Motion-based Control Interface for Intuitive and Efficient Teleoperation of Construction Robots

Usman Rasheed[1]; Xiaoyun Liang[2]; Jiannan Cai, Ph.D.[3*]; Shuai Li, Ph.D.[4]; and Yuqing Hu, Ph.D.[5]

[1]School of Civil & Environmental Engineering, and Construction Management, The University of Texas at San Antonio. Email: usman.rasheed@utsa.edu
[2]School of Civil & Environmental Engineering, and Construction Management, The University of Texas at San Antonio. Email: xiaoyun.liang@utsa.edu
[3*]School of Civil & Environmental Engineering, and Construction Management, The University of Texas at San Antonio. Email: jiannan.cai@utsa.edu (corresponding author)
[4]Department of Civil and Environmental Engineering, The University of Tennessee, Knoxville. Email: sli48@utk.edu
[5]Department of Architectural Engineering, Penn State University. Email: yfh5204@psu.edu

ABSTRACT

Robotic teleoperation is regarded as a promising method of assisting workers in performing construction activities while avoiding potentially dangerous conditions. However, the control interface of conventional teleoperation systems relies on physical devices, such as a keyboard and joystick, to input control commands, which cannot be intuitively mapped to robot movement, and thus affecting the teleoperation performance of non-expert users. In this research, we present a novel human-robot interface to map human arm movement to robotic arm motion for intuitive teleoperation. Firstly, the human arm is tracked in real-time through a motion capturing system, i.e., OptiTrack. The 3D location data is then streamed to robot operating system (ROS) and is used for motion planning of the robotic arm in MoveIt. Finally, the interface is demonstrated in a teleoperated pick-and-place task through robotic simulation and the usability of the developed framework is compared to that of a joystick-based teleoperation system.

INTRODUCTION

The construction industry is one of the largest and fastest-growing industries in the world, employing millions of workers and contributing significantly to the global economy. However, it is also considered one of the least safe industries due to its complex and dynamic working environment. Workers in the construction industry are often exposed to a variety of hazards, such as falls, electrocution, and machinery accidents. According to the Occupational Safety and Health Administration (OSHA), 4,764 workers died on job sites in the United States in 2020, and 1,282 (27%) of these were in the construction industry (statistics 2022). To improve safety and productivity in the construction industry, there has been a growing interest in the use of robotics and automation.

With years of technical development and advances in robotics, teleoperation has been regarded as a promising method of assisting workers in performing construction activities while avoiding potentially dangerous conditions (Hirche and Buss 2012). Teleoperation allows workers to operate robots from a safe distance, thereby reducing the risk of injury. However, one of the most critical steps in teleoperation is to provide a user-friendly interface to support human-robot interaction, which is less analyzed when developing robot technologies in construction

(Czarnowski et al. 2018). The control interface of conventional teleoperation systems mainly relies on physical devices, such as a joystick (Truong et al. 2017) to input control commands, which cannot be intuitively mapped to robotic movement, affecting the teleoperation performance of non-expert users (Mavridis et al. 2015). The fundamental challenge to design and optimize the robotic system's input for the efficient human-robot interaction is still unresolved (Al-Mouhamed et al. 2008) and demands efficient feedback and control methods to enable seamless integration between human goals and robotic responsiveness (Hirche and Buss 2012).

Gesture-based teleoperation, which enables human operators to communicate with robots using their natural body gestures, is regarded as one of the most intuitive types of human-robot interaction (Molchanov et al. 2016 and Köpüklü et al. 2019) in the noisy environment of construction sites (Zhang et al. 2023). Wang and Zhu (2021) proposed a vision-based system for capturing and interpreting hand gestures of construction workers to control construction machines (Mavridis, Pierris et al. 2015). Du and Zhang (2014) demonstrated the use of hand gestures to control dual robot manipulators without the use of any contact devices or markers. However, gesture disagreement may occur between users, and no standard design is currently available for human-robot collaboration on construction sites. Moreover, gesture-based teleoperation is not very effective for supporting human-robot collaboration in the construction industry since it mostly requires a user to sit or stand stationary while making gestures and could fail when the user starts moving (Wang and Zhu 2021, Zhang et al. 2023).

An alternative to gesture-based teleoperation is to use the operator's body pose as control input (Macchini et al. 2022). For instance, a human operator may control a remote robotic arm using the natural movements of their own arm (Hasegawa et al. 2020). To capture the human pose, different motion capture systems such as image processing system, electromagnetic system, inertial measurement units (IMUs) and vision-based systems are available in the market. The accuracy of image processing system is now outperformed by electromagnetic systems and the electromagnetic system is not very efficient to use for construction robots due to their low sample frequency and sensitivity to ferromagnetic disturbance (Van der Kruk and Reijne 2018). The IMUs are widely studied in literature to capture human motion for teleoperation of construction robots. Kim et al. (2009) developed teleoperation system for the excavator with movements of a human arm. Three inertial measurement units (IMUs) sensors were attached to the operator's arm to track the motion, and the operating commands for the actuators of an excavator were transmitted via bluetooth wireless communications. Škulj et al. (2021) presented a novel wireless wearable system that simply uses IMUs to determine the orientation of the operator's upper body parts. An algorithm was developed to transform measured orientations into movement commands for an industrial collaborative robot. Many researchers tracked the motion of operator's head (Rudigkeit and Gebhard 2019), leg (Moschetti et al. 2019), hand (Fang et al. 2017) and wrist (Gromov et al. 2019) using IMUs for various industrial robotic applications (Bertomeu et al. 2015 and Repnik et al. 2018).

Although IMUs are very promising for reliable human motion tracking (Galzarano et al. 2014), they are susceptible to drift and measurement noises (Zhou and Hu 2008), which may accumulate over time (Bao et al. 2017 and Xu et al. 2019). The accumulative and drift error is the biggest challenge faced by IMUs for motion tracking, causing delays in teleoperation (Lu et al. 2017). In this research, a novel teleoperation framework of a construction robot is proposed, based on the pose of human arm using vision-based motion capture system. Vision-based motion capturing system (e.g., OptiTrack system (OptiTrack 2020) used in this study) has much higher accuracy than IMUs and do not suffer from drift errors in motion tracking of human arm. The estimated poses of the human arm are then streamed to the robot operating system (ROS) and used for motion planning of the robotic arm in MoveIt. This allows users to control the robot through natural arm movements, making teleoperation more intuitive and accessible to non-expert users. A simulated construction scenario, i.e., pick-and-place materials on construction sites, was used to demonstrate the efficacy of proposed interface, which requires careful planning, coordination, and attention to safety and is crucial to the success of any construction project.

MOTION-BASED CONTROL INTERFACE

Figure 1 illustrates the proposed motion-based robot arm control interface. Six OptiTrack cameras were strategically placed around the space to capture motion. OptiTrack cameras principally use infrared light to track the movement of markers placed on the subject being tracked (human's arm in this research). The motive software then used the data collected from the OptiTrack cameras to calculate the positions and orientations of the markers in 3D space. The collected data was then streamed into ROS in real-time. Finally, a node, named classifier, was developed in ROS to subscribe the data coming from motive and classify it into different commands to map the movement of robot arm in the robotic manipulation platform named MoveIt. The ROS visualization (Rviz) was used with MoveIt to visualize robot's motion planning and execution (panda arm in this research).

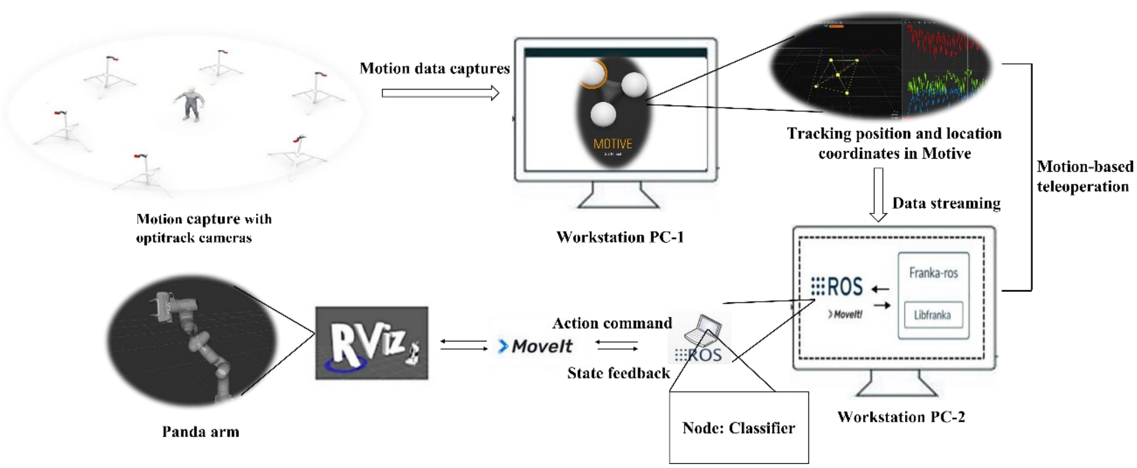
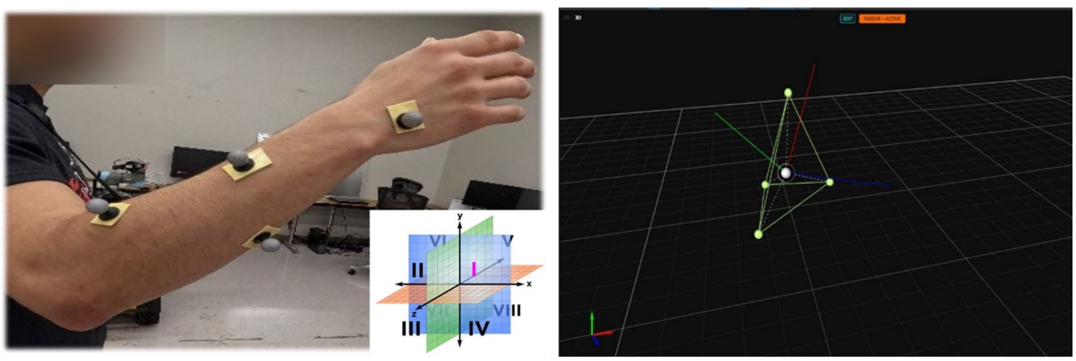


Figure 1. Overview of motion-based teleoperation

The proposed teleoperation system involved four steps. First, the OptiTrack cameras were calibrated to ensure accurate tracking and the markers were placed on the arm of the subject being tracked as shown in Figure. 2 (a). The attached markers were tracked by creating a rigid body in 3D space as shown in Figure 2 (b). In motive, a rigid body (OptiTrack 2022) was created by connecting three or more markers in space using distance constraints. The cameras captured the movement of the rigid body as the subject performed the desired actions. The rigid body was distinguishable and recognizable in the virtual 3D space and the positions and orientations of the rigid body were recorded in motive. Second, the collected data of rigid's body positions and orientations were streamed to ROS in real-time using virtual-reality peripheral network (VRPN) (Taylor et al. 2001). Third, a node called classifier was developed in ROS, which subscribed to the data coming from motive and classified it to generate different commands for the motion planning of the robotic arm. Fourth, for the motion planning and control of robotic arm, MoveIt was used considering the collision avoidance environment, which is a powerful software package built on top of the ROS and is used to control a wide range of robotic systems. The panda arm was used for motion planning and execution as it has seven degrees of freedom, making it highly flexible and capable of performing a wide range of tasks with great precision. The arm was equipped with torque sensors in each joint, which provide accurate feedback to the control system, allowing it to adjust its movements in real-time. The Rviz was used with MoveIt to visualize panda arm's motion planning and execution as it can display robot's planned trajectory, including the start and end position as well as intermediate waypoints.



(a) (b)

Figure 2. Demonstration of creating rigid body to track human arm in motive.
(a) Rigid body created with four markers on human arm; (b) Rigid body in motive

DEMONSTRATION AND DISCUSSIONS

Task Design

The task considered in this research is picking and placing items on construction sites, which involves handling a wide variety of materials, such as slabs, rods, blocks and beams, etc. A simulated construction site environment was developed in Rviz, and a set of slabs and rods were placed as shown in Figure 3. The motion of panda arm was planned to pick slabs or rods from their pick up locations and place them to their respective drop off locations. The task was performed with both motion-based and joystick-based teleoperation and the usability and intuitiveness of both interfaces are discussed.

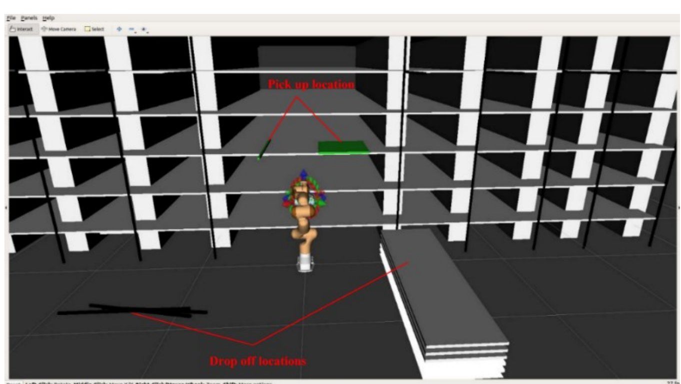


Figure 3. Simulated construction site scenario for picking and placing slabs and rods.

Prototype Demonstration

Considering the setup explained above, a prototype of the proposed teleoperation system was developed, which was based on the coordinate system shown in Figure 2 (a). The specific set of commands defined for the movement of the panda arm are explained below:

- The panda arm maintains its home position if the subject keeps his arm within the 8th octant as shown in Figure 4 (a).
- The panda arm will only move from its home position to pick up the slab when the subject places his arm within the 5th octant as shown in Figure 4 (b).
- As the subject moves his hand from the 5th octant and place in the 1st octant, the panda arm will place the slab to the drop off location as shown in Figure 4 (c).
- The panda arm will only move towards the rod and pick up the rod when the subject places his hand within the 6th octant as shown in Figure 4 (d).
- As the subject moves his hand from the 6th octant and place in the 7th octant, the panda arm will place the rod at the drop off location as shown in Figure 4 (d).
- The panda arm will return to its home position when the subject places his hand back in the 8th octant as shown in Figure 4 (e).

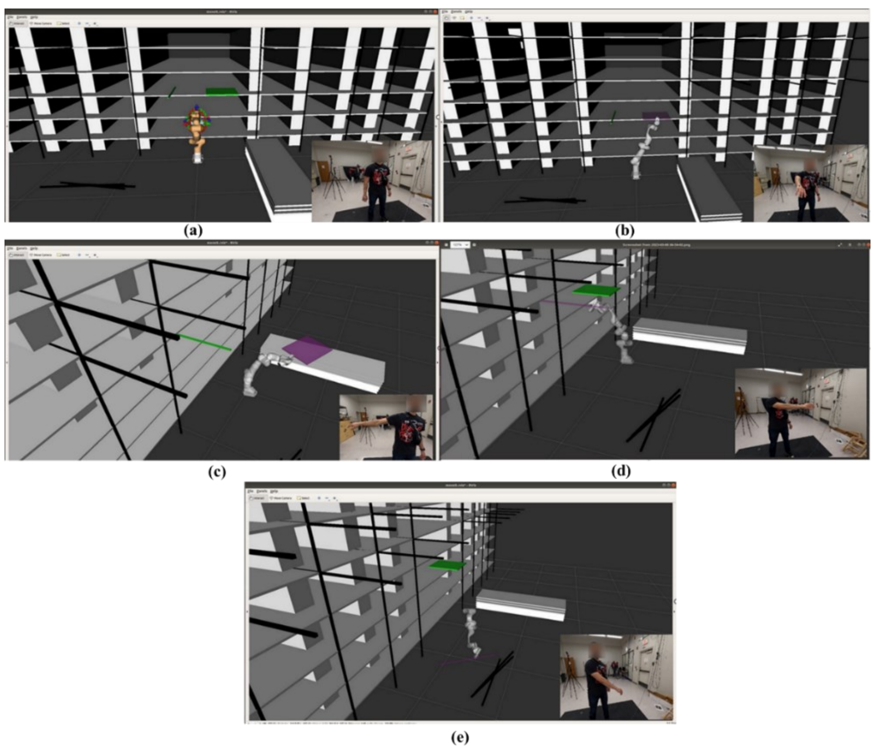


Figure 4. Demonstration of motion-based teleoperation
(a) Home position (b) Picking slab (c) Placing slab (d) Picking rod (e) Placing rod

The task of picking and placing the slab and rod was also performed with the PlayStation 3 joystick as shown in Figure 5.

Figure 5. Demonstration of joystick-based teleoperation

DISCUSSION

The demonstration shows that in terms of intuitive control and usability, the motion-based teleoperation system performs better than the joystick-based system. In contrast to the joystick-based system, the motion-based system allows the operator to control the panda arm using natural poses and movements, providing more intuitive control. The proposed motion-based teleoperation system shows high-level control as compared to the joystick. Unlike motion-based teleoperation, which has preprogrammed pick and place location and only needs a motion-based command to execute it, joystick-based teleoperation requires careful step-by-step motion planning of the panda arm, which was mentally and physically tiring. Furthermore, the proposed system is open to being integrated with other deep learning models, which can increase the precision and responsiveness of the proposed teleoperation system.

CONCLUSION

This research represented a motion-based teleoperation system, where the motion of the human arm is captured through OptiTrack cameras, and the position and location of the arm is tracked using software named motive. The tracked data is then streamed into ROS, which is coupled with MoveIt and Rviz, for the motion planning of panda arm. The simulated construction site environment is developed in Rviz, and the task of picking and placing the slabs and rods are performed with the panda arm. The intuitiveness of the proposed motion-based teleoperation system is then compared to joystick-based teleoperation system, which is widely used in the construction industry. The demonstration of both teleoperation systems in performing picking and placing tasks with panda arm showed that the motion-based interface for teleoperation is better than the joystick-based interface in terms of usability and intuitive control.

There remain some limitations that deserve future study. First, the applicability of the proposed motion-based teleoperation system is demonstrated qualitatively with only one case study in the simulated environment. In future research, a more comprehensive evaluation of the proposed system, covering aspects such as accuracy, latency, and user experience etc., will be conducted to quantitively assess the performance of the proposed interface in both virtual and physical environments. Second, simplified rules were defined to connect human pose and motion with robot motion commands, which requires extensive efforts to design the rules and may be difficult to scale to complex scenarios. For future work, the proposed system can be integrated with other deep learning models to improve planning capability and efficiency in complex construction scenarios.

ACKNOLWEDGMENTS

This research was funded by the U.S. National Science Foundation (NSF) via Grants 2138514, 2222670, 2222730, and 2222810. The authors gratefully acknowledge NSF's supports. Any opinions, findings, recommendations, and conclusions in this paper are those of the authors, and do not necessarily reflect the views of NSF, The University of Texas at San Antonio, The University of Tennessee, Knoxville, and Penn State University.

REFERENCES

Al-Mouhamed, M. A., M. Nazeeruddin and N. Merah (2008). "Design and instrumentation of force feedback in telerobotics." *IEEE Transactions on Instrumentation and Measurement* 58(6): 1949-1957.

Bao, S.-D., X.-L. Meng, W. Xiao and Z.-Q. Zhang (2017). "Fusion of inertial/magnetic sensor measurements and map information for pedestrian tracking." *Sensors* 17(2): 340.

Bertomeu-Motos, A., L. D. Lledó, J. A. Díez, J. M. Catalan, S. Ezquerro, F. J. Badesa and N. Garcia-Aracil (2015). "Estimation of human arm joints using two wireless sensors in robotic rehabilitation tasks." *Sensors* 15(12): 30571-30583.

Czarnowski, J., A. Dąbrowski, M. Maciaś, J. Główka and J. Wrona (2018). "Technology gaps in human-machine interfaces for autonomous construction robots." *Automation in Construction* 94: 179-190.

Du, G. and P. Zhang (2014). "A markerless human–robot interface using particle filter and Kalman filter for dual robots." *IEEE Transactions on Industrial Electronics* 62(4): 2257-2264.

Fang, B., F. Sun, H. Liu, D. Guo, W. Chen and G. Yao (2017). "Robotic teleoperation systems using a wearable multimodal fusion device." *International journal of advanced robotic systems* 14(4): 1729881417717057.

Galzarano, S., R. Giannantonio, A. Liotta and G. Fortino (2014). "A task-oriented framework for networked wearable computing." *IEEE Transactions on Automation Science and Engineering* 13(2): 621-638.

Gromov, B., G. Abbate, L. M. Gambardella and A. Giusti (2019). "Proximity human-robot interaction using pointing gestures and a wrist-mounted IMU." *International Conference on Robotics and Automation (ICRA)*, *IEEE* pp. 8084-8091.

Hasegawa, Y., T. Kitamura, S. Sakaino and T. Tsuji (2020). "Bilateral control of elbow and shoulder joints using functional electrical stimulation between humans and robots." *IEEE Access* 8: 15792-15799.

Hirche, S. and M. Buss (2012). "Human-oriented control for haptic teleoperation." *Proceedings of the IEEE* 100(3): 623-647.

Kim, D., J. Kim, K. Lee, C. Park, J. Song and D. Kang (2009). "Excavator tele-operation system using a human arm." *Automation in construction* 18(2): 173-182.

Köpüklü, O., A. Gunduz, N. Kose and G. Rigoll (2019). "Real-time hand gesture detection and classification using convolutional neural networks." *14th IEEE international conference on automatic face & gesture recognition (FG 2019),* IEEE.

Lu, Z., P. Huang and Z. Liu (2017). "Predictive approach for sensorless bimanual teleoperation under random time delays with adaptive fuzzy control." *IEEE Transactions on Industrial Electronics* 65(3): 2439-2448.

Macchini, M., J. Frogg, F. Schiano and D. Floreano (2022). "Does spontaneous motion lead to intuitive Body-Machine Interfaces? A fitness study of different body segments for wearable telerobotics." *31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, IEEE.

Mavridis, N., G. Pierris, P. Gallina, Z. Papamitsiou and U. Saad (2015). "On the subjective difficulty of joystick-based robot arm teleoperation with auditory feedback." *IEEE 8th GCC Conference & Exhibition*, IEEE.

Molchanov, P., X. Yang, S. Gupta, K. Kim, S. Tyree and J. Kautz (2016). "Online detection and classification of dynamic hand gestures with recurrent 3d convolutional neural network." *Proceedings of the IEEE conference on computer vision and pattern recognition.*

Moschetti, A., F. Cavallo, D. Esposito, J. Penders and A. Di Nuovo (2019). "Wearable sensors for human–robot walking together." *Robotics* 8(2): 38.

OptiTrack. (2020). "Motive Documentation - NaturalPoint Product Documentation,." Retrieved 2023, Februrary 09, from https://v22.wiki.optitrack.com/index.php?title=Motive_Documentation.

OptiTrack. (2022). "Rigid Body Tracking." Retrieved 2023, Februrary 09, from https://docs.optitrack.com/v/v2.3/motive/rigid-body-tracking.

Repnik, E., U. Puh, N. Goljar, M. Munih and M. Mihelj (2018). "Using inertial measurement units and electromyography to quantify movement during action research arm test execution." *Sensors* 18(9): 2767.

Rudigkeit, N. and M. Gebhard (2019). "AMiCUS—A head motion-based interface for control of an assistive robot." *Sensors* 19(12): 2836.

Škulj, G., R. Vrabič and P. Podržaj (2021). "A wearable imu system for flexible teleoperation of a collaborative industrial robot." *Sensors* 21(17): 5871.

statistics, U. s. b. o. l. (2022). "OSHA Data & Statistics Commonly Used Statistics." Retrieved 2023, March 10, from https://www.bls.gov/news.release/cfoi.t03.htm.

Taylor, R. M., T. C. Hudson, A. Seeger, H. Weber, J. Juliano and A. T. Helser (2001). "VRPN: a device-independent, network-transparent VR peripheral system." *Proceedings of the ACM symposium on Virtual reality software and technology*.

Truong, D. Q., B. N. M. Truong, N. T. Trung, S. A. Nahian and K. K. Ahn (2017). "Force reflecting joystick control for applications to bilateral teleoperation in construction machinery." *International Journal of Precision Engineering and Manufacturing* 18: 301-315.

Van der Kruk, E. and M. M. Reijne (2018). "Accuracy of human motion capture systems for sport applications; state-of-the-art review." *European journal of sport science* 18(6): 806-819.

Wang, X. and Z. Zhu (2021). "Vision–based framework for automatic interpretation of construction workers' hand gestures." *Automation in Construction* 130: 103872.

Xu, C., J. He, X. Zhang, X. Zhou and S. Duan (2019). "Towards human motion tracking: Multi-sensory IMU/TOA fusion method and fundamental limits." *Electronics* 8(2): 142.

Zhang, M., R. Xu, H. Wu, J. Pan and X. Luo (2023). "Human–robot collaboration for on-site construction." *Automation in Construction* 150: 104812.

Zhou, H. and H. Hu (2008). "Human motion tracking for rehabilitation—A survey." *Biomedical signal processing and control* 3(1): 1-18.