

# 具有情感的类人表情机器人研究综述

王志良 王 巍 谷学静 郑思仪

(北京科技大学信息工程学院 北京 100083)

**摘 要** 具有情感的类人表情机器人是机器人研究领域中的热点问题之一。对这方面的研究现状进行了综述。首先,梳理了国内外具有情感的类人表情机器人的研究之路,概括性地总结了具有代表性的表情机器人涉及的相关理论与技术,并对比分析了其系统构成、机械结构、传感系统、控制系统和特性;其次,详细阐述了该领域所涉及的理论与技术,包括面部编码系统理论、人工心理与人工情感理论、传感技术、图像处理技术以及语音识别与表达技术,继而针对该领域研究中存在的问题讨论了进一步的研究方向;最后,对该领域未来的发展进行了展望。

**关键词** 机器人,表情机器人,情感,交互

**中图分类号** TP24 **文献标识码** A

## Research Summarization of Humanoid Expression Robot with Emotion

WANG Zhi-liang WANG Wei GU Xue-jing ZHENG Si-yi

(School of Information Engineering, University of Science and Technology Beijing, Beijing 100083, China)

**Abstract** Humanoid expression robot with emotion is as one of hot topics in robot research. To this area, summarization was given. First, after peptination of the road of research on emotional humanoid expression robot, related theories and technologies of representative robot were summarized synoptically. In addition, their system construction, mechanical structure, sensors, controlling system and another features were compared and analyzed. Second, theories and technologies to this area, such as facial action coding system, artificial psychology and artificial emotion, sensing technology, image processing, voice recognition and synthesis were expounded in detail. Based on these, some future research directions to several existing problems were proposed. Lastly, developing of humanoid expression robot with emotion was forecasted.

**Keywords** Robot, Expression robot, Emotion, Interaction

家庭服务机器人是机器人的发展方向之一,是进一步加强机器人技术研究的突破口,是我国“十一五”期间机器人技术发展的重点<sup>[1]</sup>。个人电脑普及革命的领军人物比尔·盖茨预言,机器人即将重复个人电脑崛起的道路,“未来家家都有机器人”。机器人将与个人电脑一样,彻底改变这个时代的生活方式。韩国科学家也预测,未来 10 年间,每个家庭将至少拥有一个机器人。在 2008 年中国科协举办的“五个 10”系列评选活动中,“未来家庭机器人”入选“10 项引领未来的科学技术”,这些都表明家用机器人技术的前沿学科性与重要实用价值。随着服务机器人进入我们的生活,在其能够完成一定任务的基础上,人们已不满足对着冰冷的金属,而是需要具有情感和个性、具有丰富的表情、能够和人类通过动作、语言等进行合作的机器人。因此,对具有情感的类人表情机器人的研究就显得十分必要了。

本文对具有情感的类人表情机器人研究进行了梳理,概括了相关的理论与技术,并对比分析了 8 种表情机器人的详

细技术参数、特性。阐述了目前研究中涉及的理论与技术,在此基础上,对进一步的研究方向进行了探讨。

## 1 国内外研究现状

表情机器人是近年来机器人研究领域中兴起的一个新的研究热点,它既是人工情感的重要研究和应用对象,也是研究开发人工情感的高效实验和演示平台。

对具有情感的类人表情机器人的研究,日本、美国的大学和研究所开展较早,并且多集中在早稻田大学和麻省理工大学,且发展较快<sup>[2,3]</sup>。近期,德国与英国也开始有了相关研究<sup>[4,5]</sup>。我国自 2002 年开始,出现了对象表情机器人的研究<sup>[6]</sup>。目前国内研制的比较先进的机器人是由哈尔滨工业大学开发的“百智星”幼教机器人,它能显示一定的表情<sup>[7]</sup>。此外,还有 2004 年由中国科学院自动化研究所研制的“女童”机器人<sup>[8]</sup>,以及 2007 年由北京科技大学设计的情感机器人头部等<sup>[9]</sup>。图 1 分别对国内外具有情感的类人表情机器人研究之路进行

到稿日期:2010-01-08 返修日期:2010-04-20 本文受国家高技术研究计划(2007AA04Z218),国家自然科学基金(60903067),北京市重点学科建设项目(XK100080537)资助。

王志良(1956—),男,教授,博士生导师,主要研究方向为人工心理及情感计算、智能机器人学、和谐人机交互、3C 融合;王 巍(1983—),男,博士生,主要研究方向为表情机器人、人工心理与情感计算、智能优化算法, E-mail: wangwei8311@163.com(通信作者);谷学静(1972—),女,博士生,主要研究方向为虚拟人技术、人工心理与情感计算;郑思仪(1982—),女,博士生,主要研究方向为机器视觉、人工心理与情感计算。

了梳理<sup>[10-14]</sup>。图2概括性地对其中17个表情机器人涉及的相关理论与技术进行了总结<sup>[15,16]</sup>。

随着技术的不断进步,近年来对表情机器人的研究也逐渐升温。目前,美国 MIT 媒体实验室的研究多关注于机器人

的社会交互性<sup>[17,18]</sup>,而日本早稻田大学的研究则关注于拟人性与情绪性<sup>[19,20]</sup>。我国在此领域也取得了较大的进展<sup>[21]</sup>,所研制的机器人已经具有语音、图像、情感等交互能力,但在社会性、智能性等方面还有欠缺,有待进一步提高。

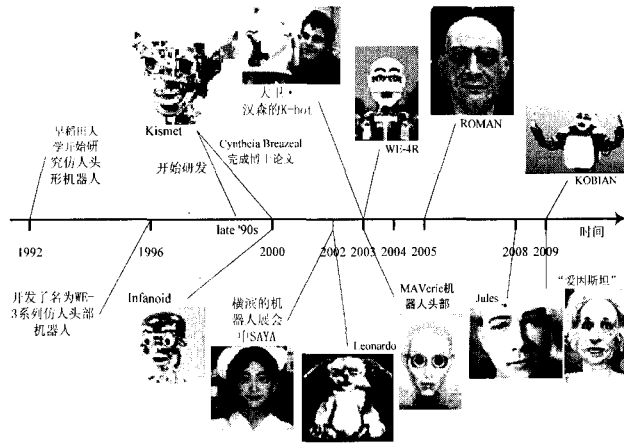


图1 国内外表情机器人发展过程

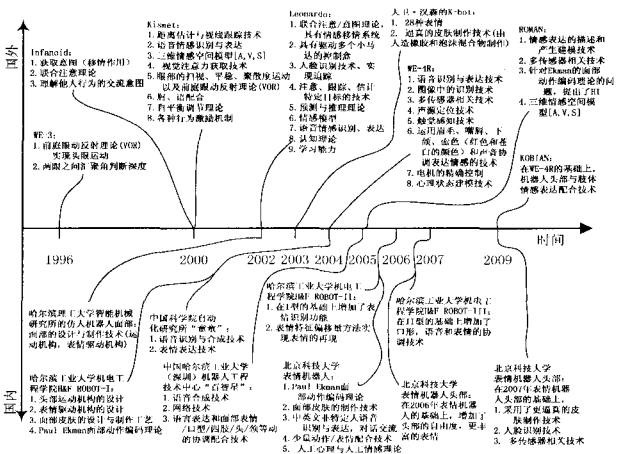
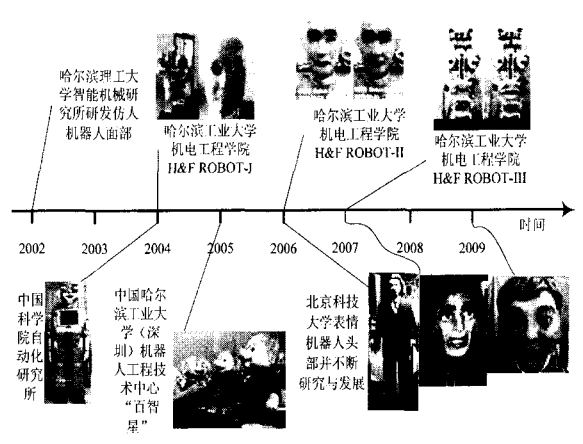


图2 国内外表情机器人理论与技术发展过程

早稻田大学的 WE 系列表情机器人,从早期的 WE-3 仅仅在前庭眼动反射理论的指导下,具有头眼协调运动功能,发展到后来的 WE-4R 具有心理状态模型,能够运用眉毛、嘴唇、下颌、面色和声音协调表达情感等多种功能,取得的进步显而易见<sup>[2]</sup>。国内的哈尔滨工业大学<sup>[22-24]</sup>和北京科技大学<sup>[9,25,26]</sup>在几代表表情机器人研发中的进步也是如此。

为了给表情机器人今后的研究奠定基础,表1、表2详细地列出了8个机器人的系统构成、机械结构、传感系统、控制系统和特性进行对比分析。以上仅对具有代表性的类人表情机器人研究进行了统计,其它具有少量表情或者仅具有类人特性的机器人未统计在内。

表1 系统构成和机械结构对比表

名称	研究机构	系统构成	机械结构
ROMAN	德国 Kaiserslautern 大学	1) 机械系统; 2) 传感器系统; 3) 控制系统; 4) 情感系统	1) 10个舵机推拉线(控制硅胶面皮); 2) 1个舵机控制下颌; 3) 8个金属板支撑骨架
Jules	英国西英格兰大学和布里斯托尔大学联合所属布里斯托尔机器人学实验室	——	34个舵机,控制软性橡胶皮肤——由大卫·汉森(David Hanson)研制
Infanoid	日本东京通信研究实验室	1) 传感——伺服系统; 2) 行为指令系统; 3) 计算系统; 4) 学习机制; 5) 注意力系统	1) 头部 5DOF; 2) 颈部 3DOF; 3) 3个电机驱动眼睛; 4) 嘴唇 2DOF
WE-4R	日本早稻田大学理工学术院高西淳夫研究室	1) 传感——伺服系统; 2) 表情表达系统; 3) 心理状态系统	1) 眼球 3DOF(俯仰轴方向1个、偏航轴方向2个); 2) 眼睑 6DOF; 3) 眉毛 8DOF; 4) 嘴唇 4DOF; 5) 下颌 1DOF; 6) 颈部 4DOF(俯、仰、转、侧); 7) 面色(脸颊)改变是运用蓝色和绿色的电致发光板 EL
KOBIAN	日本早稻田大学理工学术院高西淳夫研究室	——	双足类人机器人 WABIAN-2R 与 WE-4R 的集成: 1) 腿 6×2DOF; 2) 脚 1×2DOF; 3) 腰 2DOF; 4) 躯干 2DOF; 5) 手臂 7×2DOF; 6) 手 3×2DOF; 7) 颈部 3DOF
Kismet	美国 MIT 媒体实验室	1) 特征提取层; 2) 感知系统; 3) 注意力系统; 4) 行为系统; 5) 动机系统; 6) 电机系统	Maxon DC 伺服电机(控制眼部)
Leonardo	美国 MIT 媒体实验室	1) 感知系统; 2) 注意力系统; 3) 空间推理系统; 4) 信念系统; 5) 行为系统	1) 面部 24DOF; 2) 双目视觉 4DOF; 3) 可控耳朵 3DOF; 4) 颈部 4DOF
百智星	哈工大(深圳)机器人工程技术中心	——	身高约 35 公分

表 2 传感、控制系统和特性对比表

名称	传感系统	控制系统	其他特性
ROMAN	1) 立体视觉摄像头;2) 2 个声源定位麦克风;3) 红外传感器;4) 惯性系统;5) 气味传感器(预装)	1) 5 个 DSP(Motorola 56F803)(惯性系统 1 个、步进电机 1 个、颈部电机 2 个机、11 路舵机 1 个);2) PC	1) 颈部有 4DOF(绕垂直轴±60°绕水平轴侧向±30°绕水平轴前后向±30°头部中心水平轴点±30°);2) 具有两种控制方式(基于行为的控制、基于情感的控制);3) 能够实现 6 种表情(情感刺激输入(A, V, S)向量);4) 基于 HI 的行为产生
Jules	1) 眼部的视频摄像头;2) 麦克风	—	1) 可真正理解人类表情,甚至可以根据表情说话;2) 模仿人类的十种常见表情;3) 整合了 ASR、TTS,使用“chat-bot”工具,能够进行语音识别与理解;4) 脸部追踪与识别;5) 保持眼睛注视;6) 视线追踪;7) 能够点头、转头和倾斜头
Infanoid	每一个眼睛有两个彩色 CCD 摄像头(下面的一个是 120°范围的广角镜头、上面的一个是 20°范围的长焦镜头)	1) MCU;2) PC 机	1) 眼球能在 100ms 中 45°范围内扫视;2) 具有与对话者或物体的关注交互能力;3) 跟踪非特定人脸;4) 粗略地确定被跟踪人脸的方向;5) 跟踪具有特定颜色和纹理的物体;6) 交替地关注人脸或物体能力;7) 视线追踪(预研);8) 语音交互功能(少量语句)
WE-4R	1) 两个彩色 CCD 摄像机(眼部,图像采集卡);2) 两个麦克风(定位);3) 触觉(力感电阻,2 层结构)、温度传感器;4) 四个半导体气敏传感器作为嗅觉传感	以太网连接 3 台 PC 机(PC1: 视频、PC2: 音频、PC3: 嗅觉、压力和温度,并计算心理状态,控制电机、面色和声音)	1) 能做出六种不同的表情(每一个情感的强度分为 50 个等级);2) 眼球的角速度接近于人类的 600[deg/s];3) 眼睑开合角速度为 900[deg/s],能够 0.3s 眨眼一次;4) 眼球的俯仰运动与眼睑的开合能够联动;5) 颈部的俯、仰、转、侧最大角速度为 160[deg/s];6) 运用眉毛、嘴唇、下颌、面色(红色和苍白的颜色)和声音表达情感;7) 语音合成应用的是 LaLa-Voice 2001(TOSHIBA Corporation);8) 能够识别任意颜色、同时识别四种物体;9) 声源定位(左右);10) 识别“按、敲击、打击”触觉;11) 闻出酒精、氨水和香烟的气味;12) 具有心理状态模型(反射—情感—智能,情感层包括:情绪动态响应、心境和学习系统,需求模型采用的是马斯洛的需求层次结构)
KOBIAN	1) 六轴力传感器;2) 光电传感器;3) 磁性编码器;4) 陀螺仪	PC 机	具有了 WE-4R 的功能,且能通过身体的动作表达多种不同情绪
Kismet	1) 4 个彩色 CCD 摄像头;2) 2 个广角摄像头;3) 1 个 0.5 英寸 CCD 摄像头(具有 8mm 焦距和高分辨率);4) 1 个微型麦克风	1) 2 个 500MHz PC; 2) 1 个 450MHz PC; 3) 4 个 Motorola 68332; 4) 9 个 400MHz PC	1) 注意力处理(用户眼睛识别);2) 表情表达;3) 视线追踪;4) 机器人姿态控制;5) 语音交互功能(少量语句);6) 学习人类的能力(指令);7) 理解人类的意图;8) 具有情感;9) 调整与人流时的距离(舒适的交流)
Leonardo	1) 麦克风;2) 视觉系统,共 4 个摄像头:1 个安装于机器人后方广角镜头、1 个安装于机器人头顶天花板长焦镜头、2 个安装于机器人眼球中长焦镜头	—	1) 有联合注意功能;2) 有站在他人角度看问题的能力;3) 交流中语音与动作配合;4) 对话技术(主动请求说话、聚焦会神听话、放弃说话权等);5) 指示参考功能(视线、指向);6) 对用户头部跟踪;7) 学习能力;8) 语音识别与语义理解
百智星	—	—	1) 针对 0-8 岁儿童智力启蒙、性格塑造关键阶段;2) 集成了智能仿生、语音合成以及网络扩展;3) 在说话同时能配合做出相关面部表情和口型,四肢、头、颈等也能相应动作

## 2 所涉及的理论与技术

### 2.1 面部编码系统(Facial Action Coding System, FACS)理论

头由颅和面部两部分组成。颅内包含脑,面部则有眼、耳、鼻、舌等特殊感觉器官以及呼吸、消化系统的起始部。人体头部的表层肌肉和颈部肌肉能实现头部运动和面部表情。在人脸展现特定表情的过程中,所涉及的主要肌肉包括口轮匝肌、鼻肌、颧肌、眼轮匝肌等。唇部的状态主要由口轮匝肌完成;皱鼻主要由鼻肌收缩实现;颧骨处脸颊的提高由颧肌收缩完成;眼轮匝肌收缩导致眼睛及眼部周围状况的改变。

瑞典解剖学家约特舍在 20 世纪 60 年代首次发现并研究了人类表情的最小单元。在约特舍理论的基础上,美国心理学家 Paul Ekman 和 Friesen 较早地对脸部肌肉群的运动及其对表情的控制作用做了深入研究,开发了面部动作编码系统(Facial Action Coding System, FACS)来描述面部表情。他们根据人脸的解剖学特点,将面部划分为 44 个既相互独立又相互联系的运动单元(AU),并分析了这些运动单元的运动特征、所控制的主要区域以及与之相关的表情。这 44 个活动单元完成的是脸部肌肉的基本运动,其中某些基本运动的组合就构成特定的面部表情。例如提高面颊和翘起嘴角分别是两个活动单元的基本运动,而两者的综合就形成了微笑的面部表情。

与此类似,通过驱动表情机器人头部的各特征点,就能使机器人具有面部表情。面部整体的机械设计是以 Ekman 的

FACS 分类方法为理论基础实现的。基本面部表情的肌肉运动组合如表 3 所列。

表 3 面部表情肌肉的协同运动

表情	额头、眉毛	眼睛	脸的下半部
惊奇	眉毛抬起,变高变弯,眉毛的皮肤被拉伸	眼睁大,上眼睑被抬高	下颌下落,嘴张开
恐惧	眉毛抬起并皱在一起,额头的皱纹集中在中部	上眼睑抬起	嘴张开,嘴唇轻微紧张,向后拉,
厌恶	眉毛压低,并压低上眼睑	—	上唇抬起,嘴角下拉,脸颊抬起
愤怒	眉毛皱起,眉宇间出现竖纹	眼睛愤怒得眯着	唇紧闭,嘴角拉直,或嘴张开
高兴	眉毛有点弯	—	嘴角后拉抬高,嘴张大,脸颊被抬起
悲伤	眉毛内角皱在一起抬高	眼内角的上眼睑抬高	嘴角下拉

### 2.2 人工心理与人工情感理论

表情机器人要表达丰富的情感,自身就应具有情感系统。情感系统如图 3 所示。

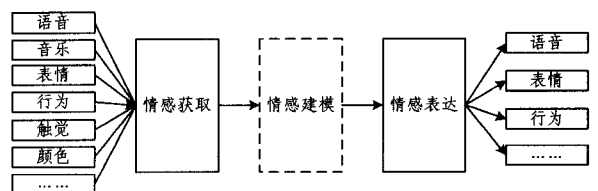


图 3 应用于表情机器人的情感系统

最早提出让计算机拥有情感的人是美国的 Minsky 教授。1985 年,他在“*The Society of Mind*”中提到:“问题不在于智能机器能否拥有任何情感,而在于机器实现智能时怎么能够没有情感”。如今这句话已经成为人工情感领域几乎人皆引用的“经典”。1995 年,美国 MIT 媒体实验室的 Rosalind W. Picard 出版了自己的专著“*Affective Computing*”,从此在人工智能和信息科学领域,有关情感计算或人工情感的研究进展就不断地引起世界各国的关注。日本从 20 世纪 90 年代就开始了感性工学(kansei engineering)的研究,它是将人类的情感和工程学相结合来进行商品制造设计的技术科学。目前日本已经形成了举国研究感性工学的高潮。北京科技大学的王志良教授于 2000 年提出了人工心理的概念,这是我国在情感计算领域的新探索<sup>[27]</sup>。2005 年 10 月在北京召开的第一届情感计算和智能交互国际学术会议,以及中国人工智能学会人工心理与人工情感专业委员会的成立,都标志着我国在此领域的研究达到了一个新的水平。

对情感进行度量的研究,吸引着心理学、认知科学和信息科学等很多学科的研究人员。不同学科的研究者从不同的角度试图模拟情感的产生和变化。由于情感本身的复杂性以及人类对情感变化规律的研究尚不完善,使得这项工作十分艰巨。但是随着对情感认识的不断深入,研究开始有了进展。目前在情感计算领域,已经出现了很多情感模型,用以有限地模拟情感。下面将介绍比较有影响的情感模型,如图 4 所示。

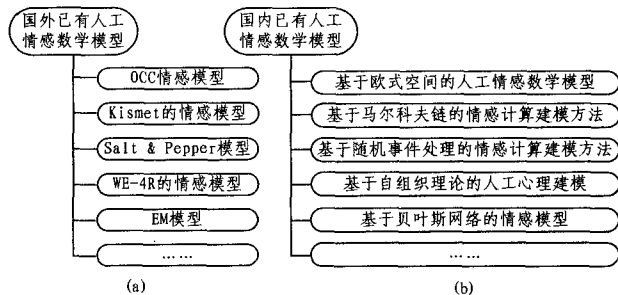


图 4 比较有影响的情感模型

OCC 模型<sup>[28]</sup>:是第一个以计算机实现为目的发展起来的模型,也是当前使用较为普遍的模型。通过分析环境中所发生的各种事件以及和其它实体交互所引发的各种情感,以规则的形式总结并归纳出对应关系。OCC 模型将情感产生的起因分为 3 大类:事件的结果、智能体的动作和对于对象的观感。

Kismet 模型<sup>[29]</sup>:Kismet 是麻省理工学院 C. Breazeal 设计的一个机器人。其情感模型将环境、内部刺激和行为动作联结起来。基于此模型,机器人综合判断外界输入刺激和内部需求,进而引起行为的各种变化。情感模型系统由情感刺激、情感评价、情感激活和情感表达 4 个部分组成,如图 5(a)所示。

基于欧式空间的情感模型<sup>[30]</sup>:以基本情绪为基向量,建立情感空间,如图 5(b)所示。情感状态描述为此空间里的离散点。图中所建立的情感空间具有比较简单的 3 种情感:喜悦、恐惧和愤怒。

基于概率空间的隐马尔可夫过程(HMM)情感建模方法<sup>[31]</sup>:在情感状态概率空间中,利用马尔可夫链和 HMM 模拟情感的基本转移(心境刺激转移、心境自发转移、情绪刺激

转移和情绪自发转移),为情感计算和机器情绪自动生成理论研究提供了一种新的方法。建立的情感空间如图 5(c)所示,其坐标轴分别表示 3 种情感类型。

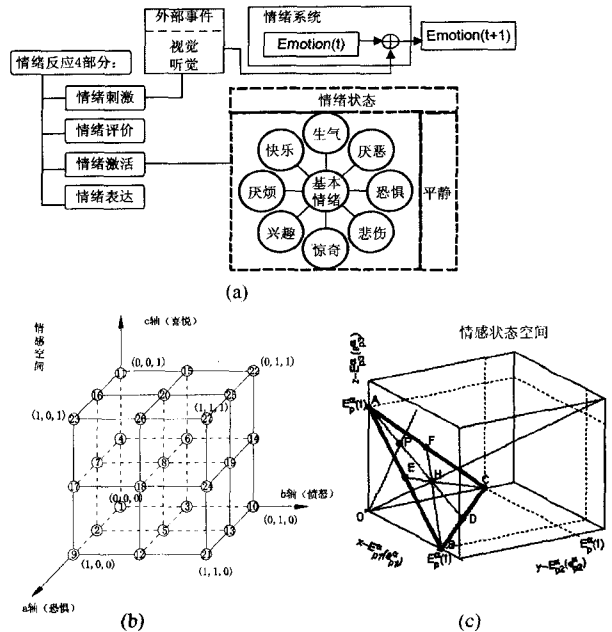


图 5 Kismet 模型框架、基于欧式空间的和基于概率空间的情感模型

### 2.3 传感技术

为了更加真实地模仿人类,表情机器人除了在外表上不断地完善,还在各种感官上接近人类水平。通过安装各种传感器,使机器人能充分地感知外界环境,进而为后续的交互打下基础。一般来说,表情机器人都具有摄像头和麦克,能够进行基于图像和语音的人机交互。

在此基础上,不同的研发者给机器人安装的传感器又存在差别,形成了具有不同感知能力的表情机器人。例如,德国 Kaiserslautern 大学开发的 ROMAN 表情机器人,其额头处安装有测量物距的红外传感器,用于交互时判断前方物体的距离<sup>[32,33]</sup>。这个机器人比较有代表性的传感器是能够分别测 3 自由度的加速和转角的惯性传感器,用于其头部的空间位置估计。日本早稻田大学理工学院高西淳夫研究室的 WE-4R 机器人,采用了具有双层结构的力感电阻作为触觉传感器,使其能够精确地识别“按、敲击、打击”的触觉。采用 4 个半导体气敏传感器作为嗅觉传感,能够区分酒精、氨水和香烟的气味。它还具有两个微型麦克,能进行声源定位,判断说话人的位置。由其和双足类人机器人 WABIAN-2R 组成的机器人 KOBIAN,具有六轴力传感器、光电传感器、磁性编码器、陀螺仪等传感设备。此外,美国 MIT 媒体实验室开发的 Leonardo 社交机器人,其视觉系统较为完善,共安装有 4 个摄像头:1 个广角摄像头,安装于机器人后方,用来获取大范围空间内的信息;1 个长焦摄像头,安装于机器人头顶天花板上,用来俯视拍摄 Leonardo 的工作区,便于人与机器人在工作区内合作完成任务时获取相关信息;另外 2 个长焦摄像头分别安装于机器人两个眼球中,用来获取近距离的人或物的信息。

### 2.4 机器视觉

在与用户交互时,表情机器人通过图像获取的环境信息量非常巨大。对由视觉传感器得来的数字图像需按一定的算

法进一步处理,使计算机能理解图像的意义。对表情机器人来说,主要完成的任务有表情识别、人脸识别、视线跟踪、头部姿势检测以及行为理解等。

Leonardo 利用头部后方和上方的摄像头,获取人脸、头部姿势、行为姿势等信息。例如在其“按钮任务”中,对于按钮的检测是通过按钮的色彩饱和度信息实现的。其过程是首先将获取的图像信息从 RGB 色彩空间转换到 HSV 色彩空间;其次,如果像素的 S 维分量超过一定阈值,就通过色相 H 维分量分隔红色按钮和绿色按钮的区域;最后,当合适的饱和度像素被提取出来后,图像标记算法对其进行扫描,赋予连接和近似连接的像素以相同标记,并将此区域作为一个独立目标。哈尔滨工业大学的 H&F Robot-II 仿人头像机器人具有表情识别机能,能够识别自然、开心、生气、悲伤、惊讶、厌恶等表情。基于人脸几何特征和知识的方法,所开发的表情识别软件包括两个部分:1)人脸区域检测;2)表情特征点提取。其它表情机器人一般也具有相关的视觉,如 Kismet,WE-4R 等,因此,机器视觉是表情机器人的一个基本支撑技术。

### 2.5 语音识别与语音合成

人们可以从触觉、嗅觉等传感信息中得到某些外界信息,但最重要、最精细的信息源只有图像和语言两种。而且,语言是人类最重要、最有效、最常用和最方便的交互形式。

语音识别通过对语音信号进行处理,判别“说话者是谁”或“说话的内容是什么”。利用这一技术可以让表情机器人听懂用户,并作出正确反应。语音合成使计算机具有了“说”的能力,可以将文字等内容转换成语音信号。用户通过“听”就可以明白信息的内容。语音识别与语音合成相辅相成,实现语音交互过程。

英国科学家研制的“朱利斯(Jules)”机器人,整合了自动语音识别技术(ASR)和语音合成技术(TTS)。它利用“chatbot”工具进行语音识别、理解与表达。WE-4R 的语音合成是基于 LaLaVoice 2001(TOSHIBA Corporation)的。Leonardo 采用的自然语言理解系统(Nautilus)是其实验室与 Navy Research Lab 合作开发的。其他的表情机器人也大多具有语音交互功能,如 Infanoid, KOBIAN, Kismet, 百智星等<sup>[34,35]</sup>。可见语音识别与表达技术在表情机器人开发中的重要性。

较为成熟的语音识别与合成软件开发包有:1) IBM ViaVoice;2) Microsoft Speech SDK;3) Pattek ASR 等。它们的性能参数对比如表 4 所列<sup>[36-38]</sup>。

表 4 三种语音识别软件对比

特性	IBM ViaVoice	Microsoft Speech	Pattek ASRV3.4
识别语言种类	英文、中文	英文、中文等	中文
是否支持泛识别	是	是	否
泛识别语言种类	英文、中文	英文、中文等	不支持
是否支持命令式识别	是	是	是
支持命令式识别种类	英文、中文	英文	中文
是否特定人	泛识别特定人 命令式非特定	泛识别特定人 命令式非特定	非特定
是否需要训练	是	是	否
识别速度	快	快	快
环境敏感度	高	高	中

在语音交互技术发展较为成熟后,语音的情感识别与表达研究也引起了研究者的关注,但现在还缺乏较为成熟的产品。

## 3 未来研究方向

### 3.1 仿人头部的研究

表情机器人仿人头部皮肤材料与机械机构的研究是增强其拟人性的基础。由大卫·汉森开发的表情机器人 K-bot,其皮肤是用人造橡胶和泡沫混合物制作的。同样由他制作的“爱因斯坦”机器人头部的皮肤也与真人相似。这种皮肤用一种特殊的海绵状橡胶材料制成。它融合了纳米以及软件工程技术,连褶皱都非常逼真。西安超人高仿真机器人“邹人偶”采用硅橡胶制作,连汗毛、皱纹、毛发、斑痣和皮下血管、皮肤的弹性等细节都纤毫毕现。因此,新材料的研究与引入,对提高表情机器人的拟人性具有较为重要的意义。

在表情机器人研究中,头部机械结构的设计较为复杂。基于 Ekman 所划分的 44 个运动单元,目前的机器人大多仅有 6 种基本表情,以及在 6 种基本表情基础上的扩展,并且都是采用某种机械结构带动面部特征点运动产生表情,虽然能够表现出一定情绪,但不够细腻化和丰富化<sup>[21]</sup>。巧妙的结构设计有两个方面:1) 骨架结构的小型化设计,可以增加头部内空间的相对大小,增加面部特征点的驱动单元数;2) 一个驱动单元驱动多个面部特征点的研究。

在皮肤材料和机械结构研究的基础上,增强表情、头部动作以及肢体动作的协调性是提高表情机器人拟人性的另一个途径。KOBIAN 与 Kismet 机器人能够通过上述 3 方面的协调驱动,加强所要表达的情感,这使得机器人的行为更接近人类。通过对人类的心理、行为的研究,建立丰富的表情、头部动作以及肢体动作协调驱动模式。

### 3.2 多通道人机交互

用户与表情机器人的交互是多通道的,遍布全身的感觉器官共同对人产生影响。随着对机器人系统性能和仿生水平要求的不断提高,在机器人身上需要装配大量的种类繁多的传感器。统一处理这些信息意味着增加了信息量,可能会出现各传感器数据组之间不协调现象。基于图像、语音、文本、触觉等多种传感器提取的可计算信息进行融合,能够增强基本特征的多源复合效用,提高表情机器人与用户的交互能力。

### 3.3 网络技术

表情机器人除了能够与用户进行自然交互,还应具有其他方面的用途,即服务模式的研究,例如作为提供信息服务的客户端等。此外,在交互与服务过程中,其自身知识的更新也是非常重要的,这些都需要以网络技术作为基础。因此,集成网络技术的表情机器人应该是未来的研究方向之一。表情机器人应该能够通过网络访问服务器获取信息并提供给用户,或者能够通过网络控制所在环境内的其他设备。

### 3.4 心理状态的认知计算与社会交互性

研究表明,在机器人与用户的交互过程中,用户越来越希望机器人社会化、拟人化。而现有的人机交互手段几乎都忽略了将用户的认知因素和认知状态考虑到分析系统中。仿人交互的实质是对人心理状态的认知过程。在人机交互中,意图理解、问题分析以及行为预测/决策等相关心理状态认知技术将推动机器人仿人交互的发展与应用,促进和谐人机交互的实现。因此,心理状态的认知计算与社会交互性将是未来的研究重点。

**结束语** 在人工智能、心理理论、计算机、材料科学、机构

学、传感技术、仿生学、控制技术等相关学科发展的推动下,具有情感的类人表情机器人的研发有了较大的进步。本文对表情机器人的国内外发展状况以及所涉及的理论与技术进行了综述,理清了与它们相关的理论与技术的发展脉络,为此领域研究获得更大突破奠定了必要的基础。并在综述的基础上,提出了表情机器人未来研究的4个方向,网络技术、认知理论与社会交互性的研究更是研究的重点。

随着“物联网”时代的来临,数字家庭的研究将步入逐渐实现阶段,具有情感的类人表情机器人将会在未来的人类生活中扮演越来越重要的角色。它在为人们提供方便舒适的服务的同时,也将成为一个新兴的快速发展的产业。

## 参考文献

- [1] 田国会. 家庭服务机器人研究前景广阔[J]. 国际学术动态, 2007(1): 29-30
- [2] <http://www.takanishi.mech.waseda.ac.jp/top/research/we/we-4rII/index.htm>
- [3] <http://www.ai.mit.edu/projects/humanoid-robotics-group/kismet/kismet.html>
- [4] Hirth J, Schmitz N, Berns K. Emotional architecture for the humanoid robot head roman[C]// Proceedings of IEEE International Conference on Robotics and Automation (ICRA). Apr. 2007; 2150-2155
- [5] Hirth J, Braun T, Berns K. Emotion based control architecture for robotics applications[C]// Proceedings of 30th Annual German Conference on Artificial Intelligence. Sep. 2007; 464-467
- [6] 于爽, 张永德. 一种仿人机器人面部的结构设计[J]. 机械科学与技术, 2004, 23(2): 196-200
- [7] [http://www.gov.cn/jrzq/2005-11/07/content\\_93123.htm](http://www.gov.cn/jrzq/2005-11/07/content_93123.htm)
- [8] <http://scitech.people.com.cn/GB/53752/4361040.html>
- [9] 李娜, 陈工, 王志良. 表情机器人设计与实现[J]. 微计算机信息, 2007, 23(12-2): 232-234
- [10] Takanishi, Matsuno, Kato. Development of the anthropomorphic head-eye robot with two eyes[C]// Proceedings of the IEEE/RSJ International Conference on Robots & Systems. Institute of Electrical and Electronics Engineers, Sep. 1997; 799-804
- [11] Takanishi, Takanobu, Kato, et al. Development of anthropomorphic head-eye robot WE-3R with an autonomous facial expression mechanism[C]// Proceedings of IEEE International Conference on Robotics and Automation (ICRA). May 1999; 3255-3260
- [12] Kobayashi H, Hara F. Facial interaction between animated 3D face robot and human beings[C]// Proceedings of IEEE International Conference on Systems, Man, and Cybernetics. Oct. 1997; 3732-3737
- [13] Kozima H, Nakagawa C, Yano H. Emergence of imitation mediated by objects[C]// Proceedings of the Second International Workshop on Epigenetic Robotics. Aug. 2002; 52-56
- [14] Kozima H. Infanoid; a babybot that explores the social environment[C]// D. K, B. A. H, C. L and E. B, eds. Socially Intelligent Agents: Creating Relationships with Computers and Robots. Kluwer Academic Publishers, 2002; 157-164
- [15] Kobayashi H, Ichikawa Y, Senda M, et al. Toward rich facial expression by face robot[C]// Proceedings of International Symposium on Micromechatronics and Human Science. Jan. 2002; 139-145
- [16] Kobayashi H, Ichikawa Y, Senda M, et al. Realization of realistic and rich facial expressions by face robot[C]// Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003). Oct. 2003; 1123-1128
- [17] Breazeal C, Scassellati B. Robots that imitate humans [J]. Trends in Cognitive Sciences, 2002(6): 481-487
- [18] Breazeal C, Edsinger A, Fitzpatrick P, et al. Active vision systems for sociable robots [J]. IEEE Transactions on Systems, Man, and Cybernetics, Part A, 31(5): 443-453
- [19] Miwa, Umestu, Takanishi, et al. Human-like robot head that has olfactory sensation and facial color expression[C]// Proceedings of IEEE International Conference on Robotics and Automation (ICRA). May 2001; 459-464
- [20] Takanishi A. An anthropomorphic head-eye robot expressing emotions based on equations of emotion [C]// Proceedings of IEEE International Conference on Robotics and Automation (ICRA). Apr. 2000; 2243-2249
- [21] 罗旃咄, 张永德, 宋继良. 机器人仿生面部系统研究综述[J]. 机器人, 2003, 25(3): 282-288
- [22] 鹿麟, 吴伟国, 孟庆梅. 具有视觉及面部表情的仿人头像机器人系统设计与研制[J]. 机械设计, 2007, 24(7): 20-24
- [23] 宋策. 具有语音口形及表情的仿人头像机器人系统研究与实验[D]. 哈尔滨: 哈尔滨工业大学, 2007
- [24] 吴伟国, 宋策, 孟庆梅. 仿人头像机器人“H&Frobot-III”语音及口形系统研制与实验[J]. 机械设计, 2008, 25(1): 15-19
- [25] 刘遥峰, 王志良. 基于情感交互的仿人头部机器人研究[J]. 机器人, 2009, 31(6): 10-17
- [26] Wang Z L, Liu Y F, Jiang X. The research of the humanoid robot with facial expressions for emotional interaction[C]// The 1st International Conference on Intelligent Networks and Intelligent Systems (ICINIS 2008). 2008; 416-420
- [27] 王志良. 人工心理学—关于更接近人脑工作模式的科学[J]. 北京科技大学学报, 2000, 22(5): 478-483
- [28] Ortony A, Clore G L, Collins A. The cognitive structure of emotions[M]. UK: Cambridge University Press, 1988
- [29] Breazeal C. Emotion and sociable humanoid robots[J]. International Journal of Human-Computer Studies, 2003, 59(1/2): 119-155
- [30] 魏哲华. 基于人工心理理论的情感机器人的情感计算研究[D]. 北京: 北京科技大学, 2002
- [31] 滕少冬. 应用于个人机器人的人工情感模型研究[D]. 北京: 北京科技大学, 2006
- [32] Berns K, Braun T. Design concept of a human-like robot head [C]// Proceedings of IEEE-RAS/RSJ International Conference on Humanoid Robots. Dec. 2005; 32-37
- [33] Berns K, Hirth J. Control of facial expressions of the humanoid robot head roman[C]// Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Oct. 2006; 3119-3124
- [34] Kozima H. An ontogeny of socially communicative robots[C]// Interactivist Summer Institute (ISI-2001). Bethlehem, PA, USA, 2001
- [35] Kozima H, Zlatev J. An epigenetic approach to human-robot communication[C]// Proceedings of IEEE International Workshop on Robots and Human Interactive Communications. Sep. 2000; 346-351
- [36] Strom V, Clark R, King S. Expressive prosody for unit-selection speech synthesis[C]// Proceedings of INTERSPEECH 2006 and 9th International Conference on Spoken Language Processing. Sep. 2006; 1296-1299
- [37] Mihelic A, Gros J Z. Efficient unit-selection in text-to-speech synthesis[C]// Proceedings of Text, Speech and Dialogue-11th International Conference. Sep. 2008; 411-418
- [38] Lam Y M, Mak M W, Leong P H W. Speech synthesis from surface electromyogram signal[C]// Proceedings of the Fifth IEEE International Symposium on Signal Processing and Information Technology. Dec. 2005; 749-754