A Fast and Intuitive Kinematics Mapping Approach for Imitating the Motion of the Human Arm with an UR5 Robot

Rong-jian Liang, Zhang Chen, Yan-bo Tao, Bin Liang, Tao Zhang, Gang Li

***Abstract*—It is quite difficult to imitate the motion of human arms using non-humanoid robots due to their dissimilar embodiments[1] (degree-of-freedom (DOF), body morphology, constraints and so on); however, it will greatly enlarge the application scope of robotic imitation since many types of robots have non-humanoid structures. This paper addresses the procedure to imitate the motion of the human arm with an UR5, a kind of commonly-used industrial robot. The motion of the human arm is obtained by an inertial motion capture system, and then the captured motion is reproduced using the UR5's own embodiment. A virtual-joint-based approach is put forward to facilitate the fast and intuitive kinematics mapping between the human arm and the UR5 robot, leading to a robotic imitation system that can imitate the tip location and configuration of the human arm simultaneously. The proposed approach is verified experimentally on a real UR5 robot and compared with the Cartesian-space-based mapping approach.**

***Index Terms*—Robotic imitation, Kinematics Mapping, Dissimilar Embodiments, Programing by Demonstration, Motion Similarity.**

I. Introduction

Robotic imitation is a powerful tool in robotic programming and teleoperation. Robots can learn from the demonstration of human operators or other robots, which is often referred to as Learning by Imitation or Programing by Demonstration[2], to avoid intricate code programing. Also, programing by demonstration is more intuitive than the code programing so it can promote natural and efficient knowledge transfer from human to robots or between robots. With respect to the teleoperation[3] field, the imitation of human operators motion by robots at the remote workspace is essential to extend the humans sensing and manipulation capacity to the remote site, especially in the master-slave control mode.

One of key challenges in robotic imitation is kinematics mapping[4], especially across dissimilar bodies. Kinematics mapping, also referred to as the correspondence problem[5], is the process of transforming the captured motion of the demonstrator into the imitators own capacity. Many of existing robotic imitation systems employ humanoid robots[6, 7] that have similar bodies to their demonstrators, such as the human arm, so that kinematics mapping can be done by directly mapping the values of the DOFs of the demonstrator to that of the imitator. However, quiet a lot of types of robots have very different embodiments from humans', for example, industrial robots, space robots and even some of humanoid robots. On the one hand, dissimilar embodiments pose great challenge to the kinematics mapping; kinematics mapping solutions for non-humanoid robots can greatly enlarge application range of robotic imitation systems, on the other hand.

Chunxu Li et al.[8] fused the human motion information captured by the Kinect and inertial measurement device and then reproduced the motion using the humanoid Baxter Research Robot. Petar Kormushev et al.[9] used a free-standing humanoid robot to imitate the whole bodily motion of human beings. Imitation learning was used for training the upper body of the humanoid robot via kinesthetic teaching, while at the same time Reaction Null Space method was adopted for keeping the balance of the robot. NAO Humanoid Robot is commonly used in robotic imitation systems. A simplified geometric approach proposed by Pourya Shahverdi et al.[10] to solve the inverse kinematic problem of the upper-body of NAO robots. A whole-body imitation system was developed by Liang Zhang et al.[11] by fitting corresponding eight kinematic chains between the NAO robot and human body. Wael Suleiman et al[12] proposed an optimization framework to generate the upper body motion of humanoid robots from human captured motions, taking into account the humanoid physical capabilities. A robotic imitation system using a small-size humanoid robot–Darwin-OP was developed by Van Vuong Nguyen et al.[13]. Joint angles of the robot were found based on geometric relations. Chih-Lyang Hwang et al developed a humanoid robot imitation system, in which the neural-network-based inverse kinematics (IK) was employed to reduce the computation time and improve the accuracy of the time response of such system.

Imitation systems using non-humanoid robots often employ the Cartesian-space-based mapping approach[15, 16]-mapping the Cartesian coordinate of the end link of the demonstrator to that of the imitator. It can be applied to the imitation of human motion with non-humanoid robots that have totally different embodiments from human beings', but it needs inverse-kinematics computation and cant make full use of the human motion data. Xia Jing et al.[17] improved the similarity in the configurations of the demonstrator and the

R. Liang, Y. Tao, B. Liang, and T. Zhang were with the Department of Automation, Tsinghua University, Beijing, 100084, China e-mail: ($rj\_liang@qq.com$)

Z. Chen and Gang Li are with National Laboratory for Information Science and Technology, Shenzhen Graduate School of Tsinghua University,518055, Shenzhen, China.

This work was supported in part by the National Natural Science Foundation of China under Grant 61703228, and Grant 61673239; in part by the China Postdoctoral Science Foundation under Grant 2017M610895; in part by the Science and technology project of Shenzhen under Grant JCYJ20160428182227081

imitator in the Cartesian-space-based mapping approach by matching their workspace. Zhong Shi et al.[18] proposed a kinematics mapping method that combined the Cartesian-space-based metrics and joint-space-based metrics to achieve a global similarity of humanrobot motion. However, the joint-space-based metrics used in [18] required the robot having the same DOFs as human arms. A generic approach for kinematics mapping across dissimilar bodies was proposed by Aris Alissandrakis et al.[19, 20]. Body mappings were formalized using a unified (linear) approach via correspondence matrices to allow kinematics mapping across different embodiments. However, correspondence matrices can only lead to linear mapping solutions which are not powerful enough to be used to deal with complex tasks. Alissandrakis et al. also studied the metrics of similarity between demonstrators and imitators systematically[1, 21]. They divided the metrics for robotic imitation into three levels, i.e. action, state and effect level, and gave numerous examples to illustrate these concepts.

The main contribution of this paper is the development of an on-line imitation system of the human arm with an industrial robot, UR5, that can track the location of the human wrist and imitate the configuration of the human arm simultaneously. The motion of human arm is obtained with an inertial motion capture system, and then the captured motion is reproduced using the UR5's own embodiment. A virtual-joint-based approach is put forward to enable fast and intuitive kinematics mapping between the human arm and UR5. Experiments are conducted on real UR5 robot and the performance of the proposed kinematics mapping approach is compared with that of the Cartesian-space-based approach.

II. Inertial Motion Capture System

The modular Perception Neuron system is used for capturing the human motion in our research. It outputs the Biovision Hierarchy (BVH) data, which records the human pose as a nested-structure of parent-joints and child-joints, containing information that specifies the location of the skeletal joint relative to its parent-joint.

The upper-limb of human beings mainly consists of the shoulder complex, the elbow complex and the wrist joint[22], whose motions are shown in Fig. 1. Basically, the shoulder joint allows 3 DOFs motions, i.e., shoulder abduction/adduction, shoulder flexion/extension and internal/external rotation. The elbow complex allows 2 DOFs motionselbow flexion/extension and supination/pronation. The wrist joint allows 2 DOFs motionswrist flexion/extension and radial/ulnar deviation. In total, the human upper-limb can be modeled as a 7-DOF structure.

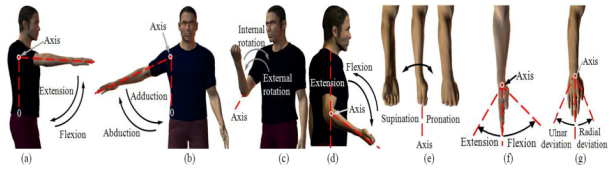


Fig. 1. motions of the human arm[22]

To facilitate further use of human data, the Shoulder-, Elbow- and Wrist-frame, are attached to the human arm (as shown in Fig. 2). The shoulder-frame moves with the shoulder complex, the Elbow-frame with the the elbow flexion/extension and the Wrist-frame with the elbow flexion/extension, wrist flexion/extension and radial/ulnar deviation. In this way, the motion of the human arm can be represented by the displacement and orientation of these frames, which can be obtained via the relative location specifications in the BVH data. l_1 in Fig. 2 refers to the length of human upper-arm and l_2 to the length of forearm. The states of the human arm can be represented as $\boldsymbol{S}^D = \left[\theta_1^D\ \theta_2^D\ \theta_3^D\ \theta_4^D\ \theta_5^D\ \theta_6^D\ \theta_7^D\right]^T$.

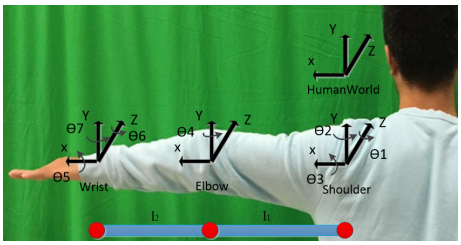


Fig. 2. frames on the human arm

III. UR5 Robot

An UR5 robot is used in this project as the imitator of the human arm. An UR5 robot has 6 links. In order to obtain the relations between successive links, the so-called Denavit-Hartenberg (DH) convention is adopted. Fig. 3 shows the frames attached to the UR5 robot according to the DH convention. Corresponding DH parameters of UR5 robot (as shown in Table I) can be obtained from the frames which are assigned on the links. The states of UR5 can be expressed as $\boldsymbol{S}^I = \left[\theta_1^I\ \theta_2^I\ \theta_3^I\ \theta_4^I\ \theta_5^I\ \theta_6^I\right]^T$.

TABLE I
DH parameters of UR5 robot

Link i	θ_i	$d_i[mm]$	$a_i[mm]$	$\alpha_i[rad]$
1	θ_1^I	89.16	0	$\pi/2$
2	θ_2^I	0	-425	0
3	θ_3^I	0	-392.25	0
4	θ_4^I	109	0	$\pi/2$
5	θ_5^I	93	0	$-\pi/2$
6	θ_6^I	82	0	0

The UR5 robot is mounted on a metal holder and the orientation of the base_link-frame of the UR5 is shown in Fig. 4. The roll, pitch and yaw (RPY) Eular-angle from the World-frame to the base_link-frame of the UR5 is $(\pi/2, 0, \theta_m)$, where θ_m is used as an optimization variable to enhance the imitation performance and would be determined in the next section.

The initial configuration is set as ${}^0\boldsymbol{S}^I = [{}^0\theta_1^I\ {}^0\theta_2^I\ {}^0\theta_3^I\ {}^0\theta_4^I\ {}^0\theta_5^I\ {}^0\theta_6^I]^T = [0\ \pi\ 0\ 0\ \frac{\pi}{2}\ 0]^T$ and shown in Fig. 4.

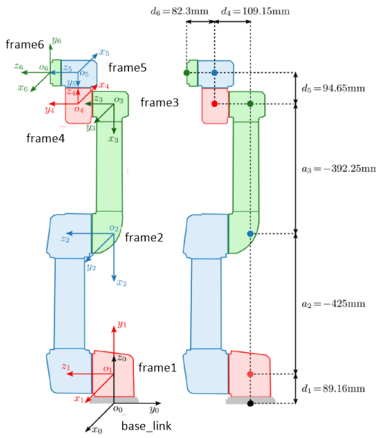


Fig. 3. frames on the UR5 robot[23]

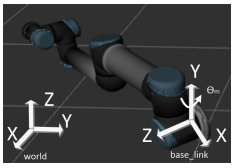


Fig. 4. initial configuration

IV. Virtual-joint-based kinematics mapping

A. General idea

From sections II and III it can be known that the human arm and the UR5 robot have very dissimilar embodiments, which poses great challenge to the kinematics mapping across them. To be specific, human arm has 7 DOFs while an UR5 robot has 6. What's more, their body morphologies are different, which means that it is difficult to find a satisfactory kinematics mapping solution even if one of the DOFs of human arm were ignored.

A virtual-joint-based approach is put forward to tackle the kinematics mapping across dissimilar embodiments. Firstly, the DOFs of the demonstrator (human arm) and the imitator (UR5 robot) are divided into the same number of groups. One group of DOFs is called one virtual joint, and a virtual joint of the demonstrator and the corresponding one of the imitator form a virtual joint pair. Then metrics of motion similarity can be defined at the virtual-joint-level. In other words, diverse types of metrics can be employed to measure the similarities between the virtual joints in one pair. And the kinematics mapping solution can be found by minimizing the mterics for all virtual joint pairs.

In the proposed system, the DOFs of the human arm and the UR5 are divided into three groups respectively.

$$\boldsymbol{S}^{D} = \left[\boldsymbol{S}^{D_1{}^T}\ \boldsymbol{S}^{D_2{}^T}\ \boldsymbol{S}^{D_3{}^T}\right]^T \tag{1}$$

$$\begin{cases} \boldsymbol{S}^{D_1} = \left[\theta_1^D\ \theta_2^D\ \theta_3^D\right]^T \\ \boldsymbol{S}^{D_2} = \left[\theta_4^D\right]^T \\ \boldsymbol{S}^{D_3} = \left[\theta_5^D\ \theta_6^D\ \theta_7^D\right]^T \end{cases} \tag{2}$$

$$\boldsymbol{S}^{I} = \left[\boldsymbol{S}^{I_1{}^T}\ \boldsymbol{S}^{I_2{}^T}\ \boldsymbol{S}^{I_3{}^T}\right]^T \tag{3}$$

$$\begin{cases} \boldsymbol{S}^{I_1} = \left[\theta_1^I\ \theta_2^I\right]^T \\ \boldsymbol{S}^{I_2} = \left[\theta_3^I\right]^T \\ \boldsymbol{S}^{I_3} = \left[\theta_4^I\ \theta_5^I\ \theta_6^I\right]^T \end{cases} \tag{4}$$

Consequently, three virtual joint pairs are formed. Then metrics of motion similarity can be defined for these three virtual joint pairs.

As pictured in Fig. 5, free vector $\boldsymbol{v}_{link2}$ is considered from the origin of the frame1 to that of the frame2 of the UR5, $\boldsymbol{v}_{link3}$ from the frame2 to the frame3, $\boldsymbol{v}_{upper-arm}$ from the Shoulder-frame to the Elbow-frame, $\boldsymbol{v}_{forearm}$ from the Elbow-frame frame to the Wrist-frame. θ_e is the angle down from the vector $\boldsymbol{v}_{link2}$ to vector $\boldsymbol{v}_{link3}$, and θ_4^D is the angle down from the vector $\boldsymbol{v}_{upper-arm}$ to the vector $\boldsymbol{v}_{forearm}$.

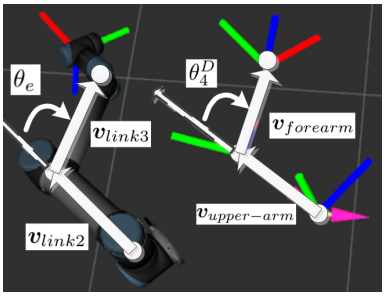


Fig. 5. Illustration of the metrics used in the first and second virtual joint pairs

The desired global performance is that $\boldsymbol{v}_{link2}$ tracks the direction of the human upper-arm, $\boldsymbol{v}_{link3}$ tracks the direction of the huamn forearm and the last three links of the UR5 robot perform a bit like a transformative hand, moving with the rotation of the human wrist. Three metrics are employed as follows to achieve the desired overall performance.

For the first virtual group pair, the metric used is presented as follows:

$$\boldsymbol{M}_1 = \arccos\left(\frac{\boldsymbol{v}_{link2}\cdot\boldsymbol{v}_{upper-arm}}{|\boldsymbol{v}_{link2}||\boldsymbol{v}_{upper-arm}|}\right) \tag{5}$$

where $\arccos(*)$ is the arc-cosine function that ranges from 0 to π.

For the second virtual joint pair, a joint-level metric is used:

$$\boldsymbol{M}_2 = ||\theta_e - \theta_4^D||_2 \tag{6}$$

$$\theta_4^D = \arccos\left(\frac{\boldsymbol{v}_{upper-arm}\cdot\boldsymbol{v}_{forearm}}{|\boldsymbol{v}_{upper-arm}||\boldsymbol{v}_{forearm}|}\right) \tag{7}$$

where $||*||_2$ means the L^2 norm.

For the third virtual joint pair, the rotation of the frame6 with respect to the frame3 of the UR5 robot is expected to equal to the rotation of the Wrist-frame with respect to Elbow-frame (as shown in Fig. 6). The metric used is presented as follows:

$$\boldsymbol{M}_3 = ||\boldsymbol{R}_c\boldsymbol{R}_{Wrist}^{Elbow}(\boldsymbol{R}_c)^{-1} - \boldsymbol{R}_{Frame6}^{Frame3}||_2 \tag{8}$$

where $\boldsymbol{R}_{Wrist}^{Elbow}$ represents the orientation of the Wrist-frame with respect to the Elbow-frame, $\boldsymbol{R}_{Frame6}^{Frame3}$ represents the orientation of the frame6 of the UR5 robot with respect to its frame3, and

$$\boldsymbol{R}_c = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \tag{9}$$

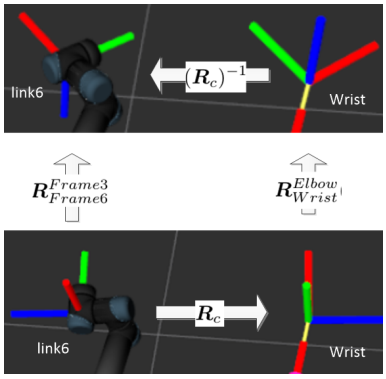


Fig. 6. Illustration of the metric used in the third virtual joint pair

B. Computational procedure

(1)For virtual joint pair#1 and the determination of θ_m

The goal of this step is to obtain the desired joint angles θ_1^I and θ_2^I by minimizing the metric (5) and also to determine θ_m.

A new frame, Joint1-frame, whose directions always pararel to those of the base_link-frame, is attacted to the origin of the frame1. Point O, A, B and P are attached to the body of UR5 as shown in Fig. 7. Then in the initial configuration of the UR5 robot, the locations of these points in the Joint1-frame are shown in Fig. 8(a). The desired joints angles θ_1^I and θ_2^I can be obstained as follows:

$$\theta_1^I = {}^0\theta_1^I - \angle AOC = -\angle AOC \tag{10}$$

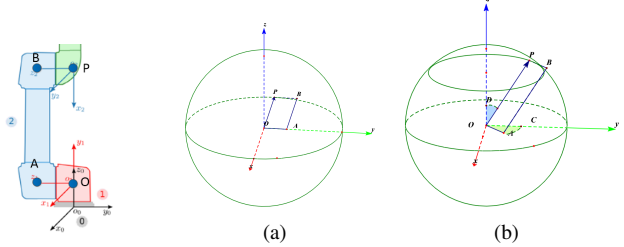
$$\theta_2^I = {}^0\theta_2^I + \frac{\pi}{2} - \angle POD = \frac{3\pi}{2} - \angle POD \tag{11}$$

$$\angle POD = \arccos\left(\frac{\boldsymbol{z}\cdot\overrightarrow{OP}}{|\boldsymbol{z}||\overrightarrow{OP}|}\right) \tag{12}$$

$$\angle AOC = \arccos\left(\frac{\boldsymbol{y}\cdot\overrightarrow{OA}}{|\boldsymbol{y}||\overrightarrow{OA}|}\right) \tag{13}$$

$$\overrightarrow{OA} = \overrightarrow{OP} * \boldsymbol{z} \tag{14}$$

where $\overrightarrow{OP} = \boldsymbol{v}_{link2}$, $\boldsymbol{z} = [0\ 0\ 1]^T$, $\boldsymbol{y} = [0\ 1\ 0]^T$, '$\cdot$' means the dot product, and '*' means the cross product.



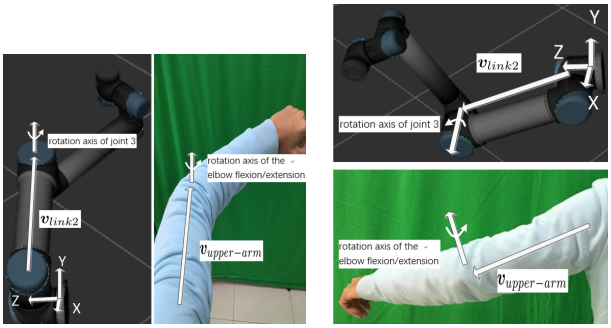
(a)

(b)

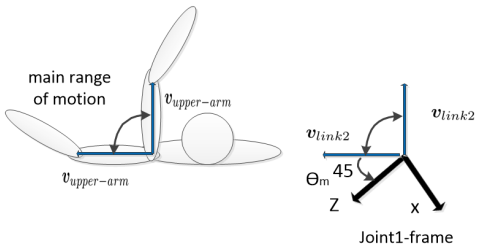
Fig. 7. Attached ponits on the UR5 virtual joint pair robot

Fig. 8. Illustration of the geometric relationship for first virtual joint pair

Though the solutions (10)-(14) presented above can bring the metric (5) to zero, it can't let the UR5 robot reproduce the shoulder internal/external rotation. The underlying cause lies in one less DOF in the first virual joint of the UR5 robot than that of the human arm. Ignoring the shoulder internal/external rotation will lead to great tracking errors of the location of the human wrist. To alleviate this effect, on the one hand, the human operator is required to avoid the shoulder internal/external rotation; on the other hand, a proper θ_m should be chosen to make the rotation axis of the elbow flexion/extension motion parallel to that of the joint 3 of UR5 as nearly as possible when the human arm moves in its wrokspace. We will show how to determine θ_m in the following.



(a) (b)



(c)

Fig. 9. Illustration of the determination of θ_m

According to the solutions (10)-(14), the rotation axis of the elbow flexion/extension would parallel to that of the joint 3 of the UR5 when $\boldsymbol{v}_{link2}$ moves in the x-y-plane of the Joint1-frame (as shown in Fig. 9(a)), while those two axes would be vertical to each other when $\boldsymbol{v}_{link2}$ moves in the y-z-plane (as shown in Fig. 9(b)). Consequently, its preferable that $\boldsymbol{v}_{link2}$, which follows the direction of $\boldsymbol{v}_{upper-arm}$, moves near the x-y-plane of Joint1-frame when the human arm moves in its wrokspace. As a result, the θ_m is choosen as $\pi/4$ according to the main range of human arms motion, as shown in Fig. 9(c).

(2) For virtual joint pair#2

The desired joint angle θ_3^I can be obtained directly with the following

$$\theta_3^I = {}^0\theta_3^I - \theta_e = -\theta_4^D \tag{15}$$

(3) For virtual joint pair#3

According to rotation order of the last three joints of the UR5 robot (as shown in Fig. 10), $\boldsymbol{R}^{Elbow}_{Wrist}$ has to be converted to the intrinsic X-Y-Z Euler-angle $(\alpha_W, \beta_W, \gamma_W)$ so that it can be executed by the UR5. Then the desired joint angles θ_4^I, θ_5^I and θ_6^I can be obtained as:

$$\theta_4^I = {}^0\theta_4^I + \alpha_W = \alpha_W \tag{16}$$

$$\theta_5^I = {}^0\theta_5^I + \beta_W = \pi/2 + \beta_W \tag{17}$$

$$\theta_6^I = {}^0\theta_6^I + \gamma_W = \gamma_W \tag{18}$$

V. Overall imitation procedure

The overall imitation procedure is shown in Fig. 11. Firstly, the human arms motion would be captured by the inertial Perception Neuron system, which outputs the BVH data. Then the transform data about the frames attached to the human arm can be obtained via the relative location specifications in the BVH data. Later the desired joint angles of the UR5 robot could be acquired via (10)-(18). Finally, the desired angles would be executed by the real UR5 robot.

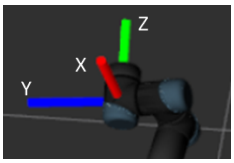


Fig. 10. Illustration of the rotation order of the last three joints of the UR5

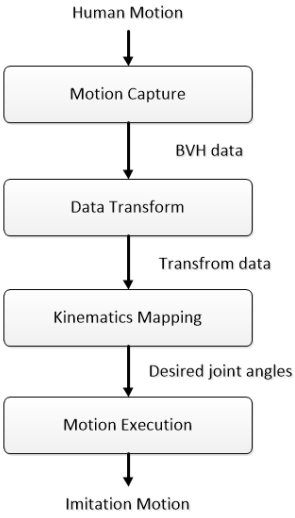


Fig. 11. Imitation procedure

VI. Experiment

A. Experiment settings

(1) System composition

The imitation system mainly consists of the human operator, the Perception Neuron motion capture system, kinematics mapping module and the UR5 robot, which is shown in Fig. 12. Communication among the latter three modules are done by using Robot Operating System ROS framework.

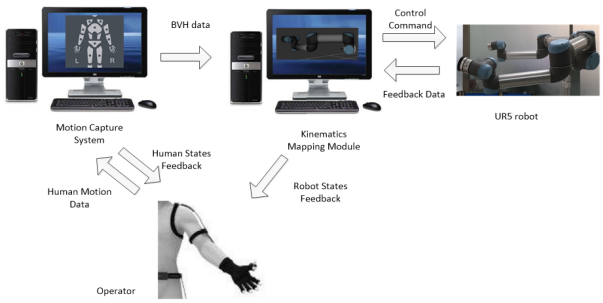


Fig. 12. Components of the imitation system

(2) Parameters settings

The Perception Neuron motion capture system will first map the human arms motion to the motion of an avatar, which then would be output via the BVH data files. In the practical experiment, the length of the upper-arm of the avatar is tuned as $l_1 = 0.7|a_2| = 297.5mm$, and the length of the forearm is tuned as $l_2 = 0.7|a_3| = 274.6mm$.

The velocity limit is set to 3.15 rad/s for every joint of the robot. The target states for the UR5 robot that violate the joint velocity limit would be abandoned.

(3) Comparion approach

The performance of the virtual-joint-based kinematics mapping approach is compared with that of the Cartesian-space-based mapping approach, which is commonly applied to imitation systems using non-humanoid robots. In the Cartesian-space-based approach, the mounting location and initial configuration are set the same as those in the proposed approach. $\boldsymbol{v}_{Wrist}^{Shoulder}$ is considered as the vector from the origin of Shoulder-frame to that of the Wrist-frame. $\boldsymbol{R}_{Wrist}^{HumanWorld}$ is orientation of Wrist-frame with respect to the HumanWorld-frame. Then the goal pose for the link6-frame of the UR5 robot is set as:

$$\begin{bmatrix} \boldsymbol{R}_{Wrist}^{HumanWorld}\boldsymbol{R}_c{}^{-1} & \frac{1}{0.7}\cdot\boldsymbol{v}_{Wrist}^{Shoulder} \\ \mathbf{0} & 1 \end{bmatrix} \tag{19}$$

where $\mathbf{0}$ is a 1×3 zero vector.

As we know, the UR5 robot has at most 8 inverse-kinematics (IK) solutions given a goal pose for its link6. For the stability and smoothness of the robot motion, the IK solution that closest to the robots current states is chosen as the desired joint angles.

(4) System input and collected data

The motion of human arms can be roughly divided into two modes, i.e., position-changing mode and orientation-changing mode. In the position-changing mode, the operator rotates his first and second virtual joint, consequently changing the location of the upper-arm and forearm, as well as the position of the Wrist-frame. In the orientation-changing mode, the operator rotates his third virtual joint, consequently changing the orientation of the Wrist-frame. Two demos of the human arms motion are recorded via the Perception Neuron system and then sent to the kinematics mapping module. In the first demo, the human arm mainly moves in the position-changing mode; while in the second demo, the human arm mainly moves in the orientation-changing mode. Both demos are used as inputs to the proposed approach and the Cartesian-space-based mapping approach.

To separate the imitation errors caused by kinematics mapping from those generated from execution, the desired joint angles obtained from kinematics mapping solutions, instead of the states of the real UR5 robot, are collected.

The following statistics are employed to evaluate the imitation performance.

a. Measuring the similarity in the tip locations:

It is of great importance to measure the similarity between tip location of the demonstrator and that of the imitator. The Cartesian coordinate of the vector $\boldsymbol{v}_{Wrist}^{Shoulder}$ is considered as

the tip position of human arm and recorded as (x^D, y^D, z^D). The orientation of Wrist-frame with respect to the World-frame is considered as the tip orientation of human arm and recorded as RPY angles (α^D, β^D, γ^D).

As for the tip position of the UR5 robot (x^I, y^I, z^I) in the proposed approach, it refers to the Cartesian coordinate of the vector from the origin of the frame1 to that of the frame3 of the UR5; while in the Cartesian-space-based mapping approach it refers to the coordinate of the vector from the origin of the base_link-frame to that of the frame6. In both approaches, the tip orientation of the UR5 robot $(\alpha^I, \beta^I, \gamma^I)$ refers to the the orientation of the frame6 of the UR5 with respect to the World-frame, transformed into the RPY angles.

The root-mean-square errors (RMSE) between the relative movement of the tip of the human arm and that of the UR5 robot can be computed as follows:

$$\begin{cases} RMSE\_dx = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n-1}(dx^D(i) - 0.7dx^I(i))^2} \\ dx^D(i) = x^D(i+1) - x^D(i) \\ dx^I(i) = x^I(i+1) - x^I(i) \end{cases} \tag{20}$$

$$\begin{cases} RMSE\_dy = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n-1}(dy^D(i) - 0.7dy^I(i))^2} \\ dy^D(i) = y^D(i+1) - y^D(i) \\ dy^I(i) = y^I(i+1) - y^I(i) \end{cases} \tag{21}$$

$$\begin{cases} RMSE\_dz = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n-1}(dz^D(i) - 0.7dz^I(i))^2} \\ dz^D(i) = z^D(i+1) - z^D(i) \\ dz^I(i) = z^I(i+1) - z^I(i) \end{cases} \tag{22}$$

$$\begin{cases} RMSE\_d\alpha = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n-1}(d\alpha^D(i) - d\alpha^I(i))^2} \\ d\alpha^D(i) = \alpha^D(i+1) - \alpha^D(i) \\ d\alpha^I(i) = \alpha^I(i+1) - \alpha^I(i) \end{cases} \tag{23}$$

$$\begin{cases} RMSE\_d\beta = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n-1}(d\beta^D(i) - d\beta^I(i))^2} \\ d\beta^D(i) = \beta^D(i+1) - \beta^D(i) \\ d\beta^I(i) = \beta^I(i+1) - \beta^I(i) \end{cases} \tag{24}$$

$$\begin{cases} RMSE\_d\gamma = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n-1}(d\gamma^D(i) - d\gamma^I(i))^2} \\ d\gamma^D(i) = \gamma^D(i+1) - \gamma^D(i) \\ d\gamma^I(i) = \gamma^I(i+1) - \gamma^I(i) \end{cases} \tag{25}$$

where n is the length of data. For the purpose of the accuracy of the statistics, the data collected when the robot keeps static are abandoned.

b. Measuring the similarity in configuration:

In the experiment, the angle $err_{upperarm}$ formed by $\boldsymbol{v}_{link2}$ and $\boldsymbol{v}_{upper-arm}$and the angle $err_{forearm}$ formed by $\boldsymbol{v}_{link3}$ and $\boldsymbol{v}_{forearm}$ are employed to measure the similarity in the configuration of the human arm and UR5 robot. The following statistics are computed:

$$\begin{cases} RMSE\_upperarm = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(err_{upperarm}(i))^2} \\ err_{upperarm}(i) = \arccos(\frac{\boldsymbol{v}_{link2}(i)\cdot\boldsymbol{v}_{upper-arm}}{|\boldsymbol{v}_{link2}||\boldsymbol{v}_{upper-arm}(i)|}) \end{cases} \tag{26}$$

$$\begin{cases} RMSE\_forearm = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(err_{forearm}(i))^2} \\ err_{forearm}(i) = \arccos(\frac{\boldsymbol{v}_{link3}(i)\cdot\boldsymbol{v}_{forearm}(i)}{|\boldsymbol{v}_{link3}(i)||\boldsymbol{v}_{forearm}(i)|}) \end{cases} \tag{27}$$

c. Measuring the smoothness of imitation motion:

The following statistic is used to measure the smoothness of the imitation motion.

$$\begin{cases} MEAN\_SumJoints = \frac{1}{n-1}\sum_{i=1}^{n-1}\sum_{j=1}^{6}|d\theta_j^I(i)| \\ d\theta_j^I(i) = \theta_j^I(i+1) - \theta_j^I(i) \end{cases} \tag{28}$$

d. Measuring the time efficiency of algorithms:

The consuming time of kinematics mapping, which consists in the time for listening to the transform data and that for the computation of the desired joint angles for the UR5 robot, is recorded with the help of the clock() function provided by the C Time library. The average of the consuming time is computed as below.

$$MEAN\_time = \frac{1}{n}\sum_{i=1}^{n} t(i) \tag{29}$$

where $t(i)$ is the consuming time of the ith kinematics mapping.

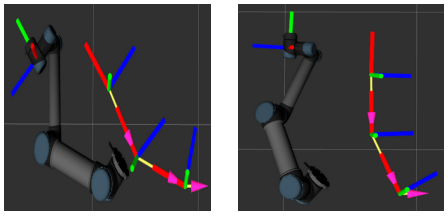
B. Experiment results

The statistical data is shown in Table.II. To visualize the mapping rsults, the poses of the human arm and the UR5 robot were shown together in RVIZ(as shown in Fig. 13 and Fig. 14).

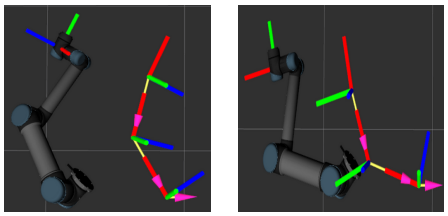
It can be known that, with the proposed mapping approach, the frame3 of the UR5 robot tracks the displacement of the Wrist-frame successfuly in both demos, and the the position tracking errors along the Z axis are a litter larger than those along the X and Y axes. The orientation of frame6 imitates that of the Wrist-frame at a high accuracy. The $RMSE\_d\alpha$, $RMSE\_d\beta$ and $RMSE\_d\gamma$ are all less than 1°. The angle $err_{upperarm}$ is close to zero all the time. The angle $err_{forearm}$ is about 5° in the demos. The $MEAN\_SumJoints$ is about 4° in the first demo and 3° in the second demo. The consuming time for kinematics mapping is less than 0.2ms in most cases. The $MEAN\_time[ms]$ is about 0.1ms.

With the Cartesian-space-based mapping approach, the frame6 of the robot tracks the location of Forearm-frame perfectly. The angle $err_{upperarm}$ and angle $err_{forearm}$ are much larger than those in the virtual-joint-based approach. The $RMSE\_err_{upperarm}$ is about 10° and the $RMSE\_err_{forearm}$ is about 23°. In the first demo, the $MEAN\_SumJoints$ in the two approaches are pretty close. But in the second demo, the $MEAN\_SumJoints$ is 13% larger in the Cartesian-space-based mapping approach. Also, it can be known that the consuming time of the Cartesian-space-based mapping approach is larger than that of the proposed approach. The average of consuming time of the Cartesian-space-based approach is about 50% larger than that in the proposed approach, though they are both less than 0.2ms.

* Approach#1 refers to the proposed mapping approach, approach#2 refers to the Cartesian-space-based mapping approach, demo#1 refers to the first demo and demo#2 refers to the second demo.

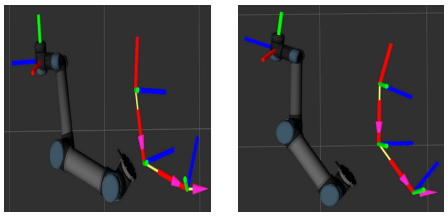


(a) pose1 (b) pose2

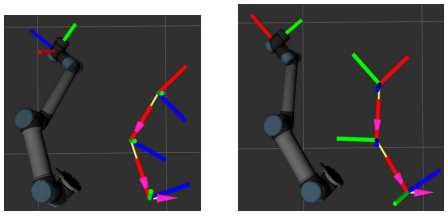


(c) pose3 (d) pose4

Fig. 13. mapping results with the Cartesian-space-based approach



(a) pose1 (b) pose2



(c) pose3 (d) pose4

Fig. 14. mapping results with the virtual-joint-based approach

TABLE II
EXPERIMENT RESULTS

statistics	approach#1 in demo#1	approach#2 in demo#1	approach#1 in demo#2	approach#2 in demo#2
$RMSE\_dx[mm]$	0.3117	5.15e-4	0.0522	5.18e-4
$RMSE\_dy[mm]$	0.5468	4.20e-4	0.2123	5.02e-4
$RMSE\_dz[mm]$	2.3279	3.22e-4	1.1346	5.24e-4
$RMSE\_d\alpha[^\circ]$	0.4171	3.84e-4	0.2666	3.07e-4
$RMSE\_d\beta[^\circ]$	0.6728	0	0.2011	0
$RMSE\_d\gamma[^\circ]$	0.3796	0	0.1527	0
$RMSE\_upperarm[^\circ]$	7.17e-5	12.7535	6.91e-5	8.2251
$RMSE\_forearm[^\circ]$	3.5605	23.4819	6.2294	22.1848
$MEAN\_SumJoints[^\circ]$	4.1739	4.2360	2.9939	3.4019
$MEAN\_time[ms]$	0.0966	0.1502	0.1144	0.1561

C. Discussions

From the experiment results it can be known that both approaches have a good performance in imitating the location of the human wrist, but the Cartesian space based approach has a higher accuracy. The location tracking errors of Cartesian-space-based approach are mainly caused by calculation errors. In the proposed approach, the location tracking errors are resulted mainly from the dissimilar embodiments in the first virtual joint pair. To be specific, the first virtual joint of the human arm has one more DOF than that of the UR5 robot, and their morphologies are different. Consequently, the shoulder internal/external rotation can't be imitated by the UR5, which would lead to tracking errors between the tip locations of the human arm and UR5 robot. The relatively larger position tracking errors along the Z direction is an evidence for this, because the shoulder internal/external rotation affects the wrists position along the Z direction most significantly.

The proposed approach does a much better job in imitating the configuration of human arm than the Cartesian-space-based mapping approach. In the proposed approach, the direction of the vector $\boldsymbol{v}_{link2}$ is the same as that of the human upper-arm, and the direction of $\boldsymbol{v}_{link3}$ is fairly close to that of the forearm. But in the Cartesian-space-based mapping approach, the similarity in the configuration of the UR5 and that of the human arm can't be guaranteed, even the size parameters of the avatar in motion capture system are well tuned and a proper initial configuration is chosen.

In the first demo, two approaches have similar performance in the smoothness of robot motion; while in the second demo, the proposed approach does a better job than the Cartesian-space-based approach in terms of the motion smoothness. In the second demo, the human operator mainly moves his third virtual joint of arm, corresponding to the third virtual joint of the UR5 robot in the proposed approach. So in the proposed approach, mainly the last three joints of the robot have to move to imitate the human arm's motion in the second demo while all six joints have to be used in the Cartesian-space-based approach.

Also, the proposed approach has a higher computation efficiency. In the proposed approach, the overall kinematics mapping problem is decomposed into finding mapping solutions for three virtual joint pairs, so that the inverse-kinematics for a 6-DOF structure is avoid. In the Cartesian space based approach, the kinematics mapping mainly consists of three parts-listening to the transform data, calculating all feasible inverse-kinematics solutions and choosing a solution closest to the current states as the desired joint angles for UR5. The time on the first part is also needed in the proposed approach, but the proposed approach needs less time to calculate the desired joint angles for UR5 robot.

In the proposed approach, the position of human wrist is tracked by the origin of the frame3 of the UR5 robot, and the last three links of robot move with the rotation of the human wrist, acting like a transformative hand. In the applications where the link6 of an UR5 robot is controlled by a person to reach a specific posture, the human arm can firstly move in the orientation-changing mode to direct the link6 of the robot to the desired orientation, then work in the position-changing mode to reach the specific position. Also, the intuitive correspondence between the virtual joints of the human arm and those of the UR5 robot in the proposed approach would help the human operator control the robot in a more natural way.

To sum up, compared with the Cartesian-space-based approach, the proposed approach has better performance on

imitating the configuration of the human arm, better motion smoothness and higher computation efficiency, at the price of the less satisfying performance in tracking the location of the human wrist. In the applications where the tracking accuracy of the tip location is concerned mainly, the Cartesian-space-based approach is preferable. But in those where the similarities in the tip location and configuration are both required, the proposed apprroach has advantages over the Cartesian-space-based approach.

TABLE III
COMPARISON OF TWO KINEMATICS MAPPING APPROACHES

	similarity in the tip locations	similarity in configuration	motion smoothness	time efficiency
approach#1		better	better	better
approach#2	better			

VII. CONCLUSION

In this paper, an on-line imitation system of the human arm with an UR5 robot is developed, in which the tip location and configuration of the human arm are imitated simultaneously with the robot. A virtual-joint-based approach is proposed to tackle the kinematics mapping between dissimilar embodiments and applied to the mapping betwen the human arm and the UR5 robot in this research. Experiments on a real UR5 robot show that, compared with the Cartesian-space-based approach, the proposed kinematics mapping approach has better performance in imitating the configuration of the human arm, better motion smoothness and higher computation efficiency, at the price of the less satisfying performance in tracking the location of the human wrist. Also, the intuitive correspondence between the virtual joints of the human arm and that of the UR5 would help the human operator control the robot in a more natural way. The authors hope that the work has some reference significance for developing robotic imitation systems using other types of non-humanoid robots.

Many interesting works still lie behind, for instance, formulating the virtual-joint-based approach into a generic kinematics mapping method across dissimilar embodiments with the help of mathematical tools, capturing the human motion using non-contact equipments, and dealing with effect of tremble in human motion on the imitation performance.

REFERENCES

[1] Alissandrakis A, Nehaniv C L, Dautenhahn K. Correspondence mapping induced state and action metrics for robotic imitation[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 2007, 37(2): 299-307.

[2] A. Billard and R. Siegwart, Robot learning from demonstration,Robotics and Autonomous Systems, vol. 47:2-3, pp. 6567, 2004.

[3] Liarokapis M V, Artemiadis P K, Kyriakopoulos K J. Mapping human to robot motion with functional anthropomorphism for teleoperation and telemanipulation with robot arm hand systems[C]//Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on. IEEE, 2013: 2075-2075.

[4] Zhao X, Huang Q, Peng Z, et al. Kinematics mapping and similarity evaluation of humanoid motion based on human motion capture[C]//Intelligent Robots and Systems, 2004.(IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on. IEEE, 2004, 1: 840-845.

[5] C. Nehaniv and K. Dautenhahn, Of hummingbirds and helicopters: An algebraic framework for interdisciplinary studies of imitation and its applications, in Interdisciplinary Approaches to Robot Learning, J. Demiris and A. Birk, Eds. World Scientific Press, 2000, vol. 24, pp. 136161.

[6] Calinon S, Billard A. Incremental learning of gestures by imitation in a humanoid robot[C]//Proceedings of the ACM/IEEE international conference on Human-robot interaction. ACM, 2007: 255-262.

[7] Lin H I, Liu Y C, Lin Y H. Intuitive Kinematic Control of a Robot Arm via Human Motion[J]. Procedia Engineering, 2014, 79: 411-416.

[8] Li C, Yang C, Wan J, et al. Neural learning and Kalman filtering enhanced teaching by demonstration for a Baxter robot[C]//Automation and Computing (ICAC), 2017 23rd International Conference on. IEEE, 2017: 1-6.

[9] Kormushev P, Nenchev D N, Calinon S, et al. Upper-body kinesthetic teaching of a free-standing humanoid robot[C]//Robotics and Automation (ICRA), 2011 IEEE International Conference on. IEEE, 2011: 3970-3975.

[10] Shahverdi P, Masouleh M T. A simple and fast geometric kinematic solution for imitation of human arms by a NAO humanoid robot[C]//Robotics and Mechatronics (ICROM), 2016 4th International Conference on. IEEE, 2016: 572-577.

[11] Zhang L, Cheng Z, Gan Y, et al. Fast human whole body motion imitation algorithm for humanoid robots[C]//Robotics and Biomimetics (ROBIO), 2016 IEEE International Conference on. IEEE, 2016: 1430-1435.

[12] Suleiman W, Yoshida E, Kanehiro F, et al. On human motion imitation by humanoid robot[C]//Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on. IEEE, 2008: 2697-2704.

[13] Lee J H. Full-body imitation of human motions with kinect and heterogeneous kinematic structure of humanoid robot[C]//System Integration (SII), 2012 IEEE/SICE International Symposium on. IEEE, 2012: 93-98.

[14] Hwang C L, Chen B L, Syu H T, et al. Humanoid Robot's Visual Imitation of 3-D Motion of a Human Subject Using Neural-Network-Based Inverse Kinematics[J]. IEEE Systems Journal, 2016, 10(2): 685-696.

[15] Ardakani M M G, Cho J H, Johansson R, et al. Trajectory generation for assembly tasks via bilateral teleoperation[J]. IFAC Proceedings Volumes, 2014, 47(3): 10230-10235.

[16] Yoon W K, Goshozono T, Kawabe H, et al. Model-based space robot teleoperation of ETS-VII manipulator[J]. IEEE Transactions on Robotics and Automation, 2004, 20(3): 602-612.

[17] Xia J, Jiang Z, Liu H, et al. Optimizing motion mapping

of robonaut's arms in its teleoperation[J]. Huazhong Keji Daxue Xuebao(Ziran Kexue Ban)/ Journal of Huazhong University of Science and Technology(Nature Science Edition), 2012, 40(12).

[18] Shi Z, Huang X, Hu T, et al. Weighted augmented Jacobian matrix with a variable coefficient method for kinematics mapping of space teleoperation based on humanrobot motion similarity[J]. Advances in Space Research, 2016, 58(7): 1401-1416.

[19] Alissandrakis A, Nehaniv C L, Dautenhahn K, et al. An approach for programming robots by demonstration: Generalization across different initial configurations of manipulated objects[C]//Computational Intelligence in Robotics and Automation, 2005. CIRA 2005. Proceedings. 2005 IEEE International Symposium on. IEEE, 2005: 61-66.

[20] Alissandrakis A, Nehaniv C L, Dautenhahn K. Imitation with ALICE: Learning to imitate corresponding actions across dissimilar embodiments[J]. IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 2002, 32(4): 482-496.

[21] Alissandrakis A, Nehaniv C L, Dautenhahn K. Action, state and effect metrics for robot imitation[C]//Robot and Human Interactive Communication, 2006. ROMAN 2006. The 15th IEEE International Symposium on. IEEE, 2006: 232-237.

[22] Gopura R, Kiguchi K. Mechanical designs of active upper-limb exoskeleton robots: State-of-the-art and design difficulties[C]//Rehabilitation Robotics, 2009. ICORR 2009. IEEE International Conference on. IEEE, 2009: 178-187.

[23] Gravdahl J T. Force Estimation in Robotic Manipulators: Modeling, Simulation and Experiments[J]. 2014.