International Journal of Human Movement and Sports Sciences 10(5): 900-905, 2022 http://www.hrpub.org
DOI: 10.13189/saj.2022.100505

Markerless Motion Capture System Based on Webcams Using OpenPose

Wootaek Lim[1,2]

[1]Department of Physical Therapy, Woosong University, Daejeon, 34606, Republic of Korea
[2]Woosong Institute of Rehabilitation Science, Woosong University, Daejeon, 34606, Republic of Korea

Received June 27, 2022; Revised August 19, 2022; Accepted September 19, 2022

Cite This Paper in the Following Citation Styles

***(a):** [1] Wootaek Lim , "Markerless Motion Capture System Based on Webcams Using OpenPose," International Journal of Human Movement and Sports Sciences, Vol. 10, No. 5, pp. 900 - 905, 2022. DOI: 10.13189/saj.2022.100505.*

***(b):** Wootaek Lim (2022). Markerless Motion Capture System Based on Webcams Using OpenPose. International Journal of Human Movement and Sports Sciences, 10(5), 900 - 905. DOI: 10.13189/saj.2022.100505.*

Abstract Recently, webcam-based motion capture systems have attracted tremendous attention as alternative low-cost motion analysis tools, owing to the rapid advancement of markerless technologies. Although the upper and lower limbs are moved simultaneously in most function activities in daily life, most existing studies focused on kinematic data observed in the hip and knee joints during gait. The purpose of this study was to measure the kinematic data of the upper and lower limb simultaneously during modified squat, using low-cost webcams and to compare them to those of the 3D marker-based motion capture system. The modified squat was performed nine times in total without rest in between. The modified squat consisted of hip and knee flexion with elbow flexion. During modified squat, the kinematic data were collected using 4 RGB cameras and 13 Vicon cameras. The positions of the joints were obtained using OpenPose and the 3D positions were reconstructed using OpenCV. Angular trajectories from the Vicon motion capture system tended to be very similar to the ones from the webcam-based motion capture system. There was no significant difference in the kinematic data at the elbow between two different capture systems. Additionally, the minimum joint angles at the hip and knee were not significantly different. It was confirmed that the markerless motion capture system based on webcam can measure angular joint motions at the same level as the Vicon 3D motion capture system with markers. The webcam-based motion capture system will enhance the usability in clinical settings.

Keywords Kinematics, Motion Detection, OpenPose, OpenCV, Vicon, Webcam

---

1. Introduction

Human motion analysis is a prerequisite for evaluating movement dysfunction. Visual observation is widely used in clinical settings, but it lacks objectivity because the subjective experience of the measurers is reflected in the interpretation [1]. With technological advancement, a variety of specialized equipment currently aids in biomechanical analysis. In the technical aspect, the three-dimensional (3D) motion capture system is the most accurate and reliable technology, thus being considered as the gold-standard [2]. However, a large space is required for equipment installation, and it has to be used in an environment where light gets blocked because of the usage of infrared markers [3]. Furthermore, it requires a lengthy advance preparation process for measurement and specialized skills for analysis. Therefore, 3D motion analysis is operated by professionals in laboratory environments and is unsuitable for daily use in clinical settings. Commercially developed products, such as Kinect, are easy to use in daily life outside the laboratory environment; however, they lack accuracy. For application in patients, the validity should be improved because an under- or over-estimated joint range of motion (ROM) is observed compared to the Vicon 3D motion capture system [4–6]. Digital video analysis of human motion is also often used, which specifies two segments and calculates the joint angle between the specified segments. However, it has disadvantages: it requires a high-definition camera, and the analysis is limited to single plane movements [7,8].

Furthermore, if there is no marker, errors may occur in joint tracking [9]. Recent advances in markerless technology suggest the possibility of webcam-based motion capture systems [10,11].

Webcam is inexpensive and easy for operation, resulting in excellent versatility and usability. Expensive equipment such as Vicon is difficult to use in common clinics that account for the absolute majority in the rehabilitation medical system because of the price burden. Because measurement and analysis using webcam-based motion captures are simple, a therapeutic program can be created on-site based on the measurement results. In the early days, webcam-based motion captures required markers for the accurate analysis of human body movements. In addition to direct attachment of sphere-shaped objects as markers on the body, certain parts that have significant contrasts to the background may be specified for tracking [12,13]. Because marker-based systems have spatial constraints, a markerless system is essential for use in natural settings. Currently, it is possible to capture markerless 3D motions if the OpenPose and OpenCV is used on images obtained through two or more webcams [14,15]. However, despite its high usability, there is a lack of research on markerless 3D motion capture systems compared to those on other motion capture systems. Many existing studies have separated the upper and lower limbs to analyze their movements, mostly focusing on lower limbs [12,13,16–18]. Because the upper and lower limbs are moved simultaneously in most function activities in daily life, unlike in the laboratory, it is important to measure and analyze them without separation.

The purpose of this study was to measure the single joint motion of the upper limb and multi-joint motion of the lower limb simultaneously, using highly versatile low-cost webcams and to compare it to the 3D marker-based motion analysis used as a gold-standard.

2. Materials and Methods

2.1. Subjects

A healthy adult (age 29 years, height 173 cm, weight 73 kg) voluntarily participated and informed consent was obtained before the experiment. The Institutional Review Board of Woosong University approved this study.

2.2. Instrumentation

2.2.1. Vicon 3D Motion Capture System

Thirteen infrared cameras (6 Vantage5, 1 MX-T160, 5 MX-T40S, 1 MX-T40, Vicon Motion Systems, Oxford, UK) and the sampling frequency of 100 Hz were used for the study. Twelve retro-reflective markers (diameter 14 mm) were attached to only the right side of the human body, and this study followed the ISB recommendation for marker placement [19,20]. A low-pass filter (3 Hz) was used for eliminating high frequency noise. The software used to process the data was Vicon Nexus 2.9.

2.2.2. Webcam-Based Motion Capture System

Four RGB cameras (IRC70, i-rocks, Gwacheon-si, Republic of Korea) were connected to the computer, with a high-end GPU (RTX2080Ti, NVIDIA, CA, USA), via USB cables. The motion data collected at a frame rate of approximately 25 Hz was subsequently interpolated to match the Vicon's frame rate of 100 Hz. The positions of the major joints (neck, shoulder, elbow, wrist, hip, knee, and ankle) of the subject were obtained using OpenPose, a software for pose estimation from 2D images [21]. The 3D positions were reconstructed using OpenCV from the joint position information analyzed in four simultaneously captured 2D images [15]. To eliminate noise, the data was filtered using a low-pass filter with a cutoff frequency of 3 Hz. To compensate the time offset between the two measurement systems, motion data collected by the webcam-based motion capture system (4DEYE, SYM Healthcare, Seoul, Republic of Korea) was automatically aligned to the reference motion data collected by the Vicon motion capture system to achieve the best correlation between the two data.

2.3. Procedures

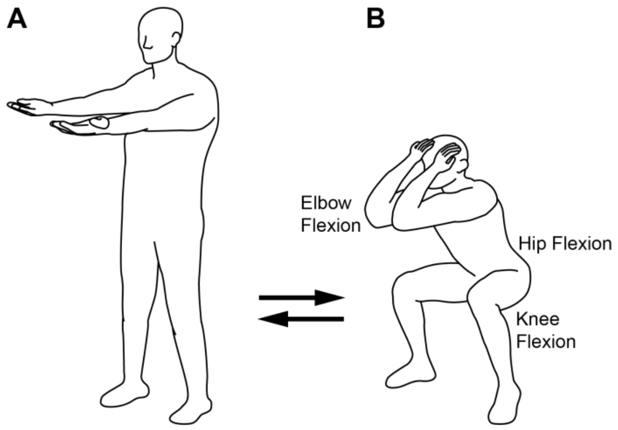

Figure 1. Modified squat including elbow flexion-extension

Before starting, both feet were spread shoulder-width apart, with both feet facing slightly outward (Figure 1A). The lower limbs maintained straightness in the hips and knees. Both arms were raised to shoulder level, with palms supinating toward the sky. With the start signal, the subject performed a modified squat. The modified squat consisted of hip and knee flexion with elbow flexion (Figure 1B). Hip and knee flexion were performed to the maximum possible range, and elbow flexion was performed until a soft end-feel was felt. After complete flexion, the subject performed the hip and knee extension by getting up at the same speed as sitting, and simultaneously, performed the elbow extension. The modified squat was performed nine times in total without rest in between.

2.4. Data Analysis

Paired T-test was used to compare the average of the angles at the elbow, hip, and knee joints obtained by the Vicon 3D motion capture system and webcam-based motion capture system. The Bland-Altman agreement analysis was conducted to determine the mean bias and limits of agreement between the two motion analysis systems. Statistical analyses were performed using SPSS Statistics 25 (IBM Corp., Armonk, NY, USA). The statistical significance levels were set at 0.05. The values were reported as the mean ± standard deviation. The figures were created using GraphPad Prism 8.01 (GraphPad Software Inc, La Jolla, CA, US).

3. Results

Angular trajectories from the Vicon 3D motion capture system tended to be very similar to the ones from the webcam-based motion capture system (Figure 2). The maximum (peak flexion) and minimum (peak extension) of joint angles were observed at almost the same time on the upper and lower limbs.

There was no significant difference in the maximum and minimum of joint angles and ROM at the elbow between two different capture systems (Table 1). Additionally, the minimum joint angles at the hip and knee were not significantly different. However, the maximum joint angles and ROM were significantly lower at Vicon motion capture system at the hip and knee. Bland-Altman analyses revealed that the mean bias was 7.610 and the standard deviation (SD) of bias was 6.972 (Figure 3). Limits of agreement ranged from -6.055 to 21.270.

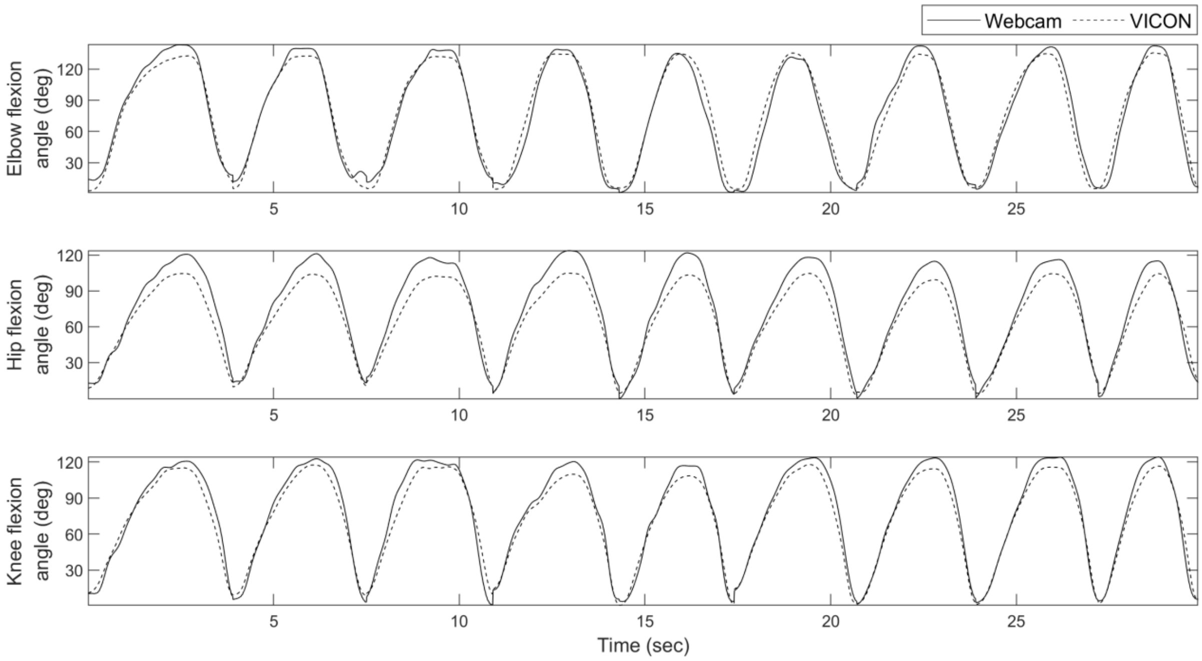

Figure 2. Comparison of sagittal plane joint angles for right elbow, hip, and knee

Table 1. Elbow, hip, and knee angles during flexion and extension of joints

		Elbow		**Hip**		**Knee**	
		Angle(°)	***p***	**Angle(°)**	***p***	**Angle(°)**	***p***
Max	Vicon	134.1±1.3	0.011	103.6±1.8	<0.001	114.5±3.3	<0.001
	Webcam	138.9±3.8		118.5±3.2		121.2±2.7	
Min	Vicon	4.5±0.7	0.085	4.7±2.4	0.710	3.4±3.1	0.424
	Webcam	7.7±4.7		4.3±4.4		4.2±2.1	
ROM	Vicon	129.6±1.5	0.253	98.8±2.6	<0.001	111.1±3.1	<0.001
	Webcam	131.2±4.6		114.2±5.6		116.9±4.2	

Max: Maximum of joint angle, Min: minimum of joint angle, ROM: range of motion, defined as the difference between the maximum and the minimum angle

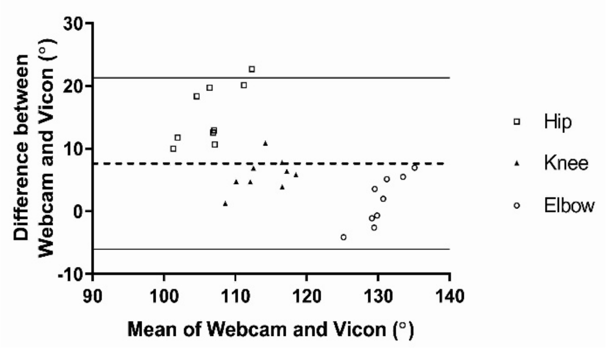

Figure 3. Bland-Altman plots comparing two different systems

4. Discussion

Most existing studies focused on and analyzed kinematic data observed in the hip and knee joints during gait. However, while the lower limbs move in an activity of daily living, the upper limbs move simultaneously and it is, therefore, important to analyze them together [22,23].

In this study, kinematic data of the upper and lower limb was analyzed after measuring the sagittal plane joint angles in real-time while the modified squat (elbow flexion during hip and knee flexion) was performed. Motions in the frontal and transverse planes are allowed limitedly during gait, whereas those in the sagittal plane are allowed in a wider range [24,25]. Because motions in the sagittal plane are limited after body damage in most cases, clinical analysis on angular trajectories in the sagittal plane is crucial for constructing a rehabilitation program. In the current study, kinematic data at the elbow, hip, and knee were obtained using the webcam-based markerless motion capture system and compared to those obtained using the Vicon motion capture system, which is considered as a gold standard.

During the nine times of the modified squat, the peak and valley of the angular trajectories measured at the elbow, hip, and knee by the two systems were found to be all quite similar. More specifically, the minimum of joint angles showed no significant difference at every joint. Comparing Vicon and Kinect, which is widely used as an alternative low-cost motion capture system, it was often observed that the measured trajectories did not match or showed a large difference between the two systems during joint flexion-extension [4,26,27]. In contrast, the webcam-based motion analysis using OpenPose and OpenCV did not show much difference in the Vicon 3D motion analysis. OpenPose allows obtaining the center of an individual joint and vectors between two separate joints through the synchronized images acquired from multi-webcams [14,21]. According to a study by D'Antonio, it is possible to obtain kinematic data of the lower limbs during gait if OpenPose is used; furthermore, it is found that not only angular kinematic data but also a significant part of the angular path is consistent in the comparison with inertial measurement units (IMUs) [10,16]. Webcams are hardly affected by spatial constraints, unlike Vicon, which requires infrared markers for motion tracking. Furthermore, they can be used in a variety of clinical settings because no professional knowledge is required to operate the devices. In the case of recently released high-end webcams, some products have a performance of 120 or 240 frames per second, facilitating the use in sports fields [28,29].

Angular trajectories during the modified squat were consistent overall, but some parts were different in detailed kinematic data between the two systems depending on the joint. In the case of the elbow, there was no significant difference in the maximum and minimum of the joint angle and ROM. In the case of the lower limb, however, there was a significant difference in the maximum of the joint angle at the hip and knee, and in the case of the hip, especially, the difference was larger. Because a significant difference occurred in the maximum of the joint angle in the state where there was no significant difference in the minimum joint angle, there was a significant difference in ROM. In previous studies, it was concluded that a full ROM of hip and knee was not observed in motion tracking using OpenPose because the peak flexion obtained by the webcam was quite lower than that obtained by the IMUs [4,16]. However, the peak flexion angle observed using OpenPose in this study was larger than that measured by Vicon. Because the validity and reliability are both rated higher in Vicon than IMUs, it can be inferred that OpenPose can technically measure full ROM. A slight difference in the kinematic data between Vicon and webcams is probably due to the difference in the trunk anatomical coordinate system's estimation method and the longitudinal axis' estimation method between the two systems. Because the hip flexion angle is calculated as the angle between the trunk and thigh, it is important to accurately estimate the coordinates of the trunk as the reference point while the thigh is moving. In the Vicon system, two markers are attached to the ventral part (CLAV and STRN) and two to the dorsal part (C7 and T8) [30]. The medio-lateral axis, supero-inferior axis, and antero-posterior axis are estimated from the four markers. In contrast, the webcam-based system confirms the positions of the neck, left and right shoulders, and left and right hip joints in the frontal plane and then uses the trunk anatomical coordinate system that has each joint center's position as a reference point. In other words, the reference position of the trunk may vary depending on the coordinate estimation method. This is not an error, but rather a difference based on the estimation method. Second, in the case of Vicon, a single rigid-body segment is recognized based on the markers attached to the skin, and the values measured according to the skin curvature may not be parallel to the actual longitudinal axis of segment. This is a systematic error and will continuously have an effect during measurement. In summary, in the case of the hip joint angle-the angle between the trunk and long bone, the difference between the two systems may be relatively larger compared to the case of the elbow and knee joints. Despite some differences, the measured angular trajectories are highly consistent between the two systems, and because the deviation is not large, it can be determined that there will be no problem with the clinical application.

This study has several limitations including a small sample size. Most functional motions are accompanied by positional movements, and in this study the experiments were conducted with the subject fixed in one place. When the subject passes in front of the webcam, the distance between the subject and webcam decreases gradually and then increases gradually. Here, the plane captured on the camera changes from the front to the back [31]. For accurate measurement of the angular motion of the segment, a previous study utilized OpenPose also restricts positional movements by having gait performed on a treadmill [16]. In the future, studies would be conducted in an environment where positional movements are free.

5. Conclusions

The technological advancement of software and hardware for markerless motion capture enables low-cost human motion capture. In this study, it was confirmed that if OpenPose and OpenCV are used, the webcam-based motion capture system-despite being markerless-can measure angular joint motions at the same level as the Vicon 3D motion capture system with markers. Because webcams are portable and consume low power, the webcam-based motion capture system will enhance the usability in natural settings, such as clinical and sports sites, beyond artificial spaces such as laboratories [32].

Acknowledgments

This research was supported by 2021 Woosong University Academic Research Funding.

REFERENCES

[1] Hazari, A.G. Maiya, T.V. Nagda. Kinematics and Kinetics of Gait, in: A. Hazari, A.G. Maiya, T.V. Nagda (Eds.), Conceptual Biomechanics and Kinesiology, Springer, Singapore, 2021: pp. 181–196.

[2] F. Roggio, S. Ravalli, G. Maugeri, A. Bianco, A. Palma, M. Di Rosa, G. Musumeci. Technological advancements in the analysis of human motion and posture management through digital devices, World J Orthop, 12, 467–484, 2021.

[3] S.R. Simon. Quantification of human motion: gait analysis-benefits and limitations to its application to clinical problems, J Biomech, 37, 1869–1880, 2004.

[4] A. Pfister, A.M. West, S. Bronner, J.A. Noah. Comparative abilities of Microsoft Kinect and Vicon 3D motion capture for gait analysis, J Med Eng Technol, 38, 274–280, 2014.

[5] Bilesan, S. Behzadipour, T. Tsujita, S. Komizunai, A. Konno. Markerless Human Motion Tracking Using Microsoft Kinect SDK and Inverse Kinematics, in: 2019 12th Asian Control Conference (ASCC), 2019: pp. 504–509.

[6] F. Schlagenhauf, S. Sreeram, W. Singhose. Comparison of Kinect and Vicon Motion Capture of Upper-Body Joint Angle Tracking, in: 2018 IEEE 14th International Conference on Control and Automation (ICCA), 2018: pp. 674–679.

[7] Z. Lu, G. Nazari, J.C. MacDermid, S. Modarresi, S. Killip. Measurement Properties of a 2-Dimensional Movement Analysis System: A Systematic Review and Meta-analysis, Arch Phys Med Rehabil, 101, 1603–1627, 2020.

[8] L. Khadilkar, J.C. MacDermid, K.E. Sinden, T.R. Jenkyn, T.B. Birmingham, G.S. Athwal. An analysis of functional shoulder movements during task performance using Dartfish movement analysis software, Int J Shoulder Surg, 8, 1–9, 2014.

[9] M. Eltoukhy, S. Asfour, C. Thompson, L. Latta. Evaluation of the performance of digital video analysis of human motion: Dartfish tracking system, Int J Sci Eng Res, 3, 1–6, 2012.

[10] E. D'Antonio, J. Taborri, E. Palermo, S. Rossi, F. Patanè. A markerless system for gait analysis based on OpenPose library, in: 2020 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), 2020: pp. 1–6.

[11] Y. Rybarczyk. 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, Springer International Publishing, Cham, 2016: pp. 731–738.

[12] A. Pusara, S. Heamawatanachai, K. Sinsurin, C. Jorrakate. Reliability of a low-cost webcam recording system for three-dimensional lower limb gait analysis, Int Biomech, 6, 85–92, 2019.

[13] C. Krishnan, E.P. Washabaugh, Y. Seetharaman. A low cost real-time motion tracking approach using webcam technology, J Biomech, 48, 544–548, 2015.

[14] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, Y. Sheikh. OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields, IEEE Trans Pattern Anal Mach Intell, 43, 172–186, 2021.

[15] G. Bradski, A. Kaehler. OpenCV, Dr Dobb's Journal of Software Tools, 3, 2000.

[16] E. D'Antonio, J. Taborri, I. Mileti, S. Rossi, F. Patané. Validation of a 3D Markerless System for Gait Analysis Based on OpenPose and Two RGB Webcams, IEEE Sensors Journal, 21, 17064–17075, 2021.

[17] T. Zult, J. Allsop, J. Tabernero, S. Pardhan. A low-cost 2-D video system can accurately and reliably assess adaptive gait kinematics in healthy and low vision subjects, Sci Rep, 9, 18385, 2019.

[18] F. Wang, E. Stone, M. Skubic, J.M. Keller, C. Abbott, M.

Rantz. Toward a passive low-cost in-home gait assessment system for older adults, IEEE J Biomed Health Inform, 17, 346–355, 2013.

[19] G. Wu, F.C. Van der Helm, H.D. Veeger, M. Makhsous, P. Van Roy, C. Anglin, J. Nagels, A.R. Karduna, K. McQuade, X. Wang. ISB recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion—Part II: shoulder, elbow, wrist and hand, Journal of Biomechanics, 38, 981–992, 2005.

[20] G. Wu, S. Siegler, P. Allard, C. Kirtley, A. Leardini, D. Rosenbaum, M. Whittle, D.D. D'Lima, L. Cristofolini, H. Witte, O. Schmid, I. Stokes, Standardization and Terminology Committee of the International Society of Biomechanics. ISB recommendation on definitions of joint coordinate system of various joints for the reporting of human joint motion--part I: ankle, hip, and spine. International Society of Biomechanics, J Biomech, 35, 543–548, 2002.

[21] Z. Cao, T. Simon, S.-E. Wei, Y. Sheikh. Realtime multi-person 2d pose estimation using part affinity fields, in: 2015 IEEE Conference on Computer Vision and Pattern Recognition, 2017: pp. 7291–7299.

[22] Wootaek Lim , "Tensile Force Transmission from the Upper Trunk to the Contralateral Lower Leg throughout the Posterior Oblique Sling System," International Journal of Human Movement and Sports Sciences, Vol. 9, No. 2, pp. 294 - 300, 2021. DOI: 10.13189/saj.2021.090217.

[23] A. Sidiropoulos, R. Magill, A. Gordon. Coordination of the upper and lower extremities during walking in children with cerebral palsy, Gait Posture, 86, 251–255, 2021.

[24] J.M. Rodda, H.K. Graham, L. Carson, M.P. Galea, R. Wolfe. Sagittal gait patterns in spastic diplegia, J Bone Joint Surg Br, 86, 251–258, 2004.

[25] A. Michelini, A. Eshraghi, J. Andrysek. Two-dimensional video gait analysis: A systematic review of reliability, validity, and best practice considerations, Prosthet Orthot Int, 44, 245–262, 2020.

[26] D. Lim, C. Kim, H. Jung, D. Jung, K.J. Chun. Use of the Microsoft Kinect system to characterize balance ability during balance training, Clin Interv Aging, 10, 1077–1083, 2015.

[27] T.M. Guess, S. Razu, A. Jahandar, M. Skubic, Z. Huo. Comparison of 3D Joint Angles Measured with the Kinect 2.0 Skeletal Tracker Versus a Marker-Based Motion Capture System, J Appl Biomech, 33, 176–181, 2017.

[28] G. Worah, A. Khan, M. Kothari, M. Naik. Monitor eye-care system using blink detection—a convolutional neural net approach, Int J Eng Res Technol, 6, 12–16, 2017.

[29] Pueo. High speed cameras for motion analysis in sports science, Journal of Human Sport and Exercise, 11, 53–73, 2016.

[30] N.P. Brouwer, T. Yeung, M.F. Bobbert, T.F. Besier. 3D trunk orientation measured using inertial measurement units during anatomical and dynamic sports motions, Scand J Med Sci Sports, 31, 358–370, 2021.

[31] R. Baker, H.M. Hart. Measuring walking: a handbook of clinical gait analysis, Mac Keith Press London, 2013.

[32] A. Badiola-Bengoa, A. Mendez-Zorrilla. A Systematic Review of the Application of Camera-Based Human Pose Estimation in the Field of Sport and Physical Exercise, Sensors (Basel), 21, 5996, 2021.