Feasibility of using depth cameras for evaluating human - exoskeleton interaction

Soocheol Yoon[1,4], Ya-Shian Li-Baboud[1], Ann Virts[1], Roger Bostelman[2], and Mili Shah[1,3]
[1]National Institute of Standards and Technology, [2]Smart HLPR LLC, [3]The Cooper Union, [4]Georgetown University

With the increased use of exoskeletons in a variety of fields such as industry, military, and health care, there is a need for measurement standards to understand the effects of exoskeletons on human motion. Optical tracking systems (OTS) provide high accuracy human motion tracking, but are expensive, require markers, and constrain the tests to a specified area where the cameras can provide sufficient coverage. This study describes the feasibility of using lower cost, portable, markerless depth camera systems for measuring human and exoskeleton 3-dimensional (3D) joint positions and angles. A human performing a variety of industrial tasks while wearing three different exoskeletons was tracked by both an OTS with modified skeletal models and a depth camera body tracking system. A comparison of the acquired data was then used to facilitate discussions regarding the potential use of depth cameras for exoskeleton evaluation.

INTRODUCTION

The exoskeleton market is expected to increase from $360 million in 2020 to $4.92 billion by 2027 (Market Reports World, 2021). This increase is largely due to the potential for exoskeletons to support work efficiency and decrease injury risks in many domains. In the automobile industry, exoskeletons have been shown to help in the performance of repetitive tasks that occur in the assembly process (Iranzo, 2020). In the defense industry, exoskeletons have been introduced for repair and maintenance of warships (Magnuson, 2015). In the healthcare industry, exoskeletons have been studied to aid in patient transport (Tröster, 2020). However, standards to measure the performance of exoskeletons using markerless, mobile, low costs techniques are lacking (Lowe, 2019). The Intelligent Systems Division (ISD) at the National Institute of Standards and Technology (NIST) is leading the development of these standards through the ASTM F48 committee (ASTM, 2022).

The actual changes in human body and motion are one of the significant factors when evaluating the effectiveness of an exoskeleton (Zanotto, 2015). There are several ways to measure human body motion and forces. For instance, marker-based methods using optical tracking systems (OTS), provide an understanding of human kinematics (Bostelman, 2020). Limitations of these methods include uncertainties in modeling and tracking of a human joint while wearing an exoskeleton.

Motion tracking using depth cameras has been considered in various situations such as static posture, gait assessment, and movement measurement (do Carmo Vilas-Boas, 2019). In addition, research has been conducted to classify dynamic objects by leveraging artificial intelligence and big data techniques to advance machine learning algorithms in computer vision, which are being applied to improve human pose estimation (Moon, 2019). There have been various devices, software, and applications for measuring the body skeletal joint positions and motion using depth cameras. These devices capture depth images based on the infrared (IR) sensor input to segment the human subjects and to estimate the 2D skeletal joint positions using machine learning algorithms such as convolutional neural networks (Liu, 2020). The depth information can be combined with model fitting, to estimate the 3D joint positions (Liu, 2020). However, many of the pose estimation algorithms were trained on datasets of human poses without exoskeletons. In addition, many industrial poses, such as squatting, can occlude certain joints or limit the camera positioning options, especially when the cameras are placed behind or to the side of the subject due to the test environment. The NIST ISD exoskeleton research team has experimented with various depth camera body tracking system (DCBTS)s for industrial exoskeleton performance tests. The DCBTSs showed different results but had common limitations caused by the factors described above.

This study systematically evaluated the feasibility of using depth cameras to detect human skeletal joints for industrial exoskeleton performance evaluation by determining the visibility of human body joints of a user performing a set of simulated industrial task poses, while wearing three different exoskeletons and without the exoskeleton from different camera positions. In addition, the consistency of measuring human joints was evaluated. These results were compared with data gathered from an OTS.

EXPERIMENT

This paper explores the use of depth cameras to measure the effects of exoskeletons on the motions of humans performing industrial tasks. Two experiments were performed in this study: the first measured the accuracy of the DCBTS to recognize human joints while a user performs industrial tasks, while the second measured how consistently a DCBTS provided joint angle data. Based on modified conventional markersets, the OTS, one of the gold standards in motion tracking, provided the skeletal joint estimations available in 6 degrees of freedom (6DoF), which were used as a reference for comparing DCBTS performance.

Experiment setup

Three types of exoskeletons were used in the experiment as shown in Figure 1. Type 1 was a full body exoskeleton consisting of metal frames and straps with mechanical knee

Figure 1. Exoskeletons of Type 1 (Left), Type 2 (Middle), and Type 3 (Right)

support. It supported the shoulder, back, and knee activities. Type 2 was an upper body exoskeleton consisting of metal frames and straps to support back and shoulder activities. Type 3 was an exosuit consisting of elastic straps providing back and knee support. Each exoskeleton was adjusted to best fit the subject which includes postural support or torque assistance for the arm, shoulder, torso, back, hip, femur, tibia, and knee. Base tests were performed first where the subject did not wear an exoskeleton.

Load alignment and applied force were chosen to represent actual industrial applications such as tire installation and grinding. The load alignment tasks required poses of squat arm-down, squat arm-forward, and stand arm-forward Figure 2 (top), while the applied force tasks required poses where the subject bends at the waist with arms extending downward (waist-bend arm-downward) and where the subject stands with arms extending-upward (stand arm-upward) Figure 2 (bottom). These five poses were tested in the first experiment. A standing pose, with legs together and arms fully extended like a 'T' (T-pose) and a standing arm-upward pose were used in the second test.

The tests were performed by one subject for the initial feasibility study of using depth cameras to control and minimize confounding factors. The subject height was 1.765 m, weighed 80 kg, and met the manufacturers' fit guidelines. The subject had two years of experience with the exoskeletons that were used in this experiment.

Measurement system and procedure

The tests were performed in a motion capture area of 9 m by 22 m between July and September 2021. The depth cameras were placed to the side, side-back, and back of the subject with respect to the exoskeleton test environment. The distance between each camera and the subject was 2.7 m, and the height of the camera was 0.87 m. Microsoft Azure Kinect cameras and the associated body tracking software development kit were used as the body tracking system. The depth images were recorded at a 640 x 576 pixel resolution and at 30 frames per second (fps).

Twenty OptiTrack Prime 41 cameras were mounted at a height of approximately 2.7 m off the ground with one overhead camera at 4.3 m relative to the ground, acquiring data at a rate of 120 fps using Motive versions 2.2 and 3.0 beta. The twenty

Figure 2. Load alignment poses (Top) and applied force poses (Bottom)

cameras were triggered using the manufacturer-provided synchronization system. The OTS was calibrated per manufacturer's documentation on a weekly basis using a calibration bar. Figure 3 illustrates the experimental setup.

For each task, the subject repeated each pose five times for a duration of 5 seconds per repetition. Between each pose, the subject returned to a neutral standing pose. The body tracking performance was evaluated based on detection of arm and leg skeletal joints by visual inspection. Each trial is considered successful when the detected skeletal joints of the arms or legs continuously align with the subject's perceived skeletal joints for at least 5 seconds. Figure 4 (a) and (b) demonstrate correct and erroneous skeletal joint estimation, respectively.

The same test was performed separately using the OTS to avoid optical interference between the OTS and the image sensor in the depth camera. The subject wore a body tracking suit with passive, 15 mm retroreflective markers. When wearing the exoskeletons, several markers were attached directly to the exoskeletons to track the subject's joints using the body tracking markersets available in the OTS software. This study used the conventional markerset as the skeletal model to minimize the number of occluded markers relative to the exoskeletons used. Figure 4 shows the subject with the

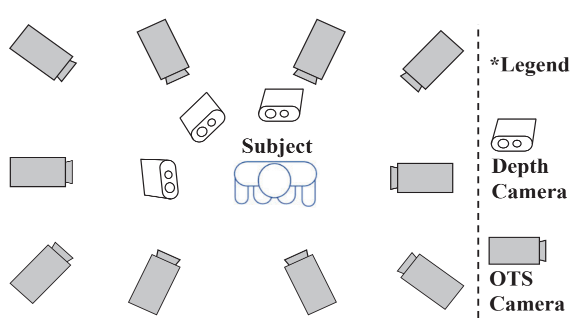


Figure 3. Optical and depth camera setup

Table 1. Body tracking usability test result

	Squat Arm-Down				Squat Arm-Forward				Stand Arm-Forward				Stand Arm-Upward				Waist-bend Arm-down			
	*S	SB	B	OTS	S	SB	B	OTS	S	SB	B	OTS	S	SB	B	OTS	S	SB	B	OTS
Arm																				
Base	0	0	0	5	0	0	0	5	4	0	5	5	0	1	2	2	0	3	4	5
Type 1	2	0	0	3	0	0	2	2	3	0	2	5	0	0	3	0	0	0	1	3
Type 2	0	0	0	4	0	0	0	5	5	0	0	5	0	2	4	5	0	1	4	5
Type 3	4	0	0	5	2	0	0	5	4	0	0	5	0	0	0	5	0	0	0	5
Legs																				
Base	5	2	3	5	5	3	5	5	5	5	5	5	5	5	5	5	1	3	4	5
Type 1	0	0	5	4	5	3	5	5	5	0	4	5	5	0	5	5	0	0	0	5
Type 2	5	1	0	5	5	5	5	5	5	2	5	5	2	4	0	5	0	1	1	5
Type 3	5	0	0	4	3	2	0	5	5	3	0	5	5	4	4	5	2	0	0	5

*Side view(S), Side-back view(SB), back view(B) of DCBTS

optical markers when wearing the marker suit and exoskeleton Type 1 with the conventional markerset (c), and the resolved skeleton tracking model (d).

A static pose test was performed to measure how accurate and constant joint angle values can be obtained through each body tracking system for both the depth camera and OTS experiments. The target poses are the T-pose and stand arm-upward test. Target joints are shoulder, knee, and torso. The subject maintained each pose for more than 10 seconds, and each system measured the subject's joint poses.

RESULTS

Body tracking usability

The experimental results are shown in Table 1. Each cell represents the number of successful body tracking for each condition. For each trial, it was considered successful if the tracked joint pose is on the actual pose from the overlaid image. In the case of tracking the arm, it was required that the shoulder, elbow, and wrists of both arms are on the actual poses. In the case of tracking the leg, the hip, knee, and heel of both legs needed to be on the actual poses. In addition, each joint detected was successful when it was continuously tracked for each task period.

The DCBTS successfully detected the subject's joint 230 out of 600 trials or 38.3% of the time. The most frequently detected pose was the stand arm-forward, and the least frequently detected pose was the waist-bend arm-down. The side perspective resulted in optimal detection of the subject's arm and leg skeletal joints. Out of 150 trials each, the Type 1 exoskeleton test resulted in the detection of the subject's arm and leg joints 50 times (or 34.0% of the time), the Type 2 exoskeleton test 57 times (or 38.0% of the time), and the Type 3 exoskeleton test 43 times (or 28.7% of the time).

The OTS performed better compared to DCBTS tracking by successfully detecting the subject's joints 182 times out of 200 trials (or 91.0% of the time). The arms resulted in the most failed detections while the subject wore the Type 1 exoskeleton.

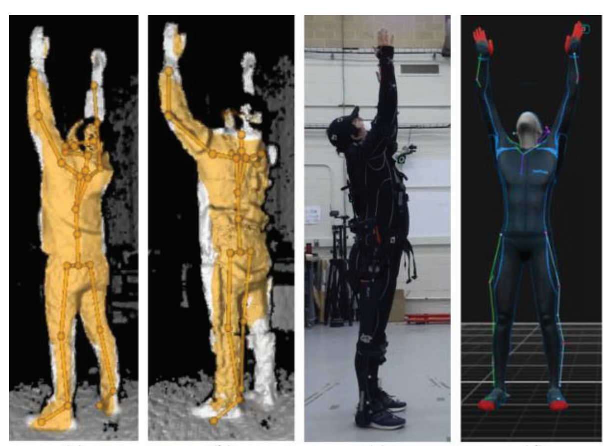
Joint angle measurement

The DCBTS collected a total of 120 images, and the OTS system collected a total of 1000 frames for the T-Pose and stand arm-upward. Using the collected body pose data, joint angles were estimated by each system. Joint angles were calculated from the DCBTS result using vector projection where the knee angle by right ankle-knee-hip, the shoulder angle by left clavicle-shoulder-elbow (sagittal plane), and the torso angle by lower spine-upper spine-neck (coronal plane). The knee angle was computed from the OTS 6DoF data using the quaternion rotation between the thigh and shin bone segments right thigh-shin (sagittal plane), the shoulder angle between the clavicle and humerus bone segments through left shoulder-upper arm (sagittal plane), and the torso angle using the Euler angle relative to the thorax (coronal plane). The results are shown in Table 2 including manual measurements using a goniometer.

For the knee flexion angle, the DCBTS measurement showed a value closer to that of the goniometer for the T-pose. The angle differences between the DCBTS and the OTS ranged 7° to 8°. DCBTS showed smaller than 1° standard deviations for both poses, while OTS showed larger than 2° in the stand arm-up test. For the shoulder angle, the OTS, the DCBTS, and the goniometer measurements showed angle differences between 6° and 9° for the T-Pose. In the case of the stand arm-upward pose, the angle differences ranged from 4° to 17°. For the torso angle, the OTS showed a value closer to that of the goniometer. Both standard deviations were lower than 2° in the T-Pose and the stand arm-upward tests.

ANALYSIS AND DISCUSSION

DCBTS and OTS body tracking results analysis



(a) (b) (c) (d)

Figure 4: Body tracking results by (a) DCBTS - all joints present, (b) DCBTS - missing joints, (c) wearing an OTS tracking suit, and (d) data from OTS system.

Table 2. Static joint angle measurement test result

Angles (°)	DCBTS mean[std]	OTS mean[std]	Gonio -meter
T-Pose			
Knee	6.3[0.9]	14.5[0.2]	8.5
Shoulder	18.9[5.7]	21.2[0.8]	13.1
Torso	10.2[1.7]	3.3[0.3]	2.9
Stand Arm-Up			
Knee	5.5[0.8]	12.2[2.1]	5.2
Shoulder	69.9[1.7]	52.5[2.0]	53.3
Torso	13.9[0.6]	3.0[1.4]	4.0

The arm recognition rate was generally lower than that for the leg. For all views, the number of successful arm detections was lower than the number of successful leg detections. The arm detection rate was significantly reduced in the squat poses. It was observed that the arm detection rate decreased when the arm was not directly visible depending on the task motion and camera position. The leg detection performance had the most significant skeletal joint tracking errors in the waist-bend and arm-down pose. Although the leg was visible to the camera, the joint detection algorithm often erred significantly.

The body tracking performance also differed depending on which exoskeleton was worn. Source of joint detection errors include marker occlusion by the exoskeleton and the modification of marker placement to minimize occlusions. For example, from the side view, the Type 3 exoskeleton test detected the arm most frequently, and all exoskeleton tests showed similar or better results in the squat arm-forward poses than the base. The Type 1 exoskeleton test showed the most frequent leg detection from the back view, and the Type 2 exoskeleton test showed the most frequent leg detection from the side-back view.

The Type 3 exoskeleton test had limited correct joint detections from the back view. This may be due to the retroreflective tape attached to the Type 3 exoskeleton, which decreased the body tracking performance. Better results were obtained when the safety tape was covered, and the same experiment was repeated with this change.

To see the effect of camera placements, the same experiments were repeated with front and front-side placements. The DCBTS performed optimally when the camera was placed in front of the subject. From 25 trials where the camera was placed in the front, successful whole body tracking results were obtained in 24 cases from the base test, 17 cases from the Type 1 exoskeleton test, 16 cases from the Type 2 exoskeleton test, and 23 cases from the Type 3 exoskeleton test. If the side or back body tracking algorithm is improved, where the joint detection is as good as the front, the value of using depth cameras in exoskeleton research is expected to increase.

Considering the number of successful skeletal joint tracking, the OTS showed more stable results than the DCBTS. Unlike the DCBTS, the OTS measurements were less sensitive to the Type 3 exoskeleton reflective tape even though it uses optical markers. The Type 3 exoskeleton test showed the highest performance through successful detection of 49 out of a total of 50 trials.

Joint angle measurement analysis

Table 3. DCBTS and OTS comparison

	Depth Camera Body Tracking System	Optical Tracking System
Body Tracking Performance	38% success rate Affected by camera position Resolution and FPS are dependent to devices, up to 1024 x 1024 and 30 Mean and std deviation of standing knee joint angle: 6.3 [0.9] °	91 % success rate 360° body tracking Less than 1 mm error, and 120 fps Mean and std deviation of standing knee joint angle: 14.5 [0.2] °
Ease of Implementation	System and subject can move anywhere Subject should be within 2 m from the cameras System requires high graphical computing power There is no requirement for subjects Calibration: not necessary when single device usage for deriving relative position. Checkerboard calibration is required when using multi-device or for absolute position.	System is fixed in certain place Subject can move anywhere in system area System requires high network volumes Marker should be attached to subjects Calibration:0.08 mm wand error, 0.65 mm 3D reconstruction error.
Price	Low-cost recording device Most of software and application are free for developers	High-cost entire system (cameras, software, and network)

The core of the body tracking system in exoskeleton research is how consistently the skeletal joints are detected for the same pose. Even if there is a difference between the OTS and the DCBTS methods, subject motion changes can be measured when the system consistently provides similar pose estimation for the same motion. The standard deviation of the static motion test indicated the consistency of the systems' joint angle measurement.

It was observed that once the DCBTS correctly recognized the subject's motion, the performance of the consistent joint information was often comparable to that of the OTS. Therefore, if the DCBTS improves the recognition rate for complex motion as well as exoskeleton-wearing subjects, it could be used more effectively for exoskeleton research.

Data acquisition from each system differs for the successfully detected skeletal joints. Although both systems provide joint 3D position and a bone segment rotation quaternion, the DCBTS provides the positions relative to the camera, while the OTS provides the absolute position in a user-defined coordinate system. In addition, the OTS can provide angle information of the subject with respect to the ground plane.

Systems comparison

In addition to the body tracking performance analysis, the DCBTS and OTS were also compared in terms of ease of use and cost and the results are shown in Table 3. The constraints on the introduction and use of each system are very different. In the case of the DCBTS, it can be easily used anywhere, but the subject should be within two meters from the camera. In the case of the OTS, it is costly, labor intensive, and time consuming to install, and the test is often constrained in the space where the system is installed. However, the subject's activity range is much wider.

While skeletal tracking performance improvement with and without exoskeletons are needed, DCBTSs have the potential to be used in exoskeleton studies. Most of all, it is easy to capture the images in the field. This means that the system can measure the subject's motion of exoskeleton tests regardless of location. Second, subjects can wear typical attire without constraints of applying markers. This reduces the burden on the subject. Lastly, the system can potentially be used in the field to evaluate the ergonomics of a human-exoskeleton system while performing real tasks without significant investment.

This study demonstrated a methodology to evaluate the feasibility of using depth cameras in industrial exoskeleton test methods. For example, if an exoskeleton test aims to measure the changes in task completion rate, depth cameras can be applied by using the subject's relative skeletal joint positions over time to detect the beginning and the end of the task. On the other hand, when measuring the precise change of joint angles as a result of wearing exoskeletons, an OTS can provide more precise joint position estimation through strategic marker placement along the joints and segments of interest.

CONCLUSION

This study evaluated the feasibility of using a state-of-the-art DCBTS as a measurement method for exoskeleton studies. Five industrial poses were chosen considering load alignment and applied force tasks. Depth cameras were placed at the side, side-back, and back of the test subject in consideration of the simulated industrial exoskeleton test environment. The results showed that the DCBTS successfully detected all joints of arms or legs 230 out of 600 times. The performance of the DCBTS depends on the position of the camera, the type of exoskeleton worn, and the industrial task motion. The results were compared to results from an OTS, which is one of the leading technologies in human motion tracking. The OTS successfully detected 182 times out of 200 trials, which is 53% better than the DCBTS under the same conditions. It was verified that the consistency of the joint angles derived using the skeletal joint position and quaternion data from the DCBTSs for the successfully detected subject were comparable to the estimates derived from the OTS based on the results from the static pose tests. Despite the lower overall joint detection performance, depth cameras have the potential to provide useful measurements as they are portable, low-cost, easy to use, and open to algorithmic adaptation to achieve improved results with test subjects wearing exoskeletons.

The NIST ISD exoskeleton research team continues to develop technologies and research body measurement techniques for exoskeleton studies. In particular, the feasibility study presented here will be extended to more industrial tasks with more poses. In addition, other body tracking systems will be considered including color image-based systems and specialized algorithms will be developed to detect the human body while wearing exoskeletons.

DISCLAIMER

Commercial products are identified in this paper to foster understanding. This does not imply recommendation or endorsement by NIST, nor that the products identified are necessarily the best available for the purpose.

REFERENCE

Market Reports World. (2021). Global exoskeleton industry research report, growth trends and competitive analysis 2021-2027, https://www.marketreportsworld.com/global-exoskeleton-industry-18235662, accessed by FEB08 2022

Iranzo, S., Piedrabuena, A., Iordanov, D., Martinez-Iranzo, U., & Belda-Lois, J. M. (2020). Ergonomics assessment of passive upper-limb exoskeletons in an automotive assembly plant. Applied Ergonomics, 87, 103120.

Magnuson, S. (2015). Power Remains Key Challenge for Building SOCOM's Iron Man Suit. National Defense, 99(738), 38-43.

Tröster, M., Wagner, D., Müller-Graf, F., Maufroy, C., Schneider, U., & Bauernhansl, T. (2020). Biomechanical model-based development of an active occupational upper-limb exoskeleton to support healthcare workers in the surgery waiting room. International Journal of Environmental Research and Public Health, 17(14), 5140.

Lowe, B. D., Billotte, W. G., & Peterson, D. R. (2019). ASTM F48 formation and standards for industrial exoskeletons and exosuits. IISE transactions on occupational ergonomics and human factors, 7(3-4), 230-236.

ASTM Committee F48, Exoskeletons and Exosuits, https://www.astm.org/get-involved/technical-committees/committee-f48/subcommittee-f48, accessed by FEB08 2022

Zanotto, D., Akiyama, Y., Stegall, P., & Agrawal, S. K. (2015). Knee joint misalignment in exoskeletons for the lower extremities: Effects on user's gait. IEEE Transactions on Robotics, 31(4), 978-987.

Bostelman, R., Li-Baboud, Y., Van, K. and Shah, M. (2020), Development of a Kinematic Measurement Method for Knee Exoskeleton Fit to a User, Technical Note (NIST TN), National Institute of Standards and Technology, Gaithersburg, MD, https://doi.org/10.6028/NIST.TN.2107

do Carmo Vilas-Boas, M., Choupina, H. M. P., Rocha, A. P., Fernandes, J. M., & Cunha, J. P. S. (2019). Full-body motion assessment: Concurrent validation of two body tracking depth sensors versus a gold standard system during gait. Journal of biomechanics, 87, 189-196.

Moon, G., Chang, J. Y., & Lee, K. M. (2018). V2v-posenet: Voxel-to-voxel prediction network for accurate 3d hand and human pose estimation from a single depth map. In Proceedings of the IEEE conference on computer vision and pattern Recognition (pp. 5079-5088).

Pavón-Pulido, N., López-Riquelme, J. A., & Feliú-Batlle, J. J. (2020). IoT architecture for smart control of an exoskeleton robot in rehabilitation by using a natural user interface based on gestures. Journal of Medical Systems, 44(9), 1-10.

Gancet, J., Ilzkovitz, M., Motard, E., Nevatia, Y., Letier, P., De Weerdt, D., ... & Thorsteinsson, F. (2012, June). MINDWALKER: Going one step further with assistive lower limbs exoskeleton for SCI condition subjects. In 2012 4th IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob) (pp. 1794-1800). IEEE.

Keselman, L., Iselin Woodfill, J., Grunnet-Jepsen, A., & Bhowmik, A. (2017). Intel realsense stereoscopic depth cameras. In Proceedings of the IEEE conference on computer vision and pattern recognition workshops (pp. 1-10).