

仿人型服务机器人双臂动作示教及评估研究

张续冲, 杨朋

(华南理工大学, 广州 510006)

摘要: **目的** 基于自主研发的仿人型双臂服务机器人, 研究体感示教及动作再现技术, 并评估机器人拟人动作的合理性。 **方法** 首先根据机器人的构型, 利用 D-H 法对机器人进行运动学建模。然后设定 6 种日常生活中常见的人体上肢动作, 利用 Kinect V2 动作捕捉设备获取真人演示动作数据, 将臂形角作为冗余量进行运动学逆解, 求得机器人所需关节角, 使机器人能够复现人体动作。之后基于恐怖谷理论, 结合层次分析法和语义分析法, 提出评估方法, 评估机器人复现的动作。 **结果** 机器人以拟人的姿态复现了人体示教动作, 评估结果显示, 不同类型的动作会让受试者产生不同的心理感受。 **结论** 对于仿人型双臂服务机器人而言, 体感示教是一种自然高效的示教方式, 由此产生的行为符合用户心理预期, 为服务机器人融入日常生活提供了理论依据。

关键词: 仿人型机器人; 体感示教; 肢体语言; 恐怖谷理论

中图分类号: TB472 **文献标识码:** A **文章编号:** 1001-3563(2021)08-0030-06

DOI: 10.19554/j.cnki.1001-3563.2021.08.004

Teaching and Evaluation of Dual Arm Action of Humanoid Service Robot

ZHANG Xu-chong, YANG Peng

(South China University of Technology, Guangzhou 510006, China)

ABSTRACT: The work aims to study the somatosensory teaching and motion reproduction technology, and evaluate the rationality of humanoid action of the robot based on the self-developed humanoid dual-arm service robot. Firstly, according to the configuration of the robot, the kinematic model of the robot is established by D-H method. Then, set six kinds of common upper limb movements in daily life, use Kinect V2 motion capture device to obtain real-life demonstration motion data, perform kinematic inverse solution with arm angle as redundancy to obtain the required joint angle of the robot, and make the robot can reproduce the human action. After that, based on the theory of Uncanny Valley, combined with the Analytic Hierarchy Process (AHP) and the Semantics Differential (SD), the evaluation method is proposed to evaluate the repetition action of the robot. The results show that the robot reproduces the teaching action of human body in a humanoid posture. The evaluation results show that different types of actions will make the subjects have different psychological feelings. For humanoid dual-arm service robot, somatosensory teaching is a natural and efficient way of teaching. The resulting behavior meets the psychological expectation of the user, which provides a theoretical basis for the service robot to integrate into daily life.

KEY WORDS: humanoid robot; somatosensory teaching; body language; uncanny valley

仿人机器人的交互设计必定在某些方面延续着人与人交互的意识形态, 因此肢体语言可以作为人机交互的媒介^[1]。在人与机器人交互 (Human-Robot Interaction, HRI) 的过程中, 机器人运动的拟人化是

实现安全有效 HRI 的关键组成部分, 有助于提高机器人的接受度和运动可预测性^[2], 人也会被机器人的肢体情绪所感染^[3]。ZHANG 等人^[4]基于模糊层次分析法研究了 HRI 的交流氛围。郝向阳^[5]利用 Kinect

收稿日期: 2021-01-15

基金项目: 广东省自然科学基金资助项目 (2020A1515010397)

作者简介: 张续冲 (1989—), 男, 山东人, 博士, 华南理工大学副教授, 主要研究方向为仿人机器人、人机交互技术。

实现了七自由度双臂机器人的拟人运动,但并未研究示教动作可能给人造成的感受。LIU W^[6]针对双臂服务机器人进行了可用性测试并制定了评估方法。钟翔伟等人^[7]以虚拟机器人为研究对象,对仿人机器人交互行为设计了评估方法,但未在实体机器人上验证。本文基于自主研发的仿人型双臂服务机器人,借助 KinectV2 动作捕捉设备获取真人动作,将动作映射到仿人型双臂机器人,以实现拟人的姿态动作,并评估机器人肢体行为表现,为 HRI 过程中肢体语言表达提供借鉴。

1 仿人型双臂机器人构型及参数

仿人型双臂机器人 Player 见图 1^[8]。人的手臂可以等效为 S-R-S 构型,具有七个自由度,冗余的自由度可以让手臂具有更丰富的姿态和更强的避障能力。机器人的双臂各由七个转动关节串联而成。前三个关节轴线相交,组成了肩关节,第四个关节为肘关节,后三个关节轴线也相交,组成了腕关节。Player 两肩距离 $2L_0=320$ mm,上臂长度 $L_1=270$ mm,前臂长度 $L_2=240$ mm,手掌中心距腕关节长度 $L_3=100$ mm。采用 D-H 法^[9]进行运动学建模,双臂对应的 D-H 参数及各关节转动范围,见表 1。

相邻关节变换矩阵表达式为:

$${}^{i-1}T_i(\theta_i) = \begin{bmatrix} c\theta_i & -s\theta_i & 0 & \alpha_{i-1} \\ s\theta_i c\alpha_{i-1} & c\theta_i c\alpha_{i-1} & -s\alpha_{i-1} & -d_i s\alpha_{i-1} \\ s\theta_i s\alpha_{i-1} & c\theta_i s\alpha_{i-1} & c\alpha_{i-1} & d_i c\alpha_{i-1} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (1)$$

其中: $i=1\sim 8$, $s\theta_i=\sin\theta_i$, $c\theta_i=\cos\theta_i$ 。全局基坐标系 $\{B\}$ 到右臂坐标系 $\{0\}$ 的变换矩阵为 ${}^B T_0^R$ 、到左臂坐标系 $\{0\}$ 的变化矩阵 ${}^B T_0^L$ 及 Player 机器人运动学逆解过程见参考文献[8],本文不再赘述。

2 基于 KinectV2 的动作示教

为实现机器人拟人动作复现,人体需要尽可能以自然的状态运动,本文选择 KinectV2 体感相机获取人体动作数据。Kinect V2 利用结构光和相关技术获取物体的深度信息,配合彩色摄像头提供的彩色信息,可以获得物体的 3D 数据^[10]。人体骨骼可以等效为全身的 25 个关节点,受试者无需穿戴设备,避免肢体动作被干扰。其主要技术参数包括:彩色相机分辨率为 1920 像素×1080 像素,帧率为 30 帧/s,视野范围为 $70^\circ\times 60^\circ$,有效测量范围为 0.5~4.5 m,满足动作捕捉要求。

Kinect V2 捕捉到的关节数据不能直接应用到机器人上。以人体自然状态下的一种姿势为例,臂形角示意图 2, Kinect V2 捕捉到的关节点 B、S、E、W 和 H 分别为颈部、肩关节、肘关节、腕关节和手掌

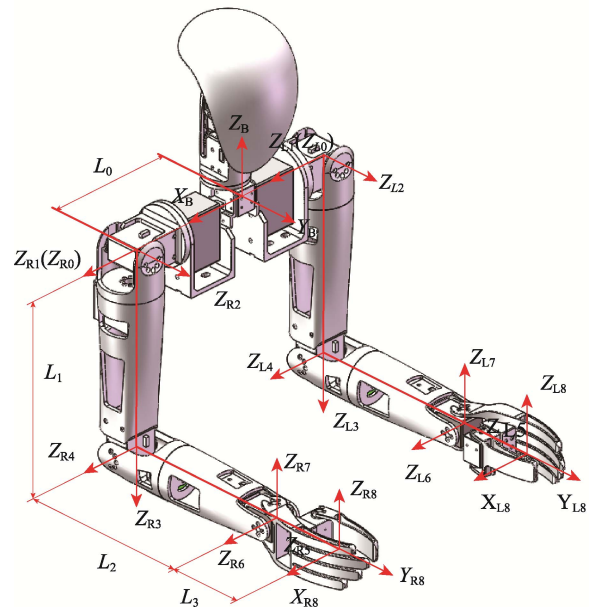


图 1 仿人型双臂机器人 Player
Fig.1 Humanoid dual-arm robot Player

表 1 D-H 参数及转角范围
Tab.1 D-H parameters and angle range

i	a_{i-1}	α_{i-1}	d_i	θ_i	θ_{θ_i} 范围	θ_{θ_i} 范围
1	0	0	0	θ_1	$[-45,180]$	$[-45,180]$
2	0	-90	0	θ_2	$[0,100]$	$[80,180]$
3	0	90	L_1	θ_3	$[0,180]$	$[0,180]$
4	0	-90	0	θ_4	$[0,125]$	$[0,125]$
5	0	90	L_2	θ_5	$[0,180]$	$[0,180]$
6	0	-90	0	θ_6	$[70,110]$	$[70,110]$
7	0	90	0	θ_7	$[-75,75]$	$[-75,75]$
8	$-L_3$	0	0	90	—	—

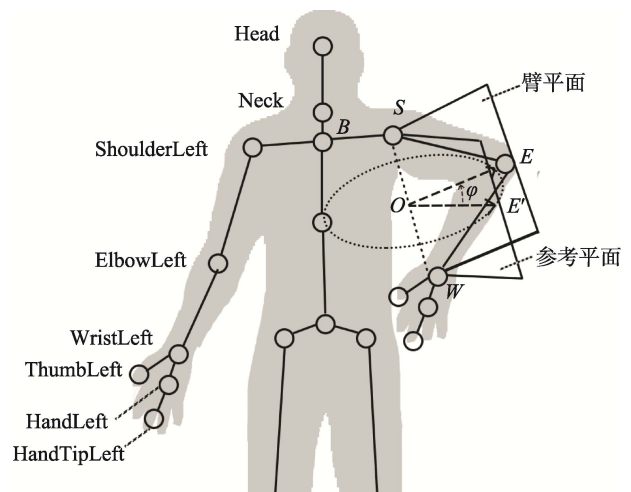


图 2 臂形角示意
Fig.2 Schematic of arm angle

的位置。肩关节 S、肘关节 E 和腕关节 W 构成了一个臂平面,在手部位置和姿态不变的情况下,人手臂可以通过调整臂平面达到避障的目的。定义关节角为

零时所在的臂平面为参考平面,实际的臂平面与参考平面的夹角为臂形角 φ 。因此,需要将捕捉的数据转化为手臂末端的位置、姿态和臂形角,以完全描述手臂的运动供机器人使用。由于人臂尺寸和机器人尺寸不一致,工作空间不一样,所以手臂末端位置映射到机器人时需要缩放,此处的缩放比例为人手臂长度与机器人手臂长度的比值。

臂形角 φ 的求解,公示推导过程以右手为例,左手类同。需要首先确定参考平面,根据公式(1)可得肘关节坐标系 $\{3\}$ 相对于肩关节坐标系 $\{0\}$ 的齐次变换矩阵:

$${}^0T_3 = {}^0T_1 {}^1T_2 {}^2T_3 = \begin{bmatrix} {}^0r_3 & {}^0p_3 \\ 0 & 1 \end{bmatrix} \quad (2)$$

当 θ_3 为零时,肘关节 E' 在参考平面上,此时 E' 在坐标系 $\{0\}$ 的位置可由下式求得:

$${}^0p_3 = \begin{cases} x_3 = L_1 c_1 s_2 \\ y_3 = L_1 s_1 s_2 \\ z_3 = L_1 c_2 \end{cases} \quad (3)$$

其中: $s_1=\sin\theta_1$, $c_1=\cos\theta_1$, θ_2 类似。 θ_1 和 θ_2 可由腕关节位置逆解求得。则参考平面的法向量计算为:

$$\mathbf{n}_1 = \overrightarrow{SW} \times \overrightarrow{SE'} \quad (4)$$

而实际臂平面的法向量计算为:

$$\mathbf{n}_2 = \overrightarrow{SW} \times \overrightarrow{SE} \quad (5)$$

则 $\mathbf{n}_1$ 、 $\mathbf{n}_2$ 的夹角即为臂形角 φ ,由此可得:

$$\varphi = \cos^{-1} \left(\frac{\mathbf{n}_1 \cdot \mathbf{n}_2}{|\mathbf{n}_1| |\mathbf{n}_2|} \right) \quad (6)$$

手部姿态求解,腕关节点 W 、手掌点 H 和拇指点 T 构成一个平面,定义手掌点 H 和拇指点 T 的中点 M 为名义手掌中心,其位置可以由 H 、 T 的坐标直接求得。记 $\mathbf{Y}_M$ 为 $\overrightarrow{WM}$ 的单位向量,记 $\mathbf{Z}_M$ 为过 M 点且垂直于 WHT 平面的单位向量,记 $\mathbf{X}_M$ 为与 $\mathbf{Y}_M$ 和 $\mathbf{Z}_M$ 正交的单位向量,方向符合笛卡尔坐标系,手部姿态描述见图3。

则手部在基坐标系 $\{B\}$ 的姿态矩阵表示为:

$${}^B\mathbf{r}_M = [\mathbf{X}_M, \mathbf{Y}_M, \mathbf{Z}_M] \quad (7)$$

通过以上推导,将 KinectV2 捕捉到的数据转换

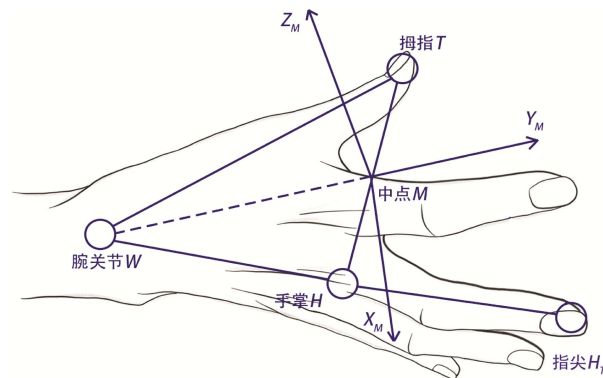


图3 手部姿态描述
Fig.3 Description of hand pose

为手臂末端的位置、姿态和臂形角。末端位置比例缩放后,经过机械臂运动学逆解,便可求解出机械臂运动所需的关节角度。仿人双臂机器人执行指令后,便可复现人演示的动作。

3 上肢拟人动作评估研究

根据先前的研究,机器人的肢体动作可以作为人机交互过程中的一种媒介存在^[11],同时机器人的肢体语言会使人获得更加舒适的交互体验^[12]。在 HRI 领域,机器人在模仿人的过程中往往会产生恐怖谷效应^[13],在设计时需要避免落入谷底。因此动作示教时,选取日常生活中常见的肢体动作,基于恐怖谷理论结合层次分析法和语义分析法构建评价标准,邀请志愿者评估仿人服务机器人复现的人体动作,并对评估结果进行分析。

3.1 示教动作确定

示教选用常见的 6 种上肢动作,其中 T01~T03 分别是“右手挥手”“主动握手”“举手示意”,均为单手动作;T04~T06 分别是“惊讶动作”“轮流出拳”“双手抱头”,均为双手动作。需要说明的是,示教的动作是连续的,如右手挥手表示举起右手并左右挥动,其他类同。截取动作片段,6 种人体上肢动作见图 4。

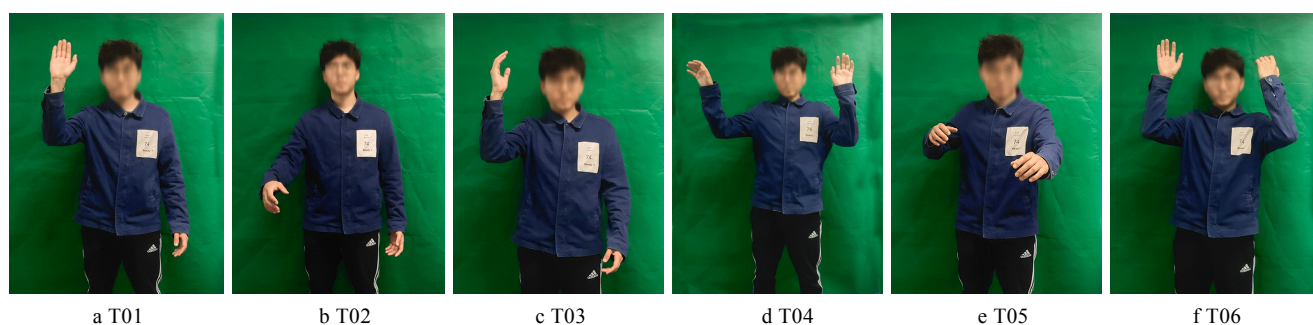


图4 6种人体上肢动作
Fig.4 Six upper limb movements

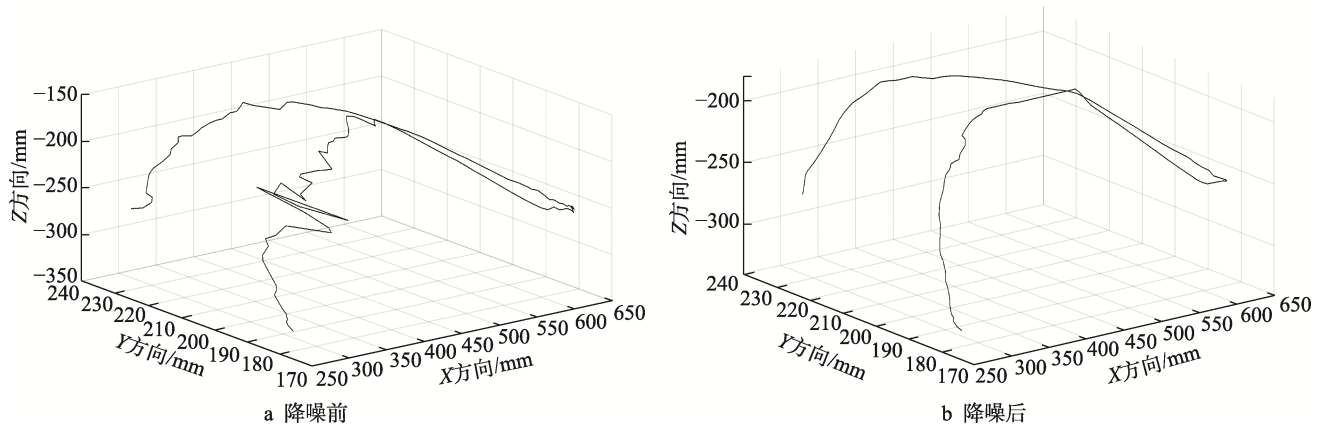


图 5 腕部关节运动轨迹
Fig.5 Motion trajectory of wrist joint

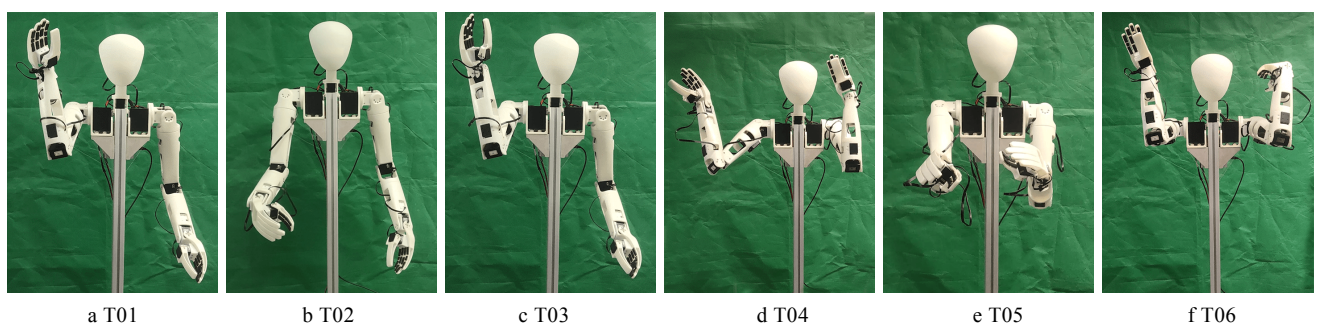


图 6 Player 复现的上肢动作
Fig.6 Upper limb movement of robot Player reproduction

3.2 轨迹降噪及动作再现

Kinect V2 捕捉的人体动作数据不可避免存在噪音，为了兼具处理效率和动作的拟人性，选用中值滤波法对动作数据进行降噪。此处以腕部关节运动轨迹为例，见图 5，降噪前腕部运动的空间轨迹如图 5a。

经过中值滤波法降噪后的腕部运动轨迹如图 5b。可以看出，降噪后的轨迹光滑流畅，并且保留了运动特征，保证了运动轨迹的拟人性。将降噪后的数据发送给 Player 机器人，完成机器人对肢体动作的复现，截取动作片段，见图 6，动作顺序与图 4 对应。

3.3 评估方法定义及计算

产生恐怖谷的原因主要包含两个方面：人的低级生物觉和人的高级生物觉^[14]。低级生物觉，指生物在进化过程中积累的本能，包括恐惧死亡、厌恶疾病等。高级生物觉，指人特有的移情思维，具体包括认知移情、运动移情和情感移情。通过语义范畴化，构建 5 种感性评价因子，将“对死亡的恐惧”概括为词对：有生命—无生命（代号 L）。将“对病态异类的排斥”概括为词对：健康的—病态的（代号 H）。将“认知移情”概括为词对：易理解—难理解（代号 U）。将“运动移情”概括为词对：易模仿—难模仿（代号 S）。将“情感移情”概括为词对：易共鸣—难共鸣（代号 R）。感性评价因子见表 2。

表 2 感性评价因子
Tab.2 Perceptual evaluation factor

分类	代号	评价因子	形容词对
低级生物觉	L	对死亡的恐惧	有生命—无生命
	H	对病态异类的恐惧	健康的—病态的
高级生物觉	U	认知移情	易理解—难理解
	S	运动移情	易模仿—难模仿
	R	情感移情	易共鸣—难共鸣

根据层次分析法拟定心理总量计算公式：

$$K_I = (\sum_{i=1}^n (L_i \times W_L + H_i \times W_H + U_i \times W_U + S_i \times W_S + R_i \times W_R)) / n \tag{8}$$

K_I 表示机器人的第 I 个上肢动作的感性总量， n 表示参与感性评估的样本人数， i 表示在此感性评估子量上打分的人数， W 为不同感性评价因子对应的权值，针对上述 5 种感性评价因子，文献[7]提出了灵活—迟钝、正常—诡异、亲切—冷漠、熟悉—陌生、可爱—可憎 5 组形容词对来表征心理量，并通过词语相关度的调研，确定出权值分别为： $W_L=0.189$ ； $W_H=0.163$ ； $W_U=0.223$ ； $W_S=0.182$ ； $W_R=0.243$ 。

3.4 评估实验及数据统计

邀请 60 名志愿者，通过在线视频平台，分别单独在指导下观看录制的机器人上肢动作视频^[15]。每看

表 3 感性评价因子的 7 级语义度量评分 (T02)
Tab.3 7-level semantic measurement score of perceptual evaluation factors (T02)

动作名称	感性评价等级						
T02	-3	-2	-1	0	1	2	3
主动握手	无生命的					√	有生命的

表 4 6 种动作的感性评价因子值及心理总量
Tab.4 The perceptual evaluation factor value and psychological total of six movements

感性评价因子	T01 右手挥手	T02 主动握手	T03 举手示意	T04 惊讶动作	T05 轮流出拳	T06 双手抱头
有生命—无生命	1.10	0.38	1.54	0.59	0.15	-0.21
健康的—病态的	0.94	0.53	1.54	0.37	0.12	-0.24
易理解—难理解	1.38	0.35	1.71	-0.07	-0.35	-0.59
易模仿—难模仿	1.50	1.46	1.96	0.51	0.13	0.41
易共鸣—难共鸣	1.25	0.51	1.65	0.15	-0.41	-0.53
心理总量	1.28	0.63	1.66	0.23	-0.06	-0.29

完一个动作，按照 7 级语义量表法评估该动作，依据主观感受对感性评价因子进行打分。如 T02 主动握手动作的 L 评价量表，见表 3。

按照机器人复现动作的时间先后顺序，6 个动作的感性心理总量，以及对应的 5 种评价因子的计算结果，见表 4。

3.5 实验结果分析

通过表 4 可以看出，6 个动作按心理总量大小排序，排名靠前的为“举手示意”“右手挥手”“主动握手”“惊讶动作”，这些动作是人与人之间交流常用的动作。说明在生活中越被频繁的动作越容易辨识，呈现出较高的心理总量，即越远离恐怖谷的谷底位置。同时单手动作相较于双手动作，具有更高的心理总量。

“举手示意”“挥手动作”的心理总量均超过 1，具有较高的心理总量，而“双手抱头”“轮流出拳”则为负值，呈现出较低水平。针对这 4 项差异性较大的动作，进一步分析其评价因子值。“举手”和“挥手”动作的感性评价因子均高于“抱头”和“出拳”动作，在日常交流过程中，“举手”和“挥手”使用频率也远远大于“抱头”和“出拳”动作，并且在缺乏具体语境的情况下，“抱头”和“出拳”动作没有具体的含义，因此很难让人理解和产生共鸣，而“举手”和“挥手”作为基本人与人之间交流的方式，具有传递信息的功能，因此在“理解”和“共鸣”两项上具有较高分值，同时也表明机器人的上肢动作可以和人的肢体语言一样传递信息。

4 结语

本文基于仿人型双臂机器人 Player，利用 KinectV2 设备对真人动作进行了捕捉，求解臂形角，完成了人体运动数据到机器人数据的映射，采用中值

滤波算法对数据进行降噪，实现了仿人型服务机器人对真人示教动作的复现。对于复现后的机器人动作，基于恐怖谷理论，采用层次分析法和语义分析法评估了复现的动作。结果表明，在日常生活中使用频次越高的动作，服务机器人复现后的肢体动作的心理总量就越大；单臂动作相较于双臂动作具有更高的心理总量；另外具有特定含义的肢体动作本身具有易于被理解的特点，因此机器人复现后仍具有较高的心理总量，容易被人理解和产生共鸣。受限于示教动作的数量，本文并未进行多种影响因素的相关性分析，如动作的使用频次、动作为单手或双手，以及动作是否包含具体含义等，后续会在此方面展开研究。

参考文献：

[1] CARPENTER J, DAVIS J M, ERWIN-STEWART N, et al. Gender Representation and Humanoid Robots Designed for Domestic Use[J]. International Journal of Social Robotics, 2009, 1(3): 261-265.

[2] AVERTA G, SANTINA C D, VALENZA G, et al. Exploiting Upper-limb Functional Principal Components for Human-like Motion Generation of Anthropomorphic Robots[J]. Journal of Neuro Engineering and Rehabilitation, 2020, 17: 1-15.

[3] XU J, BROEKENS J, HINDRIKS K, et al. Mood Contagion of Robot Body Language in Human Robot Interaction[J]. Autonomous Agents and Multi-agent Systems, 2015, 29(6): 1216-1248.

[4] ZHANG R, HE Y, LIU Z T. Communication Atmosphere in Humans and Robots Interaction Based on Fuzzy Analytical Hierarchy Process[C]. Dalian: the 2017 36th Chinese Control Conference (CCC), 2017.

[5] 赫向阳. 七自由度仿人机械臂直接示教方法研究[D]. 武汉: 华中科技大学, 2017.

- HE Xiang-yang. Direct Teaching for Humanoid Manipulator of Seven Degrees of Freedom[D]. Wuhan: Huazhong University of Science and Technology, 2017.
- [6] LIU W. Comprehensive Evaluation Model of Service Robot[C]. Gothenburg: Proceedings of the International Conference on Identification, 2018.
- [7] 钟翔伟, 留沧海, 蒋刚. 仿人机器人交互行为设计评估方法[J]. 西南科技大学学报, 2017, 32(1): 78-84.
ZHONG Xiang-wei, LIU Cang-hai, JIANG Gang. Evaluation Method of the Humanoid Robot Interaction Design[J]. Journal of Southwest University of Science and Technology, 2017, 32(1): 78-84.
- [8] 张续冲, 张瑞秋, 陈亮, 等. 仿人机械双臂的运动学建模及实验[J]. 机械设计与研究, 2020, 36(6): 24-28.
ZHANG Xu-chong, ZHANG Rui-qiu, CHEN Liang, et al. Kinematical Modeling and Experimental Research of a Humanoid Dual-manipulator[J]. Machine Design and Research, 2020, 36(6): 24-28.
- [9] SHIMIZU M, KAKUYA H, YOON W K, et al. Analytical Inverse Kinematic Computation for 7-DOF Redundant Manipulators with Joint Limits and Its Application to Redundancy Resolution[J]. IEEE Transactions on Robotics, 2008, 24(5): 1131-1142.
- [10] Microsoft. Kinect for Windows SDK. (2014-10-21) [2021-01-12]. [https://docs.microsoft.com/zh-cn/previous-versions/windows/kinect/dn799271\(v=ie8.10\)](https://docs.microsoft.com/zh-cn/previous-versions/windows/kinect/dn799271(v=ie8.10))
- [11] GELDER B D. Why Bodies? Twelve Reasons for Including Bodily Expressions in Affective Neuroscience[J]. Philosophical Transactions of the Royal Society B: Biological Sciences, 2009, 364 (1535): 3475-3484.
- [12] MA G, HUANG Q, YU Z, et al. Experiments of a Human-Robot Social Interactive System with Whole-Body Movements[C]. Cham: Springer International Publishing, 2014.
- [13] MORI M, MACDORMAN K F, KAGEKI N. The Uncanny Valley[J]. IEEE Robotics & Automation Magazine, 2012, 19(2): 98-100.
- [14] SHIMADA M, MINATO T, ITAKURA S, et al. Uncanny Valley of Androids and Its Lateral Inhibition Hypothesis[C]. NW Washington: Proceedings of the IEEE International Symposium on Robot & Human Interactive Communication, 2007.
- [15] Bilibili. Experimental Evaluation of Limb Movements of Humanoid Service Robots[EB/OL]. (2021-02-22)[2021-02-28]. <https://b23.tv/1soK3I>.

(上接第19页)

- [26] KLEINBERG J. Bursty and Hierarchical Structure in Streams[J]. Data Mining and Knowledge Discovery, 2003, 7(4): 373-397.
- [27] CHEN C M, SONG M. Visualizing a Field of Research: a Methodology of Systematic Scientometric Reviews[J]. PloS One, 2019, 14(10): 25.
- [28] CHEN C M. CiteSpace II: Detecting and Visualizing Emerging Trends and Transient Patterns in Scientific Literature[J]. Journal of the American Society for Information Science and Technology, 2006, 57(3): 359-377.
- [29] 宋歌. 共被引分析方法迭代创新路径研究[J]. 情报学报, 2020, 39(1): 12-24.
- SONG Ge. Iterative Innovation Path of Co-citation Analysis Method[J]. Journal of the China Society for Scientific and Technical Information, 2020, 39(1): 12-24.
- [30] CHEN CHAOMEI. Science Mapping: A Systematic Review of the Literature[J]. Journal of Data and Information Science, 2017, 2(2): 1-40.
- [31] 陈新华, 蒋建文, 周华, 等. COVID-19 疫情背景下的医院人工智能快速布局和发展战略探讨[J]. 中国工程科学, 2020, 22(2): 130-137.
CHEN Xin-hua, JIANG Jian-wen, ZHOU Hua, et al. Rapid Layout and Development Strategy of Hospital Artificial Intelligence During the COVID-19 Pandemic[J]. Strategic Study of CAE, 2020, 22(2): 130-137.