

信息与控制  
*Information and Control*  
ISSN 1002-0411, CN 21-1138/TP

## 《信息与控制》网络首发论文

题目: 认知-运动康复医疗机器人应用设计  
作者: 张宇, 刁亚楠, 梁升云, 叶超翔, 周艳霞, 赵国如  
网络首发日期: 2021-06-17  
引用格式: 张宇, 刁亚楠, 梁升云, 叶超翔, 周艳霞, 赵国如. 认知-运动康复医疗机器人应用设计. 信息与控制.  
<https://kns.cnki.net/kcms/detail/21.1138.TP.20210617.1612.019.html>



**网络首发:** 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式 (包括网络呈现版式) 排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

**出版确认:** 纸质期刊编辑部通过与《中国学术期刊 (光盘版)》电子杂志社有限公司签约, 在《中国学术期刊 (网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊 (网络版)》是国家新闻出版广电总局批准的网络连续型出版物 (ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

# 认知-运动康复医疗机器人应用设计

张宇<sup>1,2</sup>, 刁亚楠<sup>1</sup>, 梁升云<sup>1,2</sup>, 叶超翔<sup>1</sup>, 周艳霞<sup>3</sup>, 赵国如<sup>1,2</sup>

(1.中国科学院深圳先进技术研究院, 广东 深圳 518000; 2.中国科学院大学, 北京 100190;  
3.深圳市第二人民医院, 广东 深圳 518035)

**摘要：**脑神经损伤及认知衰退人群存在运动能力下降、言语功能退化、跌倒风险增大等问题，目前的康复机器人主要关注于被动训练，运动康复效果不佳。提出研究一款新型认知运动康复机器人，开发其“感知-认知-运动”多模态智能感知交互功能。采用 Pepper 机器人搭建了一个集认知、运动康复训练为一体的康复机器人系统：首先，设计了多模态的智能感知及交互系统，实现了与患者的视听触觉交互和言语训练；其次，利用机器人声呐、激光和视觉实现实时地图构建及人体追踪；最后，将深度学习方法应用到机器人上，实现了基于机器人视觉的人体姿态识别，为下一步通过视频图像判别人体稳定性提供了可行性方法。实验结果显示设计的康复机器人系统在认知运动协同康复训练过程中具有正作用，该研究初步证明了智能机器人在医疗辅助康复训练上的可行性，被期望未来用于居家及临床上的认知运动协同康复。

**关键词：**医疗机器人；认知-运动康复训练；人机交互；目标跟踪；姿态估计

**中图分类号：**TP242

**文献标识码：**A

## Cognitive-motion Rehabilitation Medical Robot Application Design

ZHANG Yu<sup>1,2</sup>, DIAO Yanan<sup>1</sup>, LIANG Shengyun<sup>1,2</sup>, YE Chaoxiang<sup>1</sup>, ZHOU Yanxia<sup>3</sup>,  
ZHAO Guoru<sup>1,2</sup>

(1. Shenzhen Institutes of Advanced Technology Chinese Academy of Sciences, Shenzhen 518000, China;  
2. University of Chinese Academy of Sciences, Beijing 100190, China;  
3. Shenzhen Second People's Hospital, Shenzhen 518035, China)

**Abstract:** People with brain nerve injury and cognitive decline have problems such as decreased exercise ability, degraded speech function, and increased risk of falling. The current functional rehabilitation robots mainly focus on passive training, and the effect of sports rehabilitation is not good. Proposed to study a new type of cognitive sports rehabilitation robot and develop its "perception-cognition-motion" multi-modal intelligent perception interaction function. The Pepper robot is used to build a rehabilitation robot system that integrates cognitive and sports rehabilitation training: first, a multi-modal intelligent perception and interaction system is designed to realize the audio-visual tactile interaction and speech training with the patient; secondly, the robot is used Sonar, laser and vision realize real-time map construction and human tracking; finally, the deep learning method is applied to the robot to realize the human posture recognition based on robot vision, which provides a feasible method for judging the stability of the human body through video images in the next step. The experimental results show that the designed rehabilitation robot system has a positive effect in the process of cognitive sports collaborative rehabilitation training. The research initially proves the feasibility of intelligent robots in medical-assisted rehabilitation training, and is expected to be used for home and clinical cognition in the future Exercise synergistic rehabilitation.

**Keywords:** Medical robot; Cognitive-sports rehabilitation training; Human-computer interaction; Target tracking; Pose estimation

## 1 引言 (Introduction)

随着人工智能、深度学习等信息技术的崛起，机器人技术也得到了迅速发展。目前主流的服务机器人主要有家用服务机器人、医疗服务机器人、公共服务机器人，其中医疗服务机器人技术门槛高，是医学、工学和机器人学的结合，市面上比较少见<sup>[1]</sup>。医疗服务机器人能够针对性的进行医疗服务，如导医、康复指导、监控、陪护、娱乐等，能够进行多模态的人机交互，处理采集到的各类数据，在技术上给医生和患者提供相应的服务。如今老龄化进程加剧，人们对于健康的需求不断增长，临床上多数运动障碍患者康复

过程中需要有人指导，专业的医疗康复设备有限，医疗服务机器人的需求巨大，主流的研究多为穿戴式或接触式康复机器人，典型系统如 MIT 的上肢康复机器人系统 (MIT-MANUS 系列)<sup>[2-4]</sup>，Hocoma AG 公司研发的最早用于临床的悬吊减重 (Partial Weight System, PWS) 步态训练的康复机器人 Lokomat<sup>[5]</sup>，美国加州大学伯克利分校的 eLEGS 助力机器人系统<sup>[6]</sup>和德国的本杰明·富兰克林医科大学自动化与机器人研究室提出的 Haptic Walker 的步态康复机器人<sup>[7]</sup>。而可提供指导和激励的非接触式机器人——医疗辅助康复机器人研究较少，或将成为未来功能康复机器人的重要

发展方向。本文主要面向脑神经损伤、认知-运动障碍人群，探索智能人形机器人在医疗辅助康复训练方向的应用。

为满足需求，人形机器人选择了高度开源的 Pepper<sup>[8]</sup>和 Nao 类人机器人，很多学者对于二者做了不同的研究：（1）类人机器人目标检测方面，主流的是运用传统的机器学习算法<sup>[9-10]</sup>，而该类算法多是应用到数据规模小的情况，遵循一定的标准解决问题。针对传统机器学习不适用的情况，近些年也有一些好的深度学习算法（如 YOLO）慢慢的出现在了机器人目标检测中<sup>[11-12]</sup>，这些算法都提高了目标检测准确率和速度；D.Coquin 等人<sup>[13]</sup>则是融合了外部摄像头和 Nao 机器人摄像头，同样提高了目标识别率。（2）类人机器人自主路径导航有“边线式”<sup>[14]</sup>机器人迷宫导航方法和基于语音的自主栅格地图创建方法<sup>[15]</sup>，刘成菊等<sup>[16]</sup>采用改进的快速扩展随机树（RRT）算法设计了一种能够适应机器人足球赛场动态移动障碍环境的路径规划器，并将策略移植到实体 NAO 机器人上。研究机器人的路径规划和导航技术可以广泛应用到服务类机器人中，为机器人提供更广泛的发展空间。（3）类人机器人模仿学习方面，Hsieh<sup>[17]</sup>、Avalos 等<sup>[18]</sup>提出了结合 Kinect 相机输入，使 Pepper/Nao 机器人模仿人的动作的方法，Hirschmanner 等<sup>[19]</sup>则是教 Pepper 机器人学习单词和动作的映射关系，使之能够识别单词对应的动作。（4）在医疗应用方面，徐健等<sup>[20]</sup>将 Nao 机器人应用于和患者、医护的交互，文<sup>[21]</sup>则将 Nao 应用于自闭症儿童的治疗干预中，文<sup>[22]</sup>用 Nao 机器人给手部骨折患者做康复演示；Rozanska 等<sup>[23]</sup>介绍了一种多模态情绪分析的方法，用于患者和 Pepper 机器人各种情绪下的互动，而 Ilyas 等<sup>[24]</sup>则利用深度神经网络学习识别脑外伤患者的情绪。（5）系统设计方面，Gardecki 等<sup>[25]</sup>利用 Pepper 打造了一个机器人前台接待系统，是一个 Pepper 机器人系统开发示例。

Pepper 机器人和 Nao 机器人均为日本 softbank 公司研发的类人机器人，二者具有相同的操作系统。外形方面，Pepper 机器人高 1.21m，下肢运动机构为 360° 旋转的轮子，运动灵活；Nao 机器人高 0.57m，下肢为双腿双足，运动能力较差。由于设计方案要求机器人能够跟踪人体，对机器人的灵活性要求较高，因此最后决定使用 Pepper 机器人作为此次的研究目标。

本文探索 Pepper 机器人系统的设计方案，将其应用到脑神经损伤、认知-运动功能衰退人群的康复治疗。针对该类人群存在的认知障碍、言语功能退化、运动能力下降、跌倒的风险增大的情况，扩展机器人情感交互语料库，锻炼语言表达互动能力；设置自主生活模式时钟，提醒用药及锻炼指导；智能人体追踪，实时陪护、监控及人体姿态估计，最终把 Pepper 打造成一款适用于上述人群的医疗服务机器人，以上功能均在 Pepper 机器人上实现，并在医院临床试验应用。

## 2 实验方案设计（Experimental design）

### 2.1 患者筛选及确定

在深圳市第二人民医院招募了十名患者，男性 5 人，女性 5 人，平均年龄 61.7 岁，均为专业医生诊断的认知障碍患者，纳入标准如下：（1）具有认知障碍的中老年人群，专业医师的诊断认知障碍患者，伴随着言语能力的下降；（2）年龄在 80 岁以下；（3）运动能力存在缺陷，行走缓慢且存在失稳风险；（4）仍然具备行走锻炼的能力，能够进行正常的言语交流。（5）身体健全，不穿戴假肢，能够进行自主活动，适当的可以使用辅助装置（如拐杖）；（6）下肢未经历过重大手术（如膝关节置换、截肢等）；（7）没有神经性疾病史，无服用神经性药物；（8）所有招募的患者能够自主了解实验的过程，并对实验做出自我评判，在其确认可行性的基础上方能参加实验，同时全程需取得家属的同意并签署知情同意书。

以上患者均为认知-运动能力衰退的中老年人群，相较于正常中老年人存在更高的跌倒风险。经过筛选确认了符合以上要求的 10 名患者，所有患者的简易智力检查量表 MMSE 和蒙特利尔认知评估量表 MoCA-B 得分如下表 1 所示。其中 MMSE 总分为 30 分，分数在 27-30 分为正常，分数<27 为认知功能障碍；MoCA-B 总分 30 分，分数<26 为认知功能障碍。患者招募及相关实验，均通过本单位伦理委员会审查。

表 1 患者基本信息表

Table 1 Patient general information						
患者	性别	年龄	身高 /cm	体重 /kg	MMSE 得分	MoCA-B 得分
1	女	65	154	63	26	20
2	女	60	158	51	23	19
3	女	63	162	55	20	18
4	女	57	150	44	27	17
5	女	64	150	54	21	16
6	男	79	167	49	25	23
7	男	51	175	60	23	16
8	男	70	167	60	13	8
9	男	49	172	70	28	18
10	男	59	160	50	26	20

### 2.2 实验环境及实验设置

实验环境及示意图如下图 1 所示，地点为深圳市第二人民医院神经内科康复室，场地大小为 2.4m×8m，中间有两个障碍物，相距约 3.5m。

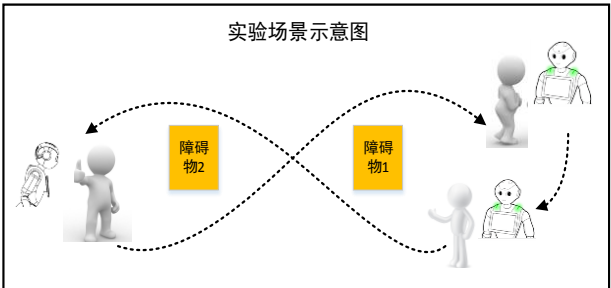


图1 实验场景

Fig.1 Experimental scene

实验开始前,由 Pepper 机器人对场地进行扫描定位,绘制出场景地图,之后按照以下步骤进行实验:

- (1) 由 Pepper 主动发起和患者进行语言交流;
- (2) 患者通过语音或者触摸给 Pepper 下达诸如人脸识别、摆拍、跳舞等指令,机器人根据指令做出相应的反应;
- (3) Pepper 在设定的时间内可以自主唤醒并提醒患者服药、锻炼等预设活动,该功能持续整个实验过程;
- (4) 在达到运动提醒设定时间以后,Pepper 提醒患者做康复运动锻炼,患者将沿着图1中的路线进行走路锻炼;
- (5) Pepper 对人体进行实时的跟踪,精确躲避障碍物,并对该行动过程做视频采集。
- (6) Pepper 将二维摄像头采集实时的视频图像数据发送到电脑端,在电脑端用 GPU 上的训练好的模型对视频图像中的人体进行关键点预测,实时估计出

人体主要关节的位置,并对运动状态做出评估。

实验结束后 Pepper 恢复自主生活模式,可随时被患者唤醒,执行相应的指令。

### 3 认知-运动康复医疗机器人的基本控制结构 (A general control architecture of the cognitive-motion rehabilitation medical robot)

认知-运动康复医疗机器人的基本控制结构框架如下图2所示。其最终能实现:(1)自主生活模式下,Pepper 机器人可以代替人与患者进行语音交流锻炼,智能提醒用药、康复锻炼、生活娱乐;(2)在既定的环境中可以进行地图的扫描构建,之后能够在患者进行行走的过程中对人体进行实时的跟踪,语音提醒,陪护锻炼;(3)在日常活动中,Pepper 还能通过头部的摄像头进行实时的视频图像信息采集,通过 TCP/IP 协议将数据传输到电脑中,在对缓存的视频图像数据进行处理,识别出患者的关节点的位置,进一步能够判断出患者身体的稳定性。

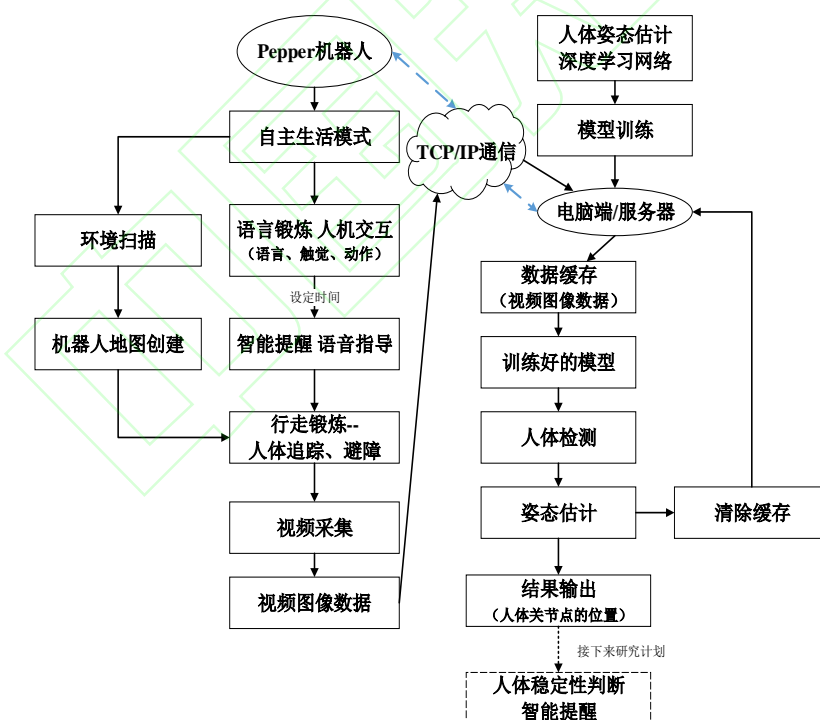


图2 认知-运动康复医疗机器人的控制结构

Fig.2 The control architecture for cognitive-motion rehabilitation medical robot

#### 3.1 认知-运动康复医疗机器人交互设计

语音交互是人机交互最重要的内容之一,机器人对话系统的基本原理如下图3所示,完成一次机器人与人的沟通需要机器人系统进行自然语言理解、对话状态追踪、策略学习以及自然语言生成等步骤,最后

给出恰当的回答。Pepper 机器人采用的是检索式对话系统,由于此次面向的是认知运动障碍人群,对话内容以任务型为主,需要构建有针对性的语料库即对话样本集,其中关键词检索是该语音交互的重点。本文采用 Pepper 机器人的 Dialog 模块根据 QiChat 语法规



则，为其编写语音交互语料库，包括不同的语速、语气、响度。此外还可以调用外部的语音交互系统<sup>[21]</sup>，由于面向对象的独特性这里并未使用。语料库内容基于医生和患者家属的建议，充分考虑到康复指导的科学性、语料内容的普适性，以及该类人群的关注点，构建了科学的情感语料库及语音指令集。

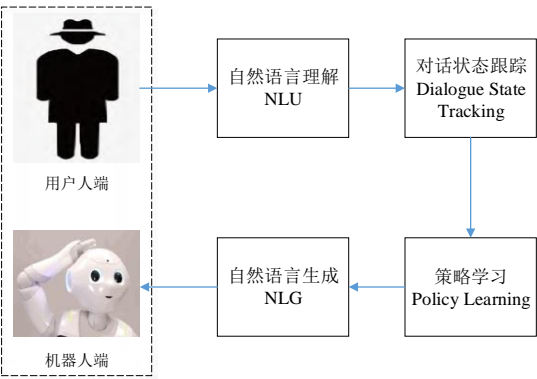


图3 机器人对话系统基本原理图

Fig.3 Basic schematic diagram of robot dialogue system

下图4所示为交互设计的实例，整个交互设计充分利用 Pepper 的听觉、触觉感知系统，在语音指令或者触摸头部和手部的时候可以执行/停止一些设定好的行为（如唱歌、人脸识别等）。此外设定不同的触发条件，可以控制 Pepper，包括触摸指令使用 Touch 模块中手部和头部的触觉传感模块，在该模块接收相应的刺激之后可以执行相应的操作，头部的触摸设置为停止命令，手部的触摸为开始指令，过程中通过触摸机器人相应位置可以开始/停止当前行为。而人脸识别使用 Face Reco 模块，首先需要 Pepper 对相关人脸信息进行学习、记忆，然后 Pepper 在识别人脸的过程中与已学习人脸数据库中的数据进行比对，最终实现对患者的准确识别。

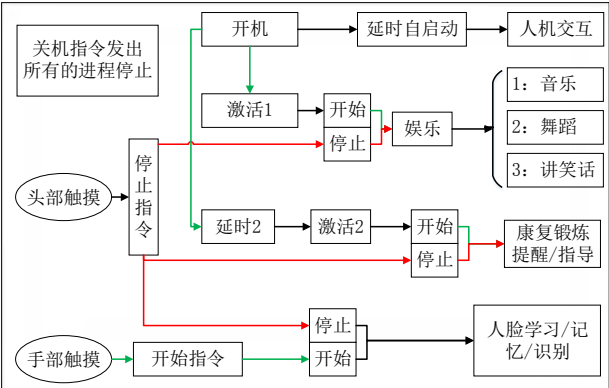


图4 认知运动-康复医疗机器人的交互设计

Fig.4 Interactive design of cognitive-motion rehabilitation robot

3.2 实时地图构建及人体追踪设计

机器人实时地图扫描的基本原理是地图栅格化

和坐标系化，算法的基本步骤为：（1）选定初始探测位置后建立坐标系；（2）探测，利用 Pepper 机器人的激光、声呐传感器对设定距离内的环境进行探测，判断出有障碍物的栅格，无障碍物的栅格（包括可移动的栅格）；（3）根据短路径在可移动栅格内移动设置距离；（4）重复上述步骤直至坐标系下均探测完成；（5）进行地图绘制。整个地图构建的算法流程图如下图5所示。

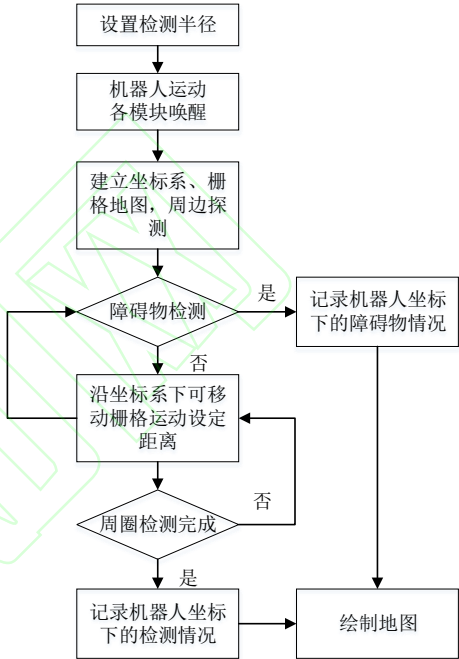


图5 Pepper 机器人绘制地图算法流程图

Fig.5 Pepper robot mapping algorithm flow chart

绘制出的实验环境地图如图6所示，黑色为障碍区域，白色为无障碍区域。由图可见图像四周为黑色障碍区域，与设定的环境区域大小、形状一致，中间的两块黑色区域为设置的障碍物，也与实验设置相符合。地图绘制后，在机器人坐标内各障碍物位置、大小明晰，可在此基础上实现机器人的精准避障，能够更好地完成人体追踪。



图6 Pepper 机器人绘制的地图

Fig.6 Maps drawn by pepper robots

基于动态摄像头的人体追踪比较复杂，如何快速的检测目标并对其进行追踪是该研究的重点。Pepper具有两个二维摄像头，可以 30 帧每秒提供  $640 \times 480$  分辨率的图像数据，如图 7 所示 Pepper 摄像头视场范围有限，容易丢失目标，针对这个问题采用正面人脸检测及背面红球检测追踪的方法，保证不同角度人体在 Pepper 视野中有目标可循。人体追踪的算法原理如下：（1）目标初始化，对运动目标进行检测；（2）特征提取和目标检测；（3）判断目标模型，这里只用到了人脸及红球的模型；（4）目标搜索及识别，在每一帧的图像中获取目标候选样本，该步骤影响跟踪效率；（5）模型更新，这步可以用来判断机器人面对的是人的正面或者背面；（6）机器人跟踪运动。

机器人在确定目标以后再进行位置及距离的计算，以便在安全合理范围内进行人体跟踪。Pepper 机器人的双目测距原理如图 8 所示，其中： $o_1$  和  $o_2$  分别为机器人上下摄像头的光心， $f$  为摄像头的焦距， $T$  为光心间距， $P$  为空间中一点， $x_1$  和  $x_2$  分别为  $P$  点在上下成像平面  $c_x^1$ 、 $c_x^2$  上的点；视差  $d = x_1 - x_2$ ，距离  $Z$  为最终所求距离，由三角形相似原理可得公式（1）：

$$Z = \frac{T * f}{T - d} \quad (1)$$

最后 Pepper 确定前进的方向以及距离，实现对人体的追踪，配合运动障碍人群走路锻炼。

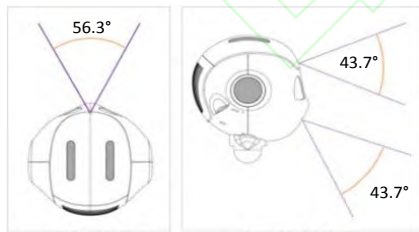


图 7 Pepper 机器人的视野范围  
Fig.7 The pepper robot's visual field

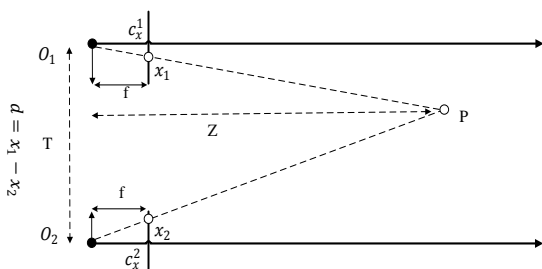


图 8 双目测距原理

Fig.8 Principle of binocular ranging

### 3.3 人体姿态估计在 Pepper 机器人上的应用设计

人体姿态估计是指在给定的图像或视频数据中恢复人体主要关节的过程。根据人体主要关节可以计算人体重心及稳定域的位置，进而能够判断人体的稳定性<sup>[27]</sup>，是通过视频图像判断人体跌倒的可行方法，这是医疗服务机器人未来发展的一个热点研究。

传统的人体姿态估计一般是基于图结构和模型，设计 2D 人体部位检测器，使用图模型建立各部件的连通性，并结合人体运动学的相关约束不断优化图结构模型来估计人体姿态，效果不好。而基于深度学习的人体姿态估计方法主要是利用卷积神经网络从图像中提取人体姿态特征，可以得到语义信息更为丰富的特征、感受野下多尺度多类型的人体关节特征向量和每个特征的全部上下文关系，对这些特征向量进行坐标回归以反映当前姿态，近些年来基于深度学习的人体姿态估计的准确率和速度已大大提高。

本文使用的是深度学习方法，其网络结构复杂，需要大量的训练样本，运算性能要求比较高，在 Pepper 机器人上无法运行。如下图 9 所示，为实现 Pepper 机器人实时人体姿态估计，在 Pepper 追踪人体的过程中将采集到的视频图像数据将通过 TCP/IP 协议传输到电脑端，将视频图像数据进行缓存，然后再将缓存的视频图像数据输入已经训练好的人体姿态估计深度学习网络中，经过人体姿态估计得到人体的主要关节的位置，进一步能够判断人体稳定性。

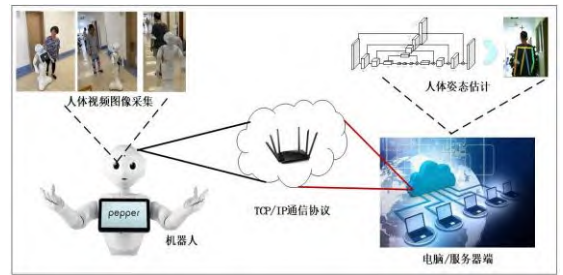


图 9 人体姿态估计在 Pepper 机器人上的应用

Fig.9 Application of Human posture estimation in Pepper robot

本文使用的人体姿态估计深度学习网络主要借鉴上海交通大学卢策吾老师团队做的 Alphapose<sup>[28]</sup>，采用自上而下（top-down）的检测方法，网络结构如下图 10 所示，训练所用数据集为 COCO 关键点数据集。

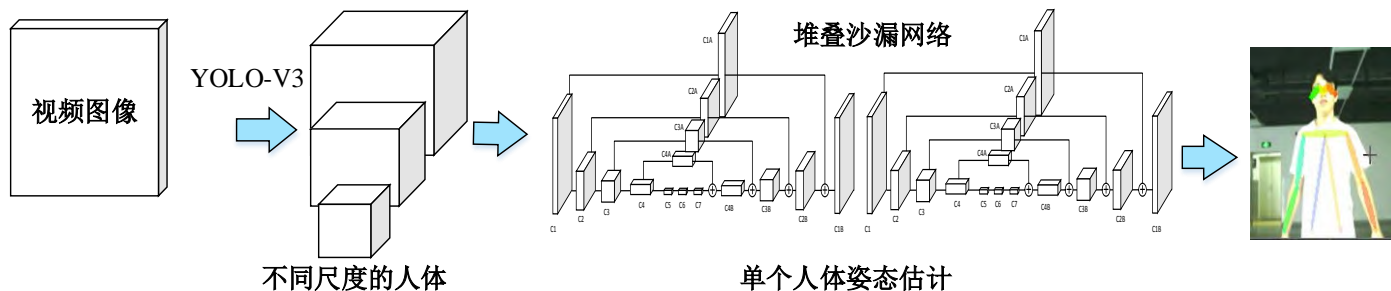


图 10 应用于 Pepper 机器人的人体姿态估计的网络结构

Fig.10 Network structure of human posture estimation for pepper robot

如图可见，在视频图像信息输入到网络以后，首先用 YOLO-V3<sup>[29]</sup> 网络进行快速的人体检测，YOLO-V3 从三个不同的尺度进行目标采样预测，能够很好地检测视频图像数据中的大、中、小尺度的人体目标，特别是对小目标的检测有着很好的效果，在此基础上再对检测到的每个人体进行单人姿态估计，这里单人姿态估计使用的是 2 层堆叠沙漏网络<sup>[30]</sup>模型，沙漏网络最主要的特点就是能够保留多层的卷积特征，利用多尺度特征来识别姿态。该方法首先使用单个沙漏网络，如图 11 所示，将特征图层层叠加，最终生成一个拥有所有层的信息，与原图大小相同的特征图 C1B，并通过  $1 \times 1$  卷积生成代表关键点概率的热力图。由于关节点之间是可以相互参考预测的，并且热力图包含了所有关节点间的相互关系，于是将第一个沙漏网络输出的热力图作为下一个沙漏网络的输入，从而在第二个沙漏网络中使用关节点间的相互关系来提升单个关节点的预测精度。通过以上方法预测关节在每个像素处存在的概率，概率最大的像素处即最终得到人体关节点的位置。

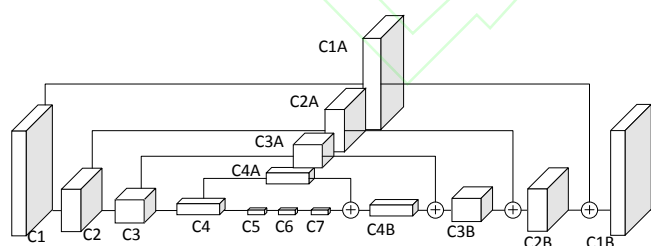


图 11 单个沙漏网络示意图

Fig.11 An illustration of a single “hourglass” module

## 4 认知-运动康复医疗机器人的临床试用结果 (Clinical trial results of cognitive-motion rehabilitation medical robots)

### 4.1 Pepper 的多模态交互结果

根据临床诊断及病历病史可见，深圳市第二人民医院十名患者皆为认知-运动障碍的患者，语言功能都

有所退化，表现为言语表达不清、肢体乏力、手脚麻木等症状，但仍然具备一定的运动能力。在使用 Pepper 辅助康复训练后，这些患者们明显有愿意多说话的迹象，10 人全部表现为全程积极参与人机互动，患者们更像是看待一个孩子一样看待机器人，愿意给予更多的时间去交流，医生给出的评价是机器人的人机交互能力能够对该类患者的言语康复起到正作用。

### 4.2 Pepper 的人体追踪结果

如下图 12 所示，多次实验中 Pepper 能准确的检测并跟踪目标。所做 10 组实验中，8/10 人次跟踪效果理想，2/10 人次存在被干扰导致某阶段目标丢失的情况，下表 2 记录了 10 组实验的 3 次误差测量结果，在 3 个不同地点的跟踪误差有区别。如图 13 所示，在测量跟踪误差时每组实验选择 3 个点进行误差测量，A 点所测的区域为直线到曲线路径的误差，B 点所测为直线误差，C 点所测为曲线误差。初始设定 Pepper 机器人到被跟踪者的距离为 50cm，表 2 所示误差均为在 50cm 基础上的误差，根据测得数据不难发现，在从直线到曲线的跟踪过程中，机器人需要调整运动方向，而导致误差较大，平均误差 20.04cm；而在直线跟踪的过程中误差最小，趋近于 0；曲线跟踪时 Pepper 机器人则行走在较小半径内，导致误差在负值。



图 12 Pepper 机器人目标追踪过程图

Fig.12 Target tracking process of pepper robot

表 2 追踪过程 3 个测量点的误差

Table 2 Errors of 3 measuring points in the tracking process



序号 / 误差	A 点/cm	B 点/cm	C 点/cm	平均误差 /cm
1	15.2	5.0	-4.6	5.20
2	26.0	0	-7.9	6.03
3	18.6	6.4	-10	5.00
4	21.6	3.7	-9.4	5.30
5	19.0	8.0	-8.2	6.27
6	20.7	5.4	-5	7.03
7	18.6	4.6	-4.5	6.23
8	19.5	2.0	-6	5.17
9	23.3	4.0	-7.2	6.70
10	17.9	7.5	-3.5	7.30
平均误差	20.04	4.66	-6.63	

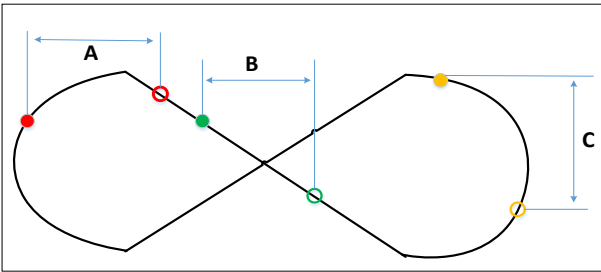


图 13 误差测量示意图

Fig.13 Error measurement diagram

在追踪过程中 Pepper 可以自动检测障碍物，能够达到避障精度为 10cm。如下图 14 所示，在正常追踪的情况下，绘制了 Pepper 与其中一个被追踪者的运动轨迹示意图，Pepper 运动轨迹在一定误差的情况下是可以与患者的运动轨迹相重合，能够完成对人体的有效追踪，追踪过程中采集视频图像数据，为 Pepper 机器人进行人体姿态估计提供运行条件。

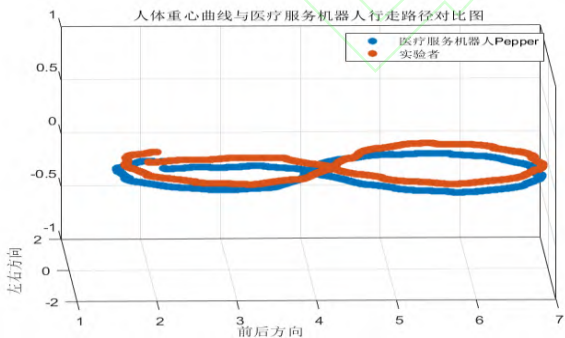


图 14 运动轨迹对比图

Fig.14 The contrast of the trajectories

#### 4.3 Pepper 的人体姿态估计结果

在人体追踪的过程中，通过 Pepper 摄像头采集视频图像信息，经过深度学习网络的判断预测，得到的部分结果如图 15 所示（Pepper 机器人摄像头像素较低）。



图 15 机器人视角的人体姿态估计结果

Fig.15 The results of human posture estimation from robot perspective

图 16 为交并比计算示意图，在目标检测中通常将检测出来的矩形区域与人工标注出来的矩形区域的交集与并集之比称为交并比（IoU），这里 IoU 取值 0.5，只有检测到的区域与真实值的 IoU 大于 0.5 就认为该模型输出结果是正确的，当模型输出多个结果时，取置信度最大的结果。根据 IoU 在 0.5 时的模型输出结果与人工标注区域的对比，计算出了整个过程每个关节的准确率（包含了未被检测出来的关节点）。

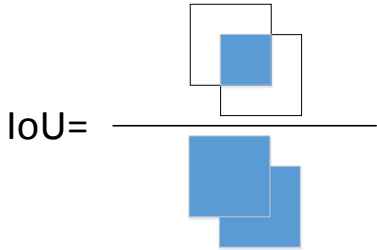


图 16 交并比计算示意图

Fig.16 Calculation diagram of IoU

如表 3 所示，在截取的 10 人共 200 帧测试数据中，头部的测试准确率最高为 91%，腕关节准确率最低有 59.5%，整体准确率受外部衣着遮盖影响。此外 Pepper 机器人进行人体姿态估计的过程主要受到 Pepper 机



器人与电脑端数据传输速度的影响，整个过程会比直接进行人体姿态估计的时间要长，运行速率能够达到2 帧每秒，最终得到的关节点的位置信息可以为人体姿态的稳定性判别提供可行性基础。

表 3 不同人体部位测试结果准确率

Table 3 Accuracy of test results of different body parts

身体部位	头	肩	肘	腕	髌	膝	踝
平均准确率/%	91.0	80.5	68.5	59.5	77.0	69.4	65.0

经过反复对比，总结了肢体部位识别准确率偏低的原因如下：

1) Pepper 机器人采集的视频图像质量不高，且实验环境（医院）灯光一边偏强一边偏弱，在机器人转身的时候有光照影响；

2) 实验者衣着宽松，有的还是穿着裙子，不利于关节点的识别；

3) 除跟踪目标外仍有其他目标人群存在，且有时人体会有相互遮挡；

4) 肘关节、腕关节等目标较小，不易识别。

接下来将会针对以上存在的问题进行改进，进一步提升识别准确率，以便下一步研究。

5 结论 (Conclusion)

本文针对脑神经损伤及认知运动衰退患者认知运动协同康复训练难，临床、居家医疗服务机器人需求大且功能单一的现实情况，主要开发 Pepper 机器人的医疗服务功能，将其设计成一款针对脑神经损伤、认知-运动障碍患者的看护、认知运动协同康复治疗机器人。首先利用多模态感知融合，打破了传统服务机器人只能进行语音交互的短板，增加了触觉交互、人脸记忆识别、跟踪行走等功能，扩展了医疗服务机器人的服务能力，丰富了机器人情感表达，特别是人体追踪将是未来服务机器人发展的一个热点，是机器人融入人类社会生活的一项基本要求。其次，利用机器人的视觉系统，引入了人工智能新技术——人体姿态估计，通过视频图像数据分析人体主要关节点的位置数据，为下一步研究通过视频图像数据来实时判断人体稳定性做出了贡献，相信越来越多的人工智能技术将会应用到医疗服务机器人上，使其更好的为人类服务。

目前该研究的不足之处在于缺乏具有针对性、真实性的语料库，如何能够渐进的锻炼认知障碍患者的语言能力是值得思考的问题；其次 Pepper 机器人与电脑端传输视频等大数据量内容速度慢，导致整个姿态估计过程的运行速率只能达到 2fps/s。接下来将继续研究通过机器人实时跟踪采集视频图像

数据，实时判断人体稳定性情况等问题，拓展认知-运动康复机器人的功能和适用性。

参 考 文 献 (References)

[1] 亿欧智库, 2020 中国服务机器人产业发展研究报告解析[J]. 机器人产业, 2020, (4): 83-100.

[2] Hogan N, Krebs H I, Charnnarong J, et al. MIT - MANUS : A Workstation for Manual Therapy and Training[J]. 1992, 1833(161-5).

[3] Krebs H I, Volpe B T, Williams D, et al. Robot-aided neurorehabilitation: a robot for wrist rehabilitation[J]. 2007, 15(3): 327-35.

[4] Krebs H, Dipietro L, Levy-tzedek S, et al. A paradigm shift for rehabilitation robotics[J]. 2008, 27(4): 61-70.

[5] Hidler J, Wisman W, Neckel N J C B. Kinematic trajectories while walking within the Lokomat robotic gait-orthosis[J]. 2008, 23(10): 1251-9.

[6] E. Ackerman. Berkeley bionics introduces eLEGS robotic exoskeleton. IEEE Spectrum[J]. 2010.

[7] Schmidt H, Hesse S, Bernhardt R, et al. HapticWalker---a novel haptic foot device [J]. 2005.

[8] 陈观养. Pepper SoftBank Robotics[J]. 中国物业管理, 2019, (6): 40-1.

[9] 梁付新, 刘洪彬, 张福雷. 基于 NAO 机器人的目标识别方法[J]. 计算机工程与设计, 2017, 38(8): 2235-9.

[10] 褚光耀, 陈中, 徐佳妮. 基于 NAO 机器人的红球识别与定位算法[J]. 卫星电视与宽带多媒体, 2019, (20): 41-2+6.

[11] 黄怡蒙, 易阳. 融合深度学习的机器人目标检测与定位[J]. 计算机工程与应用, 1-10.

[12] 马迎杰, 王佳斌. 基于 NAO 机器人的目标跟踪和人脸识别研究[J]. 现代计算机, 2019, (13): 62-6.

[13] Coquin D, Boukezzoula R, Benoit A, et al. Assistance via IoT networking cameras and evidence theory for 3D object instance recognition: Application for the NAO humanoid robot[J]. 2019, (9)

[14] 贾文川, 汪志威, 孙翊, et al. NAO 机器人在特殊迷宫环境行走导航[J]. 工业控制计算机, 2019, 32(7): 20-2+4.

[15] 苗凯尧, 刘利芳, 张梦宇, et al. NAO 机器人的语音交互与任务规划研究[J]. 电子测试, 2018, (10): 59-60.

[16] 刘成菊, 韩俊强. 基于改进 RRT 算法的 RoboCup 机器人动态路径规划 [J]. 机器人, 2017, 39(1): 8-15.

[17] Hsieh W F, Sato-shimokawara E, Yamaguchi T, et al. Enhancing the Familiarity for Humanoid Robot Pepper by Adopting Customizable Motion[M]. Iecon 2017 - 43rd Annual Conference of the Ieee Industrial Electronics Society. New York; Ieee. 2017: 8497-502.

[18] Avalos J, Cortez S, Vasquez K, et al. Telepresence Using the Kinect Sensor and the NAO Robot [M]/JULIAN P, ANDREOU A G. 2016 Ieee 7th Latin American Symposium on Circuits & Systems. New York; Ieee. 2016: 303-6.

[19] Hirschmanner M, Gross S, Krenn B, et al. Grounded Word Learning on a Pepper Robot [M]. New York: Assoc Computing Machinery, 2018.

[20] 徐健. 基于 NAO 机器人的医疗助手系统的设计与实现[J]. 信息技术, 2016, (12): 205-8.

[21] 李晓博, 于会敏, 孟昭军, et al. NAO 机器人干预治疗自闭症儿童的人机交互应用平台[J]. 机器人技术与应用, 2020, (1): 35-8.

[22] Matic D, Kovacic Z. NAO Robot as Demonstrator of Rehabilitation Exercises after Fractures of Hands[M]/BEGUSIC D, ROZIC N, RADIC J, et al. 2019 27th International Conference on Software, Telecommunications and Computer Networks. New York; Ieee. 2019: 413-8.

[23] Rozanska A, Podpora M. Multimodal sentiment analysis applied to interaction between patients and a humanoid robot Pepper[J]. Ifac Papersonline, 2019, 52(27): 411-4.

[24] Ilyas C M A, Schmuck V, Haque M A, et al. Teaching Pepper Robot to Recognize Emotions of Traumatic Brain Injured Patients Using Deep Neural Networks[M]. 2019 28th Ieee International Conference on Robot and Human Interactive Communication. New York; Ieee. 2019.

[25] Gardecki A, Podpora M, Beniak R, et al. The Pepper humanoid robot in front desk application[M]. New York: Ieee, 2018.

[26] 李群明, 熊蓉, 褚健. 室内自主移动机器人定位方法研究综述[J]. 机器人, 2003, (6): 560-7+73.

[27] Fujimoto M, Chou L S J G, POSTURE. Sagittal plane momentum control during walking in elderly fallers [J]. 2016, (45):121-6.

[28] Fang H S, Xie S Q, Tai Y W, et al. RMPE: Regional Multi-Person Pose Estimation[M]. 2017 Ieee International Conference on Computer Vision. New York; Ieee. 2017: 2353-62.

[29] Redmon J, Farhadi A. YOLOv3: An Incremental Improvement [J]. 2018.

[30] Newell A, Yang K, DENG J. Stacked Hourglass Networks for Human Pose Estimation [J]. 2016,

作者简介:

张 宇 (1995-), 男, 硕士生。研究领域: 机器人, 生物力学, 深度学习。

刁亚楠 (1994-) 男, 硕士生。研究领域: 机器学习, 智慧医疗。

梁升云 (1991-) 女, 博士生。研究领域: 模式识别, 深度学习。

叶超翔 (1997-) 男, 硕士生。研究领域: 机器人, 材料学, 机器学习。

周艳霞 (1976-) 女, 硕士, 主任医师。研究领域: 脑血管病, 神经感染免疫, 神经康复。

赵国如 (1980-) 男, 研究员, 博士生导师。研究领域: 人体生物力学。

