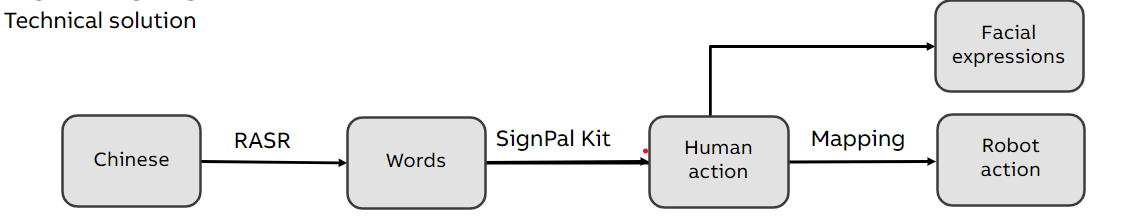
**课题名称：**机器人模拟人体上肢动作的实现原理研究

**课题来源：**生产实践

1. 课题背景及意义（课题的立题依据及研究意义）

随着科技的不断发展，人们对机器人的需求逐渐从简单的执行任务发展到更为复杂和自然的人机交互。当前，人机交互技术、康复机器人、生产自动化等领域对于机器人能够模仿人体上肢动作的需求逐渐增加。人们期望与机器人进行更自然、直观的交流。通过研究机器人模仿人体上肢动作，可以改进人机交互系统，使机器人能够更好地理解人类的意图并做出相应的响应，提高用户体验。

沟通是人与人之间，人与群体之间联系方式，该过程中思想与情感得到传递和反馈，人们通过沟通来达成思想统一和情感交流。全国残疾人调查情况数据显示：中国聋哑人数超过2000万。在日常生活中，聋哑人由于沟通不便，因此需要专业手语翻译员帮助，这可能导致隐私泄露等问题，为聋哑人生活带来困难。而手语是一种通过手部动作和手势进行沟通的语言，对于聋哑人士是一种主要的交流方式。为了促进聋哑人士的社交和信息交流，手语机器人成为一种潜在的有益工具。手语具有丰富的语法和词汇，而不同地区的手语可能存在差异。因此，手语机器人需要能够理解并模仿复杂的手势动作，以实现与聋哑人士的有效交流。通过手语机器人的研究，可以满足聋哑人士的社交需求，提高其在社会中的融入感和生活质量。手语机器人的研究可以使聋哑人士更便捷地与非手语使用者进行交流。机器人能够模仿手语，帮助非手语使用者理解和回应，从而提高交流效率。

本课题来源于上海ABB工程有限公司机器人模仿人体动作研究项目。

1. 语音识别手语机器人技术解决方案

该项目旨在完成一款语音识别手语机器人的设计。该机器人分为两个模块，一个是基于华为手语翻译工具的语音识别模块，其功能就是将输入的中文语音实时翻译转换成手语数据；另一个模块就是利用手语数据实现机器人运动定位映射。本课题将基于ABB公司的手语数据完成机器人模仿人体上肢动作的研究，实现机器人上肢运动的人机动作映射。

通过研究手语机器人，可以推动机器人技术在感知、识别和模仿上肢动作方面的发展。这对于机器人在其他领域的应用也具有启示作用，如人机交互、康复机器人等。我们将探究机器人上肢动作规划和控制算法，以实现模仿手语的自然、流畅的动作。通过这些研究方向，可以建立起一套完整的手语机器人系统，既能满足聋哑人士的交流需求，同时也为机器人模仿人体上肢动作提供了一个具体而有挑战性的研究平台。这有助于推动机器人技术在服务人类社会的过程中更深入、更广泛的应用。

1. 课题研究现状及发展趋势（课题研究领域的发展现状及可能的发展方向）

机器人模拟人体上肢动作的实现原理包括两个部分，一部分是人体动作的识别，另一部分是机器人手语动作的生成。因为本课题主要负责实现机器人上肢运动的人机动作映射，下面从手语机器人入手详细介绍这一方面的研究发展现状和未来发展方向。

1.研究现状

由于人和机器人之间存在着结构、尺寸、运动能力的差异，因此采集到的示教数据需要经过匹配映射后才能应用到机器人上。这个问题在学术上被称为人机动作映射，或者运动重定向。

运动重定向方法主要分为传统方法和深度学习方法。传统方法主要思想是找出源骨骼和目标骨骼之间的差异，建立两个骨骼之间的约束，然后进行约束求解。使用到的约束主要包括时空约束，运动学约束和动力学约束。

1)时空约束方法

时空约束方法是最先被提出的运动重定向方法。Gleicher首次提出运动重定向方法时候使用的就是时空约束方法。时空约束方法的主要原理在于对目标骨骼的末端效应器进行特定时间的位置和方向上的约束，以减少不必要和不期望的动作的出现。然后建立起一个约束求解器，把运动重定向问题变成了全局优化问题，优化时空约束求解器。基于时空约束方法的运动重定向按照时间片段来处理动作，该方法和其他方法相比，运算量大，速度较慢。另外，在时空约束下，可能无法实现一些动作，存在一定的缺陷。

2)运动学约束方法

为了解决时空约束的速度慢和部分动作缺陷问题，一些学者开始研究运动学约束方法，实时逐帧进行重定向。其中比较主要的方法就是逆向运动学方法。逆向运动学方法也是现在运动重定向最常用最主要的方法。逆向运动学首次由Le和Shin应用于运动重定向中。逆向运动学使用源骨骼和目标骨骼的骨骼长度和旋转角等运动学约束，根据目标模型末端效应器的位置，逆向计算出目标模型其他骨骼节点的位置。运动学约束的运动重定向方法生成的结果缺乏帧间约束，会产生明显的抖动。而且如果骨骼长度和比例之间差异较大，会损失一定的动作准确度。

3)动力学约束方法

为了使得运动重定向动作更加真实，很多研究开始把动力学约束加入了运动重定向过程中。不是所有的物理规则都适合加入运动重定向里面去，像动态约束等方法因为计算复杂，会出于算法的性能考虑被排除。动力学约束相比其他方法来说，最大的优点是能够自动仿真人体的真实动作，动作真实度高。但是运动学约束方法计算量大，而且在运动重定向上实际应用比较少，只用在了个别实验上，例如延伸模拟走路周期运动。

4)基于深度学习的运动重定向

随着深度学习技术的发展，运动重定向领域也开始用上深度学习技术。很多研究开始把深度学习应用到动作生成上来，而运动重定向也是一种动作生成问题，这些方法之间可以通用。使用深度学习进行运动重定向，使用无监督学习方法训练重定向网络，完成从源模型运动数据到目标模型运动数据的转变。

Villegas等提出了一种重定向网络结构，通过重定向网络可以自动的找到目标模型的关节旋转角。该重定向网络是无监督学习，不需要ground-truth信息，实现了基于关节坐标表示的源模型运动数据到不同骨骼拓扑结构的目标模型之间的运动重定向。周阳等人2列在Villegas的基础上提出了一个通用双向循环自编码器，该方法用到了运动生成技术，在仅拥有1种骨骼运动数据下也可以进行训练。Aberman等人提出了一个基于骨架的无监督学习的运动重定向网络，动作分解为位置和旋转两部分分别进行训练。

深度学习方法相对于传统的基于约束的方法来说，不需要太过去注重约束的合理性，不需要对动力学，运动学有很深的了解，通过训练网络就能得到一个较好的结果。

2.未来研究发展趋势：

越来越多的研究者已经开始探索使用深度学习模型，如卷积神经网络（CNN）和循环神经网络（RNN），对运动重定向进行建模和学习。未来的研究趋势之一是引入深度强化学习，以使系统能够通过与环境的交互不断优化运动重定向策略。随着深度学习在计算机视觉和自然语言处理等领域的成功，未来研究将更加注重多模态信息的融合，包括视觉、声音、生理信号等，以提高运动重定向的性能。

1. 研究内容及研究目标（对研究的内容进行说明，并阐明要达到的目标）

本项目依托于浙江大学的Kinematic Motion Retargeting via Neural Latent Optimization for Learning Sign Language项目研究成果。该项目名为神经潜优化的运动学运动重定向学习手语，主要针对现有方法操作繁琐、效率低下的问题，提出了基于编码解码器的手语动作生成方法。

该方法仅仅需要一个人类示教的手语单词动作轨迹数据库，生成手语语句时根据输入文本从数据库中查询出相应的单词轨迹，然后使用过渡动作生成网络和人机动作映射网络动态地对这些轨迹进行拼接和映射，自动生成机器人的手语语句动作轨迹。为了学习人机之间复杂的非线性映射关系，文中实现了基于图编码解码器的人机动作映射网络。为了估计手语单词和过渡动作切换的时刻并生成单词之间的过渡轨迹，文中设计并实现了基于时序编码解码器的过渡动作生成网络。考虑到生成的手语动作轨迹需要满足人机相似性、可执行性等要求，文中设计了多个损失函数项对网络进行联合训练。

而本项目的任务就是复现浙江大学论文研究成果，在他们的基于图卷积网络（GCN）的图编码解码器人机动作映射神经网络模型架构基础上，利用该模型训练我们自己的机器人，并最终实现从人体肢体语言到机器人具体动作的映射，要求在仿真环境中实现全流程的语音转手语的过程。

目的要求：研究基于深度学习的人体到机器人的姿态映射方法。

技术要求：在仿真环境中实现机器人对人体姿态的模仿，解决模型碰撞干涉问题。

1. 预计的研究难点（课题研究过程中可能遇到的理论难题或技术难点）

我们的任务是对看不见的手语任务执行动作重定向，其中包括复杂的双臂运动和手指运动。然而，由于人类和机器人之间的差异，这仍然是一个持续的挑战。即使是结构与人类相似的类人机械臂，在自由度、运动学参数和物理约束等方面也存在差异。再加上对相似性、安全性、快速性的各种要求，使得问题变得难以解决。

具体的理论难题或技术难点如下：

1.输入数据处理问题。浙江大学的数据通过动作捕捉获得，而我们的数据通过华为手语翻译工具获得。两种数据的类型不同，在将我们的数据集导入到图编码解码器人机动作映射神经网络模型训练之前要对数据进行处理。

2.模型适用问题。浙江大学采用的ABB的YuMi双臂机器人包含14个自由度，并配备了Insipire- Robotics公司的灵巧双手，每只手有6个自由度和12个关节。而我们采用的是YuMi双臂机器人搭配Schunk公司的SVH rechte Hand五指伺服电动机械抓手。两个机器手在自由度、运动学参数和物理约束等方面与其并不相同，包括初始偏移量、初始旋转、关节角度的上限和下限等。浙江大学的机械手模型与我们的不同，我们需要对相关参数和手部模型进行修改，重新导入神经网络中进行训练，才能使机器人在仿真环境中做出正确的手语动作。

3.碰撞干涉和超参数设置问题。由于机器人模型的不同，在进行神经网络深度学习过程中，我们需要的损失函数和超参数肯定和浙大的不同，如果直接采用浙江大学的损失函数和超参数一定会导致我们机器人在运动时手部出现碰撞干涉。所以在后续的代码编写和模型训练过程中应当调整损失函数和超参数，根据TensorBoard显示的训练效果及时修改。

4.模型优化和泛用。解决完前两个问题后，我们还可以将模型性能优化，使其能够适应更广泛具体的实际运用场景。比如浙大采用的是基于编码解码器的神经网络模型，如果有时间精力，我们可以尝试其他神经网络的组合，比如CNN和RNN等，不同神经网络的训练效果不同。

1. 创新点（选题、观点、理论、材料、方法等创新点）

创新点：

选题上：我们从手语机器人出发，没有选择传统的约束求解方法实现机器人的动作映射，而是选择了最前沿的浙江大学的基于编码器解码器的深度学习网络，并将其运用到我们的实际机器人项目中。

材料上：我们选择采用ABB公司的YuMi双臂机器人加Schunk公司的SVH rechte Hand五指伺服电动机械抓手的组合。我们选择的Schunk公司的机器手自由度等方面的参数均领先于浙大选择的Insipire- Robotics公司的灵巧双手，比他们的机器手更复杂更接近实际运用的需求。

模型数据上：浙江大学的数据集通过动作捕捉获得，我们的数据集通过华为手语翻译工具获得，数据获取更容易，数据量也更大。

实际运用：我们的机器人最终实现的功能是直接从中文语音到机器人手语动作的转换，而浙江大学等以往的研究是从人的手语动作到机器人的手语动作的转换。显然，我们的机器人更具有现实意义，更符合实际生产生活的需要，可以直接运用到相关场景中，比如说取代人工的手语翻译工作。将基于图编码解码器训练的模型运用到实际机器人上，使该方法训练的模型的实用性得到验证，为后续进一步推广应用该方法提供现实案例。

1. 进度计划（根据研究内容及研究目标所预计的进度安排）

1.1-1.5 完成任务书

1.2-1.12 完成开题报告，进行开题报告答辩

1.2-1.21 查阅文献，完成调研综述，学习相关知识

1.21-2.4 依据现有数据集，完成模型训练

2.4-3.1 在模型中加入碰撞检测模块，实现无碰撞训练

3.1-4.8 将模型性能优化，使其能够适应更广泛具体的实际运用场景

4.8-4.12 完成中期报告和中期检查表

4.12-5.18 完成全文初稿

5.18-5.25 论文第一次查重和第二次查重

5.25-5.31 完成论文修改、答辩

1. 参考文献

[1]斋藤康毅. 深度学习入门:基于Python的理论与实现[M].北京:人民邮电出版社,2018.7

[2]弗朗索瓦·肖莱. Python深度学习[M].北京：人民邮电出版社，2018.8

[3] Richard S.Sutton and Andrew G.Barto. Reinforcement Learning: An Introduction second edition [M].The MIT Press .Cambridge,Massachusetts London,England

[4] Kinematic Motion Retargeting via Neural Latent Optimization for Learning Sign Language. [J].Haodong Zhang; Weijie Li; Jiangpin Liu; Zexi Chen; Yuxiang Cui; Yue Wang; Rong Xiong.IEEE Robotics and Automation Letters, vol. 7, no. 2, pp. 4582-4589, April 2022

[5]李威杰.人机手语交互中的动作识别与生成[D].浙江大学,2022.DOI:10.27461/d.cnki.

gzjdx.2022.001848

[6]倪恋. 基于特征共享与时序建模的连续手语翻译研究[D].武汉理工大学,2022.DOI:10.

27381/d.cnki.gwlgu.2022.000214.

[7]罗文杰. 基于深度学习的手语识别研究与应用[D].南昌大学,2023.DOI:10.27232/d.cnki.

gnchu.2023.003074.

[8]高辉. 基于深度学习的手语识别系统设计[D].哈尔滨理工大学,2023.DOI:10.27063/d.

cnki.ghlgu.2023.000303.

[9]黄仰来. 基于深度学习的动态中文手语识别研究[D].东北电力大学,2023.DOI:10.27008/

d.cnki.gdbdc.2023.000381.

[10]米娜瓦尔·阿不拉,阿里甫·库尔班,解启娜等.手语识别方法与技术综述[J].计算机工程与应用,2021,57(18):1-12.

[11]韩祥祖. 基于深度学习的手语识别算法研究及应用[D].山东大学,2023.DOI:10.27272/d.

cnki.gshdu.2022.003190.

[12]李少波. 机器人的人体姿态动作识别与模仿算法[D].上海交通大学,2013.

[13]刘迪. 基于人体动作捕获的机器人姿态控制系统研究[D].吉林大学,2019.

[14]简继盼. 基于结构文法的机器人模仿学习方法[D].大连理工大学,2022.DOI:10.26991/d.

cnki.gdllu.2021.002029.

[15]谢梦依. 基于深度学习的非特定人手语识别[D].山东大学,2021.DOI:10.27272/d.cnki.

gshdu.2020.004692.

[16]Static Hand Gesture Recognition Based on Convolutional Neural Networks[J]. Pinto Raimundo F.;Borges Carlos D. B.;Almeida Antônio M. A.;Paula Iális C..Journal of Electrical and Computer Engineering,2019

[17]A Generative Human-Robot Motion Retargeting Approach Using a Single RGBD Sensor.[J]. Sen Wang;Xinxin Zuo;Runxiao Wang;Ruigang Yang.IEEE Access,2019

[18]Motion Retargeting for Humanoid Robots Based on Simultaneous Morphing Parameter Identification and Motion Optimization[J]. Ayusawa Ko;Yoshida Eiichi.IEEE Transactions on Robotics,2017

[19]Trajectory generation with multi-stage cost functions learned from demonstrations[J]. Hu Jin;Xiong Rong.Robotics and Autonomous Systems,2019

[20]Sign Language Recognition with the Kinect Sensor Based on Conditional Random Fields[J]. Hee-Deok Yang.Sensors,2014

[21]Trajectory Learning for Robot Programming by Demonstration Using Hidden Markov Model and Dynamic Time Warping.[J]. Aleksandar Vakanski;Iraj Mantegh;Andrew Irish;Farrokh Janabi-Sharifi.IEEE Trans. Systems, Man, and Cybernetics, Part B,2012

[22]冯健豪. 基于深度学习运动重定向的模型驱动系统的设计与实现[D].北京邮电大学,2022.DOI:10.26969/d.cnki.gbydu.2022.002192.

[23]葛春宝,陈益强,尹宝才等.一种新的手势运动数据重定向方法[J].计算机学报,2006(10):1850-1855.

[24]李琳,郑利平,王玉培等.基于关键姿态映射的运动重定向[J].系统仿真学报,2012,24(01):127-131.DOI:10.16182/j.cnki.joss.2012.01.036.

[25]金梅,张立国,金菊等.基于物理约束的运动重定向[J].高技术通讯,2014,24(10):1062-1068.

[26]王森. 基于RGB-D信息的人-机器人运动重定向视觉关键技术研究[D].西北工业大学,2022.DOI:10.27406/d.cnki.gxbgu.2019.000191.

[27]周阳,李书杰,朱海生等.面向关节坐标运动数据重定向的通用双向循环自编码器[J].计算机辅助设计与图形学学报,2020,32(02):315-324+333.

[28]周阳. 面向关节坐标运动数据的运动重定向方法研究[D].合肥工业大学,2021.DOI:10.27101/d.cnki.ghfgu.2020.000872.

[29]柯文德,彭志平,武春英.基于运动相似性的仿人机器人动作规划研究综述[J].广东石油化工学院学报,2014,24(03):33-37.