Evaluation of Joint Range of Motion Measured by Vision Cameras

Oky Dicky Ardiansyah Prima, Yuta Ono, Yoshitoshi Murata, Hisayoshi Ito
Graduate School of Software and Information Science, Iwate Pref. Univ.
Takizawa, Japan
prima@iwate-pu.ac.jp, g231q005@s.iwate-pu.ac.jp, {y-murata, hito}@iwate-pu.ac.jp

Takashi Imabuchi
Office of Regional Collaboration, Iwate Pref. Univ.
Takizawa, Japan
t_ima@ipu-office.iwate-pu.ac.jp

Yukihide Nishimura
Dept. of Rehabilitation Medicine
Iwate Medical University
Morioka, Japan
ynishi@iwate-med.ac.jp

Abstract— **Joint Range of Motion (ROM) can be measured through a variety of methods including the use of sophisticated devices such as goniometers and non-intrusive three-dimensional (3D) sensor devices such as motion capture systems. The Microsoft Kinect has been proposed as an affordable motion capture device as an alternative to goniometers. However, due to limited measurement range and complex setup, this device cannot be used as a self-measurement during home rehabilitation or flexibility training. With the recent progress in human pose estimation based on computer vision approaches, it has become possible to estimate human joint positions in real time from vision cameras. This study evaluates joint ROM measured by two vision cameras using 3D human pose estimation based on a single camera and a stereo camera. The ROM of major joints, which consist of shoulders, elbows, a hip, and knees was evaluated for 10 users. The stereo camera gives the best results with a small bias to the goniometer compared to the single camera and the Kinect. Vision cameras have advantages on estimating semi-occluded joint locations than the Kinect. The 3D human pose based on a single camera opens up possibilities to build Tele-Rehabilitation (TR).**

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

I. INTRODUCTION

Human movement is dependent on the amount of range of motion (ROM), the amount of motion available at a synovial joint. This movement is unique to each joint and is dependent upon the shape of the articular surfaces of bones and the integrity and flexibility of the periarticular soft tissues. ROM can be measured as either active and passive. While the former is measured by the person contracting the muscles around the joint, the latter by an external force pushing on the body around the joint. Passive motion can either limit or perform full joint ROM. This study is an extension of our previous work on the assessment of ROM from body joints estimated by vision cameras [1].

There are close relationships between joint ROM and Activities of Daily Living (ADL) [2][3]. 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. Although loss of ROM may not be associated with complete loss of function, people who have impaired ROM need to perform their activities by using compensatory strategies [4]. For example, 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.

Joint ROM can be assessed through a variety of methods including goniometers, inclinometers, photographs, and Motion Capture (MoCap) systems. The double-armed goniometer is the most common device to use, whereas ROM is measured at the end of its full range of movement. To obtain reliable measurements, clinicians are suggested to take repeated measurements. Since the universal goniometer has scale in 5° increments, the measurement fluctuation is usually expected up to ±5°. On the one hand, goniometers may introduce error into the measurement because the positions of the bones and axis points must be estimated, on the other hand, inclinometers are easier because no such alignment is required. Inclinometers have dials that indicate the angle at which the inclinometer is located with respect to the line of gravity. Photographs can be used to measure a certain joint ROM. Since most smartphones currently available are equipped with cameras, the use of smartphones as non-intrusive ROM assessment tools is increasing. DrGoniometer, a photo-based iPhone app, potentially offers an easy tool of ROM measurements [5][6]. It also has an ability storage of all related information to build up historical data for each movement for further analysis. MoCap systems can be categorized into marker-based and markerless system. Marker-based MoCap systems, such as Vicon, can accurately capture human movement. Vicon is often regarded as a standard in motion capture. These systems utilize multiple vision cameras to detect the light reflected by the marker and calculate the three-dimensional (3D) position using triangulation. Markerless MoCap systems use a depth sensor to measure the 3D position of the target within range of the sensor. Microsoft Kinect is widely used as an inexpensive markerless MoCap system that can track human movement and posture in three dimensions. Accuracy assessment of the Kinect against Vicon show that the Kinect is sensitive enough to be used as a portable MoCap system for workplace ergonomic assessments [7]. Other studies have shown that the Kinect performs well for a range of healthcare imaging applications [8][9].

Due to the lack of specialized medical institutions that provide rehabilitation services, there is a growing need for simple methods of ROM self-measurement. Marker-based MoCap systems are superior on handling occlusion with multiple vision cameras fixed in multiple directions to capture

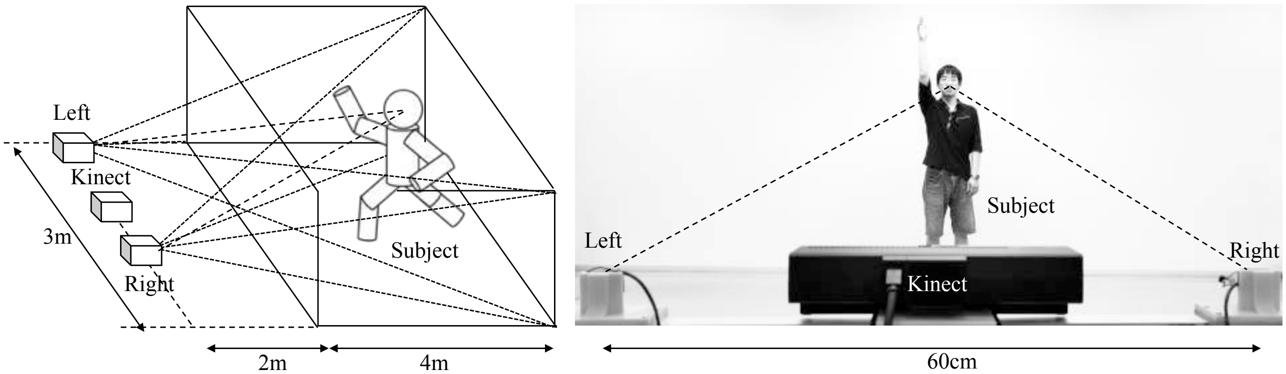


Figure 1. The experimental setup in this study.

the target. However, these systems can retrieve data only in a limited area. In addition, these systems are mostly used indoors because sunlight interferes with infrared cameras used for measurement [10]. These drawbacks also exist with markerless MoCap systems. The complexity to setup MoCap systems has prevented these systems to be used as self-measurements during home-based rehabilitation or flexibility training.

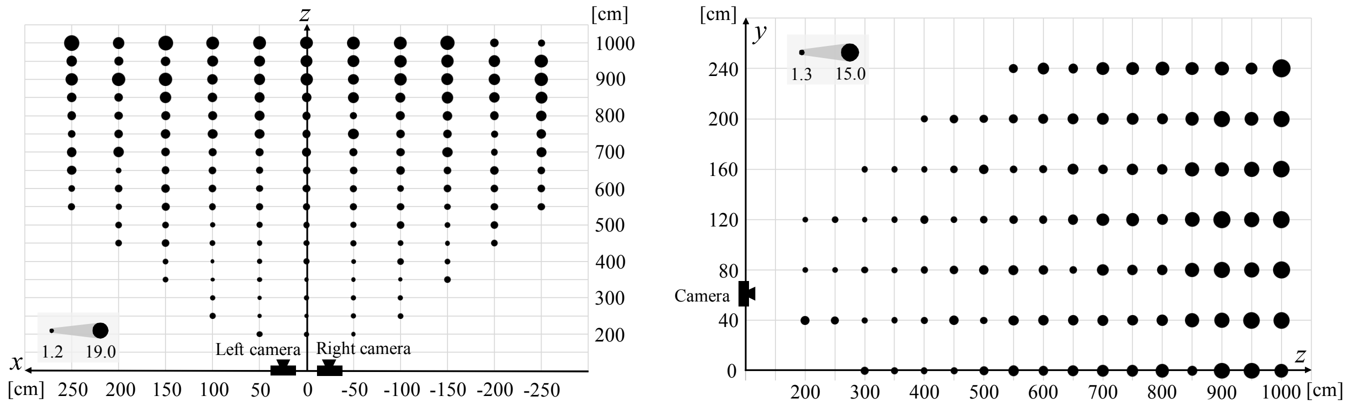
Various rehabilitation programs require a system that is capable to measure joint ROM indoors or outdoors. The basic ADL includes a functional mobility to move from one place to another while performing tasks, such as walking, getting in and out of bed, and getting into and out of a chair. This study seeks to expand the use of vision cameras: a single camera and a stereo camera, as a simple and low-cost tool to promote basic self-care by measuring joint ROM during ADL, home rehabilitation, or flexibility training. Joint ROM is measured according to the method and guidelines for joint range of motion measurement by the Japanese Association of Rehabilitation Medicine and the Japanese Orthopedic Association (JARM & JOA) [11]. ROM of the major joints, consisting of the shoulders, elbows, hips and knees is measured using the vision cameras. Measurement results are compared with the Kinect's to determine whether the vision camera is suitable for practical use of joint ROM quantification.

This paper is organized as follows. Section II describes related work on detection of 3D body joints using vision cameras. Section III describes methods to measure joint ROM using vision cameras and the Kinect based on JARM and JOA guidelines. Section IV evaluates accuracies of the resulted joint ROM obtained from each modality. Finally, Section V summarizes the results and describes the future prospects of this study.

II. Related Work

The task of estimating human pose without using a marker is attracting attention in the field of computer vision research. Many studies have been conducted to enable the use of cameras to detect various joints with complex postures. Toshev et al. (2014) use a regression model using cascade Deep Neural Networks (DNN) to detect 2D joints and associates the corresponding joints throughout the body posture [12]. Newell et al. (2016) proposed Stacked Hourglass Network (SHN) to improve detection by processing a diverse and challenging series of poses using a simple mechanism for initial prediction evaluation [13]. SHN was known to have robust performance against various challenges related to joint detection of multiple people. Both [12] and [13] require a human detection process as a pre-processing to detect joints in the body, and if the pre-processing fails to detect humans, joint detection cannot be performed. The latest method, "OpenPose", uses Part Confidence Maps (PCM) to detect joints and Part Affinity Field (PAF) to associate corresponding joints directly without the human detection process in advance [14]. SHN and OpenPose are available online as open source software for research purposes. Ono et al. (2018) applied OpenPose to the stereo camera and estimated the 3D joint from the corresponding 2D joints using a stereo vision approach [15]. 3D joint measurement based on stereo vision seems promising. As long as stereo cameras are available, patients can create a self-report ROM at home and send reports to the clinician to assess their ability to engage in ADL tasks. Measurement results show that this approach is effective in measuring joint ROM as an alternative to the Kinect. Since the stereo camera only captures visible spectrum, 3D measurements can be performed indoors and outdoors in relatively bright light conditions [1]. The rapid growth of Virtual Reality (VR) has made stereo cameras widely available in the market, making it easier to implement.

Along with the breakthrough in 2D human pose estimation, studies on 3D human pose estimation from a single camera have made significant progress. This estimation requires two steps: joint position estimation in 2D image coordinates and 3D coordinate estimation of each corresponding 2D joint. Many studies have investigated the problem of inferring 3D joints from 2D projections. These studies include traditional 2D to 3D methods that define bone length and estimate 3D joints using binary decision trees [16], or deep net based to estimate 3D joints with DNN. Martinez et al. (2017) proposed a relatively simple deep feedforward DNN [17] via Human3.6M, the largest 3D human pose dataset that includes 3.6 million human poses and corresponding images to estimate 3D joints from 2D projections [18]. Unlike the



(a) Horizontal errors (X-Z plane)

(b) Vertical errors (Y-Z plane)

Figure 2. Distribution of errors measured at ground truth points used in this study.

current MoCap system, this technique is capable of estimating body joints in semi-occluded image regions. Hence, the proposed system is much better at handling occlusion than depth sensor-based systems, such as the Kinect. This work has been made available as open source software, namely "3D-pose-baseline [19]."

Both 2D and 3D human pose estimation described above require a dedicated Graphics Processing Unit (GPU) and result in higher costs to implement. However, this calculation can be efficiently performed using GPU-accelerated cloud services over the Internet. Using services over the Internet, clinicians can measure joint ROM during physical rehabilitation with patients at home. Prima et al. (2019) proposed an IoT-based Tele-Rehabilitation (TR) framework that uses a single camera to observe the joints of a client's body in 3D when performing an ADL task [20]. Measurement results show that although the Kinect gives better results in terms of absolute accuracy, the proposed framework is less sensitive to noise than the Kinect. Further expansion of this application will enable ROM measurement in aquatic therapy and fitness pools.

III. Methods

This study measures the ROM of the major body joints, which consist of shoulders, elbows, hip, and knee joints, using three modalities: a single camera, a stereo camera, and the Microsoft Kinect V2 (hereafter, simply called the Kinect). The resulting joint ROM obtained from each device is verified using the ROM measured with a conventional goniometer. Goniometric measurements were performed in a standardized way [21]. Figure 1 shows the experimental setup for this study. Two cameras at 60cm intervals are stereo-calibrated on images with resolution of 1280×720 pixels. The Kinect is set on the middle of these cameras. The subject performs various joint movements at a distance of 4m from the center line between the two cameras. Data processing is synchronized at 30 frame per second (fps).

A. 3D Joint Measurement Using a Single Camera

The left side camera (Figure 1) is used to capture the subject's movement. For each frame, positions of joints are estimated based on PCM calculated using the OpenPose library [22]. Here, the resulting joint structure is rearranged to match the structure used in SHN. 3D coordinates corresponding to these joints are estimated using the 3D-pose-baseline. A weighted is applied to reduce noise in the resulted 3D joints.

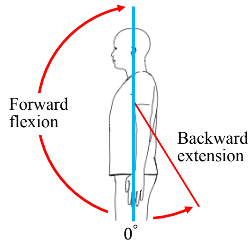
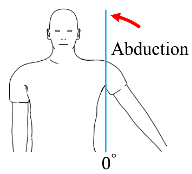
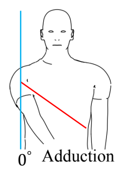
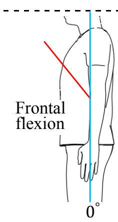
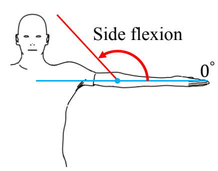
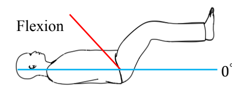
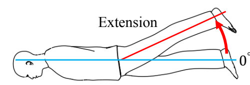
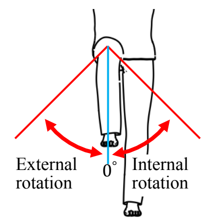
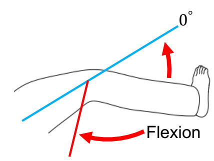
B. 3D Joint Measurement Using a Stereo Camera

The method of Ono et al. (2019) is used to measure 3D joints using a stereo camera [1]. To improve measurement accuracies, the stereo camera is re-calibrated using ground truth points. A total 1,004 ground truth points were regularly placed in an area covered by the stereo camera. This area accounts for 5m×10m. After the re-calibration process, Root Mean Square Error (RMSE) of the ground truth points was measured 8.12cm. This RMSE is considerably acceptable for our study. In the location where the subject performs joint movements (4m from the center line between the two cameras), RMSE was measured 5.07cm. Error distributions are shown in Figure 2.

C. 3D Joint Measurement Using Kinect

The Microsoft Software Development Kit (SDK) for the Kinect is used to access 3D body joints from data taken from the Kinect's sensor. Here, the resulting 3D joint is not calibrated using ground truth points. The temporal synchronization of the captured data between the Kinect and other cameras was performed using Network Time Protocol (NTP).

TABLE I. ROM MEASUREMENTS OF THE MAJOR BODY JOINTS IN THIS STUDY.

Joint	Motion	ROM	Posture
Shoulder	Forward flexion	180°	
	Backward extension	50°	
	Abduction	180°	
	Adduction	75°	
Elbow	Frontal flexion	145°	
	Side flexion	145°	
Hip	Flexion	125°	
	Extension	15°	
	Internal rotation	45°	
	External rotation	45°	
Knee	Flexion	130°	

D. Data Extraction

ROM measurements of the major body joints are shown as in Table I. Measurements are performed according to JARM and JOA guidelines. For the shoulder joints, ROM of four movements: forward flexion, backward extension, abduction, and adduction are measured. In this measurement, the torso is fixed to the wall so that the spine does not bend back and forth while moving. For the elbow joints, ROM of two movements: frontal flexion and side flexion are measured where the forearm is in the supination position. ROM measurements of hip joints include flexion, extension, internal rotation, and external rotation. The subject lies firmly on his back on a flat surface during hip flexion, while the subject lies firmly in an anatomical position during hip extension. However, in this study, ROM measurements for the internal and external rotation is performed with the subject standing and with the back fixed to the wall because these movement cannot be observed from either vision cameras or the Kinect. The knee joint flexion is performed with the hip joint in flexion. For each movement, angle values between the minimum and maximum angles are measured. The maximum angle does not represent a precise full ROM because external forces such as partner stretching are not involved during the measurement.

Joint angle measurements from data obtained with three modalities: a single camera, a stereo camera and the Kinects, were performed as the relative angle between the longitudinal axis of two adjacent segments. As an example, for elbow joint angles, the adjacent segments are the upper arm and the forearm. Whereas, for knee joint angles, the adjacent segments are the upper and the lower legs.

In this study, two types of measurements are performed: static and continuous. Static measurement measures the

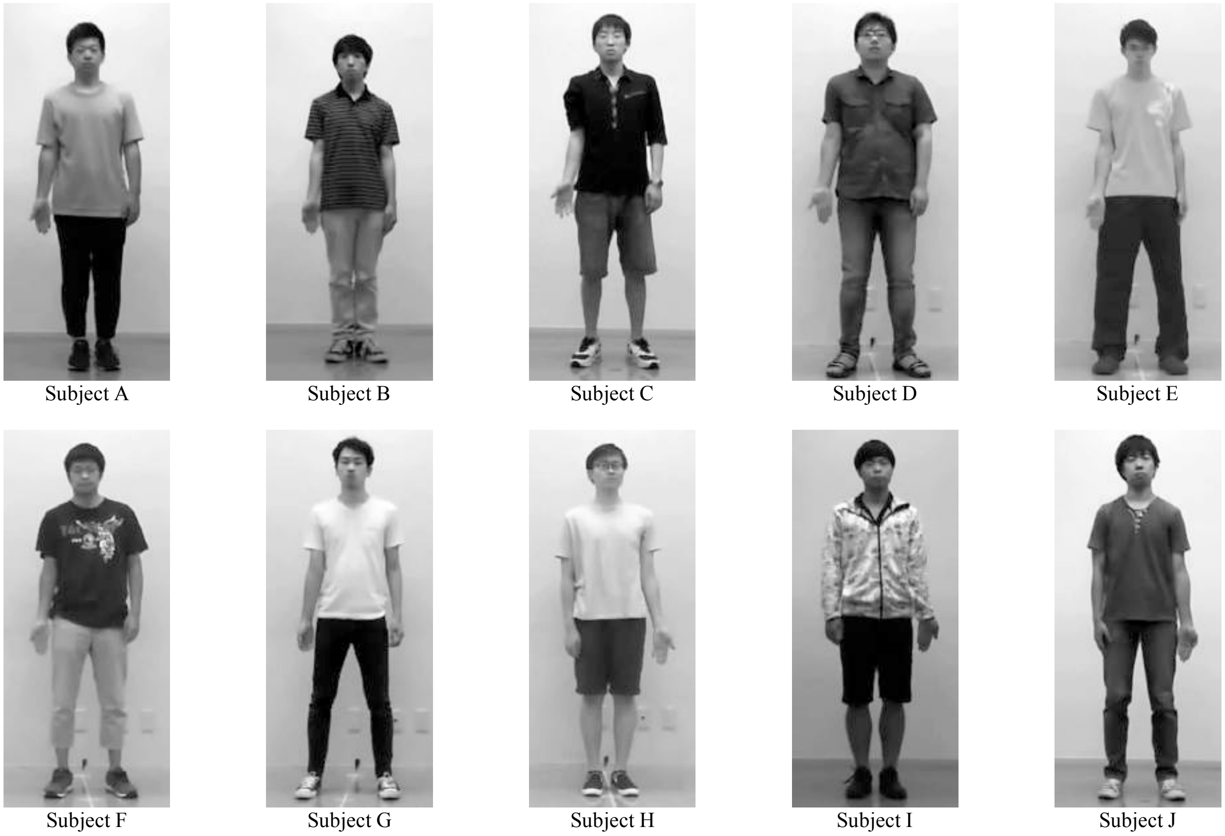


Figure 3. 10 Subjects participated in the experiment.

absolute accuracy of a particular posture. Continuous measurement will reveal how stable the values of angles are measured. Ten healthy male subjects (mean age 21.8 ± 0.92 years) were recruited for the experiments (Figure 3). All participants agreed to participate and signed the consent forms, to allow their data to be used in publications of this research. All subjects were instructed to wear normal clothing to analyze how clothes affects the 3D joints measured by the three modalities. Single cameras and stereo cameras are considered to have a greater impact than the Kinect.

a. Static measurement

Subjects were asked to perform each movement for both left and right joints as shown in Table I. By the time the subject performed a maximum ROM, the examiner aligned the goniometer to the bone and axis position, and another examiner recorded the reading. Subjects were instructed to move each joint to its maximum ability to obtain maximum ROM. The agreement of measurements from the three modalities against the goniometer were evaluate by studying the mean bias and constructing Limits of Agreement (LOA) to determine validity [23][24]. Here, the 95% LOA were defined as the mean bias to ±1.96 Standard Deviation (SD).

b. Continuous measurement

Measurements using three modalities were conducted for 15s from the start position, the position where the goniometer is aligned at 0°, to the end position, the position where full ROM is achieved. During the measurement, a progress bar is displayed on the monitor, so the subject can adjust the movement speed. The resulting measurements were individually fitted using a 4th order polynomial regression model to investigate the stability of each measurement by each device. The RMSE, which is an absolute fit to the model data, was calculated to evaluate the model.

IV. Evaluation of the Resulted Joint ROM Measured by Three Modalities

Motion data was collected for 10 subjects performing 11 motions (Table I). Each motion was performed at the joints of the left and right bodies. Therefore, each user generated 22 motion data for the experiment. For statistical analysis, the goniometer measured the maximum ROM for each motion.

a. Static measurement

Table II shows the mean bias and the 95% LOA of the maximum ROM measurements of joint angles of shoulders, elbows, hip, and knees obtained from the three modalities versus the measurements from the goniometer. Due to 5° scale in goniometer, the measurement fluctuation is usually expected up to ±5°. However, aligning the goniometer correctly with its axis on the joint axis is a difficult task. Here, we considered that a mean bias between ±10° is acceptable for the resulting measurements obtained from the three

TABLE II. MEAN BIAS AND LOA THE MAXIMUM ROM MEASUREMENTS OF JOINT ANGLES FOR SHOULDERS, ELBOWS, HIP, AND KNEE JOINTS ACQUIRED USING THREE MODALITIES AGAINST THE GONIOMETER.

Side	Joint	Motion	Single camera		Stereo camera		Kinect	
			Mean bias	95% LOA	Mean bias	95% LOA	Mean bias	95% LOA
Left	Shoulder	Forward flexion	-25.55°	-36.78 to -14.32°	**3.05°**	-8.55 to 14.66°	**2.79°**	-8.33 to 13.90°
		Backward extension	12.88°	-1.20 to 26.96°	**0.75°**	-22.72 to 24.23°	**7.62°**	-5.75 to 20.98°
		Abduction	-30.65°	-45.45 to -15.85°	**-3.28°**	-14.17 to 7.60°	**-3.78°**	-18.60 to 11.05°
		Adduction	**7.23°**	-7.23 to 21.69°	13.43°	-2.92 to 29.79°	**8.92°**	-5.40 to 23.25°
	Elbow	Frontal flexion	**-5.37°**	-19.98 to 9.24°	16.44°	-6.28 to 39.17°	11.90°	-0.20 to 24.00°
		Side flexion	-29.25°	-47.40 to -11.09°	**-7.02°**	-17.06 to 3.01°	-23.49°	-31.60 to -15.38°
	Hip	Flexion$^{*)}$	-46.38°	-67.90 to -24.87°	**1.06°**	-7.13 to 9.26°	19.32°	-79.26 to 117.90°
		Extension$^{*)}$	**-2.25°**	-24.95 to 20.44°	15.32°	-19.42 to 50.06°	78.76°	-63.94 to 221.46°
		Internal rotation	**4.68°**	-14.35 to 23.70°	17.32°	-14.20 to 48.84°	**-3.73°**	-39.02 to 31.55°
		External rotation	**-2.17°**	-15.68 to 11.34°	11.54°	-4.56 to 27.63°	**4.58°**	-6.61 to 15.78°
	Knee	Flexion$^{*)}$	-57.85°	-81.82 to -33.88°	**-2.35°**	-21.02 to 16.32°	-34.45°	-152.53 to 83.64°
Right	Shoulder	Forward flexion	-27.29°	-44.35 to -10.23°	**5.15°**	-9.23 to 19.53°	**5.04°**	-3.52 to 13.59°
		Backward extension	**3.10°**	-10.28 to 16.48°	**-4.58°**	-15.82 to 6.66°	**2.95°**	-3.36 to 9.25°
		Abduction	-35.02°	-49.37 to -20.66°	**-0.28°**	-12.01 to 11.46°	**-5.27°**	-20.31 to 9.76°
		Adduction	**-0.08°**	-16.38 to 16.22°	10.25°	-4.03 to 24.53°	10.45°	-2.84 to 23.73°
	Elbow	Frontal flexion	**1.04°**	-13.12 to 15.20°	16.81°	-1.89 to 35.51°	16.63°	3.86 to 29.40°
		Side flexion	-23.32°	-36.58 to -10.06°	**-4.79°**	-18.97 to 9.38°	-15.03°	-31.85 to 1.80°
	Hip	Flexion$^{*)}$	-26.22°	-56.36 to 3.92°	**-1.85°**	-15.08 to 11.38°	-22.00°	-132.40 to 88.39°
		Extension$^{*)}$	**7.34°**	-8.09 to 22.77°	12.86°	-12.36 to 38.08°	**8.88°**	-19.24 to 37.01°
		Internal rotation	15.30°	-18.70 to 49.30°	14.69°	-18.38 to 47.75°	**-4.34°**	-35.17 to 26.48°
		External rotation	22.86°	-6.37 to 52.09°	14.62°	0.55 to 28.69°	**4.58°**	-9.12 to 18.28°
	Knee	Flexion$^{*)}$	-66.46°	-91.15 to -41.77°	**0.01°**	-14.75 to 14.78°	-42.97°	-155.55 to 69.61°

$^{*)}$ Measurements were taken with the subject lying on the floor (sleeping posture).

modalities, as shown by bold numbers in Table II. These numbers cover 41% of the measurements obtained with a single camera and 55% with stereo cameras and the Kinect. High mean bias was observed in resulting measurements with a single camera and the Kinect, as shown by underlined numbers. 3D joint measurement using a single camera relies on the 3D human pose dataset used in the 3D-pose-baseline library. Therefore, this library may not provide optimal results when estimating 3D human postures with specific postures such as sleeping postures. Kinects also appear to be insufficient to measure body joints in these postures. Morever, the Kinect suffers from occluded joints which cannot be observed by the its depth sensor. The resulting measurements using the stereo camera indicate relatively better accuracies than those of the single camera and the Kinect. Overall, our measurements show that the 95% LOA for the discrepancy of the three modalities against the goniometer exceeded ±5°, which can be considered as clinically significant. For the stereo camera and the Kinect, this finding is consistent with [1].

ROM measurement results for all joints from each user are presented in Table III. Overall, regardless the subjects, the measurements taken with a single camera significantly show higher mean bias than those taken with the stereo camera and the Kinect, $F(1, 28)= 21.82$, $p<0.01$. There is no trend of results found with specific subjects. Differences in results are more likely depending on the type of motion and posture, as shown in Table II. Here, we consider that there is no significant difference in measurement results depending on clothes. This observation is different from what previous studies suggested [1].

b. Continuous measurement

Table IV shows RMSE for polynomial regression model fitted to the measurement results from the three modalities. Here, the smaller the RMSE, the more stable the measurement. To compare measurement stability among the three modalities, two-way Analysis of Variance (ANOVA)

TABLE III. MEAN BIAS AND LOA OF THE MAXIMUM ROM MEASUREMENTS FOR ALL JOINT ACQUIRED USING THREE MODALITIES AGAINST THE GONIOMETER.

Subject	Single camera		Stereo camera		Kinect	
	Mean bias	95% LOA	Mean bias	95% LOA	Mean bias	95% LOA
A	-10.00°	-72.20 to 52.20°	8.86°	-25.54 to 43.26°	-7.33°	-98.41 to 83.74°
B	-13.69°	-63.99 to 36.61°	4.87°	-16.70 to 26.45°	5.38°	-92.20 to 102.97°
C	-13.11°	-66.24 to 40.01°	8.29°	-16.00 to 32.58°	2.72°	-86.36 to 91.80°
D	-15.89°	-61.31 to 29.52°	3.30°	-14.53 to 21.13°	-3.24°	-42.68 to 36.20°
E	-17.05°	-63.21 to 29.10°	9.32°	-20.47 to 39.11°	-4.99°	-66.11 to 56.13°
F	-13.78°	-67.41 to 39.86°	5.09°	-9.62 to 19.80°	4.96°	-45.23 to 55.14°
G	-10.71°	-57.19 to 35.77°	2.22°	-21.53 to 25.98°	10.22°	-52.06 to 72.51°
H	-10.61°	-63.76 to 42.54°	6.06°	-9.93 to 22.06°	5.03°	-52.00 to 62.06°
I	-15.71°	-59.97 to 28.55°	3.18°	-28.30 to 34.66°	6.11°	-77.21 to 89.44°
J	-17.37°	-71.80 to 37.06°	7.51°	-14.90 to 29.91°	-6.42°	-71.31 to 58.47°

TABLE IV. RMSE FOR MODEL FITTING RESULTS.

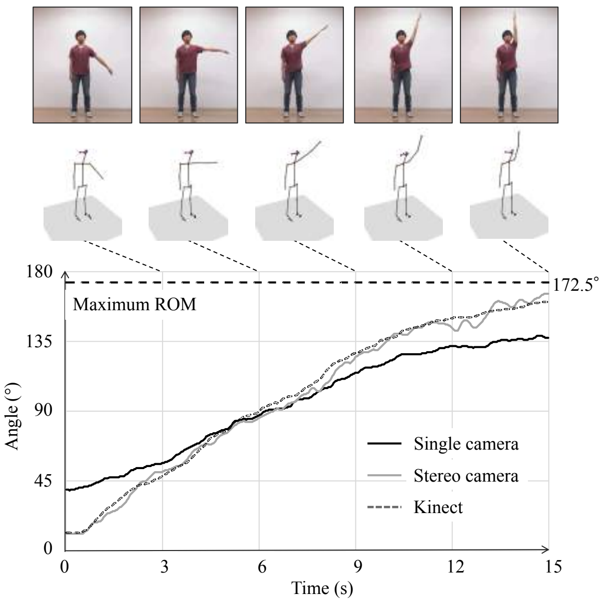
Joint	Motion	Single camera	Stereo camera	Kinect	Fluctuation
Shoulder	Forward Flexion	7.121°	7.624°	15.610°	High
	Backward Extension	0.797°	3.439°	1.131°	Low
	Abduction	2.463°	4.110°	3.384°	Low
	Adduction	3.652°	4.233°	1.993°	Low
Elbow	Frontal Flexion	5.605°	9.734°	5.791°	Moderate
	Side Flexion	2.152°	9.007°	3.918°	Moderate
Hip	Flexion	6.050°	4.129°	8.128°	Moderate
	Extension	1.110°	3.880°	6.501°	Moderate
	External Rotation	3.834°	3.935°	4.576°	Low
	Internal Rotation	2.522°	3.991°	3.050°	Low
Knee	Flexion	8.829°	7.105°	11.019°	High
Mean		4.0124°	5.5625°	5.9184°	
Standard deviation		2.57969°	2.32978°	4.28928°	

for two factors (joint movement and modality) was performed. The main effect of the joint movement was not significant, $F(1, 29) = 1.93$, $p = 0.18$. The main effect of the modality was also not significant, $F(1, 29) = 0.56$, $p = 0.46$. Hence, there are no significant difference in RMSE of measurement data with the three modalities.

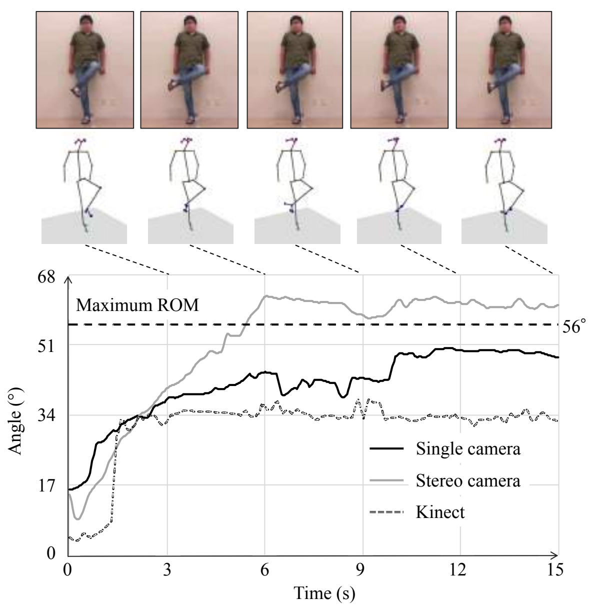
For further analysis, a visual interpretation is conducted to visualize the fluctuation in the resulting measurement during 15s motion. The amount of fluctuations were categorized according to the RMSE values: low, moderate, and high. Figures 4 ~ 6 show samples with low, medium and high fluctuations in the measurement results, respectively. The Dashed straight lines represent maximum ROM meausured by the goniometer. Figure 4 shows two samples of measurement results (shoulder abduction and hip internal rotation) with low fluctuations. Joint angles measured with the three modalities show similar trends. However, the Kinect measures the maximum ROM close to the goniometer. Figure 5 shows two samples of measurement results (elbow frontal flexion and elbow side flexion) with moderate fluctuations. Occlusion at the elbow joint has a slight effect on the Kinect measurements. Figure 6 shows two measurements of shoulder forward flexion with high fluctuations. Occlusion at the elbow joint greatly affects the Kinect measurements. During the experiment, the Kinect fails to measure some postures as indicated by the drops in the curve. Figure 7 shows Kinect's failure to measure hip extension and hip flexion in sleeping posture.

V. CONCLUSION AND FUTURE WORK

In this study, 2D human pose and 3D human pose estimation techniques were used to measure 3D body joints and calculate their ROM for various motions and postures. The former was applied to a stereo camera to estimate 3D joints using triangulation. The latter was solely applied to a single camera to estimate 3D joints by referring to the 3D human pose dataset. Based on our experiments, the stereo camera gives the best results with a small bias to the goniometer compared to the single camera and the Kinect.

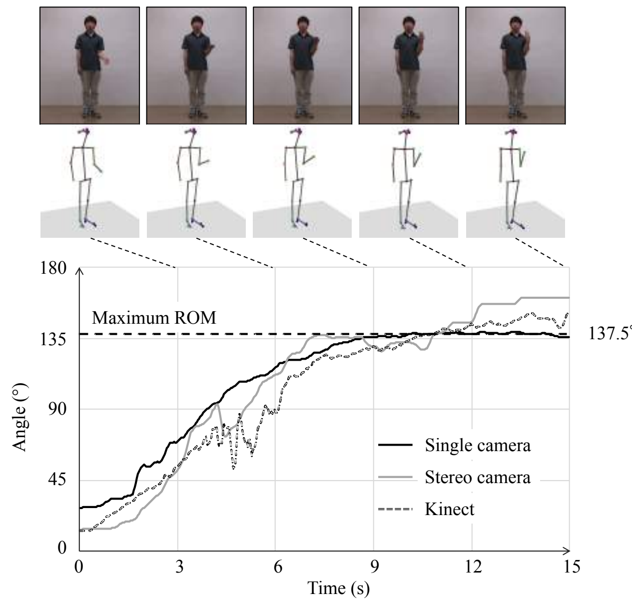


(a) Shoulder abduction

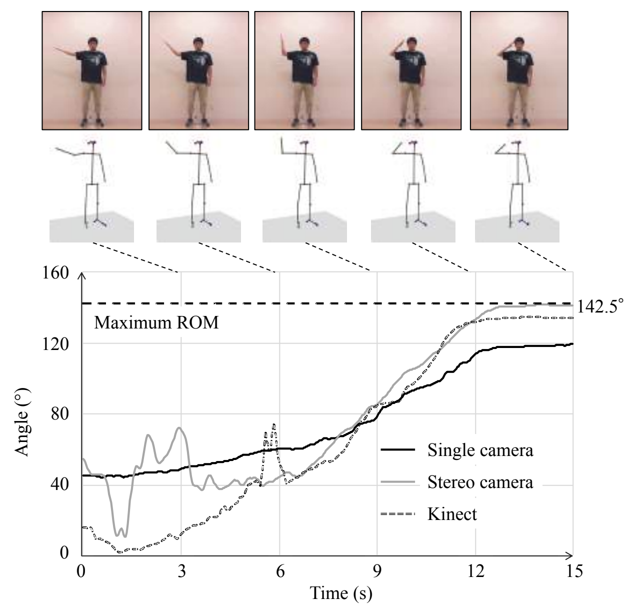


(b) Hip internal rotation

Figure 4. Samples of measurement results with low fluctuations.



(a) Elbow frontal flexion



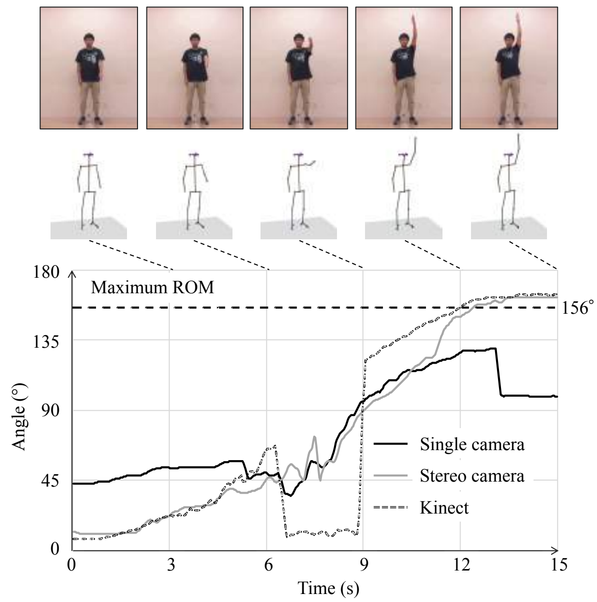
(b) Elbow side flexion

Figure 5. Samples of measurement results with moderate fluctuations.

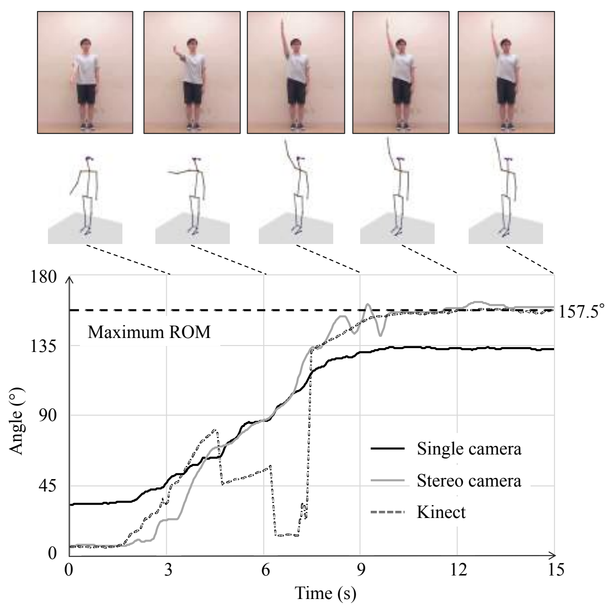
The Kinect surprisingly shows a high bias to goniometer in several measurement cases (Table II).

Occluded joints are a concern for 3D joint measurements. Since our method of calculating 3D joints with either a single camera or a stereo camera is based on the OpenPose library, in many cases, semi-occluded joint locations can be estimated to calculate the 3D locations. In contrast, the Kinect's algorithm does not take special measures against occlusion. Another drawback is that the Kinect needs to detect the posture of the whole body in order to properly detect a particular joint.

The 3D human pose based on a single camera has various applications. People can easily measure their posture everywhere using their own camera. It can be used to build a TR service, allowing clinicians to interact with patients in real-time. By promoting TR, we can expect to reduce the potential time and cost of rehabilitation services, especially for individuals who have economically disadvantaged.

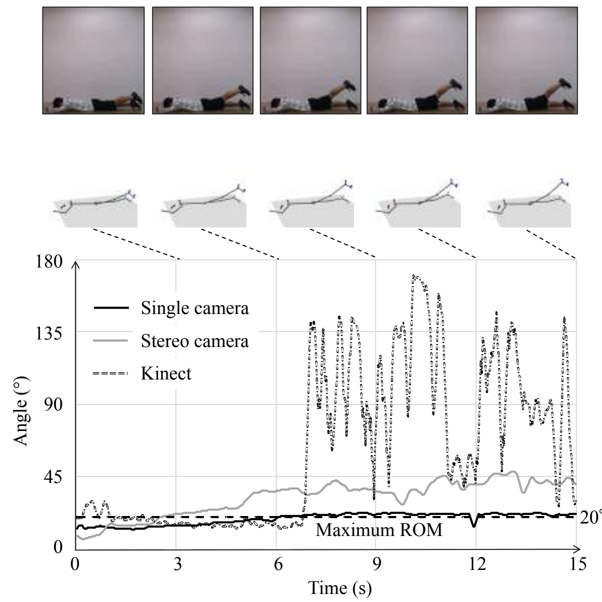


(a) Shoulder forward flexion

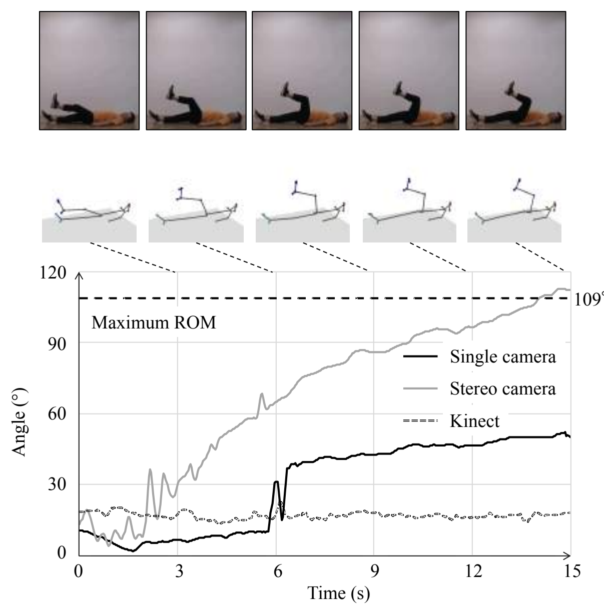


(b) Shoulder forward flexion

Figure 6. Samples of measurement results with high fluctuations.



(a) Hip extension



(b) Hip flexion

Figure 7. Samples of the Kinect's failures.

Our future work includes improving 3D human pose technology to enable body orientation measurement and detection of specific human behavior. We are also working on 3D human poses based on a 360-degree camera that allows measurement of 3D joints in all directions.

ACKNOWLEDGEMENT

We acknowledge the effort from the authors of OpenPose and 3d-pose-baseline to make 3D joint measurements using a single camera possible. This work was supported by the MIC/SCOPE #181602007. We wish to acknowledge the four anonymous reviewers for their careful reading of our manuscript and their many insightful comments and suggestions.

REFERENCES

[1] Y. Ono, O. D. A. Prima, T. Imabuchi, Y. Murata, H. Ito, and Y. Nishimura, "Assessment of Joint Range of Motion Measured by a Stereo Camera," eTELEMED 2019 : The Eleventh

International Conference on eHealth, Telemedicine, and Social Medicine, IARIA, pp. 23-29, 2019.

[2] 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.

[3] 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.

[4] 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.

[5] K. Mitchell, S. B. Gutierrez, S. Sutton, S. Morton, and A. Morgenthaler, "Reliability and validity of goniometric iPhone applications for the assessment of active shoulder external rotation," *Physiotherapy Theory and Practice*, vol. 30, no. 7, pp. 521–525, 2014.

[6] DrGoniometer. http://www.drgoniometer.com/ [retrieved: December, 2019]

[7] T. Dutta, "Evaluation of the Kinect sensor for 3D kinematic measurement in the workplace," *Applied Ergonomics*, vol. 43, no. 4, pp. 645–649, 2012, doi:10.1016/j.apergo.2011.09.011.

[8] 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.

[9] S. H. Lee, C. Yoon, S. G. Chung, H. C. Kim, Y. Kwak, H-w. Park, and K. Kim, "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.

[10] J. Spörri, C. Schiefermüller, and E. Müller, "Collecting kin-ematic data on a Ski track with optoelectronic stereophotogram- metry: A methodological study assessing the feasibility of bringing the biomechanics Lab to the field," *PloS ONE*, vol. 11, no. 8, e0161757, 2016.

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

[14] 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* (*CVPR*), 2017 IEEE Conference, pp. 7291-7299, 2017.

[15] Y. Ono, 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)

[16] H. J. Lee and Z. Chen, "Determination of 3D Human Body Postures from a Single View," *Computer Vision, Graphics and Image Processing*, vol. 30, pp. 148–168, 1985.

[17] J. Martinez, R. Hossain, J. Romero, and J. J. Little, "A Simple Yet Effective Baseline for 3d Human Pose Estimation," arXiv preprint arXiv:1705.03098, pp. 1-10, 2017.

[18] C. Ionescu, D. Papava, V. Olaru, and C. Sminchisescu, "Human3.6M: Large Scale Datasets and Predictive Methods for 3D Human Sensing in Natural Environments," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 7, pp. 1325-1339, 2014.

[19] 3D-pose-baseline. https://github.com/una-dinosauria/3d-pose-baseline. [retrieved: December, 2019]

[20] O. D. A. Prima, T. Imabuchi, Y. Ono, Y. Murata, H. Ito, and Y. Nishimura, "Single Camera 3D Human Pose Estimation for Tele-rehabilitation," eTELEMED 2019 : The Eleventh International Conference on eHealth, Telemedicine, and Social Medicine, IARIA, pp. 13-18, 2019.

[21] 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.

[22] OpenPose, https://github.com/CMU-Perceptual-Computing-Lab/openpose. [retrieved: December, 2019]

[23] 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.

[24] 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.