

Human-Machine Perception and Assistive Technology

人机感知和辅助技术研究

Zhigang Zhu

朱志刚

The City College, City University of New York

160 Convent Ave, New York, NY 10031, USA

纽约城市大学城市学院计算机系

Email: zhu@cs.ccny.cuny.edu

Abstract

Human vision is truly delicate, and the human brain is more amazing. But if either of them has problems, daily lives are very challenging. Assistive living requires emerging technology, and integrative research needs collaboration. For this end, funding support is important, and user engagement is even more critical. Further, the goals of research at higher education are not only to inspire new ideas, but also to train the next generation workforce, to do good to society. Under the support of an NSF program Emerging Frontier in Research & Innovation (EFRI), in the topic area of Man, Machine and Motor Control (M3C), human and machine vision research has been carried out at the City College of New York. This article describes some of the results and findings. It has been found that multimodal perception is the key for assistive technology, and integration of mobile and cloud computing is the backbone. Both virtual reality and

gaming can be more positive and beneficial, and if we are not constrained by conventional wisdom, substitute perception can be developed and well accepted. One of our objectives is for location-based services go indoors, with multimodal sensing and deep learning techniques. It also becomes evident that public transportation needs to be improved, to be both smart and accessible. Together with technology development and application-driven implementations, the hope is that the mystery of the human brain can be further explored as well.

引言

辅助技术（Assistive Technology），顾名思义，就是用各种技术来辅助那些需要帮助的人，特别是在感知或交流有挑战的朋友，如视障^[1]，自闭^[2]。本文着重在感知辅助技术上。人机感知，就是要结合人的感知和计算特长和机器（包括计算机、机器人、传感器等）的特长，更好地让机器服务于人类。本文以一个美国自然科学基

金（NSF）资助的跨学科人机辅助视觉研究的项目为中心，介绍一下纽约城市学院（CCNY）视觉计算实验室近几年人机视觉和辅助技术的研究。这个 NSF 资助的研究创造前沿（EFRI）项目，是两所院校五个实验室的联合：城市学院的视觉计算实验室、媒体实验室、和脑科学实验室 [3, 4, 5]，佐治亚理工的机器视觉实验室和生理实验室 [6, 7]。研究的内容涵盖网膜和脑皮层 TMS，大脑分析，舌头读图，标识识别，深度触觉，网膜移植，运动分析等，主要应用对象为视障人士（图 1）。本文只介绍和城市学院视觉视觉实验室密切相关的几

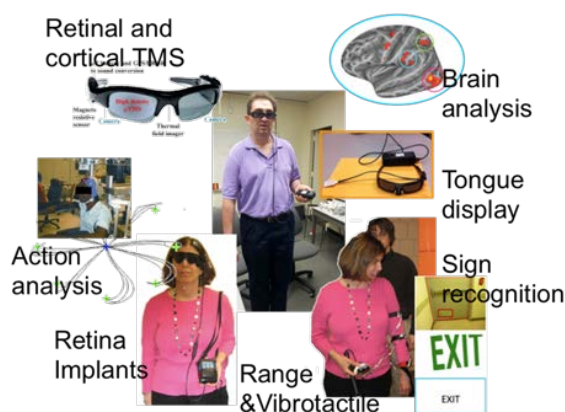


图 1 CCNY-Georgia Tech 人机感知及视觉辅助技术研究^[8]。涉及得技术包括（自左上顺时针）：网膜和脑皮层 TMS，脑分析，舌头读图，标识识别，深度触觉，网膜移植，动作分析。

个方面，并试图从辅助视觉扩展道涵盖更广的感知辅助技术的研究和教育^[8-11]，在应用方面服务于视力障碍、听力障碍，运动障碍、失忆症、老年人、自闭症等，在科学研究方面探索人脑的奥秘，包括感知、表情和社会交流。所报道的研究结果尚离期望有不少差距，但求本文所描述的科研

探索能启发更多的研究思路并导致更成熟的研究成果。

下面，本文先以一个人机视觉研究的学术研讨会为主，讨论视觉辅助技术研究和开发方面的最新进展点滴和一点思考。然后，再介绍视觉实验室关于人机感知和辅助技术的五个主要专题：

- (1) 视觉导航技术：包括双目、单目、全相和三维视觉的辅助视障应用。
- (2) 替代感知方法：包括全身目设计，群帮眼尝试，机电辨声初探，舌头“看”图案测试。
- (3) 面部计算研究：包括人脸识别、表情识别、深度学习和人脑感知研究。
- (4) 测试应用平台：包括一个幻境与游戏测试和训练平台，和一个智能无障碍交通平台。
- (5) 教育和培训项目：包括一个暑期的培训活动和一個本科的联合实习课程。

为讨论简洁起见，在概述以上五个方面的专题时，每一个专题配一首小诗表达核心思想，并以图示和参考文献做补充，从而读者一方面无需一下子进入细节，另一方面，若有兴趣深入了解可参阅有关文章。如若词不达意，先请见谅，但希望耐心细读，领会大意，必得些许启发。

跨学科视觉辅助技术研究

根据“2012 年世界卫生组织报告”^[1]，有超过 2.85 亿视障人士，其中 3900 万人失

明。视力受损人士中，约有 65% 是 50 岁以上，而这一年龄组占世界人口的 20% 左右。随着许多国家老年人口的增加，更多的人将面临年龄相关的视力障碍。对多模式和替代感知的研究将对社会的健康和福利产生长期影响，不仅是对视力有挑战者，而且对于经常在危险环境中工作的人员，如消防员，司机和士兵也将如此。由 NSF 资助、本文作者和多位各个领域专家主持的 2013 年 IEEE 视力障碍多模态与替代感知研讨会 (MAP4VIP 2013, 图 2)^[12]，是关于这个话题的第一个 IEEE 研讨会。MAP4VIP 2013 旨在使来自多个学科 (包括计算机视觉，神经科学，多媒体计算，传感器技术和辅助技术应用) 的研究人员和从业者讨论视觉感知、计算智能、神经科学和视觉假体移植的基本问题，以帮助盲人和视障人士和在有视觉挑战的环境中工作的人，并进一步促进人脑感知功能的科学研究。研讨会的主题包括：(1) 帮助视觉上有挑战者的计算机视觉、感知和学习算法，包括电子旅行援助，电子导向辅助，标志检测，触觉显示，脸部识别，道路和障碍物检测；(2) 视觉假肢和人机交互，包括视网膜植入物，视觉皮层刺激，生物模拟和生物系统工程探索；(3) 神经科学/生物力学/心理/社会研究，包括人类感知，神经计算，视觉编码，感觉运动模型，多感觉整合，系统评估，科学技术政策和未来趋势。

研讨会将包括以下内容：主题演讲 (一个侧重人的视觉，另一个侧重视觉假肢)；专案讨论 (包括约 6-7 名来自学术界，行业，政府和社区的专家)；口头报导



图 2 IEEE /NSF 多模式替代感知视障辅助国际研讨会^[12](资助单位：IEEE，美国自然科学基金，微软研究院，纽约城市学院)

（包括特邀演讲和公开征文）；以及看版/演示（年轻/学生研究人员的研究结果）。专案讨论《视觉辅助：需求，挑战和希望》的议题包括：（1）您认为在未来 5 年内可能开发的视障辅助装置的期望的质量是什么？这种设备目前能允许视障人士做些什么？（2）视障人士在日常生活中面临的一些最大挑战和障碍是什么？（3）开发视觉辅助技术当前的一些限制是什么？（4）有哪些注意事项来确定辅助技术设备是否将被采用并被广泛使用（例如成本，尺寸）？（5）我们可以和视障人士可以进行哪些类型的标准化和科学测试，以如何量化视觉装置的性能提高以及各种辅助设备对提升生活质量的影响？

研讨会各行专家发表了自己的见解；详见^[12]。一些有代表性的认识包括：在不熟悉或复杂环境里的视觉导航技术仍有很大的挑战，需要跨领域的研究，包括电脑、人脑并生理诸方面；视觉是一个眼见、脑算后行动的综合过程，计算视觉读标识、寻找日常生活中的物件大有需要；辅助视障的研究需要比研讨会更进一步的具体合作；视觉网膜移植装置需要更高的分辨率，比如在五年内达到 16x16，硬件方面需要有突破，软件方面也需要密切配合；计算机视觉和群帮（Crowdsourcing）技术为基础的智能手机和穿戴装置的应用软件（Apps）可能会大显用场；有可持续性的政府和民间的资助必不可少（如城市学院视觉实验室的研究就是个例子^[8-11]）；研究不仅需要跨学科、并结合各方面的优势，更需要用户的密切参与；虽然辅助技术需求很大，但由于视障用户大都属于低收入的群体，新的技术难以拓展市场，所以可及性

（accessibility）的真正实现需要和提升各级政府以及民众的意识相配合；辅助视觉装置的实用功能的可靠性和使用的容易度是视障用户最关心的两大因素，但系统的性能价格比和人机智能的有效结合也是重要的考虑因素；装置的标准化、科学测试，和大规模的用户试用和反馈非常重要；我们随后的调研也发现，替代视觉的途经很多，包括物理、数字和医学等三方面的途经^[13]。现有的研究和开发是百花齐放，包括由耳、舌^[14]、身、手感知的替代方法，更有机器视觉理解的方法和网膜移植^[15]，但科学测试标准和用户试用的方法和平台仍然缺少^[13]。

另一个人机视觉可以大显用场的领域是帮助自闭患者的社交和学习。统计表明^[2]，在美国 68 个人中，就有一个在自闭谱系障碍（ASD）上，这主要是大脑工作的特异性问题。和视障者类似，他们也有严重的行动和交流上的困难。当然还有其他有挑战者和老人，辅助技术特别是视觉辅助可以提供很大的帮助，包括导航、读脸、找地方等。在这些方面，多种新兴技术都可用上，包括多种传感、定位技术、移动计算、深度学习等。总之，人眼功能甚精细，大脑运作更奥秘；千万精英齐努力，成果频出需继续！

机器视觉技术

辅助视觉的第一个主要方面，就是机器视觉传感器的选择和视觉算法的实现。由于需要人身穿戴或携带，而且视障人士大都是低收入，所以小型化、轻便、省电、便宜是主要的考虑。当然从性能方面，实时和可靠是基本的要求。这两方面往往是

相互矛盾的，所以折衷可行的方案便是研究的重点。视觉辅助可有多个层次、多个方面。从导航来说，有近距离的安全要求（可行通路和避障）和远距离的方向指示（地方和路标识别等）。我们下面的三个技术着重与近、中距离的安全导航，即定位、定向和避障。我们首先介绍一个室内导航研究的各种方案中^[16-18]的一个特别的、经济的方法，就是在智能手机上装载一个简易的 360 度全向镜头（图 3），建立一个室内场景的图像数据库，然后用图像检索的方法定位^[16, 19]。为了解决实时计算速度和大数据量的要求，我们采用了移动计算和云端计算相结合的方案^[16]，同时在建模时融合了其它的技术，包括地图预读，路径规划等^[20]。第二个技术是双目视觉的智能采样方法（图 4），主要是为解决现今低分辨率视觉替代装置有效视觉信息的提供问题（如网膜移植产品 Argus II^[15] 只有 6x10，Brainport 的舌头读图板^[14] 只有

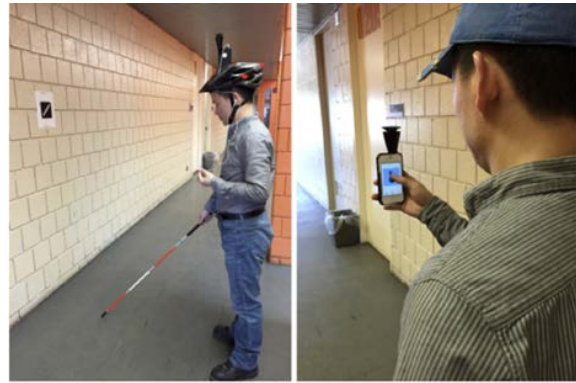


图 3 智能手机全向辅助视觉 (左：头盔系统；右：手持系统)

20x20)。其中核心是针对室内场景纹理少的情况的一个基于彩色图像分割的立体视觉匹配算法^[21, 22]和在其基础上的一个智能采样和感兴趣目标增强的算法^[23, 24]。最后一个辅助视觉方案是用单幅图像进行三维推理的方法，目标是用智能手机的摄像和计算功能，即可进行楼道或街道方向的三维估计和近、中甚至远距离通道（如门口）

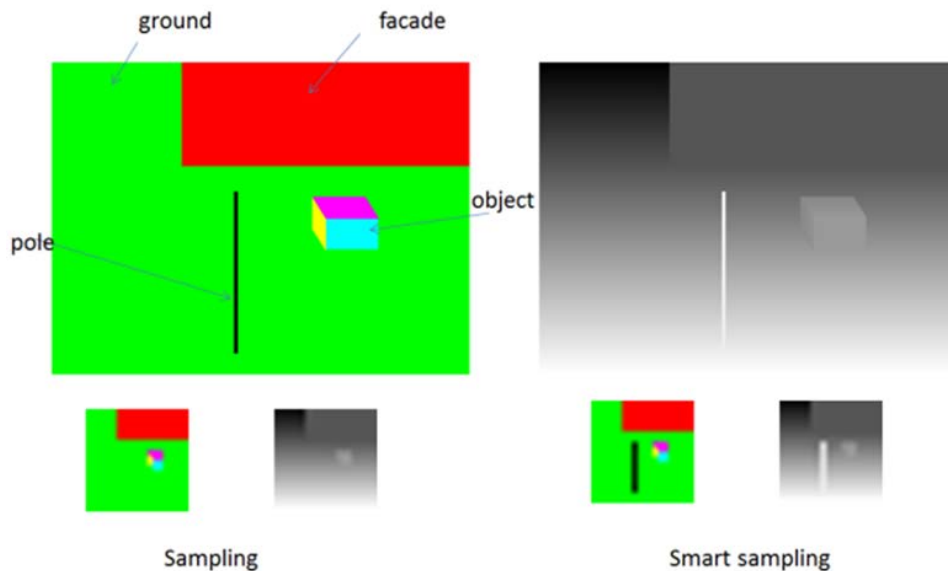


图 4 双目视觉智能采样 (左上：模拟彩色图像，包括地面，楼面，方形物，细杆；右上：深度图像；左下：20x20 普通采样细杆丢失；右下：20x20 智能采样细杆保持)

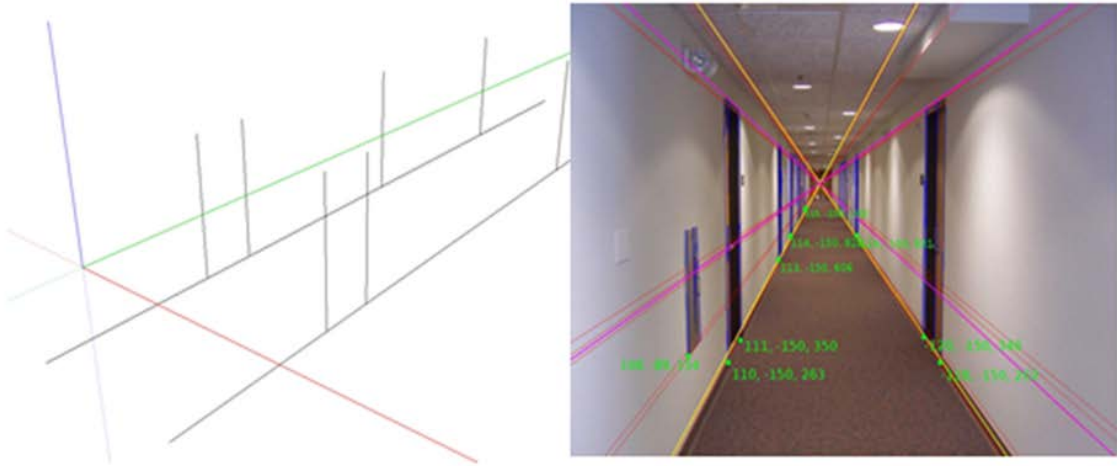


图 5 霍夫变换单目测距^[25]（右：原图上标注走廊和房门的检测结果，并三维定位数据；左：检测结果的三维线画图模型）

的检测和距离估算。我们尝试过两种方法：有效的定位、方向信息，除了可以用网膜移植的方法直接传递给用户的视神经外，用语音反馈恐怕是最容易想到也最有效的方法。但是太频繁占用语音频道会影响视障用户行之有效的听力导航功能，所以其它的替代感知传递方法也曾广为研究。我们尝试或测验了四个方面的替代感知技术：（1）

- 智能手机观全向，图像检索定方位；^[16]
- 移动三维更快捷，建模定位已实时。^[17, 18]
- 室内简洁少图案，立体匹配有难处；
- 图像分为自然块，灵巧采样数据精。^[21-24]
- 智能手机拍照易，计算功能也超强；
- 人眼单目可测距，缘何不让机器做？^[25, 26]
- 单项技术总有限，多种融合更完善；
- 地图预读先规划，智能手机联云端。^[18-20]

替代感知技术

辅助视觉的第二个主要方面，就是如何将采集和处理的视觉信息传递给视障用户，以及探索有无除了机器视觉以外其它

“全身目”的设计和最优优化：尝试用皮肤各个部位来感知震动的强弱为用户提供四周的距离感知信息^[29]，并决定传感器的多少和位置^[30]。为此我们设计了各种模块化的距离-震动感知单元（图 6）^[31]，配以适当的穿戴部件，并用 BLE 低耗蓝牙技术将其与一智能手机连接，从而不仅方便设置参数，而且容易收集数据。（2）“群帮眼”试验系统和测试（图 7）：机器视觉是最佳选择，但是目前小巧低成本的解决方案仍待开发。作为一个替代方案，我们尝试用目前非常流行的群帮（Crowdsourcing）技术，来帮助视障用户实时导航^[32, 33]。（3）肌电图（EMG）辨音：我们也研究了用置

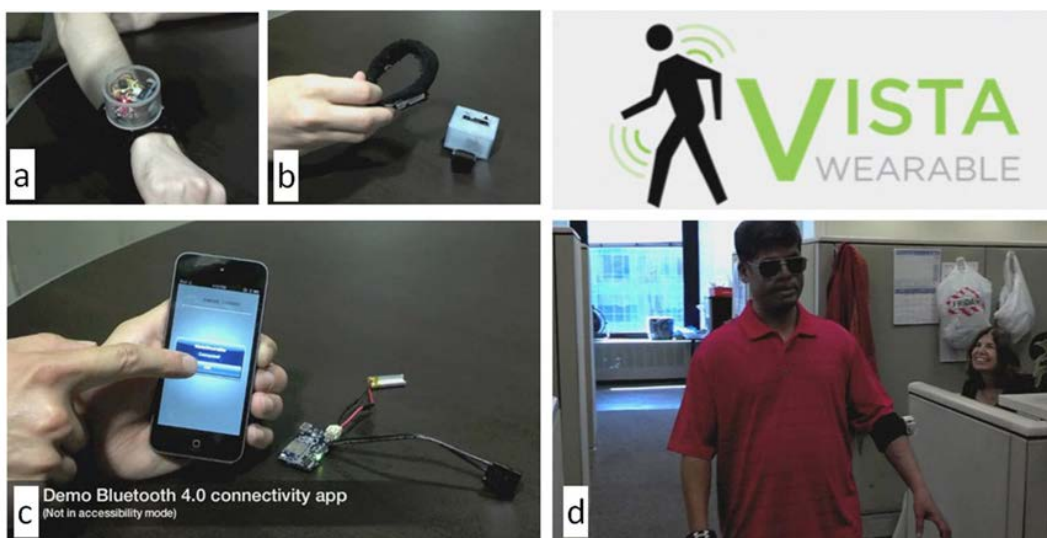


图 6 可穿戴全身目的原型设计^[31]。a. 第一代圆形设计； b. 第二代扁形设计； c. 第三代智能手机蓝牙遥控设计； d. 一用户试戴。

于面部的多个 EMG 电极和深度 Bayesian 网络 (DBN) 来试验可否识别讲话者 (图 8)。从一个 20 余人的实验中, 我们发现, 讲话者只讲一个英文单字, 用 EMG 可以达到和用语音同样的识别率^[34]。这个实验说明 EMG 信号采集了话者相关的信息, 将来可能为视障和其它残障用户提供感知帮助。我们正在进行同时用肌电图 (EMG) 和脑电图 (EEG), 来鉴别实验者的表情变化, 从而更好地认识人脑的工作机理。(4) 舌头“看”的测试。我们也和脑科学和感知专家合作, 测试了 Brainport 的舌头读图^[14], 虽然我们发现用其识别图案仍非常困难, 但也许可以用其实时指示用户行走的方向 (图 9)^[35], 从而可以空出用户的耳朵和手等, 以便其更自然的行走。

- 全身是目有预测, 皮肤能“看”见小说^[27, 28]; 全身沉浸距离场, 振动强弱可表达^[29-31]。

- 肌电信号来辨音, 深度学习做比较; 一字识得讲话者, 效果堪比用语音^[34]。
- 舌头看图初测试, 识别图案待商榷; 或可巧用指方向, 空出手、耳做其它^[35]。
- 群帮已成新时尚, 何不用其助导航? 用户视频传网上, 网友众目来群帮^[32, 33]。

面部识别和深度学习

对于视障和自闭的用户, 除了满足他们基本的导航和指向需要外, 更进一步的机器视觉技术可帮助他们增强与他人交流的能力, 从而提高他们的生活质量。这可以作为辅助技术的第三个主要方面。其中, 人之间一个主要的交流需要是人脸识别和表情的识别。我们曾用 RGBD 传感器, 用彩色和三维信息为盲人提供人脸鉴别的功能 (图 10)^[36]。我们随后更多的研究集中在人脸表情的识别上, 以便为视障和自

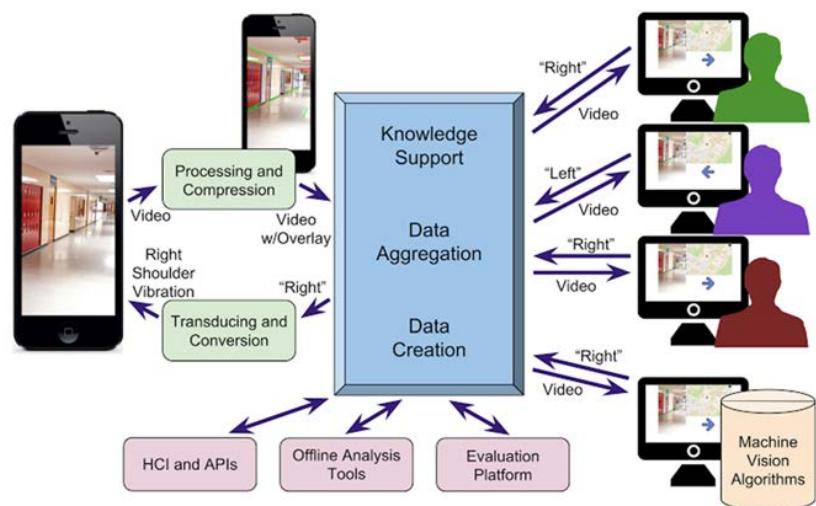


图 7 群帮眼的原理示意图^[32]：视障用户的视频经预处理和压缩传往互联网，网上志愿者和机器算法提供实时行走指示，经处理系统过滤集成后，用声音或触觉等指示用户导航。收集的数据可供人机视觉研究用。

闭用户提供更有效的与他人交流的工具。比如，我们设计了一个智能手机 App，通过一个简单的用表情控制的游戏，训练自闭患者的表情识别和反馈（图 11）^[37]。为了这些应用的有效性，我们进行了系列的研究：用主要的图像搜索引擎收集和清理出了一个有一万多张自然的人脸表情图像的数据库 CIFE^[38]；基于该数据库训练了一个七类表情识别的深度卷积网络（CNN），以此为引擎，设计了一个简洁的表情可控

游戏，在一个月就收集到 100 多个游戏用户的超过一万五千张表情图像，生成了表情更接近日常生活、七个类别的图像数目也更平衡的 GaMo 表情图像库^[39]；图像、声音、视频等多种模式融合、人工设计和

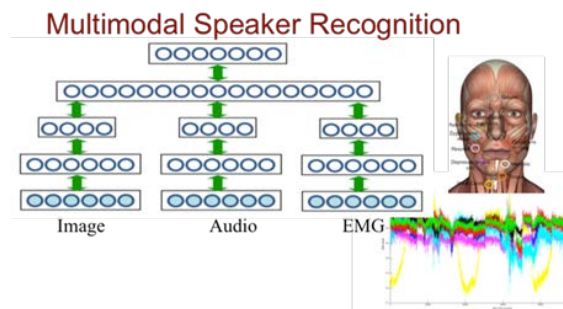


图 8. 多模式说话者识别^[34]：图像、语音和肌电图信号经由深度神经网络集成后，说话者讲一个字即可识别。

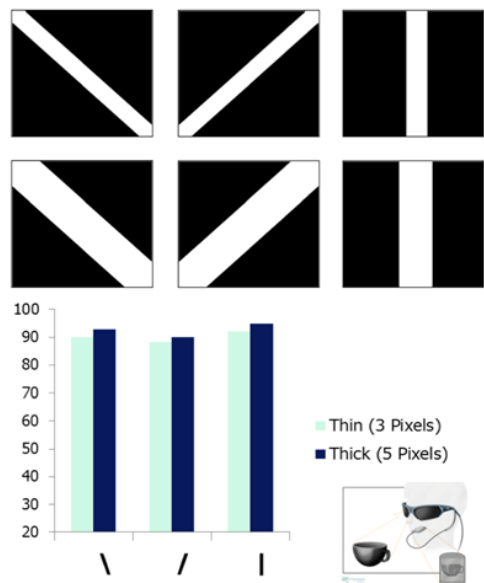


图 9 测试结果表明^[35]，舌头看图较难用于形状识别。但是可以很好的辨别三个方向的线条，因此可用于导航指示。

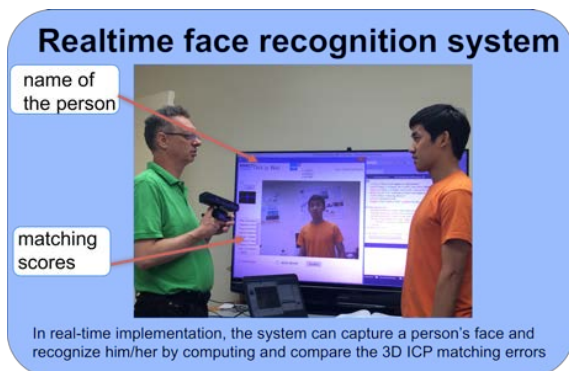


图 10 实时三维人脸识别初步试验^[36]。目标是用于视障用户的交互。

深度学习相结合的特征提取、和多核学习的方法，使得我们设计的算法在 2015 年的 ACM 表情识别挑战 (EmotiW) 的 60 个参赛中，名列第五^[40]。为了进一步研究人脸表情的机理，我们研究了基于深度学习的表情单元 (Action Unit) 识别，设计了综合深度卷积学习转移 (Learning Transfer)，多层次兴趣编码 (Attention Coding) 和长短期记忆 (LSTM) 时空集成的模型，使我



图 11 帮助自闭症用户训练表情的智能手机游戏系统界面。

们的表情单元识别性能比目前最好的结果提高了 10% 以上^[41, 42]。

- 喜怒哀乐写脸上，面部识别很重要；无奈视障不能见^[36]，自闭患者亦作难^[37]。
- 深度学习识表情，网上收集众脸谱

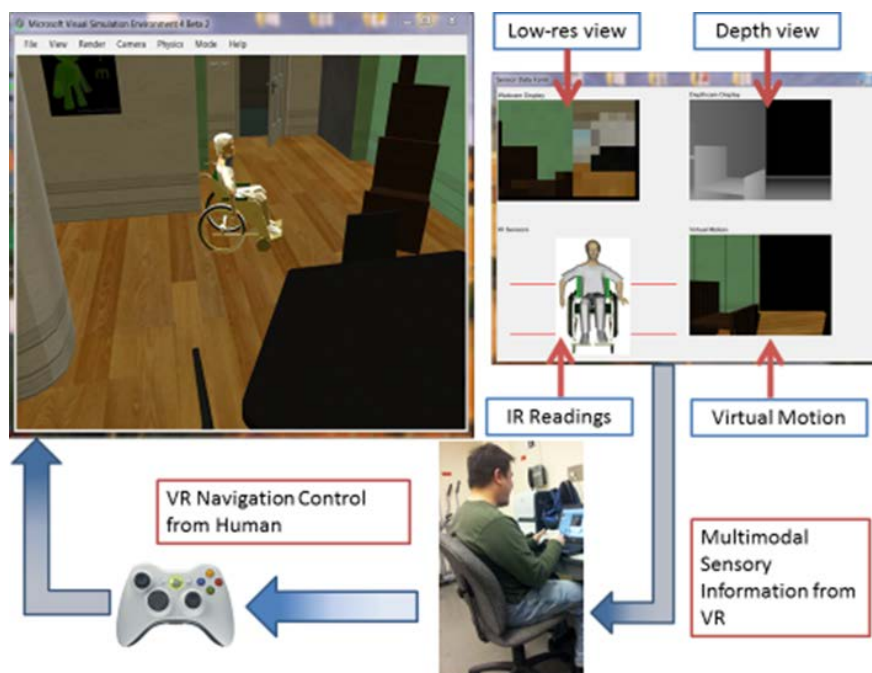


图 12 虚拟现实传感器仿真和导航测试。

[38]; 卷积网络为引擎, 设计游戏采数据 [39]。

- 多种模式来集成, 实战表情列前茅 [40]; 人机智能相结合, 多核学习最给力。
- 表情单元是核心, 学习转移省训练; 兴趣编码显效果 [41], 时空集成更提升 [42]。

测试和应用平台

辅助视觉和其它辅助技术的第四个主要方面是测试和应用平台。我们这里介绍两个平台: (1) GIVE-ME (Gamification In Virtual Environment for Multimodal Evaluation): 这是一个结合游戏化和虚拟现实的多模式测试和训练平台 [43, 44]。多种传感信息可实时仿真, 包括低分辨率图像, 深度图像, 红外单点测距, 运动目标模拟。传感信息经处理后, 转换成视障用户可感知的形式传给用户, 用户进而控制其替身

在虚拟环境中导航 (图 12)。(2) SAT-Hub (Smart and Accessible Transportation Hub): 这是一个由纽约城市大学(CUNY)、若歌大学(Rutgers University)和其它单位合作所提出的智能无障碍交通枢纽平台, 2016 年开始成为全球城市团队挑战 (GCTC) 的一个行动组 (Action Cluster)。关键技术包括: 移动定位, 人群分析, 三维语义模型, 用户定制导引。(图 13) [45-47]。

- 训练模拟在幻境, 灵活有趣无危险; 仿真各类传感器, 研究开发成本低。 [43, 44]。
- 其一模拟舌头看, 其二模拟群帮眼 [35, 33]; 其三模拟全身目, 其乐融融育博士 [31, 44]。
- 交通智能又可及, 城市、若歌双联合; 安全、交通兼服务, 集成一体利全民 [45-47]。
- 数字模型识环境, 人群分析知动态;

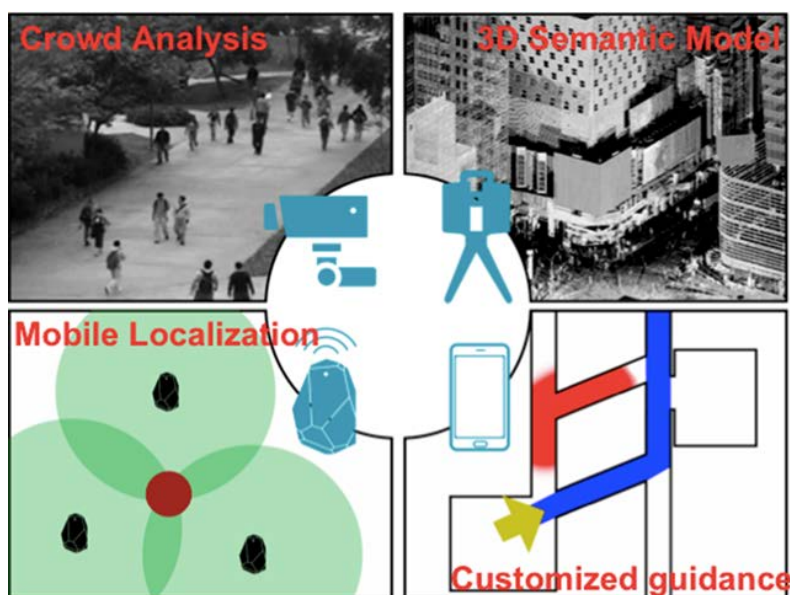


图 13 智能可及交通枢纽关键技术示意图。

多种传感定方位，户内室外均便行。

教育和培训

辅助技术的第五个主要方面是教育和培训。我们不仅培养博士生进行算法和系统的研究和开发，我们还开设了以下两个由 NSF 和工业界资助的教育和培训项目：

(1) 横跨三个专业——计算机科学、计算机工程和电子工程的辅助技术本科实习课（图 14）^[9, 10, 47]。学生项目组不仅屡次获得创业奖，而且研究文章亦频出^[17, 29, 48-53]。

(2) 主要针对少数群体、但招收从社区大学教授、大学生、高中老师到高中生的研究经历和指导（REM）项目，历经四年，训练学员 70 余人（图 15）^[8, 53]。



图 14 作者(右)、Green Jacket 项目组的学生和 NYSID 总裁 Ron Romano (左)在纽约州首府 Albany 的 2016 年度 CREATE 展示会^[51]

- 本科联合实战课，成功运作已六载；电子、计算和工程，培养学生逾百人^[48]。
- 启于技术帮视障，智能生活渐顾及；始于用户真需要，学于实践有动力。
- 研究指导新尝试，少数群体得帮助^[53]；四年培训愈七十，各种层次成果出^[37, 39, 54]。

- 暑期十周密培训，日日小组细研讨；每周大组做报告，来春全国年会秀^[55]。

结束语

人机感知和辅助技术研究，我们有三方面的思考。首先，研究的意义，我们强调科学价值、社会效益及学生培训的结合。我们认为，视觉甚精细，人脑更神奇；其一若有障，生活成问题。科研要联合，辅助需科技；多方来资助，用户亦参与。高校做研究，启发新创意；培养有用才，工作有意义。

其次，我们研究的路线，采取跨学科的联合和观念的突破：创新最前沿，机脑



图 15 2015 年 NSF 研究经历指导（REM）梯队。指导者包括项目负责人（教授）和博士生；学员包括：社区大学教授，本科生，高中老师和高三学生。

并生理；感知加计算，移动联云际。幻境与游戏，皆能更有益；不拘旧观念，感知互代替。导航到户内，深度来学习；交通在改进，智能又可及。

最后，我们在技术的各个点上，也追求有创新：双目加智能，单目可测距；全向有优势，三维更迅捷。设计全身目，尝试群帮眼；肌电能辨声，舌头“看”图案。

机器读人脸，盲者识表情；人脑之奥妙，或许可探知？

致谢

致谢以下资助和合作单位：NSF (Awards (#1137172, #1160046, #1327236, #1416396), DHS (Award #DE-SC0014664), VenturWell (#10087-12), NYSID, Goodwill NY/NJ, Lighthouse Guild, NYSCB, Georgia Tech, Rutgers 等。也感谢多位家人、朋友仔细阅读并提宝贵修改意见，在此不一一提名。本文所述观点只代表作者本人。

参考文献

- [1] W. H. Organization, “Visual Impairment and Blindness. Fact Sheet No. 282. Updated August 2014,” <http://www.who.int/mediacentre/factsheets/fs282/en/>, accessed: 2017-05-20.
- [2] Autism Speaks, Facts about Autism, <https://www.autismspeaks.org/what-autism/facts-about-autism>, accessed: 2017-05-20
- [3] Z. Zhu, The City College Visual Computing Lab (Ccvcl), <http://visionlab.engr.ccny.cuny.edu/ccvcl>
- [4] Y. Tian, Media Lab, The City College of New York. <http://media-lab.ccny.cuny.edu/wordpress/>
- [5] T. Ro, The Ro Lab, The City University of New York. <http://rolab.ws.gc.cuny.edu/tro/>
- [6] K.-M. Lee, Advanced Intelligent Mechatronics Research Laboratory (AIMRL), Georgia Tech, <http://aimrl.gatech.edu/>
- [7] B. Prilutsky, Biomechanics Group, Georgia Tech, <http://www.bioengineering.gatech.edu/people/boris-prilutsky>
- [8] Z. Zhu, K.-M. Lee, B. Prilutsky, T. Ro and Y. Tian, EFRI-M3C: Mobility Skill Acquisition and Learning through Alternative and Multimodal Perception for Visually Impaired People, NSF Award #1137172, EFMA Emerging Frontiers & Multidisciplinary Activities, ENG Directorate For Engineering (Program Manager: Radhakisan S. Baheti), 2011-2018. https://www.nsf.gov/awardsearch/showAward?AWD_ID=1137172
- [9] J. Xiao, Z. Zhu, Senior Design Program on Assistive Technology to Aid Visually Impaired People, National Science Foundation (Award #1160046), 06/01/2012 – 05/31/2017. https://www.nsf.gov/awardsearch/showAward?AWD_ID=1160046
- [10] Z. Zhu, J. Xiao and T. Ro, Course and Program Grant (10087-12): Human and Machine Intelligence - Perception, Computation and Action, VentureWell (formerly NCIIA), 06/01/2013- 08/31/2017.
- [11] E. Molina, Z. Zhu, T. Ro, Wearable 3D Navigation for the Blind and Visually Impaired, NSF SBIR Phase I (Award #1416396), NSF ENG IIP Div Of Industrial Innovation & Partnership, 07/01/2014 – 03/31/2015. https://www.nsf.gov/awardsearch/showAward?AWD_ID=1416396
- [12] Z. Zhu, Z. Zhang, K.-M. Lee, Y. LeCun, Y. Tian, T. Ro and S. Kelly, IEEE/NSF Workshop Multimodal and Alternative Perception for Visually Impaired People (MAP4VIP), NSF CISE (Award 1327236), July 15th, 2013, San Jose, USA. <http://www-cs.engr.ccny.cuny.edu/~zhu/MAP4VIP/>
- [13] W. L. Khoo, Z. Zhu, Multimodal and Alternative Perception for the Visually Impaired: A Survey, Journal of Assistive Technologies, Vol. 10 Iss: 1, pp.11-26, 2016. <http://dx.doi.org/10.1108/JAT-04-2015-0014>
- [14] Brainport Technologies, <http://www.wicab.com/>
- [15] Second Sight, <http://www.secondsight.com/>
- [16] F. Hu, Z. Zhu and J. Zhang, Mobile Panoramic Vision for Assisting the Blind via Indexing and Localization, Second Workshop on Assistive Computer Vision and Robotics, Zurich, Sept 12, 2014
- [17] Q. Chen, M. Khan, C. Tsangouri, C. Yang, B.

- Li, J. Xiao and Z. Zhu, CCNY Smart Cane, IEEE-CYBER 2017-The 7th Annual IEEE Int. Conf. on CYBER Technology in Automation, Control, and Intelligent Systems, July 31 – August 4, 2017, Hawaii, USA
- [18] H. Feng, N. Tsering, H. Tang, and Z. Zhu. Indoor Localization for the Visually Impaired Using a 3D Sensor. *Journal on Technology and Persons with Disabilities*. Vol. 4. Nov. 2016
- [19] F. Hu, H. Tang, A. Tsema, Z. Zhu, "Computer Vision for Assistive Indoor Localization", *Assistive Computer Vision*, M. Leo and G. Farinella (Eds.), the *Computer Vision and Pattern Recognition Series*, Elsevier, to appear in 2017 (invited)
- [20] H. Tang, N. Tsering, F. Hu, and Z. Zhu. Automatic Pre-Journey Indoor Map Generation Using AutoCAD Floor Plan. *Journal on Technology and Persons with Disabilities*. Vol. 4. Oct. 2016
- [21] H. Tang, Z. Zhu, A Segmentation-based Stereovision Approach for Assisting Visually Impaired People, the 13th International Conference on Computers Helping People with Special Needs (ICCHP), July 11-13, 2012, Linz, Austria.
- [22] H. Tang and Z. Zhu, Content-Based 3D Mosaics for Representing Videos of Dynamic Urban Scenes, *IEEE Transactions on Circuits and Systems for Video Technology*, 22(2), 295-308, 2012. DOI: 10.1109/TCSVT.2011.2178729
- [23] H. Tang, T. Ro, Z. Zhu, Smart Sampling And Transducing 3d Scenes For The Visually Impaired, IEEE International Conference on Multimedia and Expo (ICME), July 15 to 19, 2013.
- [24] H. Tang, Z. Zhu, M. Vincent, T. Ro, From RGB-D to Low-Resolution Tactile: Smart Sampling And Early Testing, IEEE Workshop on Multimodal and Alternative Perception for Visually Impaired People (MAP4VIP), in conjunction with IEEE International Conference on Multimedia and Expo (ICME), San Jose, California, USA, July 15 - 19, 2013
- [25] G. Olmschenk, Z. Zhu, 3D Hallway Modeling Using A Single Image. The Fourth IEEE Workshop on Mobile Vision, in conjunction with CVPR2014
- [26] G. Olmschenk, Z. Zhu. Mobile Real-Time Single Image 3D Corridor Reconstruction Using J-Linkage. The 14th IAPR Conference on Machine Vision Applications (MVA 2015), Tokyo, May 18-22, 2015
- [27] John Apostle, Revelation, 4:8. Holy Bible, New International Version, ©2011, Biblica Inc.
- [28] Steven Millhauser, The Wizard of West Orange, Alfred A. Knopf, 2008
- [29] F. Palmer, Z. Zhu, T. Ro, Wearable Range-Vibrotactile Field: Design and Evaluation, the 13th International Conference on Computers Helping People with Special Needs (ICCHP), July 11-13, 2012, Linz, Austria.
- [30] W. L. Khoo, F. Plamer, J. Knapp, Z. Zhu, T. Ro, Designing and Testing Wearable Range-Vibrotactile Devices, *Journal of Assistive Technologies*, Vol. 7 Iss: 2, 2013, pp.102 – 117
- [31] E. Molina, W. L. Khoo, F. Palmer, L. Ai, T. Ro, Z. Zhu, Vista Wearable: Seeing through Whole-Body Touch without Contact, The 12th IEEE International Conference on Ubiquitous Intelligence and Computing (UIC 2015), August 10-14, 2015, Beijing China
- [32] G. Olmschenk, C. Yang, Z. Zhu, H. Tong and W. H. Seiple, Mobile Crowd Assisted Navigation for the Visually Impaired. The 12th IEEE International Conference on Ubiquitous Intelligence and Computing (UIC 2015), August 10-14, 2015, Beijing China
- [33] W. L. Khoo, G. Olmschenk, Z. Zhu and T. Ro, Evaluating Crowd Sourced Navigation for the Visually Impaired in a Virtual Environment. IEEE 4th International Conference on Mobile Services (MS 2015), June 27 - July 2, 2015, New York, USA
- [34] F. Abtahi, W. Li, Z. Zhu, T. Ro, Multimodal Speaker Recognition using Deep Belief Networks, CVPR Women in Computer Vision Workshop, 2015
- [35] M. Vincent, H. Tang, W. L. Khoo, Z. Zhu, T. Ro. Shape Discrimination using the Tongue: Implications for a Visual-to-Tactile Sensory Substitution Device. *Multisensory Research*, 29(8), October 2016, pp. 773-798. DOI: 10.1163/22134808-00002542

- [36] W. Li, X. Li, M. Goldberg and Z. Zhu, Face Recognition by 3D Registration for the Visually Impaired Using a RGB-D Sensor, Second Workshop on Assistive Computer Vision and Robotics, Zurich, Sept 12, 2014
- [37] C. Tsangouri, W. Li, Z. Zhu, F. Abtahi and T. Ro, An Interactive Facial-Expression Training Platform for Individuals with Autism Spectrum Disorder, 2016 IEEE MIT Undergraduate Research Technology Conference (URTC), November 4-6, 2016 at MIT, Cambridge USA.
- [38] W. Li, Z. Su, M. Li, Z. Zhu. A Deep-Learning Approach to Facial Expression Recognition with Candid Images. The 14th IAPR Conference on Machine Vision Applications (MVA 2015), Tokyo, May 18-22, 2015
- [39] W. Li, F. Abtahi, C. Tsangouri and Z. Zhu, Towards an "In-the-Wild" Emotion Dataset Using a Game-based Framework. In Affect in the Wild Workshop, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR'16) Workshops (CVPRW).
- [40] W. Li, F. Abtahi, Z. Zhu, A Deep Feature Based Multi-kernel Learning Approach for Video Emotion Recognition, Emotion Recognition in the Wild (EmotiW) Challenge 2015, the 17th ACM International Conference on Multimodal Interaction (ICMI 2015), Seattle, USA. November 9-13th, 2015 (Top 5 in terms of performance among 60 submissions)
- [41] W. Li, F. Abtahi, Z. Zhu, L. Yin. EAC-Net: A Region-based Deep Enhancing and Cropping Approach for Facial Action Unit Detection. The 12th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2017), May 30 –June 3, 2017 in Washington, DC.
- [42] W. Li, F. Abtahi, Z. Zhu. Action Unit Detection with Region Adaptation, Multi-labeling Learning and Optimal Temporal Fusing, IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017), July 21-26, 2017, Honolulu, Hawaii, USA
- [43] W. L. Khoo, E. L. Seidel and Z. Zhu, Evaluating Multimodal Sensing in Virtual Environment for Assisting the Visually Impaired, the 13th International Conference on Computers Helping People with Special Needs (ICCHP), July 11-13, 2012, Linz, Austria.
- [44] W. L. Khoo, "GIVE-ME: Gamification In Virtual Environments for Multimodal Evaluation - A Framework," PhD Dissertation (advisor: Z. Zhu), Department of Computer Science, The CUNY Graduate Center, 2016.
- [45] J. Gong, C. Feeley, H. Tang, G. Olmschenk, V. Nair, Y. Yu, Z. Zhou, K. Yamamoto, and Z. Zhu. Building Smart Transportation Hubs with Internet of Things to Improve Services to People with Special Needs. Transportation Research Board (TRB) 96th Annual Meeting, January 8-12, 2017
- [46] Zhu, Z., Gong, J., Tang, H., Feeley, C., Olmschenk, G., Nair, V. and Lysicatos, M. (2016) Towards Smart Transportation Hub: Services to Persons with Special Needs Requiring Minimal New Infrastructure, GCTC Expo, June 13-14, 2016 (YouTube video: https://youtu.be/_hxdXqTVG7I)
- [47] Z. Zhu, C. Kelnhofer-Feeley, J. Gong, W. Seiple, C. Brakewood. SCC-Planning: Towards Smart and Accessible Transportation Hub - Research Capacity Building and Community Engagement. NSF Award #1737533, Smart and Connected Community (S&CC) Program, Division Of Computer and Network Systems (CNS), CISE (Program Manager: Ralph Wachter). 2017-2018. https://www.nsf.gov/awardsearch/showAward?AWD_ID=1737533
- [48] Z. Zhu and J. Xiao, CCNY Joint Senior Design Program in Assistive Technology Across the Department Boundaries. The 2017 IEEE Integrated STEM Education Conference (ISEC'17), March 11, 2017, Princeton, New Jersey
- [49] F. Palmer, D. Zuleta, C. Rodriguez, J. Montesino and L. Ai (Mentors: Z. Zhu and T. Ro), V.I.S.T.A.: Vibro-Tactile Intelligent System for Travelling Aid - A Wearable Alternative. First Place Kaylie Prize of the 2012 CCNY Kaylie Entrepreneurship Competition.
- [50] L. Disla, S. Liu, O. Naeem, T. Panna. Mentors: Z. Zhu and H. Tang. GesTherapy: Using Gesture Based Therapy to Remodel Physical Rehabilitation. Second Place Dean's Prize of

the 2013 CCNY Kaylie Entrepreneurship Competition.

- [51] B. Tan, B. Xu, N. Yao and X.-B. Zhao. Mentors: J. Xiao and Z. Zhu. HAST: Health and Support Technology for Families using Open Source Wearable Devices. 2015 Social Innovation Prize Winner, Zahn Innovation Center.
- [52] A. Khan, J. Lopez, F. Moideen, W. L. Khoo and Z. Zhu, KinDectect: Kinect Detecting Objects, the 13th International Conference on Computers Helping People with Special Needs (ICCHP), July 11-13, 2012, Linz, Austria.
- [53] Z. Zhu, W. L. Khoo, C. Santistevan, Y. Gosser, E. Molina, H. Tang, T. Ro and Y. Tian, EFRI-REM at CCNY: Research Experience and Mentoring for Underrepresented Groups in Cross-disciplinary Research on Assistive Technology. The 6th IEEE Integrated STEM Education Conference (ISEC), March 6, 2016, Princeton, New Jersey.
- [54] C.-S. Lin, Z. Zhu, and T. Ro, Dynamic Project-based STEM Curriculum Model for a Small Humanities High School. The 2016 IEEE Integrated STEM Education Conference (ISEC), March 5, 2016, Princeton, New Jersey.
- [55] ERN – Emerging Researchers National Conference in STEM, <http://new.emerging-researchers.org/>

作者简介



Dr. Zhigang Zhu is currently a Herbert G. Kayser Chair Professor of Computer Science, at The City College and The Graduate Center, The City University of New York. He is Director of the City College Visual

Computing Laboratory (CcvCL), and Co-Director of the Center for Perceptual Robotics, Intelligent Sensors and Machines (PRISM) at CCNY. Previously he was Associate Professor at Tsinghua

University, and a Senior Research Fellow at the University of Massachusetts, Amherst. His research interests include 3D computer vision, multimodal sensing, virtual/augmented reality, video representation, and various applications in assistive technology, education, environment, robotics, surveillance and transportation. He has published over 160 technical papers in the related fields. His research at CCNY has been supported by AFOSR, AFRL, ARO, DARPA, DHS, NSF, NY State as well as industry. He is an Associate Editor of the Machine Vision Applications Journal, Springer, and IFAC Mechatronics Journal, Elsevier. He was a Technical Editor of the ASME/IEEE Transactions on Mechatronics from 2010 to 2014.

朱志刚博士，现为美国纽约城市大学终身教授，赫伯特·G·凯泽 (Herbert G. Kayser) 计算机科学讲席教授。曾任清华大学计算机系副教授，美国安城麻州大学计算机系资深研究员。先后主持参与 30 余项科研项目，由美国国家自然科学基金、国防部、国安部以及在美国工业界，和中国自然科学基金、863 计划等资助。发表论文 160 余篇，专著和译著 3 部，在世界各地做特约或专题报告 40 余次，研究的领域集中在人机视觉和多模态感知。现任计算机视觉期刊 Machine Vision and Applications 和 IFAC 机电一体化期刊副主编，曾任 ASME/IEEE 机电一体化汇刊 (Transactions on Mechatronics) 学术编辑。曾有幸分别与计算机视觉领域的两位泰斗 Thomas S. Huang 和 Takeo Kanade 合作编著《多模式监视：传感器、算法和系统》(Multimodal Surveillance: Sensors, Algorithms and Systems) 一书和《International Journal of Computer Vision》大型三维景物建模特辑。