Performance evaluation of an inverse kinematic based control system of a humanoid robot arm using MS Kinect

Ha Xuan Nguyen[2], Hung Chan Nguyen[1], Ngoc-Anh Mai[3], Hung Trung Nguyen[1], Duc Xuan Tran[1], Lam Bao Dang[2], Hai Minh Pham[2] and Luong Nguyen Dinh[1]

[1] Vietnam Research Institute of Electronics, Informatics and Automation, 156A Quan Thanh str., Hanoi, Vietnam
[2] School of Mechanical Engineering, Hanoi University of Science and Technology, 1 Dai Co Viet str., Hanoi, Vietnam
[3] Advanced Technology Center, Le Quy Don Technical University, 236 Hoang Quoc Viet str., Hanoi, Vietnam
Email: hungnc@vielina.com, hungnc@gmail.com

***Abstract* — This paper presents an inverse kinematic based control system of a novel self-developed humanoid robot using a Microsoft Kinect as 3D sensor. Based on the processed sensory data from the Kinect, human gestures and motions are captured, analyzed and mapped into joint variables by using inverse kinematic equations. A significant improvement in the inverse kinematic algorithm has been made to achieve better perception of human gestures. The performance of the proposed system is evaluated in a newly developed test-bed under real-world conditions, which effectively shows perception-related error control in order to further improve the system.**

Keywords—humanoid robot; kinect camera; inverse kinematic solution; robot arm; VIEBOT

I. INTRODUCTION

Humanoid robot is a robot shaped like human with basic components such as head, neck, torso and arms. Normally, humanoid robots are equipped with the ability to observe the surrounding environment using cameras as well as avoiding obstacles using sensors acting as human senses such as hearing, seeing and touching. This is the reason why humanoid robots have excellent perceptions about the working environment.

In addition, humanoid robots are capable of performing many complex operations thanks to their ingenious hands. In other words, humanoid robots can perform various sophisticated tasks because they are able to mimic complex human behaviors such as turning on/off light switches, opening/closing doors, and pushing/pulling carts. Therefore, so many humanoid robots have been developed for different life areas such as health care, education and training, or advertising services [1, 2], or "… helping elderly people live in their own homes to responding to disasters" [3].

Kinect is a device with the ability to recognize gestures and motions through captured images taken from a RGB camera and a depth sensor. The new version of Kinect, MS Kinect V2, has a striking feature of reproducing high definition of RGB and depth images with low cost. Thanks to this, the device is utilized for tracking human gestures so that the robot can imitate human gestures in real time or play back recorded gestures.

Based on the results of the process of kinetic image regeneration of the limb's skeletal structure, the system determines positions of human hand joints that, in turn, are used with inverse kinematic equations to control movements of the robot's hand joints accordingly.

The basic tasks of humanoid robot control are normally divided into two main groups: i) Control of movements for moving in a given or unknown terrain and for avoiding obstacles; ii) Control of flexible hands and arms for getting a job done. *In this article, the authors focus on explaining the group ii on a humanoid robot named VIEBOT*. This robot uses a Kinect as the main sensor for recognizing and tracking human objects.

To control the robot arms, inverse kinematic and dynamic equations must be solved. In case of high-level flexibility of the arms, it would be difficult for the control system to solve those equations. In fact, if the movement speed of the arm is low, dynamic effects can approximately be neglected. Thus, the inverse kinematic solution is usually used to calculate the joint variables. There are three main methods for solving the inverse kinematic equations, including interactive method [4, 5], geometric method [6, 7, 8] and inverse-transformation method [9, 10]. Due to the mechanical structure and joint variable constraints, multiple solutions are always complicated issues and there has not been any commonly used method able to apply for different kinematic structures of the arm yet.

To detect the motions and gestures of a tracked human, two methods can be applied for VIEBOT: i) Using a motion capture that suits having several angular position sensors for the tracked person to measure the joint angles as in [11]; ii) Using a 3D camera that converts motions and gestures into skeleton points, and then maps them to joint variables through an inverse kinematic solution [12, 13, 14, 15]. In reality, the second method is used more often. However, the data collected by the 3D camera is just exploited to calculate joint variables. To the author's knowledge, the use of the data collected by the camera for solving multiple-solution-related problems has not been published yet.

In this research, we use the method presented in [10] as an inverse kinematic solution with a modification of extension on

978-1-5386-3742-5/17/$31.00 © 2017 IEEE

the constraints of joint variables suitable to VIEBOT's real mechanical structure. The combination of the arm and the hand movements are inferred from motion and gesture recognitions resulted from a MS Kinect V2. The collected data is not only used as the inputs of inverse kinematic equations, but also exploited to choose a desired solution in case of having multiple ones.

This paper is organized as follows: The next section describes the architecture and several new features of VIEBOT. In the third section, the inverse kinematic control mechanism with our new extension is presented. In the fourth section, the experiments in a real environment to evaluate the quality of the control system are illustrated along with the results and discussions on some quality characteristics of the system. The final section is the conclusion with future research directions.

II. VIEBOT ARCHITECTURE

VIEBOT is a humanoid robot developed by the Vietnam Research Institute of Electronics, Informatics and Automation (http://vielina.com). The VIEBOT's 3D model and its actual skeleton are shown on the left and right side of Fig. 1, respectively. Its architecture is illustrated on Fig. 2.

In Vietnam, VIEBOT is considered as a new robot generation inherited from several Industry 4.0 technologies such as Cloud Analytics, Micro Service-based software architecture, and 3D printing.

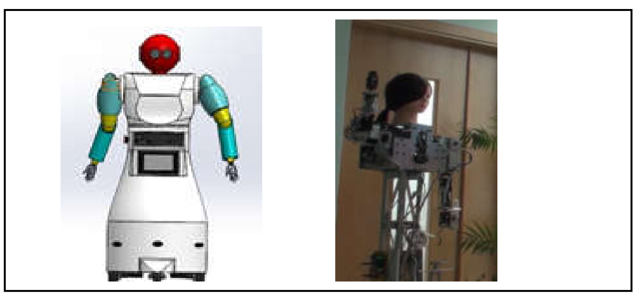

Fig. 1. VIEBOT: 3D design and skeleton.

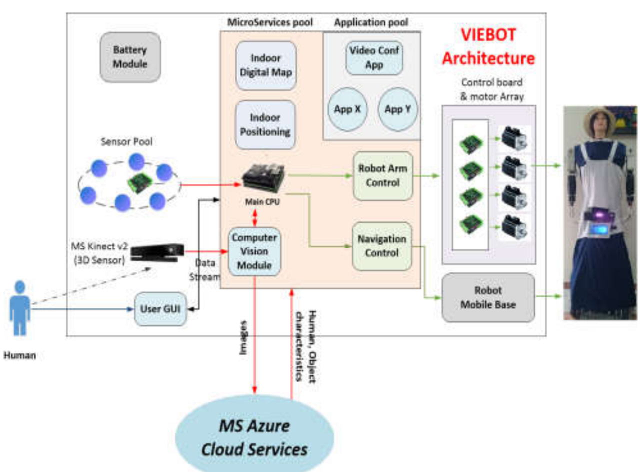

Fig. 2. VIEBOT architecture.

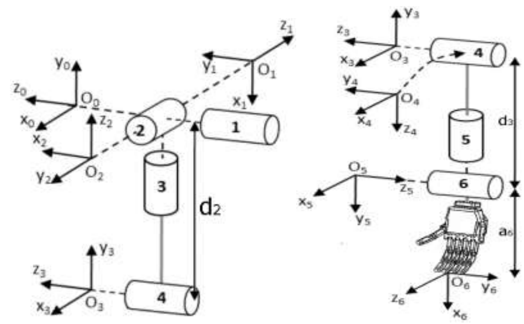

Fig. 3. Coordinate systems of VIEBOT arm and hand.

Joint	θ_i	α_i	a_i	d_i
1	q_1-90°	90°	0	0
2	q_2+90°	-90°	0	0
3	q_3+90°	90°	0	$-d_2$
4	q_4	90°	0	0
5	q_5	90°	0	d_3
6	q_6+90°	90°	a_6	0

Fig. 4. Kinematic parameters of VIEBOT arm.

VIEBOT is designed with two arms and hands. Each arm has six degrees of freedom (DOF), while the hand has one DOF to hold things. The arm consists of 3 clusters: Shoulder; Upper arm and elbow; Forearm, wrist, and five-finger hand. The coordinate systems of the arms and hands are shown in Fig. 3. The Denavit – Hartenberg (DH) parameter table of the arm and hand is shown in Fig. 4 used as the basis for forward and inverse kinematic calculations, where:

- θ_i is an angle between axes $x_{(i-1)}$ and x_i about axis $z_{(i-1)}$.
- α_i is an angle between axes $z_{(i-1)}$ and z_i about axis x_i.
- a_i is a distance between $O_{(i-1)}$ and O_i along x_i.
- d_i is a distance between $O_{(i-1)}$ and O_i along $z_{(i-1)}$.

One of VIEBOT's fundamental tasks is to mimic gestures of human actions. For this reason, VIEBOT has to recognize the human gestures and then calculate DH parameters to produce corresponding motions of the arm and hand.

In addition, VIEBOT is designed to be market-oriented, so it is flexible enough to be transformed into other kinds of robots with different functions such as indoor mobile cars for hospitals, service robots, advertising robots or educational robots with low-cost. For the first purpose of being an advertising robot, it has a similar shape to a young woman with 1.6m height and 55kg weight including batteries. The brain of the robot contains an Intel core i7 Mini-PC with 16GB RAM running MS Windows 10. The robot connects with MS Azure cloud through 4G LTE mobile network to use advanced Microsoft Azure cloud services including Face API, Computer Vision and Machine learning that support VIEBOT to recognize human face, gender, age and surrounding objects. Thanks to its high speed wireless connection, VIEBOT can also be utilized as a mobile high definition Video & Audio conferencing system that allows remote users to interact with people in VIEBOT's zone of operation.

III. Inverse Kinematic Control Mechanism

Figure 5 shows the principle diagram of the inverse kinematic control system of VIEBOT. The system consists of four functional modules: Identification of Joint Pose - IJP, Computation of Inverse Kinematics - CIK, External Quality Assessment - EQA and Joint motion controller - JMC.

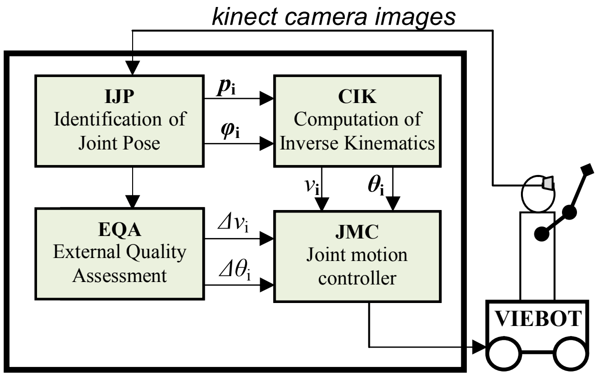

Fig. 5. Inverse kinematic control system.

In our research, the robot has two arms each of which has six degrees of freedom. Because of the real time operating condition of the arms, a closed-form solution for the joint variables is of advantage. Thus, we use the inverse-transformation method proposed in [10] to find out closed-form solution for the joint variables. Additionally, since there are multiple solutions, it is vital to choose the best solution fitted with the real mechanical structure of the robot arm. Here, the data obtained from Kinect camera is used as reference parameter for choosing the solution.

The invese kinematic equation has the form:

$$T={}^{0}A_1.{}^{1}A_2.{}^{2}A_3.{}^{3}A_4.{}^{4}A_5.{}^{5}A_6=\begin{bmatrix} n_x & s_x & a_x & p_x \\ n_y & s_y & a_y & p_y \\ n_z & s_z & a_z & p_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (1)$$

where the left-hand side is the multiplication of transformation matrixes obtaining joint variables q_i $(i=1\div 6)$ and the right-hand side is the matrix representing the position and orientation of the hand.

Due to the real mechanical structure of the arm and anthropomorphic motion, the joint variables are constrained to

$$\begin{array}{ll} 0^0 \le q_1 \le 180^0 & 0^0 \le q_4 \le 120^0 \\ 0^0 \le q_2 \le 180^0 & -90^0 \le q_5 \le 90^0 \\ -90^0 \le q_3 \le 90^0 & -90^0 \le q_6 \le 90^0 \end{array} \quad (2)$$

For the position and orientation of the hand (end-effector), we proposed to use Kinect to get coordination of three points on the hand D_6, D_7, D_8 as shown in Fig. 6. Based on this data, the end-effector matrix is determined as follows:

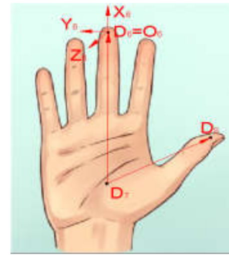

Fig. 6. Coordinate system for determing position and orientation of the hand.

- The position O_6 is identical to D_6
- The orientation unit vector of the coordinate system $O_6x_6y_6z_6$ is calculated from the three points

$$\begin{cases} \overrightarrow{O_6x_6}=\overrightarrow{D_7D_6} \\ \overrightarrow{O_6z_6}=\overrightarrow{D_7D_8}\wedge\overrightarrow{D_7D_6} \\ \overrightarrow{O_6y_6}=\overrightarrow{O_6z_6}\wedge\overrightarrow{O_6x_6} \end{cases} \quad (3)$$

- Projecting this vector system to the coordinate system $O_0x_0y_0z_0$ attached to the arm's base, we obtain the end-effector matrix

$$T=\begin{bmatrix} n_x & s_x & a_x & p_x \\ n_y & s_y & a_y & p_y \\ n_z & s_z & a_z & p_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \Rightarrow T^{-1}=\begin{bmatrix} n'_x & s'_x & a'_x & p'_x \\ n'_y & s'_y & a'_y & p'_y \\ n'_z & s'_z & a'_z & p'_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4)$$

We transform the kinematic equation as follows:

$$T={}^{0}A_6 \Rightarrow T^{-1}={}^{0}A_6^{-1} \Rightarrow {}^{5}A_6.T^{-1}={}^{0}A_5^{-1} \quad (5)$$

The left-hand and right-hand sides of (5) are respectively deployed in (6) and (7) as follows:

$${}^{5}A_6.T^{-1}=\begin{bmatrix} g_{511} & g_{512} & g_{513} & p'_zc_6-(a_6+p'_x)s_6 \\ g_{521} & g_{522} & g_{523} & p'_zs_6+(a_6+p'_x)c_6 \\ g_{531} & g_{532} & g_{533} & p'_y \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (6)$$

$${}^{0}A_5^{-1}=\begin{bmatrix} g_{511} & g_{512} & g_{513} & d_2c_5s_4 \\ g_{521} & g_{522} & g_{523} & -d_3-d_2c_4 \\ g_{531} & g_{532} & g_{533} & d_2s_4s_5 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (7)$$

Compare elements in matrixes of (6) and (7) we have:

$$\begin{cases} p'_zc_6-(a_6+p'_x)s_6=d_2c_5s_4 \\ p'_zs_6+(a_6+p'_x)c_6=-d_3-d_2c_4 \\ p'_y=d_2s_4s_5 \end{cases} \quad (8)$$

$$\Rightarrow c_4 = \frac{r^2 + p_y'^2 - d_2^2 - d_3^2}{2d_2d_3} \Rightarrow s_4 = \pm\sqrt{1-c_4^2} \quad (9)$$

$$\text{Because } 0^0 \le q_4 \le 120^0 \text{ so } s_4 = \sqrt{1-c_4^2} \Rightarrow q_4 \quad (10)$$

substitute q_4 into (8) we have:

$$s_5 = \frac{p_y'}{d_2 s_4} \Rightarrow c_5 = \pm\sqrt{1-s_5^2} \quad (11)$$

with $-90^0 \le q_5 \le 90^0$ we obtain

$$c_5 = \sqrt{1-s_5^2} \Rightarrow q_5 \quad (12)$$

substitute q_4, q_5 into (8):

$$\begin{cases} s_{\psi 6} = \dfrac{d_2 c_5 s_4}{r} \\ c_{\psi 6} = \dfrac{d_3 + d_2 c_4}{r} \end{cases} \Rightarrow q_6 \quad (13)$$

Similarly, to determine the *l* unknown joint variables we transform (5) to:

$$T = {}^0A_6 \Rightarrow T^{-1} = {}^0A_6^{-1} \Rightarrow {}^3A_6 . T^{-1} = {}^0A_3^{-1} \quad (14)$$

Compare the elements of matrixes of the two sides of (14):

$$s_2 = -g_{423} \Rightarrow c_2 = \pm\sqrt{1-s_2^2} \quad (15)$$

To choose a suitable solution in this case, we propose to use the kinect to detect the position of the shoulder D_0 and elbow D_4 as the deciding factor to achieve the desired solution.

$$\text{If } y_{D_4} \ge y_{D_0} \Rightarrow 90^o \le q_2 \le 180^o \text{ then } c_2 = \sqrt{1-s_2^2} \quad (16)$$

$$\text{If } y_{D_4} \le y_{D_0} \Rightarrow 0^o \le q_2 \le 90^o \text{ then } c_2 = -\sqrt{1-s_2^2} \quad (17)$$

From (14) we can also obtain q_1, q_3 :

$$\begin{cases} -c_2 s_3 = g_{413} \\ c_2 c_3 = g_{433} \end{cases} \Rightarrow q_3 \quad (18)$$

$$\begin{cases} -c_2 s_1 = g_{421} \\ c_1 c_2 = g_{422} \end{cases} \Rightarrow q_1 \quad (19)$$

IV. Experiments and Results

The experiments are conducted in a typical office environment with office furniture and equipment, where people moving freely around the room. The experiments concerning Kinect image processing are carried out under normal office LED-lighting conditions.

A. *The objectives of the experiments*

The main objective is to evaluate the performance of the whole process from detecting human gestures to controlling the robot arm to improve the accuracy of the process of imitating human gestures. Among various performance metrics, such as accuracy, delay, jitter, etc, we only focus on the accuracy in this stage.

To achieve this target, a test-bed is setup to collect the data of movements from both a human and the robot that mimics the human under the same conditions to compare the results.

B. *Test-bed setup*

The test-bed, shown in Fig. 7, includes VIEBOT, an external MS Kinect to analyze both the human and the robot movements in the same conditions and a PC to run some customized software modules of the robot itself for experimental purposes.

Microsoft Kinect Studio is used as *Data Collector Tool* to store human gestures in a database. This data used for controlling VIEBOT's arms & hands later so that the real-world situation of the robot mimicking the human is precisely emulated.

The *Robot Arm Controller* module (developed in C# on MS Visual Studio 2013) of VIEBOT is customized to perform the following tasks:

1) Getting data stream from MS Kinect Studio; 2) Calculating angles of the arm through inverse kinematic equations; 3) Controlling robot arms & hands; 4) Exporting data to text files in a *File Pool* (for the purpose of collecting test data in experiments).

From the file pool, data can be matched and processed either manually or automatically by automatic scripts to form and update comparative graphs.

C. *Experimental implementation*

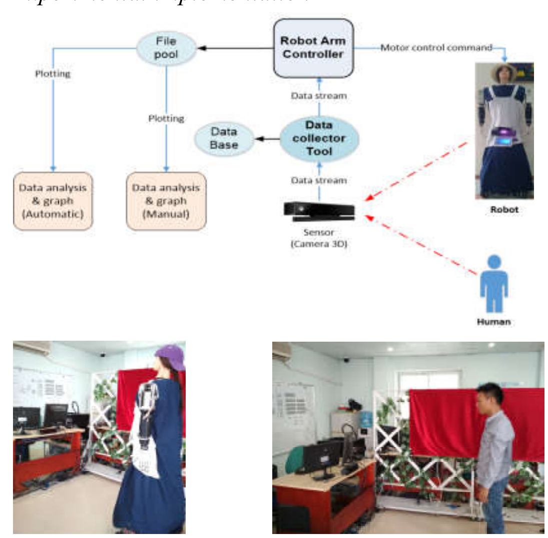

Fig. 7. Experiments in a real environment: Test-bed (above), VIEBOT (left below) and person (right below).

The experimental procedure consists of four steps:

Step 1: Run MS Kinect Studio to collect the data of movements from a person standing in front of the Kinect, and store it in a database of *Xef* format.

Step 2: Transfer the database to VIEBOT's control system where it is processed to control the robot hands & arms, while using MS Kinect Studio on the PC to capture the robot motion data and export it into another database file.

Step 3: Run the customized *Robot Arm Controller* software on the PC using the human gesture database (from step 1) as the input to export a human data text file and then use the robot gesture database (from step 2) to export a robot data text file.

Step 4: Compare the two text files, merge and analyze them using MS Excel. Some parts of this process can be automated using VBA scripts.

The output of these steps is the differences of the human and robot movements, e.g. perception-related errors, are shown in Fig. 8 for the shoulder perception, in Fig. 9 for the elbow perception and in Fig. 10 for the wrist perception. One of the advantages of this method is that it can eliminate several sources of perception-related errors from the testing environment because both datasets might be influenced by several factors such as ambient noise.

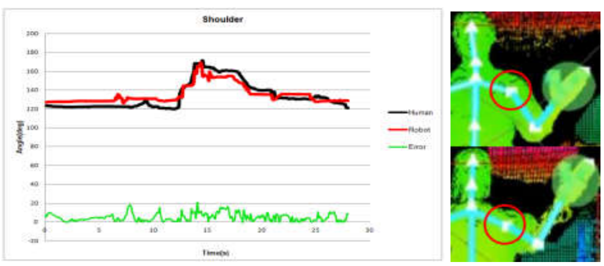

Fig. 8. The difference between VIEBOT's shoulder and human shoulder.

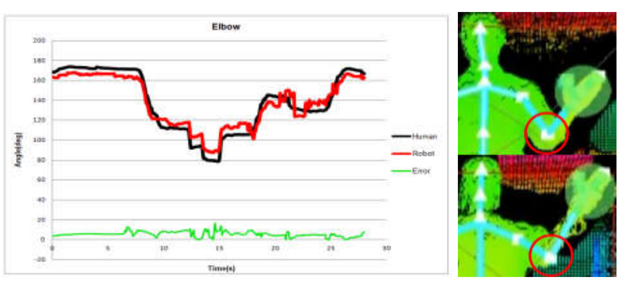

Fig. 9. The difference between VIEBOT's elbow and human elbow.

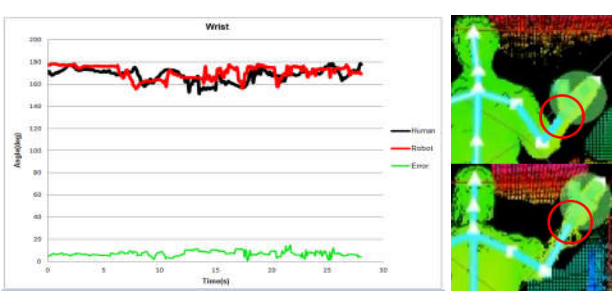

Fig. 10. The difference between VIEBOT's wrist and human wrist.

Base on the errors, their mean errors are computed at 4,32%, 5,73% and 3,12%, respectively. The mean errors demonstrate the effectiveness of our inverse kinematic control algorithm.

Fig. 11 to Fig. 13 show the dependence of the errors on the input angles and Fig. 14 compares all these dependencies.

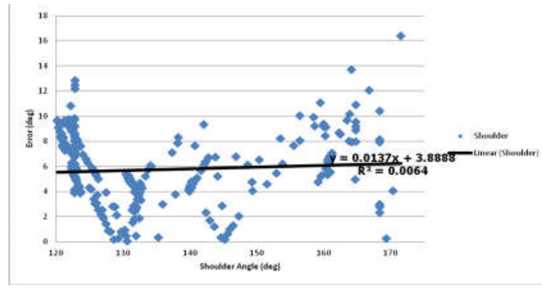

Fig. 11. The relationship between the errors and the shoulder angles.

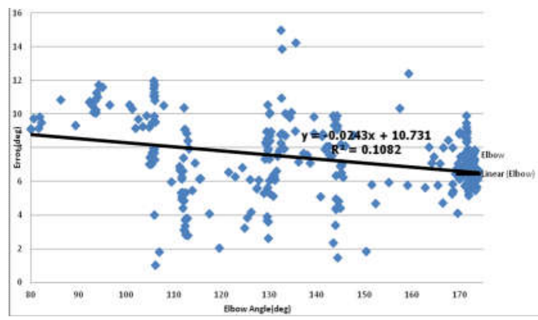

Fig. 12. The relationship between the errors and the elbow angles.

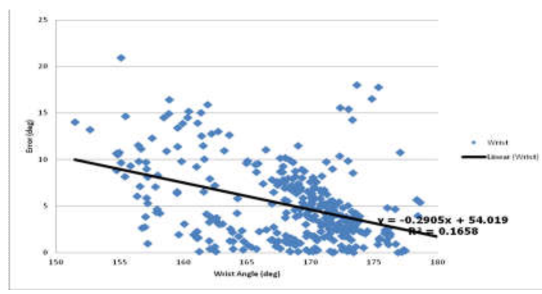

Fig. 13. The relationship between the errors and the wrist angle.

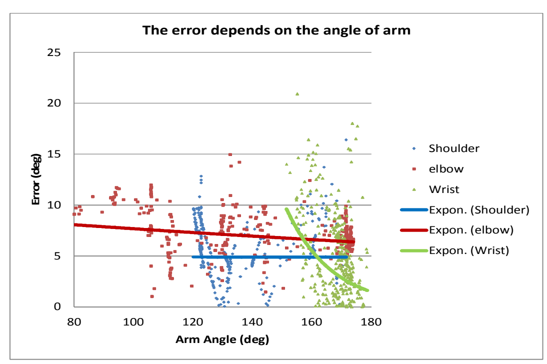

Fig. 14. Comparison of the dependencies of error on input angle of shoulder, elbow and wrist.

From our study, there are three main sources of the errors including: 1) Mechanical errors come from looseness at the joints (couplings, cone gear transmission, or belt transmission); 2) Image processing errors may result from the Kinect image with ambient light noises; 3) Operation errors come from the robot arm vibrations which occur because inertia may lead to fluctuations in the collected data.

From Fig. 14, by observing the three trend lines of the shoulder, elbow and wrist, it is possible to conclude that the dependence on the input angles of wrist joint errors is the strongest, while that of shoulder joint errors is the weakest.

D. Error histogram assessment

Stemming from the experiment results, a histogram of shoulder errors is formulated and shown in Fig. 14. It is noticeable that the highest frequency area of shoulder errors is within the range from 6 to 7 degree. Based on these results, in the future we will implement the mechanism to minimize the errors in this zone.

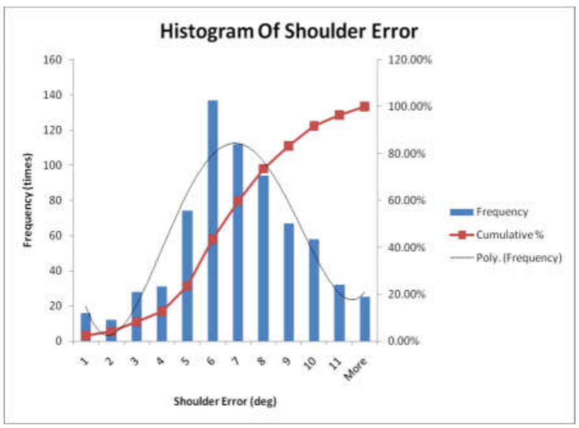

Fig. 15. Histogram of shoulder error.

V. Conclusion

This paper presents the structure, robotic paradigm and performance evaluation of an inverse kinematic controller for a humanoid robot namely VIEBOT. Utilizing a MS Kinect V2, VIEBOT has a good ability to detect, track a human and imitate his/her arm gestures by using inverse kinematic solutions. Some experiments in an office environment with humans and office furniture have proven the feasibility of the control mechanism. It is possible to claim that we have successfully developed a new humanoid robot prototype exploiting inverse kinematic mechanism for controlling robot arms and hands.

The effective test-bed for evaluating the performance of the control mechanism is illustrated and can be used for further improvements of the robot arm control algorithms.

In the future, we will improve the algorithms to mimic more complex gestures and faster activities. Besides, we will optimize the mechanical design to improve the precision of joint movement and use fuzzy logic in adaptive control to increase the flexibility with prioritized parameters.

Acknowledgment

The authors express their appreciation to the Ministry of Industry and Trade (MoIT) and the project World Bank FIRST (Fostering Innovation through Research, Science and Technology) for their financial supports.

References

[1] Riadh Zaier, "*The Future of Humanoid Robots - Research and Applications*", Text book, ISBN 978-953-307-951-6, 310 pages, Publisher: InTech, 2012. (Edited Volume)

[2] Youtube link: https://www.youtube.com/watch?v=CKW-CtyzzUk#t=1015.067468

[3] E. Guizzo and E. Ackerman, "DARPA Robotics Challenge: Here Are the Official Details", Retrieved from http://spectrum.ieee.org/automaton/robotics/humanoids/darpa-robotics-challenge-here-are-the-official-details.

[4] G. Tevatia and S. Schaal, "*Inverse kinematics for humanoid robots*," in Proc. IEEE International Conference on Robotics and Automation (ICRA), vol. 1, 2000, pp. 294–299.

[5] J. Wang and Y. Li, "*Inverse kinematics analysis for the arm of a mobile humanoid robot based on the closed-loop algorithm*," in Proc.of the IEEE Information and Automation (ICIA), 2009, pp. 516–521.

[6] T. Asfour and R. Dillmann, "Human-like motion of a humanoid robot arm based on a closed-form solution of the inverse kinematics problem", in Proceeding of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003), 2003.

[7] J. I. Zannatha and R. C. Limon, "*Forward and inverse kinematics for a small-sized humanoid robot*," in Proc. of the IEEE Intl. Conf. on Electrical, Communications, and Computers, 2009, pp. 111–118.

[8] Yuting Wang and Panagiotis Artemiadis, "Closed-Form Inverse Kinematic Solution for Anthropomorphic Motion in Redundant Robot Arms" Journal of Advances in Robotics & Automation, Vol. 2, Issue 3, 2013.

[9] R. P. Paul, B. E. Shimano, and G. Mayer, "*Kinematic control equations for simple manipulators*," in IEEE Transactions on Systems,

[10] Muhammad A. Ali, H. Andy Park, and C. S. George Lee, "Closed-Form Inverse Kinematic Joint Solution for Humanoid Robots", The 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems October 18-22, 2010, Taipei, Taiwan.

[11] Christopher Stanton, Anton Bogdanovych and Edward Ratanasena, "*Teleoperation of a humanoid robot using full-body motion capture, example movements, and machine learning*" in Proceedings of Australasian Conference on Robotics and Automation, 3-5 Dec 2012, Victoria University of Wellington, New Zealand.

[12] Halit Bener Suay and Sonia Chernova, "*Humanoid robot control using depth camera*", in Proceedings of the 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 2011.

[13] Weibo Song, Xianjiu Guo and Fengjiao Jiang, "Teleoperation Humanoid Robot Control System Based on Kinect Sensor", in Proceedings of the 4th International Conference on Intelligent Human-Machine Systems and Cybernetics, 2012.

[14] Ismail Almetwally and Malik Mallem, "Real-time tele-operation and tele-walking of humanoid Robot Nao using Kinect Depth Camera", in Proceedings of the 10th IEEE international conference on networking, sensing and control (ICNSC), 2013.

[15] Rizqa Afthoni, Achmad Rizal and Erwin Susanto, "Proportional derivative control based robot arm system using Microsoft Kinect", in Proceedings of the International Conference on Robotics, Biomimetics, Intelligent Computational Systems, 2013.