

ORIGINAL ARTICLES

A Real-Time Upper-Body Robot Imitation System

Zhijun Zhang^{*1,2}, Yaru Niu^{1,2}, Lingdong Kong^{1,2}, Shuyang Lin¹, Hao Wang²

¹School of Automation Science & Engineering, South China University of Technology, China

²Center for Brain Computer Interfaces & Brain Information Processing, South China University of Technology, China

Received: November 19, 2018

DOI: 10.5430/ijrc.v2n1p49

Accepted: January 13, 2019

URL: <https://doi.org/10.5430/ijrc.v2n1p49>

Online Published: January 29, 2019

ABSTRACT

An upper-body robot imitation (UBRI) system is proposed and developed to enable the human upper body imitation by a humanoid robot in real time. To achieve the imitation of arm motions, a geometry-based analytical method is presented and applied to extracting the joint angles of the human and mapping to the robot. Comparing to the traditional numerical methods of inverse kinematic computations, the geometrical analysis method generates a lower computational cost and maintains good imitation similarity. To map the human head motions to the head of the humanoid robot, a face tracking algorithm is employed to recognize the human face and track the human head poses in real time. A hand extraction and hand state recognition algorithm is proposed to achieve the hand motion mapping. At last, the completion rate and similarity evaluation experiments are conducted to verify the effectiveness of the proposed UBRI system.

Key Words: Imitation, Motion mapping, Hand detection

1. INTRODUCTION

Robot imitation is an efficient technique that endows the robot with the ability to transmit and reproduce human motions. It provides a novel and effective way of controlling the robot, especially the humanoid robot with a high degree of freedom. Through imitation, a robot can reproduce a user's motions simultaneously from a long distance,^[1] and folk dances can be reconstructed.^[2]

Robot imitation can be divided into three stages, i.e., observation, representation and reproduction.^[3] For observation, Riley et al. used a 3D vision system which consisted of external cameras and head-mounted cameras to capture human motions.^[4] They put several color marks on a human body and then calculated and recorded the positions of these markers. Similar work using a marker-based visual capture system can be found in Ref.^[5] Another popular motion cap-

ture device is wearing device, such as “Xsens MVN” and “ShapeTape”.^[6–8] A more user-friendly and low-cost way is to use markerless visual capture system such as the Microsoft Kinect sensor.^[9–14] They achieved robot imitation by using the skeleton data provided by the Kinect.

In the representation stage, the human motions should be mapped to the robot. The traditional methods for motion mapping are numerical approaches of inverse kinematic computations.^[4, 9, 13–15] The problem is then converted to calculating joint angles according to the positions of the end effectors. This kind of method is effective in the work space of an imitation task, but it needs higher computational cost since only the information of the end effectors is considered and employed. Some imitation system used geometry-based analytical method to extract joint angles of the human and apply them to the robot.^[9–11, 16] They fully used the body information and obtained the joint angles in a more direct

^{*}Correspondence: Zhijun Zhang; Email: auzjzhang@scut.edu.cn; Address: School of Automation Science & Engineering, South China University of Technology, China.

way and with a lower computational cost.

Head motions enables the humanoid robot to generate more human-like behaviors. Additionally, if the humanoid robot has a head-mounted camera, the head motions will help monitor the environment around the robot in a task such as teleoperation. However, most of the existing work do not consider the head motions.

Similarly, hand motions can improve the imitation similarity and are essential for some tasks using imitation such as grabbing and releasing objects. The hand motions are also ignored by many researchers. In Ref.,^[17] the humanoid robot can imitate human arm motions in real time for a pen shopping task. Its hand closing and opening are controlled via speech signals. Obviously, if the robot's hand state can follow that of the human, the "pen shopping" will be more convenient and enjoyable.

To take advantage of the existing work, solve the problems mentioned above, and improve the performance of the upper-body imitation, we propose and develop an upper-body robot imitation (UBRI) system integrating the methods of arm motion mapping, head motion mapping and hand motion mapping. The system framework is depicted in Figure 1. In our work, deep sensor Kinect is served as the human motion capture device and imitation behaviors reproduced by the Nao robot. In the observation stage, human motions are captured by the Kinect sensor, and the skeleton data, the depth data and the color data are acquired. In the motion mapping stage, the arm motion mapping, head motion mapping and the hand motion mapping method are developed and applied, so the human upper-body motions can be mapped to the robot. At last in the reproduction stage, the humanoid robot reproduce the upper-body motions in accordance with the mapping results. Therefore, the robot can imitate various upper-body motions in real time.

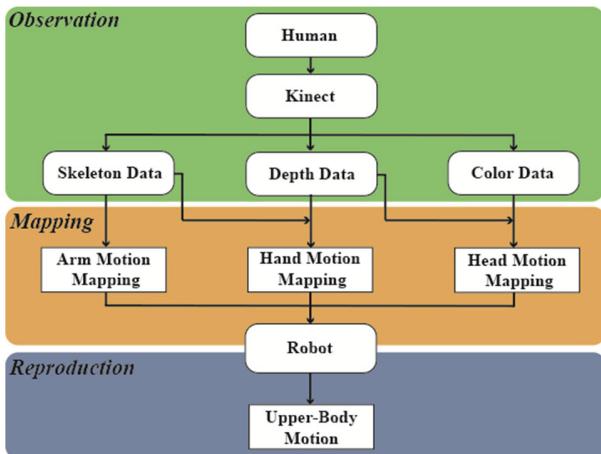


Figure 1. The framework of the proposed UBRI system

The remainder of this paper is organized as follows. In Section 2, a geometry-based analytical arm motion mapping method is proposed and presented. The face tracking algorithm for head motion mapping is described in Section 3. In Section 4, the hand motion mapping method using a hand extraction and hand state recognition algorithm is proposed. The experiments and results are described and analyzed in Section 5. Section 6 concludes the paper with some final remarks.

2. ARM MOTION MAPPING

To map arm motions of the human to the robot, some upper-body link vectors are built through the skeleton points in the human skeleton model, as shown in Figure 2. A link is defined as a rigid connecting two skeleton points. Human's upper body can be simplified as a geometrical structure consisting of these link vectors. Therefore, human motions can be represented by the motions of the dynamic geometrical structure and a joint angle can be extracted from the angle between two vectors. For the convenience of mapping, the joint angles should be calculated in accordance with the joint structure of the robot. The joint structures and joint limits of Nao's arms can be seen in Figure 2 and Figure 3, and Shoulder Pitch, Shoulder Roll, Elbow Roll, Elbow Yaw joints are considered for arm motions. In this paper, we take the left arm as an example to present how to extract joint angles used for arm motion mapping. The ranges of the joint angles are also taken into consideration.

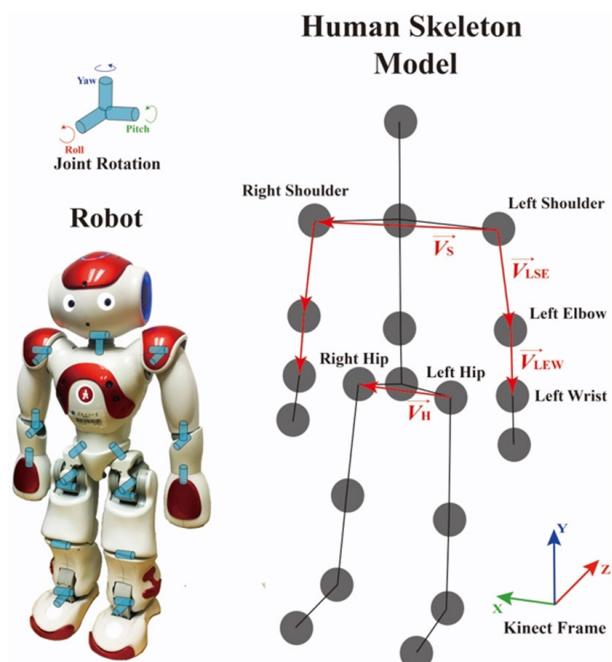


Figure 2. The joint structures of Nao and the human skeleton model with link vectors

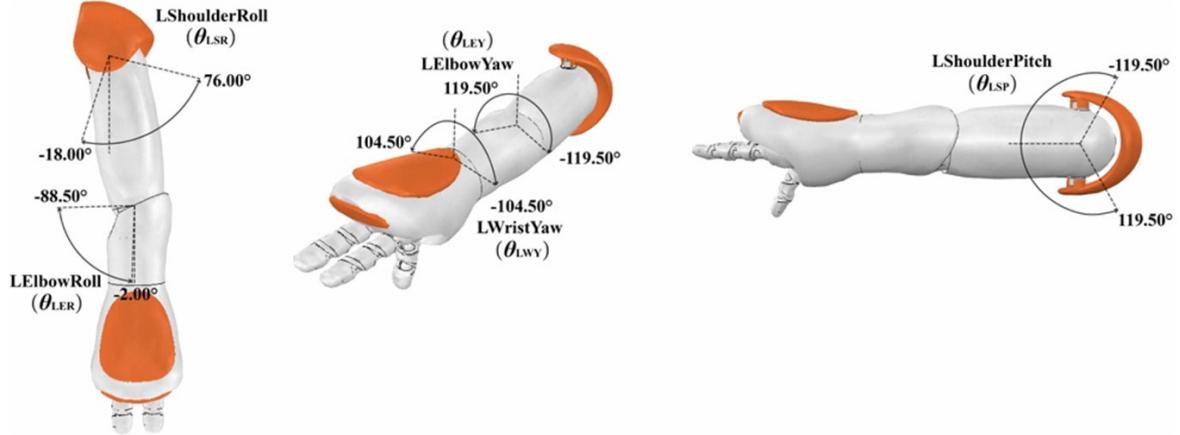


Figure 3. Left arm joints and joint ranges of the Nao robot

As shown in Figure 2, supposing that the demonstrator is facing the Kinect sensor, the calculation is discussed in the Kinect frame of which the Z axis points to the demonstrator, the Y axis points upwards and the X axis points to the right side from the view of the demonstrator. The vector \vec{V}_H should be parallel to the X axis. For the convenience of discussion, a frame similar to the Kinect frame is set up corresponding to the Nao robot. As depicted in Figure 4 (b), (c) and (d), the X axis is parallel to the vector pointing from the left hip to the right hip, the Y axis points upwards and the Z axis points to the robot. The joints of Nao's arms reach zero positions when its arms raise forwards horizontally, parallel to the Z axis. The zero position of the left arm is marked by the red lines in Figure 4 (a), (b) and (d).

Shoulder Pitch joint enables arms to swing forwards and backwards as shown in Figure 4 (a). We suppose that the link vector pointing from the left shouder to the left elbow in the Kinect frame is $\vec{V}_{LSE} = (x, y, z)$. In accordance with the definition of the joint angle of Nao's left arm, the left Shoulder Pitch joint angle \vec{V}_{LSP} can be obtained as follow:

$$\vec{V}_{LSP} = \begin{cases} \arctan\left(\frac{y}{z}\right) & z < 0 \\ -\arctan\left(\frac{z}{y}\right) + \frac{\pi}{2} & z > 0 \text{ and } y < 0 \\ -\arctan\left(\frac{z}{y}\right) - \frac{\pi}{2} & z > 0 \text{ and } y > 0. \end{cases} \quad (1)$$

Shoulder Roll joint enables arms to swing leftwards and rightwards as shown in Figure 4 (b). To calculate the angle of this joint, a vector pointing from the left hip to the right hip, \vec{V}_H is introduced. The vector that points from the left to the right shoulder denoted as \vec{V}_S is not used because it is not stable enough, for its orientation is easily affected by the movements of the upper body. Therefore, the left Shoulder Roll joint angle, denoted as θ_{LSR} is obtained minus $\pi/2$ from the angle between \vec{V}_H and \vec{V}_{LSE} :

$$\theta_{LSR} = \arccos\left(\frac{\vec{V}_H \cdot \vec{V}_{LSE}}{|\vec{V}_H| |\vec{V}_{LSE}|}\right) - \frac{\pi}{2} \quad (2)$$

where \vec{V}_{LSE} is the link vector pointing from the left elbow to the left wrist.

As depicted in Figure 4 (c), the angle of the left Elbow Yaw joint can be regarded as the angle between the plane formed by left shoulder, right shoulder and left elbow (Plane BDE) and the plane formed by left shoulder, left elbow and left wrist (Plane ABD). The normal vectors of these two planes, \vec{N}_1 and \vec{N}_2 respectively, can be defined as:

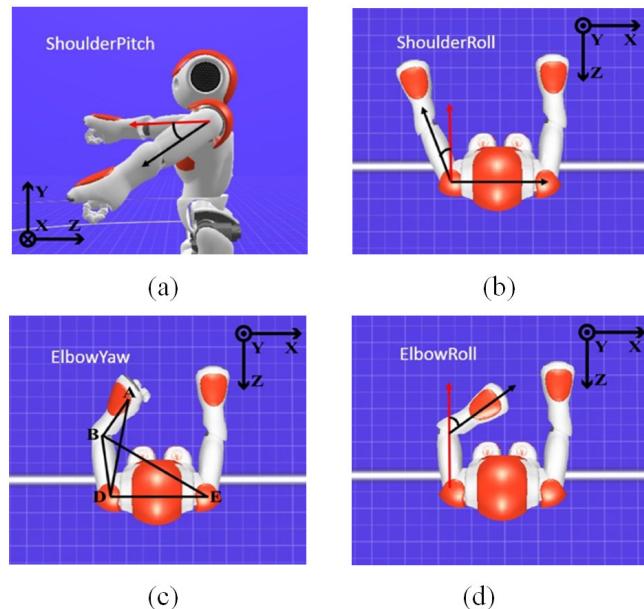


Figure 4. Diagrams of the calculation of each joint angle of the left upper body limb

$$\vec{N}_1 = \vec{V}_{\text{LEW}} \times \vec{V}_{\text{LSE}} \quad (3)$$

$$\vec{N}_2 = \vec{V}_{\text{S}} \times \vec{V}_{\text{LSE}} \quad (4)$$

where \vec{V}_{LEW} is the left lower arm vector, and \vec{V}_{LSE} is the left upper arm vector.

Suppose that the X coordinate of \vec{N}_1 is x_1 and the angle can be derived as:

$$\theta_{\text{LEY}} \begin{cases} -\arccos\left(\frac{\vec{N}_1 \cdot \vec{N}_2}{|\vec{N}_1||\vec{N}_2|}\right) & x_1 < 0 \\ \arccos\left(\frac{\vec{N}_1 \cdot \vec{N}_2}{|\vec{N}_1||\vec{N}_2|}\right) & x_1 > 0 \end{cases} \quad (5)$$

The angle of the left Elbow Roll joint is the angle between the left lower arm vector \vec{V}_{LEW} and the left upper arm vector \vec{V}_{LSE} as shown in Figure 4 (d), so it can be calculated by:

$$\theta_{\text{LER}} = -\arccos\left(\frac{\vec{V}_{\text{LSE}} \cdot \vec{V}_{\text{LEW}}}{|\vec{V}_{\text{LSE}}||\vec{V}_{\text{LEW}}|}\right) \quad (6)$$

Then the extracted joint angles can be applied to the arm of the robot. If the value of any is beyond the limit of the corresponding robot joint, the angle will be assigned a marginal value.

This arm motion mapping method can be also extended to

the lower limb, subjecting to the whole-body balance control, thus the whole-body imitation can be achieved. In this paper, the balance control is not discussed and we focus on the upper-body imitation.

3. HEAD MOTION MAPPING

Robot's imitation of human's head motions was usually ignored by researchers. In fact, head motions promotes the completeness of the robot imitation and makes the robot act more like a human. Besides, mapping head motions provides a new way to control the orientation of the camera located on the robot's head so that the view of the robot changes following human's head, which contributes to monitoring the real-time environment around the robot.^[18]

Kinect SDK supports detecting and tracking the user's face. With inputs of color and depth images, the 3D head pose can be derived, the principle of which is presented in Refs.^[19, 20] The head of the Nao robot has two joints, Head Yaw and Head Pitch, and these joints of the human can be obtained through this face tracking algorithm. The quality of input images affects the tracking quality so the tracking result will be better if the human is in a brighter environment.

As it can be seen in Figure 5, the black rectangular frame shows the detected position of the human face which can follow the human face when it turns upwards, downwards, leftwards and rightwards even at a large angle. The angles of Head Yaw and Head Pitch are output and applied to the robot continuously in this process.

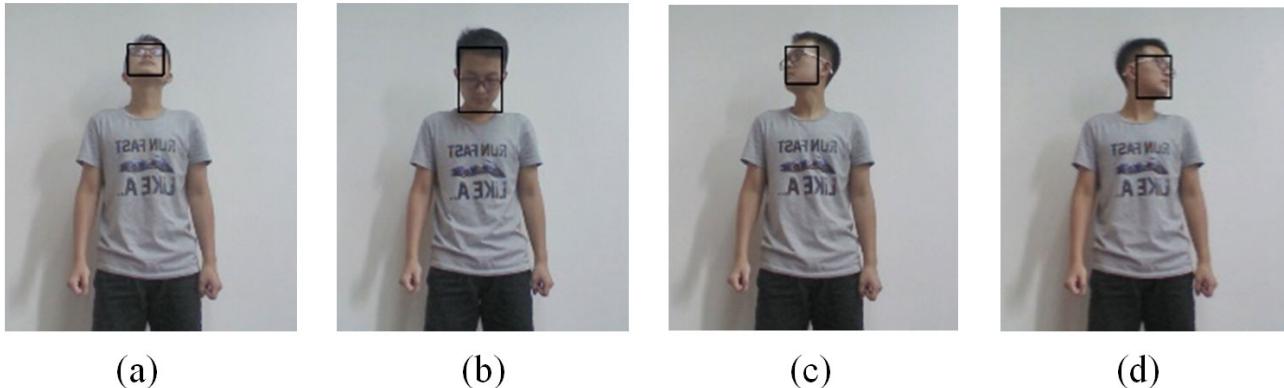


Figure 5. The face can be tracked well when the head is turned upwards, downwards, leftwards or rightwards at a large angle.

4. HAND MOTION MAPPING

Hand motions enable the robot to generate more human-like behaviors. Most importantly, hand motions are essential for task operations such as grabbing and releasing objects in a

teleoperation task. Therefore, hand motions are considered in our UBRI system, and the human hand motions are mapped to the robot. The method of extracting the hands from the depth image and recognizing the hand state are introduced in this section, and the algorithm can be seen in Algorithm 1.

Algorithm 1: Algorithm for hand motion mapping

```

Initialize left hand open times  $t_{l1} \leftarrow 0$ , left hand closed
times  $t_{l2} \leftarrow 0$ , right hand open times  $t_{r1} \leftarrow 0$ , right hand
closed times  $t_{r2} \leftarrow 0$ 
Initialize hand imitation state  $\leftarrow$  True
Initialize depth threshold  $d$ , area threshold  $S$ , time
threshold  $t$ , distance threshold  $h_0$ , lower bound of pixel
number  $N_{\text{lower}}$ , upper bound of pixel number  $N_{\text{upper}}$ 
while hand imitation state = True do
  if  $h \geq h_0$  then
    Obtain the depth value of the human body's
    nearest point to the Kinect sensor  $D_{\min}$ 
    Set  $D_{\text{hand}} \leftarrow D_{\min} + d$ 
    Count the number of pixels of the extracted left
    part  $N_l$  and that of the extracted right part  $N_r$  in
    the region  $[D_{\min}, D_{\text{hand}}]$ 
    if  $N_{\text{lower}} < N_l < N_{\text{upper}}$  then
      Obtain the area of the minimum circle which
      surrounds the left hand  $S_l$ 
      if  $S_l > S$  then
         $t_{l1} \leftarrow t_{l1} + 1$ 
         $t_{l2} \leftarrow 0$ 
        if  $t_{l1} > t$  then
          Open the left hand
           $t_{l1} \leftarrow 0$ 
        end if
      else
         $t_{l2} \leftarrow t_{l2} + 1$ 
         $t_{l1} \leftarrow 0$ 
        if  $t_{l2} > t$  then
          Close the left hand
           $t_{l2} \leftarrow 0$ 
        end if
      end if
    end if
    else
       $t_{l2} \leftarrow t_{l2} + 1$ 
       $t_{l1} \leftarrow 0$ 
      if  $t_{l2} > t$  then
        Close the left hand
         $t_{l2} \leftarrow 0$ 
      end if
    end if
  end if

```

```

end if
end if
end if
if  $N_l \leq N_{\text{lower}}$  or  $N_l \geq N_{\text{upper}}$  then
   $t_{l1} \leftarrow 0$ 
   $t_{l2} \leftarrow 0$ 
end if
if  $N_{\text{lower}} < N_r < N_{\text{upper}}$  then
  Obtain the area of the minimum circle which
  surrounds the right hand  $S_r$ 
  if  $S_r > S$  then
     $t_{r1} \leftarrow t_{r1} + 1$ 
     $t_{r2} \leftarrow 0$ 
    if  $t_{r1} > t$  then
      Open the left hand
       $t_{r1} \leftarrow 0$ 
    end if
  else
     $t_{r2} \leftarrow t_{r2} + 1$ 
     $t_{r1} \leftarrow 0$ 
    if  $t_{r2} > t$  then
      Close the left hand
       $t_{r2} \leftarrow 0$ 
    end if
  end if
  end if
  if  $N_r \leq N_{\text{lower}}$  or  $N_r \geq N_{\text{upper}}$  then
     $t_{r1} \leftarrow 0$ 
     $t_{r2} \leftarrow 0$ 
  end if
end if

```

4.1 Hand Extraction

First of all, the depth information of the human body is separated from the raw depth image in accordance with the user ID to remove interference of other people in the view of the Kinect, then we obtain the image M_{body} . Next, we traverse all the pixels in M_{body} to figure out the human body's nearest point to the Kinect with the depth value D_{\min} . After adding an appropriate value d on D_{\min} , we can get D_{hand} :

$$D_{\text{hand}} = D_{\min} + d \quad (7)$$

which is a bit larger than the maximal depth value of human's hand. Therefore, the hands can be extracted from the human body using the depth interval $[D_{\min}, D_{\text{hand}}]$ as shown in Figure 6. We use the method of counting the number of pixels of the extracted hand in a rectangle region, as can be seen in Figure 7, to judge if the hand is detected. The

rectangles' positions are determined by the positions of the hands in the depth image. Their length and width are set to cover all pixels of the hand. If the number of the pixels is between a lower bound and an upper bound, which are obtained from experiments, the hand in this rectangular region will be considered to be tracked successfully. If d is too small, the hand extraction may fail when the hand is not opposite to the Kinect. If d is too large, other parts of the body may be extracted. Also, other parts of the human body will be extracted if the hand is too close to them, which will affect the correctness of judging the hand state. To avoid errors, the distance between the hand and the torso h is calculated by 3D coordinates derived from skeleton information. If the distance is shorter than a certain threshold distance h_0 , the extraction will not be conducted. The extracted hand can be seen from Figure 8 (a).

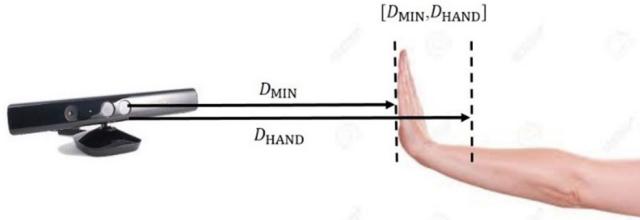


Figure 6. Hand extraction through the depth interval $[D_{\min}, D_{\text{hand}}]$

4.2 Hand State Recognition

After extracting the hand, the next step is to judge the hand state from the extracted parts, in other words, to distinguish between the palm and the fist. The problem can be simplified as comparing the area of the palm's and the fist's planar graph, since only two states need to be distinguished and the two areas are obviously different. To make this difference more remarkable, we get a minimum circle which surrounds one hand and then calculate the area of this circle, S_l for the left hand and S_r for the right hand. If the area is larger than a certain threshold S , the hand state is regarded as “open” otherwise is regarded as “closed”. As shown in Figure 8, the area of the palm is quite larger than the fist and this difference is more obvious after drawing the minimum circle. To improve accuracy, a hand’s state will be finally confirmed only when the same hand state is detected continuously for several times. If we suppose that the occurrence times of the open and closed state of the left hand and those of the left hand are t_{l1}, t_{l2}, t_{r1} and t_{r2} , respectively, the hand states will be finally determined only if they are larger than t , which is a threshold of the occurrence times.

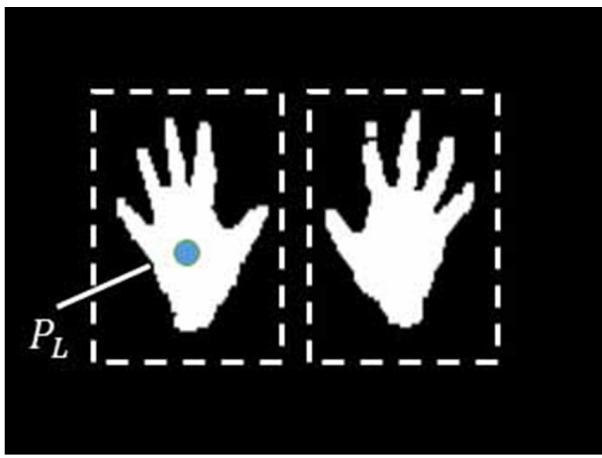


Figure 7. The rectangle is the region containing all the pixels of the corresponding extracted hand. P_L represents the position of the left hand.

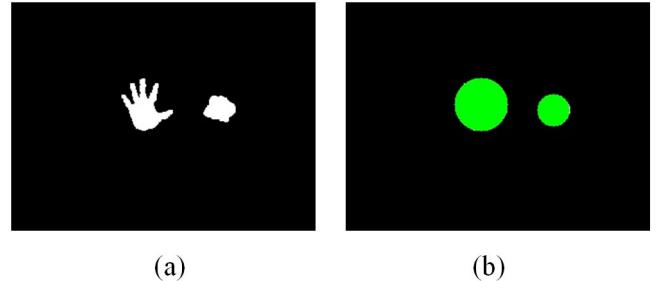


Figure 8. The minimum circles that surround the hand

5. EXPERIMENTS AND EVALUATION

In the experiment, the Microsoft Kinect v1 was used as the human motion capture device and the imitation task was executed by Nao v5. The program was run on a laptop. We collected human body information from the Kinect and these data were processed on the laptop. Then joint angles and the hand states were transferred to Nao through WIFI. In case that several people appear in the view of the Kinect, which will disturb the experiment, we always select the person who is nearest to the Kinect as the demonstrator.

Two indexes, the completion rate and similarity, are designed to evaluate the performance of our system. Eight volunteers were invited to finish the experiments. First, the poses of the upper limb, open-and-close motions of each hand and motions of moving head upwards, downwards, leftwards and rightwards. Some of the specified poses and the imitation results are shown in Figure 9, Figure 10, and Figure 11. One successful completion would be recorded when the robot managed to make a correct similar pose in real time. Specifically, a correct similar pose is the pose whose arms’ direction, face’s direction, and hand state are the same as those of the human, and which is regarded as similar to the human pose by three observers. Then the completion rate can be obtained by calculating the ratio of the number of successful completions to all trials. In this experiment, the completion rate of the arm motion imitation is 95%, and both head and hand motion imitation reach 91%. It is worth pointing out that the arm motion mapping method of the proposed system has a lower computational cost compared to the traditional numerical inverse kinematic method such as Ref.^[15] which requires 1.04 ms for posture mapping on average. We set a timer in the program and got the computational time of the arm and head motion mapping in the above experiment. Then we calculated the average time of the trials for different volunteers and obtained the result which was 0.034 ms.

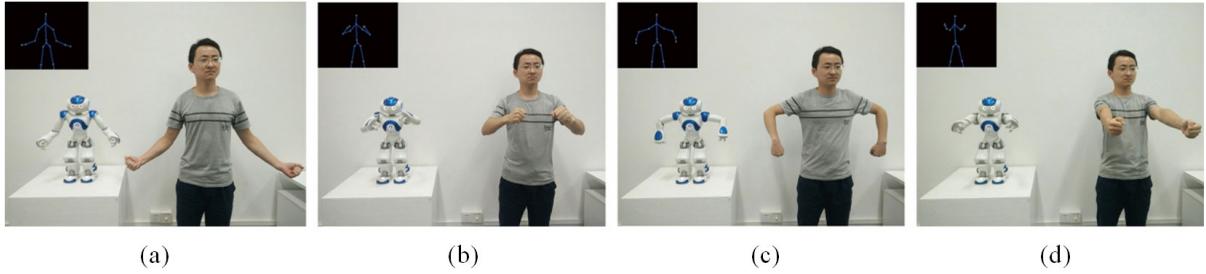


Figure 9. Arm motion imitation and corresponding skeleton information

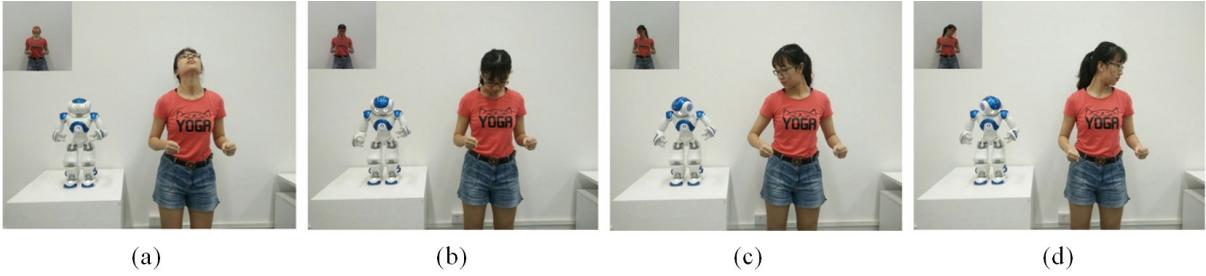


Figure 10. Head motion imitation and corresponding face tracking

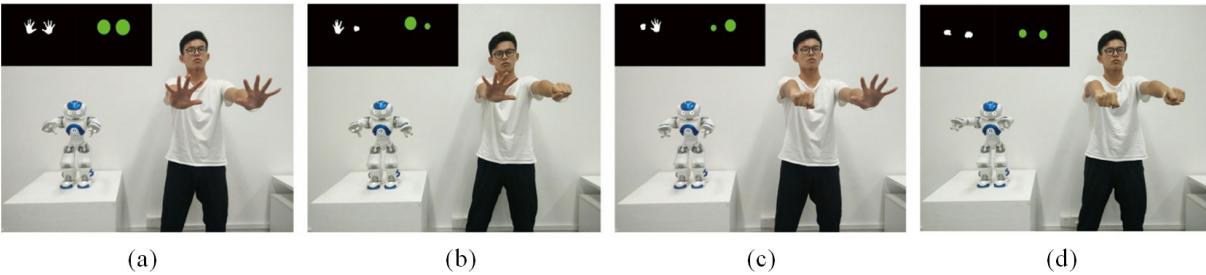


Figure 11. Hand motion imitation and corresponding hand extraction and hand state recognition

To evaluate the similarity, Mean Opinion Score (MOS) method^[12,21] was applied. The volunteers were asked to experience the imitation system freely and then finished the questionnaire in which they were asked to access the similarity ranging from 1 to 5, corresponding to terrible, bad, fair, good, and excellent, respectively. The similarity index, denoted as I_S , can be obtained by averaging the scores collected from the volunteers, which can be expressed as:

$$I_S = \frac{1}{n} \sum_i^n s_i \quad (8)$$

where n is the number of the volunteers, and s_i denotes the score of the i^{th} volunteer. In this experiment, the similarity index is 4.

According to the experimental results, the completion rates are all more than 90%, which verifies the effectiveness of the proposed UBRI system. There are some negative factors affecting the completion rates. The joint limitations of Nao

reduce the quality of imitation, that is, it makes Nao fail to imitate some human motions. For example, people can clap their hands while Nao can not do that due to the joint limits of Shoulder Roll. Besides, when raising arms forwards horizontally, the hand, the elbow and the shoulder of one side may overlap in the view of the Kinect, so the skeleton data may be not stable and the calculated joint angles may be not consistent with the actual human motion. As for the head imitation, fast movements of the head sometimes will make face tracking fail and as a result, Nao's head will not move.

Regarding to the imitation similarity, 80% of the full similarity index value is a good performance, considering the different sizes and configurations between the human and the robot. As mentioned above, the incorrect imitation will reduce the similarity during imitation. In addition, the similarity can be also affected by the smoothness and responding speed of Nao's motions, because the humanoid robot can not

generate motions as naturally as the human.

6. CONCLUSION

In this paper, a UBRI system is proposed and presented. To realize upper body motion imitation on humanoid robots, different methods are applied to three parts of the upper body including the arms, the hands, and the head. Specifically, a geometrical analytical method based on link vectors is developed and presented to map human arm motions to the humanoid robot. It has an evidently lower computational

cost than the typical numerical inverse kinematic mapping method. The head pose with joint angles of the human is deduced by a face tracking algorithm and then applied to the robot. To control the open-and-close motion of the robot's hands following the human, the human hand state is obtained by utilizing the depth and skeleton information. At last, the completion rate and similarity evaluation experiments are conducted to verify the effectiveness of the proposed UBRI system.

REFERENCES

- [1] Choi Y, Ra S, Kim S, Park SK. Real-time arm motion imitation for human–robot tangible interface. *Intelligent Service Robotics*. 2009; 2(2): 61-69. <https://doi.org/10.1007/s11370-009-0037-8>
- [2] Ikeuchi K, Sato Y, Nakaoka S, Kudoh S, Okamoto T, Hu H. Task modelling for reconstruction and analysis of folk dances. *Dance Notations and Robot Motion*. Springer; 2016. p. 187-207.
- [3] Bakker P. Robot see, robot do: An overview of robot imitation. *Proc. Aisb96 Workshop: Learning in Robots and Animals*. 1996. p. 3-11.
- [4] Riley M, Ude A, Wade K, Atkeson CG. Enabling real-time full-body imitation: a natural way of transferring human movement to humanoids. *2003 IEEE International Conference on Robotics and Automation (ICRA)*. 2003; 2: 2368-2374. <https://doi.org/10.1109/ROBOT.2003.1241947>
- [5] Liu HY, Wang WJ, Wang RJ, Tung CW, Wang PG, Chang IP. Image recognition and force measurement application in the humanoid robot imitation. *IEEE Transactions on instrumentation and measurement*. 2012; 61(1): 149-161. <https://doi.org/10.1109/TIM.2011.2161025>
- [6] Aleotti J, Skoglund A, Duckett T. Position teaching of a robot arm by demonstration with a wearable input device. *International Conference on Intelligent Manipulation and Grasping (IMG04)*. 2004. p. 1-2.
- [7] Cheng G, Kuniyoshi Y. Real-time mimicking of human body motion by a humanoid robot. *Proceedings of the Sixth International Conference on Intelligent Autonomous Systems (IAS2000)*, Venice, Italy, Citeseer. 2000. p. 273-280.
- [8] Naksuk N, Lee CG, Rietdyk S. Whole-body human to humanoid motion transfer. *5th IEEE-RAS International Conference on Humanoid Robots*. 2005. p. 104-109.
- [9] Wang F, Tang C, Ou Y, Xu Y. A real-time human imitation system. *World Congress on Intelligent Control and Automation (WCICA)*. 2012. p. 3692-3697.
- [10] Lei J, Song M, Li ZN, Chen C. Whole-body humanoid robot imitation with pose similarity evaluation. *Signal Processing*. 2015; 108: 136-146. <https://doi.org/10.1016/j.sigpro.2014.08.030>
- [11] Franz FS, Nolte-Holube R, Wallhoff F. *Nao follows me*: tracking, reproduction and simulation of human motion. Jade University of Applied Sciences, Germany; 2013.
- [12] Zuher F, Romero R. Recognition of human motions for imitation and control of a humanoid robot. *Robotics Symposium and Latin American Robotics Symposium (SBR-LARS)*, 2012 Brazilian. p. 190-195.
- [13] Poubel LP, Sakka S, Cehajic D, et al. Support changes during online human motion imitation by a humanoid robot using task specification. *IEEE International Conference on Robotics and Automation (ICRA)*. 2014. p. 1782-1787.
- [14] Avalos J, Cortez S, Vasquez K, Murray V, Ramos OE. Telepresence using the kinect sensor and the nao robot. *IEEE 7th Latin American Symposium on Circuits & Systems (LASCAS)*. 2016. p. 303-306.
- [15] Koenemann J, Burget F, Bennewitz M. Real-time imitation of human whole-body motions by humanoids. *2014 IEEE International Conference on Robotics and Automation (ICRA)*. 2014. p. 2806-2812.
- [16] Lee JH, et al. Full-body imitation of human motions with kinect and heterogeneous kinematic structure of humanoid robot. in *IEEE/SICE International Symposium on System Integration (SII)*. 2012. p. 93-98.
- [17] Pausch S, Hochdorfer S. Teleoperation of the robot nao with a kinect camera [Internet]. 2012. Available from: <http://www.zafh-servicorobotik.de/>
- [18] Zhang Z, Niu Y, Yan Z, Lin S. Real-Time Whole-Body Imitation by Humanoid Robots and Task-Oriented Teleoperation Using an Analytical Mapping Method and Quantitative Evaluation. *Applied Sciences*. 2018; 8(10): 2005. <https://doi.org/10.3390/app8102005>
- [19] Cai Q, et al. 3D deformable face tracking with a commodity depth camera. *European conference on computer vision*. Springer, Berlin, Heidelberg; 2010.
- [20] Zhang ZY. Microsoft kinect sensor and its effect. *IEEE Multimedia*. 2012; 19(2): 4-10. <https://doi.org/10.1109/MMUL.2012.24>
- [21] Itu-T P. Methods for subjective determination of transmission quality. *ITU-T Recommendation P.800*; 1996.