

# 人工智能之情感计算

## Research Report of Affective Computing

2019年 第6期



清华大学人工智能研究院  
北京智源人工智能研究院  
清华-工程院知识智能联合研究中心  
2019年9月

# 目录

## CONTENTS

### 第一篇 概述篇

1.1 情感计算的产生及发展 .....	1
1.2 概念定义 .....	3
1.3 情感计算的理论观点 .....	10
1.4 情感计算的认知困境 .....	10

### 第二篇 技术篇

2.1 传统的研究 .....	13
2.2 新兴的研究 .....	25
2.3 情感计算国际会议 .....	27
2.4 情感计算获奖论文 .....	32

### 第三篇 人才篇

3.1 学者情况概览 .....	36
3.2 典型学者 .....	40
3.3 研究机构 .....	54

### 第四篇 应用篇

4.1 课堂教学 .....	57
4.2 情绪监测 .....	58
4.3 医疗康复 .....	59
4.4 舆情监控 .....	61

### 第五篇 趋势篇

5.1 论文研究发展趋势 .....	63
5.2 情感计算技术预见 .....	66
5.3 情感计算未来发展 .....	67

参考文献 .....	69
------------	----



# 图表目录

图 1 情感计算的研究内容.....	5
图 2 手势识别的基本框架.....	23
图 3 情感计算全球专家分布.....	36
图 4 情感计算专家国家数量分布.....	37
图 5 情感计算领域学者 h-index 分布.....	37
图 6 情感计算专家迁徙图.....	38
图 7 情感计算国内学者分布.....	39
图 8 情感计算学者分布国内省份 TOP10 .....	40
图 9 EQ-Radio 无线信号监测 .....	59
图 10 NAO 机器人.....	60
图 11 情感计算发展趋势.....	62
图 12 情感计算 2007 年经典论文.....	63
图 13 1970-2019 论文研究方向趋势图 .....	64
图 14 情感计算技术预见图.....	65
图 15 情感计算预测热词图.....	66
表 1 十个文本情感分析 API .....	15
表 2 常用的语音情感数据库.....	19
表 3 脸部表情运动特征具体表现.....	22
表 4 常用的表情分析数据库.....	24
表 5 常用的手势分析数据库.....	25
表 6 AVEC 2018 Sub-Challenge Winners.....	28
表 7 AVEC 2017 Sub-Challenge Winners.....	29
表 8 AVEC 2016 Sub-Challenge Winners.....	30
表 9 EmotiW 2017 Sub-Challenge Winners .....	31
表 10 EmotiW 2016 Sub-Challenge Winners .....	32
表 11 情感计算相关期刊、会议.....	35
表 12 情感计算荐领域中国与各国合作论文情况.....	38
表 13 1970-2019 各阶段前十位研究关键词 .....	64



# 1 概述篇

## 1.1 情感计算的产生及发展

40 多年前，诺贝尔奖得主 Herbert Simon 在认知心理学方面强调，解决问题论要结合情感的影响。情感的识别和表达对于信息的交流和理解是必需的，也是人类最大的心理需求之一。人类的认知、行为等几乎都要受到情感的驱动，并影响着人际互动以及群体活动。在人与人的交往中，情感的交流还常被用来完成人的意图的传递。因此，在智能人机交互的研究中，拥有对情感的识别、分析、理解、表达的能力也应成为智能机器必不可少的一种功能。作为人工智能创始人之一的美国麻省理工学院 Marvin Minsky 教授首次提出让计算机具有情感的能力，他在其专著《The Society of Mind》中强调情感是机器实现智能不可或缺的重要能力。20 世纪 90 年代初，耶鲁大学心理学系的 Peter Salovey 教授提出了情感智能的概念，并开展了一系列的研究。该概念随后被 Daniel J. Goleman 发展为与智商（IQ）相对的情商（EQ），随着 Goleman 的赋予计算机情感能力，并让计算机能够理解和表达情感的探讨与研究引起了计算机界众多专家的兴趣，他们在情感研究的理论和实验应用方面积累了很多经验。

学术界较早对情感进行系统研究的是美国麻省理工学院媒体实验室的 Rosalind W. Picard 教授。1995 年，Picard 首先提出通过识别人体的情感信号，来创建一种能够感知、识别和理解人的情感，并且能够做出智能、灵敏和友好反应的计算机系统。1997 年，Picard 出版专著《Affective Computing》，书中给出了情感计算的定义，即情感计算是关于情感、情感产生以及影响情感方面的计算，并对情感计算的研究做了系统介绍。情感赋予计算机像人一样的观察、理解和生产各种情感特征的认知能力。在情感表达和识别方面，随着多媒体技术和人工智能技术的不断发展和广泛应用，机器的识别水平在不远的将来一定会有长足的进步。

日本在 20 世纪 70 年代提出了“感性工学”的概念，在文部省主导下，日本从 20 世纪 90 年代开始了对感性工学的研究。所谓感性工学，就是将感性与工程结合起来的技术，是在感性科学的基础上，通过分析人类的感性，把人的感性需要加入到商品设计和制造中去，它是一门从工程学角度研究能带给人喜悦和满足的商品制造的技术科学。日本在人工智能、心理学、认知、情报处理等方面都展开了相关的研究。其中之一是由筑波大学原田昭教授主持的一个超大型特别研究项目——“感性评价构造模式的构筑”。这一计划从 1997 年 7 月启动，为期三年，集中了约 50 位各国研究人员，包括工业设计、机器人工程、控制工程、资讯工程、信息管理、认知科学、美学、艺术等众多领域的专家团队，分为感性评价、程序与感性数据库和机器人系统三组。该计划在全球几个重要的美术馆，设置附有摄影机的机器人；然后让受测者在其他地

方，通过网际网络计算机联机，远距离遥控机器人来观赏艺术品；对于受测者操纵机器人的整个过程，进行纪录并分析整理，以了解观赏者在鉴赏艺术品时，如何建立其感性评价的心理机制。

欧盟国家也在积极地对情感信息处理技术（表情识别、情感信息车辆、可穿戴计算等）展开研究。欧洲许多大学成立了情感与智能关系研究小组，其中较为著名的有日内瓦大学 Klaus Soberer 领导的情绪研究实验室、布鲁塞尔自由大学的 D Canamero 领导的情绪机器人研究小组以及英国伯明翰大学的 A Sloman 领导的 Cognition and Affect Project。在市场应用方面，德国 Mehrdad Jaladi-soli 等在 2001 年提出了基于 Embassi 系统的多模型购物助手。Embassi 是由德国教育及研究部（BMBF）资助并由 20 多个大学和公司共同参与的，以考虑消费者心理和环境需求为研究目标的网络型电子商务系统。

我国对人工情感和认知的研究始于 20 世纪 90 年代，并逐步得到重视。国家自然科学基金早在 1998 年就将和谐人机环境中的情感计算理论研究列为当年信息技术高技术探索第六主题。2003 年 12 月，在北京召开了第一届中国情感计算及智能交互学习会议，标志着国内学术界对情感信息处理研究的肯定和认同。2004 年，国家自然科学基金委批准资助了重点基金项目情感计算理论与方法。这标志着我国在人工情感领域的研究达到了一个新的水平，呈现出方兴未艾的发展势头，研究队伍迅速扩大，研究领域急速拓展。2005 年 9 月，我国 40 多名专家教授在北京召开了中国人工智能学会首届全国人工心理与人工情感学术会议，并倡议成立中国人工智能学会人工心理与人工情感专业委员会，开展相关方面的学术活动。2005 年 10 月，中国人工智能学会同意并上报国家民政部，批准成立了中国人工智能学会人工心理与人工情感专业委员会。随后，人工心理与人工情感专业委员会及其成员组织召开了全国第一届人工心理与人工情感学会会议（北京科技大学，2005）和首届国际情感计算与人机交互国际会议（中科院自动化所，2005）<sup>[1]</sup>。2007 年 12 月，中国人工智能学会人工心理与人工情感专业委员会在哈尔滨 CAAI-12 届年会上举行了正式成立大会，这是国内在电子信息科学领域的首个情感计算学会。学会集合了国内一流的人工心理与人工情感的研究专家，他们获得了第一个关于人工情感计算的国家自然科学基金重点基金（清华大学）；第一个“973”项目中的和谐人机交互理论与技术的研究课题（中科院软件所）。随着研究的不断深入，关于人工情绪和情感计算的课题越来越多。

## 1.2 概念定义

### 1.2.1 情感

情感(emotion)一词源于希腊文“pathos”，最早用来表达人们对悲剧的感伤之情。达尔文(Darwin)认为，情感源于自然，存活于身体中，它是热烈的、非理性的冲动和直觉，遵循生物学的法则。理智则源于文明，存活于心理。《心理学大辞典》将情感定义为“人对客观事物是否满足自己的需要而产生的态度体验”。Antonio Damasio 在其神经生物学研究结果的基础上将情感至少分为两类，即原发性情感和继发性情感。原发性情感被认为是与生俱来的，被理解为一岁儿童情感这种典型的情感类型，继发性情感被假设为从更高的认知过程中产生。而 James Russell)则从两个方面构造情感：核心情感和心理建构，前者表示神经系统的状态，如昏昏欲睡；后者表示行动，如面部表情、音调，以及行动之间的关联。由于情感的复杂性，研究情感的相关学者对情感的定义至今也未达成一致，记载的相关理论就有 150 多种。

而“emotion”一词由前缀“e”和动词“move”结合而来，直观含义是从一个地方移动到另一个地方，后来逐渐被引申为扰动、活动，直到近代心理学确立之后，才最终被詹姆斯(William James)用来表述个人精神状态所发生的一系列变动过程。Picard 曾在其书中专门对情感和情绪方面术语进行了区分，她认为相对情感而言，情绪表示一个比较长的情感状态。情感影响我们的态度、情绪和其他感觉、认知功能、行为以及心理。同时情感容易在多次情绪体验的基础上实现，当人们多次觉得完成一项任务很高兴，就会爱上这个任务。相比情绪而言，情感更具有深刻性和稳定性。在自然语言处理中，Myriam D 等人结合韦氏字典以及他们的相关研究得到的结论是，在语言中情感是无意识的，并且很难将其定义，从文本中可以检测到的是有意识的情感，是情绪表征。而情绪这一复杂心理学现象几乎不能从文本中全部检测出，能检测到的是情绪的构成因素。许多关于情感计算的研究并没有完全区分情绪和情感（包括本文引用的大部分论文），本文统一使用“情感”一词。

情感具有三种成分：

- 主观体验，即个体对不同情感状态的自我感受；
- 外部表现，即表情，在情感状态发生时身体各部分的动作量化形式。表情包括面部表情（面部肌肉变化所组成的模式）、姿态表情（身体其他部分的表情动作）和语调表情（言语的声调、节奏、速度等方面的变化）；
- 生理唤醒，即情感产生的生理反应，是一种生理的激活水平，具有不同的反应模式。

### 1.2.2 情感计算

让计算机具有情感能力的观点并不新鲜，它与“机器人”一词几乎同时出现。1985年，人工智能的奠基人之一 Minsky 就明确指出：“问题不在于智能机器能否有情感，而在于没有情感的机器能否实现智能”。但当时，赋予计算机或机器人以人类式的情感，主要还是科幻小说中的素材，在学术界罕有人关注。1995年情感计算的概念由 Picard 首次提出，并于1997年正式出版《Affective Computing（情感计算）》。在书中，她指出“情感计算就是针对人类的外在表现，能够进行测量和分析并能对情感施加影响的计算”，开辟了计算机科学的新领域，其思想是使计算机拥有情感，能够像人一样识别和表达情感，从而使人机交互更自然。

情感研究可以从两个方面来理解，一是基于生理学的角度，通过各种测量手段来记录人体的各种生理参数，比如，人体运动数据，脸部表情、心理、脉搏、脑电波等，并以此为根据来计算人体的情感状态；二是基于心理学的角度，通过各种传感器接收并处理信息，并以此为根据计算人造机器所处的情感状态。Picard 以人类情绪的生理信号处理为基本出发点，研究取得了诸多进展，其应用领域也在逐渐扩大。当然，并不是所有的研究者都同意 Picard 的想法。例如 Sengers、Gaver、Dourish 和 Kristina Hook 等学者借鉴现象学并且把情感看作人与人、人与机互动中的成分。情感互动方法认为应从一个对情感建设性的、人文决定性视角展开，而非从认知和生物学这一更传统的角度出发，这种方法将重点放在使人们获得可以反映情感的体验并以某种方式来修改他们的反应。

简单来说，情感计算研究就是试图创建一种能感知、识别和理解人的情感，并能针对人的情感做出智能、灵敏、友好反应的计算机系统。显然，情感计算是个复杂的过程，不仅受时间、地点、环境、人物对象和经历的影响，而且要考虑表情、语言、动作或身体的接触。在人机交互中，计算机需要捕捉关键信息，觉察人的情感变化，形成预期，进行调整，做出反应。例如通过对不同类型的用户建模（如操作方式、表情特点、态度喜好、认知风格、知识背景等），以识别用户的情感状态，利用有效的线索选择合适的用户模型，并以适合当前用户的方式呈现信息。在对当前的操作做出及时反馈的同时，还要对情感变化背后的意图形成新的预期，并激活相应的数据库，及时主动地提供用户需要的新信息。举例来说，麻省理工学院媒体实验室的情感计算小组研制的情感计算系统通过记录人面部表情的摄像机和连接在人身体上的生物传感器来收集数据，然后由一个“情感助理”来调节程序以识别人的情感。假设你对电视讲座的一段内容表现出困惑，情感助理会重放该片段或者给予解释。而目前国内情感计算的研究重点在于通过各种传感器获取有人的情感所引起的生理及行为特征信号，确定情感类别的关键特征，建立“情感模型”，从而创建个人情感计算系统。

情感计算是一个高度综合化的研究和技术领域。通过计算科学与心理科学、认知科学的结合，研究人与人交互、人与计算机交互过程中的情感特点，设计具有情感反馈的人与计算机的交互环境，将有可能实现人与计算机的情感交互。情感计算研究将不断加深对人的情感状态和机制的理解，并提高人与计算机界面的和谐性，即提高计算机感知情境，理解人的情感和意图，做出适当反应的能力，其主要研究内容如下图所示：

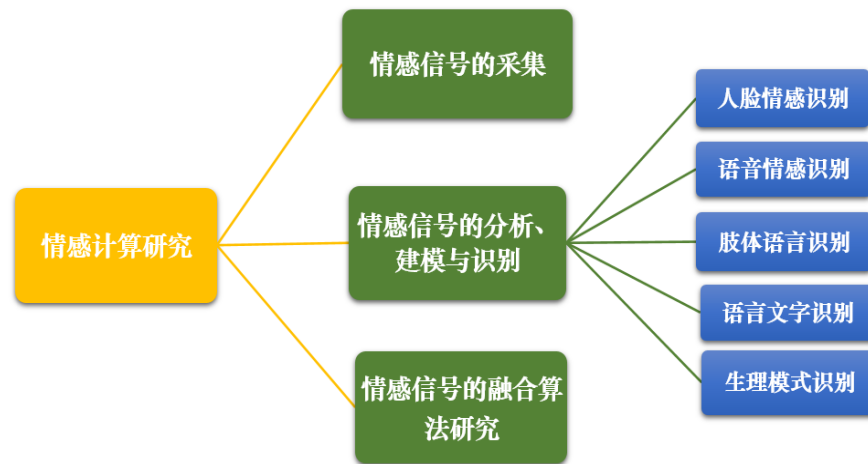


图 1 情感计算的研究内容

### 1.2.2.1 情感信号的采集

情感信号的获取研究主要是指各类有效传感器的研制。它是极为重要的环节，没有有效的传感器，就没有情感计算的研究，因为情感计算的所有研究都是基于传感器所获得的信号。各类传感器应具有如下的基本特征：使用过程中不影响用户（如重量、体积、耐压性等），经过医学检验对用户无伤害；数据的隐私性、安全性和可靠性；传感器价格低、易于制造等。美国麻省理工学院媒体实验室的传感器研制较为先进，已研制出多种传感器，如脉压传感器、皮肤电流传感器、汗液传感器及肌电流传感器等。其中，皮肤电流传感器可实时测量皮肤的导电系数，通过导电系数的变化可测量用户的紧张程度；脉压传感器可时刻监测由心率变化而引起的脉压变化；汗液传感器是一条带状物，可通过其伸缩的变化时刻监测呼吸与汗液的关系；肌电流传感器可以测得肌肉运动时的弱电电压值。

情感信号的获取必须通过一定形式的情感测量技术来完成，情感测量包括对情感维度、表情维度和生理指标三种成分的测量。例如，我们要确定一个人的焦虑水平，可以使用问卷测评其主观感受，通过记录和分析面部肌肉活动测量其面部表情，并用血压计测量血压，对血液样本进行化验，检测血液中肾上腺素水平等。确定情感维度对情感测量有重要意义，因为只有确



定了情感维度，才能对情感体验做出较为准确的评估。情感维度具有两极性，例如，情感的激动性可分为激动和平静两极，激动指的是一种强烈的、外显的情感状态，而平静指的是一种平稳、安静的情感状态。心理学的情感维度理论认为，几个维度组成的空间包括了人类所有的情感。但是，情感究竟是二维，三维，还是四维，研究者们并未达成共识。情感的二维理论认为，情感有两个重要维度：一是愉悦度，通过惊反射的方式测量其生理指标；二是激活度，通过皮肤电反应方式测量其生理指标。

### 1.2.2.2 情感信号的分析、建模与识别

一旦各类有效传感器获得了情感信号，下一步就是将情感信号与情感机理相应方面的内容对应起来，这里要对所获得的信号进行建模和识别。由于情感状态是一个隐含在多个生理和行为特征之中的不可直接观测的量，不易建模，部分可采用诸如隐马尔可夫模型、贝叶斯网络模式等数学模型。美国麻省理工学院媒体实验室给出了一个隐马尔可夫模型，可根据人类情感概率的变化推断得出相应的情感走向。

计算机对从传感器采集来的信号进行分析和处理，从而得出对方（人）正处在的情感状态，这种行为叫做情感识别。从生理心理学的观点来看，情绪是有机体的一种复合状态，既涉及体验又涉及生理反应，还包含行为，其组成至少包括情绪体验、情绪表现和情绪生理三种因素。目前对于情感识别有两种方式，一种是检测生理信号如呼吸、心律和体温等，另一种是检测情感行为如面部特征表情识别、语音情感识别和姿态识别。

#### ● 人脸情感识别

在生活中，人们很难保持一种僵硬的脸部表情，通过脸部表情来体现情感是人们常用的较自然的表现方式，其情感表现区域主要包括嘴、脸颊、眼睛、眉毛和前额等。人在表达情感时，只稍许改变一下面部的局部特征（譬如皱一下眉毛），便能反映一种心态。1972年，著名学者 Ekman 提出了脸部情感的表达方法（脸部运动编码系统 FACS），通过不同编码和运动单元的组合，即可以在脸部形成复杂的表情变化，譬如幸福、愤怒、悲伤等。该成果已经被大多数研究人员所接受，并被应用在人脸表情的自动识别与合成方面。随着计算机技术的飞速发展，为了满足通信的需要，人们进一步将人脸识别和合成的工作融入到通信编码中。最典型的便是 MPEG4 V2 视觉标准，其中定义了 3 个重要的参数集：人脸定义参数、人脸内插变换和人脸动画参数。表情参数中具体数值的大小代表人激动的程度，可以组合多种表情以模拟混合表情。在目前的人脸表情处理技术中，多侧重于对三维图像更加细致的描述和建模。通常采用复杂的

纹理和较细致的图形变换算法，达到生动的情感表达效果。在此基础上，不同的算法形成了不同水平的应用系统。

随着人脸的计算机处理技术（包括人脸检测和人脸识别）不断完善，利用计算机进行面部表情分析也就成为可能。由于各种面部表情本身体现在各个特征点运动上的差别并不是很大，而表情分析对于人脸的表情特征提取的准确性和有效性要求比较高，因而难以顺利地实现。具体的表情识别方法主要有三个：一是整体识别法和局部识别法，二是形变提取法和运动提取法，三是几何特征法和容貌特征法。当然，这三个发展方向不是严格独立的，恰恰相反，它们是相互联系和影响的，它们从不同侧面提取所需要的表情特征，都只是提供了一种分析表情的思路。例如：嘴巴张开并不代表就是笑，也有可能是哭和惊讶等。所用到的识别特征主要有：灰度特征、运动特征和频率特征三种。灰度特征是从表情图像的灰度值上来处理，利用不同表情有不同灰度值来得到识别的依据；运动特征利用了不同表情情况下人脸的主要表情点的运动信息来进行识别；频域特征主要是利用了表情图像在不同的频率分解下的差别，速度快是其显著特点。

通常根据视频识别要比根据静态图像识别更准确，视频能捕捉某种表情形成过程的面部动作。当人通过视觉器官把他人面部的刺激信号接收并传递到人的大脑之中，大脑就会进行人脸检测、人脸图像预处理、人脸特征提取等程序，然后，把以前存储在大脑中的若干基本表情的人脸特征（即脸谱）提取出来，进行对比分析和模糊判断，找出两者的人脸特征最接近的某种基本表情。这时，大脑皮层就会接通该基本表情所对应的兴奋区与边缘系统的神经联系，从而产生愉快或痛苦的情感体验。同时，大脑皮层还会接通该基本表情所对应的兴奋区与网状结构的神经联系，从而确定愉快或痛苦的强度。

### ● 语音情感识别

在人类的交互过程中，语音是人们最直接的交流通道，人们通过语音能够明显地感受到对方的情绪变化，例如通过特殊的语气词、语调发生变化等等。在人们通电话时，虽然彼此看不到，但能从语气中感觉到对方的情绪变化。例如同样一句话“你真行”，在运用不同语气时，可以使之成为一句赞赏的话，也可以使之成为讽刺或妒忌的话。目前，国际上对情感语音的研究主要侧重于情感的声学特征分析这一方面。一般来说，语音中的情感特征往往通过语音韵律的变化表现出来。例如，当一个人发怒的时候，讲话的速率会变快，音量会变大，音调会变高等，同时一些音素特征（共振峰、声道截面函数等）也能反映情感的变化。语音情感识别是指由计算机自动识别输入语音的情感状态。不同语言声调表情的信号在其时间构造、振幅构造、基频构造和共振峰构造等特征方面也有着不同的构造特点和分布规律。只要把各种具体模式的语言声调表情在时间构造、振幅构造、基频构造和共振峰构造等方面的特点和分布规律进行测算和分析，并以此为基础或模板，就可以识别出所有语言声调中所隐含的情感内容。中国科学

院自动化研究所模式识别国家重点实验室的专家们针对语言中的焦点现象，首先提出了情感焦点生成模型。这为语音合成中情感状态的自动预测提供了依据，结合高质量的声学模型，使得情感语音合成和识别率先达到了实际应用水平。

语音中的情感特征化比面部表情的情感特征化要难。面部表情信号传达了个人特征和表情，一般不传达语言信息。另一方面，语音信号包含的是混合信息，包括说话者特征、情感和说话内容中强调的词汇和语法。计算机在语音情感的识别和合成方面的进展很慢。随着计算机多媒体技术的不断发展，能处理包含在媒体中的情感信息的拟人化的多媒体计算机系统的研究越来越引起人们的兴趣。因为语音信号既是多媒体人机交互的主要利用方式，又是转载情感信息的重要媒体，所以对于包含在语音信号中的情感信息进行计算机处理研究就显得尤为重要。当人通过听觉器官把他人的语言声调信号接收并传递到人的大脑之中，大脑就会对其时间构造、振幅构造、基频构造和共振峰构造等方面的特点和分布规律进行检测、预处理和特征提取，然后，把以前存储在大脑中的若干基本表情的语言声调信号的时间构造、振幅构造、基频构造和共振峰构造等特征方面的构造特点和分布规律提取出来，进行对比分析和模糊判断，找出两者声音特征最接近的某种基本表情。

### ● 肢体语言识别

人的姿态一般伴随着交互过程而发生变化，这种变化表达了一些信息。例如手势的加强通常反映一种强调的心态，身体某一部位不停地摆动，则通常具有情绪紧张的倾向。相对于语音和人脸表情变化来说，姿态变化的规律性较难获取，但由于人的姿态变化会使表述更加生动，因而人们依然对其表示强烈的关注。科学家针对肢体运动专门设计了一系列运动和身体信息捕获设备，例如运动捕获仪、数据手套、智能座椅等；国外一些著名的大学和跨国公司，例如麻省理工学院、IBM 等则在这些设备的基础上构筑了智能空间；同时也有人将智能座椅应用于汽车的驾座上，用于动态监测驾驶人员的情绪状态，并提出适时警告。

### ● 语言文字识别

语言和文字属于第二信号系统，由于它们脱离了现实事物对于人的条件反射活动的直接参与，从而可以更为广泛地、普遍地、直接地、快速地、灵活地、多样地、深入地对各种事物进行认识，从而可以使人类能够对更为抽象、更为本质、更为遥远、更为间接、更为广泛、更为模糊、更为变化莫测的事物的价值关系进行认识和反应。

当人通过听觉器官把他人的语言信号接收并传递到人脑中，或者通过视觉器官把文字信号接收并传递到人脑中，大脑就会对其进行语义分析，对它们所描述事物的价值关系的目标指向、

变化方式、变化时态、对方的利益相关性等进行判断，从而确定和选择情感表达的某种基本模式。

### ● 生理模式识别

从生理信号中抽取出来的特征模式可以用来识别情感。计算机在人做出表情（如悲痛或愤怒）的时候，观察多种收集到的信号，然后分析哪种生理信号模式和特定情感状态关系最密切。然后计算机系统应用先前分析的结果，根据收集到的原始数据来识别出包含在信号中最有可能的情感。这方面的研究才刚刚开始。通常通过观察所有数据的子集能判断出哪种特征值的辨别能力最强。试过所有可能的三种情感、两种特征值的组合，发现愤怒、悲痛和尊敬的组合或愤怒、高兴和尊敬的组合中，情感是最容易识别出的。这两种情况中，都有一个最佳的识别特征值，即 EMG 信号的平均值，但是另一个最佳特征值却是不同的。愤怒、悲痛和尊敬的组合中，标准化信号的第一次差分绝对值的平均值就是最佳的；而对愤怒、高兴和尊敬的组合中，最佳选择还是第一次差分绝对值的平均值，但这次是从 EMG 信号中得到的。

### 1.2.2.3 情感信号的融合算法研究

某一情感的变化对应于多种生理或行为的变化，不能从单一生理或行为的变化引出相应的情感的变化，因此，情感的判定是融合各种生理或行为的过程，应该研究有关情感计算的信息融合理论与技术。

信息融合又称数据融合，也可以称为传感器信息融合或多传感器信息融合，是一个对从单个和多个信息源获取的数据和信息进行关联、相关和综合，以获得精确的位置和身份估计，以及对态势和威胁及其重要程度进行全面及时评估的信息处理过程；该过程是对其估计、评估和额外信息源需求评价的一个持续精练（Refinement）过程，同时也是信息处理过程不断自我修正的一个过程，以获得结果的改善。

信息融合技术可概括为：利用计算机技术对按时序获得的若干传感器的观测信息在一定准则下加以自动分析、综合处理，以完成所需的决策和估计任务而进行的信息处理过程。按照这一定义，多传感器系统是信息融合的硬件基础，多源信息是信息融合的加工对象，协调优化和综合处理是信息融合的核心。也可以认为，信息融合或数据融合技术是利用计算机技术对来自多传感器（同类或者不同类）探测的多源信息，按一定规则进行自动分析和综合后，自动生成人们所期望的合成信息的信息处理技术。它对多类型、多源、多平台传感器所获得的各种情报信息（如数据、照片、视频图像等信息）进行采集、传输、汇集、分析、过滤、综合、相关及合成，快速进行情报处理和自动图形标绘。



## 1.3 情感计算的理论观点

心理学领域有关情感体验的理论观点主要有两类：范畴观和维度观。

持范畴观的学者将情绪或情感分成独立的范畴，这些相互离散的情绪范畴在外部表现、生理唤醒模式上都存在一定的差异<sup>[2]</sup>。对于情感计算研究，基于范畴观来描述和测量情感的优势主要有两点。首先，用情绪范畴描述情绪和情感符合人们的直觉和常识，有利于情感计算的成果在现实生活中的推广和应用。其次，基于情绪范畴来进行情感计算，有利于智能系统中识别情绪后，进一步推理与之相联系的特定心理功能和可能的原因，然后做出适当的响应。但范畴观也给情感计算研究带来了制约和问题。首先，哪些情绪范畴对于情感计算来说是必要的，目前研究者对此并没有统一的认识。其次，情绪范畴是对情绪的定性描述，无法用量化的数字表达主观的情绪体验。

维度观则认为情感具有基本维度和两极性。例如，近代科学心理学之父威廉·冯特提出的三维说，认为情绪情感有三个维度组成，每个维度的变化幅度都存在两种相互独立的状态。Osgood 发现，情绪体验可以从评价、力度、活跃性三个方面来进行评价<sup>[3]</sup>。Mehrabian 和 Russell 基于 Osgood 的思路，提出了 PAD 情绪模型，此模型不但给出了对情感空间进行描述的理论构想，同时采用量化的方法试图建立情感空间中各种情绪范畴的定位和关系，因此被广泛地应用在情绪心理学、人格心理学、社会心理学等基础领域，以及市场营销、产品满意度等应用领域<sup>[4]</sup>。基于维度的趋向为情感计算提供了量化的理论和方法基础，但是，究竟需要多少维度才足以充分地表达人类的情感及哪些维度是必要的，这仍是一个难题。

## 1.4 情感计算的认知困境

情感计算一经提出，迅速受到学术界的关注和企业界的迅速反应。英国电信公司（British Telecom）和 IBM 已成立了专门的情感计算研究小组，学术界如美国麻省理工学院的媒体实验室、剑桥大学、中国科学院等也在进行情感计算方面的研究。虽然情感计算的研究如火如荼地进行着，并且也取得了很大的成果，但对于情感计算的一些基本问题在认识上仍然存在分歧<sup>[6]</sup>。

### 1.4.1 情感计算可能性

情感可不可以计算实际上是一个哲学终极问题，这个问题衍生出如下问题：情感能计算吗？如果能计算，应该如何计算？计算的方法可行、可靠吗？对此明显分为两种截然相反的观点：

一种是对情感计算的怀疑与否定，认为情感计算是不可能的和没必要的，主要体现在：

1) 认为情感计算在理论上是不可能的。情绪是人类本身专有的，一旦它脱离人就不存在了或者说变成假的了，一台机器不可能自发地产生那些根本不属于它的特性；情感是不能制造的，模拟永远是不真实的，机器是永远不会产生像人与人相互交流时的那种情感。神经科学家达马西奥研究一批大脑受损的病人，他们尽管在数学和语言能力上没有任何缺陷，但他们的日常生活能力却大受影响。因此，长于数学和语言能力的机器难以处理情感。

2) 认为情感计算在研究方向与方法是不可行的。情绪表达在模式与方法上是非常广泛的，而对这些是不可能进行准确把握的，如大脑活动、血液化学、神经传递物质等，因此在收集必要的方面变得不太可能或者不可行。并且，情绪比较随意和时常变化，要从一些数据准确地识别个体的情感状态也是办不到的。另一方面，要对情感进行计算，那么就必须要建立情感认知模型，但是要建立精细、多模态以及适合个体习惯的情感模型是非常困难的，至少在以后相当长的时间内是不可能的。

但以 R.W Picard 为代表的计算机学家们却认为机器具有情感是可能的，在方法上也是可行的。他们认为情感并非如心理学家所认为的那样仅仅是人的内心体验状态，而是可以通过人生理、表情、语言等体现出来，可以通过一定的技术进行测量与处理。虽然人的情感表现具有一定的差异性，但是通过他们使用的可穿戴计算机的实验却发现情感识别能够达到 80% 左右的准确性。另一方面他们认为建立在实验基础上的情感建模也取得了不少进步，很多实验室与研究所根据自己的研究都建立了比较有价值的情感模型，如美国麻省理工学院媒体实验室首先给出了一个隐马尔可夫模型、日本早稻田大学建立了 WE-3RV 的 3D 情感空间模型、计算机学家 Sumedha Kshirsagar 等建立的多层情绪—心情—性格模型等。

## 1.4.2 情感计算的理论基础

无论是情感计算的开创者 R.W Picard，还是后来的计算机学家对情感的计算处理，他们的理论是建立在心理学基础之上的。心理学把情感分为快乐、焦虑、悲伤、愤怒和厌恶，其它情感都是由此派生出来的；把情感的表现局限于心理感受强度、表情特征和生理指标。当前建立起来的情感模型主要考虑的也是这五种基本情感（见前面所论述的情感模型），建模方法大都是基于维度论的方法，通过测量生理特征和行为特征（如面部表情、语音等）来推测情感状态，但是这样的模型总存在着这样或者那样的不足；同时，对于同一类型情感，无论是情感感受强度，还是情感表达强度和情感生理唤醒，都可以采用不同的生理指标进行计算和测量，究竟应该选用哪一个生理指标？对于同一类型情感采用同一个生理指标进行测量和计算时，由于受到许多环境因素以及人体其它生理因素和精神因素的影响，其测量值的差异性和波动性如何消除？因此，一些人认为把情感计算的理论基础建立在心理学之上是不妥的，而应该把它的理论基础

建立在“统一价值论”之上。统一价值论认为任何主观意识都不是无中生有的，也不是孤立存在的，都是对某一客观存在的反映，总会有某一客观存在与之相对应；情感作为一种特殊的主观心理活动，反映了一种特殊的客观存在——人与事物之间的价值关系，情感与价值的关系在本质上就是主观与客观的关系。因此，统一价值论把情感诠释为“人脑对事物的价值关系的一种主观反映”，以数学形式对情感进行精确定义，建立情感和价值观的数学分析模型，实现了情感的合并运算和合成运算，根据价值的不同变化特征对情感进行分类，从而将“情感计算”转化为“价值计算”。

### 1.4.3 情感计算技术的两面性

人工智能与计算机领域对此持乐观的态度，为我们展现了在不久的将来机器具有情感的广阔前景。中国科学院心理研究所傅小兰研究员认为“情感计算研究将为人类提供更加人性化、情感化的服务和产品，创造更加美好的生活”，清华大学计算机系张迎辉等相信“情感计算的发展必将极大地改变人们的生活，自然和谐的人机交互的梦想会离我们越来越近”，情感计算将在如下几个方面产生深刻影响：

- 对人类的情感进行识别和响应，帮助人们理解自己和他人的情感世界。
- 在远程教育平台中，进一步优化计算机辅助学习的功能。
- 推进可穿戴的计算机系统的发展，以及智能、便携式个人身体保健与监护系统的开发。
- 帮助我们增加使用设备的安全性和增加自动感知人们的情绪状态的功能，可以在娱乐、游戏等方面提供更好的服务。

但是在社会与伦理领域的人士看来，如果机器具有像人一样的情感，将会对人类以及人类社会带来巨大的影响。首先，人与人之间的两维关系将被彻底打破，如何处理人与机器的感情问题，这将会产生一系列的伦理道德问题。其次，社会结构与秩序将会重组，人与机器在社会关系中的角色如何维持与平衡。因此他们认为机器具有了情感带给人类的不仅仅是福音，还带来麻烦与问题，并且麻烦与问题多于好处。

## 2 技术篇

情感计算是一个多学科交叉的崭新的研究领域，它涵盖了传感器技术、计算机科学、认知科学、心理学、行为学、生理学、哲学、社会学等方面。情感计算的最终目标是赋予计算机类似于人的情感能力。要达到这个目标，许多技术问题有待解决。这些技术问题的突破对各学科的发展都产生巨大的推动作用。本篇分别从情感计算的傳統研究方法和新兴研究方法对技术发展进行了探讨。此外，本篇对情感计算方向较有代表性的获奖论文进行了归纳整理，更加清晰地展示了目前该领域的研究成果。

### 2.1 传统的研究

传统的情感计算方法是按照不同的情感表现形式分类的，分别是：文本情感分析、语音情感分析、视觉情感分析。以下对三种传统的分析方法进行详细介绍。

#### 2.1.1 文本情感计算

20世纪90年代末，国外的文本情感分析已经开始。早期，Riloff和Shepherd在文本数据的基础上进行了构建语义词典的相关研究。McKeown发现连词对大规模的文本数据集中形容词的语义表达的制约作用，进而对英文的形容词与连词做情感倾向研究。自此之后，越来越多的研究开始考虑特征词与情感词的关联关系。Turney等使用点互信息的方法扩展了正负面情感词典，在分析文本情感时使用了极性语义算法，处理通用的语料数据时准确率达到了74%。在近些年的研究中，Narayanan等结合各种特征及其相关联信息，提出了基于分句、整句、结果句的分类方案，获得了很好的效果。Pang等以积极情感和消极情感为维度，对电影评论进行了情感分类。他分别采用了支持向量机、最大熵、朴素贝叶斯算法进行分类实验，发现支持向量机的精确度达到了80%。随着研究的不断深入，学者在对情感分析算法进行改进的同时，也将其应用到不同的行业中进行了实践。

文本情感计算的过程可以由3部分组成：文本信息采集、情感特征提取和情感信息分类。文本信息采集模块通过文本抓取工具（如网页爬虫工具）获得情感评论文本，并传递到下一个情感特征提取模块，然后对文本中自然语言文本转化成计算机能够识别和处理的形式，并通过情感信息分类模块得到计算结果。文本情感计算侧重研究情感状态与文本信息之间的对应关系，提供人类情感状态的线索。具体地，需要找到计算机能提取出来的特征，并采用能用于情感分类的模型。因此，关于文本情感计算过程的讨论，主要集中在文本情感特征标注（信息采集）、情感特征提取和情感信息分类这三个方面<sup>[7]</sup>。



## ● 文本情感特征标注

情感特征标注是对情感语义特征进行标注，通常是将词或者语义块作为特征项。情感特征标注首先对情感语义特征的属性进行设计，如褒义词、贬义词、加强语气、一般语气、悲伤、高兴等等；然后通过机器自动标注或者人工标注的方法对情感语义特征进行标注，形成情感特征集合。情感词典是典型的情感特征集合，也是情感计算的基础。在大多数研究中，有关情感计算的研究通常是将情感词典直接引入自定义词典中。

运用情感词典计算出文本情感值是一种简单迅速的方法，但准确率有待提高。在实际的情感计算中，会因为具体的语言应用环境而有所不同。例如，“轻薄”一词通常认为是否定词，但是在电脑、手机却被视为肯定词汇。同时，文本中常会出现否定前置、双重否定以及文本口语化和表情使用等，这些都将会对文本情感特征的提取和判断产生较大的影响。因此在进行文本情感提取时，需要对文本及其对应的上下文关系、环境关系等进行分析。

## ● 情感特征提取

文本包含的情感信息是错综复杂的，在赋予计算机以识别文本情感能力的研究中，从文本信号中抽取特征模式至关重要。在对文本预处理后，初始提取情感语义特征项。特征提取的基本思想是根据得到的文本数据，决定哪些特征能够给出最好的情感辨识。通常算法是对已有的情绪特征词打分，接着以得分高低为序，超过一定阈值的特征组成特征子集。特征词集的质量直接影响最后结果，为了提高计算的准确性，文本的特征提取算法研究将继续受到关注。长远看来，自动生成文本特征技术将进一步提高，特征提取的研究重点也更多地从对词频的特征分析转移到文本结构和情感词上。

## ● 情感信息分类

文本情感分类技术中，主要采用两种技术路线：基于规则的方法和基于统计的方法。在 20 世纪 80 年代，基于规则的方法占据主流位置，通过语言学家的语言经验和知识获取句法规则，以此作为文本分类依据。但是，获取规则的过程复杂且成本巨大，也对系统的性能有负面影响，且很难找到有效的途径来提高开发规则的效率。20 世纪 90 年代之后，人们更倾向于使用统计的方法，通过训练样本进行特征选择和参数训练，根据选择的特征对待分类的输入样本进行形式化，然后输入到分类器进行类别判定，最终得到输入样本的类别。

下表整理了 10 个文本情感分析 API，供读者参考：

表 1 十个文本情感分析 API<sup>[8]</sup>

名称	简介	地址
<b>IBM Watson Tone Analyzer API</b>	IBM 研发的一款先进的机器学习模型，可以检测文本中的交流语气，检测结果可以划分为情感语气、语言语气和社交语气三类。	<a href="https://cloud.ibm.com/apidocs/tone-analyzer">https://cloud.ibm.com/apidocs/tone-analyzer</a>
<b>Qemotion</b>	该 API 能检测以下内容：演讲中表达的情绪指数、文本中表达的关键和主要情绪、文本情感自动化分析。用该 API 提交文本后，算法就会检测文本所传递的主要情感。	<a href="https://rapidapi.com/Qemotion/api/text-to-emotions/details?utm_source=mashape&amp;utm_medium=301">https://rapidapi.com/Qemotion/api/text-to-emotions/details?utm_source=mashape&amp;utm_medium=301</a>
<b>AYLIEN API</b>	一款自然语言处理工具包，能够从文本和视觉内容中提取语义信息，能够完成情感分析、情感标记、语言检测等 15 种 NLP 任务。	<a href="https://docs.aylien.com/textapi/using-the-api">https://docs.aylien.com/textapi/using-the-api</a>
<b>PreCeive API</b>	一款文本处理工具，能执行情感分析、情绪分析等工作。	<a href="https://theysay.docs.apiary.io/#introduction">https://theysay.docs.apiary.io/#introduction</a>
<b>Indico API</b>	能够从文本中检测到生气、高兴、恐惧、悲伤和惊讶等 5 种情感，而且能将检测结果输出为整洁的 Json 文件。	<a href="https://indico.io/docs/#emotion">https://indico.io/docs/#emotion</a>
<b>Microsoft Cognitive Service-Text Analytics</b>	微软开发的文本分析工具，可完成情感分析、关键词提取、语言检测等工作。	<a href="https://docs.microsoft.com/en-gb/azure/cognitive-services/Text-Analytics/overview">https://docs.microsoft.com/en-gb/azure/cognitive-services/Text-Analytics/overview</a>
<b>Bitext</b>	提供市场上基于情绪的最精准的多语言话题。目前有 4 种语义服务：实体及概念提取，情感分析及文本分类。API 可以处理 8 种语言。	<a href="https://www.bitext.com/api/#How-accurate-is-the-analysis">https://www.bitext.com/api/#How-accurate-is-the-analysis</a>
<b>Google Cloud Natural Language API</b>	谷歌的文本处理工具，可分析文本数据的语义和结构，包括情感分析、实体识别和文本注释。	<a href="https://cloud.google.com/natural-language/docs/reference/rest/">https://cloud.google.com/natural-language/docs/reference/rest/</a>
<b>NlpTools</b>	一个用于自然语言处理的简单 API，可以对在线新闻媒体进行情感分析和文本分类。	<a href="http://nlptools.atrilla.net/web/api.php">http://nlptools.atrilla.net/web/api.php</a>
<b>Sentigem</b>	面向英语文档和文本块的情感分析工具，简单易用。	<a href="https://sentigem.com/#!">https://sentigem.com/#!</a>

## 2.1.2 语音情感计算

最早真正意义上的语音情感识别相关研究出现在 20 世纪 80 年代中期，它们开创了使用声学统计特征进行情感分类的先河。紧接着，随着 1985 年 Minsky 教授“让计算机具有情感能力”观点的提出，以及人工智能领域的研究者们对情感智能重要性认识的日益加深，越来越多的科研机构开始了语音情感识别研究的探索。在 20 世纪 80 年代末至 90 年代初期，麻省理工学院多媒体实验室构造了一个“情感编辑器”对外界各种情感信号进行采集，综合使用人体的生

理信号、面部表情信号、语音信号来初步识别各种情感，并让机器对各种情感做出适当的简单反应；1999 年，Moriyama 提出语音和情感之间的线性关联模型，并据此在电子商务系统中建造出能够识别用户情感的图像采集系统语音界面，实现了语音情感在电子商务中的初步应用。整体而言，语音情感识别研究在该时期仍旧处于初级阶段，主要侧重于情感的声学特征分析这一方面，作为研究对象的情感语音样本也多表现为规模小、自然度低、语义简单等特点，虽然有相当数量的有价值的研究成果相继发表，但是并没有形成一套被广泛认可的、系统的理论和研究方法。进入 21 世纪以来，随着计算机多媒体信息处理技术等研究领域的出现以及人工智能领域的快速发展，语音情感识别研究被赋予了更多的迫切要求，发展步伐逐步加快。2000 年，在爱尔兰召开的 ISCA Workshop on Speech and Emotion 国际会议首次把致力于情感 and 语音研究的学者聚集在一起。近 10 余年来，语音情感识别研究工作在情感描述模型的引入、情感语音库的构建、情感特征分析等领域的各个方面都得到了发展。下面将从语音情感数据库的采集、语音情感标注以及情感声学特征分析方面介绍语音情感计算。

### ● 语音情感数据库的采集

语音情感识别研究的开展离不开情感语音数据库的支撑。情感语音库的质量高低，直接决定了由它训练得到的情感识别系统的性能好坏。评价一个语音情感数据库好坏的一个重要标准是数据库中语音情感是否具备真实的表露性和自发性。目前，依据语音情感激发类型的不同，语音情感数据库可分为表演型、诱发型和自发型三种。

具体来说，表演型情感数据库通过专业演员的表演，把不同情感表达出来。在语音情感识别研究初期，这一采集标准被认为是研究语音情感识别比较可靠的数据来源，因为专业演员在表达情感时，可以通过专业表达获得人所共知的情感特征。比如，愤怒情感的语音一般会具有很大的幅值和强度，而悲伤情感的语音则反之。由于这一类型的数据库具有表演的性质，情感的表达会比真实情感夸大一点，因此情感不具有自发的特点。依据该类型数据库来学习的语音情感识别算法，不一定能有效应用于真实生活场景中。第二种称之为诱发型情感数据库。被试者处于某一特定的环境，如实验室中，通过观看电影或进行计算机游戏等方式，诱发被试者的某种情感。目前大部分的情感数据库都是基于诱发的方式建立的。诱发型情感数据库产生的情感方式相较于表演型情感数据库，其情感特征更具有真实性。最后一种类型属于完全自发的语音情感数据库，其语料采集于电话会议、电影或者电话的视频片段，或者广播中的新闻片段等等。由于这种类型的语音情感数据最具有完全的真实性和自发性，应该说最适合用于实用的语音情感识别。但是，由于这些语音数据涉及道德和版权因素，妨碍了它在实际语音情感识别中的应用。

## ● 语音情感数据库的标注

对于采集好的语音情感库,为了进行语音情感识别算法研究,还需要对情感语料进行标注。标注方法有两种类型:

离散型情感标注法指的是标注为如生气、高兴、悲伤、害怕、惊奇、讨厌和中性等,这种标注的依据是心理学的基本情感理论。基本情感论认为,人复杂的情感是由若干种有限的基本情感构成的,就像我们自古就有“喜、怒、哀、乐、恐、悲、惊”七情的说法。不同的心理学家对基本情感有不同的定义,由此可见,在心理学领域对基本情感类别的定义还没有一个统一的结论,因此不同的语音情感数据库包含的情感类别也不尽相同。这不利于在不同的语音情感数据库上,对同一语音情感识别算法的性能进行评价。此外,众所周知,实际生活中情感的类别远远不止有限几类。基于离散型情感标注法的语音情感识别容易满足多数场合的需要,但无法处理人类情感表达具有连续性和动态变化性的情况。在实际生活中,普遍存在着情感变化的语音,比如前半句包含了某一种情感,而后半句却包含了另外一种情感,甚至可能相反。例如,某人说话时刚开始很高兴,突然受到外界刺激,一下子就生气了。对于这种在情感表达上具有连续和动态变化的语音,采用离散型情感标注法来进行语音情感识别就不合适了。因为此时语音的情感,已不再完全属于某一种具体的情感。

维度情感空间论基于离散型情感标注法的缺陷,心理学家们又提出了维度情感空间论,即对情感的变化用连续的数值进行表示。不同研究者所定义的情感维度空间数目有所不同,如二维、三维甚至四维模型。针对语音情感,最广为接受和得到较多应用的为二维连续情感空间模型,即“激活维—效价维”(Arousal-Valence)的维度模型。“激活维”反映的是说话者生理上的激励程度或者采取某种行动所作的准备,是主动的还是被动的;“效价维”反映的是说话者对某一事物正面的或负面的评价。随着多模态情感识别算法的研究,为了更细致地描述情感的变化,研究者在“激活维—效价维”(Arousal-Valence)二维连续情感空间模型的基础上,引入“控制维”,即在“激活维—效价维—控制维(Arousal-Valence/Pleasure-Power/Dominance)”三维连续情感空间模型上对语音情感进行标注和情感计算。需要强调的是,离散型和连续型情感标注之间,它们并不是孤立的,而是可以通过一定映射进行相互转换。

## ● 情感声学特征分析

情感声学特征分析主要包括声学特征提取和声学特征选择、声学特征降维。采用何种有效的语音情感特征参数用于情感识别,是语音情感识别研究最关键的问题之一,因为所用的情感特征参数的优劣直接决定情感最终识别结果的好坏<sup>[9]</sup>。



**声学特征提取。**目前经常提取的语音情感声学特征参数主要有三种：韵律特征、音质特征以及谱特征。在早期的语音情感识别研究文献中，针对情感识别所首选的声学特征参数是韵律特征，如基音频率、振幅、发音持续时间、语速等。这些韵律特征能够体现说话人的部分情感信息，较大程度上能区分不同的情感。因此，韵律特征已成为当前语音情感识别中使用最广泛并且必不可少的一种声学特征参数除了韵律特征，另外一种常用的声学特征参数是与发音方式相关的音质特征参数。三维情感空间模型中的“激发维”上比较接近的情感类型，如生气和高兴，仅使用韵律特征来识别是不够的。音质特征包括共振峰、频谱能量分布、谐波噪声比等，不仅能够很好地表达三维中的“效价维”信息，而且也能够部分反映三维中的“控制维”信息。因此，为了更好地识别情感，同时提取韵律特征和音质特征两方面的参数用于情感识别，已成为语音情感识别领域声学特征提取的一个主要方向。谱特征参数是一种能够反映语音信号的短时功率谱特性的声学特征参数，Mel 频率倒谱系数（Mel-scale Frequency Cepstral Coefficients, MFCC）是最具代表性的谱特征参数，被广泛应用于语音情感识别。由于谱特征参数及其导数，仅反映语音信号的短时特性，忽略了对情感识别有用的语音信号的全局动态信息。近年来，为了克服谱特征参数的这种不足之处，研究者提出了一些改进的谱特征参数，如类层次的谱特征、调制的谱特征和基于共振峰位置的加权谱特征等。

**声学特征选择。**为了尽量保留对情感识别有意义的信息，研究者通常都提取了较多的与情感表达相关的不同类型的特征参数，如韵律特征、音质特征、谱特征等。任意类型特征都有各自的侧重点和适用范围，不同的特征之间也具有一定的互补性、相关性。此外，这些大量提取的特征参数直接构成了一个高维空间的特征向量。这种高维性质的特征空间，不仅包含冗余的特征信息，导致用于情感识别的分类器训练和测试需要付出高昂的计算代价，而且情感识别的性能也不尽如人意。因此，非常有必要对声学特征参数进行特征选择或特征降维处理，以便获取最佳的特征子集，降低分类系统的复杂性和提高情感识别的性能。

特征选择是指从一组给定的特征集中，按照某一准则选择出一组具有良好区分特性的特征子集。特征选择方法主要有两种类型：封装式（Wrapper）和过滤式（Filter）。Wrapper 算法是将后续采用的分类算法的结果作为特征子集评价准则的一部分，根据算法生成规则的分类精度选择特征子集。Filter 算法是将特征选择作为一个预处理过程，直接利用数据的内在特性对选取的特征子集进行评价，独立于分类算法。

**声学特征降维。**特征降维是指通过映射或变换方式将高维特征空间映射到低维特征空间，已达到降维的目的。特征降维算法分为线性和非线性两种。最具代表性的两种线性降维算法，如主成分分析 PCA（Principal Component Analysis）和线性判别分析 LDA（Linear Discriminant Analysis），已经被广泛用于对语音情感特征参数的线性降维处理。也就是，PCA 和 LDA 方

法被用来对提取的高维情感声学特征数据进行嵌入到一个低维特征子空间，然后在这降维后的低维子空间实现情感识别，提高情感识别性能。近年来，新发展起来的基于人类认知机理的流形学习方法比传统的线性 PCA 和 LDA 方法更能体现事物的本质，更适合于处理呈非线性流形结构的语音情感特征数据。但这些原始的流形学习方法直接应用于语音情感识别中的特征降维，所取得的性能并不令人满意。主要原因是他们都属于非监督式学习方法，没有考虑对分类有帮助的已经样本数据的类别信息。尽管流形学习方法能够较好地处理非线性流形结构的语音特征数据，但是流形学习方法的性能容易受到其参数如邻域数的影响，而如何确定其最佳的邻域数，至今还缺乏理论指导，一般都是根据样本数据的多次试验结果来粗略地确定。因此，对于流形学习方法的使用，如何确定其最佳参数，还有待深入研究。

下表为常用语音情感数据库，供读者参考：

表 2 常用的语音情感数据库<sup>[10]</sup>

数据库名称	简介
<b>Belfast 英语情感数据库</b>	Belfast 情感数据库由 Queen 大学的 Cowie 和 Cowie 录制，由 40 位录音人（18 岁~69 岁，20 男 20 女）对 5 个段落进行演绎得到。每个段落包含 7~8 个句子，且具有某种特定的情感倾向，分别为生气/anger、悲伤/sadness、高兴/happiness、恐惧/fear、中性/neutral。
<b>柏林 EMO-DB 德语情感语音库</b>	DMO-DB 是由柏林工业大学录制的德语情感语音库，由 10 位演员（5 男 5 女）对 10 个语句（5 长 5 短）进行 7 种情感（中性/neutral、生气/anger、害怕/fear、高兴/joy、悲伤/sadness、厌恶/disgust、无聊/boredom）的模拟得到，共包含 800 句语料，采样率 48kHz（后压缩到 16kHz），16bit 量化。语料文本的选取遵从语义中性、无情感倾向的原则，且为日常口语化风格，无过多的书面语修饰语音的录制在专业录音室中完成。经过 20 个参与者（10 男 10 女）的听辨实验，得到 84.3% 的听辨识别率。
<b>FAU AIBO 儿童德语情感语音库</b>	FAU AIBO 录制了 51 名儿童（10 岁~13 岁，21 男 30 女）在与索尼公司生产的电子宠物 AIBO 游戏过程中的自然语音，并且只保留了情感信息明显的语料，总时长为 9.2 小时（不包括停顿），包括 48401 个单词，由 DAT-recorder 录制，48kHz 采样（而后压缩到 16kHz），16bit 量化。为了记录真实情感的语音，工作人员让孩子们相信 AIBO 能够对他们的口头命令加以反应和执行，而实际上，AIBO 则是由工作人员暗中人为操控的。标注共涵盖包括 joyful、irritated、angry、neutral 等在内的 11 个情感标签。该数据库中的 18216 个单词被选定为 INTERSPEECH 2009 年情感识别竞赛用数据库。
<b>CASIA 汉语情感语料库</b>	该数据库由中国科学院自动化研究所录制，由 4 位录音人（2 男 2 女）在纯净录音环境下（信噪比约为 35db）分别在 5 类不同情感下（高兴、悲哀、生气、惊吓、中性）对 500 句文本进行的演绎得到，16kHz 采样，16bit 量化，经过听辨筛选，最终保留其中 9600 句。
<b>ACCorpus 系列汉语情感数据库</b>	该系列情感数据库由清华大学和中国科学院心理研究所合作录制，包含 5 个相关子库： 1) ACCorpus_MM 多模态、多通道的情感数据库；

	<p>2) ACCorpus_SR 情感语音识别数据库;</p> <p>3) ACCorpus_SA 汉语普通话情感分析数据库;</p> <p>4) ACCorpus_FV 人脸表情视频数据库;</p> <p>5) ACCorpus_FI 人脸表情图像数据库。</p> <p>其中, ACCorpus_SR 子库共由 50 位录音人 (25 男 25 女) 对 5 类情感 (中性、高兴、生气、恐惧和悲伤) 演绎得到, 16kHz 采样, 16bit 量化。每个发音者的数据均包含语音情感段落和语音情感命令两种类型。</p>
<b>VAM 数据库</b>	<p>VAM 数据库是一个以科学研究为目的的无偿数据库, 通过对一个德语电视谈话节目“Vera am 美国麻省理工学院 tag”的现场录制得到, 数据库包含语料库、视频库、表情库这 3 个部分。谈话内容均为无脚本限制、无情绪引导的纯自然交流。以 VAM-audio 库为例, 该子库包含来自 47 位节目嘉宾的录音数据 947 句, wav 格式, 16kHz 采样, 16bit 量化。所有数据以句子为单位进行保存 (1018 句), 标注在 Valence、Activation 和 Dominance 这 3 个情感维度上进行, 标注值处于 -1~1 之间, 最终的情感值是相关标注者的平均值。</p>
<b>Semaine 数据库</b>	<p>Semaine 数据库是一个面向自然人机交互和人工智能研究的数据库 (<a href="http://semaine-db.eu/">http://semaine-db.eu/</a>)。数据录制在人机交互的场景下进行, 20 个用户 (22 岁~60 岁, 8 男 12 女) 被要求与性格迥异的 4 个机器角色进行交谈 (实际上, 机器角色由工作人员扮演)。这 4 个角色分别是:</p> <ol style="list-style-type: none"> <li>1) 温和而智慧的 Prudence;</li> <li>2) 快乐而外向的 Poppy;</li> <li>3) 怒气冲冲的 Spike;</li> <li>4) 悲伤而抑郁的 Obadiah。</li> </ol> <p>录音过程在专业配置录音室内进行, 同时有 5 个高分辨率、高帧频摄像机和 4 个麦克风进行数据的收集, 其中, 音频属性为 48kHz 采样, 24bit 量化, 数据时长在 7 小时左右。标注工作由多个参与者借助标注工具 FEELTRACE 在 Valence、Activation、Power、Expectation 和 Intensity 这 5 个情感维度上进行。该数据库中的部分数据被用于 AVEC 2012 的竞赛数据库。</p>
<b>Maribor 数据库</b>	<p>该数据库由斯洛文尼亚 Maribor 大学录制, 包含 8 类情感语料 (厌恶/disgust、惊奇/surprise、高兴/joy、恐惧/fear、生气/anger、悲伤/sadness、快高中性/fast loud neutral、低慢中性/low soft neutral), 通过 4 种语言 (英语、斯洛文尼亚语、法语、西班牙语) 演绎, 每类情感各有 186 句语料。</p>
<b>Kids' Audio Speech Corpus NSF/ITR Reading Project</b>	<p>该数据库由科罗拉多大学的 Cole 教授及其助理录制, 旨在从儿童那里收集足够的音频和视频数据, 以便实现听觉和视觉识别系统的发展, 使与电子教师面对面的对话互动成为可能。</p>
<b>Emotional Prosody Speech and Transcripts</b>	<p>该数据库由宾夕法尼亚大学录制, 包含 15 类情感语料, 例如 s hot anger, cold anger, panic-anxiety, despair, sadness, elation, happiness, interest, boredom, shame, pride, disgust and contempt</p>

### 2.1.3 视觉情感计算

表情作为人类情感表达的主要方式, 其中蕴含了大量有关内心情感变化的信息, 通过面部表情可以推断内心微妙的情感状态。但是让计算机读懂人类面部表情并非简单的事情。人脸表情识别是人类视觉最杰出的能力之一。而计算机进行自动人脸表情识别所利用的主要也是视觉

数据。无论在识别准确性、速度、可靠性还是稳健性方面，人类自身的人脸表情识别能力都远远高于基于计算机的自动人脸表情识别。因此，自动人脸表情识别研究的进展一方面依赖计算机视觉、模式识别、人工智能等学科的发展，另一方面还依赖对人类本身识别系统的认识程度，特别是对人的视觉系统的认识程度。

早在 20 世纪 70 年代，关于人脸表情识别的研究就已经展开，但是早期主要集中在心理学和生物学方面。随着计算机技术的发展，人脸表情识别技术逐渐发展起来，至上世纪 90 年代，该领域的研究已经非常活跃。大量文献显示表情识别与情感分析已从原来的二维图像走向了三维数据研究，从静态图像识别研究专项实时视频跟踪。下面将从视觉情感信号获取、情感信号识别以及情感理解与表达方面介绍视觉情感计算。

### ● 视觉情感信号获取

表情参数的获取，多以二维静态或序列图像为对象，对微笑的表情变化难以判断，导致情感表达的表现力难以提高，同时无法体现人的个性化特征，这也是表情识别中的一大难点。以目前的技术，在不同的光照条件和不同头部姿态下，也不能取得满意的参数提取效果。

由于三维图像比二维图像包含更多的信息量，可以提供鲁棒性更强，与光照条件和人的头部姿态无关的信息，用于人脸表情识别的特征提取工作更容易进行。因此，目前最新的研究大多利用多元图像数据来进行细微表情参数的捕获。该方法综合利用三维深度图像和二维彩色图像，通过对特征区深度特征和纹理彩色特征的分析 and 融合，提取细微表情特征，并建立人脸的三维模型，以及细微表情变化的描述机制。

### ● 视觉情感信号识别

视觉情感信号的识别和分析主要分为面部表情的识别和手势识别两类：

**对于面部表情的识别**，要求计算机具有类似于第三方观察者一样的情感识别能力。由于面部表情是最容易控制的一种，所以识别出来的并不一定是真正的情感，但是，也正由于它是可视的，所以它非常重要，并能通过观察它来了解一个人试图表达的东西。到目前为止，面部表情识别模型都是将情感视为离散的，即将面部表情分成为数不多的类别，例如“高兴”、“悲伤”、“愤怒”等。1971 年，Ekman 和 Friesen 研究了 6 种基本表情（高兴、悲伤、惊讶、恐惧、愤怒和厌恶），并系统地建立了上千幅不同的人脸表情图像库。六种基本表情的具体面部表现如表 3 所示。1978 年，他们研究了情感类别之间的内在关系，开发了面部动作编码系统（FACS）。系统描述了基本情感以及对应的产生这种情感的肌肉移动的动作单元。他们根据人脸的解剖学特点，将其划分成大约 46 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及其所控制的主要区域以及与之相关的表情，给出了大量的照片说明。面部识别器一般要花五分钟来处理一种面部表情，准确率达到 98%。

马里兰大学的 Yeser Yacoob 和 Larry Davis 提出了另一种面部表情识别模型，它也是基于动作能量模版，但是将模版、子模版（例如嘴部区域）和一些规则结合起来表达情感。例如，愤怒的表情在从眼睛区域提取的子模版中，特别是眉毛内敛、下垂，在嘴巴区域子模版中，特别是嘴巴紧闭，两个子模板结合起来，就很好表达了愤怒这一情感。后续的研究总体上结合生物识别方法及计算机视觉进行，依据人脸特定的生物特征，将各种表情同脸部运动细节（几何网格的变化）联系起来，收集样本，提取特征，构建分类器。但是目前公开的用于表情识别研究的人脸图像数据库多是采集志愿者刻意表现出的各种表情的图像，与真实情形有出入。

表 3 脸部表情运动特征具体表现

表情	额头、眉毛	眼睛	脸的下半部
惊奇	1) 眉毛抬起，变高变弯 2) 眉毛下的皮肤被拉伸 3) 皱纹可能横跨额头	1) 眼睛睁大，上眼皮抬高，下眼皮下落 2) 眼白可能在瞳孔的上边和/或下边露出来	下颌下落，嘴张开，唇和齿分开，但嘴部不紧张，也不拉伸
恐惧	1) 眉毛抬起并皱在一起 2) 额头的皱纹只集中在中部，而不横跨整个额头	上眼睑抬起，下眼皮拉近	嘴张，嘴唇或轻微紧张，向后拉；或拉长，同时向后拉
厌恶	眉毛压低，并压低上眼睑	在下眼皮下部出现横纹，脸颊推动其向上，并不紧张	1) 上唇抬起 2) 下唇与上唇紧闭，推动上唇向上，嘴角下拉，嘴唇微凸起 3) 鼻子皱起 4) 脸颊抬起
愤怒	1) 眉毛皱在一起，压低 2) 在眉宇间出现竖直皱纹	1) 下眼皮拉紧，抬起或不太起 2) 上眼皮拉紧，眉毛压低 3) 眼睛瞪大，可能鼓起	1) 唇有两种基本的位置：紧闭，嘴角拉直或向下，张开，仿佛要喊 2) 鼻孔可能张大
高兴	眉毛参考：稍微下弯	1) 下眼睑下边可能有皱纹，可能鼓起，但并不紧张 2) 鱼尾纹从外眼角向外扩张	1) 唇角向后拉并抬高 2) 嘴可能被张大，牙齿可能露出 3) 一道皱纹从鼻子一直延伸到嘴角外部 4) 脸颊被抬起

悲伤	眉毛内角皱在一起，抬高，带动眉毛下的皮肤	眼内角的上眼皮抬高	1) 嘴角下拉 2) 嘴角可能颤抖
----	----------------------	-----------	----------------------

对于手势识别来说，一个完整的手势识别系统包括三个部分和三个过程。三个部分分别是：采集部分、分类部分和识别部分；三个过程分别是：分割过程、跟踪过程和识别过程。采集部分包括了摄像头、采集卡和内存部分。在多目的手势识别中，摄像头以一定的关系分布在用户前方。在单目的情况下，摄像头所在的平面应该和用户的手部运动所在的平面基本水平。分类部分包括了要处理的分类器和结果反馈回来的接收比较器。用来对之前的识别结果进行校正。识别部分包括了语法对应单位和相应的跟踪机制，通过分类得到的手部形状通过这里一一对应确定的语义和控制命令。分割过程包括了对得到的实时视频图像进行逐帧的手部分割，首先得到需要关注的区域，其次在对得到的区域进行细致分割，直到得到所需要的手指和手掌的形状。跟踪过程包括对手部的不断定位和跟踪，并估计下一帧手的位置。识别过程通过对之前的知识确定手势的意义，并做出相应的反应，例如显示出对应的手势或者做出相应的动作，并对不能识别的手势进行处理，或者报警或者记录下特征后在交互情况下得到用户的指导。手势识别的基本框架如下图所示：

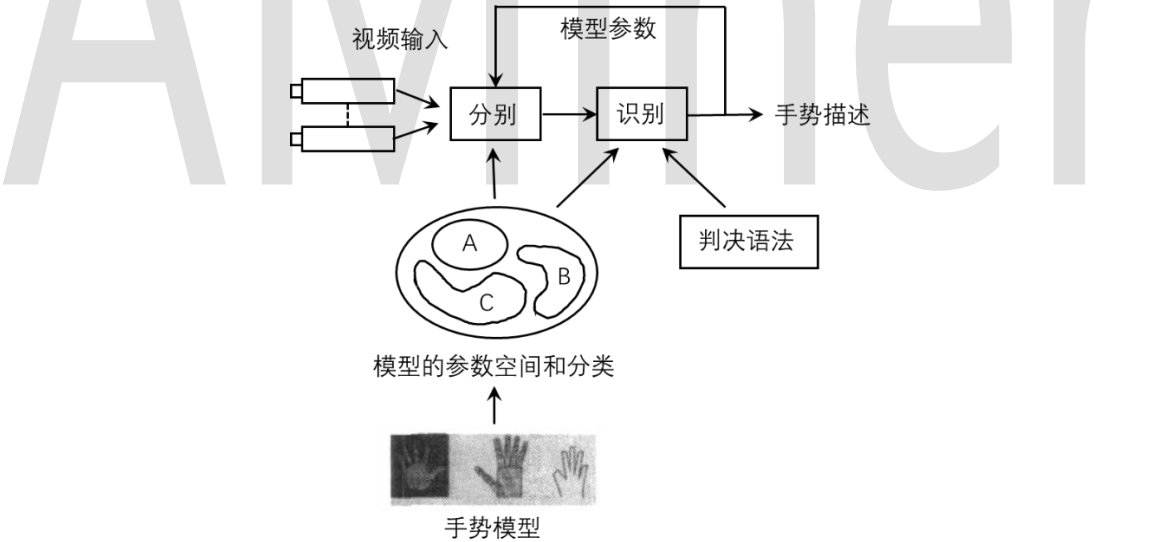


图 2 手势识别的基本框架

● 情感理解与表达

通过对情感的获取、分析与识别，计算机便可了解其所处的情感状态。情感计算技术的实质是通过各种传感器获取由人的情感所引起的表情生理变化信号，利用“情感模型”对这些信号进行识别和分析，从而理解人的情感并做出适当的响应。因此，这部分主要研究的是如何根据情感信息的识别结果，对用户的情感变化做出最适宜的反应。随着情感信息捕获技术的提高



和情感数据资源的扩大,对多特征融合的情感理解模型的研究能力将会有进一步的提高和突破。前面的研究是从视觉面部表情或者是行为特征来推断情感状态。情感表达则是研究其反过程,即给定某一情感状态,研究如何使这一情感状态在一种或几种视觉面部表情或行为特征中体现出来。例如,如何在手势合成和面部表情合成中得以体现,使机器具有情感,能够与用户进行情感交流<sup>[12]</sup>。

下表为常用视觉情感数据库,供读者参考:

表 4 常用的表情分析数据库

数据库	简介	地址
KDEF 与 AKDEF (karolinska directed emotional faces) 数据集	这个数据集最初是被开发用于心理和医学研究目的。它主要用于知觉,注意,情绪,记忆等实验。在创建数据集的过程中,特意使用比较均匀,柔和的光照,被采集者身穿统一的 T 恤颜色。这个数据集,包含 70 个人,35 个男性,35 个女性,年龄在 20 至 30 岁之间。没有胡须,耳环或眼镜,且没有明显的化妆。7 种不同的表情,每个表情有 5 个角度。总共 4900 张彩色图。尺寸为 562*762 像素。	<a href="https://link.zhihu.com/?target=http%3A//www.emotionlab.se/kdef/">https://link.zhihu.com/?target=http%3A//www.emotionlab.se/kdef/</a>
RaFD 数据集	该数据集是 Radboud 大学 Nijmegen 行为科学研究所整理的,这是一个高质量的脸部数据库,总共包含 67 个模特:20 名白人男性成年人,19 名白人女性成年人,4 个白人男孩,6 个白人女孩,18 名摩洛哥男性成年人。总共 8040 张图,包含 8 种表情,即愤怒,厌恶,恐惧,快乐,悲伤,惊奇,蔑视和中立。每一个表情,包含 3 个不同的注视方向,且使用 5 个相机从不同的角度同时拍摄的。	<a href="http://www.socsci.ru.nl:8180/RaFD2/RaFD?p=main">http://www.socsci.ru.nl:8180/RaFD2/RaFD?p=main</a>
Fer2013 数据集	图片的分辨率比较低,共 6 种表情。分别为 0 anger 生气、1 disgust 厌恶、2 fear 恐惧、3 happy 开心、4 sad 伤心、5 surprised 惊讶、6 normal 中性。	<a href="https://github.com/npinto/fer2013">https://github.com/npinto/fer2013</a>
CelebFaces Attributes Dataset (CelebA) 数据集	该数据集是商汤科技的一个用于研究人脸属性的数据集,一个包含超过 200K 名人图像的大型人脸属性数据集,每个数据集都有 40 个属性注释。该数据集中的图像涵盖了大型姿态变化和复杂背景。CelebA 的多样性很好,有约 10 万张带微笑属性的数据。	<a href="http://mmlab.ie.cuhk.edu.hk/projects/CelebA.html">http://mmlab.ie.cuhk.edu.hk/projects/CelebA.html</a>
Surveillance Cameras Face Database	是人脸静态图像的数据库。图像是在不受控制的室内环境中使用五种不同品质的视频监控摄像机拍摄的。数据库包含 130 个主题的 4160 静态图像(在可见和红外光谱中)。	<a href="http://www.scface.org/">http://www.scface.org/</a>

Japanese Female Facial Expression (JAFPE) Database	该数据库包含由 10 名日本女性模特组成的 7 幅面部表情（6 个基本面部表情+1 个中性）的 213 幅图像。每个图像被 60 个日语科目评为 6 个情感形容词。	<a href="http://www.kasrl.org/jaffe.html">http://www.kasrl.org/jaffe.html</a>
--	--	---

表 5 常用的手势分析数据库

数据库	简介	地址
LISA_HD_Static	这个数据库主要用来检测车内司机的行为，对车内司机的手、手机、方向盘进行了标注，用来分析司机是否在驾驶途中使用手机。	<a href="http://cvrr.ucsd.edu/vivachallenge/index.php/hands/hand-detection/">http://cvrr.ucsd.edu/vivachallenge/index.php/hands/hand-detection/</a>
11k Hands	<p>11k Hands 数据集涵盖了来自 190 个 18-75 岁之间的实验对象的手部图像（1600 x 1200 像素）。每个实验对象都要张开或是握紧左右手的手指，然后，在统一的白色背景，且手距离相机位置大致相同的情况下，分别从每只手的背侧和手掌侧进行拍摄。</p> <p>元数据的记录包括：（1）实验对象 ID、（2）性别、（3）年龄、（4）肤色、以及（5）有关所捕捉到的手的一组信息，即右手或左手、手的部位（手背或手掌）以及逻辑指标，即指示手部图像是否包含配饰、指甲油或不符合规则处。所提出的数据集具有大量附加更为详尽的元数据的手部图像。</p> <p>该数据集是免费的，可用于合理的学术研究直接使用。</p>	<a href="https://arxiv.org/abs/1711.04322">https://arxiv.org/abs/1711.04322</a>
Visual Geometry group Hand Dataset	数据库是由 OXFORD Google 发布的，并发布了相关的代码。数据集的目的是对静止图片的手进行检测定位。	<a href="http://www.robots.ox.ac.uk/~vgg/data/hands/index.html">http://www.robots.ox.ac.uk/~vgg/data/hands/index.html</a>
Thomas Moeslundts	这个数据集发布于 1996 年，实验者在深色背景前做不同手语字母。	<a href="http://www-prima.inrialpes.fr/FGnet/data/12-MoeslundGesture/database.html">http://www-prima.inrialpes.fr/FGnet/data/12-MoeslundGesture/database.html</a>

## 2.2 新兴的研究

### 2.2.1 网络海量数据的情感计算

随着时代的发展，网络赋予情感计算新的、更大的数据平台，打开了情感计算的新局面。网络系统由于沟通了人类的现实世界和虚拟世界，可以持续不断地对数量庞大的样本进行情感跟踪，每天这些映射到网络上的情绪不计其数，利用好这些数据反过来就可以验证心理学结论，甚至反哺心理学。由于大数据的分布范围极其广泛，样本数量非常庞大，采用单一的大数据处理方法往往得不到有效的情感要素，统计效果较差。但是，如果将大数据和心理学结合起来，

局面就会大不一样：心理学中，不同情感可以采用维度标定，如冷暖或软硬，同时各种心理效应影响人类对事物的情感判断，如连觉效应、视觉显著性、视觉平衡等，在大数据中引入心理学效应和维度，对有效数据进行心理学情感标准划分，使得数据具有情感维度，这样就会让计算机模拟人类情感的准确性大大提升。网络海量数据的情感主要有以下几个社会属性：

- 情感随群体的变化：在社交网络，如论坛、网络社区等群体聚集的平台上流露出群体的情感，通过这些情感展现可以达到影响其他个人的行为。
- 情感随图片的变化：在社交媒体出现大量的图片，这些图片的颜色、光度、图片内容等各不相同。图片的特征直接影响到了观看者的情感。
- 情感随朋友的变化：在社交平台上，朋友发表的微博、微信状态等容易展现个人的情感。朋友间的关系比陌生人间的关系更加深入，所以朋友的情感更容易引起情感变化，在海量数据中，个人情感容易优先受朋友情感的影响。
- 情感随社会角色的变化：在社交网络中，个人在不同的群体所处的角色也不一样，个人情感流露时也会跟着所处的角色不一样而展现不同的情感。
- 情感随时间的演变：人的情绪是变化无常的，所处的环境不一样，则表现出来的情感也将不一样。即使是同一件事，不同的情景下展现的情感也会不一样。另外，事件的发展是个动态的过程，随着事件的演变，人的情感也会跟着变化。

以全球 Top5 初创企业之一 Affectiva 为例，将实验环境从受控实验室转移到现实世界，搜集真实世界的的数据，用来训练算法在有限条件下识别面部表情，这已成为 Affectiva 的一个竞争优势。该公司建立了一个独特的数据库，包含在 75 个国家拍摄的超过 500 万条视频，也就是 20 亿个表达真实情绪的面部框架。海量数据帮助 Affectiva 利用最新的机器学习方法提高准确性。鉴于数据收集的广度，Affectiva 可以量化文化、性别、年龄甚至环境（例如人们在家中放松还是在开车）对情绪表达方式的影响。

### 2.2.2 多模态计算

虽然人脸、姿态和语音等均能独立地表示一定的情感，但人的相互交流却总是通过信息的综合表现来进行。所以，只有实现多通道的人机界面，才是人与计算机最为自然的交互方式，它集自然语言、语音、手语、人脸、唇读、头势、体势等多种交流通道为一体，并对这些通道信息进行编码、压缩、集成和融合，集中处理图像、音频、视频、文本等多媒体信息。多模态计算是目前情感计算发展的主流方向。每个模块所传达的人类情感的信息量大小和维度不同。

在人机交互中，不同的维度还存在缺失和不完善的问题。因此，人机交互中情感分析应尽可能从多个维度入手，将单一不完善的情感通道补上，最后通过多结果拟合来判断情感倾向。

在多模态情感计算研究中，一个很重要的分支就是情感机器人和情感虚拟人的研究。美国麻省理工学院、日本东京科技大学、美国卡内基梅隆大学均在此领域做出了较好的演示系统。目前中科院自动化所模式识别国家重点实验室已将情感处理融入到了他们已有的语音和人脸的多模态交互平台中，使其结合情感语音合成、人脸建模、视位模型等一系列前沿技术，构筑了栩栩如生的情感虚拟头像，并积极转向嵌入式平台和游戏平台等实际应用。

目前，情感识别和理解的方法上运用了模式识别、人工智能、语音和图像技术的大量研究成果。例如：在情感语音声学分析的基础上，运用线性统计方法和神经网络模型，实现了基于语音的情感识别原型；通过对面部运动区域进行编码，采用 HMM 等不同模型，建立了面部情感特征的识别方法；通过对人姿态和运动的分析，探索肢体运动的情感类别等等。不过，受到情感信息捕获技术的影响，以及缺乏大规模的情感数据资源，有关多特征融合的情感理解模型研究还有待深入。随着未来的技术进展，还将提出更有效的机器学习机制。

## 2.3 情感计算国际会议

本节对国际会议 ACM MM、ACM ICMI 中涉及情感计算的研讨会及其代表性论文进行简单介绍。

### ● ACM MM-- AVEC

ACM MM (ACM International Conference on Multimedia) 是 ACM 多媒体领域的顶级会议，也是中国计算机学会推荐的 A 类国际学术会议。自 1993 年首次召开以来，ACM MM 每年召开一次。ACM MM 的热门方向有大规模图像视频分析、社交媒体研究、多模态人机交互、计算视觉、计算图像等。ACM MM 包含的 workshop 有 SUMAC、MMSports、MAHCI、AVEC、HealthMedia 等。其中，AVEC (International Audio/Visual Emotion Challenge and Workshop) 关注的领域是多媒体信息处理和使用机器学习方法来进行情感计算。AVEC 挑战的目标是提供一个多模态的数据集来汇集多个领域的专家来尝试比较各种方法在情感计算上面的效果，以下是 AVEC 近年代表性论文。

#### **AVEC 2018 “Bipolar Disorder and Cross-cultural Affect”**

AVEC 2018 的主题为 “Bipolar Disorder and Cross-cultural Affect”，这一挑战赛研究将多媒体处理和机器学习方法用于自动音频、视觉和视听健康和情感感知上的应用比较。挑战的目标是为多模态信息处理提供一个通用的基准测试集，并将音频、视觉和视听情感识别社区聚集在

一起, 比较自动健康和情感分析方法的相对优点。挑战的另一动机是推进健康和情感识别系统, 以便能够处理大量未分段、非原型和非预选数据中的完全自然行为, 因为人机 / 人机器人通信接口必须面对现实世界, 而现实世界中多媒体和数据类型大多未分段、非原型和非预选。此届挑战任务包括: Bipolar Disorder Sub-Challenge、Cross-cultural Emotion Sub-Challenge 和 Gold-standard Emotion Sub-Challenge。

表 6 AVEC 2018 Sub-Challenge Winners

Sub-Challenge	Paper	Authors
Bipolar Disorder	Bipolar Disorder Recognition with Histogram Features of Arousal and Body Gestures.	Le Yang, Yan Li, Haifeng Chen, Dongmei Jiang, Meshia Cédric Oveneke, Hichem Sahli
	Bipolar Disorder Recognition via Multi-scale Discriminative Audio Temporal Representation.	hengyin Du, Weixin Li, Di Huang, Yunhong Wang
	Multi-modality Hierarchical Recall based on GBDTs for Bipolar Disorder Classification.	Xiaofen Xing, Bolun Cai, Yinhu Zhao, Shuzhen Li, Zhiwei He, Weiquan Fan
	Automated Screening for Bipolar Disorder from Audio/Visual Modalities.	Zafi Sherhan Syed, Kirill A. Sidorov, A. David Marshall
Cross-cultural Emotion	Speech-based Continuous Emotion Prediction by Learning Perception Responses related to Salient Events: A Study based on Vocal Affect Bursts and Cross-Cultural Affect in AVEC 2018.	Kalani Wataraka Gamage, Ting Dang, Vidhyasaharan Sethu, Julien Epps, Eliathamby Ambikairajah
	Multimodal Continuous Emotion Recognition with Data Augmentation Using Recurrent Neural Networks.	Jian Huang, Ya Li, Jianhua Tao, Zheng Lian, Mingyue Niu, Minghao Yang
	Multi-modal Multi-cultural Dimensional Continues Emotion Recognition in Dyadic Interactions	Jinming Zhao, Ruichen Li, Shizhe Chen, Qin Jin
Gold-standard Emotion	Towards a Better Gold Standard: Denoising and Modelling Continuous Emotion Annotations Based on Feature Agglomeration and Outlier Regularisation.	Chen Wang, Phil Lopes, Thierry Pun, Guillaume Chanel:
	Fusing Annotations with Majority Vote Triplet Embeddings.	Brandon M. Booth, Karel Mundnich, Shrikanth Narayanan:

## AVEC 2017 “Real-life Depression and Affect”

AVEC 2017 的主题为 “Real-life Depression and Affect”，其作为第七届挑战，旨在用于情绪和抑郁的自动视听分析的多媒体处理和机器学习方法比较。挑战的目标是在严格的可比条件下，比较用于视听情感识别和评估抑郁症的严重程度的方法的相对优点，以及在何种程度上方法的融合是可行并有益的。比赛的动机之一是推动情绪识别和抑郁严重程度估计的多媒体检索，以实现人与人、或人与人现实交互过程中能够可靠地感知。AVEC 2017 有助于通过挑战来提高情绪和抑郁水平的检测标准，并基于现实生活条件下捕获的视听数据进行检测，它能够弥合关于情绪和抑郁认知的优秀研究和低可比性结果之间的差距。此届挑战任务包括：Affect Sub-Challenge, Depression Sub-Challenge。

表 7 AVEC 2017 Sub-Challenge Winners

Sub-Challenge	Paper	Authors
Affect	Continuous Multimodal Emotion Prediction Based on Long Short Term Memory Recurrent Neural Network.	Jian Huang, Ya Li, Jianhua Tao, Zheng Lian, Zhengqi Wen, Minghao Yang, Jiangyan Yi
	Multimodal Multi-task Learning for Dimensional and Continuous Emotion Recognition	Shizhe Chen, Qin Jin, Jinming Zhao, Shuai Wang
	Investigating Word Affect Features and Fusion of Probabilistic Predictions Incorporating Uncertainty in AVEC 2017.	Ting Dang, Brian Stasak, Zhaocheng Huang, Sadari Jayawardena, Mia Atcheson, Munawar Hayat, Phu Ngoc Le, Vidhyasaharan Sethu, Roland Goecke, Julien Epps
Depression	Depression Severity Prediction Based on Biomarkers of Psychomotor Retardation.	Zafi Sherhan Syed, Kirill A. Sidorov, A. David Marshall
	Hybrid Depression Classification and Estimation from Audio Video and Text Information.	Le Yang, Hichem Sahli, Xiaohan Xia, Ercheng Pei, Meshia Cédric Oveneke, Dongmei Jiang
	Multimodal Measurement of Depression Using Deep Learning Models.	Le Yang, Dongmei Jiang, Xiaohan Xia, Ercheng Pei, Meshia Cédric Oveneke, Hichem Sahli
	A Random Forest Regression Method With Selected-Text Feature For Depression Assessment.	Bo Sun, Yinghui Zhang, Jun He, Lejun Yu, Qihua Xu, Dongliang Li, Zhaoying Wang
	Topic Modeling Based Multi-modal Depression Detection	Yuan Gong, Christian Poellabauer



## AVEC 2016 “Depression, Mood and Emotion”

AVEC 2016 的主题为 “Depression, Mood and Emotion”，其作为第六届挑战，旨在比较用于自动音频、视觉和生理抑郁和情绪分析的多媒体处理和机器学习方法的比较。挑战的目标是为多模态信息处理提供一个共同的基准测试集，并将抑郁和情感识别社区以及音频、视频和生理处理社区聚集在一起，以比较各自相对优点。此届挑战任务包括：Emotion Detection, Depression Recognition。

表 8 AVEC 2016 Sub-Challenge Winners

Sub-Challenge	Paper	Authors
Depression recognition	Detecting Depression using Vocal, Facial and Semantic Communication Cues.	James R. Williamson, Elizabeth Godoy, Miriam Cha, Adrianne Schwarzentruher, Pooya Khorrami, Youngjune Gwon, Hsiang-Tsung Kung, Charlie K. Dagli, Thomas F. Quatieri
	Staircase Regression in OA RVM, Data Selection and Gender Dependency in AVEC 2016.	Zhaocheng Huang, Brian Stasak, Ting Dang, Kalani Wataraka Gamage, Phu Ngoc Le, Vidhyasaharan Sethu, Julien Epps
	Depression Assessment by Fusing High and Low Level Features from Audio, Video, and Text.	Anastasia Pampouchidou, Olympia Simantiraki, Amir Fazlollahi, Matthew Pediaditis, Dimitris Manousos, Alexandros Roniotis, Giorgos A. Giannakakis, Fabrice Mériau, Panagiotis G. Simos, Kostas Marias, Fan Yang, Manolis Tsiknakis
	DepAudioNet: An Efficient Deep Model for Audio based Depression Classification.	Xingchen Ma, Hongyu Yang, Qiang Chen, Di Huang, Yunhong Wang
	Multimodal and Multiresolution Depression Detection from Speech and Facial Landmark Features.	Md. Nasir, Arindam Jati, Prashanth Gurunath Shivakumar, Sandeep Nallan Chakravarthula, Panayiotis G. Georgiou
	High-Level Geometry-based Features of Video Modality for Emotion Prediction. 51-58	Raphaël Weber, Vincent Barrielle, Catherine Soladié, Renaud Séguier
Emotion Detection	Online Affect Tracking with Multimodal Kalman Filters.	Krishna Somandepalli, Rahul Gupta, Md. Nasir, Brandon M. Booth, Sungbok Lee, Shrikanth S. Narayanan
	Continuous Multimodal Human Affect Estimation using Echo State Networks.	Mohammadreza Amirian, Markus Kächele, Patrick Thiam, Viktor Kessler, Friedhelm Schwenker

	Multimodal Emotion Recognition for AVEC 2016 Challenge.	Filip Povolný, Pavel Matejka, Michal Hradis, Anna Popková, Lubomir Otrusina, Pavel Smrz, Ian Wood, Céile Robin, Lori Lamel
	Exploring Multimodal Visual Features for Continuous Affect Recognition.	Bo Sun, Siming Cao, Liandong Li, Jun He, Lejun Yu

### ● ACM ICMI--EmotiW

ACM ICMI (ACM International Conference on Multimodal Interaction) 是多模式人、人和人机交互, 界面和系统开发的多学科研究的首要国际论坛。会议重点关注理论和实证基础, 组件技术以及定义多模态交互分析, 界面设计和系统开发领域的组合多模式处理技术。ICMI 包含一系列 workshop 与 challenge。Challenge 主要包含三个, 其中, challenge 中参与人数最多、举办时间最长的是 Emotion Recognition in the Wild Challenge (EmotiW)。EmotiW 挑战重点关注无约束条件下的情感感知和基于嵌入式音频视频的情感分类和基于图像的组级面部表情识别, 以模仿现实世界的情况, 以下是 EmotiW 近年代表性论文。

表 9 EmotiW 2017 Sub-Challenge Winners

Sub-Challenge	Paper	Authors
Audio-Video Sub-challenge	Learning Supervised Scoring Ensemble for Emotion Recognition in the Wild	Ping Hu, Dongqi Cai, Shandong Wang, Anbang Yao and Yurong Chen
	Convolutional neural networks pretrained on large face recognition datasets for emotion classification from video	Boris Knyazev, Roman Shvetsov, Natalia Efremova, Artem Kuharenko
	Temporal Multimodal Fusion for Video Emotion Classification in the Wild	Valentin Vielzeuf, Stéphane Pateux and Frederic Jurie
Group-based Sub-challenge	Group Emotion Recognition with Individual Facial Emotion CNNs and Global Image Based CNNs	Lianzhi Tan, Kaipeng Zhang, Kai Wang, Xiaoxing Zeng, Xiaojiang Peng and Yu Qiao
	Group-Level Emotion Recognition using Deep Models on Image Scene, Faces, and Skeletons	Xin Guo, Luisa Polania and Kenneth Barner
	A New Deep-Learning Framework for Group Emotion Recognition	Qinglan Wei, Yijia Zhao, Qihua Xu, Liandong Li, Jun He, Lejun Yu and Bo Sun

表 10 EmotiW 2016 Sub-Challenge Winners

Sub-Challenge	Paper	Authors
Video based emotion recognition	Video-based Emotion Recognition Using CNN-RNN and C3D Hybrid Networks	Fan, Yin, Xiangju Lu, Dian Li, and Yuanliu Liu
	HoloNet: towards robust emotion recognition in the wild	Yao, Anbang, Dongqi Cai, Ping Hu
	Emotion Recognition in the Wild from Videos using Images	Bargal, Sarah Adel, Emad Barsoum, Cristian Canton Ferrer, and Cha Zhang
	Multi-cue Fusion for Emotion Recognition in the Wild	Yan, Jingwei, Wenming Zheng, Zhen Cui, Chuangao Tang, Tong Zhang, and Yuan Zong
Group based emotion recognition	Happiness Level Prediction with Sequential Inputs via Multiple Regressions	Li, Jianshu, Sujoy Roy, Jiashi Feng, and Terence Sim
	Group happiness assessment using geometric features and dataset balancing	Vonikakis, Vassilios, Yasin Yazici, Viet Dung Nguyen, and Stefan Winkler
	LSTM for Dynamic Emotion and Group Emotion Recognition in the Wild	Sun, Bo, Qinglan Wei, Liandong Li, Qihua Xu, Jun He, and Lejun Yu

## 2.4 情感计算获奖论文

本节对近年 ACM MM 中涉及情感计算的获奖论文进行简单介绍，欢迎读者交流补充。

### ● 2018-ACM MM Best Demo

**篇名：**AniDance: Real-Time Dance Motion Synthesize to the Song.

**作者：**Taoran Tang, Hanyang Mao, Jia Jia

**单位：**Tsinghua University

**概述：**本篇论文介绍了一个名为 AniDance 的系统，它可以实时地将舞蹈动作与旋律相结合。当用户在手机中播放一首歌或者在 AniDance 中播放一首歌时，歌曲旋律将驱动 3D 空间角色跳舞，创造出一个生动的舞蹈动画。在实践中，研究人员通过捕捉真实的舞蹈表演，来构成 3D 空间舞蹈运动数据集，使用 LSTM 自动编码器来识别音乐和舞蹈之间的关系。基于这些技术，用户可以提升他们的学习能力以及对舞蹈和音乐的兴趣。

**下载地址：**<https://hcsi.cs.tsinghua.edu.cn/Paper/Paper18/MM18-TANGTAORANDEMO.pdf>

## ● 2016-ACMMM Best Paper

**篇名:** Multi-modal Multi-view Topic-opinion Mining for Social Event Analysis

**作者:** 钱胜胜、张天柱、徐常胜

**单位:** 中国科学院自动化研究所

**概述:** 本篇论文针对社会事件数据中呈现的多视角多模态等属性,提出一种多模态多视角的主题观点挖掘方法。该方法利用不同新闻媒体对同一个社会热点事件观点的差异性,研究不同视角对应的观点。通过本方法,既能够细致地分析不同媒体对社会热点事件的观点,又能够对社会热点事件进行全方位展示,便于探究热点事件的全局态势与舆情走势。本篇论文从多个社交平台上检测社会热点事件,并将其全面地展现出来,不仅切合政府和民众的需求,也是国际学术的研究热点。获奖论文敏锐地利用不同社交平台的在传播速度和传播内容上的特点,提出了一个有效的跨平台跨模态的社会热点事件监测方法。选题的新颖性、解决思路的创新性、以及技术的有效性,均受到评审专家的一致认可。

**下载地址:** <http://www.millenniumsoftsol.com/courses/IEEETitles/Java/Multi-Modal-Event.pdf>

## ● 2014- ACM MM Best Student Paper Award

**篇名:** Say Cheese vs. Smile: Reducing Speech-Related Variability for Facial Emotion Recognition

**作者:** Yelin Kim and Emily Mower Provost.

**单位:** Electrical Engineering: Systems, University of Michigan

**概述:** 作者认为,情感识别系统的应用非常有意义,比如社会和情感人机交互系统及与健康相关的系统可以帮助个体更好地记录他们的情感世界。但情感的表达是非常复杂的,可以通过面部表情、语音语调以及肢体语言的变化来体现。情感识别系统必须能够“解码”这些外在的变化来了解其代表的真实情感。但是,外在表情和行为的变化不仅仅受到情绪的影响,比如面部表情的变化有时是由于语言发音引起的。有效的情感识别系统必须能够辨别人的表情及行为变化是由于语言发音导致的,还是由于情感变化导致的。在本文中,作者描述了一种面部表情分析模型,目的是减少语言发音对于面部表情的影响,更加准确地识别表情变化背后的情感变化,提高面部表情的情感识别率、

**下载地址:** <https://cn.aminer.cn/archive/say-cheese-vs-smile-reducing-speech-related-variability-for-facial-emotion-recognition/555045d945ce0a409eb5a118>

● 2012- ACM MM Grand Challenge 1st Prize

**篇名:** The Acousticvisual Emotion Guassians Model for Automatic Generation of Music Video

**作者:** Ju-Chiang Wang, Yi-Hsuan Yang, I-Hong Jhuo, Yen-Yu Lin, Hsin-Min Wang:

**单位:** Research Center for Information Technology Innovation Academia Sinica

**概述:** 音乐研究中最具挑战性的工作是建立能够比较音乐情感内涵，并能根据情感组织音乐的计算模型。本文中，作者提出了一种新型声学情感高斯模型（AEG）来界定音乐情感感知的过程。AEG 作为一种生成模型，对学习过程做出了简单直接的解释。为了架起声学特征空间与音乐情感空间的桥梁，作者从数据中提取了一系列潜在特征集，并将其引入到模型中，执行两个空间端到端语义映射。AEG 模型适用于音乐情感自动标注与基于情感的音乐检索。为了深入了解 AEG 模型，文中提供了模型学习过程的说明。为了表明 AEG 模型具有更高的精度，作者对两种音乐语料库 MER60 和 Muturk 进行了综合性能研究。结果显示，AEG 模型在音乐情感注释方面胜过以前的方法。另外，文章首次提出了基于音乐情感检索的定量评估。

**下载地址:** <https://cn.aminer.cn/archive/the-acousticvisual-emotion-guassians-model-for-automatic-generation-of-music-video/53e9ad4fb7602d9703736369>

### 3 人才篇

为了了解情感计算专家学者的分布，我们统计了最近 10 年（2009 年至 2018 年），情感计算相关的期刊和会议，并按照相关关键词进行挖掘，找到了在以上限定范围内发文量最高的 1000 余名情感计算学者。

研究选择的期刊和会议包括以下：

表 11 情感计算相关期刊、会议

全称	分类	简称	中文	网址
American Association for Applied Linguistics	会议	AAAL	美国人工智能协会人工智能会议	<a href="https://www.aaal.org/">https://www.aaal.org/</a>
International Joint Conference on Artificial Intelligence	会议	IJCAI	国际人工智能联合大会	<a href="https://www.ijcai.org/">https://www.ijcai.org/</a>
IEEE Transactions on Multimedia	期刊	TMM	电气和电子工程师协会多媒体汇刊	<a href="https://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=6046">https://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=6046</a>
IEEE Transactions on Automatic Control	期刊	TAC	电气和电子工程师协会自动控制汇刊	<a href="https://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=9">https://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=9</a>
IEEE Transactions on Knowledge and Data Engineering	期刊	TKDE	电气和电子工程师协会知识与数据工程汇刊	<a href="https://ieeexplore.ieee.org/xpl/RecentIssue.jsp?reload=true&amp;punumber=69">https://ieeexplore.ieee.org/xpl/RecentIssue.jsp?reload=true&amp;punumber=69</a>
IEEE Transactions on Affective Computing	期刊	TAC	情感计算汇刊	<a href="https://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=5165369">https://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=5165369</a>
International Conference on Acoustics, Speech and Signal Processing	会议	ICASSP	国际声学、语音和信号处理会议	<a href="http://www.ieee-icassp2017.org/">http://www.ieee-icassp2017.org/</a>
Conference on Knowledge Discovery and Data Mining	会议	KDD	知识发现与数据挖掘会议	<a href="https://www.kdd.org/">https://www.kdd.org/</a>
IEEE International Conference on Multimedia and Expo	会议	ICME	电气和电子工程师协会国际多媒体与博览会会议	<a href="http://www.icme2019.org/">http://www.icme2019.org/</a>
ACM International Conference on Multimedia	会议	ACMMM	多媒体国际会议	<a href="https://www.acmmm.org/2019/">https://www.acmmm.org/2019/</a>



筛选的关键词包括: expression、emotional、emotion、emotive、affect、affective、sentiment、sentimental、mental、health、stress、depression、aesthetics 等。

### 3.1 学者情况概览

#### 3.1.1 全球学者概况

- 学者地图

学者分布地图对于进行学者调查、分析各地区竞争力现状尤为重要，图 3 为情感计算领域全球顶尖学者分布状况。其中，颜色越趋近于红色，表示学者越集中；颜色越趋近于绿色，表示学者越稀少。从地区角度来看，北美洲、欧洲是情感计算领域学者分布最为集中的地区，亚洲东部地区次之，南美洲和非洲学者极为匮乏。从国家角度来看，情感计算领域的人才在美国最多，中国次之，意大利、法国等洲国家也有较多的学者数量，整体上讲其它国家与美国的差距较大，学者数量排名前十的国家如图 4 所示



图 3 情感计算全球专家分布

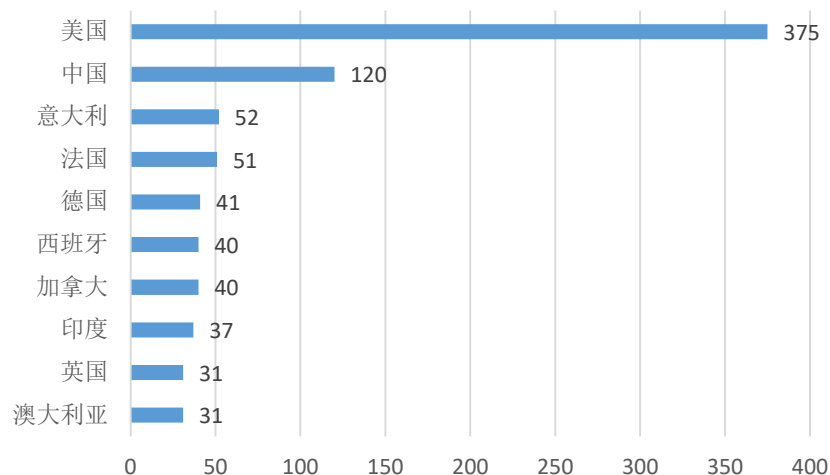


图 4 情感计算专家国家数量分布

#### ● 性别比例

在性别比例方面，情感计算领域中男性学者占比 92.7%，女性学者占比 7.3%，男性学者占比远高于女性学者。

#### ● 学者 h-index 分布

情感计算领域学者的 h-index 分布如图 5 所示，分布情况整体呈阶梯状，大部分学者的 h-index 分布在中低区域，其中 h-index 在 <10 的区间人数最多，有 524 人，占比 43.4%，50-60 区间人数最少，有 46 人，占比 3.8%。

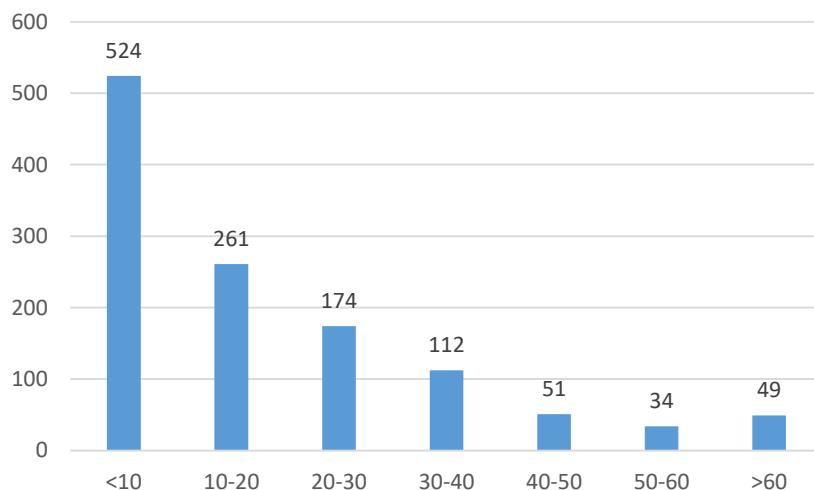


图 5 情感计算领域学者 h-index 分布

● 人才迁徙

我们对情感计算领域 TOP 学者的迁徙路径做了分析。由图 6 可以看出，各国情感计算 TOP 学者的流失和引进是相对比较均衡的，其中美国是情感计算领域人才流动大国，人才输入和输出幅度领先于其他国家，且从数据来看人才流出大于人才流入。英国、加拿大和印度等国人才迁徙流量小于美国；中国人才流入略高于人才流出。人才的频繁流入流出，使得该领域的学术交流活动增加，带动了人才质量提升的同时，也促进了领域理论及技术的更新迭代，逐渐形成一种良性循环的过程。

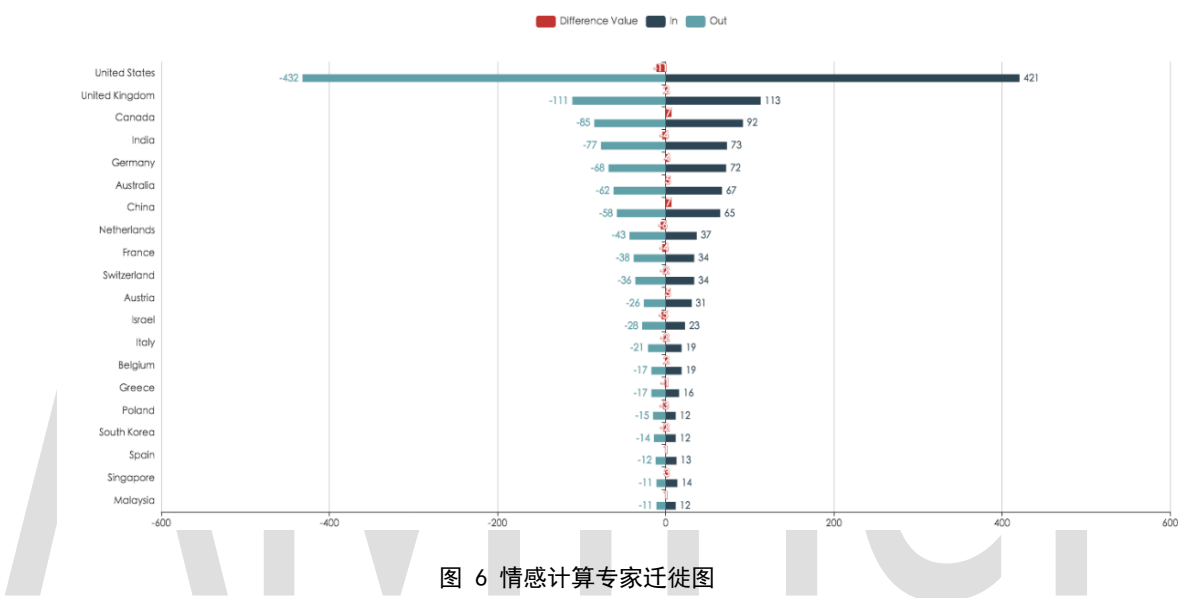


图 6 情感计算专家迁徙图

● 中外合作

中国与其他国家在情感计算领域的合作情况可以根据 AMiner 数据平台分析得到，通过统计论文中作者的单位信息，将作者映射到各个国家中，进而统计中国与各国之间合作论文的数量，并按照合作论文发表数量从高到低进行了排序，如表 12 所示。

表 12 情感计算领域中国与各国合作论文情况

合作国家	论文数	引用数	平均引用数	学者数
中国--美国	178	8603	48	468
中国--英国	28	1437	51	66
中国--澳大利亚	22	579	26	55
中国--新加坡	18	786	44	48
中国--加拿大	14	440	31	38
中国--日本	12	359	30	27
中国--瑞士	8	897	112	10
中国--德国	8	257	32	24

中国—瑞典	6	328	55	10
中国—法国	4	296	74	8

从上表数据可以看出，中美合作的论文数、引用数、学者数遥遥领先，表明中美间在情感计算领域合作之密切；从地域角度看，中国与欧洲的合作非常广泛，前 10 名合作关系里中欧合作共占 5 席；中国与瑞士合作的论文数虽然不是最多，但是拥有最高的平均引用数说明在合作质量上中瑞合作达到了较高的水平。

### 3.1.2 国内学者概况

- 学者地图

AMiner 选取情感计算领域国内专家学者绘制了学者国内分布地图，如图 7 所示。通过下图我们可以发现，京津地区在情感计算领域的人才数量最多，东部及南部沿海地区的也有较多的人才分布，相比之下，内陆地区信情感计算领域人才较为匮乏，这也从一定程度上说明了情感计算领域的发展与该地区的地理位置和经济水平都是息息相关的。同时，通过观察中国周边国家的学者数量情况，特别是与日本、东南亚等亚洲国家相比，中国在情感计算领域顶尖学者数量方面具有较为明显的优势。图 8 是我国情感计算领域顶尖学者最多的 10 个省份。

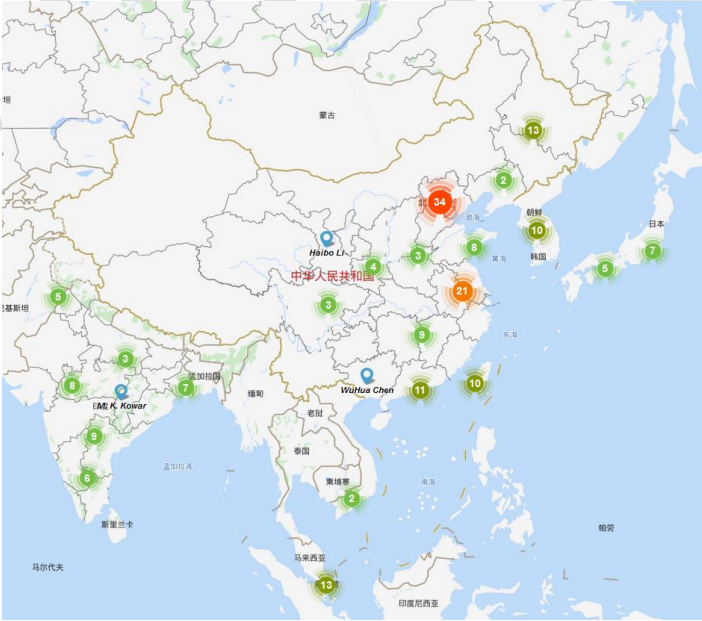


图 7 情感计算国内学者分布

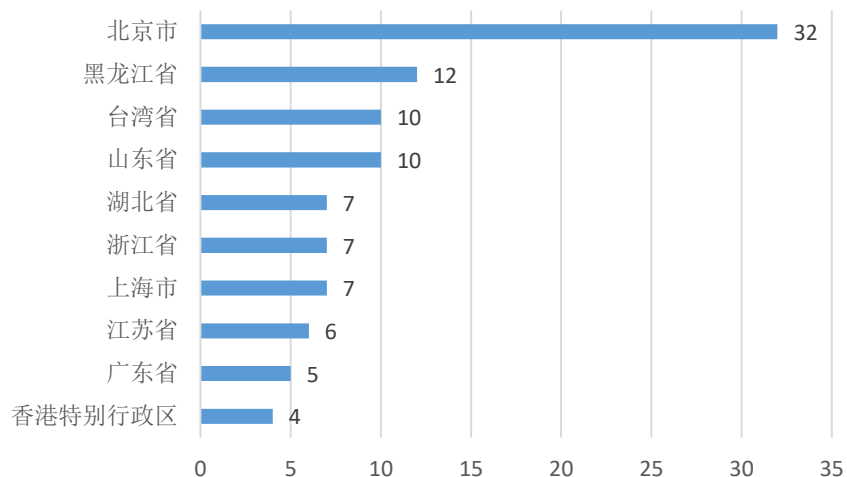


图 8 情感计算学者分布国内省份 TOP10

## 3.2 典型学者

我们在选取的期刊会议中，对所涉学者及其论文关键信息进行抽取，下面将对国内及国外人才部分代表性进行介绍。本文中“国外学者”指在国外科研单位供职的学者。

### 3.2.1 国外代表性学者

#### ● Rosalind W. Picard



Rosalind W. Picard

H 87 A 53.35 S 15.54 c 35352 P 420

Professor

Massachusetts Institute of Technology (MIT)

Affective Computing

Facial Expression

Wearable Computer

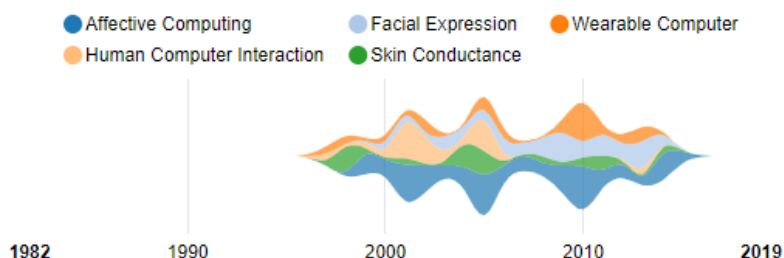
Human Computer Interaction

Skin Conductance

Image Texture

Electrodermal Activity

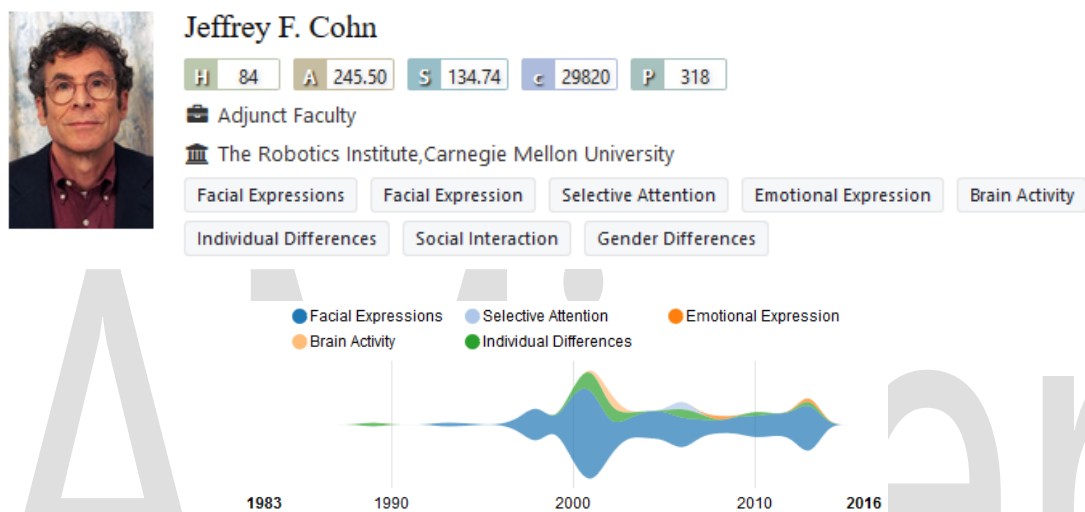
Image Processing



Rosalind W. Picard，美国麻省理工学院媒体实验室情感计算研究组创始人兼主任，美国电气电子工程师学会（IEEE）院士，创业公司 Affectiva 与 Empatica 的联合创始人。她于 1995 年

正式提出“情感计算（Affective Computing）”的概念，为人工智能奠定了理论基础。她指出“情感计算就是针对人类的外在表现，能够进行测量和分析并能对情感施加影响的计算”，开辟了计算机科学的新领域，其思想是使计算机拥有情感，能够像人一样识别和表达情感，从而使人机交互更自然，她所在的实验室通过非常动态技术的推进来追踪不同的情感，并理解情感的发生。

### ● Jeffrey F. Cohn

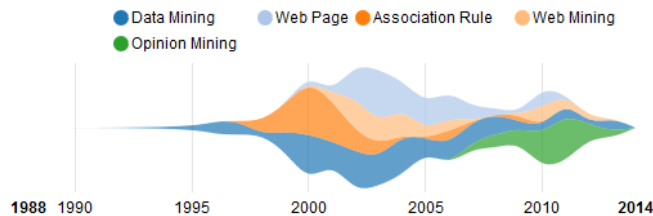


Jeffrey F. Cohn，匹兹堡大学心理学、精神病学和智能系统教授，卡内基梅隆大学机器人研究所计算机科学副教授。他领导跨学科和跨机构的努力，开发先进的自动分析和综合面部表情和韵律的方法，并将这些工具应用于人类情感、社会发展、非语言沟通、精神病理学和生物医学的研究。他的研究得到了美国国立卫生研究院、国家科学基金会、自闭症基金会、海军研究办公室和国防高级研究计划局等赞助机构的资助。曾担任 FG2020、FG2017、FG2015、FG2008、情感计算与智能交互国际会议（ACII 2009）、多模态界面国际会议（ACM 2014）总主席。他过去的合作编辑 IEEE 在情感计算（TAC）和有“特殊问题对情感计算 IEEE 交易模式分析与机器智能,杂志的形象和视觉计算、模式识别字母,计算机视觉和图像理解和 ACM 交易互动的智能系统。

### ● 刘兵

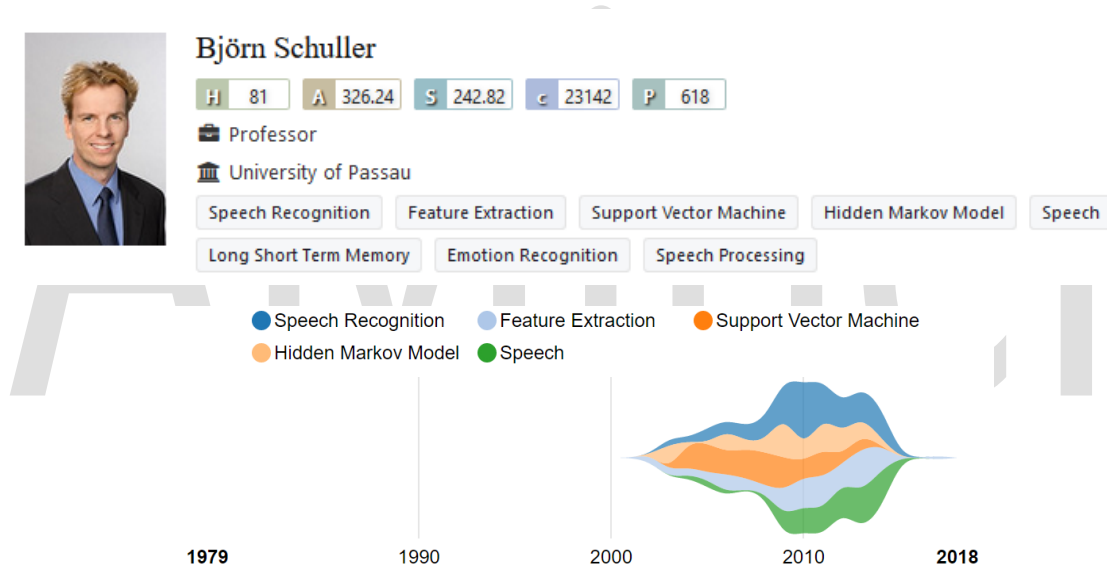






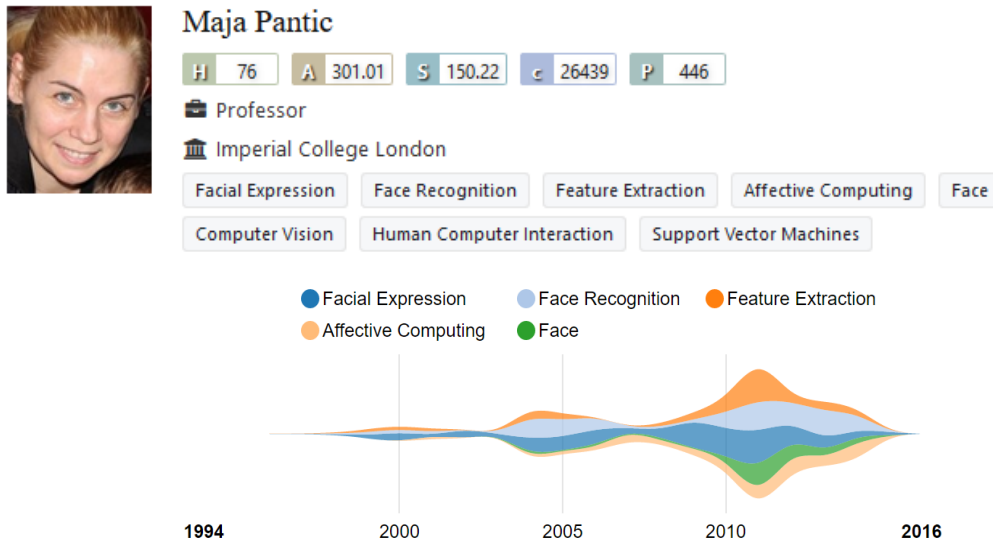
刘兵，伊利诺斯大学芝加哥分校教授，于爱丁堡大学获得了人工智能博士学位，主要研究情感分析、机器学习、数据挖掘和自然语言处理等。曾担任 KDD、ICDM、CIKM、WSDM、SDM 和 PAKDD 等会议的主席，并担任 TKDE、TWEB、DMKD 等的编辑。他是 ACM 研究员，AAAI 研究员，IEEE 研究员。在顶级会议和期刊上发表了大量文章。他的两篇论文获得了 KDD 为期 10 年的测试时间奖。他还写了三本书：两本关于情感分析，一本关于网络数据挖掘。

### ● Björn Schuller



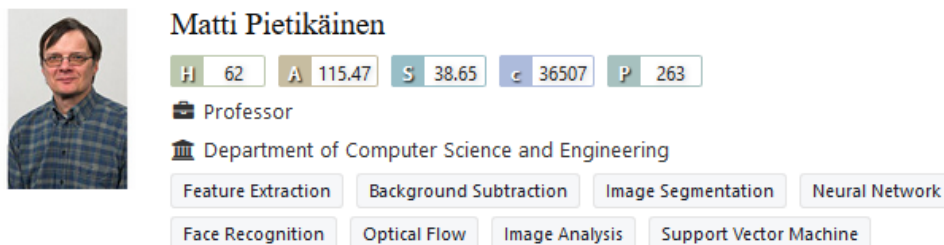
Björn Schuller，德国帕绍大学 Complex and Intelligent Systems 机构的全职教授和主席，audEERING UG 的联合创始首席执行官。他曾担任情感计算促进协会主席（2013-2015），并且还担任 IEEE 情感计算汇刊的主编。2015 年，他被世界经济论坛评为 40 岁以下的 40 位杰出科学家之一，他的研究已获得超过 700 万欧元的外部资金。他最著名的研究为：应用智能视听和复杂信息系统工程推进机器学习，以及人机/机器人交互和多媒体检索的情感计算。他还被邀请担任哈尔滨工业大学计算机科学与技术学院的常任客座教授，和瑞士日内瓦国立大学的客座教授。

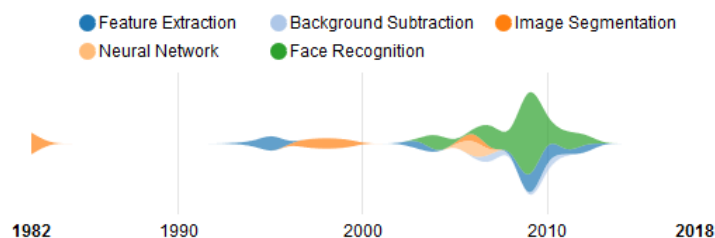
## ● Maja Pantic



Maja Pantic, 伦敦帝国理工学院情感和行为计算教授，同时也是剑桥三星人工智能中心（SAIC）的研究主管。她致力于自动分析人类面部表情，称为面部信息高级互动（FIFAI），由荷兰研究理事会初级奖学金（NOW-VENIEW-639.021.202）资助，为期三年（02-01-2003 至 02-01-2006）。该补助金每年颁发给荷兰 7 位最优秀的精确科学研究人员。此外，她还研究对人类自然主义非语言行为进行多模态分析，名为 MAHNOB，由欧洲研究理事会在启动资助计划中资助了五年（01-09-2008 至 01-09-2013）。该奖学金授予 2007 年欧洲任何研究领域 2% 的最佳初级科学家（ERC-StG-2007-奖项）。2011 年，她获得了 BCS Roger Needham 奖，该奖项每年颁发给 1 位英国研究员，Maja Pantic 获得该奖是由于其在就读博士学位十年内为计算机科学做出了杰出贡献。

## ● Matti Pietikäinen





Matti Pietikäinen 教授为芬兰奥卢大学机器视觉研究中心主任，IEEE Fellow，IAPR founding fellow。曾担任计算机视觉领域顶级期刊 IEEE Transactions on Pattern Analysis and Machine Intelligence 和重要期刊 Pattern Recognition 的副主编，目前担任 Image and Vision Computing 和 IEEE Transactions on Forensics and Security 的副编辑。他在局部二进制模式（LBP）算法、基于纹理的图像和视频分析以及人脸图像分析等很多方面都做出了开创性的贡献。2014 年，其因在计算机视觉基于 LBP 人脸描述研究的重要贡献获得 Koenderink 奖。

### ● Thierry Pun



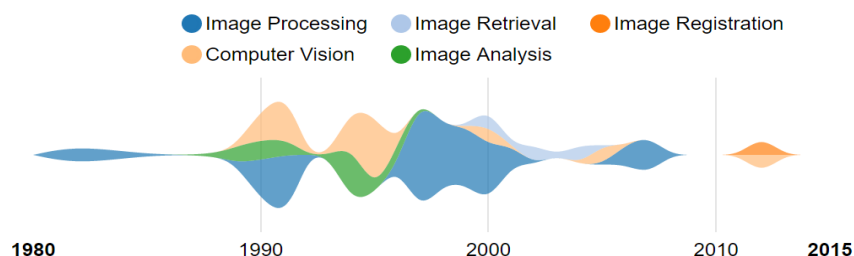
Thierry Pun

H 52 A 4.15 S 0.03 c 14806 P 376

Professor

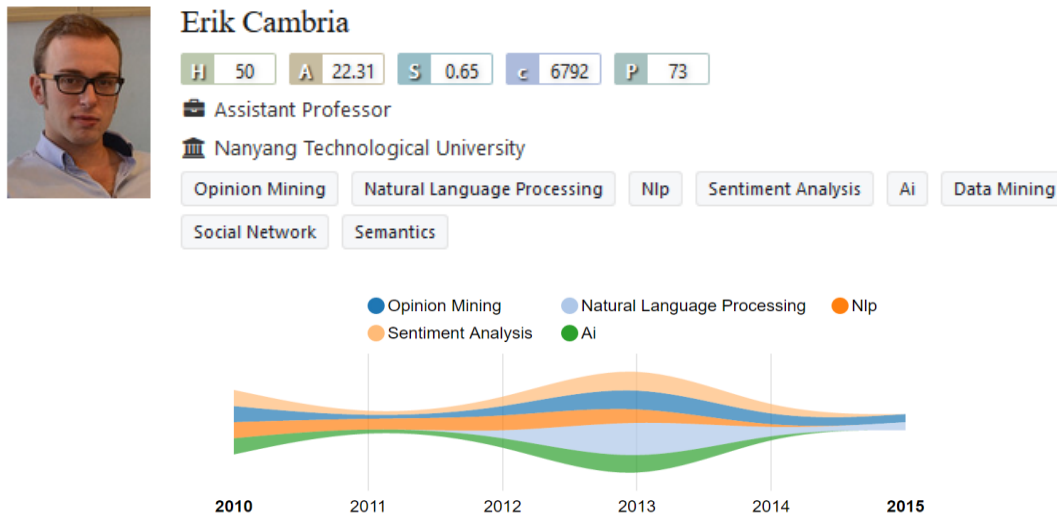
university of geneva

Image Processing Image Retrieval Image Registration Digital Image Processing Computer Vision  
Image Analysis Image Compression Text Retrieval



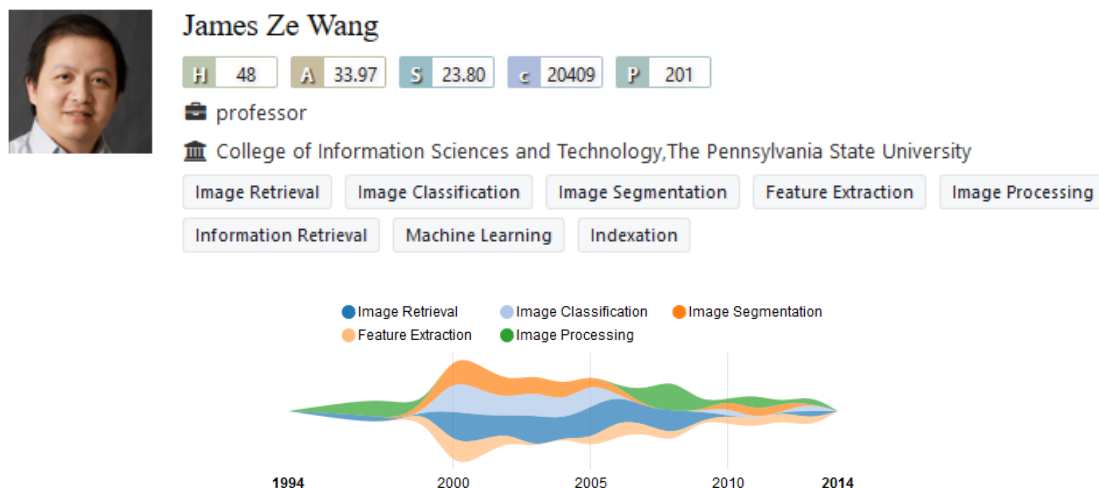
Thierry Pun，瑞士日内瓦大学计算机科学系的全职教授。他撰写了 200 多篇计算机视觉处理领域的论文，获得了 7 项专利，并领导或参与了多个国家研究项目。他所在团队为 Computer Vision and Multimedia Laboratory (CVML) - The Multimodal interaction group (MMI)。MMI 是计算机视觉和多媒体实验室（CVML）的一部分，主要致力于研究使用各种模式进行人机交互和情感计算。目前的研究模态包括视觉（眼睛注视跟踪）、听觉（3D 声音）、触觉、生理信号等。关于使用生理信号和 eInclusion 进行交互的三个主要研究和开发主题包括：基于脑电图的脑机交互（BCI）和脑源重建（前向和反向问题）；情感计算和情绪评估；视力残疾人士及长者的电子辅助设备等等。

## ● Erik Cambria



Erik Cambria, 南洋理工大学计算机科学与工程学院的助理教授, SenticNet 创始人。他是 Elsevier KBS and IPM, IEEE CIM, Springer AIRE, Cognitive Computation 等杂志的副编辑, 以及 the IEEE IS Department on Affective Computing and Sentiment Analysis 的编辑。他曾获得多个奖项, 例如 2018 年的 AI's 10 to Watch 和 2019 年 IEEE 杰出创业奖。他还是脑科学基金会的成员 (the Brain Sciences Foundation), 该基金会由已故的图灵奖获得者 Marvin Minsky 和 SenticNet Ltd. 的创始人共同创立和指导。

## ● James Ze Wang



James Ze Wang, 在斯坦福大学与 1997 年取得计算机及数学硕士学位, 于 2000 年取得医学信息学博士学位, 现为宾夕法尼亚州立大学信息科学与技术系教授。目前主要研究方向为生物医学影像信息学、计算心理学、气象大数据、视觉艺术、基础可解释学习方法等。曾获美国国

家科学基金会终身成就奖及 PNC 技术终身发展教授。他曾担任 IEEE 模式分析和机器智能特刊《真实世界图像注释与检索》的首席客座编辑，ACM 多媒体信息检索活动的总主席。

### ● Wijnand Ijsselsteijn



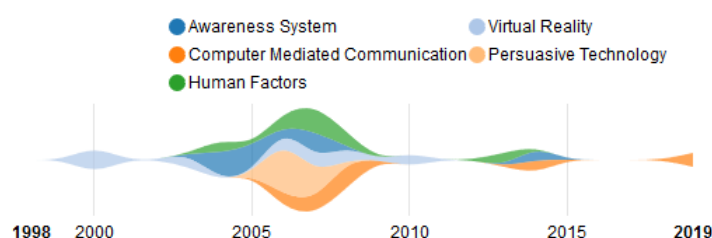
Wijnand Ijsselsteijn

H 47 A 13.42 S 3.57 c 9663 P 145

Professor

Eindhoven University of Technology

Awareness System Virtual Reality Human Factors Computer Mediated Communication  
Persuasive Technology Image Quality Affective Benefit Psychology



Wijnand Ijsselsteijn, 是埃因霍芬理工大学人机交互认知与影响专业的全职教授。他积极研究媒体技术对人类心理的影响，并利用心理学改进技术设计。他的研究重点是概念化和测量人类体验与数字环境（沉浸式媒体、严肃游戏、情感计算、个人信息学）的关系，为人类学习、健康和福祉服务。他对数据科学、人工智能和心理学之间的关系有着浓厚的兴趣，并致力于技术创新（如支持传感器的移动技术、虚拟环境），这些创新使人类行为跟踪的新形式成为可能，将方法论的严谨性与生态有效性相结合。

### ● Elisabeth André



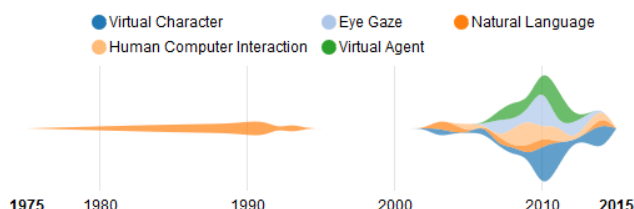
Elisabeth André

H 43 A 25.54 S 0 c 5994 P 238

Professor

Institute of Computer Science University of Augsburg

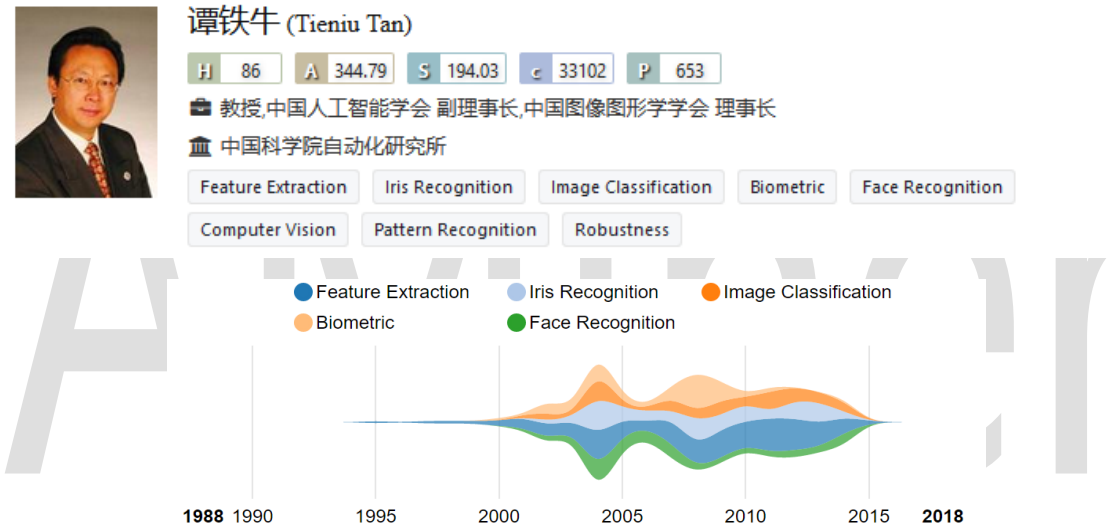
Virtual Character Eye Gaze Virtual Agent Natural Language Human Computer Interaction  
Multimodal Interaction Augmented Reality User Interface



Elisabeth André, 奥格斯堡大学应用信息系计算机科学学院教授、多媒体系主任。研究领域为多通道用户界面、多模态分析（生理数据、注视、言语、手势）、情感计算和社会信号处理、社交机器人&嵌入式会话代理、技术强化学习、有形和感知的用户界面等。担任《IEEE 情感计算学报》主编，《物联网交易》（TIOT）、ACM 智能交互系统（TIIS）副主编，自主代理和多代理系统国际会议（AAMAS 2018）、第十八届人工智能教育国际会议（AIED 2017）等会议联席主席。2019 年，德国国家信息学学会评选出德国人工智能史上最具影响力的十大人物之一。

### 3.2.2 国内代表性学者

● 谭铁牛



谭铁牛，中国科学院自动化研究所所长兼中国科学技术大学自动化系主任。在图象处理、计算机视觉和模式识别领域成就显著，主张数字化人类情感、和谐人际交互环境中的情感计算探索。1996 年底，他入选中国科学院“百人计划”，同时担任模式识别国家重点实验室主任等职务。此外，他还担任中国图象图形学会常务副理事长、中国计算机学会和中国自动化学会副理事长、国家高技术研究发展计划（863 计划）信息技术领域专家委员会成员、国际权威期刊《Pattern Recognition》、《Image and Vision Computing》以及国内核心期刊《自动化学报》、《中国图象图形学报》等多本刊物的副主编或编委。作为 IEEE 北京分会副主席和 IEEE 国际视觉监控系列研讨会创始人与会议主席，谭铁牛为众多的国际学术会议和国际学术期刊担当审稿人。目前主持多项由“973”计划、“863”计划、国家自然科学基金委、国家杰出青年科学基金、中科院“百人计划”等资助的科研项目。

## ● 刘挺



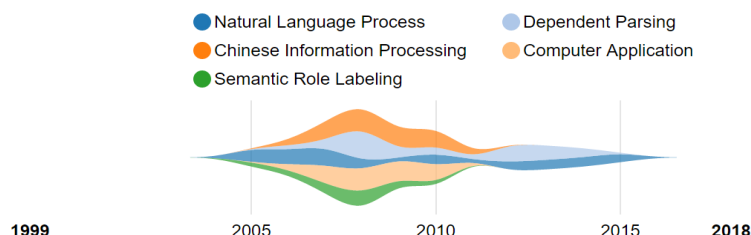
刘挺 (Ting Liu)

H 42 A 208.48 S 133.92 c 8870 P 480

Professor

哈尔滨工业大学计算机学院

Natural Language Process Dependent Parsing Chinese Information Processing Computer Application  
Semantic Role Labeling Search Engine Information Retrieval Word Sense Disambiguation



刘挺，哈尔滨工业大学教授，人工智能研究院副院长，计算机学院社会计算与信息检索研究中心主任，国家“万人计划”科技创新领军人才。他致力于攻克机器人理解文字语言情感的问题，探索文本情感分析这一重要技术。刘挺教授是国家“973”课题、国家自然科学基金重点项目负责人，多次担任国家“863”重点项目总体组专家、基金委会评专家，中国计算机学会理事，中国中文信息学会常务理事/社交媒体处理专委会（SMP）主任，曾任国际顶级会议 ACL、EMNLP 领域主席。2012-2017 年在自然语言处理领域顶级会议发表的论文数量列世界第 8 位（据剑桥大学统计）。他主持研制的“语言技术平台 LTP”、“大词林”等科研成果被业界广泛使用，曾获国家科技进步二等奖、省科技进步一等奖。

## ● 纪荣嵘



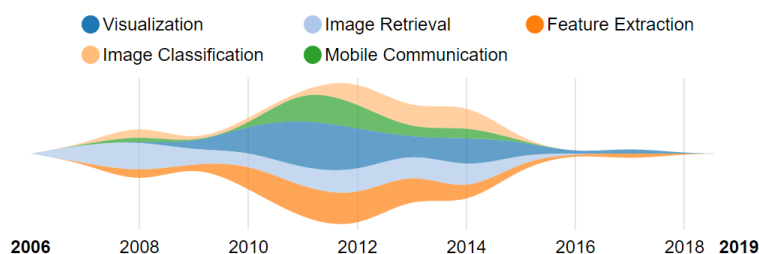
纪荣嵘 (Rongrong Ji)

H 41 A 301.46 S 166.22 c 7014 P 290

Professor

School of Information, Xiamen University

Visualization Image Retrieval Feature Extraction Image Classification Mobile Communication  
Indexing Semantics Computer Vision





纪荣嵘，福建省“闽江学者”特聘教授，厦门大学教授、博士生导师、信息学院院长助理，2015 届国家基金委优秀青年奖获得者。曾获得 2007 年微软学者奖、2011 年 ACM Multimedia 最佳论文奖、2015 年省自然科学二等奖，2016 年教育部技术发明一等奖。他担任 Neuro computing、Multimedia Tools and Applications、The Visual Computer、Frontiers of Computer Science 等国际期刊的副编辑，Valse（视觉与学习青年学者研讨会）2017 大会主席、IEEE MMSP 2015、ACM ICMR 2014、IEEE VCIP 2014、ACM MMM 2015、IEEE ISM 2015 等国际会议本地\专题\领域主席、AAAI 2015、CVPR 2013、ICCV 2013、ACM Multimedia 2010-2015 等四十余个国际会议的程序委员会委员、IEEE Multimedia Communication 技术委员会 Visual Analysis and Content Management for Communications 专家组主席（2014 至今）。此外，他还担任 IEEE 高级会员、ACM 高级会员、中国计算机学会会员、中国计算机学会计算机视觉专业组委员、中国计算机学会多媒体技术专委会委员、中国图形图像学会青工委委员、YOCSEF 厦门分论坛候任主席、福建省人工智能学会常务理事。

## ● 张艳宁



张艳宁 (Yanning Zhang)

H 32 A 265.46 S 122.92 c 4566 P 775

教授

西北工业大学计算机学院

Feature Extraction

Wavelet Transform

Support Vector Machine

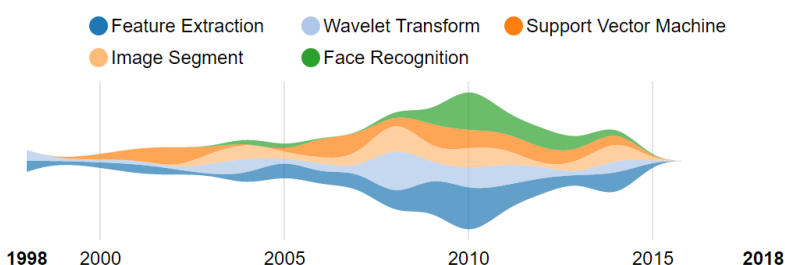
Image Segment

Face Recognition

Object Detection

Genetic Algorithm

Computer Vision

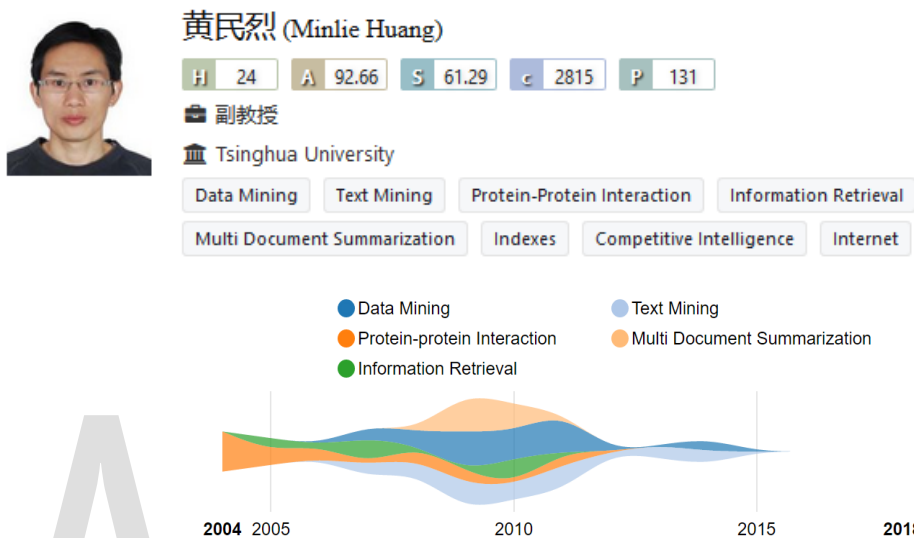


张艳宁，西北工业大学教授，计算机学院院长，是“长江学者”特聘教授、国防 973 项目首席、中组部首批“万人计划”科技创新领军人才入选者。她长期致力于图像处理、模式识别、计算机视觉与智能信息处理等的研究，并与航天、航空等方面的国家重大需求相结合。先后承担国防“973”项目、国家自然科学基金重点项目、国家/国防“863”项目、总装预研等国家级项目 40 余项。在 IEEE TPAMI、IEEE TIP、PR、IEEE TSMC-B、Info. Fusion、CVPR、ICCV 等国内外本领域权威期刊和重要国际会议上发表论文百余篇，出版专著 3 部。获国家/国防授权发



明专利 50 余项，陕西省科技进步一等奖 1 项，国防技术发明一等奖 1 项，国家教学成果二等奖 1 项、省部级科技进步奖 3 项。

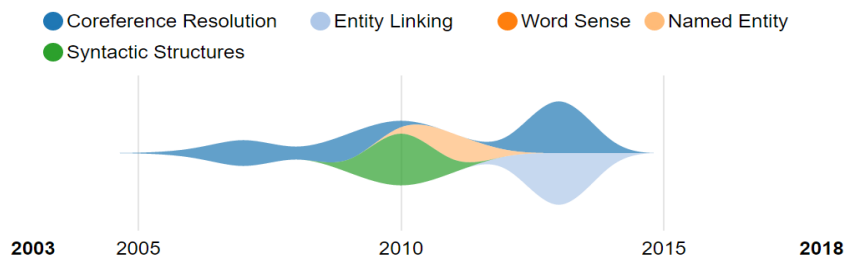
● 黄民烈



黄民烈，清华大学计算机科学与技术系副研究员。研究小组曾多次参加信息处理领域的国际评测 TREC/TAC，在多文档摘要、文本推理等任务上取得了国际领先成果。在多文档摘要方向上，提出了基于信息距离的摘要方法；同时还在观点排序、观点分类、多标签分类上提出了新的方法。开发了多个多文档摘要的算法，在多种场景下，授权给多个公司使用。黄民烈在情感分析与评论挖掘方面进行了深入研究，在观点抽取、情感极性分类、观点摘要等多方面取得了大量成果，并开发了 cReviewMiner 的中文产品评论深度挖掘系统。相关成果发表在 ACL、IJCAI、AAAI、COLING、NA-AACL、ICDM、CIKM、PAKDD 等国际会议及 Knowledge and Information System、Genome Biology、Bioinformatics、BMC Bioinformatics 等期刊上。作为 PC 成员或领域主席，他服务于 ACL、EACL、COLING、EMNLP（领域主席）、NAACL、CIKM、ICDM 等国际会议，担任期刊 BMC bioinformatics 的副编委，Bioinformatics, Knowledge-based Systems, TALIP 的审稿人。

● 亲兵





秦兵，哈尔滨工业大学计算机学院教授、博士生导师，哈尔滨工业大学社会计算与信息检索中心副主任。同时担任中国中文信息学会语言与知识计算专委会副主任，社交媒体处理专委会情感分析专业组组长。秦兵教授多年来一直从事自然语言处理研究。在顶级国际会议 ACL、EMNLP、IEEE TKDE 等国内外重要期刊及会议上发表论文 60 余篇，她带领的团队在 SemEval 2014 “Twitter Sentiment Analysis” 评测任务获第二名。此外，她还担任多个会议领域主席以及多个期刊和会议的审稿人，主持多项国家自然科学基金以及国家科技部“863”项目，同时多项研究成果进入企业产品。曾获黑龙江省科技进步一等奖等奖项。

## ● 贾珈



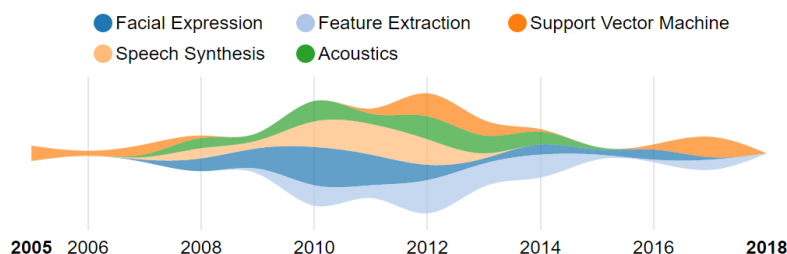
贾珈 (Jia Jia)

H 20 A 121.93 S 66.24 c 1334 P 140

Associate Professor

清华大学计算机科学与技术系

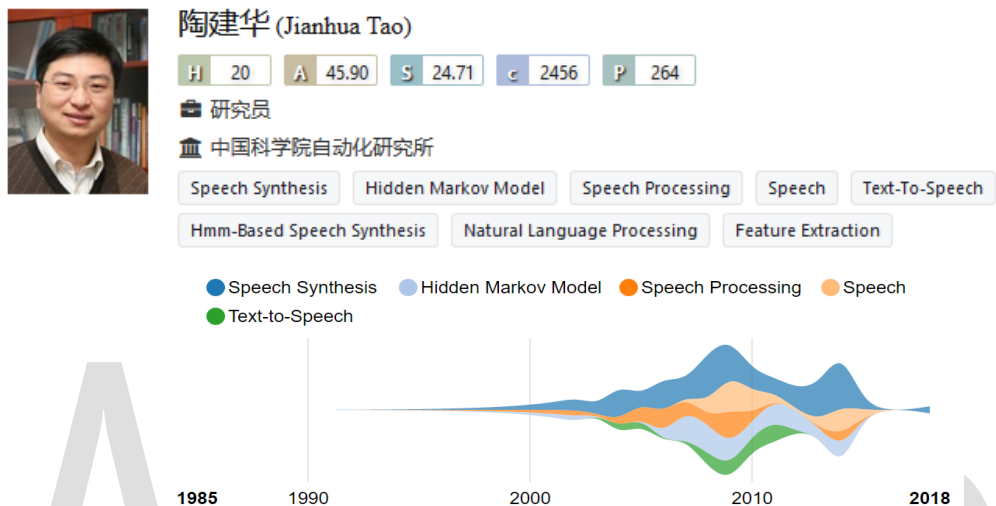
Facial Expression Feature Extraction Speech Synthesis Support Vector Machine  
Natural Language Processing Speech Acoustics Speech Recognition



贾珈，清华大学计算机系副教授，中国计算机学会高级会员，IEEE、ACM、International Speech Communication Association (ISCA) 会员，担任中文信息学会语音专业委员会秘书长、中国中文信息学会青年工作委员会委员、中国计算机学会青年工作委员会通讯委员、全国人机语音通讯学术会议常设机构委员、中国图象图形学会多媒体专委会委员、以及全国信标委用户界面分委会语音交互工作组副组长。她在 IEEE Transactions on Affective Computing、IEEE Transactions on Audio Speech and Language Processing、IEEE Transactions on Multimedia、ACM Multimedia、AAAI、IJCAI 等领域内主流学术刊物和会议上发表论文 60 余篇，并与腾讯、SOGOU、华为、西门子、

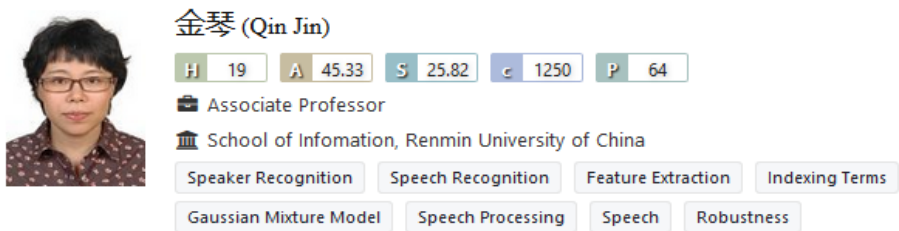
MSRA、BOSCH 等国内外同领域企业保持密切合作。贾珈曾获得 2009 年教育部科技进步奖和 2016 年教育部科技进步奖（第一完成人），2012 年 ACM Multimedia Grand Challenge Prize，2018 年 ACM Multimedia Best Demo Award。

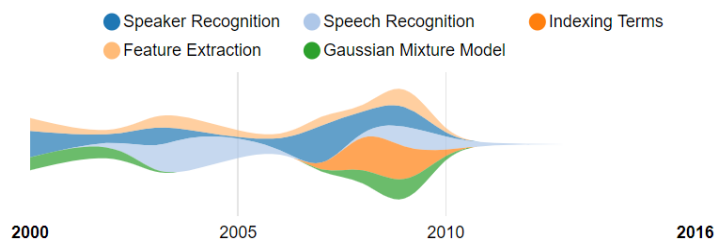
● 陶建华



陶建华，中科院自动化所模式识别国家重点实验室副主任，国家杰出青年基金获得者。目前还担任 ISCA SIG-CSLP 副主席、HUMAINE 学会执行理事、中国计算机学会常务理事、中国人工智能学会理事、中国中文信息学会理事、中国声学学会理事、中重中文信息学会语言资源建设与管理委员会秘书长等职务。他先后负责和参与国家级项目（“863”计划、国家自然科学基金、国际合作）40 余项，在 IEEE Trans、ICASSP、ICSLP、ICCV、ICIP 等国内外学术期刊和会议上发表论文 200 余篇，申请国内发明专利 15 项，国际专利 1 项，编著英文学术著作 2 部，两次获得北京市科技进步二等奖等。

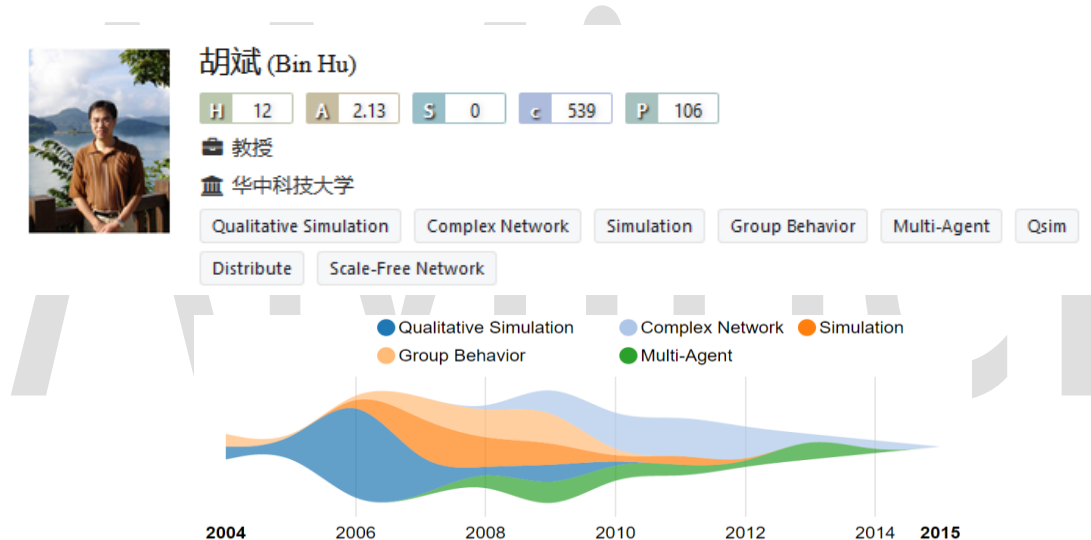
● 金琴





金琴，中国人民大学信息学院副教授，博士生导师，在国内外一流期刊和会议上发表百余篇论文。金琴老师团队在 2017-2018 连续两年获得 Audio-Visual Emotion Challenge AVEC 竞赛第一名，2016-2017 连续两年蝉联 ACM Multimedia (Video to Language) Grand Challenge 第 1 名，2017 获得 NIST TrecVID Video to Text (VTT) 国际竞赛第 1 名，2016 年获得 IBM SUR Award，2015 年获得 Image CLEF 评测 (Image Sentence Generation) 第 1 名。

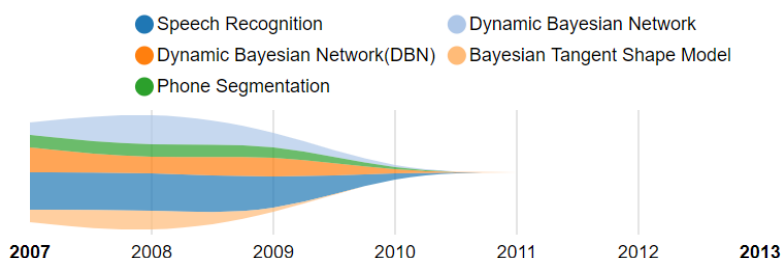
## ● 胡斌



胡斌，华中科技大学管理科学与信息管理系现代化管理研究所所长。他主持国家自然科学基金重点项目1项，国家自然科学基金项目4项，参加国家自然科学基金重点项目2项。共发表7部专著，上百篇论文。曾获得上海市科技进步三等奖、湖北省自然科学三等奖、中国大学出版社图书奖首届优秀学术著作二等奖等。

## ● 蒋冬梅





蒋冬梅，西北工业大学计算机学院教授。2005 年 11 月起担任西工大与布鲁塞尔自由大学“听视觉信号处理联合实验室”的中方联系人。她主持多项国家自然科学基金、陕西省自然科学基金、国际合作项目，以及其他应用类课题，在国内外学术刊物和会议上发表学术论文 30 多篇。主要研究方向为听视觉语音信号处理，包括听视觉语音识别与合成、说话人面部动画、听视觉情感描述与表达、三维虚拟声场，以及音视频内容分析与检索等。蒋冬梅教授所处的西北工业大学音频、语音与语言处理研究组隶属于陕西省语音与图像信息处理重点实验室。研究组成立于 1995 年，经过近 20 多年的快速发展，已形成了人机语音交互、语音与音频信号处理、音视频多模态信息处理、多媒体内容分析等主要研究方向。

### 3.3 研究机构

目前国内从事情感计算研究的代表性实验室主要有以下机构：

- **清华大学人机交互与媒体集成研究所**

清华大学人机交互与媒体集成研究所简称“媒体所”，在媒体信息智能处理、人机交互、普适计算等方面开展高水平研究，现任所长为史元春教授。媒体所建有多个学术基地，如清华信息科学与技术国家实验室普适计算研究部（普适计算教育部重点实验室）、清华大学计算机系—腾讯互联网创新技术联合实验室、网络多媒体北京市重点实验室等。媒体所近年主持多项本学科领域重要项目（“973”、“863”重点、NSFC重点等），在顶级刊会发表大量学术论文并获多篇最佳论文奖，获得国家级科技奖励5项。主要研究方向：智能信息处理（计算机视觉、图形学、多媒体编码与检索等）、和谐人机交互（情感计算、语音交互等）、普适计算环境（普适计算模式、主动服务等）。

- **中国科学院模式识别国家重点实验室**

中科院模式识别国家重点实验室筹建于 1984 年，1987 年 8 月正式对外开放，是由国家计委投资筹建的第一批国家重点实验室之一。实验室依托于中国科学院自动化研究所，第一任主任是马颂德研究员，第二任主任是谭铁牛院士，现任主任是刘成林研究员。实验室目前承担着百余项科研项目，其中包括国家重点基础研究计划“973”项目、国家自然科学基金重大项目、

国家高技术计划“863”项目，国家科技支撑计划项目及国际合作项目等。近年来，实验室获国家科技进步二等奖 1 项，中科院自然科学二等奖 2 项。主要研究方向：模式识别基础理论（认知机理与计算模型、模式分类与机器学习、模式描述与结构理解等）、图像处理与计算机视觉（三维视觉和场景分析、文档图像分析等）、语音与语言信息处理（语音识别、话语理解等）、脑网络组研究（脑网络拓扑结构、脑网络的动力学属性等）。

#### ● 哈尔滨工业大学社会计算与信息检索研究中心

哈尔滨工业大学社会计算与信息检索研究中心（HIT-SCIR）成立于 2000 年 9 月，研究中心主任是刘挺教授。已完成或正在承担的国家 973 课题、国家自然科学基金重点项目、国家 863 重点项目、国际合作、企业合作等课题 60 余项。在这些项目的支持下打造出“语言技术平台 LTP”，获 2010 年钱伟长中文信息处理科学技术一等奖。主要研究方向包括语言分析、信息抽取、情感分析、问答系统、社交媒体处理和用户画像 6 个方面。

#### ● 华中科技大学智能媒体计算与网络安全实验室

华中科技大学智能媒体计算与网络安全实验室近五年来承担了国家重点研发计划、国家支撑计划、国家自然科学基金、国家 863 重点项目、湖北省杰出青年基金、武汉市青年科技晨光计划、博士点基金和和企业合作项目等 50 余项科研项目，在计算机学报、IEEE Transaction on Parallel and Distributed System、IEEE Computer Graphics and Application、IEEE Transaction on Multimedia、等国内外权威期刊和重要国际会议上发表论文 150 余篇，其中被 SCI、EI 权威索引论文 100 余篇次，获得发明专利和软件著作权 50 余项，获得湖北省科技进步二等奖 1 项、日内瓦发明奖 1 项。目前主要的研究方向包括数字媒体处理与检索、视频情感计算、增强现实、网络安全与大数据处理、多核计算与流编译和教育信息化等。

#### ● 合肥工业大学情感计算与系统结构研究所

合肥工业大学情感计算研究所于 2011 年成立，主任为孙晓教授。主要从事先进智能、情感计算、大规模数据与知识获取的基础理论研究工作。研究所构建了丰富的数据资源，包括文本语料库、面部表情库、动作-情感库等，其中全球最大规模的中文情感语料库，已授权全球近 300 家高校、科研机构使用。构建了基于大规模多源情感数据库的情感感知、推理与交互的总体研究思路与方法体系，取得了系列突破性研究成果，使我国在先进智能及情感机器人达到了世界先进水平。主持并参与国家 973 预研项目、863 项目，国家自然科学基金项目、安徽省自然科学基金项目、企业委托项目等多项。2011 年 12 月获批建设“情感计算与先进智能机器安徽省重点实验室”，系统展开情感计算与先进智能机器的研究，是国内首个以情感计算命名

的重点实验室。研究方向包括：情感计算；自然语言处理；听觉信息认知计算；视觉信息认知计算；情感可穿戴计算；机器人云理论及其应用。

- 东南大学情感信息处理实验室

东南大学情感信息处理实验室（Affective Information Processing Lab, AIPL）创建于 2004 年，隶属于东南大学生物科学与医学工程学院和儿童发展与学习科学教育部重点实验室（东南大学），任为郑文明。实验室深耕情感计算领域，主持包括 973 计划、国家自然科学基金重点项目在内的多项国家和省部级课题，在包括 IEEE Transactions 系列期刊和 ICCV, PR, ECCV, NIPS, IJCAI 和 AAAI 等计算机领域顶级会议上已发表论文百余篇，部分研究成果获国家技术发明二等奖 1 项（2018 年）、教育部自然科学二等奖 2 项（2015 年和 2008 年）和江苏省科技进步二等奖 1 项（2009 年）。研究方向包括：情感计算、模式识别、计算机视觉和机器学习及其在儿童智能发展、教育和医疗等方面的应用研究。

AMiner



## 4 应用篇

近年来, Picard 领导的美国麻省理工学院多媒体实验室相继提出了近 50 种情感计算应用项目。例如, 将情感计算应用于医疗康复, 协助自闭症者, 识别其情感变化, 理解患者的行为; 在教育中应用情感计算, 实现对学习状态的采集及分析, 指导教学内容的选择及教学进度进行; 还可以将情感计算应用于生活中, 计算机能够感知用户对音乐的喜好, 根据对情感反应的理解判断, 为用户提供更感兴趣的音乐播放等。本篇将对情感计算在课堂教学、情感检测和医疗康复中的运用做详细介绍。

### 4.1 课堂教学

#### ● 面部识别测量学生理解程度

在美国, 公立学校的预算限制引发大规模的教师裁员和教室拥挤不堪。教师工作时间紧张, 还要考虑和满足每个学生的需求。结果就是, 那些课业困难的孩子容易受到忽视。因为只要孩子不提出问题, 老师就不会关注到他。

在过去三年里, 有企业把面部识别技术应用到了第一线教学当中。在 SensorStar 实验室, 他们用相机捕捉学生上课反应, 并且输入到计算机里面, 运用算法来确定学生注意力是否转移。通过面部识别软件 EngageSense, 计算机能够测量微笑、皱眉和声音来测定学生课堂参与度。孩子们的眼睛是专注于老师的吗? 他们是在思考还是发呆? 他们是微笑还是皱着眉头? 或者他们只是觉得困惑? 还是无聊? 测量之后, 老师将会收到一份反馈报告, 基于面部分析, 报告会告诉老师他们的学生学习兴趣何时最高、何时最低。这样, 老师能够对自己的教学方案做出调整, 满足更多学生的需求。此外, 比尔和梅林达盖茨基金会资助了传感器手镯 (sensor bracelets) 的开发, 这可以用来追踪学生的参与水平。腕部设备能够发送小电流, 通过在神经系统响应刺激时测量电荷的细微变化便可以得知学生的课程兴奋程度。

心理学家 Paul Ekman 将面部识别技术研究提升到了一个新的层次。他对 5000 多种面部运动进行了分类, 以帮助识别人类情绪。他的研究为 Emotient Inc、Affectiva Inc 和 Eyeris 等公司提供了帮助, 这些公司将心理学和数据挖掘相结合, 检测人的细微表情, 并对人的反应进行分类。目前为止, 面部识别技术的重点是协助联邦执法和市场调研。不过, 圣地亚哥市的研究人员也在医疗行业试用这项技术, 测定孩子接受外科手术之后的疼痛程度<sup>[13]</sup>。

#### ● 机器学习定制学生课堂学习内容



TechCrunch 公司的员工设计了在线教育平台，来提供一对一指导和精熟学习（mastery learning）。这是应用创新型思维，通过实时的评估和定制化的学习方式，有效地解决本杰明提出的著名的“Sigma 2 Problem”。深度学习系统将学生学习效果数据进行分类，并且在此基础上制定相关的教学内容。该系统还可以推荐附加练习，并且根据学生个人能力和教学要求，实时推荐课程内容，调整教学速度<sup>[14]</sup>。

北卡罗来纳州州立大学研究员开发了一种软件，通过摄像头捕捉和分析学生面部表情，以此改变在线课程。目前，大多数情感计算技术还仅仅停留在学术研究领域。但也已经有公司开始应用这项技术，并能成功地分辨学生表情，并根据他们的学习能力和方式，来自动调整适合的学习内容和环境。英特尔公司正是这其中的一员。有了这些学生表情数据，可以让“Emoshape”这样的情感计算智能系统，自动分析情感，并做出适当回复。这些系统具备了解决个体问题的能力，也使老师能够提供高度个性化的内容来激发学生的学习兴趣。

人工智能和大数据已经促成了大部分行业的技术革新，从电子商务到交通、金融、医疗。人工智能和大数据已经在教育方面取得进展。尽管有些反对的声音，比如说如何保护学生隐私、如何提高教学效率等，但需要指出的是，这些技术的应用并不是要代替老师，而是扮演辅助老师的角色，识别学生的个体需求，以制定更加智能的教学方案。

## 4.2 情绪监测

为了深度挖掘人类情感的奥秘，美国麻省理工学院计算机科学与人工智能实验室打造了用无线信号监测情绪的 EQ-Radio<sup>[15]</sup>。在没有身体感应器和面部识别软件辅助的情况下，EQ-Radio 通过测量呼吸和心跳的微小变化，利用无线信号捕捉到一些肉眼不一定能察觉的人类行为，判断一个人到底处于以下四种情绪中的哪一种：激动、开心、生气或者忧伤，正确率高达 87%。美国麻省理工学院教授和该项目的负责人 Dina Katabi 预测，这个系统会被运用于娱乐、消费者行为和健康护理等方面：电影工作室和广告公司也可以用这个系统来测试观众实时的反应；而在智能家居的环境中，该系统可以通过捕捉与人的心情有关的信息，调节室内温度，或者建议你应该呼吸一些新鲜空气。

现有的情绪监控方法大多依赖于视听设备或者是安装在人身上的感应器，这两种技术都有缺点：面部表情并不一定符合内心状态，而安装在身上的感应器（比如胸带和心电监护仪）会造成各种不便，而且一旦它们的位置稍微移动，监测到的数据就不精确了。

EQ-Radio 会发送能监测生理信息的无线信号，该信号最终会反馈给设备本身。其中的算法可以分析心跳之间的微小变化，从而判断人们的情绪。消极情绪会被判定为“忧伤”，而正面

且高涨的情绪会被判定为“激动”。尽管这样的测量会因人而异，但其中还是有内在统一性。通过了解人们处于不同的情绪状态下，他们的心跳会如何变化，我们就可以对他所处的情绪状态进行有效的判断。

在他们设计的实验中，参与实验者选择他们记忆中最能代表激动、开心、生气、忧伤以及毫无情感的一段视频或音乐。在掌握了这段时长两分钟的视频里的五种情绪设置后，EQ-Radio可以精确地通过一个人的行为判断他处于这四种情绪中的哪一种。与微软研发的基于视觉和面部表情的 Emotion API 相比，EQ-Radio 在识别喜悦、忧伤和愤怒这三个情绪上精确度更高。同时，这两种系统在判断中性情绪时的精准度差不多，因为毫无情绪的脸总是更容易被识别。

目前，对美国麻省理工学院计算机科学与人工智能实验室而言，最艰巨的任务就是摆脱不相关数据的干扰。比如，为了分析心率，他们要抑制呼吸可能带来的影响，因为呼吸时，人的肺部起伏比他心跳时的心脏起伏要大。

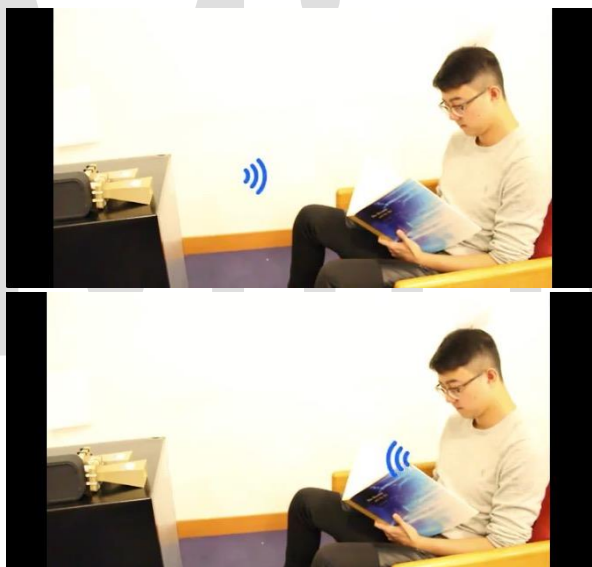


图 9 EQ-Radio 无线信号监测

### 4.3 医疗康复

近年来，情感计算运用于自闭症治疗得到越来越多的关注。例如，美国麻省理工学院情感计算团队正在开发世界上第一个可穿戴的情感计算技术设备：一个具有社交智能的假肢，用来实时检测自闭症儿童的情感，帮助机器人使用自闭症儿童独有的数据，来评估这些互动过程中每个孩子的参与度和兴趣。这个装置用一个小型照相机，分析孩子的面部表情和头部运动来推断他们的认知情感状态。还有一种叫“galvactivator”的工具，通过测量穿戴者的皮肤电流数据，推断孩子的兴奋程度。这个像手套一样的设备可以利用发光二极管描绘出人体生理机能亢奋程

度的图谱。这种可视化的展现方式，能够清晰地展示出人的认知情感水平。NAO 机器人和个性化的机器学习在治疗自闭症患者上也表现出很大的优越性：

### ● NAO机器人

人类治疗师会向孩子展示一张照片或者闪存卡片，用来表示不同的情绪，以教会他们如何识别恐惧、悲伤或喜悦的表情。治疗师随后对机器人进行编程，向孩子们展示这些相同的情绪，并且在孩子与机器人交往时观察孩子。孩子们的行为提供了宝贵的反馈信息，机器人和治疗师可以根据反馈信息继续学习。

研究人员在这项研究中使用了 SoftBank Robotics NAO 类人机器人。NAO 将近 2 英尺高，类似于装甲超级英雄，通过改变眼睛的颜色、肢体的运动以及声音的音调来表达不同的情绪。参加这项研究的 35 名自闭症儿童中，有 17 人来自日本，18 人来自塞尔维亚，年龄从 3 岁到 13 岁不等。他们在 35 分钟的会议中以各种方式对机器人做出反应，从看起来无聊和困倦，到在房间里兴奋地跳来跳去，拍手，大笑或触摸机器人。研究中的大多数孩子对机器人的看法是，它不仅仅是一个玩具，应该尊重 NAO，因为它是一个真实的人。另外，人类用许多不同的方式改变自己的表情，但机器人则通过同样的方式来改变表情，这对孩子来说更加有利，因为孩子可以通过非常有条理的方式学习如何表达表情<sup>[16]</sup>。



图 10 NAO 机器人

### ● 个性化的机器学习

麻省理工学院的研究小组意识到，具有深度学习能力的治疗机器人能够更好感知儿童的行为的。深度学习系统使用分层的多层数据处理来处理其任务，每一个连续的层都是对原始数据抽象的表示。

尽管自 20 世纪 80 年代以来深度学习的概念已经出现，但直到最近才有足够的计算能力来实现这种人工智能。深度学习已被用于自动语音和对象识别程序中，这种应用非常适合解决面部、身体和声音等多重特征的问题，从而更好地理解抽象的概念，如儿童的参与感。

对于治疗机器人，研究者构建了一个个性化框架，可以从收集的每个孩子的数据中学习。研究人员拍摄了每个孩子的脸部表情、头部和身体动作、姿势和手势，记录了儿童手腕上显示器的心率、体温和皮肤汗液反应作为数据。这些机器人的个性化深度学习网络是根据这些视频、音频和生理数据的层次，针对孩子的自闭症诊断和能力、文化和性别的信息构建的。研究人员将机器人对儿童行为的估计与五位人类专家的估计数字进行了比较，这些专家连续对孩子的录像和录音进行编码，以确定孩子在会议期间高兴或不安程度，是否感兴趣以及孩子的表现。比较发现，机器人对儿童行为的估计要比专家更加具体清晰。

## 4.4 舆情监控

网络调查法、统计规则法和文本内容挖掘是三种经常被使用的网络舆情分析方法。大数据时代的来临使传统的舆情分析方式发生改变，大数据时代数据量突增、数据产生的速度极快、冗余信息占比高的特性不仅给舆情分析带来新的发展机遇，也带来了新的难度和挑战。基于简单调查和统计的舆情分析方法将无法适用于大数据环境下的网络社区文本。当前国内外对舆情分析技术的研究也大多以大数据环境为背景，与传统舆情分析技术相比，大数据时代网络社区的舆情分析技术更多地集中于对数据的获取，并采取文本数据分析、数据挖掘、语义分析等技术获取舆情信息。当前国内外的舆情分析技术研究主要集中于话题识别与话题跟踪、意见领袖识别以及情感倾向判别这三个方面。

话题识别与话题跟踪首先在文本中识别出新话题，接下来在一段时间内检测并实时跟踪话题，实现该话题的再现，研究其随时间发展的演化过程。聚类方法常用于进行话题识别。在国外研究中，话题检测与跟踪（TDT）是了解社交媒体热点话题及其演变过程的重要手段。

意见领袖的发现和识别重点在于评价指标的制定以及模型的构建。例如，曹玖新等将网络社区用户看作一个个节点，根据节点之间信息的交互和传播过滤，从用户结构、行为和情感三个特征维度挖掘意见领袖。

情感倾向判别在舆情研究中最常见，首先收集 web 金融领域的文本数据属性，接下来构建金融领域的情感词典，最后结合语义分析，将语义规则应用到情感及情感强度识别当中，提升了分类器的准确率 M。王永等人将倾向分析应用到客户评论信息挖掘当中，结合情感词之间的依存关系计算面向产品特征的情感倾向得分，从网络评论中获取有价值的商业信息。国外针对 Twitter 的情感倾向分析研究居多，用以获取有价值的信息和舆论导向，例如，结合语言规则特征可以分别获取正面和负面的 Twitter 文章，反应公众的舆情态度<sup>[17]</sup>。

## 5 趋势篇

### 5.1 论文研究发展趋势

Trend analysis (<http://trend.aminer.cn>) 基于 AMiner 的 2 亿篇论文数据进行深入挖掘, 包括对技术来源、热度、发展趋势进行研究, 进而预测未来的技术前景。技术趋势分析描述了技术的出现、变迁和消亡的全过程, 可以帮助研究人员理解领域的研究历史和现状, 快速识别研究的前沿热点问题。

下图是当前情感计算领域的热点技术趋势分析, 通过 Trend analysis 分析挖掘可以发现当前该领域的热点研究话题 Top10 是 Affective Computing、Social Robot、Emotion Recognition、Human Computer Interaction、Feature Extraction、Support Vector Machine、Facial Expression、Human Robot Interaction、Behavioural Sciences Computing、Face Recognition。

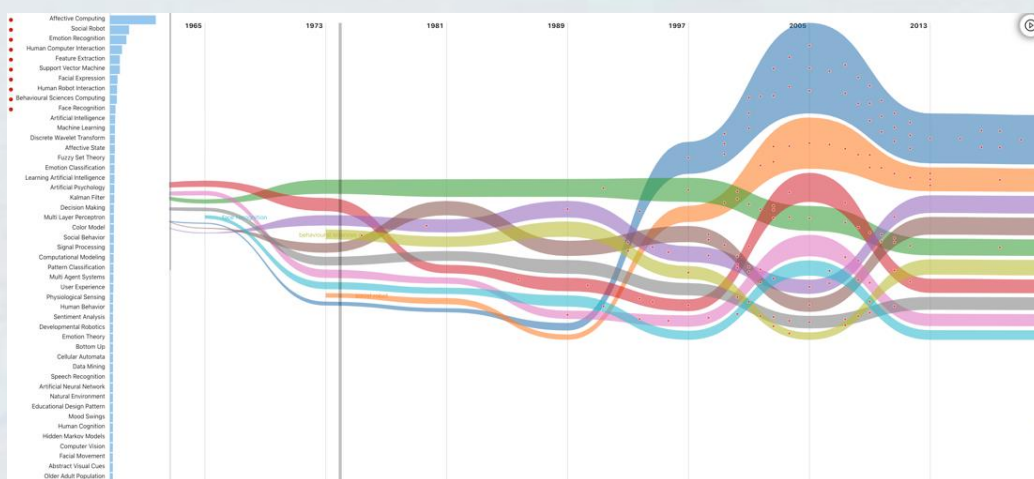


图 11 情感计算发展趋势

根据 Trend analysis 的分析我们可以发现, 该领域当前最热门的话题是 Affective Computing, 从全局热度来看, Affective Computing 的话题热度虽然有所起伏, 但从 20 世纪 90 年代开始, 热度迅速上升, 甚至在五年内超过了此前的话题 Top 1 Emotion Recognition, 并且至今其话题热度始终保持在 Top1, 论文的发表数量也较多; Social Robot 的研究热度跟随 Affective Computing 同期上升, 近几年话题热度更是超越 Emotion Recognition 成为 Top2 话题; 另外, 前期比较热门的 Feature Extraction 经过了一段时间的低迷期后, 也回到了 Top3 的位置。



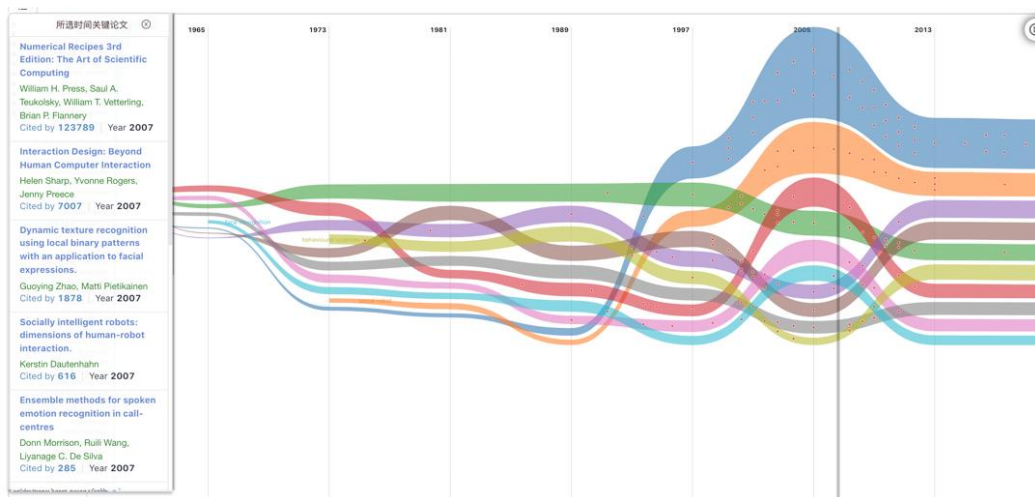


图 12 情感计算 2007 年经典论文

通过拖动趋势分析页面的时间轴，我们可以在上图左框中看到 2007 年同期的经典论文，引用量前五的论文分别如下：

1. 论文标题: *Numerical Recipes 3rd Edition: The Art of Scientific Computing*

论文作者: William H. Press, Saul A. Teukolsky, William T. Vetterling, Brian P. Flannery

引用量: 123789

2. 论文标题: *Interaction Design: Beyond Human Computer Interaction*

论文作者: Helen Sharp, Yvonne Rogers, Jenny Preece

引用量: 7007

3. 论文标题: *Dynamic texture recognition using local binary patterns with an application to facial expressions*

论文作者: Guoying Zhao, Matti Pietikainen

引用量: 1878

4. 论文标题: *Socially intelligent robots: dimensions of human-robot interaction*

论文作者: Kerstin Dautenhahn

引用量: 616

5. 论文标题: *Ensemble methods for spoken emotion recognition in call-centres*

论文作者: Donn Morrison, Ruili Wang, Liyanage C. De Silva

引用量：285

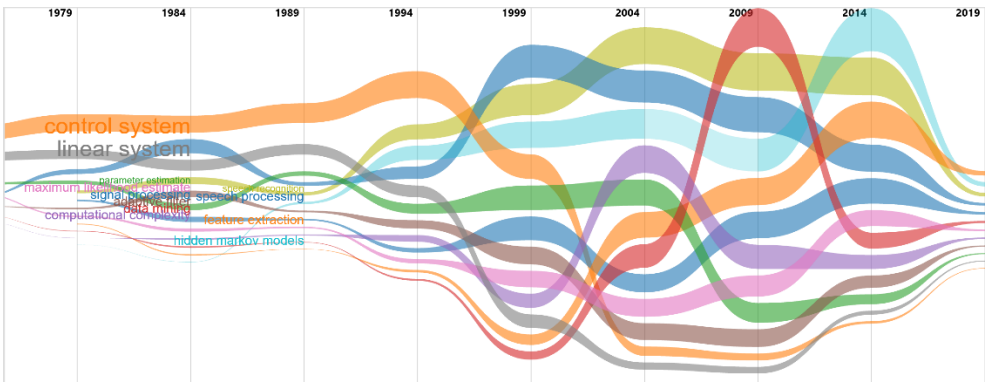


图 13 1970-2019 论文研究方向趋势图

AMiner 通过对 1970 年至 2019 年发表的情感计算相关论文进行数据分析，绘制出论文研究方向的发展趋势图。数据分析的论文主要来自于情感计算领域的顶级学术会议，如 ACMMM、KDD 的相关 workshop。从图像中可以看出，随着时代发展及研究的深入，control system 方向的论文数量在逐渐减少，speech recognition、hidden markov model、signal process、feature extract 等方面的研究成果不断增加。具体的论文关键词请见下表：

表 13 1970-2019 各阶段前十位研究关键词

1970-1979 年	1980-1989 年	1990-1999 年	2000-2009	2010-2019
control system	control system	control system	speech recognition	hidden markov model
optimal control	linear system	speech recognition	signal process	feature extraction
linear system	transfer function	signal process	data mining	speech recognition
automatic control	signal process	hidden markov model	hidden markov model	signal to noise ratio
transfer function	optimal control	parameter estimation	feature extraction	signal process
differential equations	automatic control	speech process	computational complexity	speech process
eigenvalues and eigenfunctions	digital filter	adaptive filter	parameter estimation	computational modelling
stochastic process	sufficient condition	neural network	speech process	compressed sensing
state estimation	eigenvalues and eigenfunctions	linear system	maximum likelihood estimation	support vector machines
stochastic system	stochastic process	transfer function	adaptive filter	noise measurement

## 5.2 情感计算技术预见

研究者根据情感计算领域近十年的相关论文，利用大数据分析、机器学习、人工智能等技术手段，建立算法模型及研发 demo 系统，分析挖掘出该领域的技术发展热点。

技术预见图中点的大小表示该技术的热点（主要由相关论文数量的多少决定，相关论文越多，热度越高，点越大），各技术之间的连线表示 2 个技术关键词同时在 N 篇论文中出现过（当前 N 的取值为 5）。

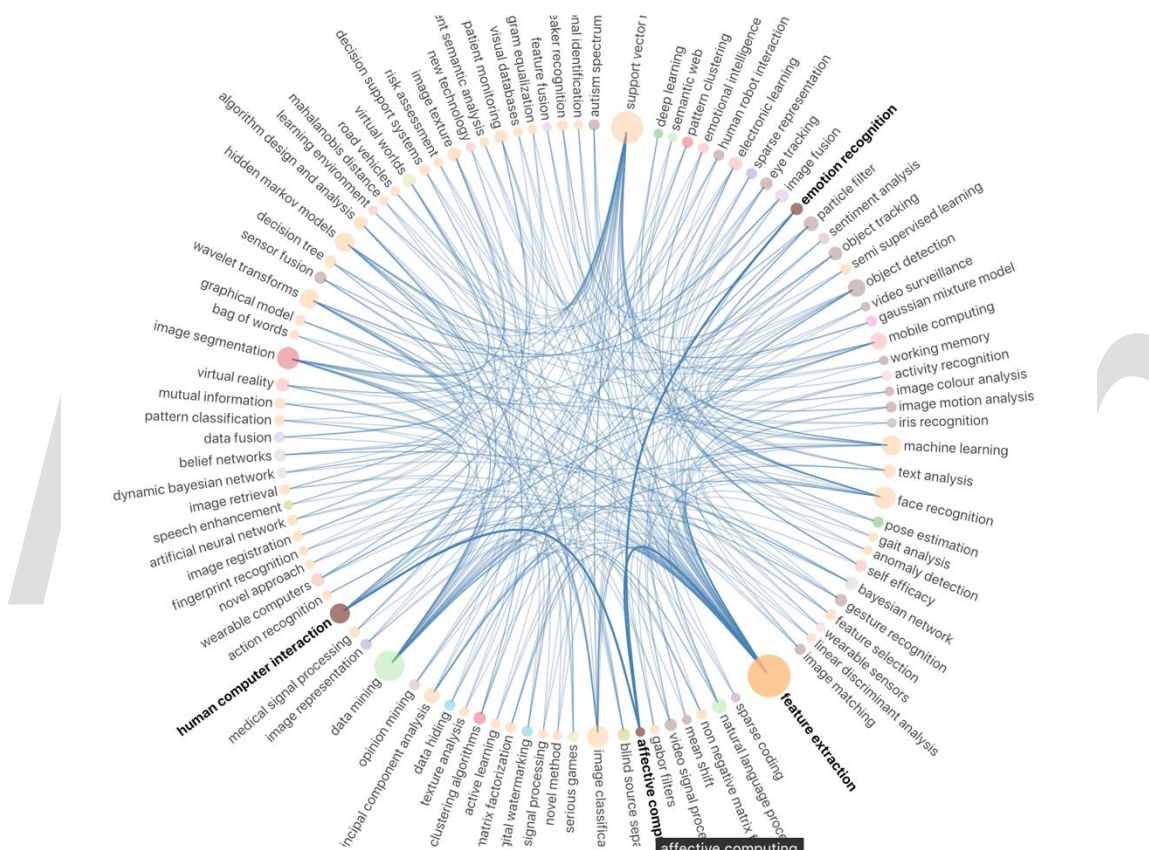


图 14 情感计算技术预见图

根据情感计算技术预见图（图 14），可以得出情感计算领域相关度最高的技术有 3 项，分别为：feature extraction、human computer interaction 和 emotion recognition。

按照技术前沿度，可以列出相关的主要技术关键词，以及该技术历年的变化趋势（论文发表数量变化趋势），及重要代表性成果。具体如下图所示：



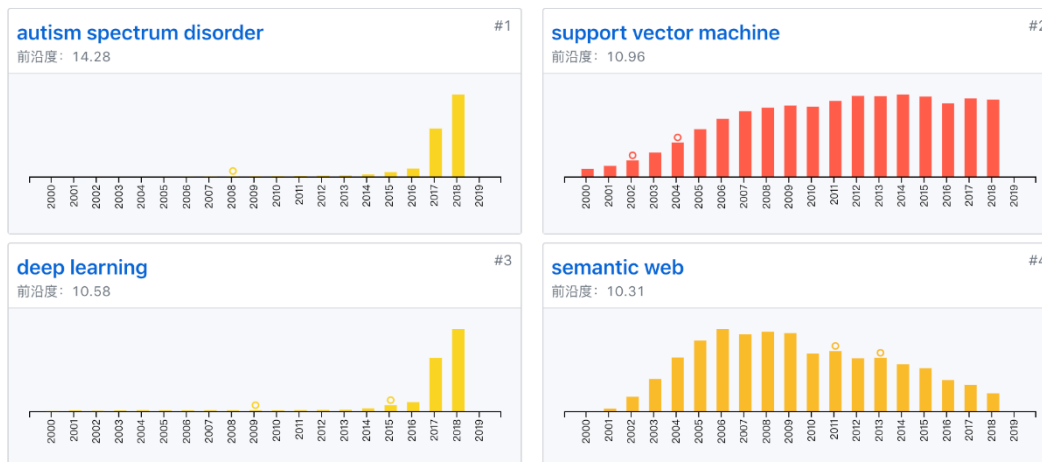


图 15 情感计算预测热词图

从图 15 中我们可以看出，情感计算领域预测前沿度比较高的前四热词有：autism spectrum disorder（前沿度为 1428）、support vector machine（前沿度为 1096）、deep learning（前沿度为 1058）和 semantic web（前沿度为 1031）。

## 5.3 情感计算未来发展

### ● 数据库的建立

现在已实现的情感计算大部分原型情感的识别来源单一。数据库本身存在短板，如训练分类的样本数少，体态识别大多依赖于一组有限的肢体表达（跳舞、手势、步态等），只关注内部效度而缺少外部效度。因此识别方面，未来研究应在情感分类方面继续努力，创建新的数据库，尤其是婴幼儿及儿童数据库的建立。

### ● 神经科学与情感计算的结合

神经科学方面，人类大脑情感过程的神经解剖学基础极其复杂并且远未被理解，因此该领域还不能为开发情感计算模型提供充足的理论基础。

### ● 情感的测试

人机交互或人与人交互过程中，人的情感变化是变速的。虽然 Picard 在《情感计算》一书中用两个比喻将情感复合分成两类：微波炉加热食物时开关间断循环与冷热水混合，两者通过不同方式的复合使物体达到“温”的状态，前者类似于“爱恨交加”所对应的情感状态，后者类似于 Russell 等人提出的环形情感模型中愉快与低强度结合为轻松的这种新状态。但基于不同情境下的情感复合远不止这两类。将动态的数字平台做成模型很难实现，因为情感的输入、输出应该在不同的情境下产生和测试。

目前国外已经有一部分研究者开始关注深度情感计算方面的研究，如 Ayush Sharma 等人利用语言数据联盟（Linguistic Data Consortium, LDC）中情绪韵律的语音和文本，研究韵律特征提取与分类的深层情感识别。随着后续情感方面的深度研究，多模型认知和生理指标相结合、动态完备数据库的建立以及高科技智能产品的加入等将会成为情感计算相关研究的新趋势，从而更好地实现对用户行为进行预测、反馈和调制，实现更自然的人机交互<sup>[18]</sup>。

# AMiner

## 参考文献

- [1] 王志良,解仑.我国人工心理与人工情感研究现状与进展[J].中国科学基金,2013,27(01):14-17.
- [2] Izard CE. Basic emotions, natural kinds, emotion schemas, and a new paradigm[J]. Perspectives on psychological science, 2007,2(3): 260-280.
- [3] Osgood CE. Dimensionality of the semantic space for communication via facial expressions[J]. Scandinavian journal of psychology, 1996,7(1):1-30.
- [4] Bradley MM, Lang PJ. Measuring emotion: the self-assessment manikin and the semantic differential[J]. Journal of behavior therapy and experimental psychiatry, 1994,25(1):9-59.
- [5] Backs RW, Da Silva SP, Han K. A comparison of younger and older adults' self-assessment manikin ratings of affective pictures[J]. Experimental aging research,2005,31(4):421-440.
- [6] 李佳源.情感计算的研究现状与认知困境[J].自然辩证法通讯,2012,34(02):23-28+125.
- [7] 张瑞.文本情感计算研究综述[J].管理观察,2017(13):28-30.
- [8] 文本情感分析有什么好资料、网站、工具推荐呢? <https://www.zhihu.com/question/20631050/answer/512693661>
- [9] 张石清,李乐民,赵知劲.人机交互中的语音情感识别研究进展[J].电路与系统学报,2013,18(02):440-451+434.
- [10] 韩文静,李海峰,阮华斌,马琳.语音情感识别研究进展综述[J].软件学报,2014,25(01):37-50.
- [11] Ververidis D, Kotropoulos C. A review of emotional speech databases[C]//Proc. Panhellenic Conference on Informatics (PCI). 2003, 2003: 560-574.
- [12] 钟新波.基于计算机视觉的情感计算[J].现代计算机(专业版),2008(07):44-47.
- [13] Does Emotive Computing Belong in the Classroom? <https://www.edsurge.com/news/2017-01-04-does-emotive-computing-belong-in-the-classroom>.
- [14] 在美国,情感计算是如何应用到课堂教学中的? <https://www.jianshu.com/p/6e79a65c1267>.
- [15] Detecting emotions with wireless signals. <http://news.mit.edu/2016/detecting-emotions-with-wireless-signals-0920>
- [16] MIT 将个性化深度学习网络应用于机器人,让 AI 更有效地评估和治疗自闭症儿童. <http://wemedia.ifeng.com/66986311/wemedia.shtml>.
- [17] 张璐.基于情感计算的网路社区舆情分析预警技术研究[D].北京邮电大学,2018.
- [18] 认知心理学与人工智能的交叉:情感计算综述. [http://www.sohu.com/a/83873777\\_297710](http://www.sohu.com/a/83873777_297710).

## 版权声明

AMiner 研究报告版权为 AMiner 团队独家所有，拥有唯一著作权。AMiner 咨询产品是 AMiner 团队的研究与统计成果，其性质是供用户内部参考的资料。

AMiner 研究报告提供给订阅用户使用，仅限于用户内部使用。未获得 AMiner 团队授权，任何人和单位不得以任何方式在任何媒体上（包括互联网）公开发布、复制，且不得以任何方式将研究报告的内容提供给其他单位或个人使用。如引用、刊发，需注明出处为“AMiner.org”，且不得对本报告进行有悖原意的删节与修改。

AMiner 研究报告是基于 AMiner 团队及其研究员认可的研究资料，所有资料源自 AMiner 后台程序对大数据的自动分析得到，本研究报告仅作为参考，AMiner 团队不保证所分析得到的准确性和完整性，也不承担任何投资者因使用本产品与服务而产生的任何责任。

AMiner

顾 问：金 琴 唐 杰

责任编辑：刘 佳

编 辑：景 晨

封面设计：边云风

