人机交互中的体态语言理解

Body Language Understanding for Human Computer Interaction

徐光祐　陶霖密　邸慧军　著

電子工業出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
http://www.phei.com.cn

国家科学技术学术著作出版基金
工业和信息产业科技与教育专著出版资金 资助出版

人机交互中的体态语言理解

Body Language Understanding for Human Computer Interaction

徐光祐 陶霖密 邸慧军 著

電子工業出版社
Publishing House of Electronics Industry
北京 • BEIJING

内 容 简 介

以人为中心的人机交互，要求人机交互从目前占主导的，由用户直接操作进行的“显式交互”方式扩展到包括通过体态语言进行的“隐式交互”方式。体态语言理解是其中的关键问题。本书论述了与此相关的心理学和脑神经学基本概念。

通过多模态信息处理来识别和理解体态语言是从非结构化的多模态传感器数据到高层语义的多层次特征检测和推理过程，也是一个约束不充分的逆向求解问题。因此本书对觉察上下境计算支持的视觉处理和理解做了系统的论述，同时也介绍了上下境定义、上下境模型和觉察上下境计算的基本概念。主要包括：基于广义弹性运动跟踪的人体动作分析，三维空间人体定位与体态估计，容忍视角和距离变化的人体动作识别，日常生活中动作（ADL）识别和理解，基于动态上下境模型的群体交互行为分析，支持觉察上下境计算的分布式多模态信息处理系统。典型应用是面向老人生活和健康看护的“日常生活动作识别”（ADL），及以会议自动分析为代表的群体行为分析。

本书创新性强，内容系统、全面，深入浅出。目前国内相关领域的理论著作尚属于空白，在国际上也还缺乏系统的理论。本书的出版将对我国人机交互领域、体态语言理解的理论研究和学科发展具有重要的参考价值和指导意义。

图书在版编目（CIP）数据

人机交互中的体态语言理解/徐光祐，陶霖密，邸慧军著. —北京：电子工业出版社，2014.8
ISBN 978-7-121-23625-9

Ⅰ. ①人… Ⅱ. ①徐… ②陶… ③邸… Ⅲ. ①人-机系统－研究 Ⅳ. ①TB18

中国版本图书馆 CIP 数据核字（2014）第 136507 号

策划编辑：曲 昕（quxin@phei.com.cn）
责任编辑：曲 昕
印 刷：涿州市京南印刷厂
装 订：涿州市京南印刷厂
出版发行：电子工业出版社
北京市海淀区万寿路 173 信箱 邮编 100036
开 本：787×1092 1/16 印张：31.75 字数：812.8 千字 彩插：8
版 次：2014 年 8 月第 1 版
印 次：2014 年 8 月第 1 次印刷
定 价：128.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：（010）88254888。

质量投诉请发邮件至 zlts@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线：（010）88258888。

序　一

半个多世纪以来，随着计算机技术的迅猛发展以及性能的不断提高，与其相关的人机接口（人机界面、输入输出模式）也产生了巨大的变化。主机时代，当一台计算机被多个用户分享时，其工作程式是这样的，用户按预先约定的时间进入机房，把指令（程序）输入计算机，并等待机器给出计算的结果。这一切在当时都是通过打在穿孔纸带或者穿孔卡片上的符号（信息）实现的，当时把完成此项任务的设备，通称为主机之外的“外部设备”。进入台式机时代，又称 PC（个人计算机）时代后，一人用一台计算机，人机接口有了变化。用户的指令通过键盘、鼠标与（屏幕）视窗输入，从屏幕显示或打印机打印输出计算的结果。当我们使用手持式平板电脑，进入所谓移动服务时代时，人机接口又进化成为触摸式屏幕。人机接口的巨大变化，固然给使用者带来越来越便捷的操作方法。但遗憾的是，这些变化都仅仅是表面上的，而通过接口对计算机进行操作的模式并没有改变。在这类接口模式下，人们总是预先定义好一批计算机容易识别的指令，并通过不同的接口（穿孔纸带、键盘、触屏等）输入机器，计算机严格地执行这些指令，最后输出预期的结果。这种输入输出的模式对计算机来讲是很方便的，因为机器很容易识别这些约定好的指令。但这种方式对用户并不方便，他们需要事先熟悉指令的含义，学会如何操作；此外，由于机器只能执行事先定义的指令，因而不便于用户与机器的交互。总之，这种以计算机为中心的人机接口，普通百姓很难掌握。

在网络时代，当千百万台计算机被网络连接起来，数以千万计的用户涌向计算机的终端时，计算机真正进入了千家万户，人机关系也由此发生了质的变化。通过网络（计算机）可以提供多种多样的信息服务，如信息检索、购物、娱乐、聊天等。为了保证服务质量，并满足广大普通用户的需要，以机器为中心的人机接口需要做出根本性的变革，变为“以人为中心”的人机交互模式。在新的模式下，机器需要准确把握用户的意图、需求与兴趣，从而提供个性化的服务。此外，在网络环境下，计算机不仅需要与用户交互，用户还将通过计算机（网络）实现人际之间的交流与互动。换句话讲，人—机与人—人的交互发生在物理与信息这两个联合空间中。

主机与它的接口形影相随地变化，提出了一个看似“鸡与蛋”的难题，即究竟是主机的换代促成接口的变化，还是接口的创新带来主机的换代？无论怎样，不容质疑的一点是，人机交互在计算机科技中的地位不断上升，已经成为它的重要组成部分。本书的出版是及时的，正好满足了这一日益增长的需求。作者对人机交互的发展趋势做了如下的阐述：为使人机交互摆脱必须在计算机面前操作的束缚，扩展到人们生活的三维物理空间中去的必要条件是，使三维空间成为“物理—信息对偶空间”，同时把人机交互从单纯的显式人机交互，扩展到包括隐式人机交互。在这些分析的基础上，作者进一步指出以人为中心的人机交互正在走向现实。

本书围绕“体态语言理解”的中心展开对人机交互基础理论与关键技术的论述。在此我想强调一下此书出版的科学意义。20 世纪 30 年代，英国数学家图灵提出图灵机（现代计算机的数学模型）的概念时，他用“可计算实数”（computable numbers）来定义计算机的可计算性，这清楚地表明，计算机仅是一台从事数字（符号）机械计算的机器。虽然，随后人工

智能学者们，包括图灵本人，提出各种机器（计算机）智能的概念，如著名的图灵测试，那也只是从机器的表观行为上定义“智能”，与智能的本质并无关系。事实上，在“智能”的框架下，计算机依然只扮演数字计算的角色。正因如此，传统的计算理论总是避开“语义”这一难题，把理论建立在与“内容无关”（meaning independent）的假设之上。然而，当我们讨论人机交互时，“语义”却成为一个绕不开的话题，“体态语言理解”不就是这个话题吗？当用户通过语言或者肢体动作表达他的意见、诉求、喜好和愿望时，计算机能否从这些行为中获得其背后隐含的语义，预测其他用户对此行为的反应，并评估可能产生的影响？要完成这些任务，计算机至少需要解决三个层面上的问题：第一，识别这些信号，包括语音、图像、表情、手势等，可归结为模式识别问题；第二，识别了这些信号，即“听清”或“看清”这些信号之后，还需要发现隐藏在信号背后的语义，如信号发送者的意图，信号发送端所提供的信息等；第三，预测受众对这些信息的反应，以及所产生的影响。回顾传统信息处理理论的发展历史，我们可以看到，过去所从事的有关研究工作，都仅集中在解决第一个层面的问题上。以脸部表情理解为例，首先需要了解的是，“他的脸部有何变化（动作）”，嘴唇上翘，还是闭合？眼睛睁开，还是眯缝？即表情识别问题。即使这样看似简单的问题，由于光照、噪声、阴影与视角等因素的影响，在传统信息处理理论里，就已经是一个难解的不适定（病态，ill-posed）问题。因此仅就模式识别而言，目前仍有许多难题尚未解决。至于说到“理解”，我们需要解决的科学问题是，依据信息发送者发出的信号，预测和构造他本身以及信息接收者的认知模型。显然，如果没有周围物理与社会环境的信息，没有用户心理状态的知识，仅仅依靠建立在数学模型基础上的数据处理，是不可能完成这一任务的。本书介绍的上下境模型以及觉察上下境的计算范式，是一个很好的解决方案。该计算范式把从底向上的数据驱动与自顶向下的知识导引结合起来，即将传统的信息处理与人工智能的方法结合起来。本书汇集了国内外、本书作者及其团队在这个领域的最新研究成果，从人体定位与体态估计、人体动作识别与分析，一直到群体行为分析、体态语言理解；并结合两个典型场合，阐述它的解决方案和相关实验系统。其中一个是，面向会议自动分析的基于动态上下境模型的群体交互行为分析；另一个是，日常生活动作（ADL）识别和理解系统。本书内容丰富、充实，既有基础理论与关键技术，又有实验系统与实际应用，理论联系实际。

目前已有的出版物大多侧重于语言通信，对在人际交往中占主导地位的非语言通信，即体态语言缺乏系统论述，为数不多的也仅限于人体动作识别，缺乏对用户行为的分析和理解。本书虽然仅限于讨论体态语言理解，但它所涉及的理论与技术是普适性的，可以推广和应用到语言通信的分析与理解。

徐光祐教授是清华大学计算机系的资深教授，是我几十年的同事。20 世纪 70 年代他在计算机外部设备教研组工作，很早就与人机接口打交道。随着人机接口模式的改变，他及时地调整了研究方向，成为国内从事普适计算、多模态信息处理的先行者之一，在这些领域都有很深的学术造诣和丰富的实践经验。这本书汇集了作者及其所领导的团队十多年来取得的系统科研成果。

中国科学院院士
清华大学教授 张钹

2014 年 4 月

序　二

随着计算机进入普适计算时代，体态语言已成为人机交互的重要形式，被学术界和产业界关注。从生物学、心理学和社会学的角度来描述和解释体态语言的本质，就是根据人体对外部和内部的刺激集合做出反应时所显示的体态语言信号，来理解用户的交互意图、态度和情绪。体态语言理解是人机交互领域中的核心技术。

作者徐光祐教授长期从事图像图形处理、普适计算和人机交互的研究。围绕以人为中心的人机交互，作者所在课题组十余年来以多模态信息处理和理解为重点，开展了系统的基础研究，取得了重要科研成果，把人机交互从单纯的显式人机交互，扩展到包括隐式人机交互，尤其在人机交互的体态语言理解方面，成果突出。

在上述成果基础上，本书分析了人机交互领域的研究现状，阐述体态语言的关键问题，不仅对当前人机交互、人工智能、普适计算、计算视觉、环境智能等领域中的共性基本概念给予了清晰的解释，而且对体态语言理解的相关技术进行了较详尽的论述，同时反映了课题组在技术和算法上的创新成果，提出了“物理—信息对偶空间”理论，给出了隐式人机交互的科学的定义，并论述体态语言识别和理解的关键技术和实现方法。本书涉及隐式人机交互的理论、方法和应用，理论有深度，内容全面，对从事隐式人机交互和体态语言理解的同行有很好的参考价值，能推进我国人机交互领域的科研发展。

中国科学院软件研究所研究员

戴国忠

2014 年 5 月

前　言

在传统的人机交互模式下，用户需要在计算机面前，通过对键盘、鼠标这样的设备进行显式的操作才能得到服务或信息。以人为中心的计算的标志之一是使用户能在生活的三维物理空间中无需专门的操作，通过隐式人机交互方式得到服务或信息。在人与人之间的人际信息交流（交互）中，人们无意识的非语言行为（non-verbal behavior）——体态语言（body language）传递了比语言通信更为丰富的信息，并与语言信息结合传递完整的语义。这是隐式人机交互的依据和基础，因此首要研究的是人际交互中由语言通信和非语言通信组成的通信机制。要在人们生活的物理空间中得到信息服务，还需要使物理空间与通过网络相互连接的信息空间融合成为一个整体，这就是“物理—信息对偶空间”。从 20 世纪 90 年代开始，人机交互研究的重点已经开始转向计算机支持下的人们相互之间（人际）的通信，也就是说计算机将参与人际交互，因此在人机交互中还需要参照人类的社交行为规范，引入社交智能（social intelligence）。以上内容已成为在 21 世纪中研究和开发人机交互技术的基础。

基于计算机的体态语言理解是涉及人机交互、计算机视觉、普适计算和人工智能的跨学科研究课题，国内还没有相关的理论著作，在国际上也还缺乏系统的理论。本书是作者结合所在研究组十余年以来多项国家重要研究项目，总结跨学科研究成果的基础上撰写而成的。所支撑的创新成果包括：基于混合变换隐马尔科夫模型（MTHMM）的广义弹性运动跟踪，实时三维人体姿势估计和跟踪，基于包络形状和 R 变换的体态表示，基于 ADL-DBN 模型的行为在线推理，事件驱动的多层次动态贝叶斯网络模型，基于动态上下境的多层次事件自适应检测方法，动态上下境中群体动作分析等。希望本书能够给有关同仁提供一些参考，并借此抛砖引玉，推动此领域的理论和方法研究工作。

全书共 9 章，分析了人机交互领域的技术研究现状和关键问题，围绕解决体态语言理解中的语义鸿沟，以觉察上下境的视觉计算理论为指导，通过与国内外同领域的技术成果进行广泛比对和分析，系统总结了本课题组多年的研究成果，提出了具有创新性的技术和算法。

第 1 章讨论了以人为中心的人机交互、隐式人机交互、非语言通信，以及体态语言及其在人与人之间的交互，即人际交互中的作用。并在此基础上讨论体态语言的识别和理解及在以人为中心的人机交互，其中包括计算机的社交智能中的作用。

研究表明人类的感知与感觉运动（sensorimotor）机制紧密相关。理解另一个人动作的必要条件是同时在个人内部和个人之间形成闭合的“动作—感知”回路（action-perception loops）。探索“动作—感知”回路的运行机制是建立识别人体动作的计算机信息系统的理论基础。为此第 2 章将讨论动作理解的心理和神经机制基础。

体态语言是人类对外部和内部刺激集合的反应。体态语言理解就是根据观察到的体态语言线索和所显示的体态语言信号，来理解用户的交互意图、态度和情绪。这是一个约束不充分的逆向求解问题，因此需要应用觉察上下境的视觉处理方法。目前基于上下境信息的计算机视觉算法已受到高度重视，但觉察上下境的视觉处理还没有受到应有的重视。通过对体态

语言理解的研究来推动这方面的研究也是本书的目的。在第 3 章中将对上述问题进行讨论。

从多模态的传感器数据到人体语言理解这样的高层语义之间存在巨大的语义鸿沟。需要解决的关键问题包括：人体运动分析、动作识别和人体行为理解。第 4～7 章将对这些关键技术分别进行讨论。

人体运动分析通过检测和跟踪获取关于人体动作的“时—空”信息，是进行体态语言理解的基础。现有基于高层语义模型的方法难以适应现场人体动作识别时面临的人体运动和成像条件的多样性，为此第 4 章提出基于广义弹性运动跟踪的人体运动分析方法。采用自底向上的方式，在不依赖关于人体的特定先验模型的条件下，从弹性运动的角度分析人体的整体运动规律。

在人机交互应用中，要求在现场在线地识别人与人以及与周围物体之间的交互行为。需要解决“人—物”的空间关系分析（即人与哪些物体发生了交互）以及人体自身动作的识别（即人做了什么）这两个基本问题。与此相关的关键课题是：人体定位以及从粗到细、多层次的人体体态估计。为了探索新的人体姿势估计和跟踪的方法，本书第 5 章介绍了基于多摄像机的人体三维空间定位和人体头肩部轮廓的三维姿势估计和跟踪。通过实时检测人体三维定位、姿态和三维朝向并按照分层的检测和推理策略，在上下境信息的指引下实现人体姿势的实时估计和运动跟踪。人体动作识别方法若要工作在现场环境下，必须具备处理以下因素：视角变化、位置变化、遮挡的能力。目前还缺乏相应的成熟方法。第 6 章讨论可容忍视角、位置变化的人体动作识别。提出的基于“包容形状”的表示和相应的动作识别方法，不需要进行对应点匹配并具有良好的容忍视角变化能力。把包容形状表示与时域中的 R 变换（R-transform）相结合提出的具有视角和尺度恒常性的人体动作 VSI-Surf 表示以及分层的识别策略，在多视角动作库 IXMAS 的数据上取得了高性能的实验结果。把上述双摄像机应用扩展到具有多摄像机的实际工作环境，可一体化解决人体自由运动造成的视角变化、观察视野受限、遮挡等问题。

体态语言理解是一个综合性、跨学科的长远研究课题。目前在人体检测、动作识别、觉察上下境计算等方面虽然已经开展很多基础研究，但离体态语言理解的要求还有较大距离。本书作者所在课题组十余年以来，围绕两个具有代表性的典型场景，开展以人为中心的人机交互和体态语言识别和理解的研究，建立了相应的实验环境，取得了一些有发展前途的研究结果，为进一步的研究打下了良好基础：（1）智能家居中的老人看护场景下的人体日常生活（activity of dairy living， ADL）动作识别与行为分析；（2）会议场景下群体交互行为的在线分析。这两项应用分别是第 7 章和第 8 章的内容。本书通过典型应用场景和实验环境检验和测试了各项相关方法和技术的可行性。

第 9 章探讨将分布式计算结构用于以人为中心计算模式的觉察上下境应用系统的途径，介绍了支持觉察上下境计算的分布式多模态信息系统。

本书全面整理了近年来最新的科研成果，创新性强，内容系统、全面，深入浅出，对从事多媒体信息处理，特别是计算机视觉、人机交互、人工智能和普适计算领域研究的科技工作者来说是很有用的参考书。这本书也可作为研究信息处理、人机交互和人工智能科学的研究生教材。

本书的作者在研究中长期合作，相互分工配合。徐光祐总体负责，负责 1～3，6～8 章的撰写并全书统稿。陶霖密和邸慧军分别在组织实施和方法探索方面发挥了重要作用，邸慧军负责第 4、5 章，陶霖密负责第 9 章。

完成相关项目的教师合作者有陶霖密、史元春。参与项目研究，对本书的内容做出贡献的，有从事人体动作识别和理解的博士后张翔、邸慧军；博士生戴鹏、黄飞跃、金国英、董力赓、孙洛、曹媛媛、白雪生、王强、任海兵、彭振云、张辉、谢峰、柳杨华、叶航军、王磊、杨雨东，硕士生李昕、王焱、朱蓝天、宋刚、庄莉、罗明、刘亚等，还有参与智能环境和分布式信息系统的博士生谢伟凯、王国意、董轩民、谭焜、索岳、喻纯、邹怀荣、陈恩义、王国建，硕士生王垚、赵彦钧、蒋长浩、毛雁华等。由于这些同学的刻苦钻研，以及开创性的工作和贡献才使本书的出版成为可能。在此谨向他们表示我们诚挚的感谢。本文在撰写过程中也参考了很多国内外同行的研究成果和资料，一并向他们表示感谢！

体态语言理解是一个多学科交叉的新兴研究课题，同时由于水平有限，书中如有疏漏或错误之处，敬请读者不吝指正。

徐光祐

2014 年 5 月于清华园

目　录

第 1 章　以人为中心的人机交互与体态语言理解

计算机、网络和数字技术正在深刻地改变人们的生活，但阻碍这些以计算机为核心的信息技术真正融入人们的生活，成为生活中的必需品的最根本的障碍，是计算机还不能根据传感器数据来识别和理解人们的情绪、态度、意愿等内心活动，从而无法以人们所习惯的方式与人们进行信息交流和提供主动的服务。也就是说，把人机交互从以计算机为中心转变为以人为中心的方式是信息技术能为人类社会发挥更大作用的关键。

在传统的人机交互模式下，用户需要在计算机面前通过对键盘、鼠标这样的设备进行显式的操作才能得到服务或信息。以人为中心的计算标志之一是使用户能在生活的三维物理空间中无须进行专门的操作，通过隐式人机交互方式得到服务或信息。在人与人之间的人际信息交流（交互）中，人们无意识的非语言行为（non-verbal behavior）——体态语言（body language）传递了比语言通信更为丰富的信息并与语言信息结合传递完整的语义。这是隐式人机交互的依据和基础。因此首先需要研究人际交互中由语言通信和非语言通信组成的通信机制。要在人们生活的物理空间中就能得到信息服务，还需要使物理空间与通过网络相互连接的信息空间融合成为一个整体，这就是“物理—信息对偶空间”。从 20 世纪 90 年代开始，人机交互研究的重点已经开始转向计算机支持下的人们相互之间（人际）的通信，也就是说计算机将参与人际交互。因此在人机交互中还需要参照人类的社交行为规范和引入社交智能（social intelligence）。以上这些内容已成为在 21 世纪中研究和开发人机交互技术的基础。

为此本章将分别讨论什么是以人为中心的人机交互、隐式人机交互、非语言通信和体态语言及其在人与人之间的交互，即人际交互中的作用。并在此基础上讨论体态语言的识别和理解及在以人为中心的人机交互（其中包括在计算机的社交智能）中的作用，并以此做为本书的序论。

1.1　以人为中心的人机交互

信息和通信是任何社会的两个支柱。嵌入式计算、无线通信技术和互联网技术的迅速发展促进了计算机、互联网、移动通信、数字媒体技术的融合，从而使得人们的生活前所未有地充满了数字和信息。数字是这些领域中使用的共同语言，因此有人认为我们正在进入一个数字化的时代。在数字化时代中无论是信息和通信都是以计算为核心，所以也可以说计算正在引起一场革命——我们所做的每件事正在以我们从未经历过的速度改变着。我们可确信的是计算正在影响我们与其他人通信、交互的方式，设计和建造我们自己的家居和城市，影响我们的学习、通信、娱乐的方式。总而言之，计算技术正在日益影响我们的日常生活的每个方面。但是，不幸的是计算的发展是一把双刃剑，这些改变不都总是积极的，主要表现为以下两个方面。

1．数字鸿沟

目前各种类型的信息、产品、服务、人们、地图等不断地成为数字生态系统的一部分。获取信息服务成为提高生活质量的必经之路，而掌握计算机的能力已成为改善生活的关键。

但荒谬的是，我们发展的计算技术既是通向各种信息源的通路，又是访问这些信息的瓶颈。这是因为：目前的计算模式和技术是笨拙的、对用户不友好的、不自然的，人机交互方式不符合人类自然的交互习惯。因此，只能是有文化的人，而且需要花费大量的时间和精力来学会使用，才能利用计算机带来的方便和好处；对于没有机会接受相关教育的人群来说，生活的某些方面正变得更复杂和费力。使得情况更为严重的是，目前计算技术的发展主要从技术本身的改进和需要出发，也就是以技术为中心，而不是以人为中心的方法。单纯从技术发展的角度看，计算机技术取得了巨大的进步：体积越来越小，功能越来越强，而其价格更为低廉。存储容量和传输速度也已不是阻碍其发展的关键因素，但与此同时操作越来越复杂，需要更多的训练才能掌握。如何解决数字鸿沟问题？扩大计算机教育的范围可能缓解但不能解决这个问题。一种根本性的解决方法是丢弃以计算机为中心的设计方法，转向以人为中心的人机交互的设计方法。

2．桌面计算模式和显式人机交互

计算机的计算模式经历了主机计算—桌面计算—移动计算的发展阶段。初期的主机计算（main frame），人们要到计算机中心，才能使用计算机，是许多人利用分时的方式共享使用一台主机；在桌面计算模式下，每个用户单独使用一台或几台桌面计算机，其中还可能包括笔记本电脑，这是目前用户使用计算机的主要方式。随着无线通信和移动计算的发展，人们不仅在办公室、家居这样的固定场所使用计算机，还可以在汽车或在旅行中方便地使用计算机。但到目前为止，无论在哪种计算模式下，用户都是通过键盘、鼠标、显示器或相应的替代物，以直接向计算机发送命令和信息的方式来操作计算机。这就是所谓的显式人机交互方式。这种交互模式的特点是：（1）计算机只是被动地等待命令和信息，否则它不会工作。因此，与计算机交互必须有相应的接口。在桌面计算模式下，用户需要在计算机面前通过接口设备才能使用计算机。（2）它无视用户的状态和需求，不会主动地提供服务。（3）计算机对用户的响应或服务是事先定义的，难以按照用户当前的状态和需求做必要的调整。（4）计算机只接受它所能接受的命令，也就是符合计算机规定格式的命令，而不论用户的文化背景和习惯如何，包括所使用的文字。

这样的人机交互方式，在微电子技术迅速发展的今天，又会遇到新的问题。著名的 Moore 定律（Moore's Law）认为硅片上系统的集成密度每 18 个月翻一番（Noyce，1977）。这样的趋势持续了 30 年，为半导体技术的发展提供了一个清晰的预测。现在 Moore 定律正沿着一维—二维—三维的方向发展。三维电路中可以开发极强功能的纳米机械系统或在一个包装中集成了传感器、执行器、计算和通信功能的单个纳米电子系统（micro electronic mechanical system，MEMS；system in a package，SIP）。电子设备的设计和制造已达到了微型化的水平，这将允许把用于处理、通信、存储、显示和访问的系统与任何实际的物体（如衣服、家俱、汽车和家居）相集成。也就是说，显式人机交互所需的接口可能已经难以找到。

为此，人们必须探索新的计算模式和以人为中心的人机交互方式，这就是普适计算（ubiquitous/pervasive computing，UPC）和背景智能（ambient intelligence，AmI）。普适计算和背景智能的出现，一方面提出了人机交互必须向以人为中心的方向发展，同时也使计算本身发生深刻的变化，即用户所面对的已不再是一台孤立的计算机，而是通过互联网联结成的信息空间（cyber space）。这个信息空间通过背景智能已与人们生活的物理空间融合在一起。这就是“物理–信息对偶空间”，它为人机交互带来了全新的环境。因此，在以下的章节中我们将不仅需要