Assessment of Joint Range of Motion Measured by a Stereo Camera

Yuta Ono, Oky Dicky Ardiansyah Prima, Takashi Imabuchi, Yoshitoshi Murata, Hisayoshi Ito
Graduate School of Software and Information Science
Iwate Prefectural University
Takizawa, Japan
e-mail: g231q005@s.iwate-pu.ac.jp, prima@iwate-pu.ac.jp, g236o001@s.iwate-pu.ac.jp, {y-murata, hito}@iwate-pu.ac.jp

Yukihide Nishimura
Department of Rehabilitation Medicine
Iwate Medical University
Morioka, Japan
e-mail: ynishi@iwate-med.ac.jp

***Abstract*—Many studies have been conducted to measure joint Range of Motion (ROM) to approach Activities of Daily Living (ADL) assessments using nonintrusive three-dimensional (3D) sensors, such as motion capture devices and the Kinect, as alternatives to goniometer. While these sensors have been widely used to measure joint angles, the measuring range is restrictive and the complexity to setup has prevented these devices to be used as self-measurements during home-based training. With the recent progress on human pose estimation using computer vision approach, locations of joints can be detected from a vision camera in real time. This achievement opens a possibility to measure ROMs in a wide area and in any locations as long as the subject is located inside the camera's Field of View (FOV). This study extends the current human pose estimation to capture joints in 3D using a stereo camera and compares the ROMs of shoulder and elbow derived from the proposed method with those derived from the Kinect. Measurements of ROM using both devices show good agreement indicating that the proposed method is valid to measure ROMs, as a replacement for the Kinect.**

Keywords-rehabilitation; computer vision; range of motion; activities of daily living; 3D human pose estimation.

I. INTRODUCTION

Many studies have reported the close relationship between joint Range of Motion (ROM) and Activities of Daily Living (ADL) [1][2]. The loss of ROM may occur at all ages due to injuries, diseases, surgery and normal aging, giving a direct effect on posture and movement. For those who have impaired ROM, the activity needs to be performed by using compensatory strategies [3]. ROM is commonly evaluated as the degree of maximum range of motion. However, the individual difference is big, such as a patient with impaired shoulder flexion motion may not be able to raise his upper limb but may still be able to conduct most ADL tasks.

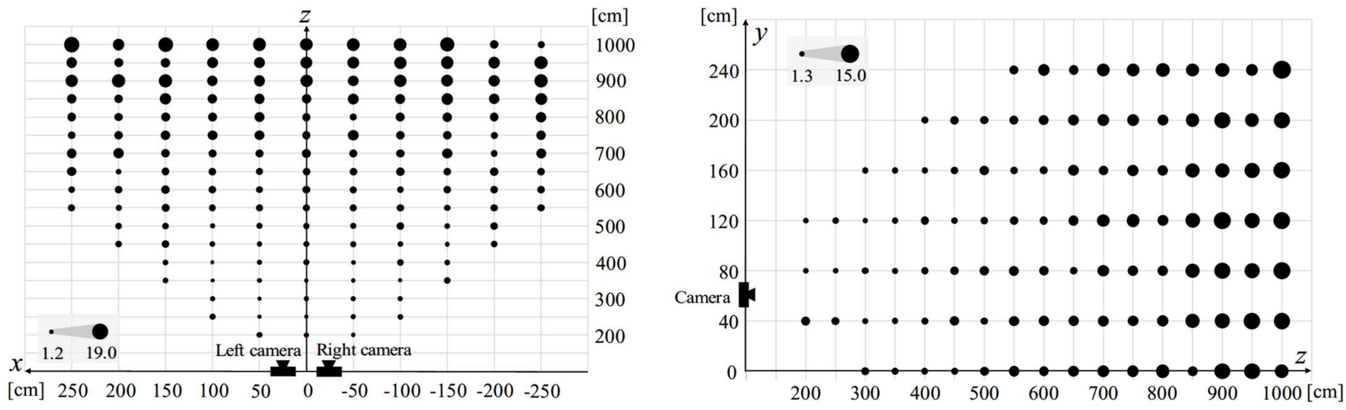
Traditional methods to measure ROM use apparatus, such as goniometers, inclinometers, and video in several specified directions. To obtain reliable measurements, clinicians are suggested to take repeated ROM measurements. Since the universal goniometer has scale in 5° increments, the measurement fluctuation is usually expected up to ±5°. The use of motion capture devices to measure angles has increased the reliabilities of ROM measurements. With these devices, ROM can be measured while the ADL tasks were performed by a subject.

Some portable and low-cost devices equipped with depth sensors, such Microsoft Kinect and Intel RealSense can be used to track human motions and capture human postures in 3D. Previous studies have shown that these devices show performance adequate for a range of healthcare imaging applications [4][5]. Unlike motion capture systems, ROM measurements using these depth sensors suffer from occlusion problems. Therefore, these sensors have to be positioned in an appropriate location and direction to avoid these problems. However, the tracking accuracy depends on having a perfect 3D model of the subject which requires pose estimation using multiple cameras [6].

The state-of-the-art computer vision techniques have enabled the detection of 2D human limb joints using a single camera. These techniques utilize fine-tuned convolutional network architectures. The "Stacked Hourglass Network (SHN)" was known to have a robust performance to a variety of challenges on joint detections for multiple people [7]. The most recent technique, "OpenPose" uses part affinity fields for the fast detection of multiple people [8]. SHN and OpenPose are available online as open source software for research purposes. Using dual cameras mounted side-by-side (a stereo camera), 3D human limb joints can be calculated by triangulating the corresponding 2D joints detected in each side camera [9]. The fast growing in Virtual Reality (VR) has made stereo cameras available on the markets. Some smartphones have been equipped by stereo cameras to produce perspective effects to the photos.

This study attempts to expand the application of stereo vision as an easy and low-cost tool to measure ROM on performing ADL and to promote basic self-care. Two kinds of ROM are measured using 3D joints detected from a stereo camera and the Kinect, respectively. The measurements of ROM were performed according to the methods and guidelines for the measurement of joint range of motion by the Japanese Association of Rehabilitation Medicine and the Japanese Orthopaedic Association (JARM & JOA) [10]. ROM from each source is analyzed to reveal whether stereo vision is adequate for practical use on quantifying ROM. This study provides a basic framework to build a ROM measurement system. The framework minimalizes the hardware requirement since the measurement of 3D joint is performed at a server.

This paper is organized as follows. Section II describes the related work on human joint detections using the Kinect and



(a) Horizontal errors (X-Z plane)

(b) Vertical errors (Y-Z plane)

Figure 1. Distribution of errors measured at 1,004 control points used in this study.

cameras. Section III describes methods to measure ROM using a stereo camera and the Kinect based on guidelines of JARM and JOA. Section IV shows measurement results of the accuracy of ROM derived from a stereo camera approach. Finally, Section V concludes the achievements and discusses the future prospective of this study.

II. RELATED WORK

The task of estimating the posture of the human body without using markers has been attracting attention in the research area of computer vision. The estimation methods can be broadly classified as either methods using depth or visible cameras.

A low-cost depth camera, the second version of Microsoft Kinect, the Kinect V2, has been used to measure ROM. Since its predecessor, the Kinect has been applied to the rehabilitation field. The Kinect V2 captures depth images between the range of 0.5~8.0m in 30 frame per second of speed. The Software Development Kit (SDK) enables developers to detect 25 joints of up to 6 persons concurrently. Hereinafter, we simply referred both versions as the Kinect. The skeletal joints of the Kinect can be considered as an adequate tool for supporting rehabilitation. However, using these data in clinical applications where precise angle measurements are required needs a significant concern [11].

Studies have been conducted to enable the detection of various joints in complex postures using a camera. Toshev et al. (2014) [12] detects 2D joints using a regression model with a cascade Deep Neural Network (DNN) and associates corresponding joints across the body posture. Newell et al. (2016) [7] proposed SHN to improve the detection by handling a diverse and challenging set of poses with a simple mechanism for reevaluation and assessment of initial predictions. Both [7] and [12] require a person detection process as a preprocessing before detecting body joints, causing detection of joints cannot be done if the preprocessing failed to detect a person. Cao et al. (2017) [8] proposed the OpenPose that creates “Part Confidence Maps (PCM)” to

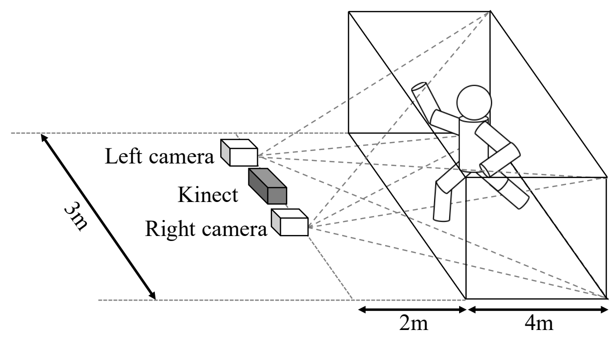


Figure 2. The experimental setup in this study.

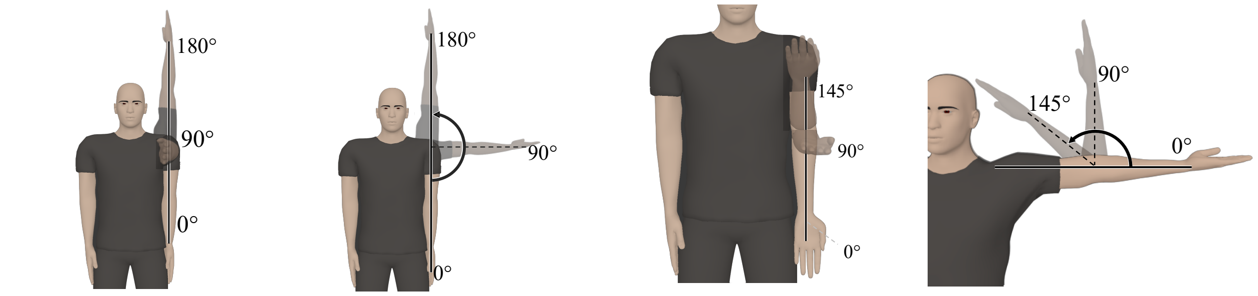
detect joints and “Part Affinity Fields” to associate corresponding joints directly without detecting a person beforehand. OpenPose is capable to detect every joint of the human body in real time.

Ohno et al. (2018) [9] applied OpenPose to a stereo camera and constructed 3D joints from 2D joints detected from each camera image using a stereo vision approach. Triangulation is used to find the optimal 3D joint by refining the location of joints from each image. The refining process uses the value of PCM to select a more reliable joint between two images and determine the corresponding joint from the counterpart image using template matching.

3D joint measurements based on stereo vision seem to be more promising on measuring ROM. It will be possible for patients to create self-reported ROM at home as long as dual cameras are available, and to send the report to clinicians to assess the ability to engage in ADL tasks.

III. METHODS

This study evaluates two kinds of ROM measured by a stereo camera and the Kinect, respectively. Stereo camera and the Kinect are connected to a single computer to capture



(a) Shoulder (Forward flexion) (b) Shoulder (Abduction) (c) Elbow (Frontal flexion) (d) Elbow (Side flexion)

Figure 3. Body postures measured in this study.

TABLE I. JOINTS MEASURED IN THIS STUDY

Joint	Pose	Incremental Measurement	Max. ROM
Shoulder	Forward flexion	30°, 60°, 90°, 120°, 150°	180°
	Abduction	30°, 60°, 90°, 120°, 150°	180°
Elbow	Frontal flexion	30°, 60°, 90°, 120°	145°
	Side flexion	30°, 60°, 90°, 120°	145°

the movement of the subject concurrently. The capturing speed is set to 30fps.

A. 3D Joint Measurement Using a Stereo Camera

The method of Ohno et al. (2018) [9] is adapted to measure 3D joints. Two cameras were calibrated and placed at intervals of 60cm. Each camera has a resolution of 1280px×720px. Errors were measured at 1,004 control points regularly placed in an area of 5m×10m. Here, Figure 1 shows distribution of errors on each control points. For all points, Root Mean Square Error (RMSE) was measured 8.12cm. For points located up to 6m against the cameras, the RMSE was measured 5.07cm. To get optimal results, subjects will be positioned between 2m and 6m from the center line between two cameras. This distance is considerably sufficient to measure ROM while engaging in ADL tasks.

B. 3D Joint Measurement Using Kinect

The Kinect is mounted on the middle between the stereo camera described above. Skeleton tracking function provided by the Kinect for Windows SDK 2.0 is used to measure 3D joints. No calibration is made on the resulted 3D joints. Here, Figure 2 shows the experimental setup for this study.

C. Data Extraction

With the stereo camera and the Kinect in frontal view, ROM for shoulder and elbow are measured to perform some ADL tasks. The measurements of shoulder and elbow, as shown in Figure 3, are conducted for subjects standing with upright postures for ROM measurements in this study. Values of angles between minimum and maximum angles on performing each posture are measured to be used as necessary data for the ADL [2]. Table I shows joints measured in this study whereas the maximum ROM for each joint movement was determined as in [11]. Here, we perform two types of measurements: incremental and continuous. The incremental measurement measures the absolute accuracy for a given posture. On the other hand, the continuous measurement will reveal how stable the values of angles are measured. Two healthy subjects (A and B) were participated in this study. Both subjects were required to wear ordinary clothes and to stand at 4m from the Kinect.

1) Incremental measurement: Subjects were asked to take each posture as shown in Figure 3, where angles for each pose is assigned as in Table I and set using a goniometer. Goniometric measurements were performed using standardized methods [13]. After the goniometer was aligned to the shoulder motion by the examiner, a second examiner read and recorded the measurement. Once the pose was set, measurements were recorded simultaneously using the Kinect and the stereo camera. For the maximum ROM, subjects were instructed to move each joint to their maximum capability. These procedures were repeated three times for both left and right joints. The agreement of two measurements against a goniometer were assessed by studying the mean bias and constructing Limits of Agreement (LOA) to determine validity [10][14]. Here, the 95% LOA were defined as the mean bias to ±1.96 Standard Deviation (SD). Since the goniometer has scale in 5° increments, the 95% LOA for the discrepancy exceeded ±5° can be defined as clinically significant.

2) Continuous measurement: During continuous measurement, subjects were asked to stand upright and slowly perform shoulder abduction and elbow frontal flexion as shown in Figure 3(b)-(c) and move back to the upright pose. During this movement, measurements were recorded using the Kinect and the stereo camera, concurrently. The resulted measurements were fitted separately using fourth-order polynomial regression models to investigate the stability of each measurement by each device. RMSE and R-squared (R^2) were calculated to evaluate the model. RMSE indicates the

TABLE II. MEAN BIAS AND LOA MEASUREMENTS OF SHOULDER AND ELBOW JOINT ANGLES OBTAINED USING THE KINECT AND THE STEREO CAMERA AGAINST A GONIOMETER

Subject	Joint	Pose	Kinect vs goniometer		Stereo camera vs goniometer	
			Mean bias	95% LOA	Mean bias	95% LOA
A	Shoulder	Forward flexion	3.39°	-16.62° to 23.39°	-1.65°	-26.74° to 23.44°
		Abduction	-2.82°	-32.41° to 26.78°	-3.90°	-17.51° to 9.71°
	Elbow	Frontal flexion	10.64°	-4.11° to 25.40°	9.91°	-14.29° to 34.11°
		Side flexion	-20.01°	-58.55° to 18.53°	-3.51°	-18.96° to 11.94°
B	Shoulder	Forward flexion	-10.95°	-75.50° to 53.59°	-0.22°	-20.80° to 20.35°
		Abduction	-4.96°	-17.93° to 8.01°	-6.29°	-20.88° to 8.31°
	Elbow	Frontal flexion	1.97°	-23.05° to 27.00°	1.27°	-26.28° to 28.82°
		Side flexion	-17.85°	-31.35° to -4.16°	-4.74°	-20.13° to 10.64°

TABLE III. MODEL FITTING AND ERROR ESTIMATION

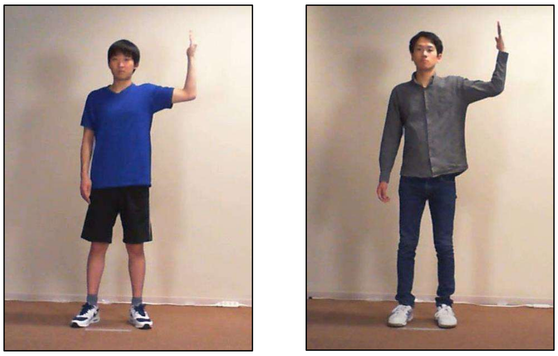
Subject	Joint	Pose	RMSE (R^2)	
			Kinect	Stereo camera
A	Left shoulder	Abduction	8.86° (0.9704)	8.59° (0.9708)
	Right shoulder		7.14° (0.9084)	8.0° (0.9735)
	Left elbow	Frontal flexion	7.13° (0.9801)	9.06° (0.9672)
	Right elbow		6.85° (0.9784)	6.66° (0.9767)
B	Left shoulder	Abduction	2.78° (0.9971)	3.13° (0.9960)
	Right shoulder		3.81° (0.9947)	5.01° (0.9913)
	Left elbow	Frontal flexion	8.54° (0.9759)	4.79° (0.9886)
	Right elbow		5.13° (0.9917)	4.66° (0.9938)

absolute fit of the model to the data whereas R^2 is a relative measure of fit.

IV. RESULTS

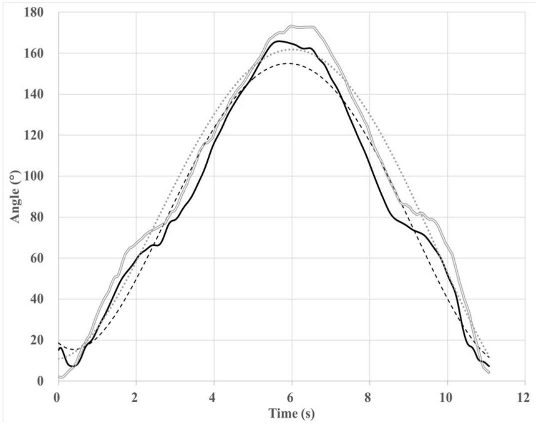
1) Incremental measurement: Joints with poses shown in Table I were measured for left and right shoulders and elbows from each subject. Table II shows the mean bias and the 95% LOA of measurements of the Kinect and the stereo camera against those of the goniometer. The analysis results derived from each user are presented because the difference of the clothes is considerably affecting the measurement results. Subject A and B, as shown in Figure 4, wore short and long-sleeved shirts, respectively. The mean bias indicates that the stereo camera had relatively better accuracies than the Kinect. Especially, the measurement of poses from frontal view (forward flexion of shoulder and frontal flexion of elbow) where those joints were partly occluded from the devices. However, the 95% LOA for the discrepancy of both devices against the goniometer exceeded ±5°, which was defined as clinically significant. For the Kinect, this finding is consistent with [10].

2) Continuous measurement: Continuous movement during shoulder abduction and elbow frontal flexion were measured from each subject. Subjects were requested to perform these tasks within 20 seconds. No start or stop signs were given to the subjects because the necessary data can be extracted manually from the recorded scenes. Table III shows RMSE and R^2 for each model fitted to the measurement results from the Kinect and the stereo camera. RMSE values from data derived by the stereo camera and the Kinect are comparable. High values of R^2 were achieved from both model fitting. Here, Figure 5 and 6 visualize the measurement data and models fitted to the data.

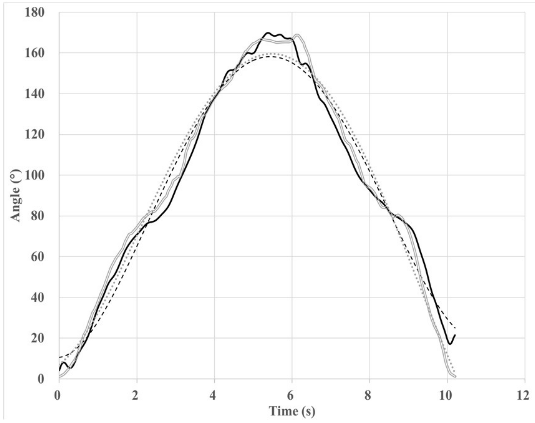


(a) Subject A (b) Subject B

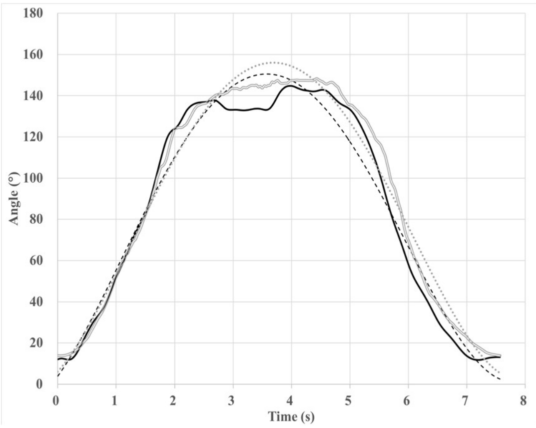
Figure 4. The looks of subjects in this study.



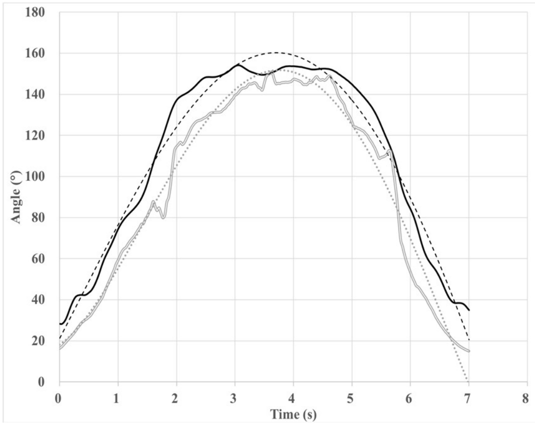
(a) Left shoulder (abduction)



(b) Right shoulder (abduction)



(c) Left elbow (frontal flexion)



(d) Right elbow (frontal flexion)

—Stereo camera - - -Stereo camera (fitted curve) —Kinect ·····Kinect (fitted curve)

Figure 5. Continuous measurement for abduction and frontal flexion (subject A).

V. CONCLUSION AND FUTURE WORK

In this study, we measured 3D joints remotely using a stereo camera and the Kinect. Our experiments show that the stereo camera had relatively better accuracies than the Kinect on measuring a pose at a given angle. On the other hand, the stereo camera was observed to be as stable as the Kinect on measuring joint angles continuously. However, before the stereo camera, as well as the Kinect can be used to measure ROM, it is important to understand their limitations in accuracy for the measurement of specific joint motions against a goniometer.

The stereo camera used in this study is superior to the Kinect because it was based on OpenPose that doesn't require the detection of full body posture to detect particular joints. Whereas, the Kinect may fail to detect joints when other part of the body is occluded. The stereo camera will enable us to observe various ADL tasks, such as dressing, eating, and bathing where a part of the body may be hidden easily.

ACKNOWLEDGMENT

We acknowledge the effort from the authors of OpenPose to make ROM measurements using a stereo camera possible. This work was supported by the MIC/SCOPE #181602007.

REFERENCES

[1] A. M. Oosterwijk, M. K. Nieuwenhuis, C. P. van der Schans, and L. J. Mouton, "Shoulder and elbow range of motion for the performance of activities of daily living: A systematic review," Physiotherapy Theory and Practice, vol. 34, no. 7, pp. 505-528, 2018.

[2] A. M. Oosterwijk, M. K. Nieuwenhuis, H. J. Schouten, C. P. van der Schans, and L. J. Mouton, "Rating scales for shoulder and elbow range of motion impairment: Call for a functional approach," PLOS ONE, vol. 13, no. 8, pp. 1-13, 2018, https://doi.org/10.1371/journal.pone.0200710.

[3] B. P. Pereira, A. Thambyah A, and T. Lee, "Limited forearm motion compensated by thoracohumeral kinematics when performing tasks requiring pronation and supination," Journal of Applied Biomechanics, vol.28, pp. 127–138, 2012.

[4] S. Aleesandro, C. Andrea, M. Matteo, and M. T. Lorenzo, “Kinect V2 Performance Assessment in Daily-Life Gestures: Cohort Study on Healthy Subjects for a Reference Database for Automated Instrumental Evaluations on Neurological Patients,” Applied Bionics and Biometrics, pp. 1-16, 2018, https://doi.org/10.1155/2017/8567084.

[5] S. H. Lee et al. , “Measurement of Shoulder Range of Motion in Patients with Adhesive Capsulitis Using a Kinect,” PLOS ONE vol. 10, no. 6, pp. 1-12, 2015, doi:10.1371/journal.pone.0129398.

[6] R. R. P. Kumar, S. Muknahallipatna, J. McInroy, M. McKenna, and L. Franc, “Real-time Range of Motion Measurement of Physical Therapeutic Exercise,” Journal of Computer and Communications, vol. 5, pp. 19-42, 2017.

[7] A. Newell, K. Yang, and J. Deng, “Stacked Hourglass Networks for Human Pose Estimation,” Computer Vision – ECCV 2016, pp. 483–499, Amsterdam, Netherlands, 2016.

[8] Z. Cao, T. Simon, S.E. Wei, and Y. Sheikh, “Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields,” Computer Vision and Pattern Recognition 2017, pp. 7291-7299, 2017.

[9] Y. Ohno, O. D. A. Prima, and H. Ito, “3D Human Pose Estimation for Motion Analysis,” The 80th Nation Convention of Information Processing Society of Japan, pp. 263-264, 2018. (in Japanese)

[10] M. E. Huber, A. L. Seitz, M. Leeser, and D. Sternad, “Validity and Reliability of Kinect Skeleton for Measuring Shoulder Joint Angles: a Feasibility Study,” Physiotherapy, vol. 101, no. 4, pp. 389–393, 2015.

[11] K. Yonemoto, S. Ishigami, and T. Kondo, “The Method Guidelines for Range of Motion Measurement,” The Japanese Journal of Rehabilitation Medicine, vol. 32, no. 4, pp. 207–217, 1995. (in Japanese)

[12] A. Toshev and C. Szegedy, “DeepPose: Human Pose Estimation via Deep Neural Networks,” Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference, pp. 1653–1660, 2014.

[13] N. B. Jain, R. B. Wilcox, J. N. Katz, and L. D. Higgins, “Clinical Examination of the Rotator Cuff,” Physical Medicine and Rehabilitation, vol. 5, pp. 45–56, 2013.

[14] J. M. Bland, and D. G. Altman, “Statistical Method for Assessing Agreement between Two Methods of Clinical Measurement,” Lancet, vol.327, pp. 307-310, http://dx.doi.org/10.1016/S0140-6736(86)90837-8, 1986.

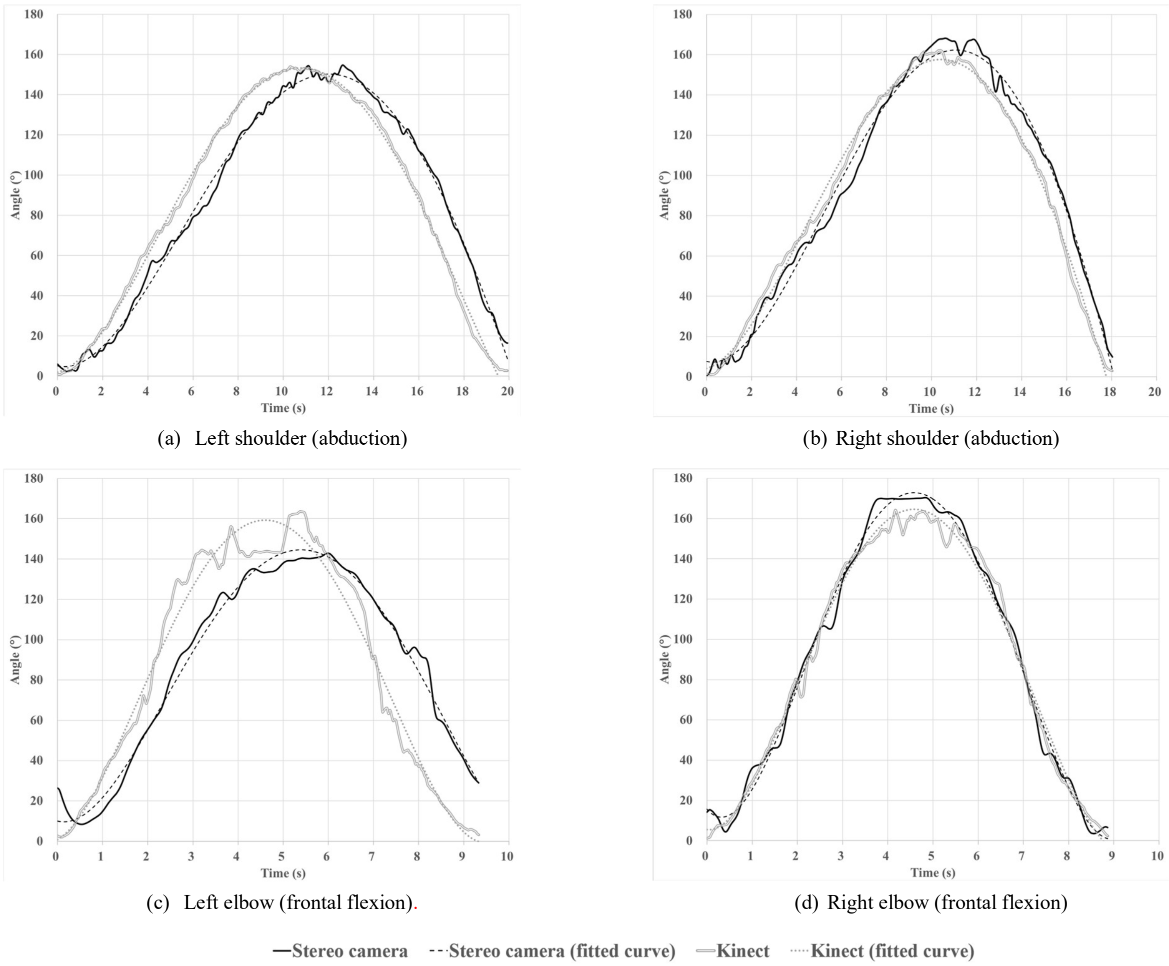


Figure 6. Continuous measurement for abduction and frontal flexion (subject B).