PSEUDO-3D BINARY SILHOUETTE FOR AUGMENTED GOLF COACHING

Kam Yuk (Leona) Chan and Boris Bacic

School of Engineering, Computer and Mathematical Sciences, Auckland University of Technology, Auckland, New Zealand

The purpose of this paper is to report the augmented golf coaching implications of novel annotations solutions and silhouette-based video transformations from commonly used video sources. The annotations include ‘one-click’ 3D grid drawings and automated club head tracking. The developed software is based on computer vision approaches to facilitate golf coaching from beginner to intermediate skill-level players and was tested on two driving ranges under different lighting conditions. The end-of-session reporting enables coaches to provide visual annotations of elements of swing performance with privacy preservation via the conversion of captured media files into a two-colour pseudo-3D silhouette. Future work is aimed at advancements of preserving video streaming and technology transfer to advance diverse sports disciplines, healthcare and sports science.

KEYWORDS: augmented video coaching, qualitative movement diagnosis, performance

INTRODUCTION: Augmented golf coaching technology can provide information on hard-to-see movement sequences. Commonly used video sources such as mobile technology combined with computer vision approaches are able to automate video capture and swing detection aimed at providing evidence to support *qualitative movement diagnosis* (QMD) of a golf swing. However, the growing popularity of on-line video coaching and other *augmented coaching systems and technology* (ACST) present challenges regarding privacy concerns. To address these challenges, the initial study relying on Microsoft Kinect depth video demonstrated a 2D silhouette video transformation was sufficient to support QMD of golf swing (Bacic, Meng, & Chan, 2017). The objective of this research is to present an advanced privacy-preserving pseudo-3D silhouette-based augmented coaching solution that can utilise a variety of commonly-used video sources such as mobile and action cameras.

METHODS: This study is based on the *design science* approach and external validation of produced *artefacts* supporting augmented golf coaching. The design science approach and artefacts production included various techniques from computer vision and video and image processing supporting users, tasks and environments (Mingers, 2001). The equipment and recording setup specific to golf data collection included the following:

- Video sources and hardware: iPhone (7 plus and 5S), a compact camera Casio Exilim (EX-ZR100 and EX-FH25) and GoPro Hero 3 for algorithm development and testing. High-speed cameras are required for capturing high-speed golf club movement.
- Video settings: 30 to 240 frames per seconds (fps).
- Fixed location monocular camera view: side view – sagittal plane, camera height at approximate hip and waist level allowing capture of a club movement during the swing.
- Software: Matlab 2016b and computer vision related toolboxes for algorithm development.

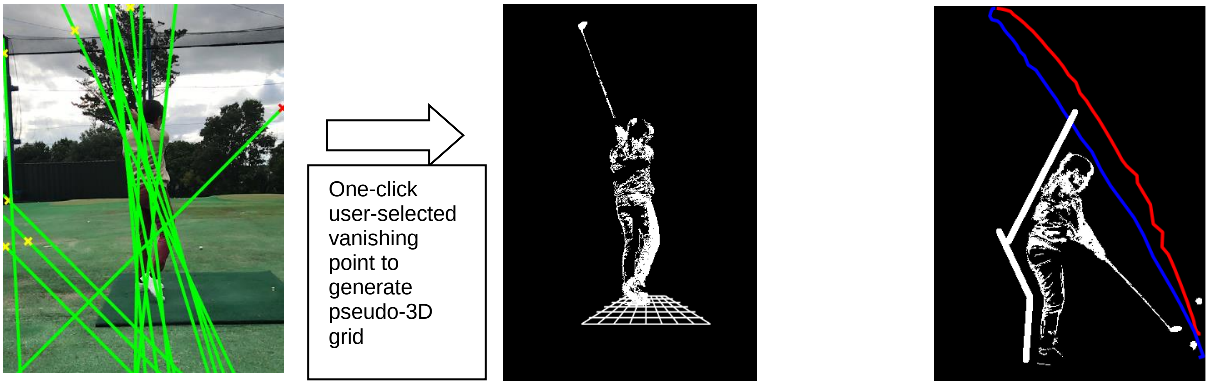
Each of the three sessions was recorded at different times of the day and lighting conditions in two different golf driving range environments. The collected video dataset also captured evidence of improved elements of swing technique, fostering the need for personalised annotation and augmented coaching feedback. All captured and transformed videos of the first author’s golf activities are presented in sagittal view.

The purpose of the developed foreground-background separation algorithm is to extract the golfer’s activity while filtering the redundant information from video frames. Compared with an indoor setting, in the case of outdoor golf sports activities, the scene includes static and dynamic foreground and background information, which represented a computational challenge. For example, the dynamic background may be caused by moving clouds, shadows and leaves. In addition, the golfer can also represent the static foreground assuming a stance position before the swing, resulting in minor change of pixel values. In contrast, high-speed club movement represents dynamic foreground information. Relying on cyclic development over three repeated data collection/training sessions (using an incrementally improved augmented coaching prototype) it was possible to develop a *pseudo-3D silhouette* algorithm whose performance is depicted in Figures 1(a-f). The algorithm developed is also based on an approach used in a tennis case study in the area of computer vision (Kelly et al., 2010) on uniformly coloured surface backgrounds. To achieve foreground-background separation, the algorithm classified each pixel to a background or foreground based on differences in video frames, brightness and colour. Figure 1a shows the original RGB image and the result of the pseudo-3D binary silhouette solution alternatives in black (Figure 1b) and white (Figure 1c).

(a) RGB image from video

(b) Developed solution is suitable for screen annotations and viewing

(c) Reversed black and white image from (b) for a hard-copy end-of session report



(d) Computer-generated lines related to possible vanishing point on the image

(e) 3D volume representation

(f) computer-generated club head trajectory, –backswing, –downswing.

Figure 1: Extracted foreground image and interactive annotation functions.

The main objective of the solution in Figure 1, is to provide privacy preservation in video while maintaining visual information needed to support the coaching process and diagnostic capabilities. Compared to 2D silhouette from Kinect sensor depth video, the pseudo-3D binary silhouette solution developed is able to work with a range of commonly-used video sources and can provide additional visual representation of a golfer's activity (Table 1).

Table 1: Comparisons of 2D silhouette vs. pseudo-3D silhouette for augmented golf coaching.

Key features	Description
Frame rate	Kinect frame rate is limited to 30 fps, which is insufficient to capture fast movements of a golf club. The higher frame rates (240 fps) of e.g. smartphone video can capture fast movements such as the moment of impact with the ball or the last frame position of the backswing.
Silhouette image	Kinect's 2D silhouette solution was originally designed to maximise privacy preservation and to be universally applicable in healthcare, elderly care, resident care and smarthomes. The enhanced black and white pseudo-3D silhouette can display a greater level of details on the image needed for QMD.
Golf club tracking	In Kinect-based video, most of the time the golf club movement is invisible. In pseudo-3D silhouette video, the golf club can be traced in high frame rates.
Minimising coaches' bias	Pseudo-3D binary silhouette can minimise potential bias including: (1) Loose clothing (e.g., determine the knee angle with loose trousers); (2) Separation of shoe and ground when the feet of the golfer are moving throughout the swing.

Table 2: Practical implications of pseudo-3D silhouette video filtering for coaching practice.

Coaching cue and visual focus	Attention to body segments	Static	Dynamic
1. Grip	Grip – hands holding a club	x	x
2. "Eyes on the ball"	Eye movement and fixation	x	x
3. Head position	Head movement	B	B
4. Stance	Body and posture assessment	✓	N/A
5. Knee angle and movement	Knees flection angle (wearing loose clothing)	✓	B
6. Upper body or spine angle	Trunk flection angle	✓	✓
7. Balance	Static and dynamic balance	✓	✓
8. Hands and arms softness	Relative estimate (including joint angles of non-occluded arm)	✓	✓
9. Swing plane (e.g. 'too steep' or 'too flat')	Club swing tracking through the swing phases (Figure 1 f).	✓	✓
Augmented feedback functions	**Description**	**Static**	**Dynamic**
1. Swing plane and joint angles visual information	Angles annotation in 2D view	✓	N/A
2. Estimation of vanishing point	Computer-generated sets of intersected lines	✓	N/A
3. Interactive pseudo-3D (perspective) grid annotation	Pseudo-3D grid drawing by choosing the intersection point.	✓	x

Note: ✓… Yes; **x** … No; **B** … potential bias; N/A … not applicable; Static … frame; Dynamic … movement video

RESULTS AND DISCUSSION: In comparison with prior work on 2D silhouette (Table 1), the reported results of external evaluation of the pseudo-3D binary silhouette in golf coaching environments, include functionality for QMD tasks and coach-specific feedback (Table 2). The developed software tools can support QMD and facilitate coach-learner communication. The solution developed also includes two functions to support a coach to minimise his/her cognitive load associated with interactive tasks involved in providing annotations of QMD of a golf swing. The first function is interactive annotation of a 3D perspective grid on a 2D image (Figure 1e), which is manual drawing on RGB before conversion into pseudo-3D. The drawing of overlay grid can help the golfer and end user to visualise the ground and 3D volume perspective, given that a silhouette visually represents a 2D plane. The existing video coaching tools such as Kinovea (www.kinovea.org) already provide four interactively adjustable points for grid drawing function, which is time-consuming and requires skills in 3D perspective drawing. The one-click interactive pseudo-3D grid drawing function can automatically generate groups of lines crossing the possible vanishing point (Figure 1d). The second interactive annotation function represents the golf club head trajectory (Figure 1f) during the swing. The function is based on the monocular markerless video tracing of the golf club head to compute the club head trajectory combining the foreground image extraction and the multi-layered processing approach. Computer-generated club trajectory on a selected frame after the ball impact allows coaches to interpret the intended backswing and downswing annotations in red and blue. The experimental work suggests that a high-frame rate video (240 fps or higher) is strongly recommended for golf club head tracing. Unlike with the naked eye, captured video can provide more details of recorded movement than it is possible to observe in traditional coaching. The solution developed to facilitate golf coaching relies on video evidence representing captured human movement associated with golf swing activity. External validity of silhouette-based augmented coaching implications is achieved via consensus of the literature and expert opinions, which is also aligned with logical validity of QMD (Knudson 2013, p. 34). Table 2 summarises practical implications for golf coaches and shows to what degree the solution developed can aid augmented coaching analysis of the elements of performance that are typical for beginners to intermediate skill-level players.

CONCLUSION: The privacy-preserving computer-generated pseudo-3D silhouette, was used to convey feedback annotations of golf coaching activity recorded in two different golf driving range settings. The reported list of annotating elements is considered sufficient for producing feedback on common mistakes – that are typical for novices to intermediate skill-level golfers. The novel proof of concept for the one-click interactive pseudo-3D grid drawing representing the ground level was achieved by computer-generated visual clues 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 club head movement during the swing that allows coaches to draw intended swing plane over the produced static image. In addition to changes in colour, contrast, and light conditions, the major challenge was that each video pixel can represent static or dynamic foreground or background information, which is also a challenge for commonly used methods in surveillance, autonomous driving cars and computer vision in general. The produced pseudo-3D binary silhouette represents a visual alternative that can be used for coaching the golf swing while preserving privacy. Future work will include multi-camera views, enhancing 3D information and transfer to other sport disciplines, coaching environments and related problem areas.

REFERENCES

Bacic, 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

Mingers, J. (2001). Combining IS research methods: Towards a pluralist methodology. *Information Systems Research, 12*(3), 240-259.

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.

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

ACKNOWLEDGEMENTS: The authors wish to thank Auckland University of Technology (AUT) for the extended use of the golf driving range and Takapuna golf course driving range for allowing patrons to record their own golf swing videos.