Journal of Bodywork & Movement Therapies (2014) 18, 220–227

Available online at www.sciencedirect.com

SciVerse ScienceDirect

journal homepage: www.elsevier.com/jbmt

MOVEMENT IMAGING

# Hybrid markerless tracking of complex articulated motion in golf swings

**Sim Kwoh Fung, Kenneth Sundaraj, Nizam Uddin Ahamed*, Lam Chee Kiang, Sivadev Nadarajah, Arun Sahayadhas, Md. Asraf Ali, Md. Anamul Islam, Rajkumar Palaniappan**

*AI-Rehab Research Group, Universiti Malaysia Perlis (UniMAP), Kampus Pauh Putra, 02600 Arau, Perlis, Malaysia*

Received 6 March 2013; received in revised form 26 April 2013; accepted 2 May 2013

**KEYWORDS**
Golf swing;
Hybrid tracking;
Markerless tracking;
Motion tracking

**Summary** Sports video tracking is a research topic that has attained increasing attention due to its high commercial potential. A number of sports, including tennis, soccer, gymnastics, running, golf, badminton and cricket have been utilised to display the novel ideas in sports motion tracking. The main challenge associated with this research concerns the extraction of a highly complex articulated motion from a video scene. Our research focuses on the development of a markerless human motion tracking system that tracks the major body parts of an athlete straight from a sports broadcast video. We proposed a hybrid tracking method, which consists of a combination of three algorithms (pyramidal Lucas-Kanade optical flow (LK), normalised correlation-based template matching and background subtraction), to track the golfer's head, body, hands, shoulders, knees and feet during a full swing. We then match, track and map the results onto a 2D articulated human stick model to represent the pose of the golfer over time. Our work was tested using two video broadcasts of a golfer, and we obtained satisfactory results. The current outcomes of this research can play an important role in enhancing the performance of a golfer, provide vital information to sports medicine practitioners by providing technically sound guidance on movements and should assist to diminish the risk of golfing injuries.


* Corresponding author. Tel.: +60 49767399; fax: +60 49851695.
*E-mail address:* ahamed1557@hotmail.com (N.U. Ahamed).

1360-8592/$ - see front matter 
http://dx.doi.org/10.1016/j.jbmt.2013.05.011

## Introduction

The study of human motion is the interpretation of human activity for the description, analysis and assessment of human movement (Blake and Shiffrar, 2007; Luo et al., 2003; Castrodad and Sapiro, 2012). The application of human motion analysis is limitless and can be categorised into three major fields: surveillance, control and analysis (Moeslund and Granum, 2001; Moeslund et al., 2006; Couceiro et al., 2012). The increasing technological advancements in medicine have drawn attention to the development of human motion analysis in the fields of sports and biomedicine. In addition, human motion analysis can also be used as a training database that athletes can use to compare their techniques with a particular professional athlete who executes the same movement in a similar sport. Motion analysis can also be used in the rehabilitation phase of an injured athlete. Hence, incorrect movements of injured athletes can be used to recover the correct movement pattern with the help of a concerned athlete. To solve this problem, many researchers are attempting to develop a motion tracking system. The main challenge of this work is to retrieve important information of human trajectories from video content and to properly configure each articulated limb to mimic the human pose shown in the video. Thus, the majority of the existing approaches rely on videos with certain conditions, such as a constant moving camera and a static background (Gouwanda and Senanayake, 2008; Erp et al., 2006; Ali et al., 2012). However, most sport activities are captured from pre-recorded sport broadcasts with many unknown parameters. As a result, the athletes might present unexpected movements and extensive body articulations that may rapidly effect changes, self-occlusion and heavy motion blur. The cluttered background in sports videos also increases the complexity of the tracking process. These challenges make the results of the direct application of the existing tracking algorithms unsatisfying. Hence, the detection and tracking of players in broadcast sports videos remains an unsolved problem.

One of the most popular types of video processing methods for the extraction of independent foreground motion is background subtraction (Liu and Sarkar, 2004; Moeslund et al., 2006; Sim and Sundaraj, 2010; Gall et al., 2010). However, the problem associated with this method is the acquisition of an instantaneous reference image, especially when a large part of the background is occluded by the moving background objects or by parts of the background that were never observed and identified. To solve this problem, a single cue or feature is not sufficient to perform human motion tracking, and studies have shown that combining evidence from difference sources yield better tracking results compared with the single cue approach (Ruixuan, 2004; Yang et al., 2005; Tony and Matsuyama, 2008; Karliga and Hwang, 2007). The common approach used to track human body poses uses an a priori human model to guide the tracking process (Urtasun et al., 2005; Park et al., 2006). However, this method requires that the motion data library learn different types of motion patterns. To avoid the use of a human model, model-free approaches were introduced to perform direct human body posture estimations from a single image (Moeslund et al., 2006; Wang et al., 2003; Karliga and Hwang, 2007; Ruixuan, 2004). This method can build a simple model to represent the pose in the initialisation stage as an interpretative guide for the tracking process.

Even though golf has not been familiar as a game related with injuries, scientific studies related to golf state that back and elbow injuries are most frequent in male recreational golfers (McCarroll, 1996; Meira and Brumitt, 2010). In addition, wrist and lower back injuries occur in male and female golf professionals. On the other hand, the elbow and the lower back are the most common areas of injury associated with female amateur golfers. These injuries are mainly associated to poor conditioning, overuse, lack of technical information and improper swing mechanics. Therefore, we envisage that this study, which embarks on the use of a hybrid markerless tracking approach, will produce outcomes that will contribute to prevention of golf-related injuries.

The primary goal of this project was to develop a human motion tracking system that is able to track the motion of a professional golfer performing a full swing from a sports broadcasting video. The tracking process was performed using a genuine sports competition video; hence, video-related problems, such as the complexity of the background, illumination changes and shadows, were our main concern. In addition, no special markers were attached to the subject's body to aid the tracking. We wanted our tracking system to track the major body parts, including the head, body, hands, shoulders, knees and feet, as the golfer performed a full swing. Therefore, we combined both motion and feature cues by introducing a hybrid tracking approach to track the subject's points of interests in each consecutive frame. The tracking outputs are represented in an articulated human skeleton that mimics the exact pose of the subject during the execution of the motion.

## Materials and methods

In this work, we used an unconstrained sports video as our input data. The video was obtained directly from the broadcast and contains several types of sports actions with a cluttered background (Reddy, 2010). In fact, one of the most challenging problems related to the tracking of the golfer concerns the occlusion of the body movements and the rapid movements of the swing. Our human motion analysis follows the functional taxonomy summarised in the survey work by Moeslund and Granum (2001), which is shown in Fig. 1.

### Initialisation

In our work, we utilised the initialisation stage to determine the location of the golfer's body parts. This allowed us to separate the foreground human subject from the background and provided useful information for the tracking. The LK algorithm was then used to track the motion of the body parts over time. However, this algorithm yields a very poor result when problems, such as occlusion and rapid changes of the swing speed, arise. To solve this problem,

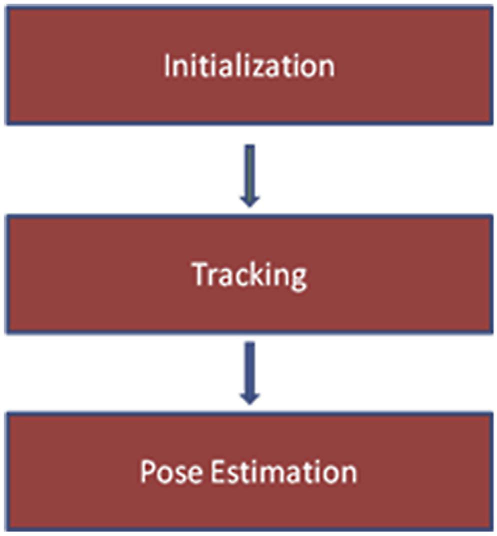


**Figure 1** General structure of the proposed motion tracking system.

we proposed a hybrid tracking method that consists of the combination of three algorithms (pyramidal LK, normalised correlation-based template matching and background subtraction) to enhance the tracking result. Finally, we matched, tracked and mapped the output result onto a 2D articulated human stick model to represent the pose of the golfer using the model initialisation obtained from the first frame of the video scene. In our approach, we placed 10 artificial markers onto the subject body; eight moving markers and two fixed markers. The two fixed markers were placed at the centre of the torso and at the centre of the lower limb. The remaining eight artificial markers were placed in the following specific order; the head, the left/right shoulders, the left/right knees, the left/right foot and the binding hands. The ($x$,$y$) coordinates of the moving markers served as the input data for the subsequent

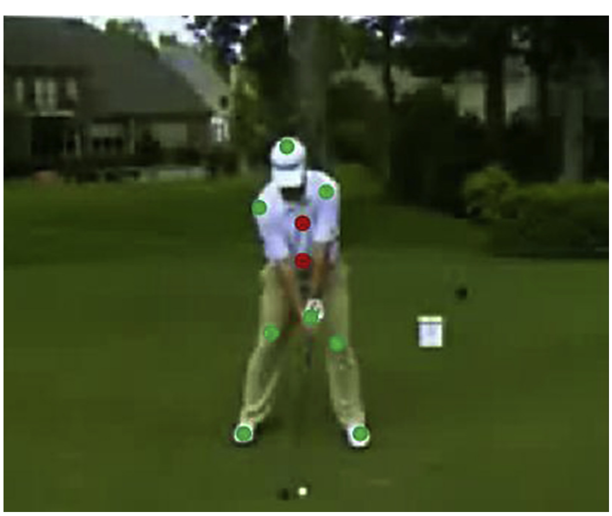

**Figure 2** Artificial marker placement: Static markers (red) and moving markers (green). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

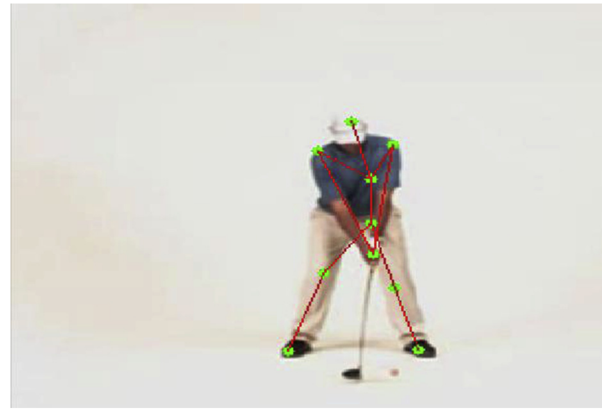

**Figure 3** Placement of 10 artificial markers on the golfer's body to form an articulated stick model. The green points represent the skeleton joints, and the red lines are the limbs. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

tracking algorithm. A schematic of the marker placement on the subject's body is shown in Fig. 2.

## Tracking (hybrid motion tracking)

To obtain the trajectories of the player, we needed to take into account the detection and tracking of the subject. However, due to the complex movement of the athlete, we used motion (optical flow and background subtraction) and feature (template matching) cues to track the head, shoulder, knees, and feet of a golfer performing a full swing from a broadcast competition video. We applied the pyramidal LK algorithm to track each point of interest of the golfer. The coordinates and the intensity value of these points were obtained from the initialisation process and fed into the pyramidal LK algorithm. This pyramidal LK algorithm exhibited some limitations when part of the golfer's body was occluded during the golf swing action.

**Table 1** Specifications of the test videos.

| No | Types of action | Video descriptions | Video complexity |
|---|---|---|---|
| 1 | Golf swing 1 | • Plain white background<br>• Indoor conditions | Low |
| 2 | Golf swing 2 | • Moderate background without spectators<br>• Outdoor conditions<br>• Competition broadcast video | Medium |
| 3 | Golf swing 3 | • Complex background with spectators<br>• Outdoor conditions<br>• Competition broadcast video | High |

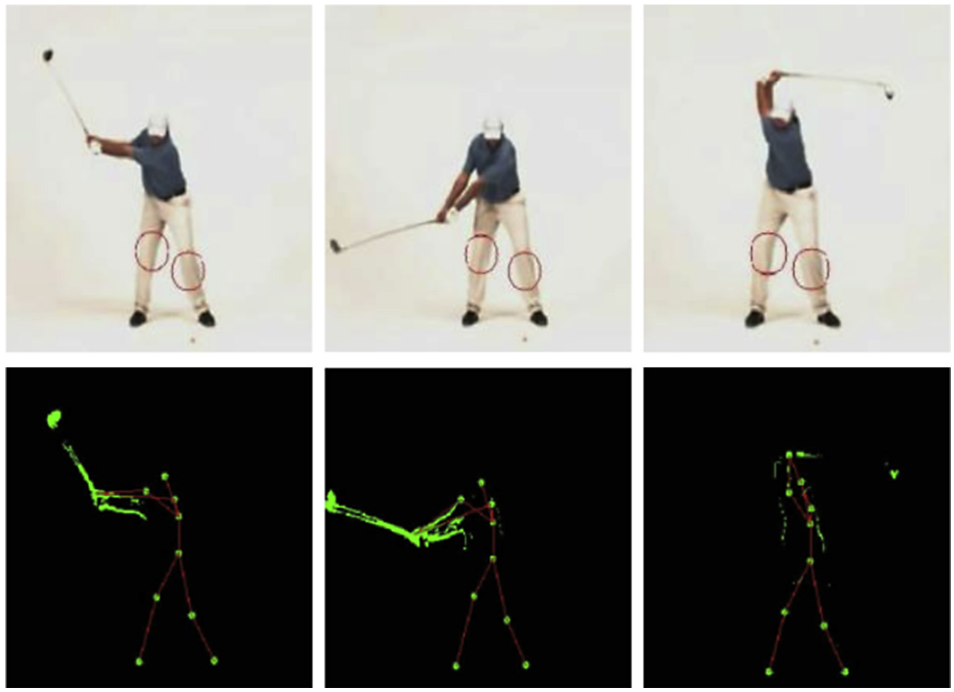

**Figure 4** Selected tracking results obtained during the complete golf swing 1 video.

We therefore wanted to overcome the occlusion problem that is caused by the golf club overlapping the golfer's knees during the full swing action. Therefore, we applied normalised correlation-based template matching to track the left and right knees of the golfer. The matching process moves the template images around the image space in every sequence of the frame to identify the best match of the template in the search image. However, no human performs their motions at the same speed. Thus, the use of the optical flow algorithm combined with normalised correlation-based template matching gave unsatisfactory results for the tracking of the golf club during the downward swing sequence. To solve this particular problem, we used a background subtraction technique to track the golf club during swing.

## Pose estimation

Pose estimation is the process used to identify how human limbs are configured in a given scene. However, we used a human stick model-free approach to obtain knowledge of the subject's body configuration in the primary stage of the initialisation process. Using this model, we transferred the resulting motion data of the golfer onto a human skeleton to obtain more accurate pose estimations during the entire tracking process. We had previously assigned 10 artificial markers on the golfer in the first frame of the video to obtain information of the human body. Each artificial marker represents a junction of the body and provides its coordinates; this data is then linked to form a 2D human articulate stick model, as shown in Fig. 3.

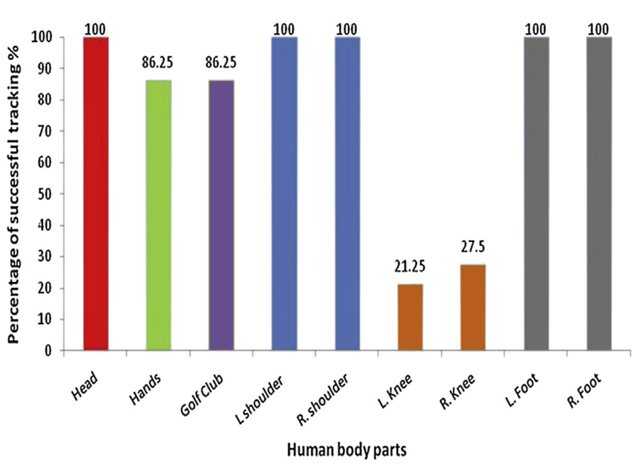


**Figure 5** Histogram of the successful tracking of each point on the golf swing 1 video using the pyramidal LK algorithm.

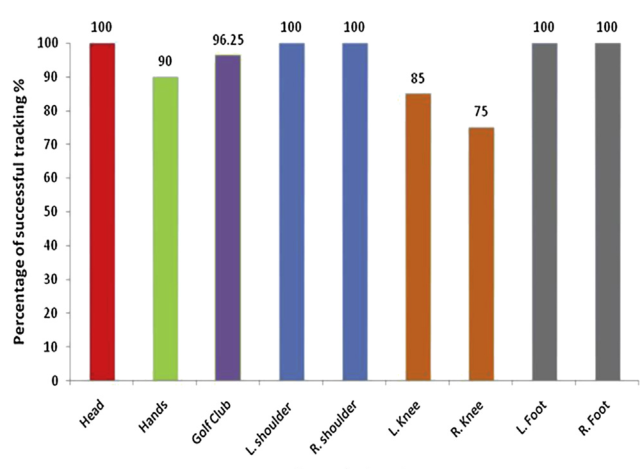


**Figure 6** Histogram of the successful tracking of each point on the golf swing 1 video using our proposed hybrid system.

**Figure 7** Selected tracking results obtained during the complete golf swing 2 video.

## Results

Three experiments were conducted using our approach to validate the robustness of the proposed motion tracking system. A detailed list of the test videos used in our experiments is provided in Table 1.

Selected tracking results obtained during the complete golf swing 1 video is presented in Fig. 4. In the first experiment, the golf swing 1 video was used as the input data to track the golfer's head, left/right shoulders, left/right knees, left/right feet, binding hands and the golf club in a simple background. In this experiment, only the pyramidal LK algorithm was used to track the complex golf swing motion. The percentage by which each point on the golf swing 1 video was successfully tracked is shown in Fig. 5.

In the second experiment, the golf swing 1 video was used as the input data for the validation and benchmarks of our proposed hybrid system. Three tracking algorithms (LK algorithm, normalised correlation-based template matching, and background subtraction) were executed simultaneously to track the golfer's head, binding hands, left/right shoulders, left/right knees, left/right feet and the golf club. The percentage by which each point on the golf swing 1 video was successfully tracked is shown in Fig. 6.

Selected tracking results obtained during the complete golf swing 2 and golf swing 3 videos are presented in Fig. 7 and Fig. 9, respectively. In the third experiment, we validated the robustness of our tracking system. Two broadcast videos (golf swing 2 and golf swing 3) were selected to test the background complexity, illumination changes and shadows. The settings used for Experiment 2 were used for the calibration parameters. Our proposed tracking system was then evaluated using these broadcast videos. Figs. 8 and 10 show the results obtained using the golf swing 2 and the golf swing 3 videos respectively. A summary of the results of the hybrid tracking system is given in Table 2.

## Discussion

Golf swing is a complex movement of the entire body to produce power to a golf ball to force the ball to great distances with accuracy (McHardy and Pollard, 2005; Coleman and Anderson, 2007; Bunker et al., 2011). It is also

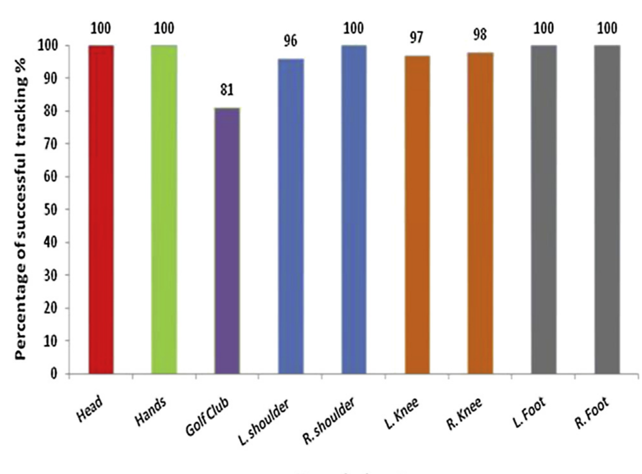


**Figure 8** Histogram of the successful tracking of each point on the golf swing 2 video.

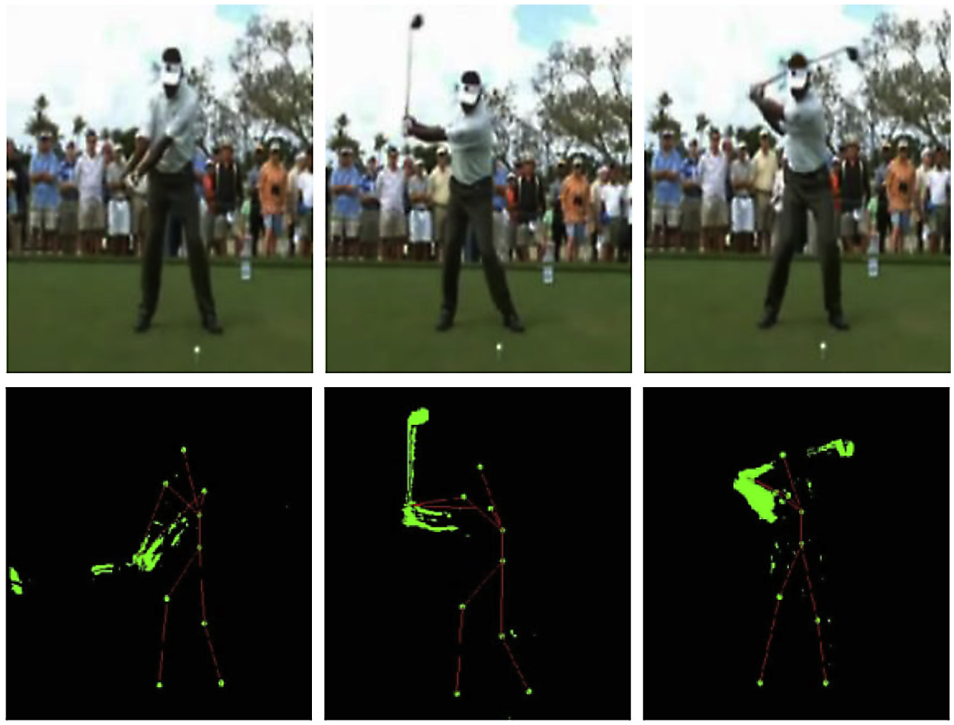

Figure 9 Selected tracking results obtained during the complete golf swing 3 video.

a fast movement sport, which makes it complicated for the golfer to identify and characterize accurately his swing motion (Negoro et al., 2011). Therefore, if players are not competent enough to find out about their swing motion by themselves, they would have lost precious information to improve their swing, increase scores and avoid injuries. Up to the present many experts have offered many novel approaches to the game, however a review of the scientific literature reports limited research evaluating the actual biomechanics of the golf swing in comparison with entire body movement and markerless tracking system (Egret et al., 2003; McHardy et al.,, 2006).

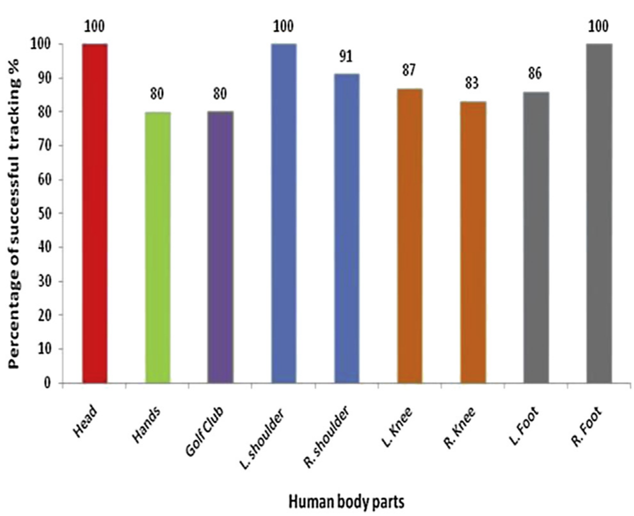


**Figure 10** Histogram of the successful tracking of each point on the golf swing 3 video.

In this paper, we hypothesise that relying on only a single cue for robust tracking will exhibit limitations. Therefore, we enhanced the tracking algorithm using hybrid tracking techniques. Because we used an unconstrained sports video input, the video parameters were not taken into consideration. Our proposed method was able to correctly track each body part. The following experiments were conducted to validate our proposed method.

## Experiment 1: motion tracking using a single feature or cue (optical flow)

In Experiment 1, the golf swing 1 video was tested using the pyramidal LK algorithm only. The user manually placed two static markers and eight moving markers on the subject's body during the first initialising frame. Although the results exhibited an overall success rate of 80.1%, the results of Experiment 1 were clearly incorrect for the knees. This unsuccessful result is due to the occlusion that occurred when the golf club's shaft overlapped the knees. Based on our observations, the tracking of the golf club also failed due to the very high speed thrust during the downswing action. In conclusion, the pyramidal LK algorithm is unable to perform robust tracking of complex human motions.

**Table 2** Summary of the results of the hybrid tracking system.

| Input | Successful tracking rate (%) |
|---|---|
| Golf swing 1 | 94.03 |
| Golf swing 2 | 96.89 |
| Golf swing 3 | 89.67 |

**Table 3** List and description of tracking failures.

| Frames | Body parts | Description |
|---|---|---|
| 1—10 | Binding hands | The subject is stationary. Hence, the background subtraction algorithm fails to detect the moving arms. |
| 80—90 | Right knee | A small deviation occurred at the right knee due to occlusion by the golf club. |
| 142—157 | Right knee | A small deviation occurred due to template drift because the template image specified by the flood fill algorithm is too large. |
| 281—287 | Right knee | Mismatch problem: the right knee was mistakenly matched to the left knee region. |
| 287—302 | Left knee | A small deviation caused by template drift |
| 323—350 | Right knee | Mismatch problem: the right knee was mistakenly matched to the left knee region. |
| 357—400 | Right and left knees | The algorithm failed to perform due to the disappearance of the knee features as the leg twisted to the left. |
| 360—400 | Binding hands | Tracking loss due to the high acceleration of the ball during the hitting action. |

We therefore developed our proposed hybrid tracking algorithm, which was tested in the next two experiments.

## Experiment 2: motion tracking based on the proposed hybrid method

In the initialisation stage, the same 10 artificial markers were manually placed on the body of the subject in the golf swing 1 video. To synchronise the three algorithms and obtain optimal tracking results, the calibration process which involves setting of parameters required for the three algorithms, must be carefully performed on each video. The results of Experiment 2 demonstrate that our proposed hybrid tracking method exhibited a significant improvement over the LK algorithm used in Experiment 1. The results clearly show that our system is able to solve the two major problems that were found in Experiment 1 — the occlusion of the golfer's knees and the rapid movement of the golf club. However, a 5.97% of the total tracking error was still unaccountable (meaning the reason for the observed error could not be directly identified). Although this error rate is low, we still need to intensively investigate it further improve our proposed hybrid system. We therefore summarise all of the tracking failures in Table 3.

## Experiment 3: testing the tracking system on a complex background

Based on the results of Experiment 3, it is clear that our proposed hybrid tracking system is able to cope with an unconstrained background environment, particularly that of a broadcast video. Our tracking system exhibits a very high success rate in both a moderate background (96.89%) and a complex background (89.67%). We identified some of the intermittent problems that contributed to the remaining errors in certain frames of the golf swing 2 and golf swing 3 videos. In the golf swing 2 video, only a few mismatched problems occurred in certain frames; these problems were due to template drift and point dislocation as a result of the background colour. As a result of these problems, attention was drawn to the golf club when the system was unable to locate the golf club. Problems also occurred when the golf club comes to a halt as it builds up momentum for the downward swing. The complexity of the background in the golf swing 3 video slightly affected the overall results. The spectators in the background are likely to cause some loss in the tracking because some of the spectators might have the same contrast value as the tracked features. This impact was observed when the golfer is commencing his downward swing action. Furthermore, although our system does not have a limit of detection, it was able to track nine body parts of the subject in each video frame. The system also successfully aligned the tracked results onto a 2D articulated stick human form.

# Conclusions

Playing golf is generally considered as a low risk and injury free sports in terms of movements. However, the lack of proper technical guidance and the addiction to golf results in overly spending too much time in this sport. This very often leads to muscle and joint problems. We proposed a non-invasive solution to this issue. The core of our approach is to track a single golfer performing a full swing from a sports broadcasting video and map the tracking results onto a 2D articulate human skeleton model. We combined both motion and feature cues to extract the movement of the golfer's body using a hybrid motion tracking technique that combines three algorithms; pyramidal Lucas-Kanade optical flow, background subtraction and normalised correlation-based template matching. We decided to avoid the use of a model-based approach for the pose estimation by manually defining the body location using artificial markers in the very first frame of the video scene. A comprehensive set of experiments was performed to validate the performance of our motion tracking system. The experimental results clearly show that our system can track complex motions with various background complexities. The results of our hybrid approach provide vital information on body parts movement which can be used to reduce the injury to a golf player.

# References

Ali, A., Sundaraj, K., Ahmad, B., Ahamed, N., Islam, A., 2012. Gait disorder rehabilitation using vision and non-vision based sensors: a systematic review. Bosnian Journal of Basic Medical Sciences 12 (3), 193—202.

Blake, R., Shiffrar, M., 2007. Perception of human motion. Annual Review of Psychology 58, 47—73.

Bunker, D.J., Rhea, M.R., Simons, T., Marin, P.J., 2011. The use of whole-body vibration as a golf warm-up. The Journal of Strength & Conditioning Research 25 (2), 293—297.

Castrodad, A., Sapiro, G., 2012. Sparse modeling of human actions from motion imagery. International Journal of Computer Vision 100 (1), 1—15.

Coleman, S., Anderson, D., 2007. An examination of the planar nature of golf club motion in the swings of experienced players. Journal of Sports Sciences 25 (7), 739—748.

Couceiro, M.S., Portugal, D., Gonçalves, N., Rocha, R., Luz, J.M.A., Figueiredo, C.M., Dias, G., 2012. A methodology for detection and estimation in the analysis of golf putting. Pattern Analysis and Applications, 1—16.

Egret, C.I., Vincent, O., Weber, J., Dujardin, F.H., Chollet, D., 2003. Analysis of 3D kinematics concerning three different clubs in golf swing. International Journal of Sports Medicine 24 (6), 465.

Erp, J.B.V., Saturday, I., Jansen, C., 2006. Application of tactile displays in sports: where to, how and when to move. In: Proceedings of the Eurohaptics International Conference, EuroHaptics 2006, Paris, France, pp. 90—95.

Gall, J., Yao, A., Gool, L.V., 2010. 2D action recognition serves 3D human pose estimation. In: Proceedings of the 11th European Conference on Computer Vision, Part III, Heraklion, Crete, Greece, 6313, pp. 425—438.

Gouwanda, D., Senanayake, S., 2008. Emerging trends of body-mounted sensors in sports and human gait analysis. In: Proceedings of the 4th Kuala Lumpur International Conference on Biomedical Engineering, IFMBE Proceedings, vol. 21, pp. 715—718.

Karliga, I., Hwang, J., 2007. Extraction and integration of human body parts for 3-D motion analysis of golf swing from single-camera video sequences. In: IEEE International Symposium on Circuits and Systems, pp. 3960—3963.

Luo, Y., Wu, T., Hwang, J., 2003. Object-based analysis and interpretation of human motion in sports video sequences by dynamic Bayesian networks. Computer Vision and Image Understanding 92 (2), 196—216.

Liu, Z., Sarkar, S., 2004. Challenges in segmentation of human forms in outdoor video. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, p. 43.

McCarroll, J.R., 1996. The frequency of golf injuries. Clinical Journal of Sports Medicine 15 (1), 1—7.

McHardy, A., Pollard, H., 2005. Muscle activity during the golf swing. British Journal of Sports Medicine 39 (11), 799—804.

McHardy, A., Pollard, H., Luo, K., 2006. Golf injuries: a review of the literature. Sports Medicine 36 (2), 171—187.

Meira, E.P., Brumitt, J., 2010. Minimizing injuries and enhancing performance in golf through training programs. Sports Health: A Multidisciplinary Approach 2 (4), 337—344.

Moeslund, T.B., Granum, E., 2001. A survey of computer vision-based human motion capture. Computer Vision and Image Understanding 81 (3), 231—268.

Moeslund, T.B., Hilton, A., Krüger, V., 2006. A survey of advances in vision-based human motion capture and analysis. Computer Vision and Image Understanding 104 (2), 90—126.

Negoro, H., Ueda, M., Watanabe, K., Kobayashi, K., Kurihara, Y., 2011. Measurement and analysis of golf swing using 3D acceleration and gyroscopic sensors. In: Proceedings of the IEEE Conference on the Society of Instrument and Control Engineers (SICE), pp. 1111—1114.

Park, M.J., Choi, M.G., Shinagawa, Y., Shin, S.Y., 2006. Video-guided motion synthesis using example motions. ACM Transactions on Graphics-TOG 25 (4), 1327—1359.

Reddy, 2010. [Imageworld] Dataset: Action Recognition Realistic Videos-UCF 50. Available from: http://lists.diku.dk/pipermail/imageworld/2010-October/002753.html (accessed 06.10.10.).

Ruixuan, W., 2004. 3D-2D Spatiotemporal Registration for Human Motion Analysis. Available from: www.comp.nus.edu.sg/∼leowwk/thesis/wangruixuan-thesis.pdf (accessed 20.10.10.).

Sim, K.F., Sundaraj, K., 2010. Human motion tracking on broadcast golf swing video using optical flow and template matching. In: Proceedings of the IEEE Conference on Computer Applications and Industrial Electronics (ICCAIE), pp. 169—173.

Tony, T., Matsuyama, T., 2008. Human motion tracking using a color-based particle filter driven by optical flow. In: Proceedings of the 1st International Workshop on Machine Learning for Vision-based Motion Analysis, pp. 1—12.

Urtasun, R., Fleet, D.J., Fua, P., 2005. Monocular 3D tracking of the golf swing. In: Proceeding of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 932—938.

Wang, L., Hu, W., Tan, T., 2003. Recent developments in human motion analysis. Pattern Recognition 36 (3), 585—601.

Yang, C.G., Duraiswami, R., Davis, L., 2005. Fast multiple object tracking via a hierarchical particle filter. In: Proceedings of the 10th IEEE International Conference on Computer Vision, pp. 212—219.