

白皮书

英特尔中国研究院

# 自适应机器人交互

英特尔中国研究院首席研究员 张益民

英特尔中国研究院高级研究员 任海兵

英特尔中国研究院高级研究员 杨华

2018 年 5 月



## 摘要



个人服务机器人在近年来随着人工智能的快速发展开始得到不少关注。个人服务机器人是否真正能够作为消费级产品大规模进入家庭？其中的技术突破口又在什么地方？作为英特尔中国研究院近两三年来重点投入的一个研究方向，我们对相关问题进行了深入的研究，尤其是相关的应用模式的分析，我们得出了自适应机器人交互可以大大提高现有 AI 技术的可用性的结论。这篇文章将介绍我们对养老服务机器人的先进应用的分析，然后会重点对其中的关键技术领域自适应机器人交互进行介绍。最后，我们会分享我们对下一步工作的一些看法。随着下面两三年更深入的研究，还有基于 HERO 开放平台的共同实践，我们相信我们能够和生态系统伙伴们共同实现在未来五年内个人服务机器人进入家庭提供先进服务的愿景。



## 目录

摘要 .....	1
服务机器人的美好未来.....	3
关键能力分析和主要技术挑战 .....	5
自适应机器人交互 .....	8
自适应机器人交互软件库 v1.0.....	13
总结 .....	14

## 服务机器人的美好未来

近年来，随着人工智能技术的快速发展，服务机器人成为产业界关注的一个热点。不过服务机器人的应用方面目前和智能手机等的应用还比较相似，自主性尤其是学习能力也比较欠缺，很多时候只是预先编程的几个功能，难以满足用户的需求。如何突破现在的简单应用模式，让服务机器人变得更加智能，更加自主，成为服务机器人行业发展的一个重要问题。

我们在电影中看到了很多机器人，有些非常智能，那么个人服务机器人到底离现实还有多远呢？从我们的调研和对人工智能技术的理解来看，我们相信一些相关的新技术将有可能在几年内准备就绪，从而能够使得个人服务机器人的应用实现一个飞跃。

过去两年多，我们与世界各地的一些机器人领域的著名学者进行了广泛的交流，其中包括美国南加州大学的 Maja Mataric 教授和汉堡大学的张健伟教授。Mataric 教授是社交机器人学(Social Robotics)方面的著名学者，她特别喜欢使用机器人主题的科幻电影来谈论服务机器人潜在的市场和技术准备情况。她表示，很多电影中的机器人只是幻想而已，离现实还很远（由于人工智能技术仍然需要很长的时间才能全面超越人类的智慧），但其中少数电影中描述的一些应用场景在不久的将来更具技术现实性。其中一部她推荐的电影就是“Robot & Frank”，它是一部关于老人看护机器人的电影，她认为里面的部分场景在不久的将来是可以实现的。

从社会需求来看，我们也持有相同的意见。我们相信，老年人看护将是个人机器人的一个刚需。根据北京大学国家发展研究所 2005 年的统计，每一位老年人(60 岁以上)按人口比例换算可以得到约 6.1 名劳动力的看护支持，也就是 1 比 6.1。到 2025 年，这个比例将下降到 1 比 2.5<sup>[1]</sup>。另据最近的统计，在 2.3 亿老年人（60 岁以上）中，约 51.3% 是空巢老人<sup>[2]</sup>。因此，缺乏足够的劳动力来对老年人进行看护将会成为越来越严重的问题，个人服务机器人可以大大缓解这方面的问题。最近看到社区养老方面已经有公司开始尝试智能老人护理（使用智能设备），而服务机器人由于移动性，自主能力，更好更自然的人机交互等优点，未来将很可能成为老年人看护的理想选择。个人服务机器人并不会完全取代家人的关爱，但会大大帮助家人提高看护效率，减少对其工作和其他家庭义务的影响，而且个人服务机器人还可以通过一些应用（比如下面提及的沟通助手）来加强老人和家人之间的沟通，作为他们互相表达关爱的一个桥梁。除了一些最基础的类似于电脑手机的应用如播放歌曲，简单问答等，服务机器人其实可以完成一些更有用的服务：

- 日常沟通助手：手机对于年轻人来说是一个非常高效的沟通工具，语音视频方式都可以，但对不少老人而言，使用手机还有上面的应用还是挺困难的一件事，尤其是年纪大一点之后由于眼睛看不清、手指不灵活等原因更加困难。如果服务机器人能够根据老人的需求（比如通过语音发出命令）来和家人进行各种方式的沟通，那么这个应该是一个非常实用的功能。
- 智能导航及和谐空间共享：机器人为了完成任务是需要在家里移动的，那么首先需要更加智能的定位导航。另外因为机器人和老人共处一个家庭空间，老人又比较容易受干扰而发生跌倒等危险情况，所以机器人应该非常注意，要与老人进行和谐空间共享。不影响老人的日常行动，也体现良好的社交礼仪(如礼貌的让路等)。
- 日常起居监控：对老人的日常起居情况进行监控, 必要时进行确认，并把相应的情况尤其是异常及时通知给家人或看护人员。这些信息对于子女而言能够让他们能随时了解父母的日常起居情况，是非常有价值的一个应用。
- 紧急情况检测及救助：从多种模态如视觉，声音来检测特别需要紧急处理的异常情况(比如跌倒、痛苦的表情、很大的声音等等), 并及时提供救助服务，比如通知家人，帮助拿救心药等。紧急情况检测及救助一直是老人看护方面的一个刚需。
- 智能提醒：目前手机上都有定时提醒功能，比如到了吃药的时间提醒老人吃药, 不过生活中有不少情况简单的定时提醒并不够，而是要根据目前的情况来进行相应的提醒，比如老人出门的时候提醒带上一些出门必备的物品(手机，钥匙，老人防丢器等), 天冷的时候可以提醒带上帽子、围巾，还有不少老人开了火之后可能就会忘了而直接出门，那么在在这些情况发生的时候机器人应该能检测出来并及时进行提醒。
- 情感关怀：比如机器人和老人聊天，并能在聊天中应用以前聊到的一些信息，如老人的喜好等，相当于机器人是有记忆的。机器人还可以通过分享老人的一些情况(比如拍摄一张老人做的美食发到家人的微信群里)，加强老人和家人之间的沟通，作为他们互相表达关爱的一个桥梁。

其他的应用应该还有更多，附录 1 列出了以上应用场景和更多的一些应用场景的更详细的信息。通过我们的一些初步调研，如果能够实现其中的两三个应用，相信已经是非常受用户欢迎的产品了。除了老人看护，我们认为个人服务机器人在客户服务(比如前台服务等)，儿童教育，家庭助手等方面都会有很好的应用前景，不过技术上和老人看护都有不少共同

的需求，所以实现了老人看护这样的刚需应用，那么实现这些其他的应用也会变得相对简单。另外我们也看到智能设备现在包含了个人电脑，智能手机，以后也在往智能家居的方向发展，个人服务机器人的一些相关技术可以应用到智能家居的一些设备上，而个人服务机器人将成为智能家居中的一个重要组成部分，可以和智能家居中的其他设备和传感器形成联动，共同为用户提供更好的服务。

根据我们对市场和技术的理解，我们认为如果能够在技术上解决一些主要的瓶颈，那么服务机器人市场在几年内有个飞跃式的发展是很有可能。如果说具体的市场体量，我们做了一个比较粗略的预测，在 2020 年仅仅在养老服务机器人这块，如果能做出满足中国 4%老人(到 2020 年根据趋势预测中国大约是 2.5 亿老人)需求的产品，那么市场的总量已经达到了 1000 万台，而随着产品功能不断的提升并满足更多用户的需求，市场的空间是非常巨大的。

## 关键能力分析 & 主要技术挑战

我们对养老服务机器人的应用模式及其要用到的关键能力进行了分析，见下表。

	特定人识别	物体检测识别	情绪理解	意图理解	事件检测	知识、记忆	新任务学习
沟通助手	多角度，正确率接近100%，知道何时出错		正确率接近100%，知道何时出错，个性化情绪表达	正确率接近100%，知道何时出错	正确率接近100%，知道何时出错	需要	
起居监控	多角度，正确率接近100%，知道何时出错	新物体类别/实例，正确率接近100%，知道何时出错	正确率接近100%，知道何时出错，个性化情绪表达			需要	
紧急救助	多角度，正确率接近100%，知道何时出错	新物体类别/实例，正确率接近100%，知道何时出错	正确率接近100%，知道何时出错，个性化情绪表达			需要	
智能提醒	多角度，正确率接近100%，知道何时出错	新物体类别/实例，正确率接近100%，知道何时出错				需要	
智能导航及和谐空间共享		新物体类别/实例，正确率接近100%，知道何时出错		正确率接近100%，知道何时出错		需要	
基础家务服务	多角度，正确率接近100%，知道何时出错	新物体类别/实例，正确率接近100%，知道何时出错				需要	需要
新任务学习	多角度，正确率接近100%，知道何时出错					需要	需要
情感关怀	多角度，正确率接近100%，知道何时出错		正确率接近100%，知道何时出错，个性化情绪表达			需要	

表 1 养老服务机器人的技术能力需求分析

表 1 中列出的应用在上一节(部分在附录 1 里)进行了介绍，以下对其中所需要的一些能力进行介绍：

- 特定人识别(人员识别或身份识别)：首先机器人对老人和其他家庭成员要能够识别，这个几乎在所有的应用里都要用到。由于人脸识别技术通常只能对接近正脸的人脸进行识别，而机器人如果每次都要移动到能看到正脸的地方并不方便，所以需要多角度进行特定人识别。如果把人认错了，后面的服务也就没法正常进行，所以正确率应该尽量接近 100%。如果机器人没有把握的时候，至少应该知道有可能出错，这个时候可以主动的和用户进行确认，这样即使系统识别人出错的情况下还是可以补救并使得服务正常进行下去。所以总结下来需求就是“多角度，正确率接近 100%，知道何时出错”。
- 物体检测识别：对家里的家具还有日常用品进行识别，这个对于导航(比如如果能识别门，家具等那么就可以实现语义导航)，还有提供特定的服务都是必须的(如智能提醒中帮老人找到特定的物体，基础家务服务中对特定的物体进行抓取)。正确率和可靠性方面也类似特定人识别要求正确率高，否则会严重影响服务的进行，而且希望对新的没有训练过的物体也能够识别。
- 事件检测：这个包括老人日常行为(吃饭，看电视等)的识别，老人跌倒等异常事件的检测。这个能力在起居监控(日常行为的检测)，紧急救助(跌倒，心脏病发作等紧急情况)等应用场景里都是非常重要的。无疑这方面的正确率要求也是很高的。
- 情绪理解：如果要更好的提供用户满意的服务，能够了解用户的情绪显然是有必要的。情绪识别等在紧急情况检测及救助、情感关怀等应用中非常重要，比如识别痛苦的表情等。另外在起居监控中，老人的情绪监控也非常重要，可以判断出老人的情绪是否正常，如果有异常那么需要及时通知家人和医护人员。情绪理解也要求很高的正确率，另外由于每个人的情绪表达会有一些的个性化，所以对于个性化的情绪表达也要能够处理。
- 知识和记忆：这部分主要是存储一些知识和信息，包括一些通用的知识(比如常识知识)，也包含个人相关的一些信息(比如一些和养老看护相关的信息--吃了一些什么，喝了一些什么，甜食是否吃得多，对什么比较感兴趣等等)。通过与人的不断交互，系统对个人的信息可以不断更新，甚至知识库也是可以更新的，这样就能更好的为用户提供更智能的服务。
- 意图理解：意图分析就是预测用户的意图，比如要移动到什么地方，是否希望和机器人交互，要做什么事情等。通过对人的意图分析，就可以更主动的更安全的为人服务。比如和谐空间共享应用场景中，机器人和老人对向而行的时候，如果机器人能够预测老人的意图是往某个方向走，那么机器人可以进行更有效的避让，使得老



人在家里的行动不受到太多干扰。而当一位老人双手端着一碗菜走向冰箱的时候，机器人应该要能预测出来老人是要把这碗菜放到冰箱里面，那么它可以帮助老人提前打开冰箱的门。

- 新任务学习：这里的新任务学习指的是向用户学习做一件新的任务，比如浇花，开关家庭音响系统，一般用户演示和讲解一遍，机器人要能够完全掌握。以开关家庭音响系统为例，包括了多个步骤，开总的电源开关(如果有的话)，并打开 CD 机，功放等的开关，调到合适音量，在 CD 机放入 CD，功放这边也要调到相应的输入(比如 CD 输入)。这个目前看来还是更前沿的功能，比目前的示教学习要更复杂一些，因为牵涉了多个步骤，但如果机器人能具备这样的学习能力，那么将会使得机器人的服务更容易被扩展，所以是很值得进一步探索的一个方向。

其中前面的一些功能是基本的必须的，而后面的一些能力相对来说是更前沿的功能，所以我们的研究探索也是重点放在前面的基本功能，但后面的功能也会开始进行一些必要的探索。

近年来人工智能随着深度学习技术的广泛应用有了巨大的进步，那么实现上面提到的一些能力是否就很容易了呢？实际上并非如此简单，技术上的挑战还是很多的，这个可能不少在机器人领域做过真正产品的业界同仁也有同感。我们分析发现，主要有以下几个挑战：

- 不确定（鲁棒性不好）：目前的不少人工智能技术有时出错，而且不知道自己出错了。

目前的人工智能技术虽然有了长足的进步，但不少和服务机器人相关的技术仍然是不完美的，无论如何改进系统仍然会有一些错误。而且更为棘手的是，系统往往无法知道自己犯错了，常会犯自己觉察不到的错误，这些都使得服务质量很难保证。这个其实也是不少人工智能落地到实际应用首先碰到的一个困难点。可以看到上面的分析中对不少识别能力的正确率要求也是很高的，而且要准确的知道何时出错，所以这方面的必须问题必须有好的解决方案才能使得上面提到的一些先进的应用得以可靠的实现。

- 不认识/知道：机器人不认识新的物体或行为，也无法处理没有预先编程过的新的任务。

目前比较成熟的方法都是基于监督学习的方法，这个一般都是预先通过大量标注数据训练好的，到了应用之后一般是没有办法再改进的，除非在后台再通过新的标注过的训练样本进行重新训练。而应用中又往往会出现新的人，物体或事件等需要识别，这就会出现机器人不认识，不知道的问题。所以机器人的功能在到了用户家庭之后应该是可扩展的，机器人要能在用户的要求和指导下学习新的功能和任务的能力。加州大学洛杉矶分校的朱松纯教授也在文章中指出，服务机器人领域的应用和过去的计算机视觉应用不同，更多的是“大（数量）任务，小数据”(Big task, small data)，尤其因为个性化服务的需求，机器人学习新的能力是必不可少的功能，而且这个学习可能是要基于小数据的。

- 缺知识：新的常识，个人相关的知识

知识和记忆我们看到是不少服务机器人应用必需的一个基本能力，这个属于人工智能中的符号方法(Symbolic method)分支。过去机器人的软件一般也会用一些背景知识，这些就是基于符号方法的。人工智能历史上一度非常流行的专家系统就是基于符号方法的，而专家系统过去没有成功的一个主要问题就是可扩展性比较差，往往会发现缺少知识，从而难以处理实际应用中不断变化的新情况。在服务机器人系统中也会碰到类似问题，那么如何使得系统的知识可以动态的去更新将会非常重要。另外过去统计学习方法和符号方法在机器人领域也没有很紧密的结合起来，我们认为两个方法在服务机器人上面如果能更好的结合的话，将会大大提高系统的智能。

以上这几点我们认为是服务机器人领域人工智能技术应用的一些关键问题，如果能够解决这几个方面的不足，那么服务机器人的能力将会有有一个长足的进步，从而真正实现一些上面提到的更有价值的应用，真正的深入的去满足广大用户在老人看护等方面的刚需。

## 自适应机器人交互

我们提出的自适应机器人交互正是我们针对上面需求提出来的一个新的研究方向，目标是使机器人能够自主的进行持续学习，并且能根据需要向用户发起交互进行学习。

针对上节提到的目前人工智能的一些不足，我们认为自适应机器人交互应该主要从下面一些方面来进行改进：

- 针对不确定，出错而不自知的情况，通过下面两个方向来改进

- 通过观察和交互来更新模型从而不断提高鲁棒性。服务机器人应用的一个可以利用的特点就是机器人会反复观察到一个需要识别的人，物体或行为事件等，这就提供了一个有效的收集样本和标注样本的机会，利用这个特点就可以获取更多标注数据来改进已有的识别模型。另外一个重要的特点是服务机器人可以与人主动交互而不会让人感觉奇怪，不像其他设备如果主动与人交互有时会让人感到奇怪甚至害怕，而人对服务机器人是有这个主动交互的心理准备的。那么可以利用这一点来主动与人交互在识别不确定的情况下获得用户的确认，然后这些信息可以作为标注过的样本加入训练集进一步改进已有识别模型。

这里对人，物体等的多模态观察和识别也是非常重要的，比如在多模态人员识别中，除了人脸识别可以事先训练好(再辅以少量用户的注册照片)，其他模态往往需要不少样本(例如基于体态特征来识别别人需要注册大量的不同角度，不同姿势，不同衣服还有不同光照条件的照片才能达到较高的识别率。声纹识别也一般需要十分钟以上的注册数据)，这些很难通过短暂的用户注册过程获得。所以通过几种模态的配合来自主的获得更多数据就非常重要。

- 通过准确的置信度估计来触发主动交互确认。虽然机器人主动交互对于用户是可以接受的，但频度还是要控制，否则会造成对用户的过多干扰。我们考虑可以通过更准确的置信度估计来更准确地判断识别的可靠性，只有不确定性到一定阈值的时候才触发主动交互。过去的机器学习方法对置信度的估计都是直接利用分类器本身的置信度输出，这个往往是不准的，比如置信度高的时候也可能出错。所以这方面的实现是一个关键点，需要针对不同的识别算法来实现相应的更准确的置信度估计算法。另外在可能出错的时候通过准确的置信度估计，可以在置信度不高时主动和人交互来进行确认，防止基于错误信息的服务。
- 针对不认识的问题，我们觉得可以通过少量样本的初步训练和后续的交互来解决这个问题。这方面针对少量样本可以考虑采用更传统的机器学习方法和深度学习方法结合起来，这样把两方面的优势都能充分利用。而后续的交互则通过与用户的自然交互来解决数据标注的问题。
- 针对知识不足的问题我们主要通过知识图谱的自动更新来解决，主要从下面几个方面来考虑：



- 除了传统的通用知识图谱(如 ConceptNet<sup>[3]</sup>等), 我们认为也需要增加个人知识图谱, 这个可以用于存储用户个人相关的一些个性化知识和信息, 以便后面更好的为用户服务。
- 所有的知识图谱应该是可以进行更新的。前面提到传统的专家系统其知识库一般都是固定的, 最近的一些语音助手系统知识库也需要一个很大的团队来维护更新, 而服务机器人进入家庭后, 面对的用户千差万别, 很难用这些方式去更新知识了, 所以必须要有能够和用户配合起来自动更新知识的方法。
- 知识图谱更新可以通过符号方法和统计方法结合。符号学习也就是基于语义和逻辑来进行的学习, 一般语义信息(事实, 规则)的学习可以是直接从用户(当然也可以从文本里面)那儿学来, 也就是直接把用户说的话翻译为内部语义表达形式来增加机器人的知识, 更进一步还可以从很多事实中通过统计分析归纳出更高层的语义规则, 比如规则归纳学习(rule induction learning)。除了向人直接学习, 另外系统的知识也可以动态的通过机器人观察人和环境来获得。比如“茶杯放在桌子上”这条知识既可以通过观察得到, 也可以是人直接告诉机器人的, 而且机器人可以把这两种方法得到的知识进行互相验证, 以获得更深入的理解, 比如有时用户记错了, 那么机器人可以通过实际观察验证后告诉用户。从这一点而言, 把统计学习方法和符号学习方法结合起来也能够使得服务机器人更加智能, 而且可以弥补传统自然语言处理仅仅在符号层处理, 而没有和现实物理世界对应, 从而语义理解并不深入的缺点。

图 1 是自适应机器人交互软件库的系统框图。可以看到, 下面主要是一些通用的自适应学习的方法, 如多视图学习<sup>[4]</sup>, One-shot/Zero-shot Learning<sup>[5,6]</sup>等, 而上层则是针对机器人交互的一些能力的自适应学习库, 其中人的理解部分包含自适应认人(Adaptive Person Identification), 自适应情绪识别(Adaptive Emotion Recognition), 自适应意图理解(Adaptive Human Intention Understanding), 环境的理解部分包含自适应物体识别(Adaptive Object Detection), 自适应场景分割(Adaptive Scene Segmentation), 而中间是符号方法的部分, 包括通用知识图谱和个人知识图谱, 通用知识图谱是机器人服务相关的一些通用知识(如 Ontology, Common Sense 等), 个人知识图谱则存储和特定人相关的一些信息。推理/规划/预测功能则是基于这些语义信息进行进一步处理, 然后产生相应的行

动和交互，最后从用户还有环境中系统可以得到一些反馈，作为进一步处理的依据。为了对系统中的能力提供训练和测试数据，我们除了采用真实的数据，也在和一些大学合作建立一个高度仿真的机器人交互模拟器来产生更多的数据，方便系统算法的训练和测试。

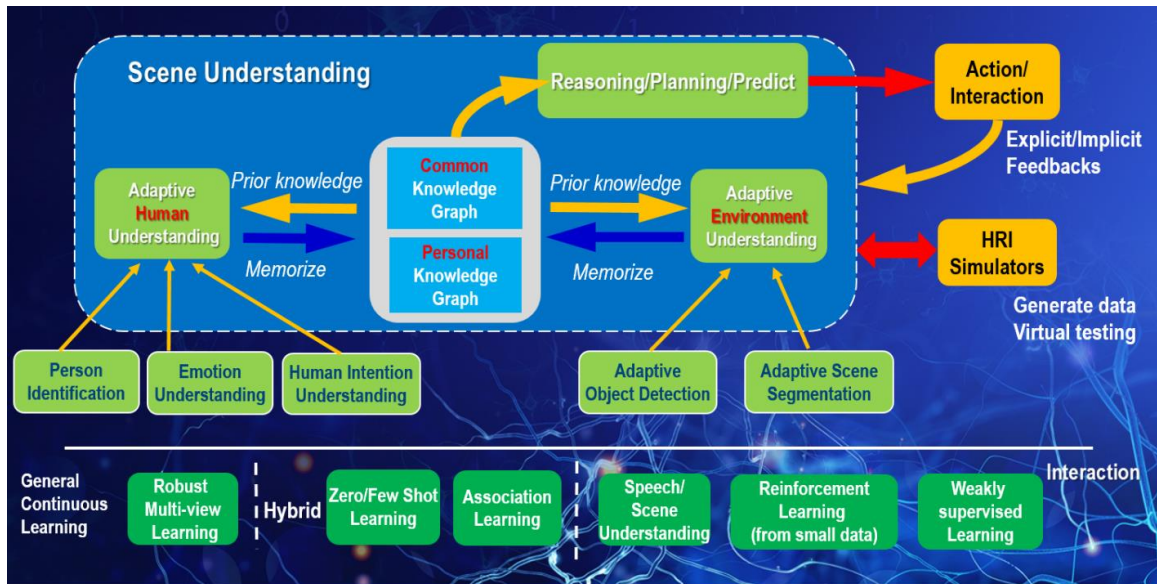


图1 自适应机器人交互软件库的系统框图

下面对几个具有代表性的能力做更为详细的介绍。

- 自适应人员识别：人员识别要求能对人员进行多角度识别。我们采用多模态融合，并且通过对人的持续观察来更新模型。利用人脸进行身份识别在人距离较远、人脸角度较大、人脸表情较大等情况下，准确性较差。在这些情况下，可以利用人体特征进行身份识别。基于人体特征的人员识别功能可以很好的弥补人脸识别的不足，从而可以实现对人员的多角度识别。值得一提的是，我们的模型是结合了深度学习和传统学习方法的，这样使得在小数据的情况下也能对模型进行持续的更新。另外我们对每个模态单独实现更准确的置信度估计，这样可以在不确定的时候通过主动发起交互来进行确认。确认的结果可以产生更多的数据对模型进行持续更新。

这样的—个自适应学习的系统可以在部署到家庭后，从刚开始只能通过人脸识别出用户，逐步过渡到能够从各个角度识别出用户。具体自适应的时间根据用户和机器人交互的频度会不同，我们希望在比较合理的时间内(几天到一周内)就能达到比较稳定的较高的识别率(接近 100%)，另外置信度的估计也尽量比较准确。而



且这个过程还可以长期不断进行，以适应用户身体的长期变化（如变胖变瘦，变老等）。

- 自适应物体检测：物体检测在不少应用模式中都是非常重要的。自适应物体检测可以学习新的物体类别或实例，通过对物体的持续观察来更新模型。比如用户可以把物体放在手上或桌子上进行展示，然后机器人就根据得到的少量样本建立一个初步的模型。以后机器人再看到该物体的时候会通过主动的观察来收集更多的样本。类似的这里的模型也结合了深度学习和传统学习方法，这样在小数据下也能更新模型。置信度的准确估计也是一个重要的部分，这样在不确定的时候通过主动交互来进行确认。
- 个人知识图谱：在和人的长期交互过程中，机器人不仅需要存储交互记录（包括多种方式，参见图 4），并且需要基于交互历史为下一次交互进行合理的推断，建立交互上下文。这部分功能无法在机器人出厂之前针对用户进行个性化定制，这使得自主学习算法在机器人领域成了不可或缺模块。以下几点是机器人需要具备的能力：

- 记录历史

历史包括但不限于对话，视频，音频，手势，表情，动作等多种模态的信息，并在时域对齐。在记录的过程中面临的挑战涉及到多个领域，例如信息的编码，一致性问题，信息安全，信息之间相关性，因果关系等。

- 从历史中观察事件和行为模式

收集历史信息完毕，机器人需要在此基础上进行归纳和判断，为环境建立长期上下文。例如，水果刀通常在厨房中，老人起床时间通常是早晨六点，出门散步要提醒她带上帽子。

- 从已知推断已知

机器人在日常观察中注意到饮水机在厨房，如果收到“端一杯水来”的命令，应该判断出需要去厨房取水。

- 从已知推断未知

同上，机器人如果在日常观察中注意到苹果一般在餐桌上，当收到“帮忙找个水果来”，或者“帮忙拿个梨子来”的命令，即使历史记录没有“水果”或者“梨子”，也可以推断出来有可能在餐桌上。

- 用户断言

在机器人的知识库中，不仅仅包含机器人在和人，环境交互的过程中记录的历史以及基于历史的推断，也包含人类直接输入的断言。在接受断言的过程中，机器人应该确保知识的一致性，以及对断言的可靠性评估。

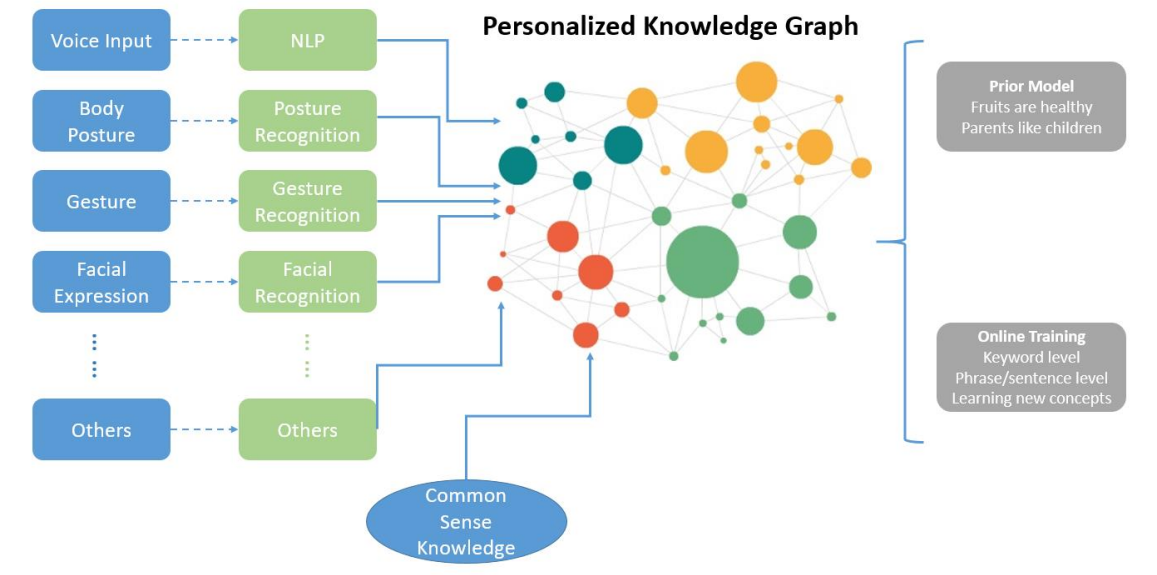


图 4 个人知识图谱

自适应机器人交互软件库 v1.0

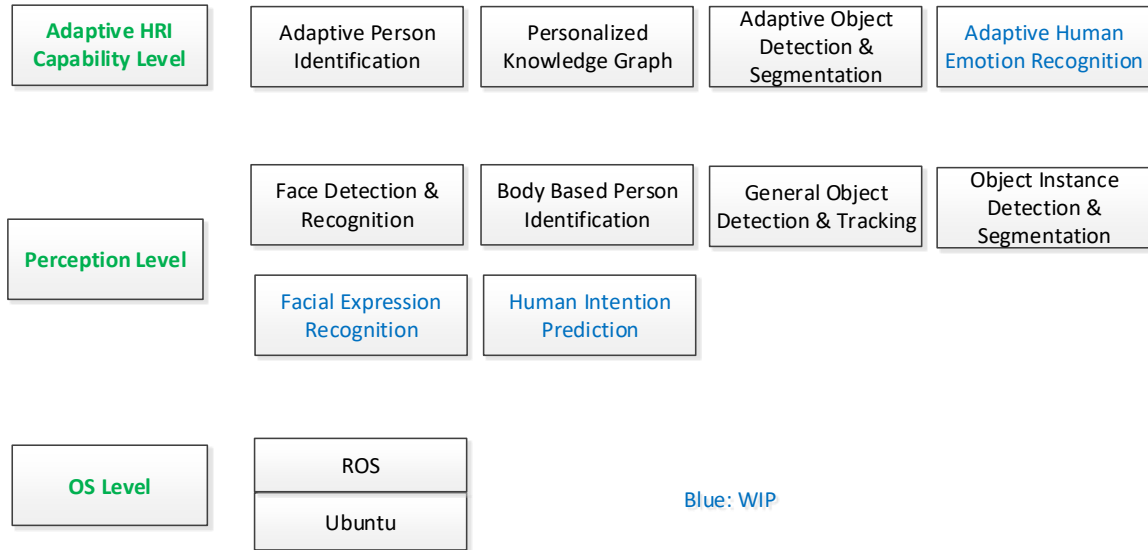
自适应机器人交互软件库(Adaptive HRI Library)是由英特尔中国研究院研究开发，专门针对服务型机器人的开发工具。机器人与人交互过程中，聪明的机器人可以自主学习用户、物体实例和事件的个性化特征，针对不同情况采用不同的处理策略。这个软件库的目的是为服务型机器人提供自适应能力，自主学习适应不同的情况。首次推出的是 Adaptive HRI library 1.0 版本，其基本框架分为 3 层：

- 操作系统层(OS Level)  
Adaptive HRI Library 1.0 是基于目前应用最广泛的机器人软件系统框架 ROS，方便开发者直接使用。
- 感知层(Perception Level)  
感知层提供了机器人感知的多种功能，包括人脸检测和基于人脸的身份识别 ( Face Detection & Recognition )、基于人体特征的身份识别 ( Body Based Person Identification )、通用物体检测和跟踪(General Object Detection & Tracking)、物体实例的检测和分割(Object Instance Detection & Segmentation)等等。



· 自适应人机交互功能层 ( Adaptive HRI Capability Level )

在感知层的基础上是自适应人机交互功能层。基于机器人的感知功能，自适应人机交互功能层提供了机器人自主学习功能，包括自适应的人员识别功能 ( Adaptive Person Identification )、个性化的知识图谱 ( Personalized Knowledge Graph )、自适应的物体检测和分割 ( Adaptive Object Detection & Segmentation ) 和自适应情绪识别 ( Adaptive Human Emotion Recognition )。



Adaptive HRI Library 具有高效灵活的计算特性，支持包括 CPU, GPU , FPGA, Movidius Neural Compute Stick 等多种加速平台。

## 总结

以上我们对个人服务机器人的美好未来尤其是养老服务机器人的先进应用和能力进行了深入的分析，并重点对其中的关键技术领域自适应机器人交互进行了介绍。我们在下面两三年会在这个方向继续进行更深入的研究，为有志于在服务机器人领域开发领先产品的生态系统伙伴提供强有力的技术支持，我们相信在广大生态系统伙伴们的共同努力下，我们一定能携手实现在未来五年内个人服务机器人进入家庭提供先进服务的愿景。



## 参考文献

- [1] 中国财经报网(来源 参考消息): 2050 年 1.6 个劳动力赡养 1 个老人  
[http://www.cfen.com.cn/zyxw/yw/yw\\_7289/201611/t20161121\\_2462618.html](http://www.cfen.com.cn/zyxw/yw/yw_7289/201611/t20161121_2462618.html)
- [2] 新浪网(来源 每日经济新闻):中国 60 岁以上老人已有 2.3 亿 养老服务总体如何?  
<http://news.sina.com.cn/o/2018-01-02/doc-ifyqcsft9142210.shtml>
- [2] ConceptNet, <http://conceptnet.io/>
- [4] Z.-H. Zhou. [When semi-supervised learning meets ensemble learning](#). In: Proceedings of the 8th International Workshop on Multiple Classifier Systems (MCS'09), Reykjavik, Iceland, LNCS 5519, 2009, pp.529-538. [\[slides\]](#) Invited plenary talk at MCS'09
- [5] L. Fei-Fei, R. Fergus and P. Perona, "One-Shot learning of object categories". *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol28(4), 594 - 611, 2006.[PDF](#)
- [6] Xian, Y., Schiele, B., Akata, Z., Campus, S. I., & Machine, A. (2017). Zero-Shot Learning - The Good, the Bad and the Ugly. In CVPR 2017.

## 附录 1

### 家庭养老服务机器人典型应用模式

- 日常沟通助手：手机对于年轻人来说是一个非常高效的沟通工具，语音视频方式都可以，但对不少老人而言，使用手机还有上面的应用还是挺困难的一件事，尤其是年纪大一点之后由于眼睛看不清、手指不灵活等原因更加困难。如果服务机器人能够根据老人的需求（比如通过语音发出命令）来和家人进行各种方式的沟通，那么这个应该是一个非常实用的功能。其中远程呈现(Telepresence) 2.0 可以让老人或家人获得远程浸入式的沟通体验（可以看屏幕或者用虚拟现实头盔），让老人远程也能感觉到浓浓的亲情。具体说来，家人可以通过控制老人家中的机器人来和老人进行沟通(也可以老人控制子女家中的机器人来和家人沟通)，和现有的较常见的远程呈现 1.0 需要用户全程进行移动控制不同的是，远程呈现 2.0 是可以发出更高级的命令的，比如“跟住我的父亲，到电视前面去”，这样用户的操作就大大简化了。

- 智能导航及和谐空间共享：机器人为了完成任务是需要在家里移动的，那么首先需要更加智能的定位导航，目前通常的 SLAM 难以解决场景动态变化等问题，语义导航则可以根据环境中物体的位置(这就是语义信息)来定位导航。另外因为机器人和老人共处一个家庭空间，老人又比较容易受干扰而发生跌倒等危险情况，所以机器人应该非常注意，要与老人进行和谐空间共享。不影响老人的日常行动，也体现良好的社交礼仪(如礼貌的让路等)。
- 日常起居监控：对老人的日常起居情况进行监控(比如什么时候起床，有没有出门，每个房间呆多久的时间，上厕所了没有，在卫生间呆的时间是否过久，起床的时间是否太晚等等)，必要时进行确认，并把相应的情况尤其是异常及时通知给家人或看护人员。比如老人在卧室/浴室呆的太久，机器人可以询问一下，确认没有问题。老人一天没出门，也可以问一下是否有什么不舒服的。 这些信息对于子女而言能够让他们能随时了解父母的日常起居情况，是非常有价值的一个应用。
- 紧急情况检测及救助：从多种模态如视觉，声音来检测特别需要紧急处理的异常情况(比如跌倒、痛苦的表情、很大的声音等等)，并及时提供救助服务，比如通知家人，帮助拿救心药等。紧急情况检测及救助一直是老人看护方面的一个刚需。
- 智能提醒：目前手机上都有定时提醒功能，比如到了吃药的时间提醒老人吃药，不过生活中有不少情况简单的定时提醒并不够，而是要根据目前的情况来进行相应的提醒，比如老人出门的时候提醒带上一些出门必备的物品(手机，钥匙，老人防丢器等)，天冷的时候可以提醒带上帽子、围巾，还有不少老人开了火之后可能就会忘了而直接出门，那么在這些情况发生的时候机器人应该能检测出来并及时进行提醒。即使提醒老人吃药也可以更加灵活，比如老人恰好在卫生间的话可能就听不到提醒，那么机器人可以等老人出来了再进行提醒。如果老人忘记了东西放在哪儿，机器人还可以通过平时的观察去帮老人记住物品放置的位置，并可以根据老人需要去取来交给老人。
- 健康生活助手：观察老人的饮食情况(吃了一些什么，喝了一些什么，有没有出门，有没有锻炼，有没有和朋友熟人聊天 ---出门回来可以问一下，记录下来)，提出改进建议(比如您这几天吃的甜食太多了，这样对身体不好，要管住嘴呀。)
- 基础家务服务：家务服务比如做饭，洗衣等一直是大家比较容易想到的服务，不过从实现成本来看(力量大的机械臂还是非常贵的)，这些可能短期还无法低成本实

现，但我们认为带有操控功能的机器人还是可以提供一些基本的家务服务的(一般需要的力量不大)，比如拉开百叶窗/窗帘，取东西(杯子/药品/眼镜/地上捡一下针)，把东西扔到垃圾桶，用电饭煲煮饭(包括取米，淘米，按键等动作)，甚至做简单的食物(比如披萨，面条)等。尤其是老人行动不便的时候，这些还是比较有用的。

- 新任务学习：很多家务服务需要一定的个性化，很难预先编程，所以如果机器人能够向用户学习做一个新任务是很有必要的。新任务学习的例子不少，比如学习花鸟鱼虫照看，学习开音响系统等。比如给花浇水就可以通过人的演示，学到操作步骤，这样机器人能够给不同的花(喜湿喜干)采用不同的浇水周期和方式。
- 情感关怀：比如机器人和老人聊天，并能在聊天中应用以前聊到的一些信息，如老人的喜好等，相当于机器人是有记忆的。机器人还可以通过分享老人的一些情况(比如拍摄一张老人做的美食发到家人的微信群里)，加强老人和家人之间的沟通，作为他们互相表达关爱的一个桥梁。



## 致谢

我们感谢在白皮书写作过程中给了我们帮助的不少同事和朋友，特别感谢英特尔中国研究院宋继强院长给予我们的一贯的支持和很多有益的指导意见，英特尔中国研究院机器人交互实验室和相关的机器人项目的各位同事，还有在应用和技术需求调研过程中很多热情参与讨论并提出宝贵意见的同仁和朋友们。

## 作者

英特尔中国研究院首席研究员 张益民

英特尔中国研究院高级研究员 任海兵

英特尔中国研究院高级研究员 杨华

本白皮书仅供参考之用。本文件以“概不保证”方式提供，英特尔不做任何形式的保证，包括对适销性、不侵权性，以及适用于特定用途的担保，或任何由建议、规范或范例所产生的任何其它担保。英特尔不承担因使用本规范相关信息所产生的任何责任，包括对侵犯任何知识产权的责任。本文不代表英特尔公司或其它机构向任何人明确或隐含地授予任何知识产权。

英特尔、Intel 标识、Intel vPro、英特尔博锐、Intel Core、英特尔酷睿、英特尔迅驰和 Xeon、至强是英特尔公司在美国和其他国家（地区）的商标。

\* 文中涉及的其它名称及商标属于各自所有者资产。

版权所有 © 2018 英特尔公司。保留所有权利。