IOP Conference Series: Materials Science and Engineering

Utilising the See-and-Follow Method for Enhancing Robot Learning Ability

To cite this article: Muh Anshar *et al* 2019 *IOP Conf. Ser.: Mater. Sci. Eng.* **676** 012002

Utilising the See-and-Follow Method for Enhancing Robot Learning Ability

Muh Anshar[1]*, D Halim[1] and C Yohannes[1]

[1]Research Lab Indonesia Australia Social Cognitive Robotics, Electrical Department, Faculty of Engineering, University of Hasanuddin

*Email: muh.anshar@gmail.com

Abstract. The rapid growth and development of robots today can be reflected from variations of research in the robotic field, and one of them that has been flourishing in the area is robot learning. Learning is one of the most important aspect needs to be implemented in a robot that allows the robot to improve its performance through experience-driven knowledge. There are several forms of learning methods and one of the most favourable approach is the see-and-follow method. This paper aims to develop robots that can learn to do the same movements with human movements through the see-and-follow method. Kinect sensor is used as a technology to support the process of recognizing the human body and its movements; which later to be transformed into the robot dimension. The results show that Kinect sensor can be used to measure the distance and recognize the joints of the human body along with the movements produced within a distance of 1 meter to 3.75 meters with an average reading error of 0.04 meter. The average delay time for the whole movement is 2.6 seconds. The average percentage of similarity between robot and human movements for the whole movement scenarios is 93.45%. Hence, high level of similarity in movement concludes that the method of see-and-follow successfully increases the learning skills of the robot.

1. Introduction

According to Mitchell [1], learning is a process to improve performance through experience. Learning is one of the important aspects that needs to be implemented in a robot, which allows the robot to improve various skills through experience. This experience can be obtained either from the robot mechanism itself or from its environment [2]. Robot can be said to learn if when doing a work, its ability increases over time correspond with its experience. One method that can be used in robot learning is to see-and-follow or typically known as mimicking method.

The method of see-and-follow models human as the object to be imitated and a robot moves according to what the human model is doing. Implementation of this method requires specific sensor that has the ability to recognize human body and human motions. These motions would be the input for the robot, which latter to be imitated during mimicking process without to manually move the robot hands to follow the human motions. Through this way, the see-and-follow method makes the Human-Robot Interaction become more user-friendly where controlling robots no longer requires a gamepad, joystick, keyboard, and mouse.

This paper aims to develop a robot framework that support learning process in order to acquire a certain skill such as motions which are performed by a human model through see-and-follow method.

Kinect sensor is used as a sensor technology to support the recognition process of the human body and its movements. Testing scenarios are designed to demonstrate the learning process by see-and-follow method and to access the performance of the learning.

2. Related Studies

The see-and-follow method has been widely implemented in the robot learning as a research area in robotics. Algan [3] in 2003 introduced a robotic arm that could mimic the movements of the human arm where the position of the arm was represented through camera images. In 2014, Jonas [4] introduced a humanoid robot that can mimic the movement of the entire human body in real time by utilizing the Xsens' MVN motion capture system. In the following year, 2015, Rajesh [5] also made a robotic arm that could mimic the movements of the human arm by utilising the Kinect sensor to determine the slope angle of the arm then transform it into joint positions of the robot.

3. Methods

In this study, the overall design of the framework robot learning system is shown in Figure 1 below. It consists of one Kinect sensor, 1 PC, 2 pieces of Arduino Mega 2560, a pair of 433 MHz RF transceiver-receiver communication module, IC 74LS241 driver, and a humanoid robot.

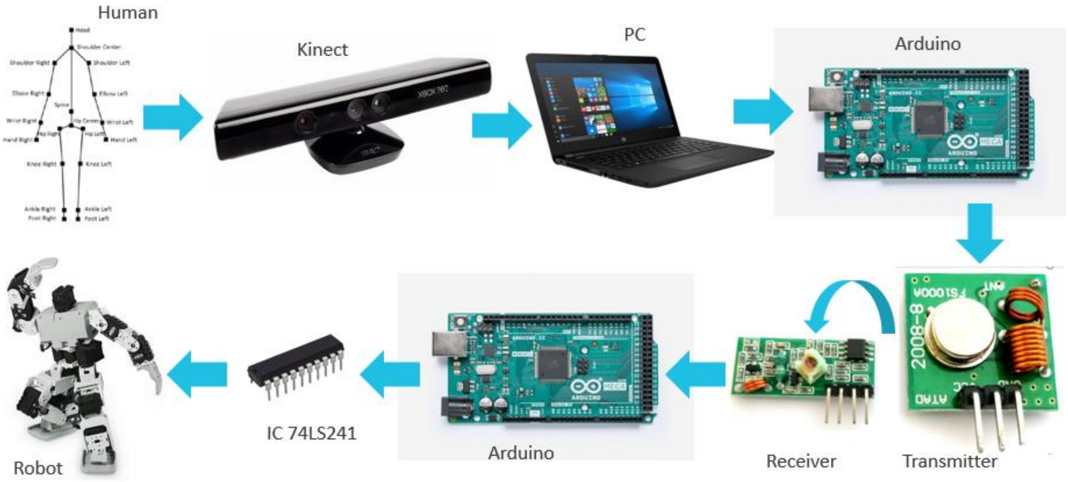

Figure 1. Overall design of robot learning system

The Kinect sensor is used for detecting the coordinates of the joint on the human body, i.e. the upper limb that constitutes to arm joints of the robot. The whole set of data is then processed by the PC for recognition phase and transformed into a sequence of joint data that associated with human arm motions. This signal is then forwarded to Arduino Mega 2560 as the local processing unit. The resulting data is sent via wireless to the robot a 433 MHz RF communication module. After the data is received, another Arduino Mega 2560 (local processor) will process the incoming data and convert the data into joint parameters, which will actuate the joint on the robot. Actuation of robot joints will produce the robot motions. The local processing at the robot side, requires an IC 74LS241, which serves as an interface that functions to sort the data whether it is addressed as Transmitter or Receiver.
The scenario of the experiment that will be conducted is the specification of the working distance from the sensor system, the minimum and maximum coordinates of 3 dimensions (x, y, z) of the left hand joint, right hand joint, and head joint, the 3 dimensional coordinates (x, y, z) of human joints when doing movement, time delay between robot movements and human movements, and the degree of similarity between robot movements and human movements. The object being tested is a human operator with a height of 175 cm standing in front of the Kinect sensor where the sensor is placed at a height of 90 cm from the floor.

3.1. Testing of sensor working distance

This experiment aims to determine the minimum and maximum working distance from the Kinect sensor to be able to read the joints of the human body. To be able to measure the distance, two types of measurement methods are used. The first is using the Kinect sensor to read the z coordinates of the operator and the second is using the meter to measure the actual distance (real) between the Kinect sensor and humans. The working distance is denoted by d measured in meters and will be compared with the Kinect working distance based on theory.

The test results are shown in Table 1.

Table 1. Sensor working distance specifications

d (real)	d (Kinect)	Joint Status	Error
0	0	not detected	0
0.25	0.3	not detected	0.05
0.5	0.54	not detected	0.04
0.75	0.79	not detected	0.04
1	1.04	detected	0.04
1.25	1.29	detected	0.04
1.5	1.54	detected	0.04
1.75	1.8	detected	0.05
2	2.04	detected	0.04
2.25	2.29	detected	0.04
2.5	2.58	detected	0.08
2.75	2.8	detected	0.05
3	3.03	detected	0.03
3.25	3.29	detected	0.04
3.5	3.54	detected	0.04
3.75	3.78	detected	0.03
4	4.04	not detected	0.04

From Table 1, the Kinect sensor is able to recognize the joint of the human body only when the operator stands between a distance of 1 meter to 3.75 meters in front of the Kinect sensor based on the actual distance calculation (real). As for the calculation using Kinect sensors, joint operators can only be identified at a minimum distance of 1.04 meters to a maximum of 3.78 meters. Overall the error reading of distance in average is about 0.04 meters.

Table 2 below shows the comparison of the sensor Kinect working distance between the theory [6] and the test results.

Table 2. Comparison of working distance between theory and test result

d min theory (meters)	d min real (meters)	Error (meters)	d max theory (meters)	d min real (meters)	Error (meters)
0.8	1	0.2	4	3.75	0.25

Figure 2 depicts the comparison of joint detection range by Kinect sensor based on the theory and test result. Based on the graph, the minimum distance required by the Kinect sensor to detect the joint operator is greater than the minimum working distance by theory. Experiment show the maximum distance of Kinect sensor work only reaches 3.75 meters which does not reach 4 meters as stated in the theory.

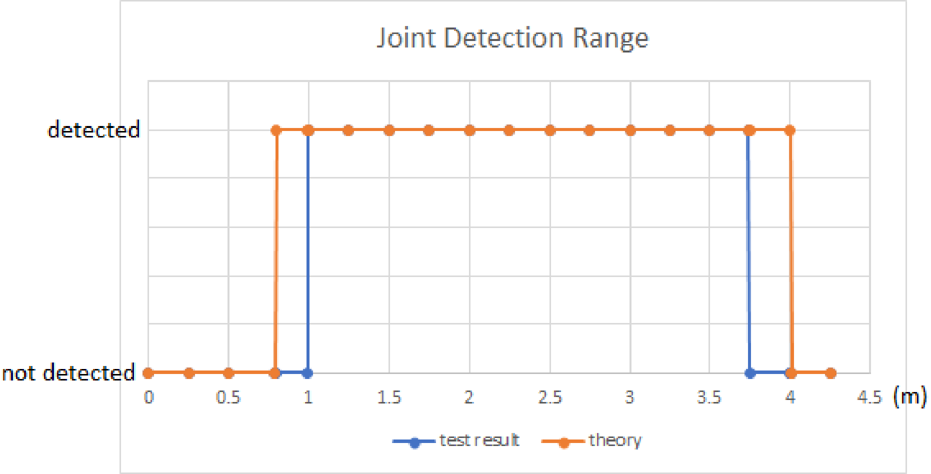

Figure 2. Joint detection range

3.2. Testing minimum and maximum 3-dimensional coordinate system of joint detection

This test aims to determine the minimum and maximum coordinates of the head joint, right hand joint, and left-hand joint of the human body that can be read by the Kinect sensor so that the operator can position themselves in front of the Kinect sensor. To be able to find out the coordinates, the measurement is done by two methods. The first is using measurements from the Kinect sensor and the second is using a meter. The Figure below shows an illustration of a 3-dimensional coordinate system (x, y, z) with detection by Kinect.

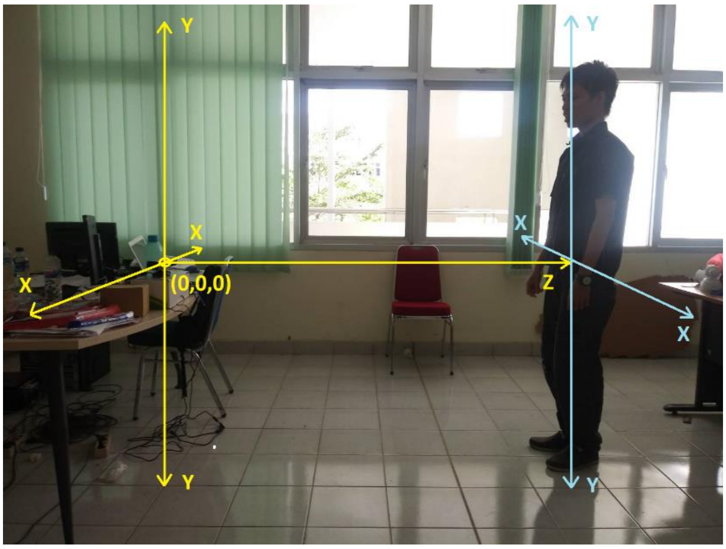

Figure 3. Illustration of 3-dimensional coordinate system

Table 4 shows that Kinect sensor can detect the right-hand joint with a maximum distance of up to 1.3 meters from the Kinect sensor (x coordinate) and left hand up to 1.32 meters from the Kinect sensor (x coordinate). The biggest error is 0.11 meters or in percentage is 9.1%.

Table 3. Minimum and maximum 3-dimensional coordinate system of joint detection (meters)

Joint	x max Kinect (meters)	x max real (meters)	Error (meters)	y max Kinect (meters)	y max real (meters)	Error (meters)
Head	0.84	0.8	0.04	0.81	0.85	0.04
Right Hand	1.3	1.36	0.06	1.1	1.2	0.1
Left Hand	1.32	1.36	0.04	1.09	1.2	0.11
Joint	z min Kinect (meters)	z min real (meters)	Error	z max Kinect (meters)	z max real (meters)	Error (meters)
Head	1.04	1	0.04	3.8	3.7	0.1
Right Hand	1.01	1	0.01	3.78	3.7	0.08
Left Hand	0.99	1	0.01	3.75	3.7	0.05

3.3. Testing coordinates of joint when doing the movement

This test aims to determine the coordinates of the head joint, right-hand joint, and left-hand joint of the human body when doing the movement. The movement includes straighten hands forward, raise hands up, lowering hands down, and stretching hands. Coordinate testing is done by Kinect sensor calculation where the results are shown in the Table 4 below.

Table 4. 3-Dimensional coordinates of joint when doing the movement

Movement	Head (meters)			Right-Hand (meters)			Left-Hand (meters)		
	x	y	z	x	y	z	x	y	z
Lowering hands down	-0.11	0.4	2.06	0.07	-0.42	1.89	-0.29	-0.41	1.91
Stretching hands	-0.32	0.4	2.14	0.4	0.32	2.1	-1.04	0.23	2.05
Raise hands up	-0.3	0.41	2.13	-0.13	0.79	2.05	-0.42	0.78	2.01
Straighten hand forward	-0.34	0.36	2.19	-0.18	0.32	1.6	-0.43	0.35	1.61

From Table 4, it can be concluded:

- When lowering hands down, the y coordinates of the right-hand joint and the y coordinates of the left-hand joint are smaller than zero
- When stretching hands, the x coordinates of righthand joints are greater than the x coordinates of head plus 0.45 meters and the x coordinates of the left-hand joint x are smaller than the x coordinates of head minus 0.45 meters
- When raising hands up, both y coordinates of right-hand and left-hand joint are bigger than the y coordinates of head
- When stretching hand forward, both z coordinates of right-hand and left-hand are bigger than the coordinates of head plus 0.45 meters

3.4. Testing time delay between human and robot movement

This test aims to find out how long it takes the robot to respond to the movements of the operator. From the Table 5 below, the smallest delay required by the robot in order to respond to the movement of the operator is 2 seconds and the biggest delay time is 4 seconds.

Table 5. Time delay between human and robot movement

Human movement	Robot movement	Time delay for experiment				
		1	2	3	4	5
Lowering hands down	Lowering hands down	2	3	2	3	2
Stretching hands	Stretching hands	2	3	3	3	2
Raising hands up	Raising hands up	3	3	2	2	2
Stretching hands forward	Stretching hands forward	2	3	4	2	4

Figure 4 below shows the average time delay between the motion of the robot and the movement of the operator for each movement.

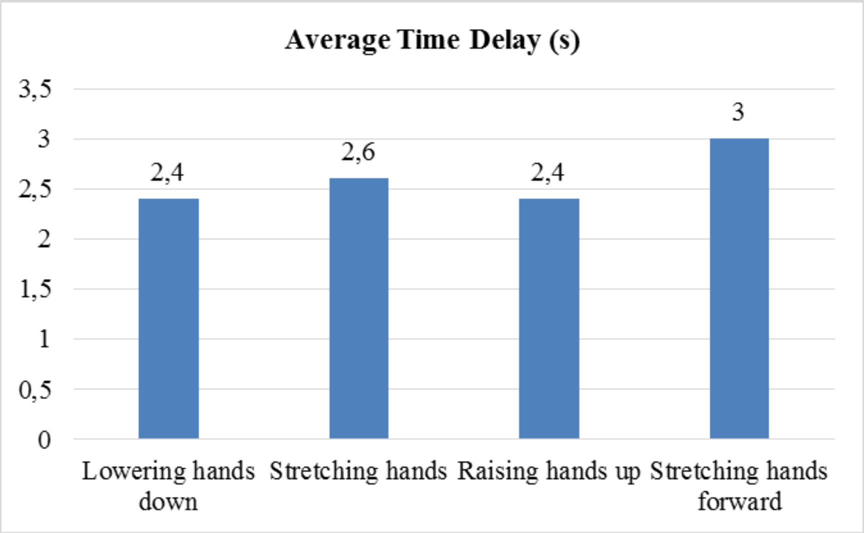

Figure 4. Average time delay for each movement

3.5. Testing the similarity between robot and human movement

The testing process is conducted by giving the opportunity to several respondents to test the robot that has been made. After testing, the questionnaire will be distributed to respondents to find out their opinions about how much the percentage of similarity of the movements produced by the robot to the movements they do.

Table 6. Percentage of similarity between human and robot movement

Respondent	Lowering hands down (%)	Stretching hands (%)	Raising hands up (%)	Stretching hands forward (%)
1	90	90	90	90
2	99	99	99	99
3	98	98	98	98
4	95	95	95	95
5	100	100	100	90
6	80	80	80	80
7	90	100	100	100
8	10	100	100	100
9	95	95	95	95
10	70	90	100	100
11	90	100	100	100
12	95	90	95	95
13	90	100	100	100
14	85	85	85	85
15	85	85	86	88

The average percentage of similarity between robot and human for each movement is above 90% where overall the average is 93.45%

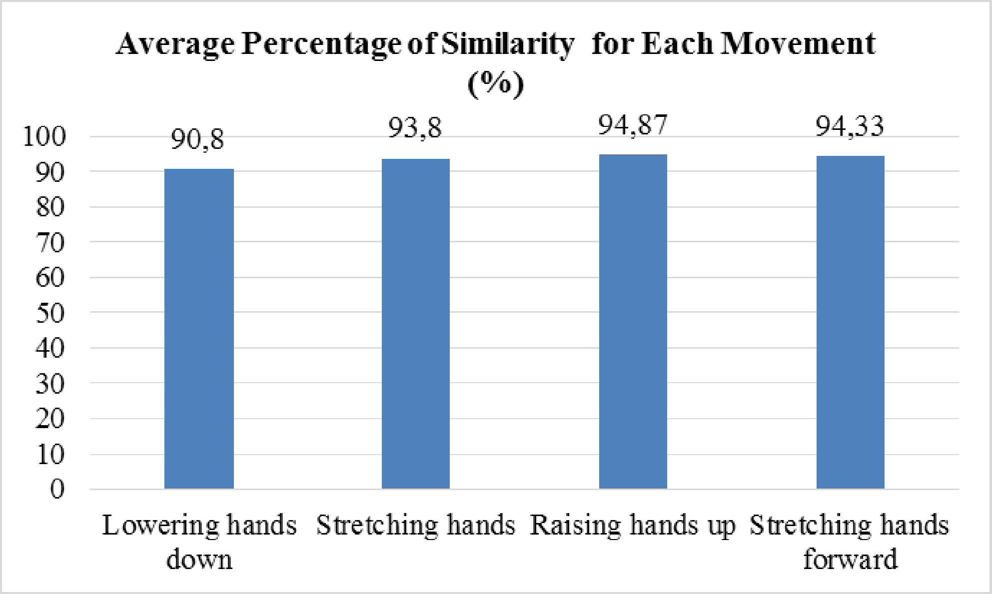

Figure 5. Average percentage of similarity between human and robot movement

4. Conclusions

The test results show that the Kinect sensor can be used to measure the distance and recognize the joints of the human body along with the movements produced within a distance of 1 meter to 3.75 meters with reading error in average is about 0.04 meters. While on the x-axis, the detection range of joint by Kinect sensor is 1.3 meters to the right and 1.32 meters to the left. The smallest delay required by a robot to respond to the movement of the operator is 2 seconds and the biggest delay is 4 seconds where the average delay time for the whole movement is 2.6 seconds. The average percentage of similarity between robot and human movements for the whole movement is 93.45%. Due to the high level of similarity in movement, it can be concluded that the method of see and follow successfully increases the ability of learning robots.

Acknowledgments

Our appreciation to all members of the Indonesia-Australia Cognitive and Social Robotics Research Lab for participation and support throughout the project. Authors also would like to thanks the Faculty of Engineering UNHAS for providing partial funds for the project under the LBE project funding scheme.

References

[1] Mitchell T 1997 *Machine Learning* (Singapore: McGraw-Hills International Editions).
[2] Anshar M 2003 *Directed Learning-based Evolutionary Approach for Legged Robot Motion* (Australia: Lambert Academic Publishing).
[3] Uskarci A, Alatan A, Dindaroglue S 2003 *Robot Mimicking: A Visual Approach for Human Machine Interaction Proc., Symp., Int., Conf., 2nd*.
[4] Koenemann J, Burget F, Bennewitz M 2014 *Real-time imitation of human whole-body motions by humanoids*, pp. 2806-2812.
[5] Rajesh K, Deepansh M, Nitin A, Nihil S 2015 *Implementation of Gesture Control in Robotic Arm using Kinect Module (Applied Mechanics and Materials* vol 786 p 378-382.
[6] Abhijit J 2012 *Kinect for Windows SDK Programming Guide* (UK: Packt Publishing).