners' motor skills (Knudson, 2013; Kooyman et al., 2013).

2.3 Computer vision and video and image processing

In the real world, the human brain perceives with apparent ease the diversity of colours, textures and shapes in the 3D structure around us. Computer vision is how computers recognise, interpret, and understand from this from 3D substance to 2D image. When computer imitates the human brain with retinal information to understand that unconstrained 3D information, its processes include segmentation, recognition, reconstruction and reorganization (Malik et al., 2016). Image processing overlaps with the computer vision in a lot of areas (Figure 2.2), where input and output are both images.

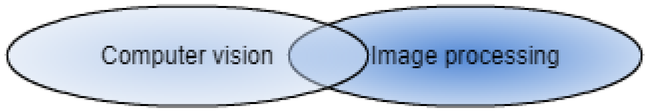


Figure 2.2: Computer vision and image processing area overlap

Human motion analysis by computer vision has increased with research. A large number of individuals would like to understand human motion from the video image (Singla, 2014). A study concluded that sports object distance is typically 5 to 20 metres and moving (Gong & Xiang, 2011, p. 133). To demonstrate, many visual surveillance applications are used to detect and classify human events by lower cost video camera (Paul, Haque, & Chakraborty, 2013). Those systems detect the moving object from the video to understand the event and explain the object's behaviour. Moreover, human detection and tracking are big tasks in most computer vision systems (Mishra & Saroha, 2016) because to detect and track human motion from a video image, the foreground-background separation is a prerequisite. A variety of foreground-background separation methods are discussed in the next sessions.

2.4 Foreground-background separation

Background is a continuing static visible object that has low variance distributions. Conversely, when a moving object comes into the background, a new distribution will be created on the existing distributions' variance. In addition, the moving object's variance should be greater than a background pixel (Stauffer & Grimson, 1999).

Foreground-background separation is a crucial pre-processing step in most vision-based computer systems (Archetti, Manfredotti, Messina, & Sorrenti, 2006). Distinguishing moving objects from the video is the essence of video and image processing. For instance, a surveillance camera counts the number of passengers entering or leaving a lift. The application needs to extract one individual alone. In general, the extracting dynamic foreground from the static background, it is required a static camera. Basically, foreground-background separation is extracting the moving object by the difference between the reference frame and a current frame that is a foreground image (Paul et al., 2013). Some literatures reports that with a dynamic background is difficult to detect foreground, especially in busy environments. With light change, shadows, running water and waving tree, it is difficult to determine whether pixel is a property of the foreground or the background

(Culibrk, Marques, Socek, Kalva, & Furht, 2007; KaewTraKulPong & Bowden, 2002). A golfer on the golf driving range will be considered a foreground object, but when the object is stationary or has only a small amount of movement, for example when preparing a swing, the object will become a background image. A simple approach to subtracting foreground is determining the different pixels between the current frame, *Ci*, the background image *Bi*, and a user-defined threshold, T. A pixel is represented by (x, y). Therefore, foreground is denoted as:

$$Ci(x, y) - Bi(x, y) > T \qquad (2.1)$$

Nonetheless, it is not as simple as formula (2.1). Much literature suggests a lot of methods, but those methods are not perfect for all situations (Archetti et al., 2006). For instance, Haar cascade based detection is object detection, which can distinguish human and non-human objects from a divided foreground. However, this can probably only work well on to the frontal face and upper body. If the human is not in the frontal pose, the system may fail to detect (H. P. Jain, Subramanian, Das, & Mittal, 2011). For foreground-background separation to allocate each pixel into the foreground or background, it is important, it is not simply categorised by moving or static. For instance, the foreground object (person) which temporary stops should be still detected. Therefore, many common methods detect the flowing water, moving clouds and waving leaves as foreground objects (Kryjak, Komorkiewicz, & Gorgon, 2014). As literature has demonstrated, foreground-background separation from a dynamic background is the challenge of computer vision (D Stalin Alex, 2014; Kryjak, 2014; Shaikh, Saeed, & Chaki, 2014). To satisfy these challenges, the next section discusses some algorithms for foreground-background separation.

2.4.1 Frame difference

Using comparison of each frame to the background frame is a most common foreground-background separation approach (Ananya, Anjan Kumar, & Kandarpa Kumar, 2015; Szeliski, 2011; Vedaldi & Fulkerson, 2010). The frame difference method is based on time intervals, and compared to the threshold (D Stalin Alex, 2014; Singla, 2014).

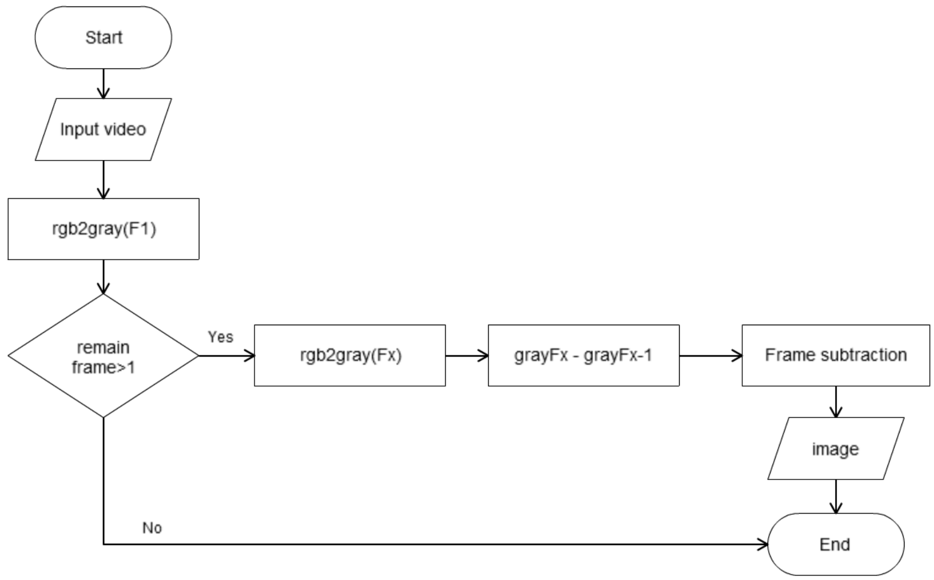


Figure 2.3: Flowchart of frame difference method, modified from (D Stalin Alex, 2014)

The frame difference method is simple but very sensitive. The motion object can be extracted the variation between the current image and the background image (Ananya et al., 2015) as per concept of Figure 2.3. The frame difference is well performed in the static background situation. Alternatively, it is evident that several critical situations including a bad quality image source, lighting change, small movements and shadow need to be handled. On the contrary, the lower computational effort is a benefit of the frame difference (Archetti et al., 2006)

2.4.2 Gaussian mixture models

Gaussian mixture models (GMM) is one of the common unsupervised classification techniques. It requires experimental parameters tuning (Archetti et al., 2006; S.-S. Huang, Fu, & Hsiao, 2007; Wang, Bunyak, Seetharaman, & Palaniappan, 2014). Object tracking failure is mainly caused by four cases: faster object moving, lower input frame rate, detection takes a very long time and the tracking region is not enough.

The background subtraction in some typical methods especially in busy environments has big problems. It is unable to differentiate between moving shadows and moving objects. A background adaptation model and shadow detection technology are needed to provide a faster and more accurate system (Stauffer & Grimson, 1999). One of the solutions is the adapted GMM which check every new pixel with existing model components to run in real-time (A. Jain, Reddy, & Dubey, 2014). If their pixels stay longer and are more static, they are defined as the background. Furthermore, GMM also tracks the lighting condition in a static scene. The static object will degrade foreground image separation (Bouwmans, El Baf, & Vachon, 2008). It is recommended to determine the Gaussian parameter estimates by a large amount of examination.

While the moving shadows cannot be identified by a tracker, GMM uses colour spaces to distinguish chromatic and brightness components. Background pixels are compared with the thresholds, shadows are the differences of chromatic and brightness components (KaewTraKulPong & Bowden, 2002).

GMM can deal with lighting changes; the mixture of adaptive Gaussians would be modelled. Because of updating the parameters over time, the non-background pixels are tracked by the multiple hypothesis trackers. GMM estimates the background by ordering the value of ω/σ, which represents the distribution and variance. After sorting, presumably, the background distributions are on top. Hence, the foreground pixels can be subtracted (Stauffer & Grimson, 1999).

2.4.3 Histogram of oriented gradients (HOG)

The *Histogram of Oriented Gradients* (HOG) is the classification technique of computer vision used for object detection. The basic idea is to compute x- and y-direction cell-wise gradient changes to produce a gradient vector. To some degree it is possible to consider that gradient vectors are inspired by human vision stimuli emphasising brightness/intensity changes close to boundaries of different intensity regions in observed image. Transformed HOG features from sub-image blocks are robust to changes in scale or rotation. The gradient estimator used in HOG transformation is also robust to the image noise (Gonzalez & Woods, 2008, pp. 40-42). The intensity-invariant and improved detection rate are important properties of a gradient-based step-edge detection. For instance, the 64 x 128 pixels window is divided into 105 blocks. Each block has 4 cells, 9-bin histogram per cell, therefore 3780 components in each concatenated histogram (Rezaei & Klette, 2017). Typically, HOG performs well in human detection tasks and as such it could also be a suitable for golfer detection tasks. Being recognized H-HOG and C-HOG (rectangular and circular) are implemented. R-HOG blocks are computed in the regular grid at a single scale and position of the detection window. It is experienced that 2x2 and 3x3 cell blocks and 6-8 pixel are the best performance for human detection. Indeed, it is observable that performed lower if the block is too big or too small. The best performance of C-HOG is small descriptors with very few redial bins. For good performance, centre and surround radial bins, and four angular bins are the requirement. The central bin, radius of 4 pixels is the best (Dalal & Triggs, 2005b).

2.4.4 Stick-figure models

Stick-figure animations are a simplified description of human movement, which are quick and easy to provide visualisation for (Bartlett, 2007; H. P. Jain et al., 2011; Mabiala et al., 2016). Stick-figure models are formulated by some rigid parts with connectivity between them by flexible joints and recognised as the whole figure. Stick-figure is one of the common models for describing human motion (Bačić, 2016a; Meeds, Ross, Zemel, & Roweis, 2008).For example, US National Aeronautics and Space Administration Anthropometric Source Book has defined the proportion of each part as fixed and the possible motion in his range. From the anthropometric ratios, the human head, neck and shoulder points of foreground segments can be estimated. The weighted-distance transform has estimated the remaining joints and limbs (H. P. Jain et al., 2011).

2.4.5 Silhouette-based human identification

Silhouette is a simplified diagram often for describing and explaining information about the object (Feldmann, 2011; Raskar & Cohen, 1999). For example, it is a benefits to presents a human body posture when preserving privacy is required (Juang, Chang, Wu, & Lee, 2009). Silhouette extraction provides a solid shape of the human body with the image object usually black in colour. For instance, the silhouette-based human identification approach that combines body height, width and body-part proportions, length of step and the amount of arm swing can handle noisy silhouettes with even a single stride of typical surveillance video data (Collins, Gross, & Jianbo, 2002). Nonetheless, the extracted silhouette boundary can be interrupted and a mistake caused by shadows and dresses. In addition, lighting and the contrast between clothing and background affects the accuracy of the identification result.

2.5 Morphological operations

The human body outline is not easy to extract even if the foreground object is segmented from the background. Most of the time the extracted silhouettes have small holes and other noises needing to be filled. Morphological operation including opening and closing are very common methods of filler for these holes and noise. (R.Hemalatha, E.Deepa, & R.Sasipriya, 2015; van den Boomgaard & van Balen, 1992; Yu & Aggarwal, 2011). Erosions and dilations occur often in morphological image processing, which result in skeletons obtained by a series of thinning (Haralick & Shapiro, 1992; Lam, Lee, & Suen, 1992). The dilation operation is the action of adding pixels around the object, therefore the image becomes wider, larger and more open reducing the size of the hole in the foreground image. Conversely, erosion is the action to remove pixels around the object, resulting in the image being gradually destroyed. Normally, erosion and dilations together refine the foreground object image

2.6 Sobel edge detector

Sobel edge detection operation the gradient measurement to emphasise image regions to edges (Mathur, Mathur, & Mathur, 2016). Sobel has two 3x3 kernels (2.2), Ix and Iy, where one is rotated by 90° relative to the other. Sobel Kernel convolution (2.2) is shown as two inverse matrices:

$$Ix = \begin{matrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{matrix} \qquad Iy = \begin{matrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{matrix} \tag{2.2}$$

Ix and Iy can be combined to find the absolute magnitude (Fisher, Perkins, Walker, & Wolfart, 2003). Therefore, gradient magnitude (2.3) is:

$$|GM| = \sqrt{Ix^2 + Iy^2} \tag{2.3}$$

2.7 Depth camera

One of the popular exergames is the Kinect device from Microsoft (Bačić et al., 2017; Rybarczyk, 2016; Wu, Quigley, & Harris-Birtill, 2017), which is also a motion capture system. The Kinect depth camera provides information to *DepthCam Sensor*. The information from the depth camera includes position and orientation of the camera, minimum and maximum value, the pixel value which is out of the minimum and maximum range and the dimension of the depth image. The raw image is in BGR24 format and provides a service to map to RGB pixels (Microsoft). Depth image benefits by avoiding the influence of environmental change including light, intensity and shadow (Bačić et al., 2017) while Kinect's sensors can provide depth maps and produce a human silhouette but their use is limited by power supply requirements and accessibility compared to mobile video (Matyunin, Vatolin, Berdnikov, & Smirnov, 2011).

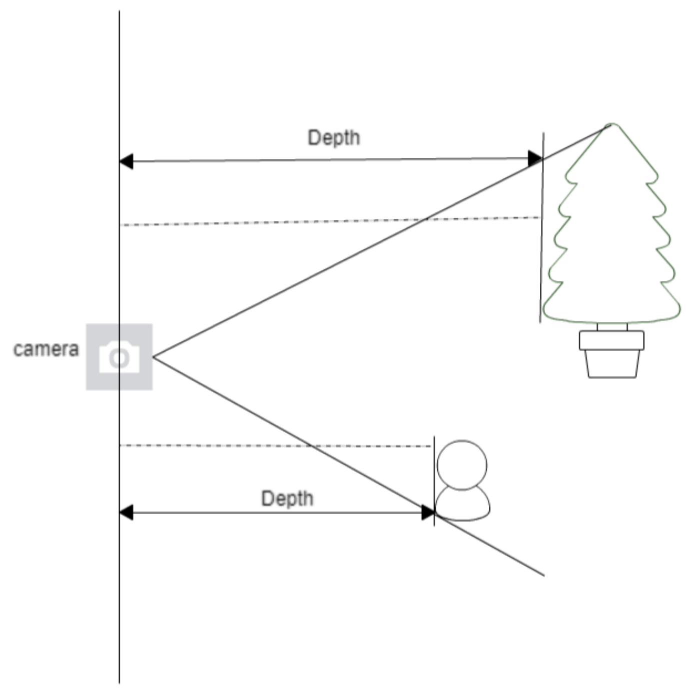


Figure 2.4: Depth camera, image modified from Microsoft depth sensor (Microsoft)

One of the studies by use ZCam from 3DV Systems, which technology seems to Kinect system. The depth sensor emits and reflects light of infra-red (IR) to calculate depth from a scene (Figure 2.4). From the depth image, the object distance can be distinguished by grey-scale proportion values(Bačić et al., 2017; H. P. Jain et al., 2011). The bright value represents the objects which are closer to the camera and, conversely, farther objects are represented by darker values (H. P. Jain et al., 2011). Evidently, the overlap issue can be solved by the horizontal and vertical grey-scale matrix, which also contains information for 3D space (Bačić et al., 2017). Evidently, the methods are employed to 3D scene reconstruction (Izadi et al., 2011; Matyunin et al., 2011).

The markerless motion capture device is adaptable which is beneficial to any environment. From author's prior co-published research (Bačić et al., 2017), Kinect can be a benefit in augmented coaching systems (*Kinect motion capture*). The user motion is simplified by displaying the human body skeleton. The recorded video provides information for rehabilitation and sports performance to analyse user motion.

2.8 Supervised and unsupervised machine learning

Unlike problem-specific types of algorithms, machine-learning algorithms are of generic nature, where obtained machine inference is derived from data in supervised, semi-supervised or unsupervised learning fashion. The examples of derived machine inference for classification purposes is data and context-specific such as handwriting recognition and spam email classification (Geitgey, 2014). In supervised learning scenarios (Figure 2.5), the machine-learning algorithms combine the input data with the output data (either as statistical ground truth or via expert labelling) to obtain machine inference that could subsequently enable model classification or prediction to autonomously solve similar problems on future input data (Geitgey, 2014; MathWorks). In computer vision and object detection contexts, *support vector machine* (SVM) is the one of the commonly used models associated with supervised classification techniques (Rezaei & Klette, 2017).

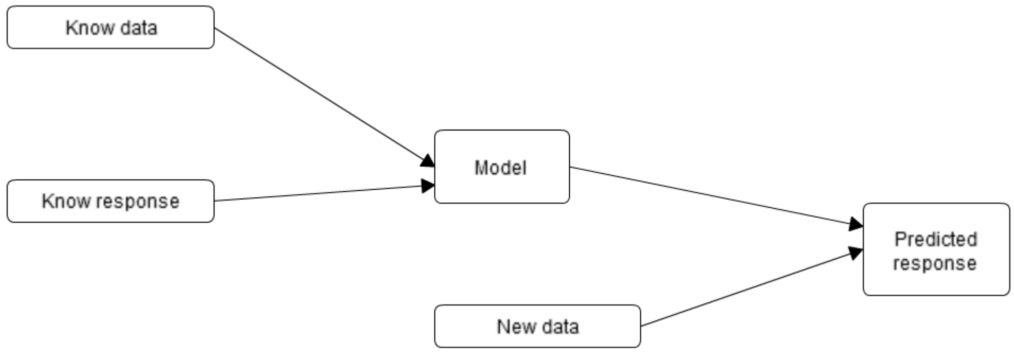


Figure 2.5: Supervised machine learning from data to build a model, and then predict the result (MathWorks)

Regarding unsupervised clustering and classification techniques popular example are: K-means clustering and Gaussian mixture models (Rezaei & Klette, 2017), which are both parts of various computer vision libraries (e.g. OpenCV and MathWorks Matlab™) commonly used for people detection and foreground-background separation.

2.9 Literature synthesis

The elements of swing performance and safety can be mutually exclusive, so in this research, it is considered that the role of a coach is to provide and communicate analytical feedback to the golf learner. In addition, captured image and video represent visual evidence of observed movement that can be used to augment coaching by facilitating analytical feedback between a coach and his/her learner.

The existing video and image processing methods covered in the literature review are expected to perform well with static backgrounds, which were considered as a starting point to help to produce a solution that could be used in augmented golf coaching context. Nevertheless, the static and dynamic pixel behaviour belonging to foregrounds or backgrounds that are typically found in the real-life environments associated with golf activities are considered a challenge in particular for people detection and foreground-background separation.

The main objectives of the intended solution include privacy preservation, elimination of redundant information, reduction of potential coaching bias and functionality for providing annotated visual feedback to improve golf coaching process. The research steps involve capturing golf activity videos and applying cyclic experimental design approaches from initial prototyping evaluations towards the development of the proof of concept.

Chapter 3 ***Methodology***

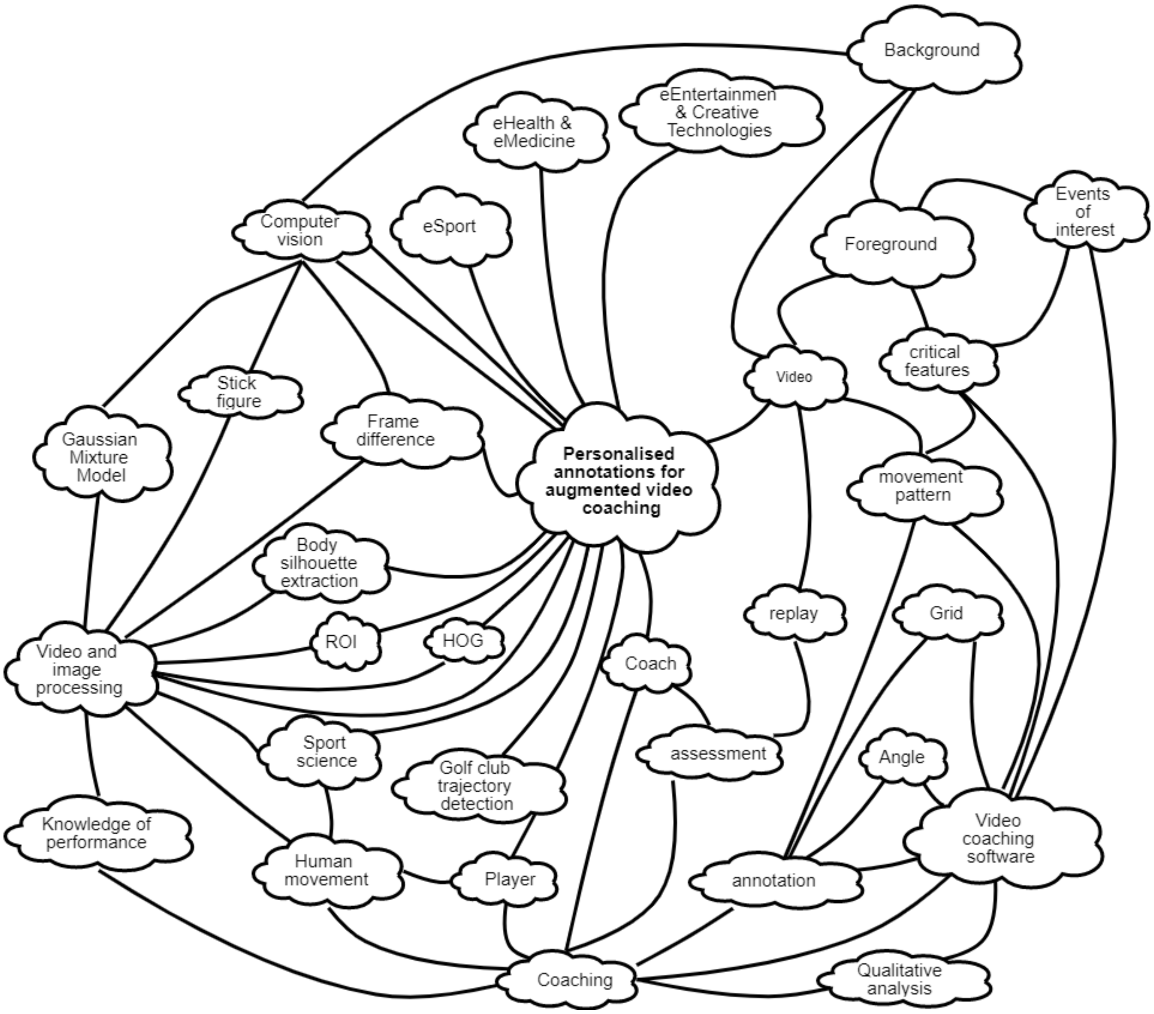


Figure 3.1: Mind map representing the complexity of the problem area and the associated multidisciplinary nature of the study

This chapter covers the data collection, experimental setups, equipment and evaluation of existing video and image processing approaches in coaching contexts covered in the literature. Aligned with design science methodology, the identified issues and drawbacks informed the experimental development stages that have led to the obtained solutions reported in the follow-up section.

The methodology is a cyclic, experimental approach that is focused on producing technological solutions to augment and facilitate golf coaching in outdoor environments (Baskerville, Pries-Heje, & Venable, 2009; Peffers, Tuunanen, Rothenberger, & Chatterjee, 2007; Wieringa, 2010). To produce the artefacts (such as algorithms, frameworks and architecture) supporting the golf coaching process, experiments were performed using various techniques from computer vision and video and image processing to present the mind map (Figure 3.1) influencing how the ideas were logically developed and organised from the literature review. This aligned with the cyclic nature of the employed experimental approach (Carlson & Bloom, 2005; Maxion, 2009), which supported evaluating the ideas and implemented proofs of concepts, from the initial stage towards achieving the final solutions (Figure 3.2).

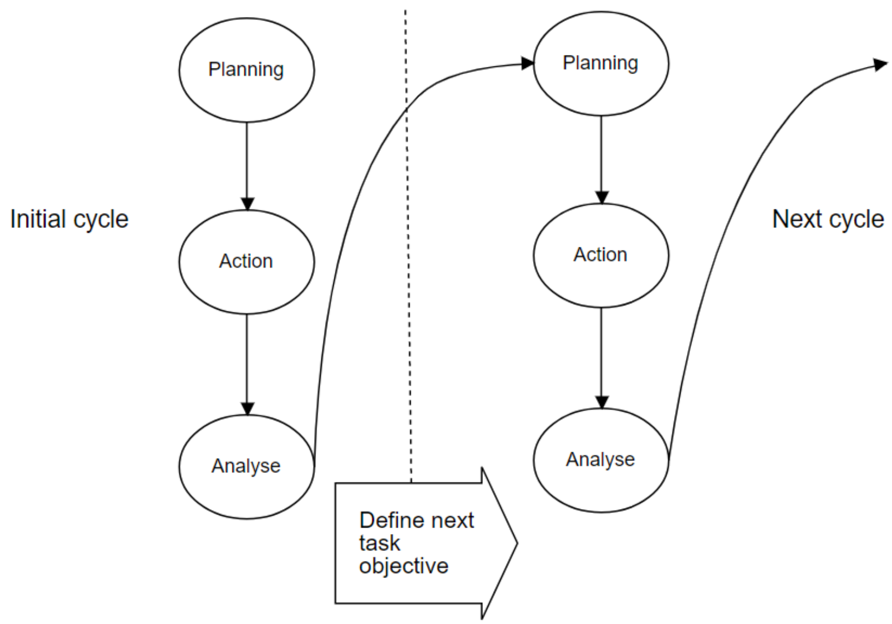


Figure 3.2: Cyclic nature of problem solving, modified from (Carlson & Bloom, 2005)

3.1 Experimental setup, tools and materials

The equipment and recording setup (Figure 3.3) specific to the golf data collection includes the following:

- Video sources and hardware: static camera using iPhone 7 plus and 5S, pocket-sized camera video Casio Exilim and GoPro Hero 3.
- Video settings:
    - Captured frame rates: 30 to 240 *frames per seconds* (fps).
    - High-speed cameras (240 fps) are required while capturing golf club high-speed movement (160/km/h) and 'impact' in accurate timing.
    - The experimental data include a variety of golf videos recorded in low and high resolutions and frame rates.
    - Camera view (placed on a tripod): side view – sagittal plane, camera height approximate hip and waist level allowing capture of full club movement.
- Software: Matlab 2016b for prototyping and algorithms.

(a)

(b)

Figure 3.3: Experimental setup (a) static camera using iPhone 7 plus, 5S camera, pocket-sized video including Casio Exilim and GoPro Hero 3. (b) The red line to emphasise the requirement of camera height to be adjusted to the hip level.

Note: To comply with AUT ethics regulations on self-reporting exception (http://www.aut.ac.nz/researchethics/guidelines-and-procedures/exceptions-to-activities-requiring-autec-approval-6), any video segment that would capture a person other than the author was removed from the experimental data. As a result, there were more golf swing videos captured from the rear (sagittal) view. In author's view to preserve privacy, the rear view should be more common in augmented video coaching given that the driving range settings permitting the front view video recording would require sufficient distance that would cover at least two driving range bays with separating panels to preserve customer's privacy.

3.2 Data collection and experimental design

All experimental data were collected in three golf sessions on two golf driving ranges with different backgrounds and orientations towards the sun. Each session was recorded at different times of the day and weather conditions. Data collection settings were aligned with the literature review and the golf-recording video protocol ("Leadbetter interactive," 2005). The collected video dataset also captured evidence of improved elements of swing technique promoting the need for personalised annotation and augmented coaching feedback. The cyclic evaluation involved incremental design of a set of tools to support visual annotation of performance and safety elements for augmenting video coaching.

Experimental design

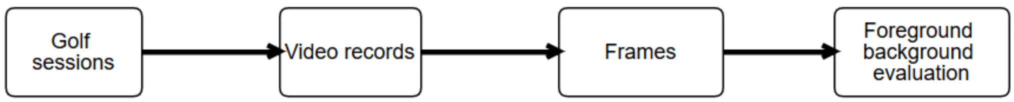


Figure 3.4: Workflow of the evaluation

The initial design was mainly focused on the foreground-background separation from the golf sessions before developing annotation capabilities of the final proof of concept. Figure 3.4 shows the workflow of evaluation of the existing foreground-background separation methods reported in Section 3.4.

3.3 Golf training sessions

Session 1 – Grip, feet and shoulders

Hold the golf club: there are three common grips: overlapping, ten-finger and interlocking, which constitute the foundation of the golf swing (T. J. Kelly, n.d.; Lamanna, 2017). Both hands must be kept solidly together but flexible when gripping the golf club. Initially, both feet should keep an even distribution of weight. Ideally, foot place distance approximates the shoulder width (Bartlett, 2007; Cole & Grimshaw, 2016; Knudson, 2013; Maddalozzo, 1987). Figure 3.5 shows the addressing phase by front view and side view.

(a) Front view specifies region of interest (ROI) on grip, not requiring privacy preservation filtering

(b) Side view (sagittal view)

Figure 3.5: Front and sagittal views of a beginner golfer during her swing preparation

Session 2 – Elements of the stance and swing basics

Based on the placement of feet linked a way to the suitable stand. The suitable narrow and width of the stand that lets the golfer turn freely, with stability and produces the greatest power (Bartlett, 2007; Maddalozzo, 1987). The right initial stance and position of posture help to maximise safety (Knudson, 2013). The golfer softly flex knees in order to maximise torso rotation and driving performance (Bartlett, 2007).

Figure 3.6: Sagittal view showing a degree of improvement regarding the knee angle

Figure 3.6 shows improved preparation and ball addressing compared to session 1.

Session 3 – Posture and swing improvements

Body posture for a beginner is the key factor in mastering golf swing. In the golf swing, good body posture as foundation, provided there is also natural adaptation regarding improvements in swing, balance, accuracy and consistency associated with the ball impact.

Figure 3.7: Addressing the ball (part of swing preparation), improved posture compared to session 1, (Figure 3.5 b)

Figures 3.5, 3.6 and 3.7 show qualitative evidence of the golfer's progress over three data collection sessions.

3.4 Evaluation of common foreground and background separation methods

Foreground-background separation algorithms play important role in many computer vision applications. The purpose is to detect the changes from the image series. In addition, the subtracted foreground only focuses on the region of interest (ROI) and removes the unwanted information as per the concept of Figure 3.8.

Compared with the indoor scenes, the outdoor golf driving range environment is difficult to produce an algorithm for separating the foreground from the background. In the case of outdoor golf sports activities, the scene includes static and dynamic foregrounds and backgrounds information. Moreover, the dynamic background is caused by shadow, moving clouds, waving water and moving leaves. In addition, the golfer can be representing static foreground. For instance, when golfer prepares to swing, there may be only a small amount of movement. In contrast, high-speed club movement represents dynamic foreground information.

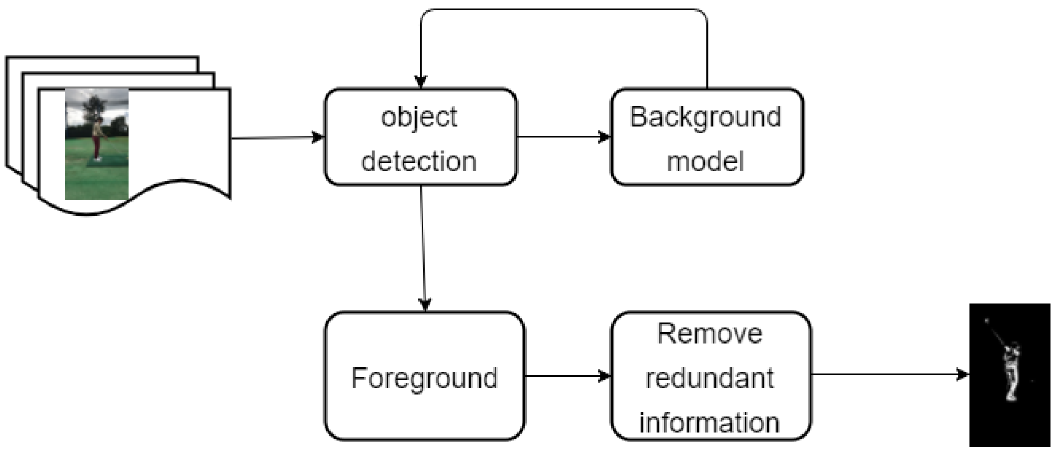


Figure 3.8: Privacy preservation filtering based on the foreground-background separation concept

3.4.1 Frame difference

Frame difference algorithm compares the current frame and a reference frame. The benefits of frame difference algorithm include relatively fast computing and simple processing requirements. Frame difference can detect changes between video frames including minor movement. Unfortunately, the algorithm can also detect minor changes in the background pixels such as running water and waving leaves.

Given that the frame difference algorithm for slow movement tends to produce dark images with little or no movement information, the approach may also include time difference parameter which is related to how many in-between frames need to be skipped for optimal results. Figure 3.9 (c) shows the extracted foreground image of the golf swing, which current frames and reference frame (a, b) are captured from the video using 240 fps.

(a) Reference image (b) Current image (c) Extracted foreground image

Figure 3.9: Extracted foreground image applying frame difference algorithm on 240 fps video

The above approach has universal application for fast frame-rate videos and slow body movement (compared to fast club movement) tracking such as found in the golf swing, supporting the concept shown in Table 5.

Table 5: Pseudo code for frame difference method

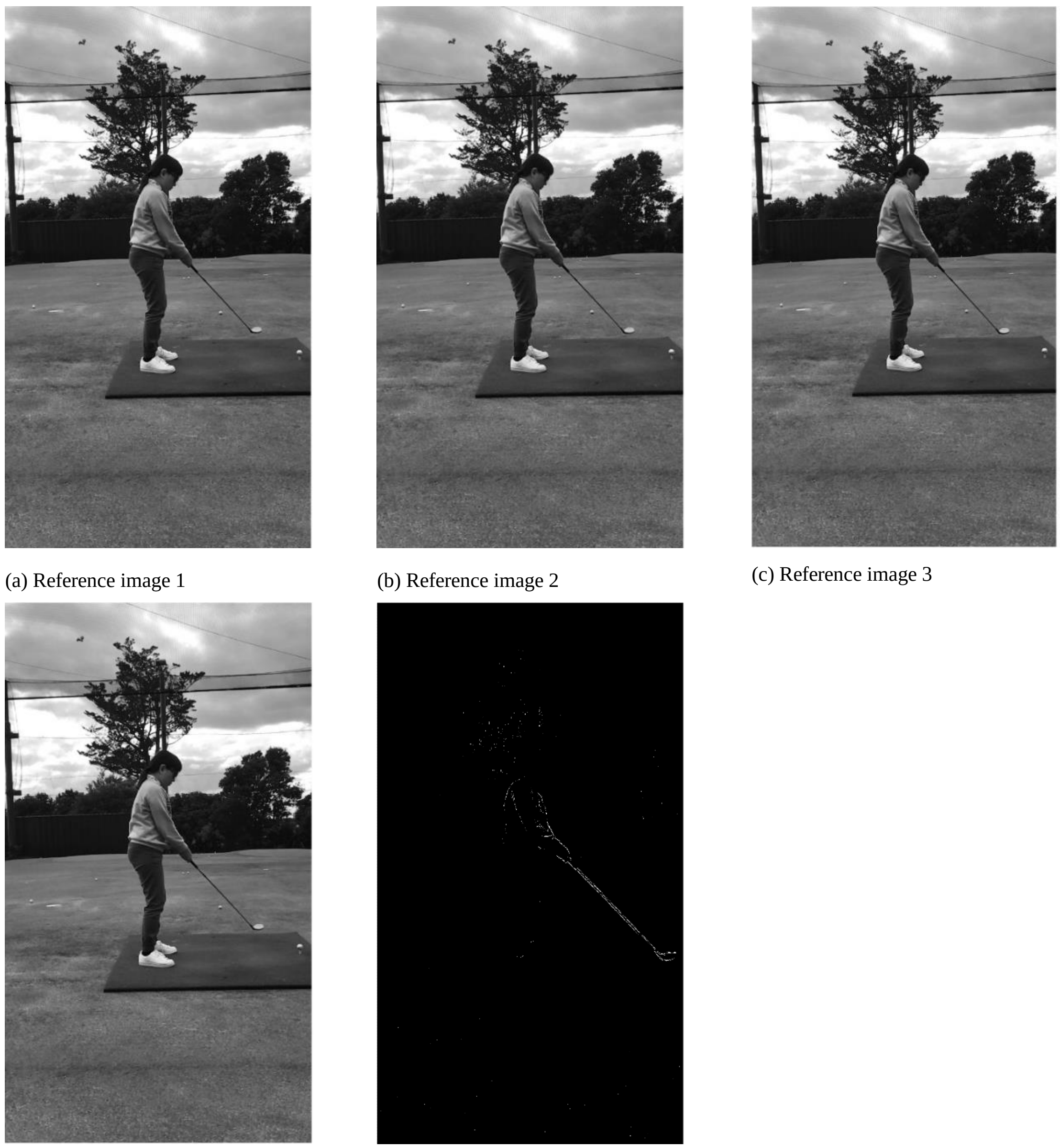
Algorithm FRAME_DIFFERENCE

input: video file video_file
parameters: bwlevel

```
videoObj = VideoReader(video_file)
grayFrame_prior = convert videoObj.frame1 to gray image
for i = videoStart to videoEnd
        rgbFrame = videoObj.read(frame(i))
        grayFrame = convert rgbFrame to gray image
        deltaFrame = grayFrame - grayFrame_prior
        grayFrame_prior = grayFrame
        bwFrame = cover deltaFrame to black and white
        foreground = remove small object from bwFrame
        show foreground image
end for loop
```

3.4.2 Difference of frame difference

The *difference of frame difference* algorithm is a variation of the frame difference algorithm, which resulting pixel values are computed by the difference between the current frame and the *reference frame*. The reference frame is the accumulative difference of two prior frames (Gonzalez & Woods, 2008, p. 627).



(a) Reference image 1

(b) Reference image 2

(c) Reference image 3

(d) Current frame

(d) Extracted foreground image

Figure 3.10: Difference of frame difference algorithm applied to extract foreground image on 120 fps video.

The *difference of frame difference* algorithm (Table 6) is also considered as an implementation of cumulative frame difference algorithm (Gonzalez & Woods, 2008, p. 627), with relatively higher sensitivity to minor changes including detection of waving leaves in the background. Figure 3.10 shows the extracted foreground image of the golf swing. Those frames are captured from the video with a frame rate of 120 frames per seconds (fps). Table 6 shows the difference of frame difference algorithm that produced 120 fps video filtering results (Figure 3.10).

Table 6: Pseudo code for time difference variation achieved by multiple frame difference

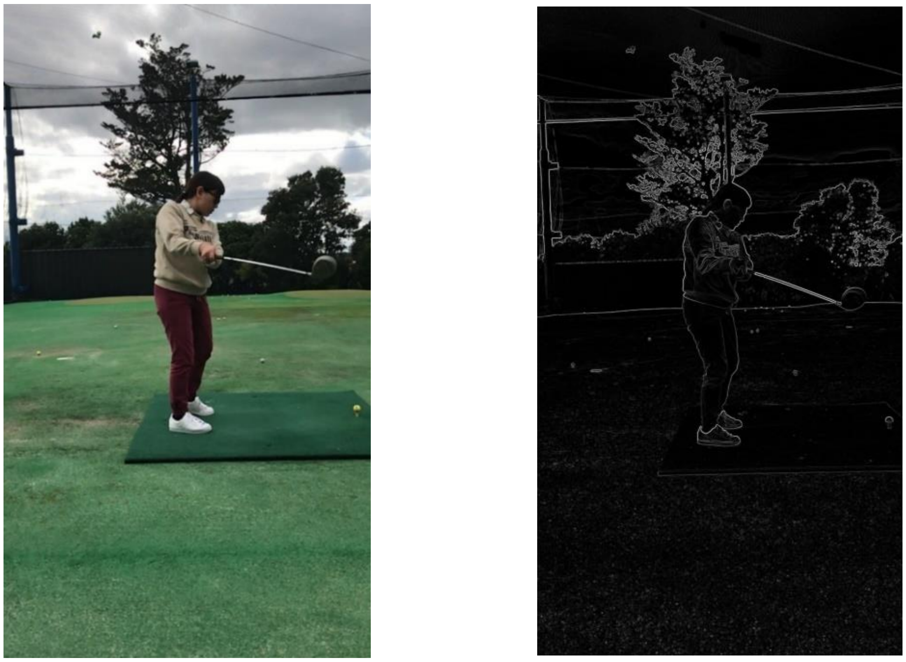
Algorithm DIFF_OF_DIFFERENCE

input: video file video_file
parameters: bwlevel

```
videoObj = VideoReader(video_file);
frame1 = rgb2gray(read(1st frame of videoObj))
frame2 = rgb2gray(read(2nd frame of videoObj))
frame3 = rgb2gray(read(3rd frame of videoObj))
for i = 4 to videoEnd
        rgbFrame = read(videoObj, frame(i))
        curFrame = rgb2gray(rgbFrame)
        bg_frame = frame2 - frame1
        deltaFrame = curFrame - frame3
        diffofDiff = deltaFrame- bg_frame
        frame1 = frame2
        frame2 = frame3
        frame3 = curFrame
        moveObj = remove small object from diffofDiff
        foreground = draw permietar of moveObj
        show the foreground image
end for loop
```

3.4.3 Edge detection by Sobel filter

Edge detection using a Sobel 3-by-3 filter emphasises recognition of horizontal edges and vertical edges. The concept is demonstrated in Table 7.



(a) Original RGB image

(b) Detected image

Figure 3.11: Edge detection using the Sobel method

The Sobel filter demonstrates good performance in detecting edges within the image. The detected golfer in the golf driving range is shown in Figure 3.11. However, the chosen golf driving range is considered a complex scene so the detected image also includes background information.

Table 7: Pseudo code for edges detection

Algorithm EDGE_DETECTION

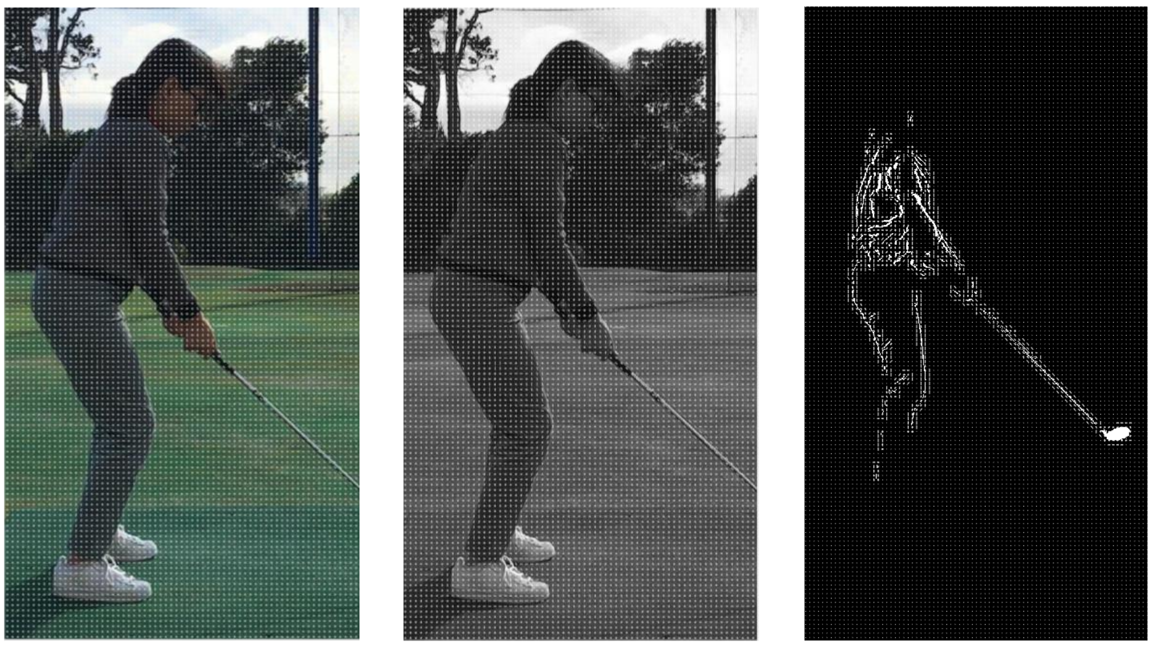
input: *R,G,B frames of an image* rgbFrame
output: *the gradient image* gradIm

Library functions:	**Purpose**
defFilter ('sobel')	*Using 'Sobel' filter*
imfilter()	*Filter the image again, this time specifying the replicate boundary option*

```
iframe = imread(rgbFrame)
grayImg = rgb2gray(iframe)
yEdge = defFilter('sobel')
xEdge = yEdge'
py = imfilter(double(I), yEdge, 'replicate')
px = imfilter(double(I), xEdge, 'replicate')
gradIm = √((px² + py²))
return gradIm
```

3.4.4 Histogram of oriented gradients (HOG)

Histogram of oriented gradients (HOG) feature implemented in Matlab (R2016b) is able to produce visualisation of gradient vectors using both true-colour, grayscale and binary image.



(a) Truecolor image plot with HOG

(b) Grayscale image plot with HOG

(c) GMM separated foreground image plot with HOG

Figure 3.12: Visualisation of histogram of oriented gradients (HOG) to separate foreground and background from a single image

The detected HOG feature visualisation can show the foreground object outline (Figure 3.12). However, in the chosen complex golf driving range environments, both HOG and Sobel edge detection would not produce favourable results from collected video frames.

3.4.5 Gaussian mixture models (GMM)

The GMM method tends to produce better results of foreground-background separation compared to two frame-difference algorithms (Figure 3.9 &Figure 3.10). However, when the foreground object (a golfer) only has minor movement over time, for example when preparing state of the golfer, the object will disappear (Figure 3.13 c).

(a) Reference frame at time position: 42.108

(b) Current frame at time position: 42.5471

(c) Extracted foreground image of (b)

Figure 3.13: Foreground detection using Gaussian mixture models showing darker foreground representation for slower movements during the swing

In Figure 3.13, image (a) and image (b) have 0.4391 seconds of time difference indicating a small amount of body movement. Depending on movement, GMM performance will vary between the frames (a vs b). The concept of Gaussian mixture models for foreground-background separation is shown in Table 8.

Table 8: Pseudo code for Gaussian mixture models (GMM)

```
Algorithm GMM_SEGMENT
Input: foreground image frame videofile
Parameters:   NUMGAUSSIANS = 20
              NUMTRAININGFRAMES = 10
              MINIMUMBACKGROUNDRATIO = 0.85
Library functions:                Purpose
Gaussians_foregroundDetector()    Foreground detection using Gaussian mixture models

foregroundDetector = Gaussians_foregroundDetector(NUMGAUSSIANS,
NUMTRAININGFRAMES, MINIMUMBACKGROUNDRATIO)
videoReader = read video file(videofile)
while videoReader hasFrame
    frame = readFrame(videoReader)
    foreground = foregroundDetector(frame)
    filled = fill image holes of foreground
    cleaned = remove small object
    show the cleaned image
end
```

Image ghosting phenomenon of GMM

The experimental results are aligned with the literature. As a known phenomenon in video and image processing, *ghost* images include the falsely detected virtual objects, which are not equivalent to actual real object(s) captured in the current frame (Kryjak et al., 2014). Ghost images are caused by the object being initially static or temporarily static and starting to move again (T. Huang, Guo, Qiu, & Ikenaga, 2009).

The experimental result on the golf driving range are shown in Figure 3.14. When the golfer prepares to swing in a static posture or with a very small amount of movement, it is classified as a background model. However, the "real" static object and the 'ghost' image share similar properties.

(a) Original image (b) Ghost image found in GMM extracted foreground

Figure 3.14: A phenomenon of "ghost" image appearance due to static foreground transitioning to movement and GMM parameter settings.

As part of the experimental work on collected golf data, it was demonstrated that the appearance of ghost phenomenon (Kryjak et al., 2014) can be solved by GMM parameter tuning.

'Ghost' image found in GMM can be adjusted by the learning rate

(a) GMM LearningRate set to **0.01**

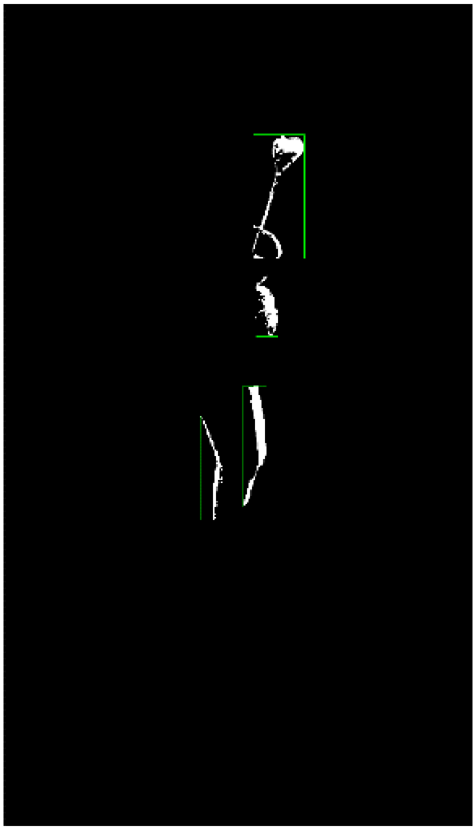
(b) GMM LearningRate set to **0.05**

Figure 3.15: Eliminating 'Ghost' image by adjusting GMM LearningRate parameter

In Matlab, foreground detection use Gaussian mixture models with *learning rate* parameter, which is specific for the implemented model. Learning rate parameter (or property) is used to control how fast the model adapts to the changing conditions. The learning rate (LearningRate) default value in Matlab (R2016b) is 0.005. Figure 3.14 shows the foreground image separation using GMM with the default LearningRate of 0.005. Figure 3.15 shows the improved results of LearningRate (a) 0.01 and (b) 0.05 achieving ghost artefact elimination. In comparison with the 'ghost' elimination problem, the visual appearance of shadows are also adjusted by increasing the LearningRate value. However, parts of the foreground image object are also likely to be removed.

3.4.6 Computing statistics for labelled regions: Blob analysis

This algorithm (Table 9) employs BlobAnalysis object to find the image connected regions in image. The MinimumBlobArea properties use to compute the lengths of the ellipses' minor axes (*vision.BlobAnalysis System object*), and can help to exclude the small objects contained in the image.



(a) The algorithm can find the moving golf club head

(b) The algorithm can find the golfer body parts

Figure 3.16: Matlab BlobAnalysis object is to compute statistics for connected regions

Figure 3.16 shows the foreground image extracted by GMM and the BlobAnalysis object returning the bounding boxes. In this experiment, the 'MinimumBlobArea' is set to 1000 to produce favourable visual results.

Table 9: Pseudo code for blob with GMM

Algorithm BOLB_GMM

Input: rgb frame video, videofile
Output: the [x y width height] bounding box coordinates, bbox
Parameters: MINIMUM_BLOB_AREA = 1000
NUMGAUSSIANS = 5
NUMTRAININGFRAMES = 150
MINIMUMBACKGROUNDRATIO = 0.85
BOUNDINGBOXOUTPUTPORT = 'true'
AREAOUTPUTPORT = false
CENTROIDOUTPUTPORT = false

Library functions:	*Purpose*
Gaussians_foregroundDetector()	*Foreground detection using Gaussian mixture models*
BlobAnalysis()	*Computes statistics for connected regions in a binary image*

```
foregroundDetector = Gaussians_foregroundDetector(NUMGAUSSIANS,
NUMTRAININGFRAMES, MINIMUMBACKGROUNDRATIO)
videoReader = vision_videoReader(videofile)
blobAnalysis = BlobAnalysis(MINIMUM_BLOB_AREA)
while videoReader hasFrame do
    frame = readFrame(videoReader)
    foreground = foregroundDetector(frame)
    filled = image fill hole(foreground)
    cleaned = image open (filled)
    bbox = blobAnalysis(foreground)
    if (bbox is not empty) then
        show the cleaned image with bbox info
    else
         show the cleaned image
    endif
end
```

3.4.7 People detector with GMM

People detector (PeopleDetector) use HOG features and trained Support Vector Machine (SVM) provided in Matlab (Table 10). Normally PeopleDetector is trained to detect upright people. It does not work well detecting people that are not in the front face.

(a) People detector used in RGB image

(b) Most of the GMM extracted foreground image cannot be detected by people detector

(c) GMM extracted foreground image with people detector

Figure 3.17: Using Matlab People detector to detect side view of upright people

From the experimental result in the golf driving range, most of the time PeopleDetector could not detect the golfer from the sagittal view shown in the Figure 3.17. In the experiment, the detector not only detected people in the RGB frame, but other moving objects such as waving leaves could also be detected in true-negative. Likewise, PeopleDetector could not achieve on the GMM extracted foreground image, only a few frames could be detected (c). The statistical data is from the experiment results, which were extracted from a part of the video, in total 263 frames, using the PeopleDetector to detect the GMM extracted foreground image and only 10 frames successfully detected the image object. Table 10 shows the pseudo code for this method.

Table 10: Pseudo code for people detection, modified from Matlab (MathWorks)

Algorithm PEOPLE_DETECTION

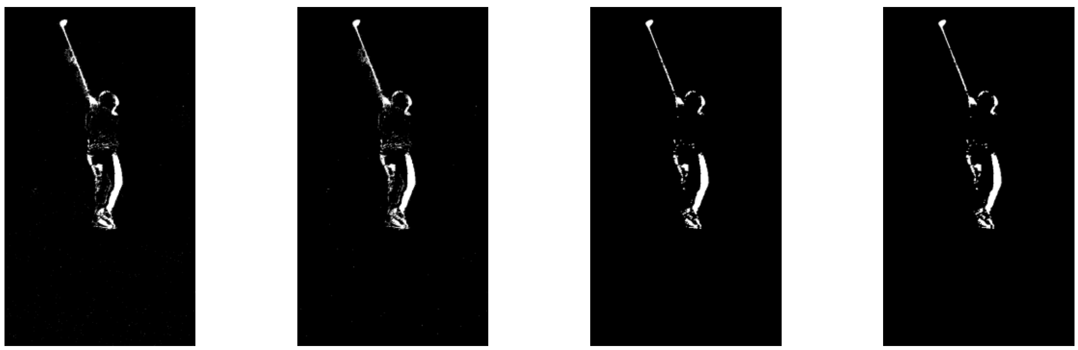
Input: rgb image video, videofile
Output: N/A
Parameters: CLASSIFICATIONTHRESHOLD = 0
MERGEDETECTIONS = true
NUMGAUSSIANS = 5
NUMTRAININGFRAMES = 60

Library functions:	*Purpose*
vision.PeopleDetector	Detect upright people using HOG features
Gaussians_foregroundDetector()	Foreground detection using Gaussian mixture models

```
peopleDetector = vision.PeopleDetector (CLASSIFICATIONTHRESHOLD,
        MERGEDETECTIONS);
foregroundDetector = Gaussians_foregroundDetector(NUMGAUSSIANS,
NUMTRAININGFRAMES)
videoReader = vision_videoReader(videofile)
blobAnalysis = BlobAnalysis(MinimumBlobArea = 500)
while videoReader hasFrame do
    frame = readFrame(videoReader)
    fground = foregroundDetector(frame)
    calculate the confidence value, set [bboxes,scores] = step(peopleDetector, ff)
    if scores and bboxes also not empty
        Img = insertObjectAnnotation(fground,'rectangle',bboxes,scores)
        show image Img
    end if
        show image fground
end
```

3.5 Evaluation of morphological operations on binary images

Morphological operations could be combined to achieve optimal foreground-background separation. Figure 3.18 shows visual results applying different combinations of morphological transformations that are categorised in Matlab's (R2016b) 'bwmorph' function is include Clean, Open, Close and Thicken operation.



(a) Clean
Removal of isolated pixels

(b) Open
Dilation after erosion

(c) Close
Erosion after dilation

(d) Thicken
Adding pixels to the outer surface of the object of interest

Figure 3.18: Evaluation of some morphological operations on collected data

Fundamental morphological processing including dilation and erosion, and the Matlab 'bwmorph' function are widely used methods (Bao, Guo, Tang, & Zhang, 2009) to remove redundancy.

$$\begin{matrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{matrix} \quad \Rightarrow \quad \begin{matrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix}$$

Figure 3.19: Using a clean operation to remove the centre pixel

The 'clean' operation of 'bwmorph' removes the isolated pixel '1' from surrounded by '0', which may be used to remove unwanted small areas from the foreground image (Figure 3.19).

3.6 Experimental Findings

Various types of algorithms used in the initial experiment may not perform well to separate the foreground from its background, in particular when a golfer produces a small amount of movement. Table 11 shows the experimental result summary of the foreground-background separation methods covered in the literature.

Table 11: Comparison between different methods for foreground-background separation

Foreground-background separation method	Experimented evidence	Positive	Negative	Reference
Frame difference		Efficiency and lower computational effort. Well performed in the static background situation.	Unwanted minor change of the background including running water and waving leaves also can be detected.	(Ananya et al., 2015; Archetti et al., 2006; D Stalin Alex, 2014; Gonzalez & Woods, 2008; Singla, 2014; Szeliski, 2011; Vedaldi & Fulkerson, 2010)

Foreground-background separation method	**Experimented evidence**	**Positive**	**Negative**	**Reference**
Gaussian mixture models	 	GMM has good performance to segment a foreground image object with boundaries. A better result can be adjusted by the initial parameter and training data. Better performance than frame difference algorithm.	The static object can cause a poor foreground estimation. Ghost image found when the object is static and moves again.	(Archetti et al., 2006; Bouwmans et al., 2008; S.-S. Huang et al., 2007; T. Huang et al., 2009; A. Jain et al., 2014; Kryjak, 2014; Rezaei & Klette, 2017; Stauffer & Grimson, 1999; Wang et al., 2014)

Foreground-background separation method	Experimented evidence	Positive	Negative	Reference
Sobel edge detection		Sobel performs well for the edge detection tasks.	The resulting includes the redundancy background information.	(Fisher et al., 2003; Kryjak et al., 2014; Mathur et al., 2016; R.Hemalatha et al., 2015)
Histogram of oriented gradients (HOG)		HOG features can extract both from a truecolor or grayscale input image.	It is not suitable for the golf driving range complex backgrounds.	(Dalal & Triggs, 2005a; Gonzalez & Woods, 2008; Rezaei & Klette, 2017)

Chapter 4 ***Developed solution and results***

This chapter presents the developed solution with insights and supporting qualitative evidence related to the advancements of augmented video golf coaching systems.

All captured and transformed videos (except for Figure 3.5 a) of golf activities are presented in sagittal view and in compliance with AUT Ethics Committee (AUTEC) requirements regarding self-recording and AUTEC approval exceptions. Sagittal view is a common and convenient camera vantage point in environments such as golf driving ranges where there is little likelihood that a random bystander may walk into the camera view.

One of the objectives of this thesis is to provide privacy preservation in video while keeping visual elements that are needed to support coaching process and diagnostic capabilities. Since a privacy-preserving solid and plane silhouette emphasises the outline only, the developed algorithms producing pseudo-3D binary silhouettes that can provide enhanced visual representation of a golfer's activity.

The benefits of developed pseudo-3D silhouette-based solutions for augmented coaching include:

- Privacy preservation
- Integration and application of widespread low-cost, portable, affordable and immediately-available video technologies such as smartphones, tablets, digital video cameras and other portable computing devices
- Ecological, paperless and paper-based reporting alternatives integrating visual evidence of captured human movement.

A developed pseudo-3D binary silhouette extracted foreground image from video is shown in Figure 4.1.

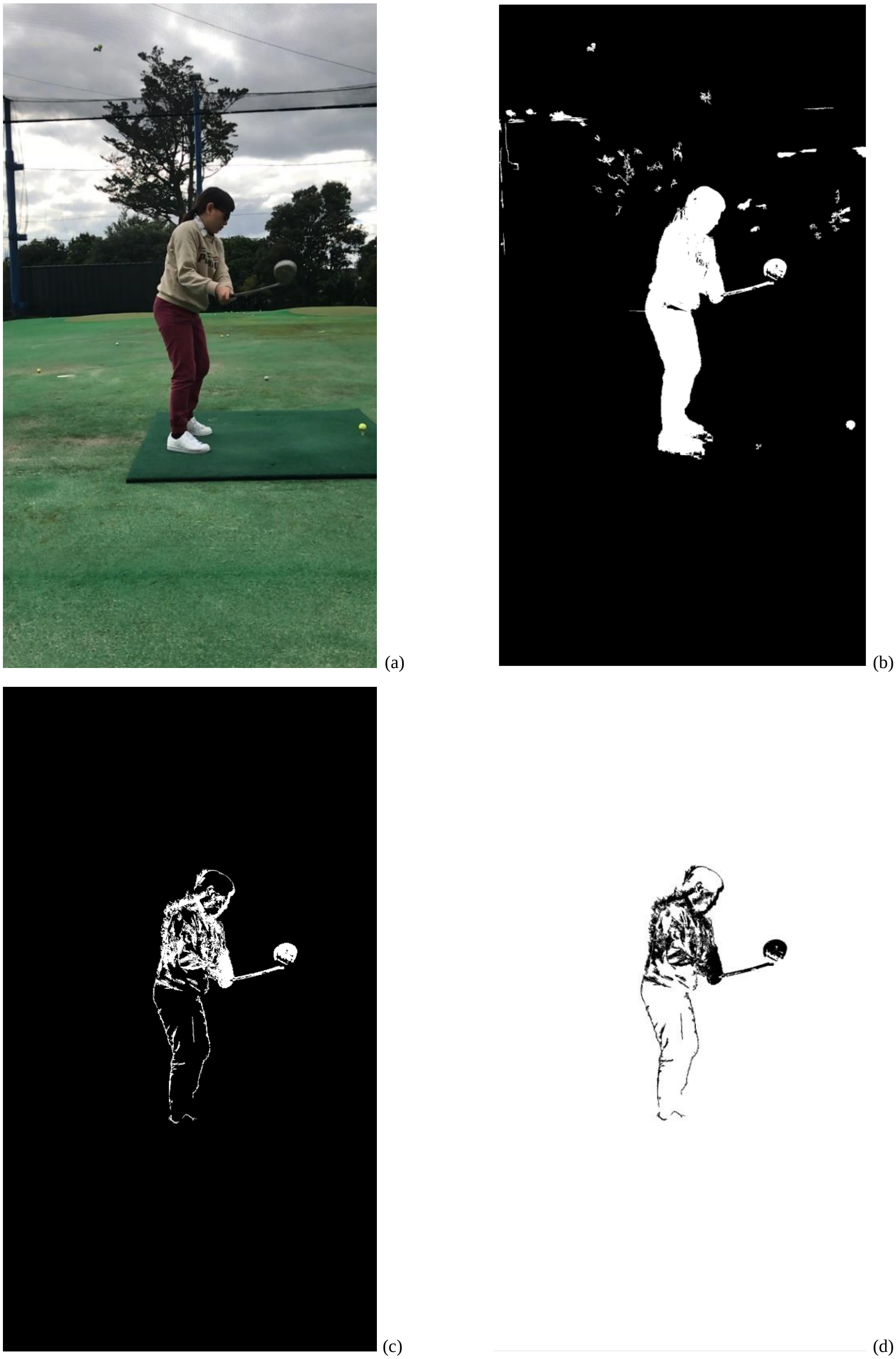


Figure 4.1: Extracted foreground image object (a) RGB image, (b) plan and solid silhouette, (c) developed solution is suitable for screen annotations and viewing, (d) inversed black and white image from (c) that is optimised for hard copy static reporting.

The plan and solid silhouette can only represent the outline of the image object. The developed solution shows the result is close to 3D proximity, which can display in additional details of the image being needed for qualitative diagnosis, for example, the ear, elbow and leg detail shown in the Figure 4.1 (c) and (d).

4.1 Function of the augmented video coaching application

The first processing stage of the application is the foreground and background segmentation from the sequence of video images, in order to remove redundant information. The follow-up user interactive stage provides annotation capabilities to help a coach to communicate and qualitatively analyse of the golfer posture. The annotation including interactive angle calculation (Section 4.4), 3D perspective grid annotation on the 2D image (Section 4.5), golf club head trajectory (Section 4.6), and the analysis report is in the final section (Section 4.7). Workflow of the developed solution for augmented video coaching is shown in Figure 4.2. In addition, Figure 4.3 depicts a screenshot of the developed GUI prototype for the augmented video solution for golf coaching.

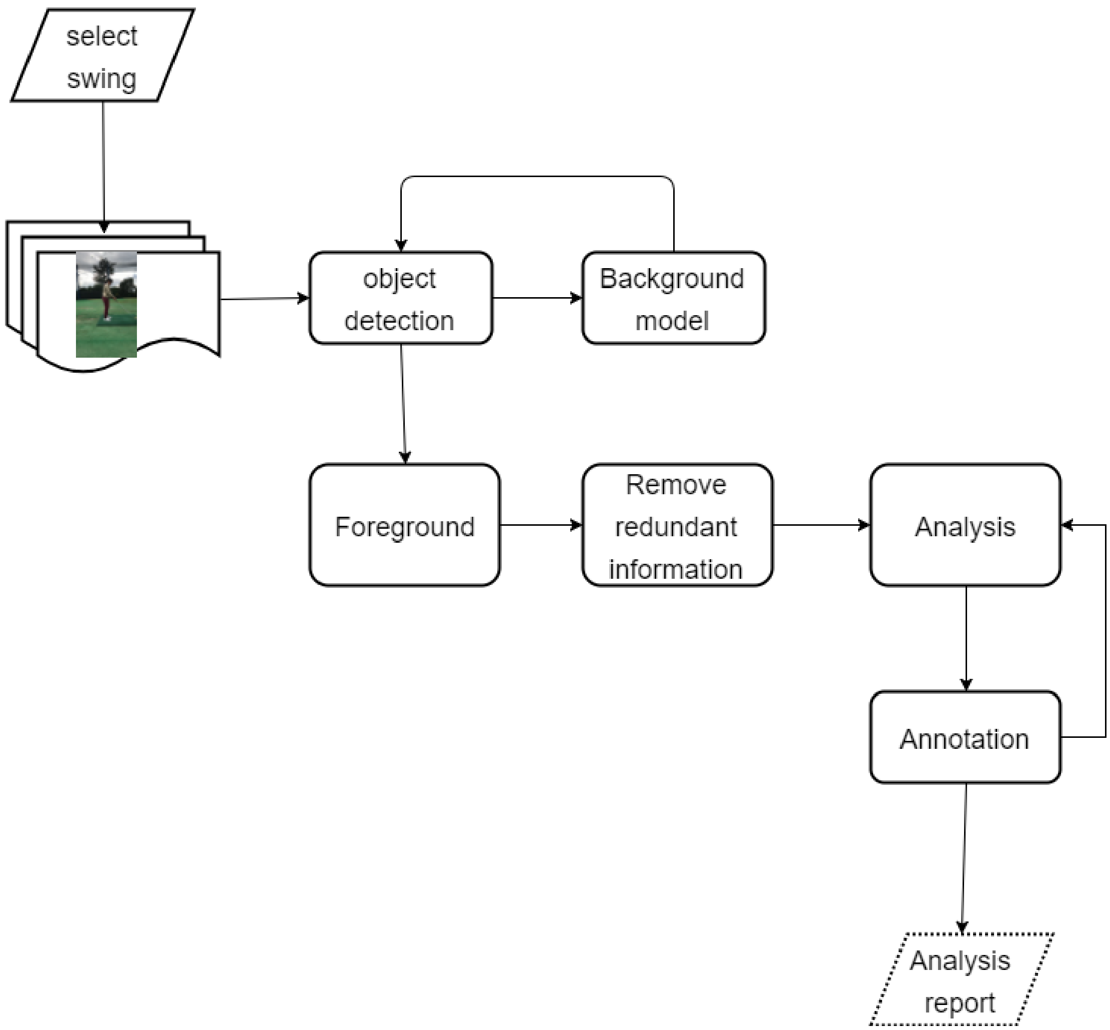


Figure 4.2: Workflow of the developed solution for augmented video coaching

The developed solution provides seven interactive functions allowing end users to:

- Select video
- Provide privacy preserving replay
- Select a swing or a frame of interest
- Draw lines and calculate joint angles
- Draw a grid
- Draw a golf club head trajectory
- Extract posture as pseudo -3D silhouette evidence
- Produce an analysis report

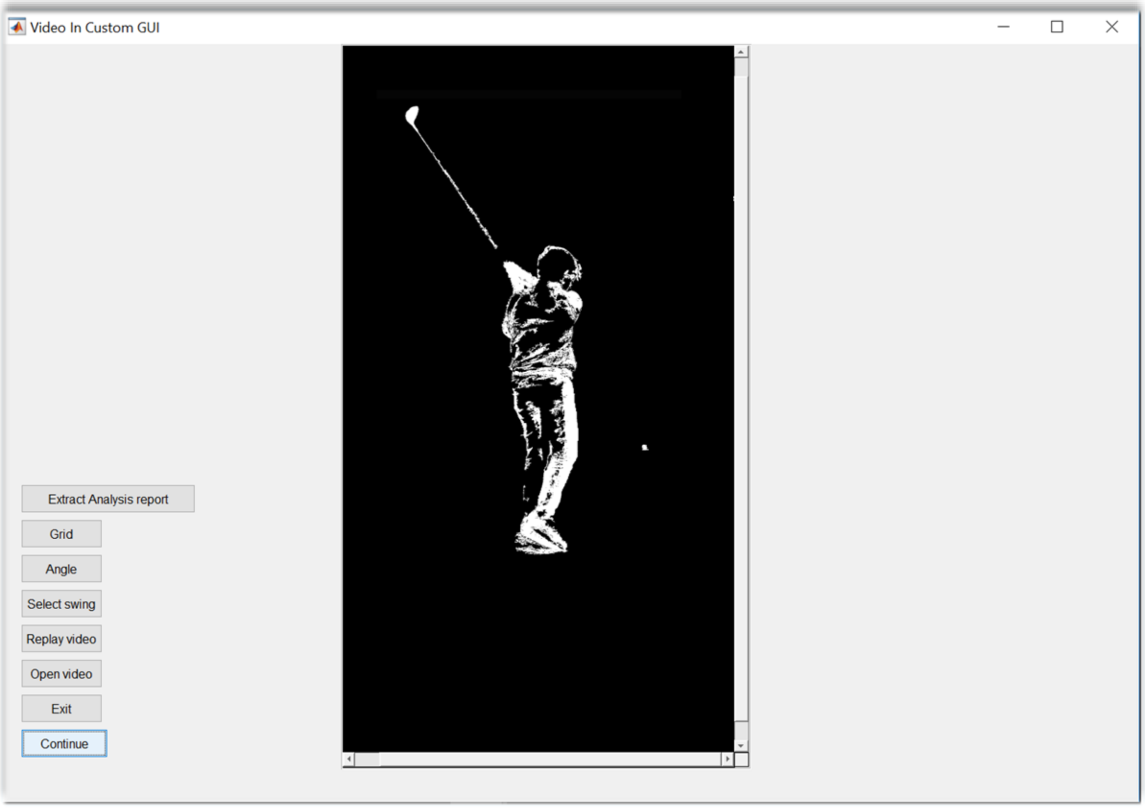


Figure 4.3: A screenshot of the developed GUI prototype for the augmented video coaching solution

4.2 Privacy preservation digitalisation visual models

The digital image record represents visual evidence of sports activities. Fast movement patterns such as golf swing can be captured by a high-speed video camera. The developed pseudo-3D binary silhouette solution in Figure 4.4 shows the frame sequence that preserves the player's privacy as well as the critical features required for video analysis and diagnostic feedback.

Figure 4.4: Digitalisation records of golf activities (a-d: backswing, d-g: downswing, h-k: Follow-through)

4.3 Foreground-background separation

Foreground-background separation is implemented as spatiotemporal processing stages (Figure 4.5).

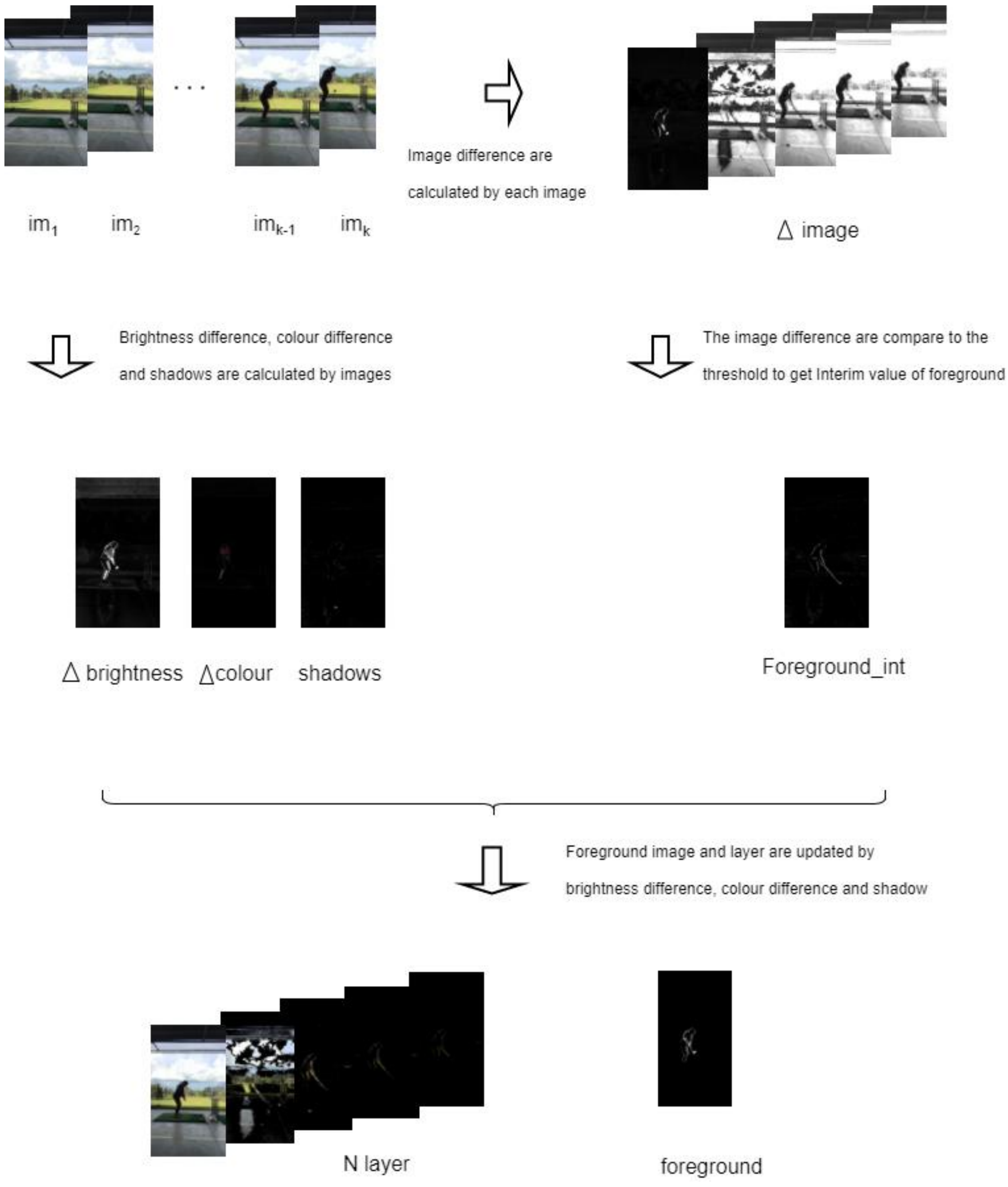


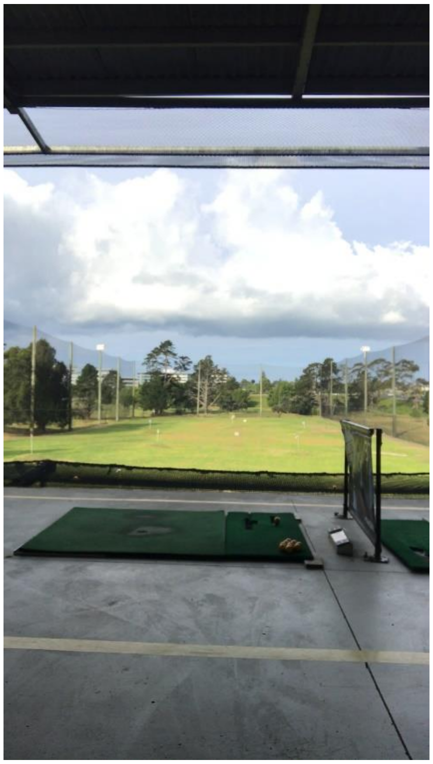
Figure 4.5: Multi-layered approach to convert the original frame sequence into a pseudo-3D binary silhouette video.

A multi-layered approach has been used to segment the foreground and background, which is already used in multimedia processes (*GIMP*; *Photoshop*; *VirtualDub*). The presented approach also represents a modified solution developed for a tennis case study (P. Kelly et al., 2010). However, outdoor golf driving ranges have more complex backgrounds than tennis courts' with uniformly coloured surface backgrounds.

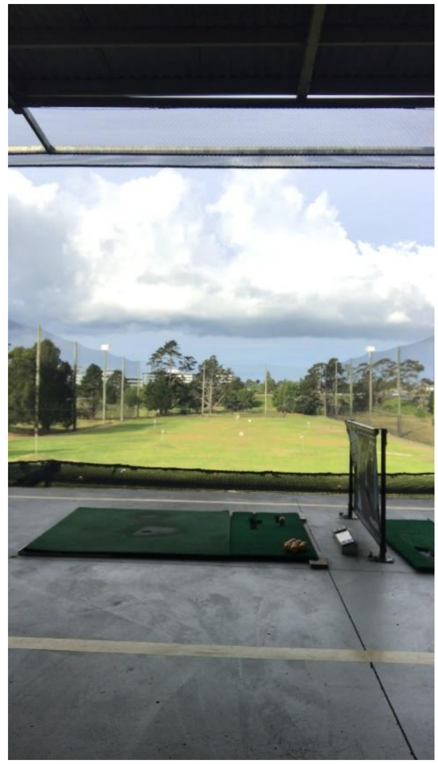
In order to achieve foreground-background separation, the approach is based on storing each frame into a 3D array. Each pixel in the 3D array will be classified to be a background or foreground relying on image differences, brightness differences and colour differences. Figure 4.5 presents the multi-layered approach segmentation of the foreground and background relying on the following:

- Image differences computed from the images
- Image differences compared to the threshold to obtain the interim value of the foreground
- Brightness differences, colour differences and shadows computed from the images' sequence
- Brightness differences, colour differences and shadows combined to update the interim foreground image and layer

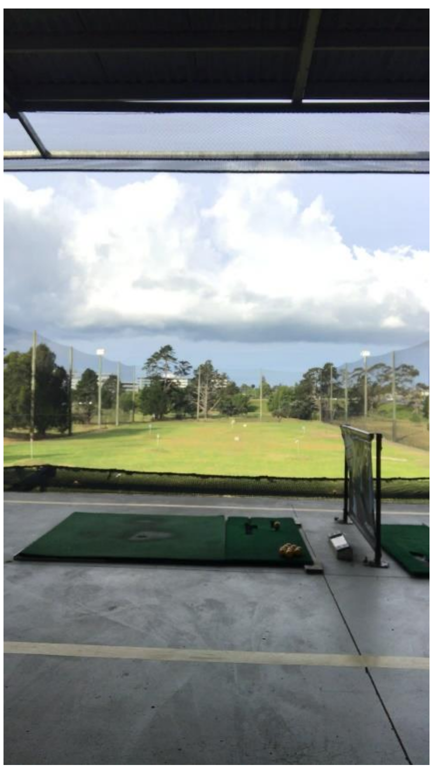
The six source images are shown in Figure 4.6 and Figure 4.7 presenting images of five layers for computing the foreground image.



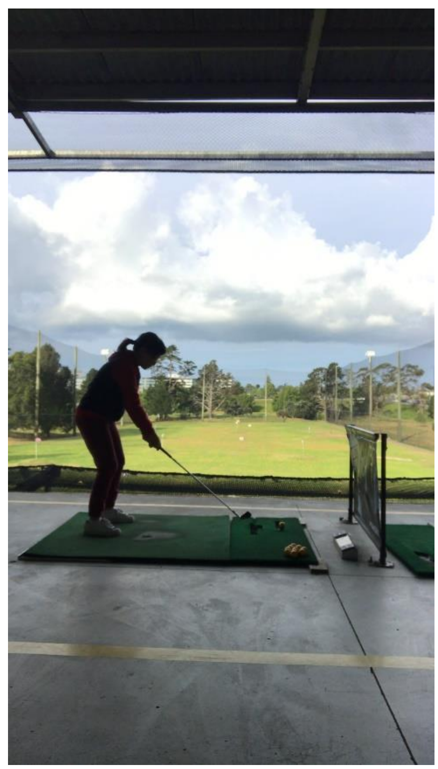
(a) 1st image



(b) 2nd image



(c) 3rd image

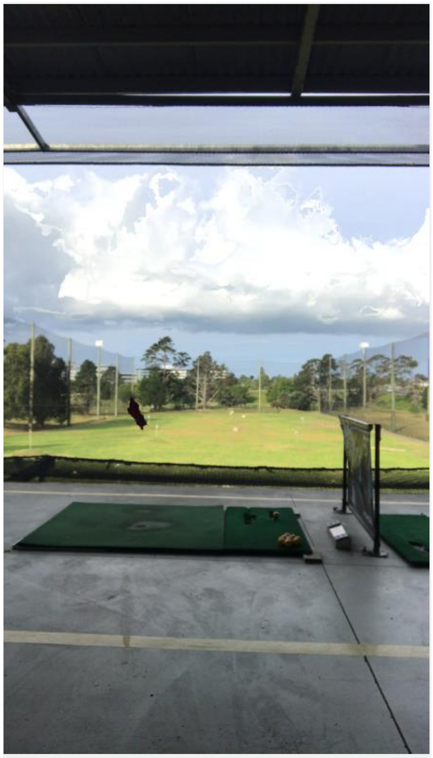


(d) 4th image

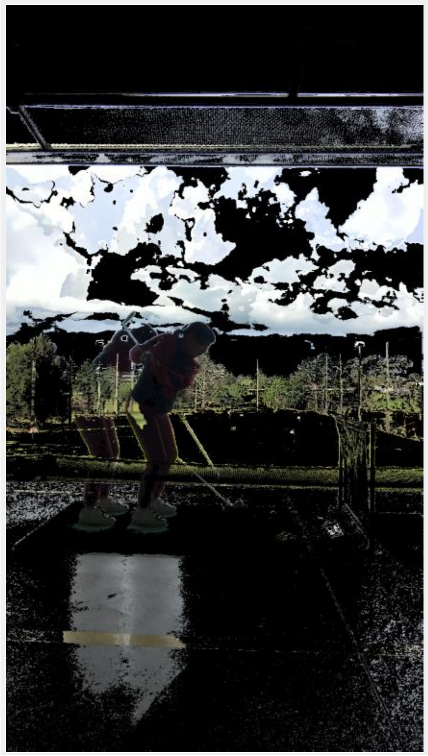
(e) 5th image

(f) 6th image

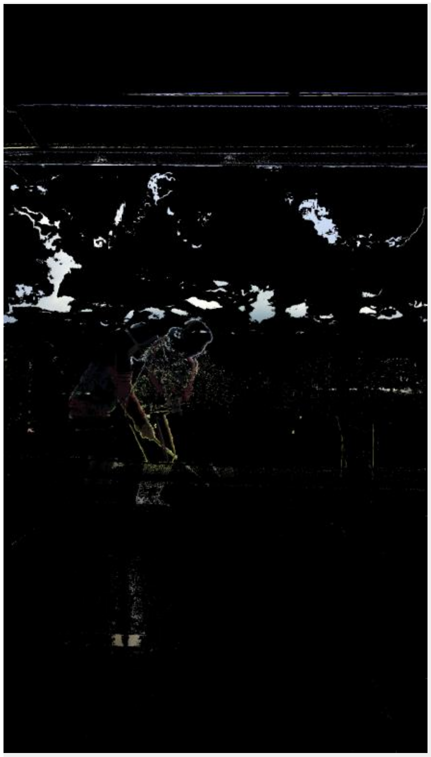
Figure 4.6: Source images



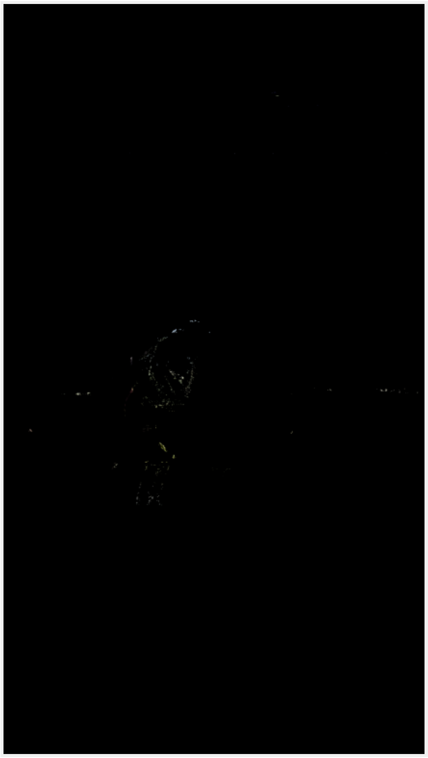
(a) 1st layer, max value = 255



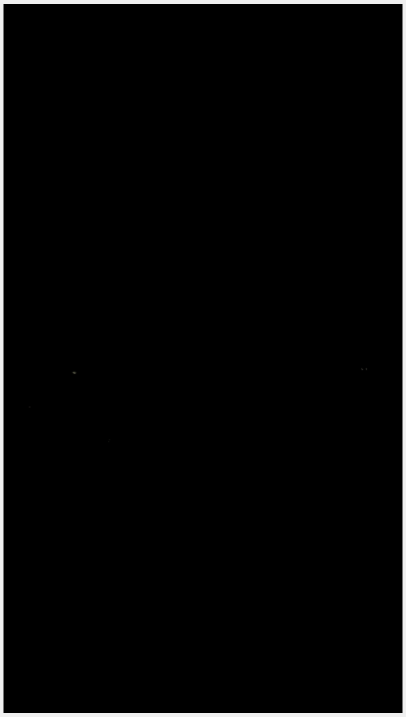
(b) 2nd layer, max value = 255



(c) 3rd layer, max value = 255



(d) 4th layer, max value = 255



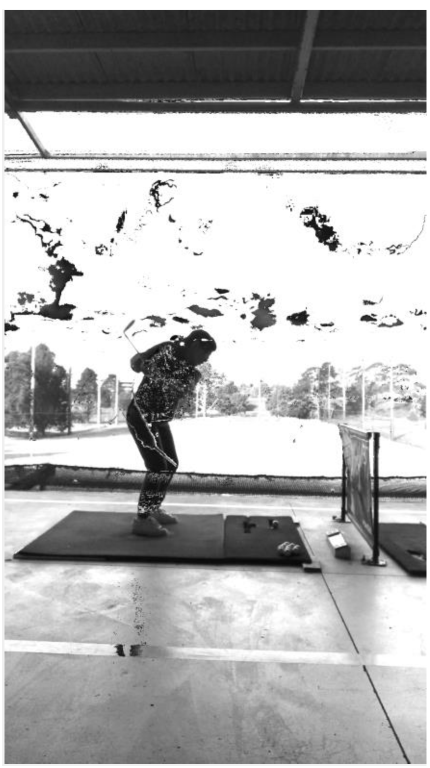
(e) 5th layer, max value =244

Figure 4.7: Five layers images

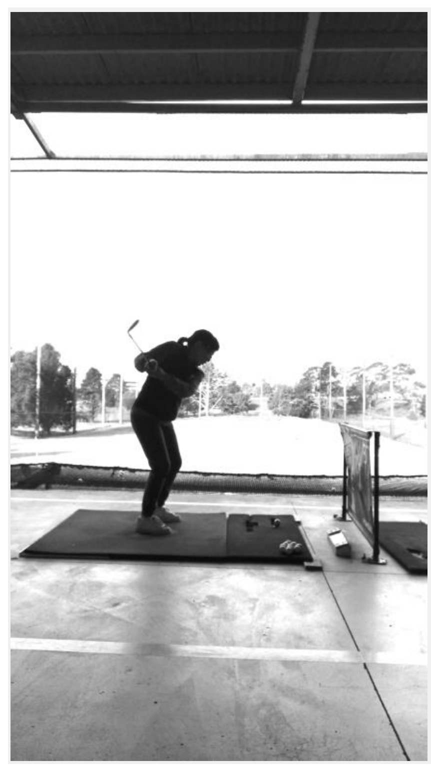
The image differences of each layer are presented in Figure 4.8, which are calculated from the source images (Figure 4.6).

(a) Image differences of layer 1

(b) Image differences of layer 2



(c) Image differences of layer 3

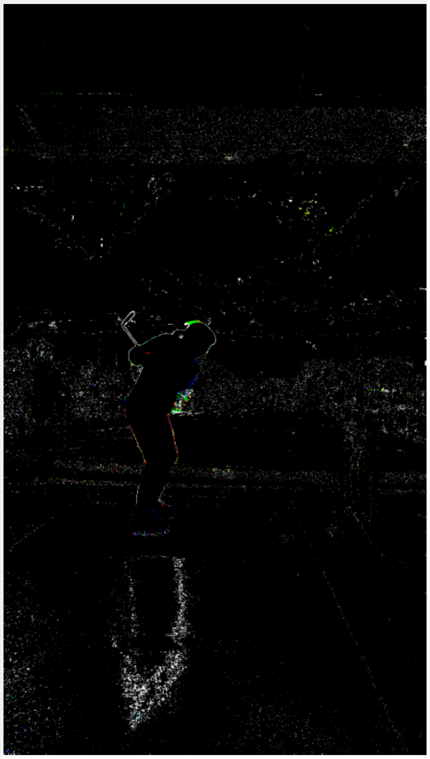


(d) Image differences of layer 4

(e) Image differences of layer 5

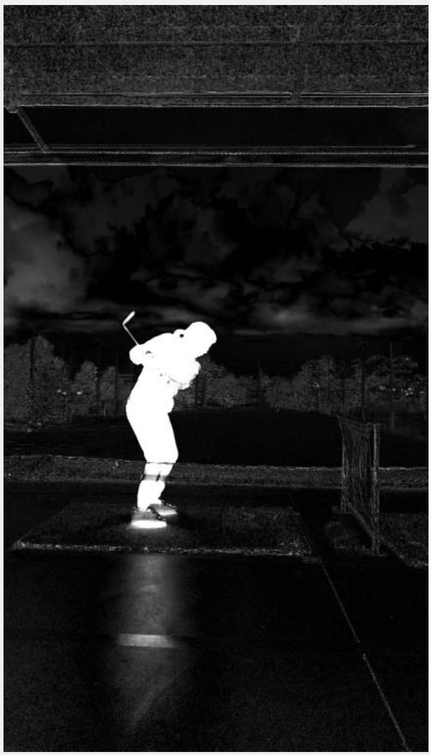
Figure 4.8: Image differences between each layer

Figure 4.9 shows the interim values of the layered approach, which includes shadows, colour differences, brightness differences, background RGB and all accumulated differences of the resulting image of the layer with the removed threshold. Figure 4.10 presents the extracted foreground image (a) and after the final step, removing the redundant information to get the clear binary image object (b).

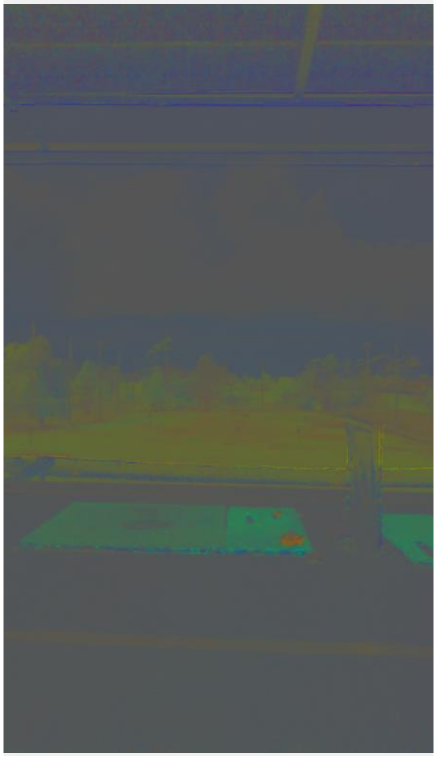


(a) Shadows

(b) Color difference



(c) Brightness difference

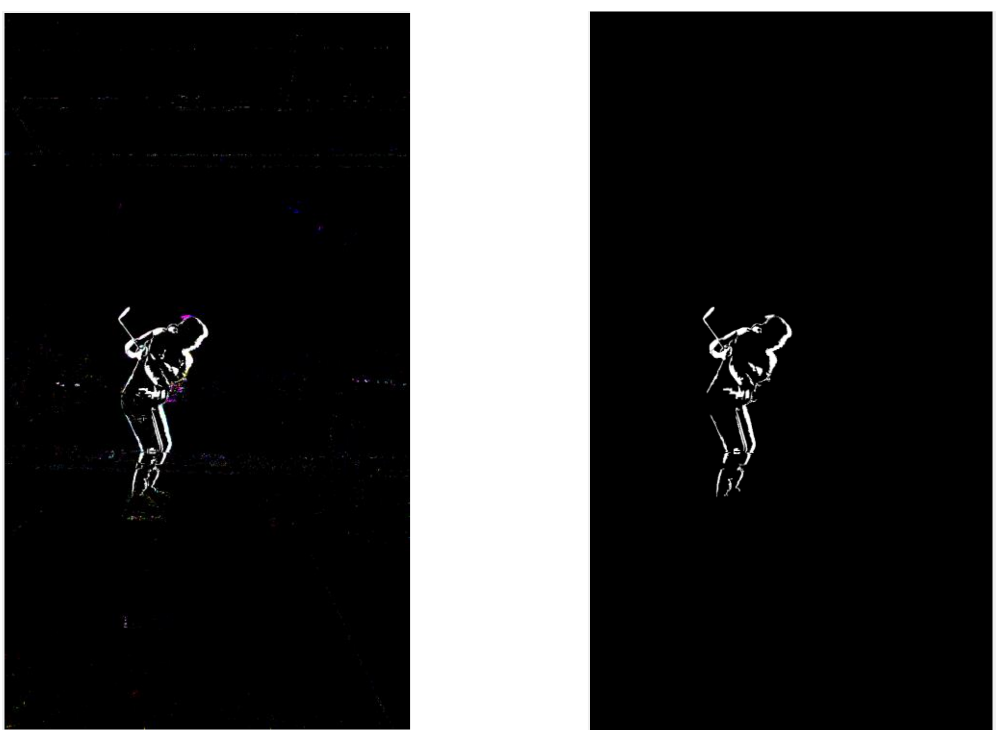


(d) Background RGB

(e) Layer remove threshold (isgood)

(f) Accumulate all difference(isok)

Figure 4.9: Interim values



(a) Foreground image

(b) Foreground image after removing the unwanted smaller regions

Figure 4.10: Extracted foreground from a multi-layered approach

As an important point, the non-motion pixels are determined by 3-frame motion differences. For example, the feet of the golfer are dynamic during the swing, while they are also static pixels during the swing preparation. Therefore, feet, as the part of background, would disappear from the foreground image.

Furthermore, the pixels representing the small amount of movement will be ignored. The purpose of this approach is to subtract clear foreground. In contrast, the object motion between frames represents substantial pixel differences contributing to obtaining a clear foreground as a result. However, shorter time intervals between all frames can also reduce background noise. For algorithm development, all frames are extracted from the video at regular time intervals. The time intervals depend on the frame rate of the captured video. For example, skipping 10 frame intervals from a captured video at 240 fps is equivalent to a 24 fps video.

4.4 Interactive annotations elements of performance and safety

Communicating critical angles in diagnostic feedback are important in helping the golfer to achieve the balanced address position needed to create consistency and generate power. The relative angles of the spine, elbows, knees and golf club are very important in golf coaching. Nevertheless, the number of degrees representing the ‘best’ angle, are typically communicated as a range of optimal values by the coach. The angles depend on the golfer’s body proportions. The annotation solution can provide angles for comparative purposes. The measured angle is only a way of augmenting coaching feedback regarding the player’s posture.

Table 12: Pseudo code for computing the angle of two annotated lines (MathWorks)

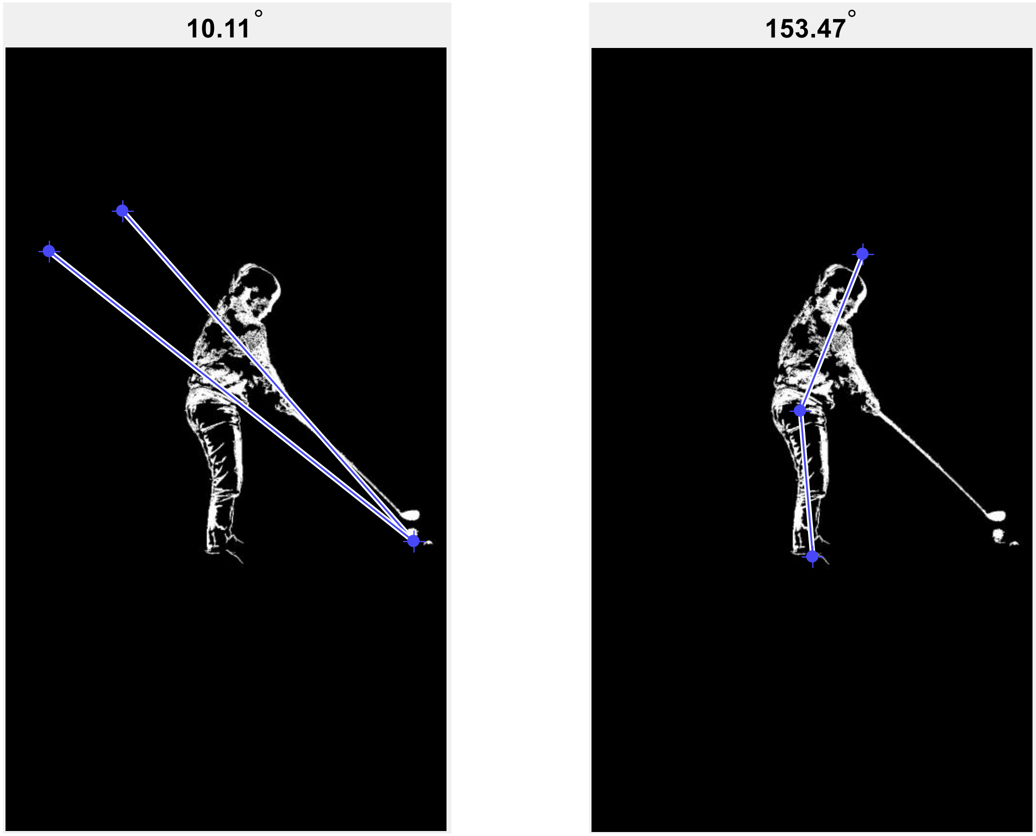
Algorithm ANGLE_OF_INTERSECTION

Input: bw foreground image, im
Output: image with calculated angle
Parameters: STRONGEST = 3

Library functions:	Purpose
detectHarrisFeatures(I)	Detect corners using Harris–Stephens algorithm and return cornerPoints object
selectStrongest(N)	Return points with strongest metrics, N is no of points
getCurPos	Returns the current position
dotV	Generate dot product for two vectors
norm	Return the 2-norm of a vector

```
Use Harris–Stephens algorithm to find the corner point, set points =
detectHarrisFeatures(im)
Find the top 3 strongest points, set strongest = points.selectStrongest(3)
Get three points location, set xy = strongest.Location
h = create a polygon by xy
find the position of h, set pos = h.getCurPos
v1 = [pos.x1 – pos.x2, pos.y1 – pos.y2]
v2 = [pos.x3 – pos.x2, pos.y3 – pos.y2]
calculate the inverse cosine in radians, set theta =
acos(dotV(v1,v2)/(norm(v1)*norm(v2)));
angle_d = (theta * (180/pi));
return angle_d
```

The annotation of angle method uses an extracted foreground image as input, detecting corners by using the Harris-Stephens algorithm. The Matlab cornerPoint object detects the three strongest points as shown in Table 12. Three-point positions on the picture are needed to draw two lines. This interactive 3-points annotation method measures the angle between two lines displayed as title. If the user moves the point, the angle will be automatically recalculated and updated in the title as Figure 4.11.



(a) Angle of swing plane (10.11°) (b) Angle of spine (153.47°)

Figure 4.11: Augmenting relative spine angle and upper body tilt using the developed annotation function over silhouette-based video frames to convey coaching feedback.

4.5 Interaction annotation of a 3D grid on a 2D image

Drawing grid can help the golfer and end user to visualise 3D volume, given that a silhouette visually represents a 2D plane. Existing video coaching tools such as open source software (*Kinovea*) already provides four interactively adjustable points for grid drawing function, which is time-consuming and not very practical for fast feedback reporting. The objective of the *one-click* interactive *pseudo-3D grid* drawing is to produce cognitive support for an end user to draw a grid relying on one-step selection of the vanishing point. The novel one-click interactive function is based on Hough transformation, that can automatically generate a subset of lines crossing the possible vanishing point. Table 13 and Table 14 show the pseudo code for the assisted grid drawing interactive function.

Table 13: Pseudo code for vanishing line

Algorithm VANISHING_LINE
Input: rgbImg, a RGB image
Output: grayImg, a gray image with hough lines

Library functions:	**Purpose**
Hough(BW)	the Hough Transform of binary image, returning: hou = Hough transform matrix, th = Angle in degrees between the x-axis and the rho vector, rh = perpendicular distance between the origin and the detecting line (MathWorks, 2017).
houghpeaks(name, value)	Identify peaks in Hough transform by parameter name and value, return row and column coordinates of peaks
houghlines	Extract line segments based on Hough transform, return a structure array of lines found. Arguments: FillGap, distance between two lines segments. MinLength, minimum line length

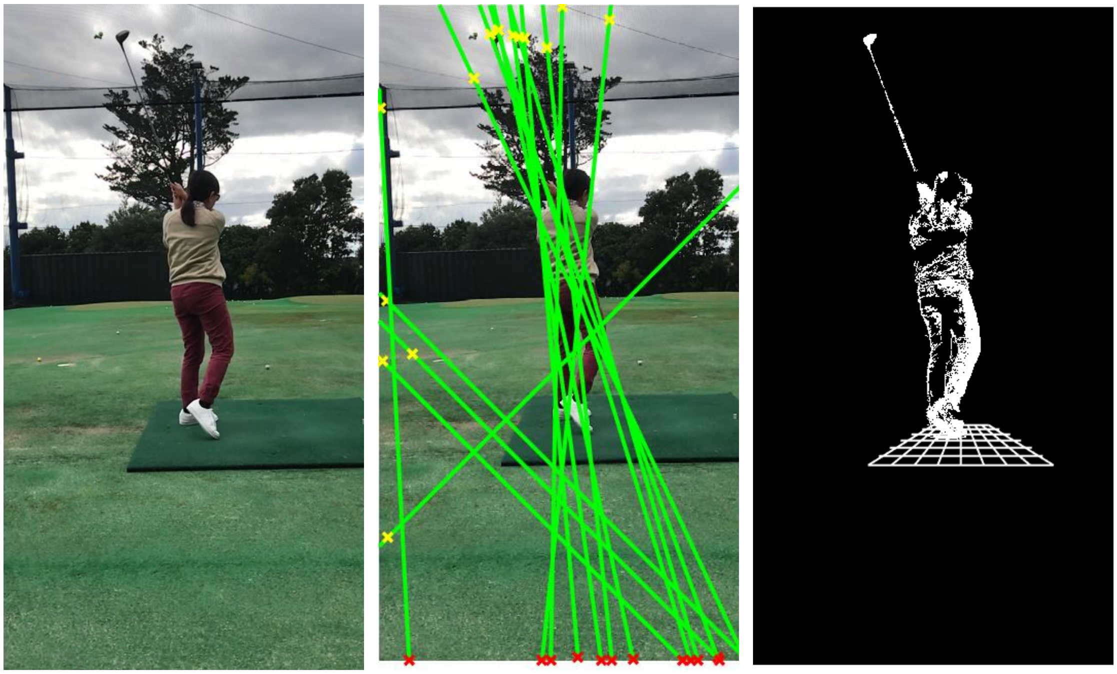
```
Img = readRGBimg
BW = rgb2bw(Img)
ts = threshold(Img)
ed = edge(BW)
H, th, rh = hough(ed)
hPeak = houghpeaks(hou, ts)
 [rows, columns] = size(BW)
lines = houghlines(BW, th, rh, hPeak, FillGap, MinLength)
For k = 1 to no of lines
        xy = [lines(k).point1; lines(k).point2]
        [x1 y1] = [xy(1,1)   xy(1,2)]
        [x2 y2] = [xy(2,1)   xy(2,2)]
        slope = (y2-y1)/(x2-x1)
        x_LeftPt = 1
        y_LeftPt = slope * (x_LeftPt - x1) + y1
        x_RightPt = columns
        y_RightPt = slope * (x_RightPt - x1) + y1
        plot the line ([x_LeftPt, x_RightPt], [y_LeftPt, y_RightPt], 'green')
        plot the point (xy(1,1),xy(1,2),'x','yellow')
        plot the point (xy(2,1),xy(2,2),'x','red')
end loop
return the image with plotted lines
```

Table 14: Pseudo code for perspective 3D grid represents the ground

Algorithm GRID_GROUND
Input: I, a RGB image; imfg, foreground image; horLine, no of row; verLine, no of column; interval, no of interval; ln_len, line length
Output: grayImg, a gray image with hough lines

```
Img = readRGBimg
Imy = size(Img,1)
Imx = size(Img,2)
[mx,my] = getInputPoint()
x1 = ceil(mx/2)
y1 = ceil((imy+my)/2)
x2 = ceil((imx+mx)/2)
y2 = ceil((imy+my)/2)
len_start = x2-x1
itv_start = len_start/verLine
xxyy = [x1 y1; x2 y2]
plot(xxyy(:,1),xxyy(:,2),'LineWidth',1,'Color','w')
st_x = x1;
st_y = y1;
en_x = x2;
en_y = y2;
for m = 1 : horLine
    st_x = st_x + ln_len
    st_y = st_y - interval
    en_x = en_x - ln_len
    en_y = en_y - interval
    xxyy = [st_x st_y; en_x en_y]
    plotLine(xxyy(:,1),xxyy(:,2),'LineWidth',1,'Color','w')
end
len_end = en_x - st_x
itv_end = len_end/verLine
en_x = st_x
en_y = st_y
st_x = x1
st_y = y1
xxyy = [st_x st_y; en_x en_y]
plot(xxyy(:,1),xxyy(:,2),'LineWidth',1,'Color','w')
for n = 1:verLine
    st_x = st_x + itv_start
    en_x = en_x + itv_end
    xxyy = [st_x st_y; en_x en_y]
    plot(xxyy(:,1),xxyy(:,2),'LineWidth',1,'Color','w')
end
```

Based on the original image (Figure 4.12 a), to calculate vanishing lines and the vanishing point presented in image (b), an end user can click on the chosen estimated vanishing point to indicate the location and relative pseudo-3D perspective orientation of the grid (c).



(a) Original image (b) Vanishing lines and vanishing point plot on the image (c) 3D volume representation

Figure 4.12: From original image to assisted grid drawing for 3D visualisations

4.6 Golf club head trajectory

Existing technology can recognise the golf club by using unique identifying stickers or wearable/ubiquitous hardware attached to the club (Marty & Edwards, 2013). In contrast, this section presents a method based on the monocular video tracing of the golf club head to compute the club head trajectory (Table 15). The experimental work suggests that a high-frame rate video (240 fps) is strongly recommended for the purpose of golf club head tracing. In the case of a high-frame rate video, the trajectory method relying on the foreground image video would be obtained by the multi-layered approach presented in Section 4.3.

Table 15: Pseudo code for golf club trajectory

```
Algorithm GOLF_CLUB_TRAJECTORY

Input: foreground image frame videofile
Parameters:     MINIMUM_BLOB_AREA = 500
                NUMGAUSSIANS = 5
                NUMTRAININGFRAMES = 5
                MINIMUMBACKGROUNDRATIO = 0.85
                BOUNDINGBOXOUTPUTPORT = 'true'
                AREAOUTPUTPORT = false
                CENTROIDOUTPUTPORT = false
Library functions: Note)               Purpose
  Gaussians_foregroundDetector()       Foreground detection using Gaussian mixture models
  BlobAnalysis()                       Computes statistics for connected regions in a binary image
  boxCombine(bbox)                     Combine the bounding boxes of overlapping (Maurer, 2016)
  roi_mask(cleaned, comBox)            Create mask for region of interest

len = videoReader.Duration * videoReader.FrameRate
path = zeros(2,len)
foregroundDetector = Gaussians_foregroundDetector(NUMGAUSSIANS,
NUMTRAININGFRAMES, MINIMUMBACKGROUNDRATIO)
videoReader = vision_videoReader(videofile)
blobAnalysis = BlobAnalysis(MINIMUM_BLOB_AREA)
while videoReader hasFrame do
    frame = readFrame(videoReader)
    foreground = foregroundDetector(frame)
    filled = image fill hole(foreground)
    cleaned = image open (filled)
    bbox = blobAnalysis(foreground)
    if (bbox is not empty) then
      combine the overlap box, set comBox = boxCombine(bbox)
      create mask on the combined box, set masked = roi_Mask(cleaned, comBox)
      erImg = erode the masked image
      cp = compute the centroid point of erImg
      if (cp is not empty)
            plot cp in red
            add cp to the path
      end if
    end if
    remove null value from the path
    find the highest point of the path, mid = highest(path)
    the last point of the path, set last = (size(path))/2
    plot the path from 1st point to mid in colour red
    plot the path mid to last point in colour blue
end while
```

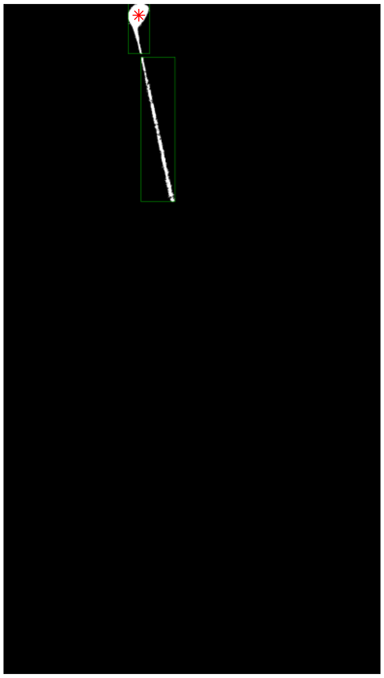
Table 16: Pseudo code for ROI mask

Algorithm ROI_MASK
Input: image im, combined box comBox
Output: Masked image maskImg

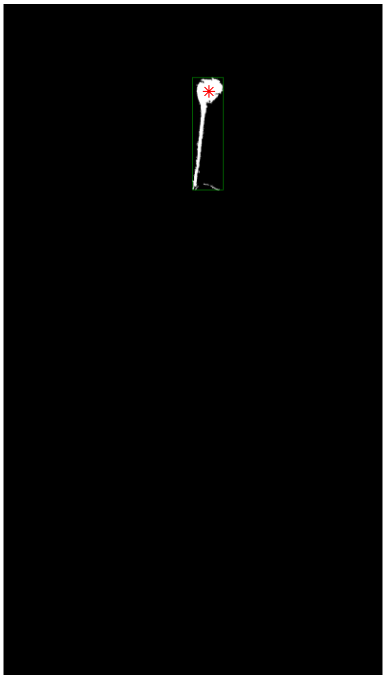
```
s = compute comBox size
imbox = 0
for I = 1:s
    box = create a rectangle of comBox
    boxMask = create a mask of box
    imboxM = im multiplication boxMask
    imbox = imbox + imboxM
end
maskImg = imbox
return maskImg
```

Note) The library functions were tested in Matlab (ver 2016b with *Computer Vision System*, *Image Acquisition* and *Image Processing* toolboxes).

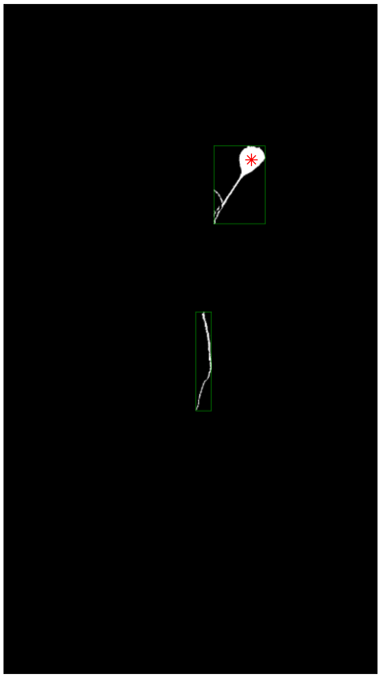
To autonomously trace the club head movement from monocular camera view, the algorithm employed combined Gaussian mixture models (GMM) with overlapping bounding boxes connected. The centroid of the golf club head from the starting point of the backswing to the impact is shown in Figure 4.13 (a) to (e). In order to trace the golf club head, a Matlab system object ‘vision.BlobAnalysis’ computes the properties of connected regions in a binary image. The property ‘MinimumBlobArea’ is to specify the size of the blob minimum size in pixels, which needs adjustments and depends on the image and video sizes.



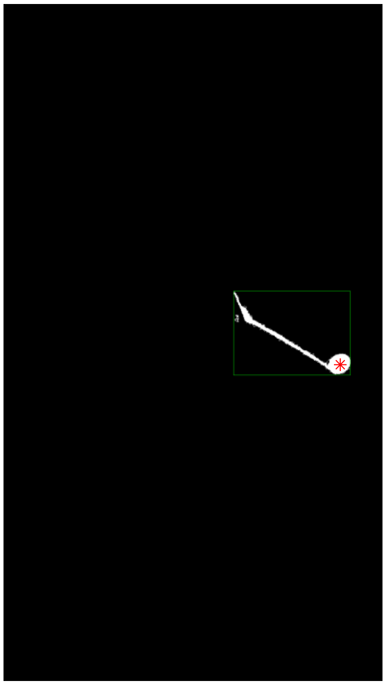
(a)



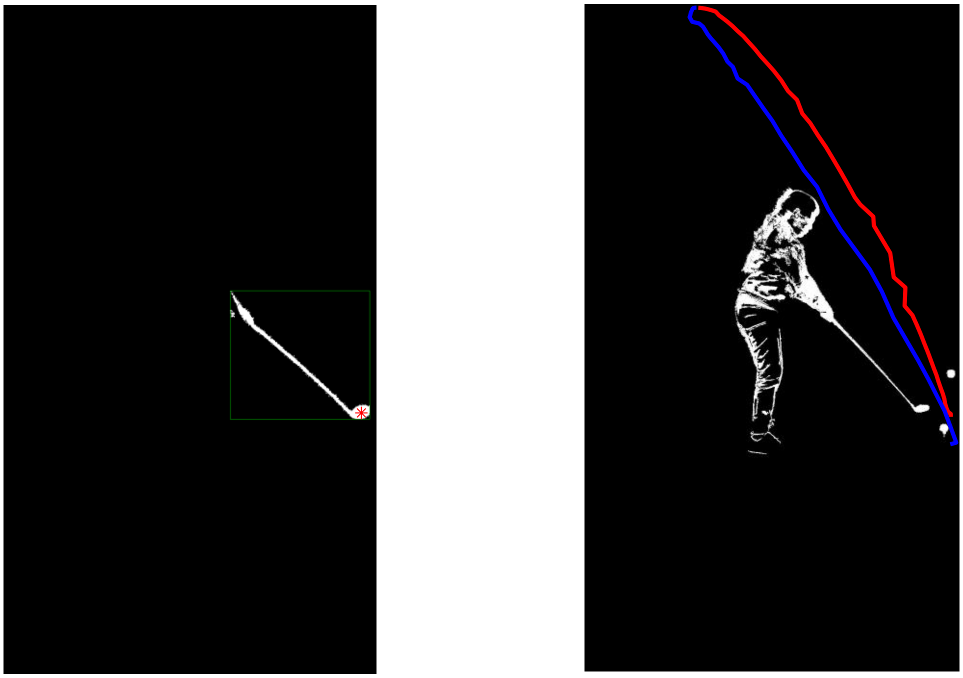
(b)



(c)



(d)



(e) (f)

Figure 4.13: Golf club head trajectory, (a-e) centroid and the bounding box of the golf club head and (f) plot backswing trajectory of a golf club in red and downswing trajectory in blue.

In some cases, the BlobAnalysis object providing the area, centroid and bounding box (Bbox) is not found or in other case there is more than one Bbox detected in a frame. Therefore, a boxCombine (Maurer, 2016) method is used to combine the overlapped bounding boxes shown in Table 16, to find the centroid of the golf club head in the largest blob. The erosion function is to combine the box until is the shaft is eroded. All traced centroids of the golf club head location are stored into an array. In the final step, the method plots the first point to the highest point in red, and the highest point to the end in blue to represent backswing and downswing (Figure 4.13 f). Nevertheless, the algorithm could not always trace the golf club head in each successive frame. For example, when the golf club head image overlapped the golfer image, the golf club head could not be traced.

4.7 Visual feedback annotation report

Part of coaching feedback requires prioritisation of what elements of performance should be corrected first. For instance, a golf beginner’s common error is looking at the golf club instead of keeping the ‘eye on the ball’. Using visual feedback can help with the performance or correction of the player (Knudson & Morrison, 2002). This visual coaching analysis report (Figure 4.14) provides a picture with the coach’s comments and instruction which shows the intended post-session report for the player. Coaches can use the developed interactive functions to perform analysis and use annotations for diagnostic feedback to support their coaching practice.

An HTML or PDF file analysis report (Figure 4.14) can be printed as a hard copy or sent by email to the player at the end of coaching session. The report is visual evidence that summarises analysis and diagnostic feedback. Moreover, the anonymised report is designed to preserve the privacy of the golfer and also to facilitate communication between coach and player as well.

Player: Anonymous

Date 1 Nov 2017
Player ID 12345678

Lesson 1: Stance

Coaching feedback

☑ **Feet**
- **Shoulders apart**
- **Weight on the balls of your feet**

☐ **Knees**
Knees should be more bent
Bend the knees around 20–25° degrees (ie. 180-25=155° of flexion) will have better result in torso rotation.

☐ **Upper body**
Slightly bent your upper body more forward with bent knees

☐ **Hands and arms**
Loose and hanging freely
Focus on more loose and elastic feel.

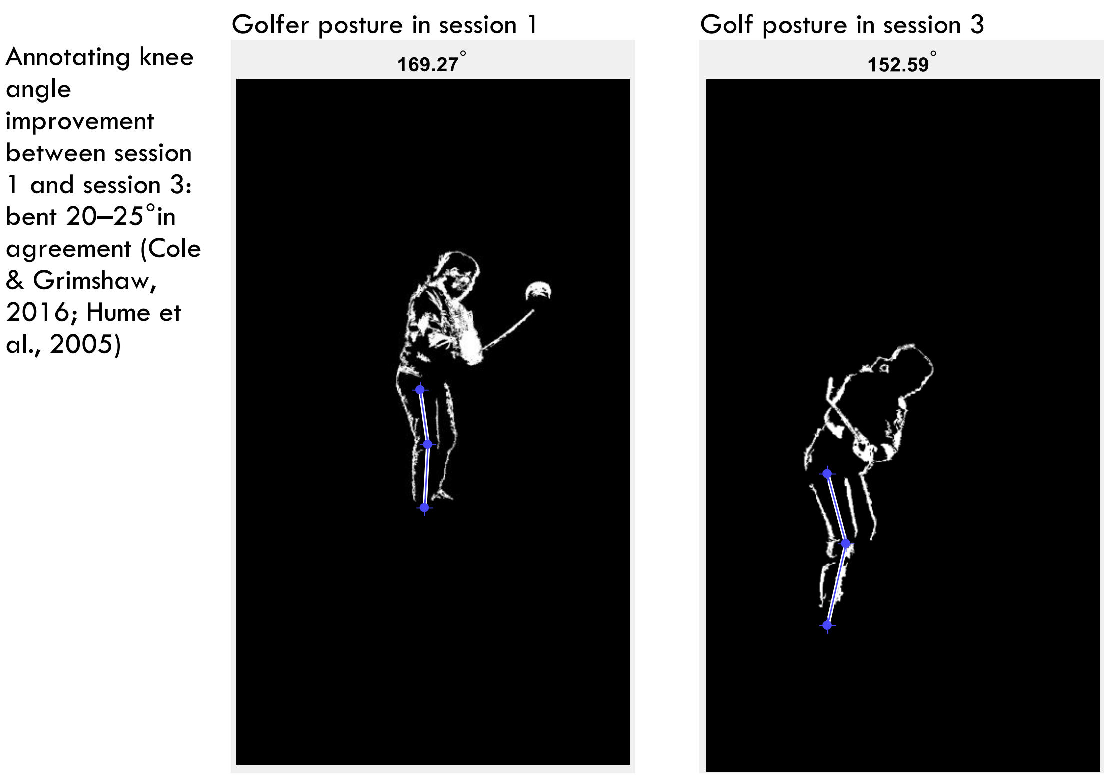
☐ **Grip**
To be checked again and adjusted if needed.

☑ … OK
☐ … Good, but needs some adjustments
☒ … Need for improvements

Figure 4.14: Swing analysis: Example of a diagnostic report showing the qualitative nature of golf coaching

4.8 Qualitative evidence of posture and swing improvement

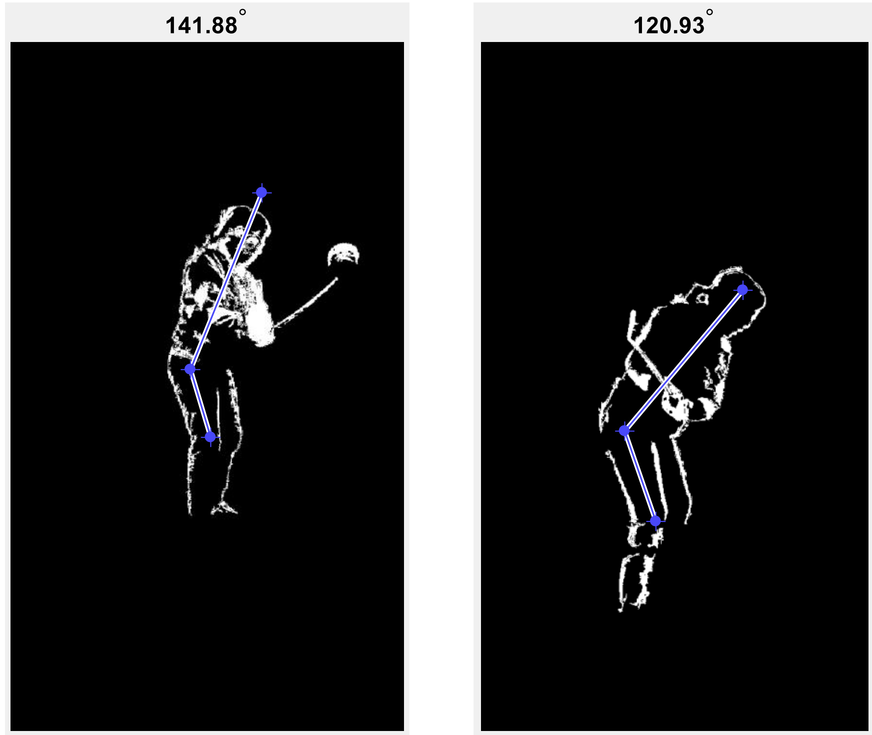
The presented silhouette-based video shows a swing with common errors and improvements that are typical at beginner level. The author has not taken any formal golf training with a coach. In this case, golf learning was supported by data collection practice sessions on golf driving ranges, video reviews and discussion with the supervisor. Figure 4.15 provides qualitative evidence showing the posture of Session 1 and Session 3.



(a) Knee angle (169.27°) of downswing in session 1

(b) Knee angle (152.59°) of downswing in session 3

Annotating trunk angle improvement between session 1 and 3



(c) Trunk angle (141.88°) with drive club in session 1

(d) Slightly exercises trunk angle (120.93°) with shorter club (iron) in session 3.

Figure 4.15: Golfer posture comparison between session 1 and session 3

4.9 Practical implications for golf coaches

The developed solution to facilitate golf coaching relied on video evidence representing captured human movement associated with golf swing activity. Unlike with the naked eye, captured video can provide more details of recorded movement, therefore augmenting traditional coaching. As external validity of silhouette-based coaching (K. Y. Chan & Bačić, 2017), Table 17 summarises practical implications for golf coaches. In support of design science methodology used in the experimental development, Table 17 also summarises to what degree the developed solution can aid to augmented coaching analysis of the elements of performance that are typical for beginners to intermediate skill-level players.

Table 17: Common analysis elements of golf swing performance from beginner to intermediate players, adapted from (K. Y. Chan & Bačić, 2017)

Elements of golf swing	Can the developed solution assist to analyse the player posture?
Grip [1)]	N
“Eyes on the ball”	N
Head position estimation before the impact relative to the body and the ball	Y
Stance assessment	Y
Knees	Y
Upper body	Y
Balance	Y
Hands and arms	Y
Swing plane	Y
Club swing tracking through the swing phases	Y
Computing swing tempo [2)]	N
Annotation angles in 2D plane views	Y
Swing plane visual information	Y
Estimation of vanishing point	Y
Interactive ground annotation via vanishing point and computer assisted grid overlay drawing	Y
Privacy preservation 2D silhouette with background elimination	Y
2D silhouette enhancement (Pseudo-3D binary silhouette)	Y
Feet and shoes [3)]	N

Note:
[1)]A coach should use visual RGB zoomed in to the region of interest of grip.
[2)]Some players tend to overswing (e.g. John Daly), so both front and sagittal videos are needed.
[3)]Learners should wear the same shoes (adequate for golf).

4.10 Results, discussion, and implications

Golf coaching is an educational process of qualitative nature where the elements of performance are not assessed as absolute values, but rather as subjective measures expressed as a degree of adherence to common-sense coaching rules. Augmented video coaching based on the annotated replay of recorded video evidence may help to communicate subjective feedback, reduce potential bias, and improve intra-rater and inter-rater consistency and reliability of diagnostic assessment.

The set of developed solutions can assist coach to qualitatively or quantitatively assess the player's hard-to-quantify elements such as:

- Automatic angle measure, which can be used to measure or compare player's posture.
- The one-click interactive pseudo-3D grid on the 2D image that can help the user to visualise 3D volume.
- Automated monocular markerless tracking of the golf club head trajectory, which indicates clubhead movement relative to the player's swing plane. Too low or too steep clubhead movement is related to the individual player's tendency to produce deviations from desired straight ball-flight pattern.

In comparing the disciplines of computation intelligence and machine learning with video and image processing it is possible to draw parallels and similarities of ideas that helped mutual interdisciplinary advancements. For example, it is possible to argue that KNN clustering with Radial Basis Function ANN share similarities with KNN clustering and Gaussian Mixture Models computing concepts. Similar interdisciplinary line of thinking in this thesis is also inspired by multi-layered ANN architectures relying on processing units (neurons) with different interconnection weights. Similar to multi-layer ANN architecture, the developed pseudo-3D silhouette solution is also based on multi-layered architecture where instead of the interconnected neurons there is a concept of layers (as a matrix of image pixels). Such matrices of pixels share a set of functions responsible for in-between-layers' temporal and spatial matrix transformations. For example, temporal transformations of cumulative frame difference share similarities with the concept of *fading memory* that can be found in *reservoir computing* (RC) paradigm (e.g. in Echo State Network and Liquid State Machines).

The underlying purpose of the implemented pseudo-3D silhouette-based solution is to separate foreground and background from captured video for the purpose of golf coaching that to some degree would be transferrable to other sport, general healthcare and rehabilitation-specific contexts. Separation of foreground and background was found to be a challenging task for the environments similar to the golf driving ranges that were chosen for experimental data collection. Regarding the limitations of the achieved performance, the silhouette-based algorithm could not produce ideal image conversion for each frame, which is also known from similar contexts covered in the literature review (Chapter 2). The extracted silhouette boundary can also be slightly distorted caused by hairstyle and loose clothing. For instance, it is not easy to determine the knee angle during the golf swing with loose trousers. Also, in some frames during the swing, a clear separation of shoes with the ground was impeded by static and shadow. For the scope of this thesis and for coaching applications, it is considered that a coach should select video frames that communicate the best feedback on recurring errors of captured golf swings.

The developed solution used mobile video, which is multi-purpose, immediately accessible and convenient technology used by most people. Compared to monocular camera utilisation, Table 18 summaries the benefits of achieved results compared to the initial Kinect prototyping evaluation for silhouette-based coaching (Bačić et al., 2017).

Table 18: Comparison between the Kinect stereo depth video (Bačić et al., 2017) and generic monocular camera view reported in the experimental design and achieved solution

Kinect solution	**Developed solution in this thesis**	**Description**
Unique hardware	Any smartphone video or digital camera	The Kinect solution specifies Kinect hardware only. In contrast, the developed solution can work on video from any digital camera and smartphone.
Fixed indoor	Portable	Smartphones and digital cameras are flexible for both indoor and outdoor, with available waterproof in some models. Kinect is comparatively bulky, designed for fixed indoor, and requires a power supply. Moreover, from the previous evaluation, Kinect I may malfunction in direct sunlight (Bačić et al., 2017)
30 fps.	30 to 240 fps	Kinect frame rate is maximum at 30fps and as such it cannot capture the fast movement of the golf swing. Conversely, the higher fps smartphone video can capture the fast movement eg. the moment of impact with the ball.
2D flat silhouette image	Between 2D and 3D	Kinect’s 2D flat silhouette solution was originally designed to hide more detail.
Invisible golf club	Traced golf club	In Kinect’s solution, most of the time the golf club movement is invisible. In this developed solution, the golf club and golf club head are traced most of the time.
Live video streaming	Video streaming is considered applicable for future work and data communication improvements e.g. 5G mobile technology.	Kinect can produce real time video streaming. Mobile phones both iOS and Android, do not provide relay on third-party tools for streaming.

Although this study focuses on the case of golf sports activities, broader implications and transferable applications are possible to diverse sports and other contexts such as smart home and rest homes and the healthcare industry in general. The concept of pseudo-3D silhouette-based filtering can be considered universally applicable to privacy sensitive contexts utilising various video sources in the near future.

Chapter 5 ***Conclusions, implications and future work***

Concluding section includes the key points achieved, general contributions, methodology as well as the implications for mixed audiences with potential transfer to diverse contexts.

The main objective of this thesis was to advance technology-mediated personalised augmented coaching by (1) preserving privacy while augmenting diagnostic elements of golf swing performance; and (2) automating elements of video annotations and post-session diagnostic feedback reporting. The enhanced black and white pseudo-3D binary silhouettes able to hide observed player's identity and keeping visual elements that are needed to support the coaching process and diagnostic capabilities. In addition, the interactive angle, 3D grid on a 2D image provide automating annotations and the diagnostic report showing the qualitative nature of golf coaching.

The privacy preservation was achieved via multi-layered spatiotemporal video and image transformations resulting in enhanced black and white 3D silhouette. Such privacy-preserving computer-generated silhouette, was used as basis to convey overlay feedback annotations of golf coaching activity recorded in two different golf driving range settings while eliminating redundancy such as surrounding background and colour variations. The enhanced 3D silhouette annotations from sagittal plane (side view, behind the golfer) did not suffer from limb occlusion such as knee separation, hand and elbow separation from the body that would lead to analytical/coaching bias. For the produced replays and post-session reporting annotations, the developed algorithms were found to be robust to change in colour, contrast and background given the different experimental settings and diverse light-changing conditions during the video recording sessions.

The list of annotated elements was considered sufficient for producing feedback on common mistakes – that are typical for golf novices. For the end-of-session report, a coach can draw lines, joint angles and other angles combining the golfer’s body with sports equipment and produce a visualisation of the golfer’s swing plane with generated angle calculations. The novel proof of concept for the one-click interactive pseudo-3D grid drawing representing the ground level was achieved by computer-generated suggestions for possible vanishing points, so that a coach could draw a grid in one step rather than manually editing all four points representing the pseudo-3D grid edges. Another novel proof of concept for backswing and downswing annotations in separate colours was achieved using computer-generated markerless monocular tracking of the clubhead movement during the swing.

The employed multidisciplinary research included combined techniques from computer vision and video and image analysis applied to sports science and golf-coaching contexts. Regarding reported data collection, experimental setup and recording protocols, all golf coaching video data included the author's self-recorded golf swings, using a monocular camera 2D view from various video sources including mobile phones and digital cameras. Pertinent to a qualitative analysis of a golf swing, there is a requirement of high frame rate videos of 120 fps, 240 fps and higher that can be found in mobile and sports/pocket cameras. All annotated examples provided in the thesis were aligned with the literature review, common coaching practice and common errors that are typical for golf amateurs i.e. novice-to-intermediate skill levels.

The initial experimental study of privacy preservation in golf coaching included the use of video sources combined with commonly used methods in video and image processing for foreground and background separation. The commonly used algorithms in computer vision in surveillance research and their combinations were tested on the experimental data collected on two different driving ranges. In addition to diverse colours, light source orientation and brightness, the main problem was with each video pixel that can represent both static and dynamic foreground or background information, therefore representing a challenge of silhouette production relying on commonly used methods. Another problem with privacy preservation and initially obtained solid-coloured silhouette filtering was a binary silhouette-based solution that would prevent coaches analysing some of the critical elements of golf swing performance. For example, it would be difficult to analyse elbow angles during the backswing from the rear camera view.

The produced pseudo-3D binary silhouette spatiotemporal transformation represents visual evidence that such a solution can be used for coaching the golf swing while preserving privacy. Furthermore, it is also possible to apply produced video overlays/annotations to visually communicate critical elements of the swing performance while minimising potential coach bias.

Consequently, the solution demonstrated that the silhouette video and image can be used for augmented video coaching and privacy preservation for:

- Amateur-level golfers (i.e. novices to intermediate skill levels)

- Qualitative analysis of swing and safety performance e.g. stance, swing phasing analysis and swing plane visualisation.
- Quantitative analysis e.g. club, joint and swing plane angles calculation.

The developed artefacts can be used to semi-automate production of coaching elements found in post-session diagnostic reporting. Given utilised monocular camera view resources with limited pixel information and lacking 3D depth recording capability, a limitation of this study is that it was not always possible to accurately separate shoes from the surrounding pixels. For online coaching, it is therefore recommended that shoes and hands holding the club are recorded separately as zoomed-in regions of interest while preserving the original pixel information. Another limitation regarding the novel interactive solution for one-click drawing is that the perspective 3D grid represents an estimate rather than accurate 3D depth visualisation. Future work will combine images from two or more cameras to construct 3D image. Also include advancing the algorithms for improving: (i) perspective 3D grid accuracy, (ii) tracing of the club movements, (iii) obtain 3D angles and (iv) ball tacking with its trajectory prediction.

Future work is aimed at further advancements of silhouette-based video streaming solutions to advance the diversity of sports disciplines, general sports science and other multi-disciplinary research contexts. Including embedded video pre-processing for smart homes, rest homes and hospitals to preserve privacy. In addition to providing sufficient activity monitoring information for computer-facilitated fall-alerting and other environmental monitoring contexts.

References

AGCAS. (Dec 2017). *Job profile sport coach*. Retrieved 5 Jun, 2017, from https://www.prospects.ac.uk/job-profiles/sports-coach

Ananya, C., Anjan Kumar, T., & Kandarpa Kumar, S. (2015). A review on vision-based hand gesture recognition and applications. In *Intelligent applications for heterogeneous system modeling and design* (pp. 256-281). Hershey, PA, USA: IGI Global. Retrieved from http://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/978-1-4666-8493-5.ch011. doi:10.4018/978-1-4666-8493-5.ch011

Archetti, F., Manfredotti, C. E., Messina, V., & Sorrenti, D. G. (2006). Foreground-to-ghost discrimination in single-difference pre-processing. In J. Blanc-Talon, W. Philips, D. Popescu, & P. Scheunders (Eds.), *Advanced Concepts for Intelligent Vision Systems: 8th International Conference, ACIVS 2006, Antwerp, Belgium, September 18-21, 2006. Proceedings* (pp. 263-274). Berlin, Heidelberg. Retrieved from https://doi.org/10.1007/11864349_24. doi:10.1007/11864349_24

Ashford, W. (2012). *Free mobile apps pose a serious threat to privacy because of their ability to capture large amounts of user information*. Retrieved 9 May, 2017, from http://www.computerweekly.com/news/2240169770/Free-mobile-apps-a-threat-to-privacy-study-finds

Bačić, B. (2016a). Echo state network for 3D motion pattern indexing: A case study on tennis forehands. In T. Bräunl, B. McCane, M. Rivera, & X. Yu (Eds.), *Image and Video Technology: 7th Pacific-Rim Symposium, PSIVT 2015, Auckland, New Zealand, November 25-27, 2015, Revised Selected Papers* (pp. 295-306). Cham: Springer International Publishing. Retrieved from http://dx.doi.org/10.1007/978-3-319-29451-3_24. doi:10.1007/978-3-319-29451-3_24

Bačić, B. (2016b). Predicting golf ball trajectories from swing plane: An artificial neural networks approach. *Expert Systems With Applications, 65*, 423-438. doi:http://dx.doi.org/10.1016/j.eswa.2016.07.014

Bačić, B., & Hume, P. (2012). Augmented video coaching, qualitative analysis and postproduction using open source software Symposium conducted at the meeting of the International Symposium on Biomechanics in Sports Retrieved from https://ojs.ub.uni-konstanz.de/cpa/article/view/5305/4877

Bačić, B., Meng, Q., & Chan, K. Y. (2017, 14-16 Jun.). Privacy preservation for eSports: A case study towards augmented video golf coaching system. *IEEE*. Symposium conducted at the meeting of the 10th International Conference on Developments in eSystems Engineering (DeSE), Paris, France. doi:10.1109/DeSE.2017.34

Bao, M., Guo, S., Tang, Q., & Zhang, F. (2009). *Optimization of the bwmorph function in the MATLAB image processing toolbox for binary skeleton computation*. presented at the meeting of the Proceedings of the 2009 International Conference on Computational Intelligence and Natural Computing - Volume 02, Retrieved from http://ieeexplore.ieee.org/document/5230976/ doi:10.1109/cinc.2009.96

Bartlett, R. (2007). *Introduction to sports biomechanics: Analysing human movement patterns* (2nd ed.). London: Routledge.

Baskerville, R., Pries-Heje, J., & Venable, J. (2009). Soft design science methodology*ACM*. Symposium conducted at the meeting of the Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology, Philadelphia, Pennsylvania. Retrieved from https://dl.acm.org/citation.cfm?id=1555631 doi:10.1145/1555619.1555631

Bentley, N. (2015). *The importance of sports coaches*. Retrieved 5 Jun, 2017, from https://www.training.com.au/ed/the-importance-of-sports-coaches/

Best, J. R. (2013). Exergaming in youth: Effects on physical and cognitive health. *Zeitschrift für Psychologie, 221*(2), 72-78. doi:10.1027/2151-2604/a000137

Bouwmans, T., El Baf, F., & Vachon, B. (2008). Background modeling using mixture of Gaussians for foreground fetection - A survey. *Recent Patents on Computer Science, 1*(3), 219-237.

Cann, A. P., Vandervoort, A. A., & Lindsay, D. M. (2005). Optimizing the benefits versus risks of golf participation by older people. *Journal Of Geriatric Physical Therapy, 28*(3), 85-92.

Carlson, M. P., & Bloom, I. (2005). The cyclic nature of problem solving: An emergent multidimensional problem-solving framework. *Educational Studies in Mathematics, 58*(1), 45-75. doi:10.1007/s10649-005-0808-x

Chan, K. Y., & Bačić, B. (2017, 15 Nov). *Producing silhouette-based augmented feedback: What elements of performance can a golf coach analyse?* presented at the meeting of the SPRINZ strength and conditioning conference – AUT SPRINZ, Auckland, New Zealand.

Chan, Z. *Swing Profile golf swing analyzer & training aid*. Retrieved 8 May 2017, from https://itunes.apple.com/nz/app/swing-profile-golf-swing-analyzer-training-aid/id1039981052?mt=8

Choi, A., Joo, S.-B., Oh, E., & Mun, J. H. (2014). Kinematic evaluation of movement smoothness in golf: Relationship between the normalized jerk cost of body joints and the clubhead. *BioMedical Engineering OnLine, 13*(1), 20. doi:10.1186/1475-925x-13-20

Cole, M. H., & Grimshaw, P. N. (2016). The biomechanics of the modern golf swing: Implications for lower back injuries. *Sports Medicine, 46*(3), 339-351. doi:10.1007/s40279-015-0429-1

Collins, R. T., Gross, R., & Jianbo, S. (2002, 21-21 May 2002). Silhouette-based human identification from body shape and gait. Symposium conducted at the meeting of the Proceedings of fifth IEEE International Conference on Automatic Face Gesture Recognition doi:10.1109/AFGR.2002.1004181

Common golf injuries and how to avoid them. (2017). Retrieved 19 Sep, 2017, from https://www.berkshireindependenthospital.co.uk/news/common-golf-injuries-and-how-to-avoid-them

Courtney, F., & Thomas, D. (2004). *Personal development, health and physical education*: Pascal Press.

Culibrk, D., Marques, O., Socek, D., Kalva, H., & Furht, B. (2007). Neural network approach to background modeling for video object segmentation. *IEEE Transactions on Neural Networks, 18*(6), 1614-1627. doi:10.1109/TNN.2007.896861

D Stalin Alex, A. W. (2014). BSFD: Background subtraction frame difference algorithm for moving object detection and extraction. *Journal of Theoretical & Applied Information Technology, 60*(3), 623-628.

Dalal, N., & Triggs, B. (2005a). Histograms of oriented gradients for human detection [Dalal2005]. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2005)*. Piscataway: IEEE.

Dalal, N., & Triggs, B. (2005b, 25-25 Jun.). Histograms of oriented gradients for human detection. Symposium conducted at the meeting of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) doi:10.1109/CVPR.2005.177

Dartfish. Retrieved 10 Sep, 2017, from http://www.dartfish.com/

David, M., & Bondu, S. (1975). *Golf tee*. United States.

Dayan, E., & Cohen, L. G. (2011). Neuroplasticity subserving motor skill learning. *Neuron, 72*(3), 443-454. doi:10.1016/j.neuron.2011.10.008

Feldmann, T. (2011). Spatio-temporal optimization for foreground/background segmentation. In R. Koch & F. Huang (Eds.), *Computer Vision – ACCV 2010 Workshops: ACCV 2010 International Workshops, Queenstown, New Zealand, November 8-9, 2010, Revised Selected Papers, Part I* (pp. 113-122). Berlin, Heidelberg: Springer Berlin Heidelberg. Retrieved from http://dx.doi.org/10.1007/978-3-642-22822-3_12. doi:10.1007/978-3-642-22822-3_12

Fisher, R., Perkins, S., Walker, A., & Wolfart, E. (2003). *Sobel edge detector*. Retrieved 30 Oct, 2017, from http://homepages.inf.ed.ac.uk/rbf/HIPR2/sobel.htm

Geitgey, A. (2014). *Machine learning is fun*. Retrieved 2 Oct, 2017, from https://medium.com/@ageitgey/machine-learning-is-fun-80ea3ec3c471

GIMP. Retrieved 24 Feb, 2018, from https://www.gimp.org/

Gong, S., & Xiang, T. (2011). *Visual analysis of behaviour : From pixels to semantics*: New York : Springer-Verlag New York Inc. Retrieved from http://www.springer.com/us/book/9780857296696. Retrieved from cat05020a database.

Gonzalez, R. C., & Woods, R. E. (2008). *Digital image processing* (3rd ed.): Prentice Hall: New Jersey.

Haralick, R. M., & Shapiro, L. G. (1992). *Computer and Robot Vision*: Addison-Wesley.

Hawk-eye. Retrieved 10 Sep, 2017, from https://www.hawkeyeinnovations.com/about

Heitner, D. (2016). The state of the golf industry in 2016. *Forbes*. Retrieved from https://www.forbes.com/sites/darrenheitner/2016/05/08/the-state-of-the-golf-industry-in-2016/#1968175633a6

History of golf. (2011, 29 January 2018). Retrieved 12 Sep, 2017, from https://en.wikipedia.org/wiki/History_of_golf

Huang, S.-S., Fu, L.-C., & Hsiao, P.-Y. (2007). A bayesian network for foreground segmentation in region level. In Y. Yagi, S. B. Kang, I. S. Kweon, & H. Zha (Eds.), *Computer Vision – ACCV 2007: 8th Asian Conference on Computer Vision, Tokyo, Japan, November 18-22, 2007, Proceedings, Part II* (pp. 124-133). Berlin, Heidelberg. Retrieved from https://doi.org/10.1007/978-3-540-76390-1_13. doi:10.1007/978-3-540-76390-1_13

Huang, T., Guo, C., Qiu, J., & Ikenaga, T. (2009, 12-14 Sept.). Temporal information cooperated gaussian mixture models for real-time surveillance with ghost detection. *IEEE*. Symposium conducted at the meeting of the 2009 Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing doi:10.1109/IIH-MSP.2009.49

Hughes, J. (2010, 13 June 2010). *How to choose a golf instructor*. Retrieved 31 May, 2017, from http://www.pga.com/golf-instruction/instruction-feature/fundamentals/how-choose-golf-instructor

Hume, P. A., Keogh, J., & Reid, D. (2005). The role of biomechanics in maximising distance and accuracy of golf shots. *Sports Medicine, 35*(5), 429-449.

Izadi, S., Kim, D., Hilliges, O., Molyneaux, D., Newcombe, R., Kohli, P., . . . Fitzgibbon, A. (2011). KinectFusion: Real-time 3D reconstruction and interaction using a moving depth camera. doi:10.1145/2047196.2047270

Jain, A., Reddy, H., & Dubey, S. (2014). Automated driving vehicle using image processing. *International Journal of Computer Science and Engine and Engineering, 2*(4).

Jain, H. P., Subramanian, A., Das, S., & Mittal, A. (2011). Real-time upper-body human pose estimation using a depth camera. In A. Gagalowicz & W. Philips (Chair), *Springer Berlin Heidelberg.* Symposium conducted at the meeting of the Computer Vision/Computer Graphics Collaboration Techniques: 5th International Conference, Berlin, Heidelberg. doi:10.1007/978-3-642-24136-9_20

Jensen, D. R. *Why you need a coach*. Retrieved 5 Jun, 2017, from http://golfchannelacademy.com/why-you-need-a-coach/

Johnson, A. (n.d.). *What is coaching*. Retrieved 6 Nov, 2017, from http://hrweb.mit.edu/learning-development/learning-topics/leading/articles/what-is-coaching

Juang, C. F., Chang, C. M., Wu, J. R., & Lee, D. (2009). Computer vision-based human body segmentation and posture estimation. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, 39*(1), 119-133. doi:10.1109/TSMCA.2009.2008397

KaewTraKulPong, P., & Bowden, R. (2002). An improved adaptive background mixture model for real-time tracking with shadow detection. In P. Remagnino, G. A. Jones, N. Paragios, & C. S. Regazzoni (Eds.), *Video-Based Surveillance Systems: Computer Vision and Distributed Processing* (pp. 135-144). Boston, MA: Springer US. Retrieved from http://dx.doi.org/10.1007/978-1-4615-0913-4_11. doi:10.1007/978-1-4615-0913-4_11

Kawasaki, J. (1996). *Golf clubs.* United States.

Kelly, P., Ó Conaire, C., Monaghan, D., Kuklyte, J., Connaghan, D., Pérez-Moneo Agapito, J. D., & Daras, P. (2010). A low-cost performance analysis and coaching system for tennis Symposium conducted at the meeting of the Multimedia Grand Challenge Track at ACM Multimedia 2010, Firenze, Italy. Retrieved from http://doras.dcu.ie/15736/1/acmmm2010_3DLifeGC.pdf

Kelly, T. J. (n.d.). The Correct Way to Grip a Golf Club. *Golfweek*. Retrieved from http://golftips.golfweek.com/correct-way-grip-golf-club-1528.html

Kim, D. H., Millett, P. J., Warner, J. J. P., & Jobe, F. W. (2004). Shoulder injuries in golf. *American Journal of Sports Medicine, 32*(5), 1324-1330.

Kinect motion capture. Retrieved 18 Sep, 2017, from https://kinetisense.com/kinect-motion-capture/

Kinovea. Retrieved 16 Dec, 2016, from www.kinovea.org

Knudson, D. V. (2013). *Qualitative diagnosis of human movement: Improving performance in sport and exercise* (3rd ed.). Champain, IL: Human Kinetics.

Knudson, D. V., & Morrison, C. S. (2002). *Qualitative analysis of human movement* (2nd ed.): Champaign, IL : Human Kinetics.

Kooyman, D. J., James, D. A., & Rowlands, D. D. (2013). A feedback system for the motor learning of skills in golf. *Procedia Engineering, 60*, 226-231. doi:http://dx.doi.org/10.1016/j.proeng.2013.07.014

Kryjak, T. (2014). Foreground object segmentation in dynamic background scenarios. *Image Processing & Communications, 19*(2/3), 25-36. doi:10.1515/ipc-2015-0007

Kryjak, T., Komorkiewicz, M., & Gorgon, M. (2014). Real-time foreground object detection combining the pbas background modelling algorithm and feedback from scene analysis module *International Journal of Electronics and Telecommunications* (Vol. 60, pp. 53).

Lam, L., Lee, S.-W., & Suen, C. Y. (1992). Thinning methodologies-a comprehensive survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*(9), 869. doi:10.1109/34.161346

Lamanna, M. (2017). Learn the Different Types of Golf Grips. *ThoughtCo*. Retrieved from https://www.thoughtco.com/types-of-golf-grips-1564427

Leadbetter interactive. (2005). New Hudson: Interactive Frontiers. Retrieved from http://www.leadbetterinteractive.com/

Lightman, K. (2016). Silicon gets sporty. *IEEE Spectrum, 55*(3), 48-53. doi:10.1109/MSPEC.2016.7420400

LongoMatch. Retrieved 5 Nov, 2017, from www.longomatch.com

Mabiala, A. C., Sousa, A. W., Roman, N. T., Bernardes, J. L., Antunes, M. M., Ortega, E. M., . . . Lima, C. A. M. (2016). Introducing a language for human stance description. In V. G. Duffy (Ed.), *Digital Human Modeling: Applications in Health, Safety, Ergonomics and Risk Management: 7th International Conference, DHM 2016, Held as Part of HCI International 2016, Toronto, ON, Canada, July 17-22, 2016, Proceedings* (pp. 179-190). Cham: Springer International Publishing. Retrieved from https://doi.org/10.1007/978-3-319-40247-5_19. doi:10.1007/978-3-319-40247-5_19

Maddalozzo, G. (1987). An anatomical and biomechanical analysis of the full golf swing. *National Strength and Conditioning Association Journal, 9*(4), 6-8, 77-79.

Malik, J., Arbeláez, P., Carreira, J., Fragkiadaki, K., Girshick, R., Gkioxari, G., . . . Tulsiani, S. (2016). The three R's of computer vision: Recognition, reconstruction and reorganization. *Pattern Recognition Letters, 72*, 4-14. doi:https://doi.org/10.1016/j.patrec.2016.01.019

MAR. Retrieved 10 Sep, 2017, from http://www.mar-systems.co.uk

Marty, A. W., & Edwards, T. A. (2013). Trajectory detection and feedback system for golf: Google Patents.

Mathur, N., Mathur, S., & Mathur, D. (2016). A novel approach to improve sobel edge detector. *Procedia Computer Science, 93*(Supplement C), 431-438. doi:https://doi.org/10.1016/j.procs.2016.07.230

MathWorks. Create Angle Measurement Tool Using ROI Objects. www.au.mathworks.com: MathWorks. Retrieved from https://au.mathworks.com/help/images/create-angle-measurement-tool-using-roi-objects.html

MathWorks. DetectPeopleExample. Retrieved from https://au.mathworks.com/help/vision/ref/vision.peopledetector-system-object.html?searchHighlight=vision.PeopleDetector&s_tid=doc_srchtitle

MathWorks. *Machine learning technique for finding hidden patterns or intrinsic structures in data*. Retrieved 2 Oct, 2017, from https://au.mathworks.com/discovery/unsupervised-learning.html

MathWorks. *Supervised learning workflow and algorithms*. Retrieved 2 Oct, 2017, from https://au.mathworks.com/help/stats/supervised-learning-machine-learning-workflow-and-algorithms.html

MathWorks. (2017). *Image processing toolbox (R2017b)*. Retrieved 27 Feb, 2018, from https://au.mathworks.com/help/images/ref/hough.html

Matyunin, S., Vatolin, D., Berdnikov, Y., & Smirnov, M. (2011, 16-18 May 2011). Temporal filtering for depth maps generated by Kinect depth camera Symposium conducted at the meeting of the 2011 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON) doi:10.1109/3DTV.2011.5877202

Maurer, W. G. (2016). *Bird and bat interaction vision-based detection system for wind turbines*. Oregon State University, OSU Libraries. Retrieved from https://ir.library.oregonstate.edu/xmlui/bitstream/handle/1957/58462/MaurerWilliamG2016.pdf?s

Maxion, R. (2009, 1-4 Sept. 2009). Experimental methods for computer science research Symposium conducted at the meeting of the 2009 Fourth Latin-American Symposium on Dependable Computing doi:10.1109/LADC.2009.29

Meeds, E. W., Ross, D. A., Zemel, R. S., & Roweis, S. T. (2008, 23-28 Jun.). Learning stick-figure models using nonparametric Bayesian priors over trees. Symposium conducted at the meeting of the 2008 IEEE Conference on Computer Vision and Pattern Recognition doi:10.1109/CVPR.2008.4587559

Microsoft. *Depth camera*. Retrieved 27 Jul, 2017, from https://msdn.microsoft.com/en-us/library/hh438997.aspx

Mishra, P. K., & Saroha, G. P. (2016, 16-18 March 2016). *A study on video surveillance system for object detection and tracking*. presented at the meeting of the 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India. Retrieved from http://ieeexplore.ieee.org/document/7724259/

Paul, M., Haque, S. M. E., & Chakraborty, S. (2013). Human detection in surveillance videos and its applications - a review. *EURASIP Journal on Advances in Signal Processing, 2013*(1), 176. doi:10.1186/1687-6180-2013-176

Peffers, K., Tuunanen, T., Rothenberger, M., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems*, *24*(3), 45-77. doi:10.2753/mis0742-1222240302

Photoshop. Retrieved 24 Feb, 2018, from https://www.adobe.com/nz/products/photoshop.html

Practice to improve your golf motor skills. Retrieved 5 Jun, 2017, from http://golf-info-guide.com/golf-tips/the-golf-swing/practice-to-improve-your-golf-motor-skills/

Quintic. Retrieved 10 Sep, 2017, from http://www.quinticsports.com/

R.Hemalatha, E.Deepa, & R.Sasipriya. (2015). Automated driving vehicle using digitalimage processing. *International Journal of Innovative Science, Engineering & Technology, 2*(9).

Raskar, R., & Cohen, M. (1999). *Image precision silhouette edges*. presented at the meeting of the Proceedings of the 1999 symposium on Interactive 3D graphics, Atlanta, Georgia, USA. doi:10.1145/300523.300539

Rezaei, M., & Klette, R. (2017). Object detection, classification, and tracking. In *Computer Vision for Driver Assistance: Simultaneous Traffic and Driver Monitoring* (pp. 51-94). Cham: Springer International Publishing. Retrieved from http://dx.doi.org/10.1007/978-3-319-50551-0_4. doi:10.1007/978-3-319-50551-0_4

Rocha, N. I. (2006). *Golf bag*. Washington, DC, U.S.

Rouillon, O., & Rouch, P. (2015). *Low back pain: Biomechanics analysis and prevention*. presented at the meeting of the 33rd International Conference on Biomechanics in Sports, Poitiers, France. Retrieved from https://ojs.ub.uni-konstanz.de/cpa/article/view/6677/0

Ruiter, K. *Golf tips for the driver*. Retrieved 2016-12-12, from http://www.free-golf-lessons.com/free-online-golf-lessons.html

Rybarczyk, Y. (2016). 3D markerless motion capture: A low cost approach. In Á. Rocha, A. M. Correia, H. Adeli, L. P. Reis, & M. Mendonça Teixeira (Eds.), *New Advances in Information Systems and Technologies* (pp. 731-738). Cham: Springer International Publishing. Retrieved from http://dx.doi.org/10.1007/978-3-319-31232-3_68. doi:10.1007/978-3-319-31232-3_68

Shaikh, S. H., Saeed, K., & Chaki, N. (2014). Moving object detection approaches, challenges and object tracking. In *Moving object detection using background subtraction* (pp. 5-14). Cham: Springer International Publishing. Retrieved from https://doi.org/10.1007/978-3-319-07386-6_2. doi:10.1007/978-3-319-07386-6_2

SiliconCoach. Retrieved 16 Dec, 2016, from www.siliconcoach.com

Simi. Retrieved 10 Sep, 2017, from http://www.simi.com/en/

Singla, N. (2014). Motion detection based on frame difference method. *International Journal of Information & Computation Technology, 4*, 1559 - 1565.

SmartSwing. Retrieved 4 Oct, 2017, from http://www.smartswinggolf.com/site/index.html

Stauffer, C., & Grimson, W. (1999). Adaptive background mixture models for real-time tracking. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR 1999)*: IEEE. Retrieved from http://ieeexplore.ieee.org/document/784637/

Stickney II, T. (2016). *3 keys you need to understand about impact to play better golf*. Retrieved 6 Nov, 2017, from http://www.golfwrx.com/387202/3-keys-you-need-to-understand-about-impact-to-play-better-golf/

Stuart, F. (2017What is human movement. Retrieved from http://blog.acpe.edu.au/index.php/health-fitness/human-movement/

SwingMoment. (2016, 10 Sep.). *The best golf swing slow motion - online golf lesson* [Video]. Retrieved from https://www.youtube.com/watch?v=2dk5vqZ_7Zg

Szeliski, R. (2011). Dense motion estimation. In *Computer Vision: Algorithms and Applications* (pp. 335-374). London. Retrieved from https://doi.org/10.1007/978-1-84882-935-0_8. doi:10.1007/978-1-84882-935-0_8

Thériault, G., & Lachance, P. (1998). Golf Injuries: An Overview. *Sports Medicine, 26*(1), 43-57.

Tinning, R. (2009). *Pedagogy and human movement : Theory, practice, research*: New York : Routledge, 2009.

Training course for coaches - Barcelona. Retrieved 15 Sep, 2017, from https://www.ertheo.com/en/trainers-formation-barcelona.htm

van den Boomgaard, R., & van Balen, R. (1992). Methods for fast morphological image transforms using bitmapped binary images. *CVGIP: Graphical Models and Image Processing, 54*(3), 252-258. doi:http://dx.doi.org/10.1016/1049-9652(92)90055-3

Vedaldi, A., & Fulkerson, B. (2010). *Vlfeat: An open and portable library of computer vision algorithms*. presented at the meeting of the Proceedings of the 18th ACM international conference on Multimedia, Firenze, Italy. doi:10.1145/1873951.1874249

VideoLAN. Retrieved 20 Feb, 2018, from https://www.videolan.org/vlc/

VirtualDub. Retrieved 20 Dec, 2017, from http://www.virtualdub.org/

vision.BlobAnalysis System object. Retrieved 1 Nov, 2017, from https://au.mathworks.com/help/vision/ref/vision.blobanalysis-system-object.html

Wang, R., Bunyak, F., Seetharaman, G., & Palaniappan, K. (2014, 23-28 Jun.). Static and moving object detection using flux tensor with split gaussian models. Symposium conducted at the meeting of the 2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops doi:10.1109/CVPRW.2014.68

Wieringa, R. (2010, 2-8 May 2010). Design science methodology: principles and practice Symposium conducted at the meeting of the 2010 ACM/IEEE 32nd International Conference on Software Engineering doi:10.1145/1810295.1810446

Winter, D. A. (2009). *Biomechanics and motor control of human movement* (4th ed.). New Jersey: John Wiley & Sons, Inc.

Wu, C.-J., Quigley, A., & Harris-Birtill, D. (2017). Out of sight: A toolkit for tracking occluded human joint positions. *Personal and Ubiquitous Computing*, *21*(1), 125-135. doi:10.1007/s00779-016-0997-6

Yu, E., & Aggarwal, J. K. (2011). Human extremity detection for action recognition. In P. S. P. Wang (Ed.), *Pattern Recognition, Machine Intelligence and Biometrics* (pp. 237-260). Berlin, Heidelberg. Retrieved from https://doi.org/10.1007/978-3-642-22407-2_10. Retrieved 2017 Oct 13. doi:10.1007/978-3-642-22407-2_10