2012 IEEE/SICE International Symposium on System Integration (SII)
Kyushu University, Fukuoka, Japan
December 16-18, 2012

Full-Body Imitation of Human Motions with Kinect and Heterogeneous Kinematic Structure of Humanoid Robot

Van Vuong Nguyen, Joo-Ho Lee, *Member, IEEE*

***Abstract*—In this work, we propose a system that has the ability to reproduce imitated motions of human during continuous and online observation with a humanoid robot. In order to achieve this goal, the problems for imitation have to be solved. In this paper, we treat two main issues. One is mapping between different kinematic structures and the other is computing humanoid body pose that satisfies the static stability generated from the human motion obtained by visual motion capture of the humanoid. The experimental results based on Webots simulation and subsequent execution by a Darwin-OP humanoid robot show the validity of the proposed system in this paper.**

I. Introduction

The interaction between robots and humans has become a key research topic in robotics. A humanoid is based on the general structure of a human, but it does not need to be seen just as a real human. A humanoid robot has potentials to support people in various environments such as homes, hospitals, offices, etc. However, if a robot has to work in a real environment, actions base on various motions which should be input by humans are essential. Thus captured human motions data [1, 3] have been used to simplify the process of programming and learning complex motions. Moreover, another purpose of whole-body imitation of human is increasing in autonomous behavior of humanoid as well as improving their reactivity. Although there are many results that are implemented by previous researchers, to imitate human motions and to control a humanoid robot are a very challenging task and still a significant topic in humanoid robotics research.

Recently, due to the manageability, expandability, and affordability, not only large or middle-size humanoid but also small-size humanoid is used in a human imitation system. In addition, they are often used for education, research, and entertainment. This is one of the main reasons why we use DARwIn-OP [2] in this research.

Several researches, which control motion of humanoid robot, were published [4, 5, 6]. However most of researches require offline process for building motion database, and many devices, high computational cost. There are constraints in enlarging the application area of a small-size humanoid. Moreover, the full-body of human consists of over 600 muscles and over 200 bones, and every joint has one or more (up to 3) degrees of freedom (DOF). [7] has proposed several methods to transform a captured motion data into the motion data that humanoid robot can execute. The transform method seems to be a significant issue because of some physical limits of humanoid robot, e.g. different DOF between human and robot, different joint velocity and different torque etc., hinder direct motion transform. The generated motions also need to satisfy the posture and dynamics constraints.

From these reasons, in this paper, with visual motion captured by Microsoft Kinect Camera and small-size DARwIn-OP humanoid robot, we will propose a human imitation system that can solve the problems of mapping between different kinematic structures. All the joint angles are calculated by coordinate from motion data, state of clavicle and geometric constraints. We will consider two core issues of imitation, known as *"what to imitate"* and *"how to imitate"*. As it was mentioned in [8], *what to imitate* refers to the problem of what actions in the world are appropriate to imitate and *how to imitate* corresponds to a problem of how the robot will perform those parts of the motion that should be imitated. On the other hand, in this paper we will focus on the imitation of upper body motion. The motion of upper body is generated while the balance of robot is keeping by optimized moment around the center of mass [3], and position of ZMP [9].

The next section describes the proposed system step by step and hardware/software platform setup. Section 3 discusses the imitation problem and approaches for a solution. Data pre-processing and transfer method are detailed in Section 4. Section 5 presents experimental results. Conclusion and our future work will be provided in Section 6.

II. System Components

Overview of our proposed system is shown in Fig. 1.

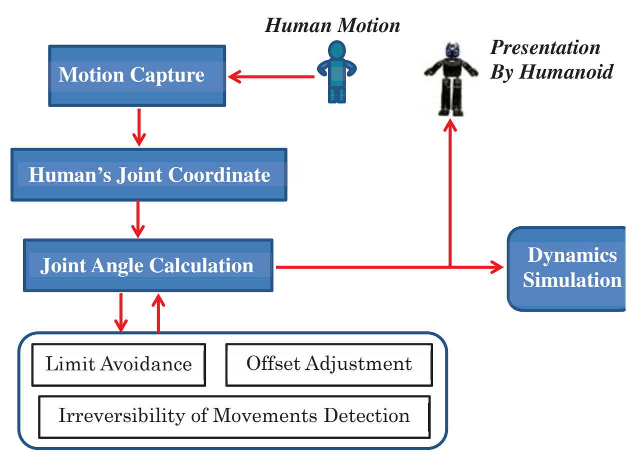


Fig. 1. System Summary

Manuscript received October 1, 2012.

V. V. Nguyen is with the College of Information Science and Engineering, Ritsumeikan University, Japan. (email: is0097rr@ed.ritsumei.ac.jp).
J.-H. Lee is with the Graduate of Science and Engineering, Ritsumeikan University, Japan. (e-mail: leejooho@is.ritsumei.ac.jp).

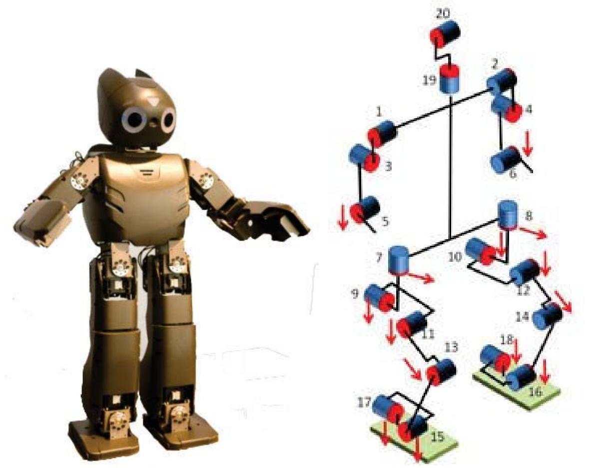


Fig. 2. Darwin-OP Humanoid and Joint Layout (Red arrows indicate zero Position on motor)

A. *Camera and Captured Motions Data*

The method for capturing the motion of human uses skeleton pose estimation with OpenNI module in each frame of Kinect Camera. All the coordinates of the body joints and their trajectory over the captured sequence is saved and transformed to the simulator.

B. *Robot and Models*

As it was mentioned in previous section, we suggested the open humanoid platform DARwIn-OP (stands for *Dynamic Anthropomorphic Robot with Intelligence–Open Platform*) as the target humanoid and its joint layout is shown in Fig. 2.

Darwin-OP was developed by ROBOTIS, consist of twenty Dynamixel RX-28M actuators. The actuators are said to offer 6-DOF for the legs (*lack 1-DOF in each ankle*), as well as dual 3-DOF arm movement (*lack of 1-DOF in the shoulder and 3-DOF in hand compared to a human arm*) and dual 2-DOF neck movement.

C. *Simulation*

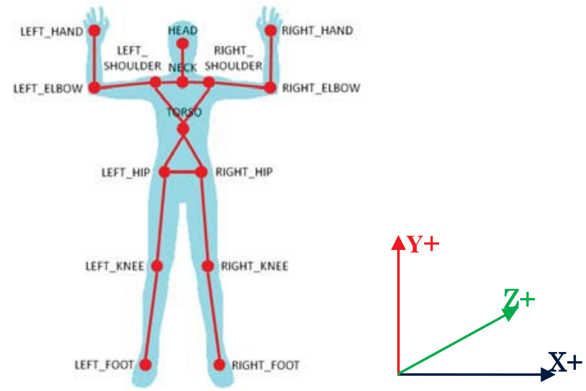
Due to demand for three-dimensional mobile robot simulator, in this case, Webots Simulator (*Developed by Cyberbotics Ltd*) was used for simulation and experiment. The simulation environment in Webots is entirely based on the Virtual Reality Modeling Language projects. Moreover, realistic simulations and fast prototyping of mobile robots help reducing the amount of development time and real hardware experiment.

D. *Data Transfer*

After performing the transformation and calculation of each joint angle on Webots Simulator, all the optimal motions, which robot can imitate will be transferred to the real DARwIn-OP humanoid robot via Transmission Control Protocol and Internet Protocol (TCP/IP) over a network.

III. IMITATION PROBLEM AND SOLUTION

The kinematic structure of the humanoid robot DARwIn-OP is given in Fig. 2. In this structure the DOFs are presented by cylinders. The structure contains 20 DOF which correspond to 20 actuators.



(a) (b)

Fig. 3. Structure of Skeleton Pose for Motion Capture

As it was mentioned in section 2, if DARwIn-OP is compared to human, this robot lacks 1-DOF in each Ankle, 1-DOF in each shoulder and 3-DOF in each hand. Thus, more difficulties arise during transforming each joint angle of human from captured motion to each motor angle of humanoid robot. In this paper, we focus on solving the issues of upper body motion but it still guarantees essential motion of lower body.

A. *Motion Data Acquisition*

As it is shown in Fig. 3(a), a body shape of human whom is standing at in front Kinect was created. From the OpenNI tracker, we can get the body coordinates that contain coordinate of 15 points correspond to 15 joints of human-body with respect to the world coordinate system is shown in Fig. 3(b).

Let C^W be the full-body coordinate matrix and $C_H^W, C_N^W, C_T^W, C_S^W, C_{UA}^W, C_{LA}^W, C_W^W, C_{UL}^W, C_{LL}^W$ represent head, neck, torso, upper arm, lower arm, waist, upper leg, lower leg coordinate:

$$C^W = [C_H^W\ C_N^W\ C_T^W\ C_S^W\ C_{UA}^W\ C_{LA}^W\ C_W^W\ C_{UL}^W\ C_{LL}^W]^T$$

In order to compute the joint angles, we need to estimate the orientation of the each rigid part of body with respect to the human torso. Hence, we transform the coordinate systems as:

$$C^T = C^W - C^O \tag{1}$$

where $C^O = [C_N^W\ C_T^W\ C_T^W\ C_S^W\ C_S^W\ C_{UA}^W\ C_W^W\ C_W^W\ C_{UL}^W]^T$.

From C^T coordinate matrix, using the X-Y-Z in 3D coordination, we can derive the roll, pitch, yaw angle of each joint on human body.

B. *Human and Robot Arm Analysis*

An overview of the difference of DOF between human and humanoid robot arm is shown in Fig. 4.

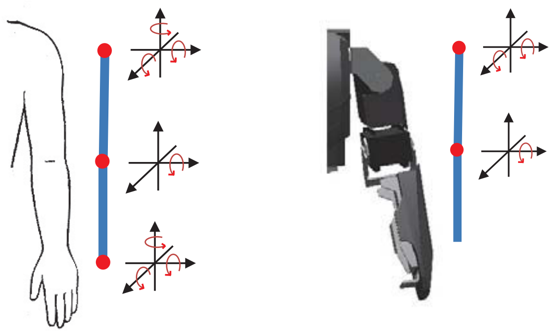
1) *Human Arm and Robot Arm Movements*

Based on clavicle [10], human arm movements around shoulder girdle are given by three types as shown in Fig. 5: abduction rotation, flexion/extension rotation and inward/outward rotation.

In case of the humanoid arm (Darwin-OP), one of the imitation problems is to overcome motions which are able for human to perform but impossible for robot to imitate such as movements as shown in Fig. 5(d).

2) *Find Rotation Angle Based on Geometric Relations*

The method of calculation of rotation angle of each human joint from its coordinate can be explained as follows:



(a) (b)

Fig. 4. Degree of freedomof the arm. (a) 7DOF Human Arm. (b) 3DOF Darwin-OP humanoid robot Arm

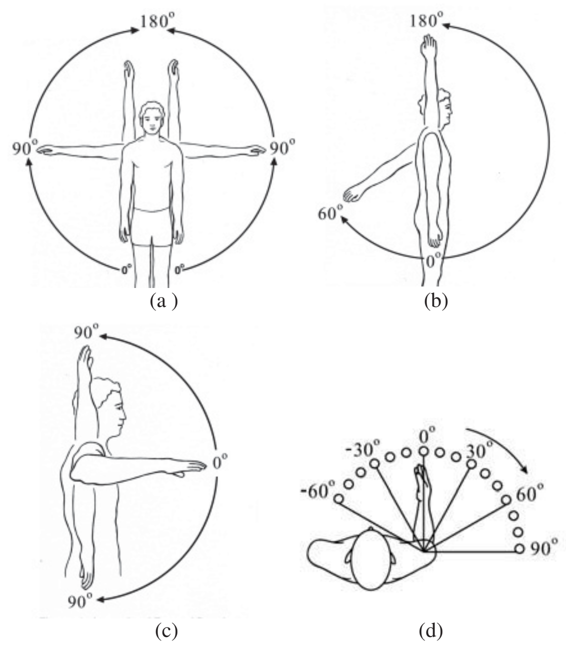


Fig. 5. Human arm movements. (a) Shoulder abduction rotation. (b) Shoulder flexion/extension rotation. (c) Shoulder inward/outward rotation . (d) A series movement with inward/outward rotation.

At first, we find elbow joint angle (θ_4) of arm. By dot product of the two vectors, the formula is as follows:

$$\overrightarrow{v_1}\cdot\overrightarrow{v_2} = |v_1||v_2|\cos(\alpha) \tag{2}$$

where $|v_1|$ is the distance from $\overrightarrow{v_1}$ to the origin, α represent the angle between two vectors $\overrightarrow{v_1}, \overrightarrow{v_2}$.

$$\alpha = \cos^{-1}\left(\frac{\overrightarrow{v_1}\cdot\overrightarrow{v_2}}{|v_1||v_2|}\right) \tag{3}$$

If l_1, l_2 are constants and represent the lengths of the upper and lower arm, $|p|$ represent the distance between start point

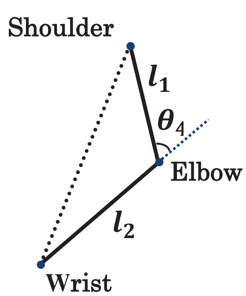


Fig. 6. A hinge joint as the elbow joint

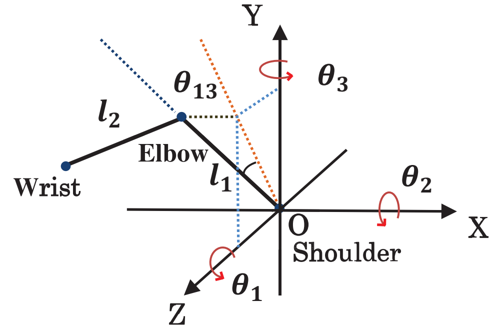


Fig. 7. Three joints with spherical joints as shoulder

$$\alpha = \cos^{-1}\left(\frac{\overrightarrow{v_1}\cdot\overrightarrow{v_2}}{|v_1||v_2|}\right) \tag{3}$$

(shoulder) and target point (wrist) as it is shown in Fig. 6. θ_4 can be found by using (3).

$$\theta_4 = \pi \pm \cos^{-1}\left(\frac{l_1^2 + l_2^2 - ||p||^2}{2l_1 l_2}\right) \tag{4}$$

Secondly, we find roll (θ_1), pitch (θ_2), yaw (θ_3) angle of shoulder joint as shown in Fig.7. By using the X-Y-Z fixed coordinate in respective global coordinate system, we assume that *Elbow* joint's coordinate is (e_x, e_y, e_z). According to the Law of Cosines in a triangle:

$$c^2 = a^2 + b^2 - 2abcos(C) \tag{5}$$

where a, b and c are sides. C is the angle opposite side c, we obtain:

$$\theta_1 = \cos^{-1}\left(\frac{e_y}{\sqrt{e_y^2 + e_x^2}}\right) \tag{6}$$

$$\theta_2 = \cos^{-1}\left(\frac{e_z}{\sqrt{e_z^2 + e_y^2}}\right) \tag{7}$$

$$\theta_3 = cos^{-1}\left(\frac{e_x}{\sqrt{e_x^2 + e_z^2}}\right) \tag{8}$$

Because of lacking yaw rotation angle (θ_3) in shoulder of humanoid robot while the arm is moving, we propose θ_{13} angle (as shown in Fig. 7.) that can be converted between yaw and roll axis.

$$\theta_{13} = \left(\frac{\sqrt{e_y^2 + e_z^2}}{\sqrt{e_x^2 + e_y^2 + e_z^2}}\right) \tag{9}$$

θ_{13} is used in case of arm movements.

Similarly, we are able to calculate other joints on the captured human skeleton.

C. *Angle Transformation between Human and Robot*

After performing the calculations based on constrained coordinate, the joint angles are mapped and transferred to robot. However, it cannot be directly input into the robot since there are physical limits, such as angle. Thus human motion has to be restricted in order to fit the robot capabilities.

1) *Angle Limit Avoidance*

In filtering and performing optimization of each joint angle in each time frame for angle limit avoidance, scaling the angle based on specifications of humanoid robot and human clavicle is used while it retains most of the individual oscillations seen in the original motion.

Scaling is performed in each initial joint:

$$\theta_i^t = f(\theta_i^t, k_i) \tag{10}$$

where k_i is rate coefficient which is estimated by range of joint angle i of robot correspond to human body (as mentioned in section B.1.). Then, all joints are filtered.

$$\hat{\theta}_i^t = \begin{cases} C_{i_{min}} & if\ \theta_i^t \leq C_{i_{min}} \\ \theta_i^t & if\ C_{i_{min}} \leq \theta_i^t \leq C_{i_{max}} \\ C_{i_{max}} & if\ \theta_i^t \geq C_{i_{max}} \end{cases} \tag{11}$$

where $\hat{\theta}_i^t$ represent configuration of θ_i joint angle calculated at time t, and it is mapped into robot joint angle. $C_{i_{min}}$ and $C_{i_{max}}$ denote the lower and upper joint angle bounds of joint i of humanoid robot. ($C_{i_{min}}$ and $C_{i_{max}}$ are derived from specifications of humanoid robot)

As it is described in (10), by estimating within the neighborhood of each motor angle limit, all angle joints in motion are modified. Hence, angle limiting can be done directly by applying lower and upper bound limit.

2) *Offset Adjustment*

In order to achieve more exact trajectory of motion, offset is used.

$$\hat{\theta}_i^t = \hat{\theta}_i^t + \sigma_i \tag{12}$$

Offset σ_i is estimated by distance bias between motors of the humanoid robot.

3) *Irreversibility of Movements Detection*

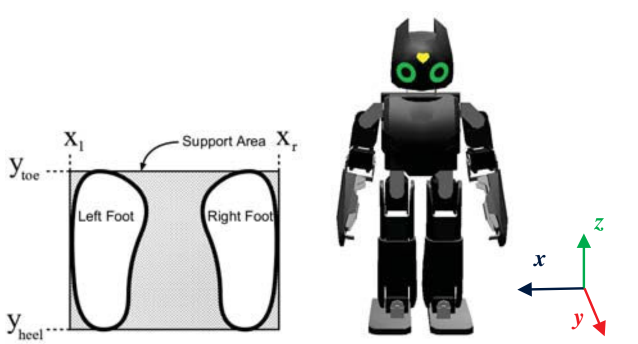
By estimating the orientation of the each rigid part of body and angle rotation around the *x*, *y* or *z* axis, we are able to detect the irreversibility of movements around shoulder. The steps of filtering are described simply as follows:

1) Input $\vec{p}$ (where $\vec{p}$ is vector motion which has $1 * n$ dimension).
2) Mapping each joint movement into joint layout in order to determine and calculate lacking rotation angles.
3) Return irreversibility of movements.

D. *Balance Control*

Humanoid robot may not be balanced if we input the motions adjusted in above section and the robot may fall down during doing some motions related to its leg. Therefore, the foot will not be controlled directly and human posture optimization to keep balancing is important.

In this paper we propose a weight shift control based on executing position feedback by Zero Moment Point (ZMP) proposed by Vukobratovic [9] and Center of Mass (COM) [3].



(a) The area (support polygon) in which ZMP can exist.

(b) Darwin-OP humanoid robot model on Simulation

Fig. 8. Support polygon and Humanoid robot model

In conventional method, the static balance control is considered important that the projection of the COM is in the foot supporting area (also called support polygon) as shown in Fig. 8(a). ZMP is a similar concept with the projection of the center of mass in the static balance control.

Let $p = [x_{ZMP}, y_{ZMP}, z_{ZMP}]^t$ be the position of ZMP, and the following equation can be obtained as follows:

$$x_{ZMP} = \frac{\sum_{i=1}^{n} m_i\{(\ddot{z}+g)x_i - \ddot{x}z_i\}}{\sum_{i=1}^{n} m_i(\ddot{z}+g)} \tag{13}$$

$$y_{ZMP} = \frac{\sum_{i=1}^{n} m_i\{(\ddot{z}+g)y_i - \ddot{x}z_i\}}{\sum_{i=1}^{n} m_i(\ddot{z}+g)} \tag{14}$$

Practically, as shown in Fig. 8(b), because the z-element of the foot's position is 0, so $p = [x_{ZMP}, y_{ZMP}, 0]^t$.

The constraint is about the area in which ZMP can exist as shown in Fig. 8(a). In order to recover projection of the COP in the foot supporting area, the constraints are written as:

$$x_l \leq x_{ZMP} \leq x_r \tag{15}$$

$$y_{\text{heel}} \leq y_{ZMP} \leq y_{toe}$$

The COM is controlled interactively by inverse kinematics to move upward around support area in a straightforward manner based on the constrained coordinate.

IV. Online Control Of Humanoid Motion

A. *Online Control*

In order to perform a series of motions continuously in real-time, it is necessary to estimate timing to transfer feasible motions to robot. In other words, it helps robot to realize imitation motions without stopping.

Since some complicated computation requires lots of time, our proposed system runs with best-effort time to guarantee all the motion state smoothly. Each acquired motion data in each frame will be transferred to robot as soon as computations are completed as shown in Fig. 9.

B. *Data Sending and Receiving*

Due to many advantages of using TCP/IP such as connections between different types of computers and servers, packet loss recovery, retransmitting missing packets as it was mentioned in Section 2.D, we use TCP/IP to transfer motion data from computer to Darwin-OP humanoid over Wireless Network.

The feasibility motion data are sent continuously through sockets as strings:

$$q_i = \{flag, \theta_0', \theta_2', \ldots, \theta_n'\}$$

where q_i is a sequence of joint angle poses, $flag$ is defined by error such as missing joint etc.

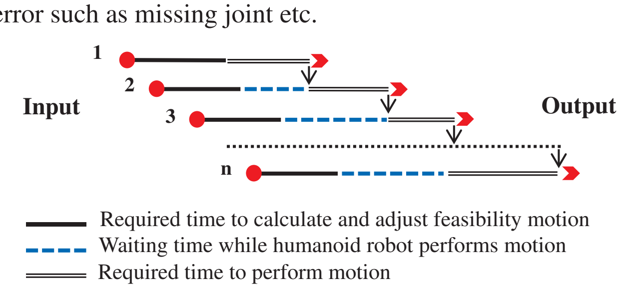


Fig. 9. Online Control Motion

All motions are performed by simulation before being executed on real humanoid robot.

V. Experiment And Result

This section describes implementation details of our system and shows the validity of the proposed system in this paper. As mentioned in section 1, the proposed system can imitate full-body .Firstly, some necessary initializations are prepared for our experiments, and then, the experiments and the results are presented in detail.

A. *Initializations*

1) *Motion Captured System*

We implemented an optical motion capture system using a Microsoft Kinect device depth camera based on PrimeSense technology that computes a skeletal model of a character because of fast calibration process and there is no need to use body markers for motion tracking.

Skeletal model of whole-human body is constructed based on the OpenNI framework.

2) *Robot Simulation*

In this case, we are now using Webots Simulation for preparing experiments with real robots. Environment and Darwin-OP humanoid robot model were developed in Webots.

Robot controller (including calculation and network related components) was implemented in C++ language based on Darwin-OP framework that is provided by ROBOTIS, and all the experiments in this section were performed under a PC with Intel processor Core 2 Duo E8500 3.16GHz CPU and 2GB of RAM on Microsoft Windows operating system.

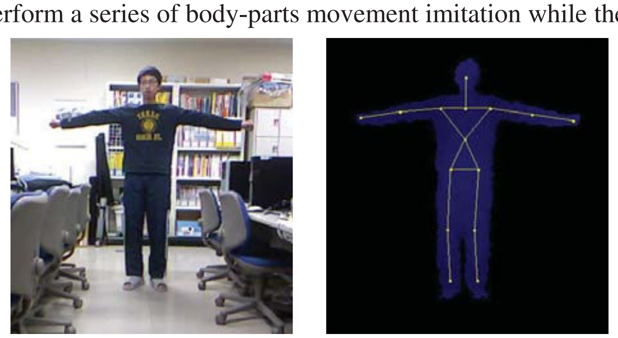
B. *Results*

Firstly, the human skeleton pose was generated as show in Fig. 10.

In this case, we created a model with 15 joints of human-body as shown in Fig. 10(b). A human is automatically segmented from the background using advanced background subtraction algorithms [14].

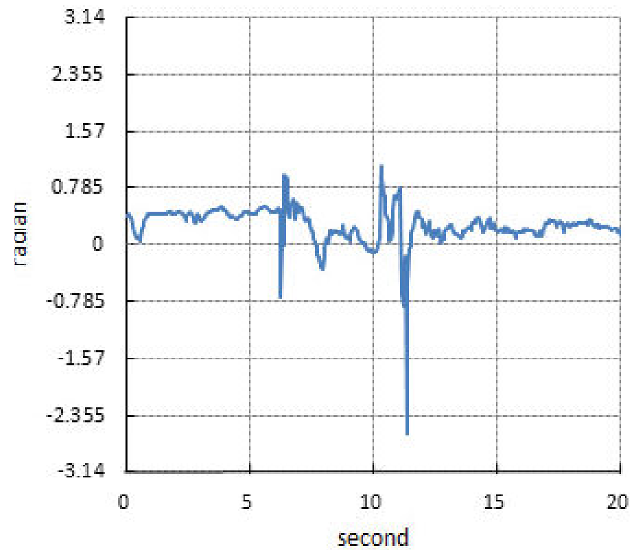
Secondly, as shown in Fig. 11, the transform methods, proposed in this paper modify input angular position trajectory. Fig. 11(a) shows the angular position trajectory of the human's right arm upper and Fig. 11(b) shows the modified trajectory after applying the transform methods explained in section 3.C.

Finally, in simulation, the virtual Darwin-OP robot could perform a series of body-parts movement imitation while the

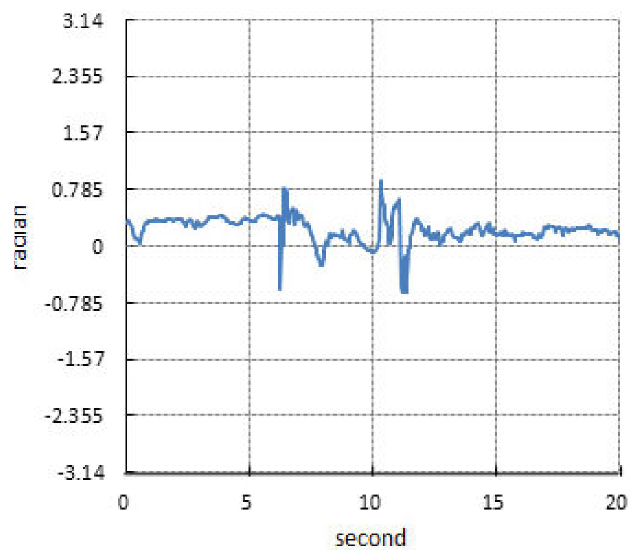


(a) Human image capture (b) The human skeleton pose

Fig. 10. Motion Tracking by Using Kinect and OpenNI



(a) Original captured motion of the right upper arm



(b) Obtained motion after adjustment

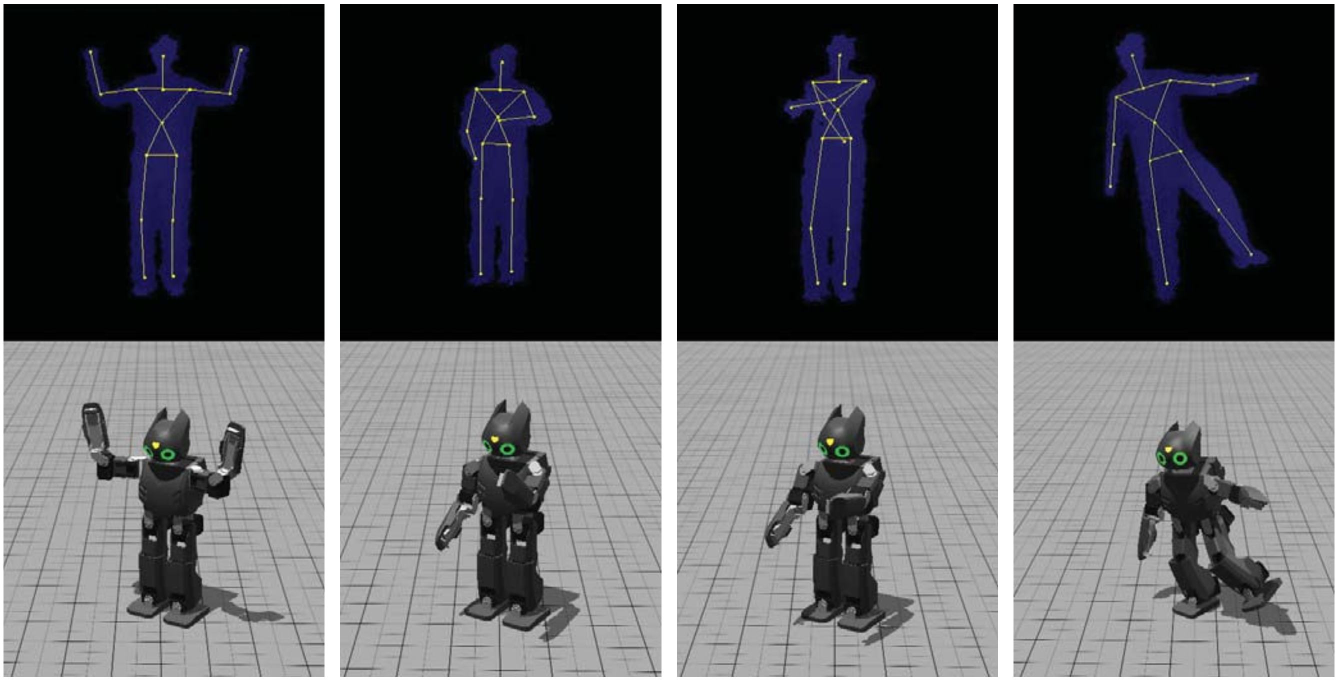
Fig. 11. Angular position of the right upper arm

robot is maintaining balance along the entire sequence. Snapshots of the human and virtual robot motion are shown in Fig. 12.

In the simulation, the feasible motions were performed by the robot and the example is shown in Fig. 12(a). In constrast, Fig. 12(b) is the case which requires inward rotation and irreversibility of movement should be determined and replaced by human-like movement. Fig. 12(c) is a case of overrunning of the limit angle of motor, humanoid robot was automatically adjusted based on the method explained in section 3.C.1. The weight shift was excuted to keep balancing while the robot was standing by one leg support as shown in Fig. 12(d).

VI. Conclusion And Future Work

This paper described the system for humanoid robot to perform full-body imitation of human motion. By using straightforward geometry and based on clavicle, our proposed



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

Fig. 12. Virtual Darwin-OP humanoid robot perform full-body imitation of human motions

method requires less time for computation of the kinematics. The experimental results showed that our system could adjust the robot motions which satisfy the mechanical constraints and dynamics consistency, feasibly.

As future works, we plan to set up experiments with real Darwin-OP humanoid and handle the motions that require fast speed or complex computation for collision free. We also plan to extract more information from motion data as key points for the primary purpose; our system can learn and predict next human motion in order to build a selecting motion strategy from data training.

REFERENCES

[1] J. B. Cole, D. B. Grimes, and R. P. N. Rao, “Learning full-body motions from monocular vision: Dynamic imitation in a humanoid robot,” in *Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2007.

[2] I. Ha, Y. Tamura, H. Asama, “Development of Open Humanoid Platform DARwIn-OP”, *International Conference on Instrumentation, Control, Information Technology and System Integration*, 2011, submitted for publication.

[3] S. Cotton, M. Vanoncini, P. Fraisse, N. Ramdani, E. Demircan, A. P. Murray, and T. Keller, “Estimation of the centre of mass from motion capture and force plate recordings: a study on the elderly.” *Applied Bionics and Biomechanics*, vol. 8, no. 2, pp. 67–84, 2011.

[4] Garam Park, Syungkwon Ra, Changhwan Kim and Jaebok Song, "Imitation Learning of Robot Movement Using Evolutionary Algorithm", *Confênce The Intênational Federation of Automatic Control*, pp.730-735, 2008.

[5] Carlos A. Acosta Calderon and Husheng Hu, "Robot Imitation from Human Body Movements", *Proceedings of the AISB05 Third International Symposium on Imitation in Animals and Artifacts*, Hatfield, United Kingdom, pp. 1-9, 2005.

[6] W. Suleiman, E. Yoshida, J.-P. Laumond, and A. Monin, “On Humanoid Motion Optimization,” in *IEEE-RAS 7th International Conference on Humanoid Robots*, Pittsburgh, Pennsylvania, USA, December 2007.

[7] E. Berglund, B. Iliev, R. Palm, R. Krug, K. Charusta, D. Dimitrov, Mapping between different kinematic structures without absolute positioning, *IEEE International Conference on Robotics and Automation (ICRA) Workshop on Autonomous Grasping*, 2011

[8] S. Calinon, F. Guenter, and A. Billard, Goal-directed imitation in a humanoid robot, in *Proceedings of the International Conference on Robotics and Automation,* Barcelona, Spain, Apr. 18-22, 2005.

[9] Vukobratovic M., Juricic D., Contribution to the synthesis of biped gait. *IEEE Transactions on Biomedical Engineering*, Vol. 16, No. 1, pp.1–6, 1969.

[10] C. Dube and J. Tapson, Kinematics design and human motion transfer for a humanoid service robot arm, *Council for Scientific and Industrial Research*, South Africa, November, 2009.

[11] Kudoh S., Komura T., Ikeuchi K., The Dynamic Postural Adjustment with the Quadratic Programming Method, *Proceedings of 2002 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp.2563-2568, January 2002.

[12] Nakaoka S., Nakazawa A., Kanehiro F., Kaneko K., Morisawa M., Hirukawa H. and Ikeuchi K., Leg task models for reproducing human dance motions on biped humanoid robots. Journal of the Robotics Society of Japan. Vol. 24, No. 3, pp.388-399, 2006.

[13] Tolani, D., Badler, N.I., Real-time inverse kinematics of the human arm, Presence 5, 393–401, 1996.

[14] OpenNI Programmer Guide. http://openni.org/Documentation/ProgrammerGuide.html