Robust Markerless Tracking of Knee Joint for Indoor and Outdoor Cycling

Oral Kaplan[1,a)] Goshiro Yamamoto[2,b)] Takafumi Taketomi[1,c)]
Alexander Plopski[1,d)] Hirokazu Kato[1,e)]

Abstract: In this work, we demonstrate the applicability of an existing deep learning based toolbox to indoor-outdoor cycling for realizing robust markerless tracking of knee joint. Cycling has become a ubiquitous physical activity worldwide for recreation, commuting, or sport. It is a low-impact non-weight bearing form of physical activity due to body weight being carried by a bicycle. However, this does not guarantee an experience free from injuries. Besides falls, the repetitive nature of cycling and monotonous loading of joints are consistently associated with overuse injuries. Among all, knee overuse remains an ill-defined injury type with anecdotal treatment approaches. Although biomedical research consider numerous factors as originators, research efforts utilize quantitative data captured through stationary indoor scenarios alone due to technological limitations. Therefore, we consider a two-part video-based framework for cycling to enable indoor-outdoor tracking of knee movement and trajectory visualizations. In this paper, we focus on former and describe our preliminary studies on tracking to demonstrate its applicability to cycling. Furthermore, we clarify the place of our work in literature by introducing the ongoing research, and formulate several future directions that may provide new insights into knee overuse injuries. We consider our approach promising for realizing cost-efficient monitoring of knee joint during indoor-outdoor cycling.

Keywords: Knee overuse injuries; Cycling, Video-based tracking, Deep learning.

1. Introduction

Cycling has developed into a ubiquitous form of sport and exercise throughout the years. Even though studies consider cycling a low-impact non-weight bearing form of exercise due to body weight being carried by a bicycle [1], it requires a high level of bodily fitness for executing the physically demanding tasks. Cycling individuals typically follow structured training regimes with high number of repetitions to achieve and maintain an adequate fitness level. This repetitiveness and monotonous loading of joints make cycling-induced knee overuse injuries frequently observed problems regardless of the expertise level of an individual [2–4]. Consequently, computer-aided sports medicine research gained a significant momentum over the last decades for enhancing our understanding of injury predictors in cycling. Yet, despite the existing significant body of knowledge, the knee joint continues to be a common site of injury with anecdotal treatment approaches [3]. Complex syndromes such as patellar tendinitis are consistently observed between cycling individuals [2]. Lateral knee pain is a common symptom of iliotibial band friction syndrome, which is attributed to repetitive flexion and extension of lower extremities [5, 6]. Furthermore, studies frequently report pain and discomfort around knee joint; making knee the second most common overuse injury region following back pain in cycling [1, 2, 7].

Studies particularly focus on two subjects for the assessment and treatment of cycling-induced knee overuse injuries. First, short and long-term load management is substantial for optimization of training and resting periods to avoid overtraining [8, 9]. Second, optimization of the closed kinetic chain formed between the man and the machine dramatically reduces the likelihood of any overuse injury. Accordingly, scholars consider various factors such as bicycle geometry, muscle imbalances, or functional and anatomical limb length discrepancies for realizing adequate cycling biomechanics [10, 11]. However, despite its dramatic change over the course of history, predominant optimization procedures offer limited solutions by entirely focusing on indoor environments [12]. Although state of the art computer-aided systems offer evidence-based strategies with high accuracy, current approaches to movement monitoring in cycling primarily support indoor environments while neglecting the significance of outdoor movement monitoring [12, 13]. Correspondingly, the existing body of knowledge remains largely unscientific and unsupported by adequate research results in terms of outdoor cycling.

[1] Nara Institute of Science and Technology, Ikoma, Nara 630–0192, Japan
[2] Kyoto University Hospital, Kyoto City, Kyoto 606–8507, Japan
[a)] oral.kaplan.nv4@is.naist.jp
[b)] goshiro@kuhp.kyoto-u.ac.jp
[c)] taketomi-t@is.naist.jp
[d)] plopski@is.naist.jp
[e)] kato@is.naist.jp

In this work, we propose a video-based motion capture framework to address the issues above. First part of our framework considers the application of an existing deep learning toolbox to stereo videos of knee movement in cycling for realizing robust and unobtrusive indoor-outdoor tracking of knee joint. Second part of our framework connects data graphs representing tracking results to input videos for simultaneously carrying out quantitative and qualitative monitoring. In this paper, we concentrate on the first part of our framework by explaining its tracking component. Furthermore, we present our preliminary studies in detail and describe several directions for future research. We consider our work encouraging for enhancing our knowledge on factors leading to cycling-induced knee overuse injuries by providing a cost-efficient and unobtrusive situated tool for indoor-outdoor knee movement monitoring.

2. Related Work

Digitally recording human movement, or commonly referred as motion capture, is a technology that has been fostering significant development in numerous fields over the last decades. From an historical perspective, early techniques used to capture movement focused on creating lifelike hand-drawn animations [14]. Originally invented by Max Fleischer, rotoscoping is an animation technique that enabled realistic human movement in cartoons through frame-by-frame tracing of motion picture footage. Fleischer used rotoscoping to create his own animated character called Koko the Clown [15] and later Walt Disney employed the same technique in 1937 Snow White and the Seven Dwarfs [16]. Although rotoscoping is seen as a proof of concept for modern motion capture systems, Lee Harrison III was the first person to capture human movement as we know it. In 1959, he used a body suit equipped with potentiometers and a cathode ray tube to capture and animate actor movements in real-time [17,18]. Current state of the art optical systems continue to utilize a similar approach where actors wear a body suit and attach passive reflective markers to capture movement through infrared cameras. These systems offer great accuracy and relatively short setup times compared to their predecessors. On the other hand, vision-based markerless approaches such as Kinect and Apple's Motion demonstrate the future direction of this technology [19].

Although it is best known to fame by its successful use in movie industry, exercise physiology is a field that's second to none when it comes to capturing and analyzing human movement [20]. Scholars of this field consistently used motion capture systems to better understand biomechanics of human movement and its clinical applications [20]. The majority of studies rely on marker or video-based approaches for gait and posture analysis [19, 21]. Individuals with physical disabilities heavily depend on prostheses developed in accordance with motion capture technology [22]. Wearable systems based on inertial measurement units are gaining popularity due to their high level of portability [23]. In cycling, these systems are frequently used for achieving

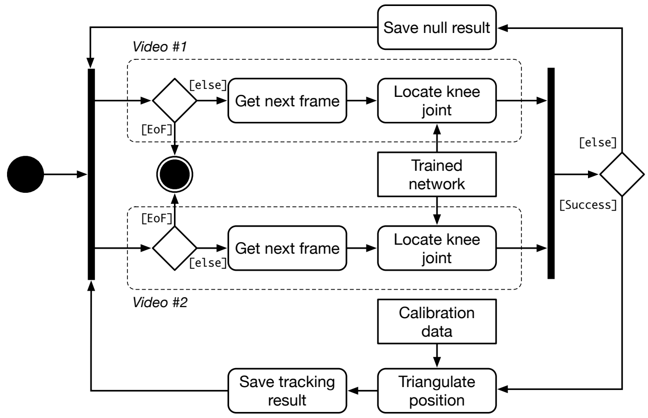


Fig. 1 Tracking process of our proposed framework.

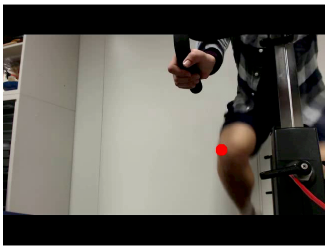
optimal bike fit and minimizing the possibility of overuse injuries [24]. Researchers track hand position to analyze its effects on pelvic motion according to gender and pedaling power [25]. As a result, it is obvious that motion capture has an important role in understanding biomechanics of human movement to eradicate overuse injuries associated with risk factors in exercise.

3. Tracking Knee Joint

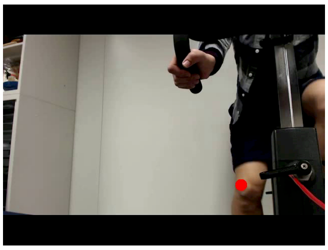
We utilized three offline stages in our framework to achieve tracking and visualization of knee movement during pedaling action. First stage included the stereo calibration of two cameras we used to track knee movement. Afterwards, we captured multiple videos via the same camera rig and used them with DeepLabCut [26] to train a network capable of unobtrusive and robust markerless tracking of knee joint in cycling. Finally, we triangulated knee joint location using the calibration information and the tracking results from the network we trained (Figure 1). We then visualize the results using common data graphs, and make use of videos to support qualitative monitoring on demand when the user clicks a point of interest on the same graphs.

3.1 Preliminary Studies

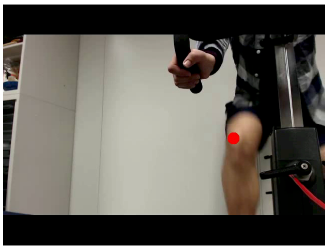
We conducted two preliminary studies to assess the applicability of our framework to cycling. We assessed the capabilities of DeepLabCut toolbox for tracking knee joint in cycling during our first study. We downloaded several indoor cycling videos that was shot through a single camera. We labeled 80% of each video to train a network and used the remaining 20% for visually confirming the tracking accuracy. The second study, however, followed the three staged process described above using several videos we captured with a calibrated stereo camera rig of our own. First, we calibrated the camera rig (Dual Logitech c920r) using standard stereo calibration functions provided in OpenCV library. Next, we captured videos of knee movement using the same rig and equally divided them into training and test groups. We used the videos from training group with DeepLabCut framework to train a network that tracks knee movement. We trained our network according to the train-



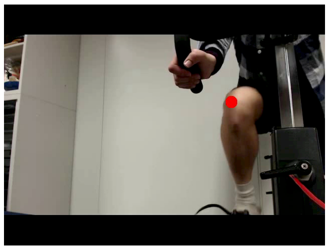
(a) Camera 1 snapshot 1



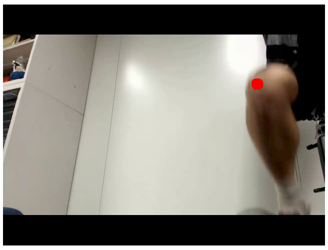
(b) Camera 1 snapshot 2



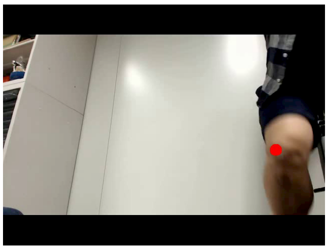
(c) Camera 1 snapshot 3



(d) Camera 1 snapshot 4

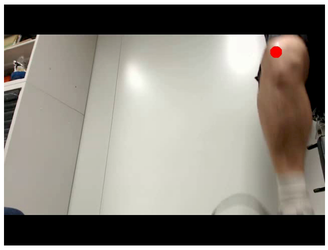


(e) Camera 2 snapshot 1



(f) Camera 2 snapshot 2

(g) Camera 2 snapshot 3



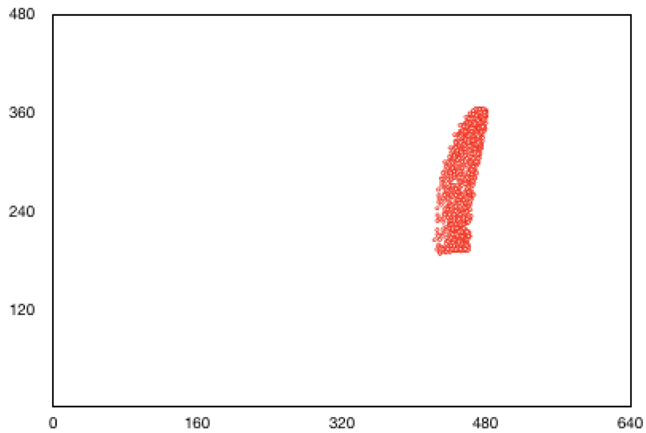
(h) Camera 2 snapshot 4

Fig. 2 Simultaneous snapshots representing detected knee joint position in tracking videos.

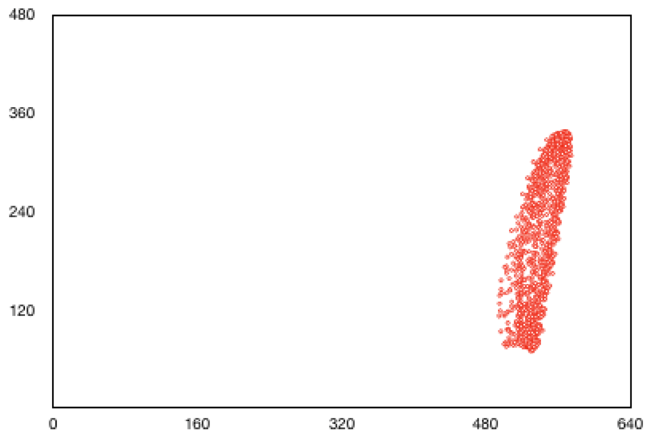
ing procedure described by Nath et al. [27]. We extracted 64 images from four 60-seconds long training videos using k-means clustering. We used these images to train our network for around 1,000,000 iterations; which takes approximately a day with Nvidia GeForce Titan X. Finally, we used this network to track knee position in test videos and combined it with stereo calibration parameters to triangulate 3D knee position. Then, we used Matplotlib with tracking results to create data graphs of knee movement.

3.2 Results

DeepLabCut demonstrated encouraging tracking performance comparable to human accuracy during both trials. Although our assessments were merely visual, currently we consider it suitable for realizing our visualization framework. We experienced no problem with detection, but observed several tracking errors during the first trial when both knees were visible. Network tracked a single knee at all times as we trained it to do so, and could not distinguish between left and right which led to a few outliers in tracking results. We believe this is due to the simplicity of labels we used to train our network, and associated problems can be easily solved by retraining it to detect both knees at the same time, or by ensuring a single knee is visible in each video feed as we did in our second study. This approach resulted in improved accuracy and a better visualization strategy which confirmed our expectations. However, although we favor the application of our visualization framework to indoor-outdoor environments, we only carried out indoor trials until the present due to hardware limitations. The placement of our stereo camera rig and a still background relatively failed to demonstrate the significance of using videos for tracking and visualization purposes. Furthermore, currently we consider naive observations from camera's perspective a mediocre ap-



(a) Pixel coordinates of knee joint in camera#1 frames



(b) Pixel coordinates of knee joint in camera#2 frames

Fig. 3 Knee joint locations detected by DeepLabCut in both camera feeds.

proach which can benefit from techniques such as structure from motion and perspective transformation.

4. Conclusion and Future Work

In this paper, we introduced the applicability of a video-based toolkit to cycling for realizing indoor-outdoor knee movement tracking. We consider our framework promising for realizing cost-effective tracking of knee joint, and promoting subjective and objective monitoring of knee movement in cycling.

We acknowledge the need for an in-depth analysis and evaluation of our framework while simultaneously considering several directions for future research. First and foremost, we are planning to modify our video capture approach by attaching dual micro cameras to a bicycle's down tube where a single knee joint is visible at all times. This indeed allows indoor-outdoor tracking of knee movement in cycling, and makes an effective comparison possible between the two for the first time. We also consider the potential of this comparison for initiating new research directions in sports science by presenting a novel perspective.

Moreover, an in-depth analysis of tracking performance is essential before placing any effort or consideration into DeepLabCut's applicability to knee tracking in cycling. We are planning to address its indoor-outdoor tracking accuracy by comparing it with an optical motion capture in a fully controlled environment. However, outdoor environments pose quite the challenge for optical systems, therefore suitable equipment, location, and adequate weather conditions must be met before any further endeavor.

Finally, we consider replacing the raw video feed approach with a model-based VR visualization of knee to reduce the possibility of information overload and introduce a more flexible visualization. This approach, however, might remove access to certain qualitative factors by discarding background information which may serve an essential role in outdoor scenarios. All these concepts require careful design and sheer attention of information and sports scientists for integrating the best of both worlds.

Acknowledgments This work was supported by the MIC/ SCOPE #162107006.

References

[1] Dannenberg, A. L., Needle, S., Mullady, D. and Kolodner, K. B.: Predictors of injury among 1638 riders in a recreational long-distance bicycle tour: Cycle Across Maryland, *The American Journal of Sports Medicine*, Vol. 24, No. 6, pp. 747–753 (1996).

[2] Callaghan, M. J.: Lower body problems and injury in cycling, *Journal of Bodywork and Movement Therapies*, Vol. 9, No. 3, pp. 226–236 (2005).

[3] Clarsen, B., Krosshaug, T. and Bahr, R.: Overuse injuries in professional road cyclists, *The American Journal of Sports Medicine*, Vol. 38, No. 12, pp. 2494–2501 (2010).

[4] Mellion, M. B.: Common cycling injuries, *Sports Medicine*, Vol. 11, No. 1, pp. 52–70 (1991).

[5] Ellis, R., Hing, W. and Reid, D.: Iliotibial band friction syndrome: A systematic review, *Manual Therapy*, Vol. 12, No. 3, pp. 200–208 (2007).

[6] Khaund, R. and Flynn, S. H.: Iliotibial band syndrome: A common source of knee pain, *American Family Physician*, Vol. 71, No. 8, pp. 1545–1550 (2005).

[7] Fukubayashi, T., Torzilli, P. A., Sherman, M. F. and Warren, R. F.: An in vitro biomechanical evaluation of anterior-posterior motion of the knee. Tibial displacement, rotation, and torque, *The Journal of Bone and Joint Surgery*, Vol. 64, No. 2, pp. 258–264 (1982).

[8] Brushøj, C., Larsen, K., Albrecht-Beste, E., Nielsen, M. B., Løye, F. and Hölmich, P.: Prevention of overuse injuries by a concurrent exercise program in subjects exposed to an increase in training load: A randomized controlled trial of 1020 army recruits, *The American Journal of Sports Medicine*, Vol. 36, No. 4, pp. 663–670 (2008).

[9] Drew, M. K. and Finch, C. F.: The relationship between training load and injury, illness and soreness: A systematic and literature review, *Sports Medicine*, Vol. 46, No. 6, pp. 861–883 (2016).

[10] Bailey, M., Maillardet, F. and Messenger, N.: Kinematics of cycling in relation to anterior knee pain and patellar tendinitis, *Journal of Sport Science*, Vol. 21, No. 8, pp. 649–657 (2003).

[11] Wilber, C. A., Holland, G. J., Madison, R. E. and Loy, S. F.: An epidemiological analysis of overuse injuries among recreational cyclists, *International Journal of Sports Medicine*, Vol. 16, No. 3, pp. 201–206 (1995).

[12] Burt, P.: *Bike Fit: Optimise your bike position for high performance and injury avoidance*, A&C Black (2014).

[13] Perin, C., Vuillemot, R., Stolper, C. D., Stasko, J. T., Wood, J. and Carpendale, S.: State of the art of sports data visualization, *Computer Graphics Forum*, Vol. 37, No. 3, Wiley Online Library, pp. 663–686 (2018).

[14] Sturman, D. J.: A brief history of motion capture for computer character animation, *Proceedings of SIGGRAPH*, ACM, p. Course 9 (1994).

[15] Fleischer, M.: Method of producing moving-picture cartoons, https://patents.google.com/patent/US1242674A/en (accessed February 6th, 2019).

[16] Pointer, R.: *The Art and Inventions of Max Fleischer: American Animation Pioneer*, McFarland (2017).

[17] Harrison, L.: Notes for early animation device, https://ohiostate.pressbooks.pub/app/uploads/sites/45/2017/09/Harrison.pdf (accessed November 1st, 2018).

[18] Schier, J.: Lee Harrison III, Animac: Hybrid graphic animation computer, http://www.vasulka.org/Kitchen/PDF_Eigenwelt/pdf/092-095.pdf (accessed November 7th, 2018).

[19] Moeslund, T. B., Hilton, A. and Krüger, V.: A survey of advances in vision-based human motion capture and analysis, *Computer Vision and Image Understanding*, Vol. 104, No. 2-3, pp. 90–126 (2006).

[20] Sigal, L., Balan, A. O. and Black, M. J.: Humaneva: Synchronized video and motion capture dataset and baseline algorithm for evaluation of articulated human motion, *International Journal of Computer Vision*, Vol. 87, pp. 4–27 (2010).

[21] Pfister, A., West, A. M., Bronner, S. and Noah, J. A.: Comparative abilities of Microsoft Kinect and Vicon 3D motion capture for gait analysis, *Journal of Medical Engineering & Technology*, Vol. 38, No. 5, pp. 274–280 (2014).

[22] Kim, J., Lee, M., Shim, H. J., Ghaffari, R., Cho, H. R., Son, D., Jung, Y. H., Soh, M., Choi, C., Jung, S. et al.: Stretchable silicon nanoribbon electronics for skin prosthesis, *Nature Communications*, Vol. 5, No. 5747, pp. 1–11 (2014).

[23] Vlasic, D., Adelsberger, R., Vannucci, G., Barnwell, J., Gross, M., Matusik, W. and Popović, J.: Practical motion capture in everyday surroundings, *Transactions on Graphics*, Vol. 26, No. 3, pp. 1–9 (2007).

[24] Burke, E. R.: Proper fit of the bicycle., *Clinics in Sports Medicine*, Vol. 13, No. 1, pp. 1–14 (1994).

[25] Sauer, J. L., Potter, J. J., Weisshaar, C. L., Ploeg, H.-L. and Thelen, D. G.: Influence of gender, power, and hand position on pelvic motion during seated cycling, *Medicine and Science in Sports and Exercise*, Vol. 39, No. 12, pp. 1–8 (2007).

[26] Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W. and Bethge, M.: DeepLabCut: Markerless pose estimation of user-defined body parts with deep learning, Technical report, Nature Publishing Group (2018).

[27] Nath, T., Mathis, A., Chen, A. C., Patel, A., Bethge, M. and Mathis, M. W.: Using DeepLabCut for 3D markerless pose estimation across species and behaviors, *bioRxiv* (2018).

