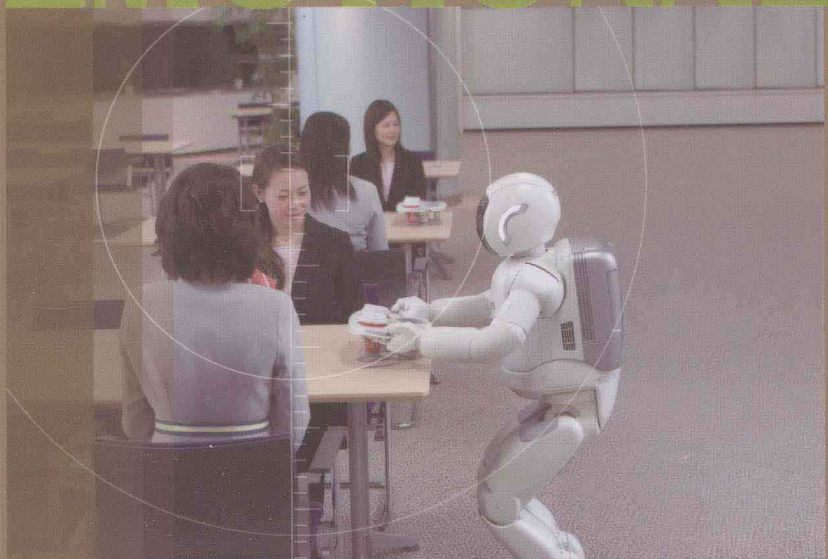


EMOTIONAL



图像情感 语义分析技术 YSIS

陈俊杰 李海芳 相洁 赵涓涓 著



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

<http://www.phei.com.cn>

图像情感语义分析技术

陈俊杰 李海芳 著
相 洁 赵涓涓

電子工業出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

本书系统地介绍了情感的认知基础及数学模型、从 fMRI 数据中解读情绪状态的方法、图像特征提取算法、神经网络和支持向量机、MPEG 中图像的描述标准、本体构建和推理方法，并在此基础上结合当前的研究动态，着重对基于强度的多层情感模型、图像特征和情感映射算法、图像特征本体的构建和推理的相关技术进行了讨论，力求向读者展示相关领域的最新研究动态，希望能为从事相关研究的广大读者提供参考，能对图像理解、人机交互技术的发展起到推动作用。

本书可作为信号处理、计算机等专业硕士研究生和博士研究生的专业教材，以及本科生的选修教材，也可作为相关领域科技工作者的参考书。

未经许可，不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有，侵权必究。

图书在版编目 (CIP) 数据

图像情感语义分析技术/陈俊杰等著. —北京: 电子工业出版社, 2011. 4
ISBN 978 - 7 - 121 - 06368 - 8

I. ①图… II. ①陈… III. ①图象 - 情感 - 语义分析 IV. ①TP391.41

中国版本图书馆 CIP 数据核字 (2011) 第 038823 号

策划编辑: 董亚峰

责任编辑: 周宏敏

印 刷: 北京东光印刷厂

装 订: 三河市皇庄路通装订厂

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 × 1 092 1/16 印张: 17.75 字数: 454 千字

印 次: 2011 年 4 月第 1 次印刷

定 价: 38.00 元

凡所购买电子工业出版社的图书, 如有缺损问题, 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 zltz@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线: (010) 88258888。

前 言

情感在人类的感知、决策等过程中扮演着重要角色，情感活动离不开周围的环境。现实世界的环境是由一帧帧图像组成的，图像中蕴涵着丰富的情感。如何利用计算机技术自动识别图像情感，从以机为本到以人为本，实现用户和多媒体信息之间的无缝连接，近年来受到许多学者的关注。图像情感语义分析是人工智能、情感计算、图像理解、模式识别等领域中一个重要的研究方向，从理论和实际应用的角度来看，国内外学者进行了广泛深入的研究，已经取得了不少成果，但尚未出现系统介绍该领域研究成果的图书。本书系统地介绍了情感的认知基础及数学模型、图像特征提取算法、神经网络和支持向量机、MPEG 中图像的描述标准、本体构建和推理方法，并在此基础上结合当前的研究动态，围绕相关技术，介绍了作者的研究内容及结果分析。主要内容如下：

通过理解脑科学有关情绪神经机制的研究进展，采用认知心理学实验手段，设计了情绪个性化认知实验。分析了现有的 fMRI 分类算法，针对目前 fMRI 数据分析工作难以揭示全脑协同工作模式的缺陷，介绍了 SVVC 数据分类方法，该方法可以作为一种通过 fMRI 数据识别情感的方法。

通过对影响情感个性化因素的研究，介绍了一种基于强度的多层情感模型。模型将性格、前一时刻情感衰减、情感间的相互影响、外界刺激和情感激活阈值有机地结合在一起，克服了 OCC 模型和基于维度的情感模型忽视非认知因素的缺点，弥补了 HMM 模型仅仅从概率上处理情感的缺陷，强化了个性化因素的影响。利用所设计情感预测系统和人对相应图像情感理解的比较表明，预测准确率达 75% 左右。

针对颜色、纹理特征提取问题，介绍了一种基于眼动跟踪的视觉注意图像特征提取算法。根据人眼观察图像的特点引入感兴趣区域概念，利用眼动仪获取感兴趣区域，并以其为中心对图像分块，采用不同区域特征向量赋予不同权重的策略，强化了感兴趣区域对情感的影响，降低了特征向量维数，提高了计算速度；采用自组织网络将已标注情感的样例图像特征进行了聚类，分析了情感分布情况，根据反馈信息调整权重，有效地优化了算法的性能。通过与颜色直方图特征提取算法比较，该算法聚类效果较好，有利于个性化图像情感理解。

针对图像低层特征与高层语义之间的复杂映射关系，总结了图像低层特征的特点，对机器学习算法进行了选择。分析了径向基概率神经网络算法具有计算复杂度低、收敛速度快的优点，同时针对此模型的不足，介绍了解决或部分解决神经网络缺陷的改进方法，根据服装情感的特点，将径向基概率神经网络算法和基于眼动跟踪的视觉注意图像特征提取算法相结合，介绍了服装图像情感分类系统，实现了服装图像的情感分类。

分析了现有的 SVM 多类分类方法，选择基于聚类的 SVM 多类分类策略，并且分析了该方法没有考虑样本的类别信息，当异类样本混杂时导致二叉树分支过多，降低分类性能的缺点。介绍了 LFCM - SVM 方法，该算法结合图像低层特征的特点，将 LDA 引入二叉树建树过程之中，在聚类之前先对样本进行投影，选择最优子空间使得同类样本聚集而异类样本松散，改进了原算法的分类性能，并通过 UCI 数据集上的实验验证了此方法的可行性。

将 LFCM - SVM 方法应用于图像语义映射, 得到较高的准确率。针对图像情感语义模糊性和不确定性的特点, 对 LFCM - SVM 进行调整, 将结果模糊化, 调整参数控制映射结果的精度。通过实验验证了该方法的合理性和可行性。

利用本体可以将图像情感领域的概念和术语规范化、刻画概念之间内在的联系以及本体所具有的语义特性, 将图像低层特征和高层语义衔接起来, 同时利用本体的推理能力, 将情感语义检索的范围和深度进一步扩大, 使基于本体的信息检索结果能更加符合用户的检索要求。介绍了一种基于 MPEG - 7 与模糊概念格的图像情感语义本体库构建方法, 利用该方法将图像情感和低层特征整合为一个整体, 建立了图像情感语义本体库。介绍了应用于图像情感本体背景下基于属性的模糊概念格生成维护算法, 较好地解决了由于图像所包含的情感的丰富性、不确定性所带来的“图像视觉特征”和“丰富情感语义”之间复杂的映射问题。

本书第 1 章 1.1 节 ~ 1.4 节由李海芳执笔; 第 1 章其余部分、第 2 章和第 5 章由陈俊杰执笔; 第 3 章、第 4 章由相洁执笔; 第 6 章、第 7 章由赵涓涓执笔; 全书由李海芳整理。在课题的研究中, 硕士贺静、焦立鹏、郭尚波、李晋、赵炜、成琳、刘嘉玲等参与了相关研究, 书中部分实验和结果由他们完成。博士研究生郭涛、郭浩, 硕士研究生王艳群、白瑀、李剑斌、郭凯、刘菲军、姜华杰完成了全书的校对, 电子工业出版社董亚峰编辑为本书的出版付出了辛勤的劳动, 在此我们表示深切的感谢! 本书所反映的研究成果得到国家自然科学基金资助项目 (项目编号: 60970059, 61070077) 和山西省自然科学基金资助项目 (项目编号: 2010011020 - 2) 的资助, 在此一并表示感谢!

由于作者水平有限, 书中错误和不妥在所难免, 恳请同行和广大读者批评指正。

目 录

第1章 绪论	1
1.1 图像语义提取技术与发展	1
1.1.1 图像语义提取技术	1
1.1.2 图像语义提取研究进展	3
1.2 情感计算的产生与发展	4
1.2.1 情感计算的产生及研究内容	4
1.2.2 情感计算的研究进展	6
1.2.3 情感计算的应用	8
1.3 图像特征和情感的关系	10
1.3.1 颜色与情感的对应	11
1.3.2 纹理、形状与情感的对应	14
1.3.3 对象及感兴趣区域与情感的对应	16
1.4 脑科学与情感计算	16
1.4.1 脑科学	16
1.4.2 脑科学的主要研究手段	17
1.4.3 脑科学研究与情感计算的关系	21
1.5 本体与图像特征的描述	22
1.5.1 MPEG 标准与本体	22
1.5.2 本体的构建	23
本章小结	24
参考文献	25
第2章 情感的认知基础	31
2.1 情绪、情感的关系	31
2.1.1 什么是情绪和情感	31
2.1.2 情绪和情感的区别和联系	32
2.2 情感的范畴表示	34
2.2.1 基本情绪	34
2.2.2 复合情绪	36
2.3 情感的维度表述	36

2.3.1	情感的维度表示	37
2.3.2	PAD 三维情感模型	39
2.4	情感的生理特征	40
2.4.1	面部表情与姿态	40
2.4.2	皮肤电反应	42
2.4.3	循环系统	43
2.4.4	内分泌系统	44
2.4.5	声音与音调	45
2.4.6	呼吸	45
2.4.7	脑电反应	46
2.5	情绪的内在神经机制	46
2.5.1	情绪的脑机制	46
2.5.2	情感的基本环路	49
2.6	从 fMRI 数据中解读情绪状态	50
2.6.1	fMRI 基本原理	50
2.6.2	fMRI 实验设计	51
2.6.3	fMRI 数据采集	54
2.6.4	fMRI 数据预处理	55
2.6.5	fMRI 数据脑功能定位分析	57
2.6.6	基于 fMRI 数据解读情绪状态	58
	本章小结	65
	参考文献	66
第3章	情感模型	69
3.1	基于认知的情感模型	69
3.1.1	OCC 模型	69
3.1.2	Roseman 模型	71
3.1.3	Scherer 模型	71
3.1.4	非认知的情感模型	72
3.2	基于概率的模型	73
3.2.1	隐马尔可夫模型	73
3.2.2	隐马尔可夫情感模型	75
3.2.3	状态转移概率矩阵的求取	77
3.3	多层情感模型	78
3.3.1	情感影响因素	79
3.3.2	情感预测系统	83
3.3.3	实验及结果分析	89
3.4	情感建模的应用	93
	本章小结	94

参考文献	94
第4章 图像特征提取算法	97
4.1 颜色特征提取算法	97
4.1.1 颜色空间	97
4.1.2 颜色直方图	102
4.1.3 颜色矩	105
4.1.4 颜色集	106
4.1.5 颜色相关图	106
4.2 纹理特征提取算法	106
4.2.1 纹理特征和纹理分析	106
4.2.2 基于统计的方法	107
4.2.3 基于结构的方法	112
4.2.4 基于频谱的方法	112
4.3 形状特征提取算法	117
4.3.1 边界描述	117
4.3.2 区域描述	120
4.3.3 几何形状描述	120
4.4 感兴趣区域特征提取算法	122
4.4.1 基于拐点的方法	123
4.4.2 基于灰度变换的方法	124
4.4.3 基于视觉注意的方法	125
4.4.4 基于眼动跟踪的方法	127
4.4.5 权重的确定	129
4.5 图像特征聚类及结果分析	135
4.5.1 实验数据	135
4.5.2 实验步骤及结果分析	136
本章小结	142
参考文献	142
第5章 图像语义提取算法	145
5.1 径向基概率神经网络算法	145
5.1.1 径向基函数神经网络	145
5.1.2 概率神经网络	147
5.1.3 径向基概率神经网络	149
5.1.4 径向基概率神经网络的应用	152
5.2 自组织映射算法	156
5.2.1 自组织映射网络	156
5.2.2 自组织映射算法	157
5.2.3 SOM 网络模型的应用	157

5.3 支持向量机	160
5.3.1 基本思想	161
5.3.2 SVM 多类分类方法	165
5.3.3 基于聚类的 SVM 多类分类方法	167
5.3.4 LFCM-SVM 多类分类方法	169
5.3.5 LFCM-SVM 图像语义映射	172
5.3.6 图像情感语义映射	178
5.4 有判定的极小极大规则学习算法	181
5.4.1 极小极大规则学习方法	182
5.4.2 有判定的极小极大规则学习算法	182
5.4.3 规则抽取及简化的实现	185
本章小结	194
参考文献	194
第6章 多媒体信息的描述标准 (MPEG 标准)	196
6.1 MPEG-1 标准与 MPEG-2 标准	196
6.1.1 MPEG-1 标准	196
6.1.2 MPEG-2 标准	198
6.2 MPEG-4 标准	201
6.2.1 MPEG-4 标准的目标	201
6.2.2 MPEG-4 标准的组成	201
6.2.3 MPEG-4 标准主要特征	202
6.3 MPEG-7 标准	206
6.3.1 MPEG-7 标准的目标	207
6.3.2 MPEG-7 标准的内容	208
6.3.3 MPEG-7 标准中的关键技术	212
6.4 MPEG-21 标准	214
6.4.1 MPEG-21 标准的目标	214
6.4.2 MPEG-21 标准框架的基本要素	214
6.5 MPEG-7 标准与图形描述	216
6.5.1 基于 MPEG-7 的颜色描述子	216
6.5.2 基于 MPEG-7 的纹理描述子	218
6.5.3 基于 MPEG-7 的其他描述子	219
6.5.4 基于 MPEG-7 的描述举例	222
本章小结	227
参考文献	227
第7章 本体的构建及推理	229
7.1 本体的定义及相关概念	229
7.1.1 本体的定义	230

7.1.2	本体的结构	231
7.1.3	本体的作用	231
7.1.4	已有本体的简介	232
7.2	本体的分类	233
7.2.1	根据应用主题的分类	233
7.2.2	根据形式化程度的本体分类	233
7.2.3	根据研究层次的本体分类	233
7.3	本体的构建	234
7.3.1	本体构建的相关知识	234
7.3.2	本体的构建工具	241
7.3.3	本体的构建方法	243
7.4	概念格理论与本体的结合	248
7.5	模糊概念格的本体约束	252
7.5.1	融合 MPEG-7 的模糊情感本体框架	252
7.5.2	模糊图像情感本体框架	259
7.5.3	模糊形式背景	261
7.5.4	模糊情感形式概念分析	262
7.6	本体推理	266
7.6.1	本体推理方法	266
7.6.2	模糊情感匹配规则算法	266
7.7	图像情感本体库构建实验及结果分析	267
7.7.1	实验数据	267
7.7.2	实验步骤及结果分析	267
	本章小结	271
	参考文献	271

第1章 绪 论

人，主宰世界的精英，情感能力是人类智能的重要标志。看，世界万物均是图像景象。“百闻不如一见”、“一目了然”，都反映了图像在传递信息中的独到之处，反映了图像自身蕴涵着丰富的情感。人类长期以来，都想设计制造一个像人一样的机器，以便能够代替人进行各种工作，在计算机技术发展日臻成熟的今天，计算机认识世界成为现实。由于人的情感具有主观性，从人类的认知机理研究图像情感语义分析技术是非常必要的。

针对图像情感语义分析技术中情感模型、图像特征提取、映射算法和图像特征描述等问题，提出新的方法，有效地建立情感和图像特征之间的对应关系，提高情感识别准确率，使计算机识别图像的能力接近人的理解水平，对于推动图像检索和图像理解具有一定的理论意义和学术价值。

1.1 图像语义提取技术与发展

1.1.1 图像语义提取技术

图像的语义，由下至上分为3个层次^[1]：特征层、对象层和语义概念层，每一层对应一个图像的语义级别（见图1.1）。其中特征层是指图像中特定的颜色、形状、纹理等视觉特征及其组合；对象层是指图像中出现的对象及对象间相互的空间关系；语义概念层是指人对图像内容的概念级反映。语义概念层由低到高又可分为：场景语义、行为语义和情感语义。其中场景语义指的是图像中的场景（如沙滩、天空等）；行为语义指的是图像中对象的行为及活动（如一场排球比赛）；情感语义指的是图像给人带来的主观感受（如高兴、生气等），它们结合在一起形成了一幅图像所要表达的意义。

大多数图像语义提取方法使用多层抽取机制^[2]：特征抽取、对象识别和基于领域知识的语义推理。

特征抽取主要包括颜色、纹理、形状等，且通常分为全局特征和局部特征两类。全局特征将整个图像视做一个单独的实体，主要优点是特征抽取和模式匹配算法的复杂度较低，而缺点是忽略了图像特征的空间位置关系，导致图像识别和检索准确率降低；

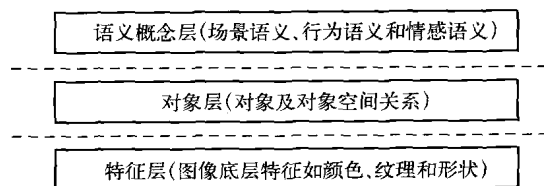


图 1.1 图像内容的层次模型



局部特征将图像分割成一系列区域，再对分割好的区域提取多种特征，每个区域表示一个潜在的用户感兴趣对象。局部特征可以被用来识别图像中的显著对象，抽取图像更多的细节信息，并且由对象识别层进行对象的识别，为提取更高层的语义创造条件。

对象识别过程通过将存储在知识库中的对象模型与抽取的图像特征相匹配来识别图像中的对象。如图 1.2 所示，先进行图像分割，然后提取区域的特征，再识别对象和对象的空间关系，这种处理过程比较直观，在一定程度上也符合人们观察事物的顺序，识别出的对象和空间关系可以成为获得更高层次语义的基础，作为获得图像场景语义和事件语义的辅助手段。在结合特定的领域知识后，这类方法可以在特定领域取得成功的应用。例如：男士西服的类别，可以根据西服是单排扣还是双排扣、扣子的数量、衬衫的颜色、领带的图案等来判断其样式是属于正式的、休闲的还是传统的。

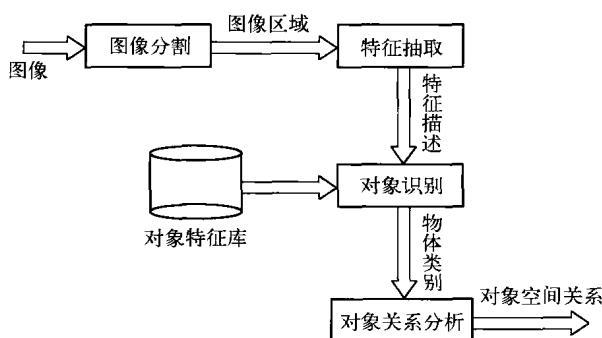


图 1.2 对象和对象的空间关系语义提取

基于领域知识的语义推理是在识别出对象的基础上，利用知识库中的规则和方法，将对象和对象之间的空间关系映射到不同层次和粒度的语义。其关键问题是图像语义知识库的建立以及如何利用知识库中的规则进行推理。根据特定的领域知识库，可以提取特定的人所关心的语义，这样可以解决某些面向特定用户的语义问题。

场景和行为语义提取是根据识别的对象、对象间的空间关系以及图像的背景，结合场景语义的知识来进行场景分类。在图像分割和对象识别技术的困难还不能有效解决的情况下，一些研究者采用了一些新的方法，这些方法的特点是绕过对象识别这一过程，将图像固定分割，然后针对图像子块分别确定其各自对应的语义类别，最后根据子块语义的组合关系来确定整幅图像的语义，如图 1.3 所示。由于每个子块相对于整个场景来说在视觉特征方面比较一致，因此在特定的领域较容易映射到相应的语义。一旦子块的语义确定了，场景语义就能有效地使用现有的统计方法从子块的相关模式获得，而特定场景对应的子块组合模式可以通过训练集来学习获得。但是，由于现实场景的多样性，目前这样的系统一般也只能对特征比较明显的少数场景有一定的效果，还不能达到通用场景识别的目的。而对于行为语义，通常无法通过单幅图像来自动提取，往往要考虑图像序列的运动信息，或者纯粹通过人工的方法来进行。

图像的情感语义相对于前面几种语义来说具有更多的主观成分，它和人的认知模型、文化背景及美学标准紧密关联，近年来受到许多国内外学者的关注。为了使计算机理解图像的能力达到人的理解水平，研究人的情感变迁过程和影响情感的关键因素，有效地建立人的情



感和图像特征之间的对应关系是非常必要的。

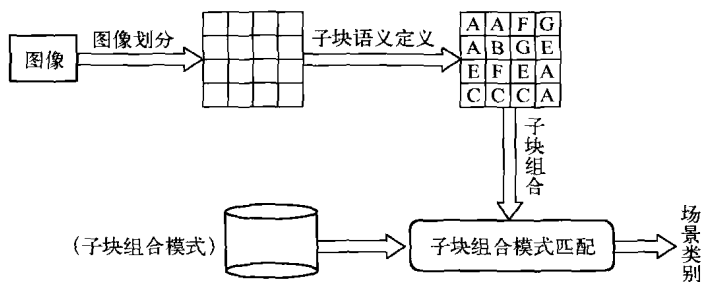


图 1.3 固定划分图像的场景语义提取

1.1.2 图像语义提取研究进展

随着多媒体技术及因特网 (Internet) 技术的发展, 可获取和使用的多媒体信息越来越多, 并逐渐改变着人们的生活和交流方式。图像和视频已成为多媒体的主要表现形式, 如何组织、表达、存储、管理、查询和检索这些大规模的图像数据库, 是对传统数据库技术的一个重大挑战。基于文本的图像注释存在两大难题^[1~3]: 一是对图像进行人工标注费时费力, 尤其是面对海量的图像库时, 人工标注工作量巨大; 二是不同人对同一幅图像有不同的理解, 文本标注存在主观性和不精确性, 直接影响到结果的准确性。基于内容的图像检索 (Content Based Image Retrieval, CBIR) 在国内外已经取得了不少成果, 技术上各种检索方法层出不穷, 学术上建立了专业期刊交流平台, 应用上也取得初步成效并得到社会的广泛关注和支 持。在因特网上已有许多关于基于内容的图像检索原型系统, 最著名的有 IBM 的 QBIC^[4,5] 系统和哥伦比亚大学开发的 Visual SEEK^[6,7] 等, 其应用领域涉及多媒体数据库、电子图书馆、商标管理、医疗图像管理、公安系统、卫星图像管理等方面的图像检索。然而, CBIR 主要利用图像本身的视觉特征来匹配图像的相似度, 难以描述图像所含的丰富语义, 对于用户的爱好、情感等主观或感性化的因素考虑较少, 没有对图像在激发情感等主观体验方面的特点给予足够的关注, 忽略了情感的影响和作用, 使计算机对图像内容的理解与人对图像内容的理解存在着巨大客观差距, 造成了“语义鸿沟”问题。

现在, 获得用户对图像主观情感的方法, 大部分采用测试者对图像进行情感调查的方法进行, 到目前为止, 还没有一个统一的模式。由于底层特征到高层语义特征的映射通常是非线性变换的, 难以建立模型, 因此让计算机准确理解图像情感语义是个难题。在图像语义获取问题上, 研究人员投入了大量精力, 其方法主要有: 机器学习、人工交互和利用外部信息源。

基于学习的语义提取方法大致可分为线性方法和非线性方法。王伟凝^[8] 等给出一个图像情感语义提取的结构框架, 指出图像中的颜色、纹理、形状和情感之间有很大的关系; 张颖^[9] 等对人类面部表情做了一定的研究; 王上飞^[10,11] 等利用图像中的颜色和形状来对情感进行标注, 并建立情感形容词对 (反义词对); Jinsub^[12] 采用自适应模糊系统建立彩色图例与情感评价间的关系, 并与神经网络方法和线性映射方法进行了比较, 具有更高的准确率; Hayashi^[13] 使用了反向传播规则的神经网络来建立图像特征与印象词间的联系, 达到了



78.8%的准确率；Chiu Chih - Yi^[14]等使用模糊聚类的方法进行纹理图像分类；高学^[15]等将支持向量机方法应用到汉字识别中取得了较好的效果；王雷^[16]等提出一种大型数据库系统中基于SVM的图像检索框架，它能够缩短反应时间和改善检索效率；王上飞、王照法^[17]等从心理学的“维量”思想出发，采用语义量化技术和因子分析的方法建立情感空间，并提出分析情感空间的相似性度量方法，通过抽取图像的颜色和形状等视觉特征，采用径向基函数将图像由特征空间映射到情感空间；余英林^[18]等根据线条与图像动感感觉之间的联系，采用了改进的边缘方向直方图作为特征，使用概率神经网络对艺术图像进行“动感”和“静感”的分类，达到了87.273%的准确率。

人工交互的语义提取方法包括对图像的预处理和反馈学习两个方面。早期的对图像库中的图像进行人工标注就是一种简单的图像预处理方式，这种方式费时费力，存在一定的缺陷。反馈学习是在提取语义的过程中加入人工干预，常见的有相关反馈方法。朱兴全^[19]等通过用户的相关反馈来调整每幅图像对应的关键词权值，即用户反馈的过程在语义上就是修改权值的过程。He^[20]等根据用户的反馈来构造语义矩阵。Liu^[21]提出一个半自动策略对图像语义进行标注，根据用户对查询结果的相关反馈来产生和修改语义注释。

基于外部信息源的语义提取方法是根据图像外部的相关信息，获取与图像语义相关的语义信息的。Shen^[22]提取网页中的文本内容作为该网页中图像的内容描述。Srihari^[23]将图像的标题作为图像的注解。张华^[24]等提出一种利用图像语义词典从网页中的图像外部文本信息中提取该网页中图像语义信息的方法。基于外部信息源的方法利用外部信息来获取图像语义，具有一定的优越性。但是当图像的外部信息不易得到，或者外部信息非常少的时候，采用这种方法获取图像语义就比较困难了。

1.2 情感计算的产生与发展

1.2.1 情感计算的产生及研究内容

情感是人类智能的一个不可分割的部分，它在人的感知、推理、决策、计划、创造及社交等诸多活动中起着不可或缺的作用^[25,26]。1995年美国MIT媒体实验室Picard教授提出了情感计算（Affective Computing）一词并给出了定义，即情感计算是关于情感、情感产生及影响情感方面的计算。其目的是赋予计算机识别、理解、表达和适应人情感的能力^[25]。情感计算研究的重点就是通过各种传感器获取由人的情感所引起的生理及行为特征信号，建立“情感模型”，从而创建感知、识别和理解人类情感，并能针对用户的情感做出智能、灵敏、友好反应的个人计算系统，缩短人机之间的距离，营造真正和谐的人机环境。

情感计算是未来人工智能理论与应用中不可缺少的重要组成部分^[27,28]。没有情感计算能力的智能机器是不能实现充分的人工智能的。然而，目前的计算机从设计原理上来讲主要是基于逻辑推理系统，这样的系统（包括所谓智能系统）大都忽视了情感能力的作用。例如，人与计算机之间的交互多少年来一直主要依靠键盘与鼠标来进行，这种交互方式的弊端是显而易见的。这种局面的形成主要是因为传统的计算机不仅“盲”（不具备视觉功能）、



“哑”（不具备语言功能），并且“聋”（不具备听觉功能），更谈不上具备“善解人意”的能力（即理解和适应人的情感或心情的能力）。没有这种情感能力，就很难指望计算机具有类似人一样的智能，也很难真正实现人机交互的和谐与自然^[29]。

根据情感计算的过程，可将情感计算的研究内容分为以下9个方面：情感机理、情感信号的获取、情感模式识别、情感的建模与理解、情感的合成与表达、情感的合成与表达、情感计算的应用、情感计算机接口、情感的传递与交流 and 可穿戴计算机^[30]，如图 1.4 所示。

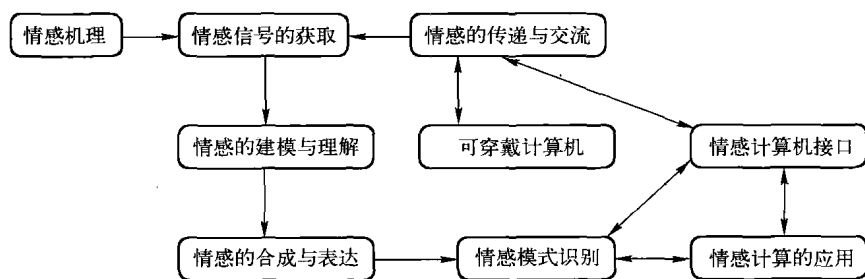


图 1.4 情感计算的研究内容

从人类情感的交流过程来讲，通过传感器直接或间接与人接触获得情感信息，通过建模对情感信息进行分析与识别，对分析结果进行推理达到感性的理解，将理解结果通过合理的方式表达出去也就完成了情感交流的全过程。根据上述过程，情感计算的研究内容主要应包括：情感机理的研究、情感信号的获取、情感信息的分析建模与识别、情感信息的理解和情感的表达。下面详细介绍各项研究内容。

(1) 情感机理的研究

情感机理的研究主要是情感状态判定及与生理和行为之间的关系。涉及心理学、生理学、认知科学等方面，为情感计算提供理论基础。任何一种情感状态都可能伴随几种生理或行为特征的变化；而某些生理或行为特征也可能起因于数种情感状态。因此，确定情感状态与生理或行为特征之间的对应关系是情感计算理论的一个基本前提，这些对应关系目前还不十分明确，需要进一步的探索和研究。

(2) 情感信号的获取

情感信号的获取研究主要是指各类有效传感器的研制，它是情感计算中极为重要的环节，没有有效的传感器，可以说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的。各类传感器应具有如下的基本特征：使用过程中不应影响用户（如质量、体积、耐压性等）；应该经过医学检验对用户无伤害；具有数据的隐私性、安全性和可靠性等特点。MIT 媒体实验室的传感器研制走在了前列，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。

(3) 情感信号的分析、建模与识别

一旦由各类有效传感器获得了情感信号，下一步的任务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，所以建立一个能够有效地在计算机上表示和处理情感的计算模型是非常关键的。目前已有一些模型，如隐马尔可夫模型、OCC 模型、贝叶斯网络模型等。情感建模是情感计算的核心研究内容之一。



(4) 情感信息的理解

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。情感计算的最终目的是使计算机在了解用户情感状态的基础上,做出适当反应,去适应用户情感的不断变化。因此,情感理解主要研究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。

(5) 情感的表达

前面的研究是从生理或行为特征来推断情感状态。情感的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生理或行为特征中体现出来。例如,如何在语音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。

(6) 可穿戴计算机

可穿戴计算机是情感计算实现的最佳平台,主要研究可穿戴计算机的传感器输入、输出装置。可穿戴计算机的研究体现在软、硬件两个方面,难点在于结构的设计。可穿戴计算机的发展必将推动情感计算的研究。同样,没有情感计算理论的武装,可穿戴计算机很难最大限度地发挥其潜能和优势。

1.2.2 情感计算的研究进展

很久之前,人们就注意到了情感对认知和智能的作用,即情感与认知、情感与智能之间的关系。1981年就有人研究人工情感问题^[31],但只是在1990年以后才开始逐渐引起人们的注意,特别是近几年来对人工情感的研究越来越受到人们的重视。

1. 国外研究进展

美国 MIT 媒体实验室的 R. Picard 教授于 1995 年提出情感计算 (Affective Computing) 的概念,并于 1997 年正式出版专著 *Affective Computing* (《情感计算》)。在该书中她定义“情感计算是与情感相关、来源于情感或能够对情感施加影响的计算”。此概念一经提出,就引起了很大反响, Picard 领导的情感计算课题组以对人类情感的生理信号处理为基本出发点,对情感计算进行了全方位的研究,研究工作取得了很大进展,其应用领域也日益扩大,在 MIT 媒体实验室的技术报告中已经涉及到约 50 种应用。

日本从 20 世纪 90 年代就开始了感性工学的研究。按照他们的定义,所谓感性工学,就是将感性(情感)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(情感),把人的感性(情感)需要加入到商品设计、制造中去,它是一门从工程学的角度实现能给人类带来喜悦和满足的商品制造的技术科学。感性工学由于可以给人们的生活带来快乐和舒适,因而也被称为“快乐而舒适”的科学,基于感性工学技术生产的商品称为“感性商品”。日本已经形成了举国研究感性工学的高潮^[32],1996 年日本文部省就以国家重点基金的方式开始支持“情感信息的信息学、心理学研究”的重大研究课题^[33],参加该项目的有十几个大学和研究单位,主要目的是把情感信息的研究从心理学角度过渡到心理学、信息科学等相关学科的交叉融合。每年都有日本感性工学全国大会召开。与此同时,一向注重经济利益的日本,在感性工学产业化方面取得了很大成功,日本各大公司竞相开发、研



究、生产了所谓的个人机器人产品系列。其中，以 SONY 公司的 AIBO 机器狗（已经生产 6 万只，获益近 10 亿美元）和 QRIO 型及 SDR-4X 型情感机器人作为典型代表。应该说，日本在人工情感技术的应用方面是领先全世界的。

欧盟国家也在积极地对情感信息处理技术（表情识别、情感信息测量、可穿戴计算等）进行研究。欧洲许多大学成立了情感与智能关系的研究小组，其中比较著名的有：日内瓦大学 Klaus Scherer 领导的情感研究实验室，布鲁塞尔自由大学的 D. Canamero 领导的情感机器人研究小组及英国伯明翰大学的 A. Sloman 领导的 Cognition and Affect Project。A. Sloman 教授提出了情感三层体系结构（CogAff Agent Architecture，认知情感体系），探讨了情感与认知的相互作用。在市场应用方面，德国 Mehrdad Jalali - Soli 等人在 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^[34]。

IBM 公司的“蓝眼计划”，可使计算机知道人想干什么，如当人的眼瞄向电视时，它就知道人想打开电视机，它便发出指令打开电视机。此外该公司还研究了情感鼠标，可根据手部的血压及温度等传感器感知用户的情感。

2. 国内研究进展

我国对人工情感的研究始于上世纪 90 年代，大部分研究工作是针对人工情感单元理论与技术实现的。哈尔滨工业大学研究多功能感知机^[35]，主要包括表情识别、人脸识别、人脸检测与跟踪、手语识别、手语合成、表情合成、唇读等内容，并与海尔公司合作研究服务机器人。清华大学进行基于人工情感的机器人控制体系结构的研究^[36]。北京交通大学进行多功能感知机同情感计算的融合研究^[37]。中国科学院自动化研究所主要研究基于生物特征的身份验证^[38]。中国科技大学开展了基于内容的交互式感性图像检索的研究^[39]。中国科学院软件所主要研究智能用户界面^[40]。南京航空航天大学研究语音情感计算。重庆大学主要研究智能服务、增强现实、环境感知、智能手表等，注重软件方面的研究。浙江大学研究 E - Teatrix 中虚拟人物及情感系统构造^[41]。不仅仅是人工智能领域的专家学者关心人工情感与人工心理的研究问题，更值得注意的是，情感心理学家对于“情感智力与人工智能中的感情计算”也进行了很深的思考^[42]。他们认为，基于情感的智力是加工、处理情感及情感信息的能力，而人工智能中的感情计算是要赋予计算机与人互动过程中情感信息的加工能力，人脑处理情感信息的能力与计算机处理情感信息的能力可以进行类比。近几年来，人工智能专家已经认识到情感智力在感情计算中的重要作用和意义，把人类识别和表达情感的能力赋予计算机，开发了具有部分情感能力的计算机。新一代情感计算机的研发和应用依赖于人工智能专家与心理学家之间的密切合作，两者的研究成果可以相互借鉴和互补。

情感建模是情感计算和人工心理研究的主要内容。情感的数学模型被认为是实现和谐人机交互的关键组成部分，其本质是对自然情绪实质的理解和表示。

1988 年 Ortony、Clore 和 Collins 等人从情感产生的认知角度，根据目标、事件和动作三个方面的评价标准将情感分为 22 种，提出 OCC 情感模型^[43]。该模型是第一个易于在计算机上实现的情感模型，很快得到计算机科学家和心理学家们的认同。1994 年，Elliott 进一步对 OCC 模型进行了扩展，将情感分类增加到 26 种^[44]。