

未来阿凡达机器人

产业分析报告

2018-2019

关于我们

清华大学-机器人技术创新创业辅修-产业前沿（必修课）

我们的团队

姜雨欣 邱楷中 宋嘉昊 高毓柯 袁乐康 陈思源

致谢

特别感谢产业前沿的任课教师顾学雍老师。顾老师在每周汇报都提出建设性的建议协助我们完成该产业分析报告。

感谢姜博成助教在整个学期的帮助。

感谢iCenter技术创新创业辅修专业给我们提供宝贵的学习机会以及硬件平台和支持。



顾学雍老师

图片来源: <http://www.icenter.tsinghua.edu.cn/faculty/guxueyong/>

声明

本报告是清华大学技术创新创业辅修专业，2017 年秋产业前沿课程下撰写的。

Copyright:

署名 -You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

商业性使用 - 您不得将本作品用于商业目的。

禁止演绎 -If you remix, transform, or build upon the material, you may not distribute the modified material.

目录

CONTENT

01

简介

- 定义
- 与传统行业的比较

03

现状

- 市场分析
- 技术分析
- 专利分析
- 案例

02

历史

- 科技史
- 人物史
- 机构史

04

未来

- 展望
- 不同国家的发展热度

Part 01

简介

一句话定义：以现有遥操作机器人为核心框架的新型机器人概念。

主要调查点

- 遥操作机器人的发展史与现有的商业产品
- 相关的技术细节发展现状及技术扩展
- 这种机器人可能的应用



//todo

Part 02

历史

“机器人” 这一词汇出现的时间并不长。在西方，机器人最早被称作自动机 (Automation) 或自行控制机器 (Self-Operating Machine)，直到1921年，捷克剧作家卡尔·恰佩克在他的剧本《罗萨姆的万能机器人》(Rossum's Universal Robots) 创造了“Robot”一词。在这个剧本中，机器人一开始是没有情感的，人们输给它什么样的程序，就进行什么样的工作，后来，机器有了感情，于是开始反抗，并且最终消灭了人类。值得玩味的是，机器人语出捷克语“Robota”，意为“奴隶”。

十四世纪至十七世纪，文艺复兴给世界带来了一场艺术与科学的革命。文艺复兴三杰之一的莱奥纳多·达芬奇，不仅是创作了《蒙娜丽莎》的世界顶级艺术家，也是一位顶尖的工程师，设计了大量的飞行器，机械车，坦克之类的超越时代的设计。其中，达芬奇设计的机器人以木头，皮革金属为材料，以齿轮作为驱动装置；可坐可站立，头部会转动，胳膊能挥舞。

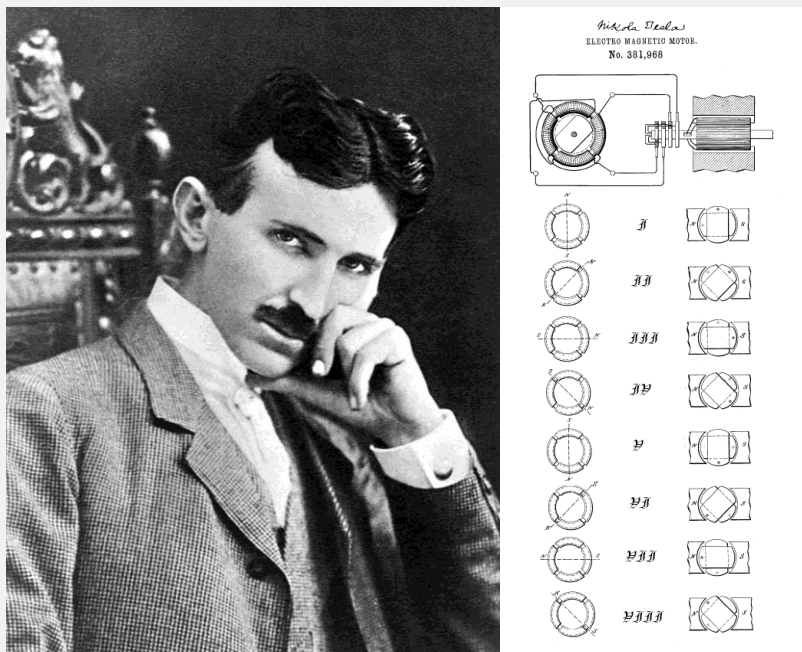
十九世纪中叶，欧洲处于剧烈变革时代。十八世纪六十年代的工业革命给这个世界带来了大量全新

的技术，让整个欧洲完成了从工厂手工业到机器工业的过渡。在这个时代，出现了大量的新型机器（包括机器人），历史学家称这个时代为“机器时代”。



达·芬奇设计的机器人
图片来源：维基百科

近代机器人研究始于20世纪中期，先后演进
出三代不同类型的机器人。第一代就是遥控操作机器人，
之后机器人的行动方式又转换为按事先编好的程序对机
器人进行控制，从而使得机器人自动完成某些重复的操
作。第三代则发展成为了智能机器人，它利用各种传感
器来获取周围的信息，然后在自己的处理器中进行决策
并做出反应。



“许多年以后，人类的机器可以在宇宙中
任何一点获取能量从而驱动机器”

——特斯拉

图片来源: tesla.stractest.org

遥控操作作为遥操作的一个雏形，较早地出现
在历史舞台上。

尼古拉·特斯拉 (Nikola Tesla, 1856--1943)
在美国密苏里州圣路易斯首次公开展示了无线电通信。而
具有历史意义的无线电发射，则是由俄国科学家波波夫和
意大利的马可尼完成。但在当时这个阶段，技术仅作为技
术本身存在，并没有任何与技术相关的无线产品出世；在
萌芽阶段，这个技术在特斯拉电气公司继续沉淀。值得留
意的是，他在1897年，在无线电工程领域注册了20项发
明专利。

最早的遥控装置追溯到1897年，在这个时间
出现了鱼雷控制装置 Brennan torpedo。它是世界上第
一个实践性的导弹控制装置。而在后来出现在世界上的设
备中，无线遥控系统无论是在机器人还是在其它领域，都
得到了普及性运用。

1898年，这项技术出现在民众面前。尼古拉·特
斯拉让人们见到，无线技术能够让人从很远的地方控制机
器运动。他在纽约麦迪逊广场花园向观众演示了一项名为
“远程自动化”的发明——一台无线电操作船。这是基于

他之前发明的感应电机模型，也是“遥操作”这一概念的
开始——他制作了第一台使用遥操作的机器。

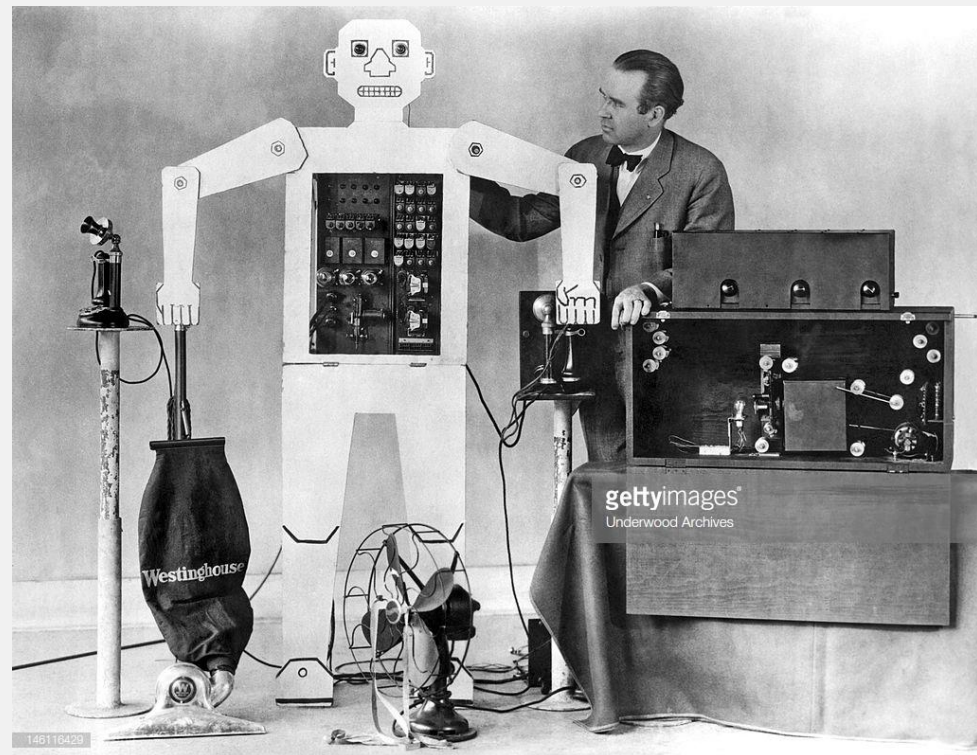
美国西屋公司工程师温兹利，则做出了第一个真正意义上的遥操作机器人。他制造的机器人名为“电报”，在纽约举行的世界博览会上展出。“电报”是一台电动机器人。它由电缆控制，脚部可以运动，装
有无线电发报机，可以回答一些问题，甚至可以抽烟。美中不足的是，该机器人不能依靠自己行走。

当时图灵测试还未被提及，但是就是在这场纽约举办的世界博览会上，“电报”第一次登台；这个身高2.1米，体重118公斤的机器人做了一个颇为风趣的自我介绍：

“女士们，先生们，我很高兴站在这里和大家讲我的故事。我比较聪明，因为我的脑袋比你们的大，里面装有48个电气继电器（electrical relays）。”

当时，“电报”只有700个单词的词汇量。

值得玩味的是，图灵1950年提出检验人工智能的铁律图灵测试，1956年Dartmouth学会提出人工智能，但是在这之前的十多年，“电报”这样的机器人就已经从诞生。



纽约博览会上的机器人“电报”

图片来源: cyberneticzoo.com

“电报”是从西屋电气与开关设备相关的业务中产生的机器人，20世纪20年代，这家公司开发出了全自动的变电站，然后一名工程师提出了能否远程调用变电站系统，于是这家公司开始了用遥操作技术来控制机器人的尝试，“电报”应运而生。

正是由于这个契机，西屋电气开始大力发展遥操作机器人。

“电报”刚开始只是西屋电气工程师Roy J. Wensley设计的一组控制单元，其一部分放置于中央电站调度员的桌上，另一部分放置在变电站。中央电站的那部分控制单元利用音叉振荡器制造不同的声音频率，不同的频率再转化成代码，接收端接受到代码以后对其进行解码，再根据代码的内容打开或关闭特定开关。

当时这个产品只是一组用声音进行远程控制的开关，但是Wensley认为，不懂技术的观众并不会对一个充满继电器和电线的盒子感兴趣。他决定对这个装满电线的盒子进行再包装，于是，这个盒子摇身一变变成了人形机器人“电报”。

1928年2月21日，在博览会之前，“电报”在纽约市的一个高级俱乐部首次亮相。

“电报”的成功引起了公众的注目，并且促进了进一步的研究。美国军方开始研究如何利用这套开关去远程控制枪支开火。而人力不足的郊区机场也在考虑利用这个系统来实现夜间飞机抵达时自动点亮跑道。

1929年，Wensley被调任负责冰箱研发和销售的协调工作。于是他来到了位于俄亥俄州曼斯菲尔德城的西屋电器部门，在那里他遇到了几个志同道合的工程师。

他们很快就推出了Katrina van Televox、Rastus以及其他款机器人，并通过不断迭代优化机器人的运动、声音、控制和技能组合。

Elektro是他们的终极作品。这款机器人有躯干、手臂和腿，由 J.M. Barnett、Jack Weeks sr.、Harold Gorsuch以及其他工程师共同创造。Elektro率先在1939年世博会登场，随后又在全国巡回演出。虽然它大量促进了西屋电器的销售，但Elektro不应该仅仅被当做一个宣传噱头——那群创造它的工程师是有情怀的，他们试图通过制造声控机器人将科幻带进现实。

操作员与Elektro的语言交互在技术上是这样实现的：Elektro并不能真正理解操作员通过麦克风对它说的话，它只能听懂在时间上按照音节精心安排的语音指令，所以操作员发出指令时说得极慢且一字一停顿。

1942到1943年和Elektro一起巡回演出的C.Bruce Hardy说，给Elektro发出指令时，开始或结束一个命令都遵循3-1-2的音节模式，且在分句间有停顿。例如，“Will you come / down / front please?”将使Elektro开始移动。而“You have come / far / enough”会让它停下来。



Elektro

图片来源: <http://cyberneticzoo.com>

操作员也可以在遵循该模式的前提下改变具体的表述。

“Tell us how / old / you are” and “Count your age / with / fingers” 都能触发相同的命令。但是操作员通常都遵循既定脚本，而非对命令改来改去。

根据1939年8月《无线电工艺》杂志上一篇文章的介绍，语音指令是精心安排的音节代码，通过电网发光管转化为电脉冲。电脉冲打开灯泡前的快门，向机器人控制单元的光电管发射闪光灯，Elektro的“电眼”再将信号转换成电流并通过电话继电器传输，最终推动Elektro 的齿轮旋转。

Elektro还在舞台上进行了走路表演，它左膝弯曲着，右腿拖在后面，依靠轮子沿着轨道滑动。

后来的媒体频频拿Elektro抽烟一事大做文章。抽烟是Elektro的经典动作，这也是西屋旗下类人机器人的一贯特色。为了支持抽烟这一功能，程序员们特地在Elektro的上嘴唇上安置了一个小洞。助手会将香烟放在这个洞里并点燃它，Elektro 就会抽上几口香烟，直到助手扑灭烟雾。每次表演之后，操作人员必须清除相关管道中的焦油。吸烟机器人至今仍然存在，只不过是放在了实验室里用于对肺部疾病的研究。

到1958年，Elektro已经有些老了。一名人才经纪人在洛杉矶附近举办的西屋公司展览会上发现了这个老化的机器人---这又给Elektro带来了进入好莱坞的机会——它获得了在1960年电影Sex Kittens Go to College中扮演一个角色的机会。

好莱坞的短暂岁月过去后，Elektro被装进板条箱，送回了它的出生地，然后几乎被遗忘了。这个机器人成

功地在废料堆中生存了下来，并在曼斯菲尔德纪念博物馆找到了一个家，在那里永久展示。

1929年，在地球的另一端，日本的生物学家西村真琴（Makoto Nishimura）在大阪制造了“学天则”机器人，这台机器人的可动性较之前有比较大的突破，这台机器人可以通过空气压力系统移动脸、头以及手。富有戏剧性的是，这个机器人在德国巡回演出的时候失踪，至今下落不明。

到了二十世纪中期，人们并没有将机器人作为机器发展的主要方向。此时大规模生产的迫切需求推动了自动化技术的发展，其结果就是机床等与工业生产直接相关的机器得到大力发展，但是其对控制系统、伺服系统等关键基础的研究成果又成为遥操作机器人坚实的技术基础。此时原子能实验室的核辐射环境要求一些操作机器能够代替人处理放射性物质。在这种技术与需求背景下，美国原子能委员会的阿尔贡研究所与1947年开发了遥控机器手，这是第一个使用遥操作技术进行操作的机器手，而后一年，1948年又开发了机械耦合的主从机器手。所谓主从机器手，是操作员控制主机械手做一连串动作时，从机器手可准确模仿主机器手的动作。这个机器臂开启了一个全新的人机交互模式，并不是以传统的复杂遥控手柄对机器臂进行操作，而是读取人

体手臂的动作操作从属机械臂，更加接近了现代的遥操作理念。

而到了二十世纪中后期，人们将无线与遥控的概念放在一边，开始发展编程机器人。1954年，第一台编程机器人问世，引导了新一轮的机器人风潮。

1950年，恩格尔伯格读到了阿西莫夫的小说集《我，机器人》，被书中内容深深吸引，机器人这个概念让他产生了一个开办一个专门制造机器人的公司的念头。而在1956年，他在一场酒会上，偶遇了发明家德沃尔（George C.Devol）。后者此时刚刚申请了一个名叫“可编程的用于移动物体的设备”的专利。两人一拍即合，决定合作创立一家生产机器人的公司。新公司取名为Unimation，意为“自动化单位”，这个公司与1958年正式运作，当年就研发出来一个可以自动完成搬运的机械手臂，但是这个产品庞大而笨重，只能完成很简单的任务，但是它时人类历史上第一个遥操作机器人商品。

恩格尔伯格和德沃尔研发产品的理念就是面向大公司，为重型制造业服务，让机器人承担对人类有危险的活动。因为在当时的机器人研发成本很高，只有大公司才能负担的起运用机器人的成本，当时美国最大的通用汽车公司，是他们的首要客户。然后，通用公司尝到了用机器人的甜头----机

未来阿凡达机器人 - 历史

机器人的甜头----机器人在生产线上的作用非常明显，动作精准，永不疲倦、不怕高温和污染。于是他们开始订购更多的机器人，并将这些机器人拓展到更多的应用领域。美国机器人协会后来评价，这个发明“彻底改变了现代工业和汽车制造的流程”。

但是随着年龄的增长，创始人之一的恩格尔伯格在技术上变得保守。这导致了他在1983年将Unimation公司以1.07亿美元的价格，卖给了上文提到的西屋公司，退出了机器人界。

受把彭博商业周刊采访时，他提到“家用机器人比工业机器人有更广阔的市场前景”，他也坚定不移的认为，未来的机器人一定无处不在。

西屋公司至今，仍然拥有如雷贯耳的名声。现在的西屋电气，已经成为一个业务领域涉及4000多种产品的大型电气集团。

而公司以外，高校也开始了对机器人研究的脚步。自20世纪60年代中期开始，美国麻省理工学院、斯坦福大学、英国爱丁堡大学等陆续成立了机器人实验室，这些大学只要是以研究智能机器人为主，但是仍然需要运用到遥控操作技术。1968年，美国斯坦福国际研究所研制出了移动式机器人Shakey，拉开了智能机器人的大幕。值得一提的是，

它仍然遵循遥控操作——通过无线通信系统由两台计算机控制，而这两台计算机在当时由一个房间那么大。



“机器人之父”——约瑟夫·恩格尔伯格

图片来源: www.eetimes.com

1967年，日本川崎重工向美国Unimation公司购买了机器人专利，正式将机器人的技术引入日本，次年边研发成功了一台通用机械手机器人。

20世纪80年代，日本的机器人产业进入鼎盛时期。

1980年，日本开始在各个领域推广使用机器人，以此来缓解其国内劳动力严重短缺的社会矛盾，与日本的快速崛起相反的是，作为发源地的美国逐步丧失了在该领域的优势，以至于目前的机器人依赖于从日本进口。

中国相比外界，机器人的相关研究晚了三十年。在1986年底，国家开始实施“863计划”，大力发展机器人，到2000年取得了突破性的进展。而其中一个显著的成果就是遥操作机器人领域：1994年中科院沈阳自动化所等单位研制成功我国第一台无缆水下机器人“探索者”，其工作深度达到1000m，甩掉了与母船之间联系的电缆，完全的实现了遥操作。1955年和1997年，在夏威夷东南海太平洋5800米水深处成功进行了两次大洋探测实验。它能进行海底录像，摄影，海底地势与水文测量和海底多金属结核丰度测量，自动记录各种数据。此时，中国成为世界上少数几个具有深海探测能力的国家，这意味着，中国的遥操作机器人技术正在步入世界前列。

而国家对于机器人的大力度支持，给机器人的发展创造了良好的环境。为了能够持续进行机器人基础科学、关键技术的研究，我国以机器人主题建立了十个实验室与研发中心，在这15年里，一共发表了学术论文2800篇，取得发明与实用新型专利100多项，培养硕士以上学历高级人才1000多名。

到了二十一世纪，遥操作机器人又进入了一个新时代，一个各项技术井喷式发展，以火箭般速度提升的时代。不仅仅是工业巨头在制造商用的机器人，由于各种技术瓶颈的突破，中小微企业也能在当时的市场上取得一隅之地。遥操作机器人也开始进入各个应用领域，形态发生了一些变化。

Part 03

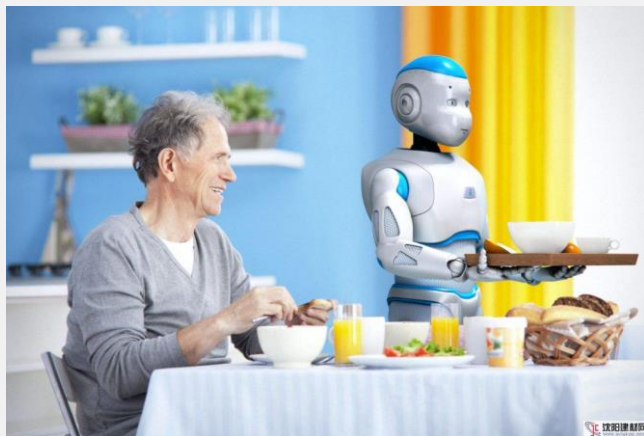
现状

远程陪伴

“有一天，任何一个在城里打工的农民母亲，只需交一点网费，就可以化做一只慈爱的机器熊来到女儿的床边，一边摇着摇篮，一边唱着儿歌。”

为缓解全国约6103万留守儿童（2013年统计）、约8450万空巢老人（2016年统计）的巨大“陪伴空缺”现状，市场上现有的“远程陪伴”产品主要有：

- i宝机器人（儿童陪伴机器人，南京阿凡达科技公司）
- “小鱼在家”、“想家宝”（主打视频通话）
- 家庭看护机器人（美国南卡罗来纳大学Jenay Beer）



2016年，国内儿童陪伴机器人总销售量大约为10万台。这个数字虽然预示了儿童陪伴机器人的市场潜力，但还远远小于应有的市场（千亿级规模）。业界分析，儿童陪伴机器人还未成为刚需，目前依然是一个小众市场。究其原因，是机器人的能力与功能尚还有限。

例如i宝机器人：为实现实际上的交流功能，i宝内置了类似微信的“童信”，让儿童可以添加好友，和亲友保持联系；家长可以和儿童视频通话，也可以远程遥控i宝，监护儿童。在这个过程中，i宝承担了传递语音、代替执行简单命令的功能，但并不是真正意义上的“阿凡达机器人”。

虽然相比简单、吸引儿童兴趣的儿童陪伴，对无法自理的老人需要更为精细的看护，目前尚未找到利用“阿凡达机器人”实现这种看护的案例；但一方面，市场上现有的看护机器人在灵活性等多方面都在为刚需所迫迅速进步着，另一方面，也确实不可否认依然有很多动作更加精细，或者需要更高层次智能的看护行为，只适合有经验的人类完成。与此同时，老人需要的也不仅是生活照顾，而也有强烈的对子女陪伴的情感需求。

可以想象，如果子女能够在工作之余，通过家中的一台阿凡达机器人，与家中老人进行亲密互动，实行贴心照顾，如一边聊天一边为老人捏肩、择菜做饭、修理坏电视等，能够弥补电话、微信产生的疏离感，给老人“近在咫尺”的陪伴。

远程陪伴

留守儿童产生的原因：

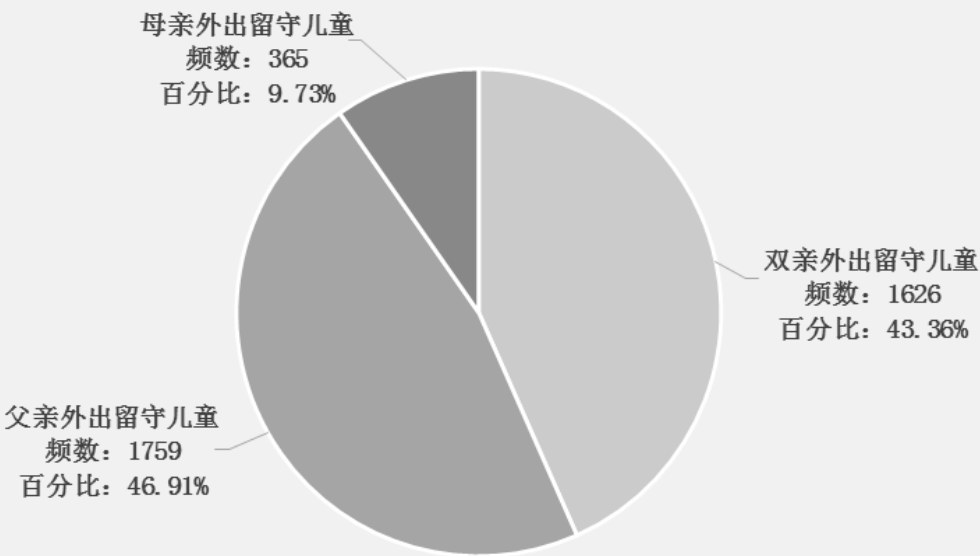
父母为了生计，离开年幼的孩子，外出打工，用勤劳获取家庭收入；孩子们却留在了农村家里，与父母相伴的时间微乎其微。包括内地城市，也有父母双双外出去繁华都市打工。

空巢老人产生的原因：

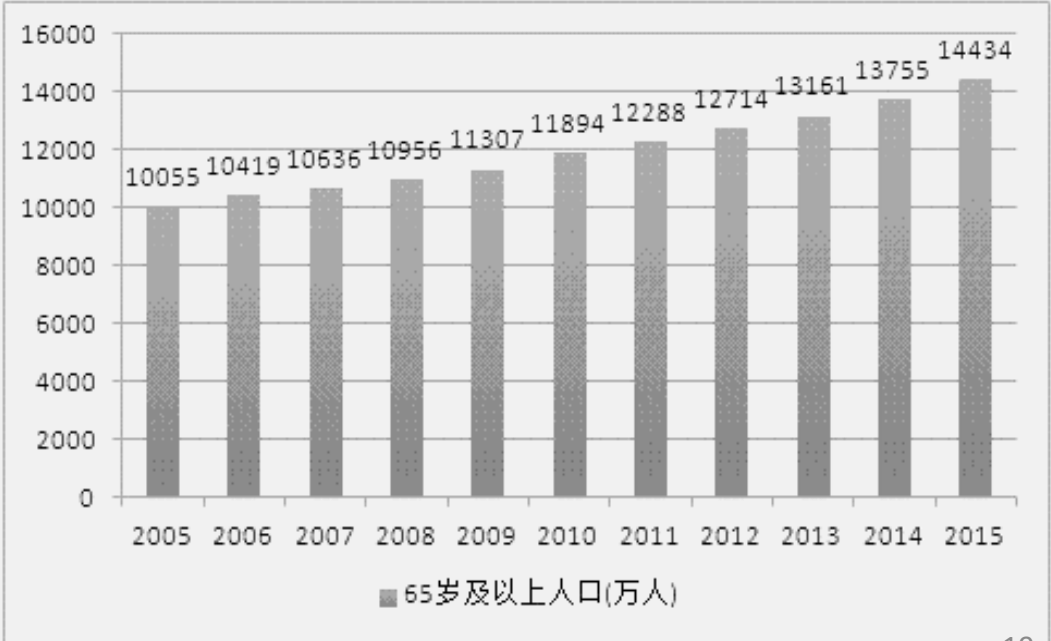
由于计划生育政策，许多父母只有一个孩子，随着父母年龄增大，子女为生活、学习等四处奔波，远离家乡。现代价值观念的冲击，使年轻人更主张追求自由、追求发展，所以离开父母外出求学工作。

阿凡达机器人可以做的：

青壮年在异地求学、工作的同时，能够时常通过家中的“阿凡达机器人”，与家人进行亲密的肢体互动（如：与孩子做游戏、共读；为老人炒菜、捏肩）。 “阿凡达机器人”消除了地理上距离远近的差异，使任何距离缩小为无线网络连接的速度。



数据来源: baijiahao.baidu.com



数据来源: news.qq.com

远程教育

“有一天，任何一位有志于乡村建设，却害怕从此交通不畅而与世隔绝，害怕从此再没有干净水洗澡的年轻教师或是资深专家，能够放下顾虑与乡村开启一段不解之缘。”

//图

//文字

//图

//文字

远程办公

//文字

//图

//文字

//文字

//图

远程救援

“有一天，任何一位邓稼先再也不用用自己肉长的手捡起弹片，任何一位无名英雄再也不用进入毒气泄漏的现场，排除了故障却献出年轻的生命。”

//图

//文字

//图

//文字

远程科考

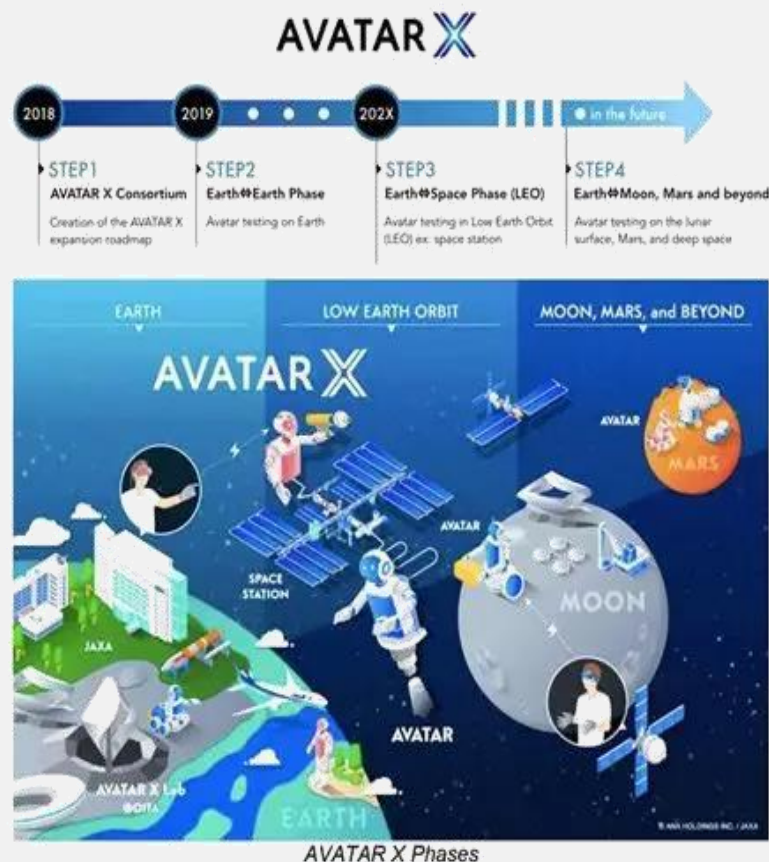
“有一天，任何一个对大自然好奇的孩子，都可以化作化作不怕水的游鱼，潜入最危险最有魅力的海底去逗弄珊瑚，一个小时之后，又准时回到桌前写周三的作业。”

现有的成熟产品：

- “探索者”（大洋探测遥操作机器人，中科院沈阳自动化所）

已开启的重大项目：

- AVATAR X项目（月球探测遥操作机器人，日本宇宙航空研究开发机构JAXA）



随着现代人经济水平和受教育水平的不断提高，科学考察已有从“少数勇者的游戏”迈向普罗大众的趋势。但如果方式仅限于传统的实地考察，是远远不够的。譬如说，一名在校大学生听说内蒙古当地的一种常见植物的花具有本地同种植物不具有的表现型，由于强烈的好奇心，希望做一对比研究。就算没有课业所迫和金钱压力，这名大学生能够买车票到内蒙古去实地考察，也要经历路途遥远，时间漫长的不便；如果碰巧这种植物花期短暂，也可能等到到达现场，已经错过。而实际上，大部分在校大学生都没有这样理想的“说走就走”的条件。如果没有阿凡达机器人的帮助，他可能首先需要联系到内蒙古当地的一位助手，再向他描述自己想要找的那种植物，以及希望助手拍摄照片的要求；而一来一回传递信息，亦会造成时间的浪费和信息的丢失。

如果在内蒙古当地事先有一个“阿凡达机器人”，大学生就可以在居住地自由地操作它，从寻找植物到采样观察，全部由自己独立完成，完全不需任何额外延迟。这样，每个对科学感兴趣的学生，都可以用极低的成本到全世界任何地方进行科学考察，来完成自己的课题。

远程旅游

“未来，旅游可能变得像读书一样随性：你有时间可以任性把书一天读完，没时间也可以每天抽空读两页。”

现有的成熟产品：

- Double Robotics (Segway式移动装置，360°VR全景相机)



//文字

//文字

//文字

动作捕捉

虽然通过手柄、数据手套、弯度传感器等多种外戴设备遥操作机器人已可以获得较高的精确度和较低的延迟，但基于纯机器视觉的遥操作依然是一个技术难点。我们认为，在此方面取得突破很有必要，因为如果不需穿戴任何设备，必将为用户带来极大的便利，减少使用遥操作机器人的（额外）时间成本与财力成本，使遥操作从娱乐产业走向日常工作和生活。

为此，我们更关注基于纯机器视觉的动作捕捉技术：通过图像分析，获取人体任意时刻的姿态信息，并转化为机器人相应的指令。

用机器视觉获取人体姿态的最有名的一篇论文发表在2014年的CVPR会议上。来自谷歌的Toshev给出了他的解决方案：利用深度神经网络（DNN）分析画面；从“整体”出发分析姿态，而不是试图在画面中定位每一个微小关节的位置；对被挡住的肢体部分根据“经验”做出一定的预测。神经网络一共分为

实现原理	计算量	对场所	精细度	便捷度
力学传感	小	要求不严格	可以很高	低
机器视觉	大	要求严格	难以很高	高

7层；为解决运算速度与输出精度之间的矛盾，Toshev创新性地采用了“姿态回归器”（Pose Regressors）。简单来说，就是由之前的层处理像素数较小的输入图片，定位每一个关节的大体位置，缩小最终输出的范围；回归层则处理像素数更大的输入图片（如原图），只求出最终输出与粗输出的残差，来得到关节等的精确定位。用这种方法，每一帧都能达到平均70%的精确度，而且即使关节的定位结果并不精确，人体的形状也几乎不会有差错。如果再考虑到帧与帧之间为保证连贯性而采取一些修正操作，从视频中检测人体姿态的准确度应会比图片更高。

微软开发的Kinect虽然目前停产，但带来了动作捕捉技术的革命；而它的影响力也远远不止于体感游戏。来自秘鲁UTEC的Jose Avalos在2016和2017两年分别在IEEE上发表了基于Kinect实时遥操作机器人的论文。除了对空间位置的精确测量，Kinect还有良好的测量深度的能力，使得人体姿态的空间信息不被丢失。由此，可以求出身体各肢段间的角度，再利用一阶微分逆运动学原理，转化为对机器人姿态的控制。软件采用了ROS（Robot Operating System）。

动作捕捉

这些先人的研究暴露出两个未解问题。首先，这种原理使得机器人的胸口正中央必须在一个固定点上，其他肢体以它为原点做各种旋转动作，因此它不能做任何空间位移。其次，由机器人自身的设计局限性，机器人尚有很多动作无法达到人体所能轻易达到的角度；超过角度限制，机器人会自动切断遥操作以自我保护。

实验证明，由于机器人自身的惯性（而不是数据处理等其他原因），动作延迟存在但小于0.3秒，基本上保持了很好的同步性。然而，当人体做大幅运动时，机器人开始跟不上人类的速度。研究人员分析这是机器人能为自身施加的扭矩有一个上限。当机器人的设计更加精巧，这种延迟也会随之消失（研究人员并不是机械方面的专家，因此没有涉及这方面的工作）。

// 图，图，图

机械系统与运动控制

机械系统是机器人的基础部分之一，它决定了机器人能否运动、能以什么样的形式运动、运动的性能如何。机器人的其他重要组成部分是传感器系统与控制算法，主要功能是感知信息并计算出期望的运动方式。

在一个机器人系统中，机械部分和算法部分的复杂度是不确定的，这一比例主要取决于机器人的用途，例如是更偏重于智能感知，还是灵活运动。我所在的主题组所需要研究的人形遥控操作机器人，在机械部分的主要目标是实现具有较高灵活度的、仿人的动作，但不需要很高的速度和精度；而在运动控制部分，期望的主要功能是模仿操作者的动作，理论上不需要很复杂的闭环控制系统，在机械结构精度足够的情况下开环控制足够，但要求有较好的实时性。

一个机器人的运动性能是由机械系统和控制系统共同决定的，使用更优化、更多功能和更高效的机械结构可以简化控制算法，而使用更完善的传感器系统和算法则能够在当前的

机械系统上达到更好的效果。然而，这两个方面需要相互协调，达到平衡，如果一个方面性能较差，有可能是无法弥补的。

机械系统——以T-HR3和NAO机器人为例

整体动作

机器人的运动主要在于关节的转动，多使用伺服电机实现。为提高转动的精度，或者是对力量的控制，还可以在关节处加入角度或扭矩传感器，形成局部闭环。

例如T-HR3机器人的29处关节和主控系统的16个部位内都安装了带扭矩传感器的扭矩伺服模块，可以精确控制动作带有的力度，并能够反馈给控制者。而规模较小、结构相对简单的NAO机器人则在肢体处安装了压力传感器和触觉传感器。

自由度

T-HR3和NAO都属于仿人机器人，具有较多的运动自由度，例如手、四肢的多个关节，而每个关节的伺服电机旋转角度与人接近，甚至具有略大于人的旋转角度。但在每个自由度上，电机旋转的速度，即关节的灵活性往往不及人。

运动控制

底层控制单元

机械系统与运动控制

仿人机器人的底层控制单元是伺服电机以及相应的控制器件，可能配备有相应的角度、转矩、压力等传感器，使得该电机输出期望的角度或力量。

加入局部闭环控制可以在一定程度上提升控制的准确性，但整体的控制性能主要还是取决于使用的电机的精度、速度、转矩等参数。轻便小巧的伺服电机往往性能较差，而如果需要实现高性能的控制，则所需的电机往往体积和重量较大，且成本高，无法安装进一些精细的结构中，例如机器人的手，因此在执行器的选择上需要进行平衡和取舍。

整体动作控制

为了让机器人实现某个特定的动作，需要让各伺服电机同时转动到设定的角度，给出设定的转矩，并且一些精度要求较高的动作还需要各电机之间有较好的同步性。

T-HR3这类直接模仿操作者动作的遥操作机器人，可以直接读取测量自操作者的关节

角度，并控制电机同步转动即可，因此对电机控制同步性和传输延迟比较敏感。而NAO机器人是完全可编程的，则需要编程者事先设定好期望的动作和角度。

自主动作的实现

T-HR3还能实现全身协调平衡控制，被人或者周围的物体碰到也能保持自身平衡，并且还配备有主动规避干扰功能，能够避免与其他机器人或者控制者接触。另外，NAO机器人也具有防自撞、倒地前使用手臂进行自我保护、摔倒后自行站立的功能。

这类复杂的功能则需要额外的整体感知器件，例如利用陀螺仪测量当前的姿态角，利用加速度计测量运动趋势，利用红外或超声波传感器感知周围的障碍物，甚至可以识别出动作上有冲突的指令并进行智能的管理。同时，还需要设计者对机器人的系统模型有较好的认识，因此才可以在一些极端或故障情况下实现一些自主动作控制。

虚拟行走

地面仿真行走与逼真行走体验领域在近十年并非技术热点。但VR近两年恰为技术热点，逼真行走体验作为VR未来发展的一部分也吸引了更多人的目光。本文将主要介绍几大VR公司的解决方案。

HTC的Room-Scale

房间追踪系统最先由HTC Vive 使用，采用 Lighthouse 定位系统，基于激光和光敏传感器来确定运动物体的位置，允许用户可以在一定的空间内自由移动，最大追踪范围可支持15英尺 X 15英尺(约 4.57米 X 4.57米)，至少需要 1.5米 X 2米的空间。而随后在Oculus Touch推出后，同样可以通过Rift+Touch+第三个传感器来实现更大的房间追踪空间。

HTC Vive上的一款游戏《Hover Junkers》便采用了第一人称视角设计，并且充分利用了Vive的房间追踪系统。该游戏开发商Stress Level Zero曾表示，他们希望玩家能够在游戏中尽量获得接近现实的移动效果。但

当在游戏的实际体验过程中你会发现，并非是让玩家在整个大地图上自由行走，而是通过一台装甲车来穿梭在游戏中。有房间追踪系统的支持玩家可以实现转身以及有限的移动，当然，只是很小面积的移动。

很显然，日常生活中我们也很难拥有一块满足这一技术的安全区域，在未来它也很难能够承载得起我们的真正体验需求。

KAT公司的KAT WALK

这套方案的核心是KAT自家的KAT SPACE（产品页面），它占地仅2平方米，比Vive需要的空间小很多。KAT SPACE使用软性束腰将使用者固定住，弹性绳索既允许用户一定幅度的做动作，同时又能保护用户不冲出机器。关于原地跑动的问题，KAT用了特殊的鞋子，鞋底与KAT SPACE之间比较光滑，跑动会比较轻松，而且感受也接近日常跑动，而不是原地跑动那种小摩擦情形。该设备目前支持行走、下蹲、坐、跳跃等动作，但尚没有翻滚闪避；兼容所有主流VR头戴设备，是玩FPS游戏的一大利器。

不过因为价格原因(52999)，一般人恐怕无法承担这样一台设备，和工作人员交流时对方也表示目前的客户主要是VR体验店。

//todo: 加入学术文献分析

VR反馈

VR传输过程中的难点

延迟与分辨率

手机屏幕的分辨率和延迟时间对于移动VR平台的限制不亚于处理器对移动VR的限制。由于手机屏幕距离玩家眼部很近，QHD (2560×1440) 分辨率的手机屏幕对于用户来说显然不够，细心一点的用户就会发现“纱窗效应”。用户想要体验更逼真的VR就必须配备4k或者分辨率更高的显示器，而这个技术门槛目前对于整个行业来说都还达不到。

除了分辨率之外，延迟问题也不容忽视。在此前的报道中，913VR曾经分析了目前Daydream手机的现状。许多原定推出Daydream平台手机的厂商最后又不得不取消计划，原因就是Daydream平台需要OLED屏幕，而该屏幕目前在产能方面还远未达到要求。厂商迫不得已只好将OLED替换成了IPS。虽然OLED有着比IPS更低的延迟，不过产能问题将会成为困扰该行业的一个棘手问题。

宽带

VR视频相对于传统视频来说体积更大，因此在传输过程中需要更稳定更快速的宽带。在VR直播中宽带对于VR体验的限制尤为突出。一些内容供应商为了能够在有限的宽带下提供流畅的VR体验，不得已降低直播视频的分辨率，因而也极大地影响了用户的体验。在VR社交中用户通常也会消耗比平常更多的流量，这对于用户的资费来说显然是一个巨大的负担。

一般视频传输中的普遍技术考虑

可扩展性

可扩展性涉及到的是对于不同传输协议中的特点。目前所使用的http框架一般是以增加延迟为代价来使得系统承担的传输的负担减小。

视频质量

视频传输过程中涉及到的编码，压缩的问题将会导致视频质量的下降，涉及到更高的带宽和更多的被压缩的比特流。在观察视频质量和视频延迟两个维度时，我们可以看到类似的矛盾。为用户实现更高质量通常会导致更高带宽要求，这是因为需要更高分辨率和帧率，在编码器端需要更多缓存时间(比如Look ahead等技术)来获得高压缩比特流。此外，减少播放器缓冲区可以减少延迟，但往往会显著影响用户体验，因为空缓冲区将导致播放停顿，特别当目标缓冲区设置较小时发生的概率较高。

VR反馈

新的压缩算法一般都会在这几个维度之间加以考虑，从而综合得到相对应的符合需求的算法——具体要根据特定应用场景做出平衡：当延迟至关重要的时，例如视频会议或视频安全监控，可以牺牲扩展性或质量；另一方面，在线性广播中大规模传送高质量媒体内容时，延迟通常会略微增加。

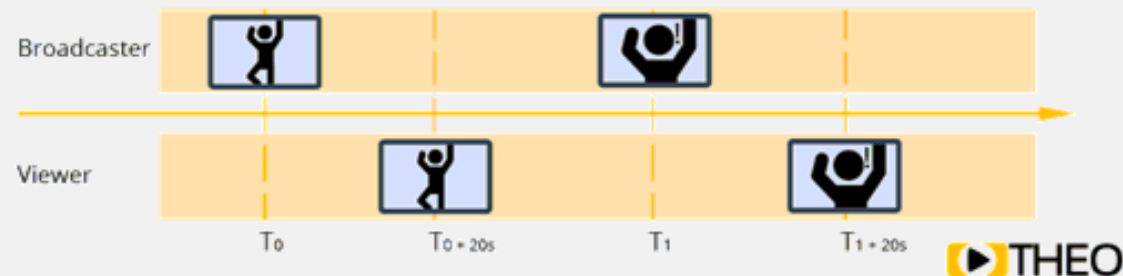
延迟

目前对于视频传输的延迟还没有一个统一的定义。但是普遍来说，延迟可以定义为“某些物理变化的原因和结果之间的时间”。

在交互式的实时体验中，延迟将会严重地影响用户的实时体验。因此对于VR的交互式设备来说延迟的解决是至关重要的一部分。

视频的传输中，大概需要经过以下五个步骤才能完成传输到设备完成播放的操作。在各个步骤中，各种原因都可能导致一定的延迟。

//todo:图片来源



编码和封装：引入延迟和参数配置、质量要求密切相关。

某些流媒体协议可能会引入额外的延迟，因为它们只有在完全接收到后才输出一大块（chunk）媒体内容。

第一英里上传（first mile upload）：将打包内容上传到CDN通常会受到商业条款的限制。

例如，与来自新闻工作室的租用线路设置相比，如果通过无线连接完成上传将会产生更大的延迟。

CDN传播：为了大规模传送内容，大多数媒体管道都利用内容传送网络(content delivery network)。

因此，内容需要在不同缓存之间传播，从而引入额外延迟。

最后一英里交付(last mile delivery)：用户网络连接可能会对延迟产生重大影响。

用户可以在家庭网络连接到wifi热点，或者使用移动连接来访问网络内容。此外，由于可能会选取不同远近的CDN端点，用户地理位置也会造成额外延迟。

播放器缓冲区：视频播放器必须缓冲媒体以确保流畅播放。

缓冲区大小通常在媒体规范中定义，但具有一定灵活性。播放缓冲是延迟的主要因素，优化缓冲区配置是常态。

VR反馈

延迟的一般性优化方法

当前的延迟的解决方案大多着眼于对于传输协议的替换，一般着眼于上述五个步骤中的第一英里传输和最后一英里传输。目前关注的主要焦点是：

- 编码器和播放器之间的视频传输快速而有效。
- 提高扩展性从而增加受众是有收益的。
- 该方法适用于整个设备，无需额外要求。

其中有一种解决方案比较常被提及：CMAF。

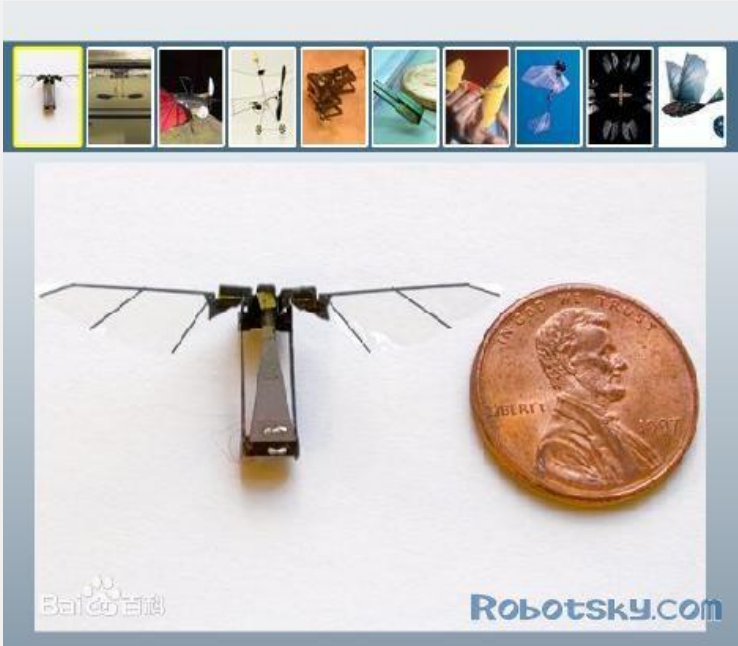
通用媒体应用框架（CMAF）可以使用fMP4容器对多个比特率配置文件中的视频进行均匀分片编码，以便通过HLS或DASH进行流式传输。CMAF提供了一个轻量级框架，它不会引入新各方案，而是以新的方式组合现有格式和标准。随着2017年实现正式标准化，行业中的以后实践就都需要考虑CMAF这种新格式。

微型化

随着MEMS技术的不断发展，机器人小型化并没有技术上的绝对门槛。

MEMS即微机电系统，是指尺寸在毫米级别的智能系统，其内部可集成微传感器、微执行器、微机械结构、微电源微能源、信号处理和控制电路、高性能电子集成器件、接口、通信等功能模块，能够在小尺度上实现机电系统的大多数功能。MEMS具有微型化、批量生产、集成化方便拓展、价格低廉（大批量时）等优势。MEMS当前技术发展主要集中在以下领域：微感知与微控制、微流动控制、微惯性测量装置、微型飞行器、可穿戴和可植入设备、纳机电谐振器、扫描隧道显微镜等。实例：在两千零几年，加州大学伯克利分校就曾研制出微型机器人苍蝇，可模拟苍蝇拍打翅膀，翅翼仅有3cm，重量越300mg。同期，上海交通大学研制出可自由前进、后退与转弯的微型六足机器人——“银甲虫一号”，尺寸约为3*3*4.2（cm）。以上均可说明机器人小型化在技术上没有绝对瓶颈。

微型化



世界上首款机器苍蝇在哈佛成功起飞
图片来源: baike.baidu.com

可能难点

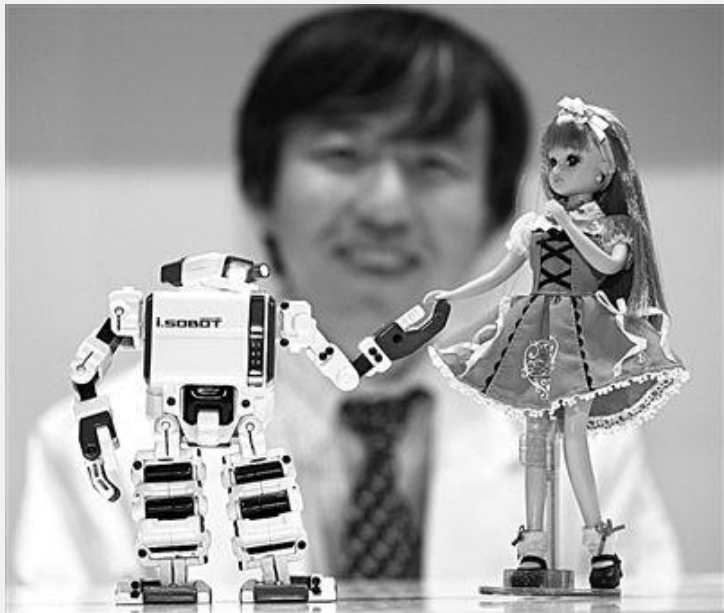
成本上：批量生产时，MEMS成本较低，但是设计成本和研发成本或某些性能要求比较高的关键元件成本较高。

耐用性上：机械结构的缩小带来了受力能力的削弱，可能会导致机器人易损坏，使用场景减少（指需要受力或剧烈运动的场景）。

结构和功能的复杂性上：由于阿凡达机器人目前尚未出现原型机，并不知道在复杂的结构和功能中是否存在某些组件不能缩小到理想的尺寸上，从而限制了整体的尺寸大小。此外，由于其具有高集成度的特性，也有在各功能组件都达到要求后仍然不能有效缩小尺寸的可能性。

供电和运算方式的选择上：如果采取电池供电，那么由于电池能量密度的限制，在电池技术没有重大突破的情况下，阿凡达机器人绝对无法有效实现小型化。如果其数据处

理单元集成在机器人本身上的话，以目前的处理器技术，可能无法有效缩小尺寸；如果将处理器技术分离出去，就额外需要一台主机和数据线，而且对使用场景也有限制。



日本多美公司推出世界最小人形机器人
图片来源: discovery.163.com

//专利分析
//一个巨大的todo



图片来源: <http://www.doublerobotics.com/double2.html>

产品简介

Double Robotics是MIT斯隆商学院开发的一款远程现实机器人。这款机器人造型简单，视觉及大部分功能基于iPad，运动功能与Segway平衡车类似。通过增加运动模块，Double Robotics让视频通话有了更多的灵活度，让用户可以在世界的另一端自由移动，更加真实地参与课堂、办公、社交、旅行。

应用案例

Peter因患癌症无法长距离通勤，只能在家上网课，感到孤独。借助Double Robot，他找回了参与真实课堂和社交的快乐，在Old Mill Patriots高中与同学们打成一片。

技术架构

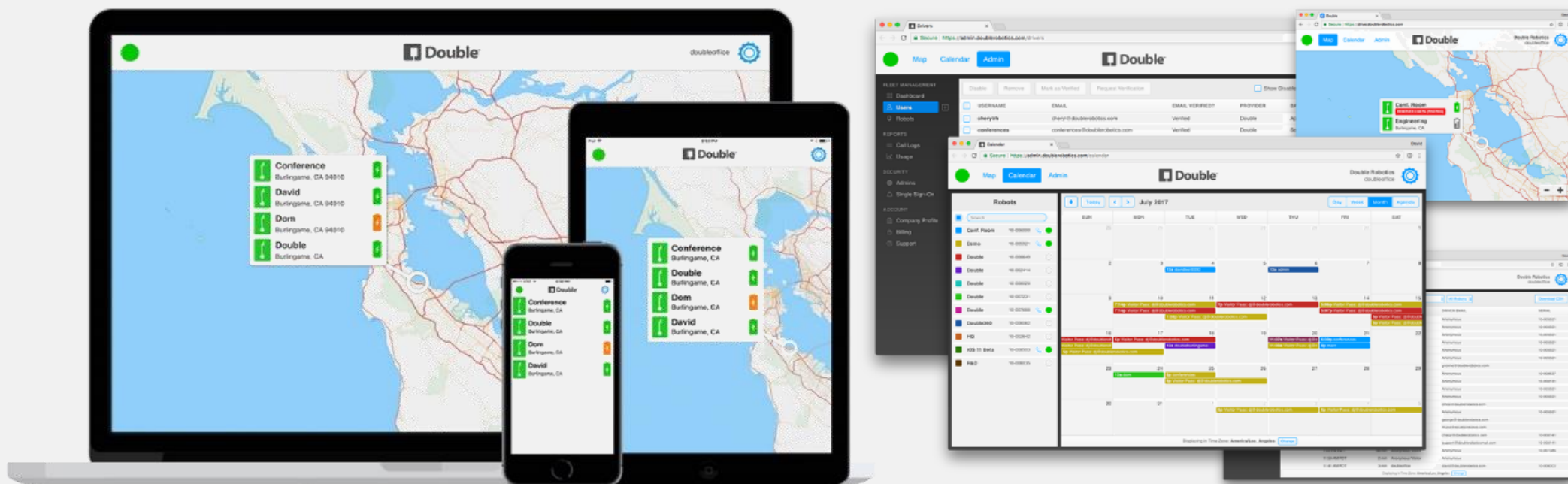
- **运动与平衡** 运动系统采取轮式，具有平稳度好、控制简单、移动效率高等优点。Double Robot继承了Segway平衡车的特点，能够主动调整角度，用闭环的控制系统保持iPad支撑杆竖直，且对复杂地形有一定适应作用：可以在电线、门槛、泥土地上平稳运行。一旦不慎与物体相撞，还有预置的自回弹算法(原文：self-bouncing algorithm)来确保其迅速恢复平衡，从视频上看几乎没有明显摇晃出现，这一点比一些平衡车有所增强。另一点增强是，当用户希望Double Robot静止不动时，可以命令其伸出双“脚”来保持固定。
- **交互与体验** 移动命令由用户手动实现，方法为电脑上的键盘按键。有前进、左转、右转三个自由度（此外），没有后退（除了撞击后自动后退），认为这是一个劣势。此外还有调节杆高的自由度，可以将视线与交谈对象水平。用户通过电脑/移动终端的屏幕获取Double Robot的视角。除iPad自带的向前摄像头外，Double Robot还提供向地面的摄像头，方便用户移动Double Robot通过狭窄区域。但实测发现用户只有向前和向下两个视角的情况下，不会主动向下看而造成视野盲区的存在，经常撞上物体或掉进坑洼，认为这是另一个劣势。如果改进为VR的无障碍视角，可能可以进一步改进用户体验。“前进”分为平速前进和快速前进（适用于长距离无障碍移动）两个命令。

市场分析

截至2017年1月18日，Double Robot已售出超过8000个机器人，产生超过2000万美元的销售输入。

北京京西学校成为首所提供交互式机器人的国际学校，其采用的就是Double Robot公司生产的K2-12机器人。伴随着越来越多的外籍人员选择在中国工作或定居，解决孩子上学和择校的问题也成为家庭的一个重中之重。许多家庭对北京京西学校比较感兴趣，想送孩子到京西上学，但又因为地域限制不能亲自到校实地考察。为满足日益增多的外籍家庭远程看校的需求，北京京西学校购置了1台K2-12。家长可控制它抵达校园的每一个地方，包括：教室、体育场、剧院、图书馆等很多场所，同时会有京西学校的招生负责人全程陪同进行讲解。整个过程简单易操作，外籍家庭可以360°无死角旋转参观，真实还原校园的学习和生活场景。

- **视频与传输** 除iPad自带的摄像头外，Double Robot在与自带摄像头等视角位置配有5兆像素的150°广角摄像头。Double Robot另一个鲜明的特点是自适应高清技术（原文：Adaptive HD），可以在移动时降低分辨率而获得流畅图像，静止时提高分辨率（最高到HD级别）而获得清晰图像。
- **网络与共享** 用户可以在Mac、PC/Windows、iPad、iPhone、iPod Touch等不同终端上远程使用Double Robot（与苹果公司合作较为密切），使用谷歌或火狐浏览器作为网络平台。用户可以在不同时间预约不同地点可供使用的Double Robot，并整理成个人的工作时间表。



产品简介

T-HR3是丰田公司在2017年发布的第三代人形机器人，于11月29日至12月2日在东京国际展览中心举行的2017年国际机器人展览会上亮相。

T-HR3是一种灵巧、具有多个自由度的人形机器人，通过遥操作和VR技术由人便捷准确地进行控制，可以实现护理、搬运、环境探索甚至外层空间考察等一系列任务，在保证人类安全的前提下充分利用了机器人的精度的强度。

T-HR3机器人高1.54米，重75公斤，它的操控系统“Master Maneuvering System”由椅子上的16个扭矩伺服模块、脚下的力与运动传感器以及位于机器人关节中的29个扭矩伺服模块组成，可以方便地控制机器人身体的运动。

机器人的视野由HTC Vive头显提供给操作者，操作者通过操纵外骨骼来控制机器人的动作。同时，即使机器人撞上障碍物，它也可以保持平衡，不会摔倒。当拿起物品时还可以控制机器人手部的力量。

技术架构

该系统主要分为两个部分，机器人本身与控制平台。

机器人以具有很高灵活性的人形机器人平台为基础，在头部安装了充当眼睛的摄像头，在脚的部位还安装有运动和力传感器。机器人本身可以接收来自控制平台的信号，接受来自人的控制，模仿操作者做出

的动作。另一方面，机器人还具有一定的自主运动能力。因为控制平台的限制，操作者无法做出行走的动作，因而机器人需要自行实现走动的功能；另外，为实现在受到撞击的情况下依然能保持平衡的功能，机器人本身还需要充分利用来自各种传感器的信息，做出自主的动作与平衡控制。此外，机器人还配备了触觉传感器，可以将受到的来自物体的力反馈给操作者，从而精确控制动作的力度。

控制平台主要有采集操作者动作和为操作者提供视觉和听觉信息的功能。视觉和听觉信息由VR头显设备提供，而动作的采集也分为上半身和下半身两个部分。上半身的动作采集主要由固定在操作者手臂和背部的转矩伺服模块完成，能够精确地检测各个关节的运动角度。而操作者腿部的动作是不完整的，即不能在操作平台上



(P37续)

完成完整的行走动作，因而腿部动作的采集由运动和力传感器完成。

总体来说，T-HR3机器人综合了（开环）遥操作和自主运动的功能，实现了以模仿操作者的动作为主、又具有一定自主性的运动与控制模式。

应用与开发目的

“合作伙伴机器人团队成员致力于在T-HR3中使用这项技术来开发与人类共存的友好和乐于助人的机器人，并协助他们进行日常生活。展望未来，为这一平台开发的核心技术将有助于通知和推动未来开发机器人，为所有人提供更好的机动性”，合伙人机器人部门总经理Akifumi Tamaoki表示。

丰田公司自20世纪80年代以来的主要开发方向是工业机器人，而这一次发布的合作伙伴机器人发挥了丰田在技术经验积累上的优势，开发了新的移动解决方案，致力于支持医生、护理人员、病人、老年人和残疾人士。

目前，这种具有很强综合能力的机器人尚未在市场上大规模出现，但就此次展览会上展示的T-HR3中的技术成熟度来看，此类机器人进入市场也不再是遥不可及的事情了。为使这种机器人大规模进入市场，进入民众的日常生活，还需要考虑的问题主要有：量产和成本控制、产品定位、功能细化等等。

产品简介

远程会议（Telepresence）目前主要为思科公司的产品，Telepresence目前主要基于2D的范围进行远程的会议交流，相关的3D产品还没有广泛地应用到商业的会议上。基于3D的远程会议被业界坚称是未来的会议的主要模式，但是受限于当下的一些技术，它尚不能成为当下流行的应用。

目前的三维全息影像的投影需要佩戴相对应的设备进行AR的技术增强：其中较为出名的有微软的HoloLens眼镜；HoloLens眼镜利用深度摄像头对物体进行捕获，通过HoloLens眼镜中的GPU实现对深度摄像头采集到的信息进行相对应的编码化处理，重构出3D图像



未来阿凡达机器人 - 现状 - 案例 - Dexmo

(P38续)

并且使其通过衍射型屏显显示在人们的眼前。

这一款AR眼镜研发的理念基于微软对混合现实(mixed reality)的坚定信念——“微软坚信，混合现实是计算的未来。HoloLens不仅代表了混合现实当今最尖端的技术，更代表了混合现实未来发展的方向。这一认知不会止步于简单的产品计划。”

Hololens主要包括三个方面的技术：一是对实时物体的三维计算，利用渲染的方式将物体进行三维化、可视化；二是利用对姿态和物体的精准确定和位置识别，利用Microsoft-Kinetic以及设备上的深度摄像头得到的相关的位置信息深度信息来实现三维物体的相对应优化，但是其在初始化的过程中需要利用红外线对整个房间进行扫面以得到相对应的模型和深度信息；三是精准的图像识别技术。

但是这一款被称作“改变未来潮流方向的眼镜”早在2015年就被微软发布，但是在17年的时候被外界怀疑已经停产。



//文案todo

Part 04

未来

“远程操控机器人”是除了AI之外，另一个机器人行业发展的趋势。

——Rich Walker

机器人技术的研究重点应从全自主方式转向人机交互方式，尤其是在未知环境中作业的机器人，这样更具有现实意义。

——Kazerooni.H

中国

东南大学人机交互遥操作机器人已在我国载人航天与探月工程中得到应用

日本

制定了一项为期8年的先进机器人研究计划实现临场感遥操作机器人系统

美国

NASA针对空间遥操作机器人关键技术进行深入研究

德国

针对空间机器人的应用开展临场遥操作机器人ROTEX研究计划