

个性化视频情感内容分析:综述

张立刚 张九龙

(西安理工大学计算机科学与工程学院 西安 710048)

摘要 个性化视频情感内容分析是近几年兴起的一个研究热点,其目的是基于视频用户的个人兴趣和爱好来向其推荐个性化的视频内容。然而,当前仍然缺乏对该研究方向的最新研究进展的综述和讨论。为解决此问题,从 3 个方面对该领域的研究成果进行论述,讨论了各类现有方法的优缺点、存在的问题和挑战,并对未来的研究方向提出一些思路和建议。

关键词 视频情感内容分析,个性化推荐,用户兴趣,综述

中图分类号 TP391.44 文献标识码 A DOI 10.11896/j.issn.1002-137X.2018.01.003

Personalized Affective Video Content Analysis: A Review

ZHANG Li-gang ZHANG Jiu-long

(School of Computer Science and Engineering, Xi'an University of Technology, Xi'an 710048, China)

Abstract Personalized affective video content analysis is an emerging research field which aims to provide personalized video recommendation to an individual viewer tailored to his/her personal preferences or interests. However, there still lacks a review about recent progress on the development of approaches in this field. This paper presented a review of state-of-the-art approaches towards building automatic systems for personalized affective video content analysis from three perspectives of audio-visual features in video content (e. g. light, color, and motion), physiological response signals from viewers (e. g. facial expression, body gesture, and pose), and personalized recommendation techniques. It discussed the advantages and disadvantages of existing approaches, and highlighted several challenges and issues that may need to be overcome in future work.

Keywords Affective video content analysis, Personalized recommendation, Viewer interest, Review

1 引言

信息技术,特别是互联网宽带和流媒体技术的快速发展极大地促进了视频创建、发布、传播及视频用户的增长,视频浏览已成为人们生活 and 娱乐的重要组成部分。以 YouTube 为例,2016 年每月平均有超过 60 亿小时的视频被观看,每分钟有 100 小时的视频被发布,类似情况也存在于国内网络视频平台,如搜狐视频、优酷、土豆等。

面对快速增长的视频数据和不断扩大的视频用户群体,传统的针对普通用户的视频分析方法越来越不能满足个体用户复杂多变的实时兴趣、爱好和需求。因此,研究以个体用户兴趣和需求为中心的个性化视频分析和推荐技术,已成为未来计算机发展的必然趋势。

当前对个性化视频情感内容分析的研究还处于起步阶段。文献[1]从基于视频内容和基于用户反应两方面来总结当前视频情感内容的研究方法,而文献[2]则从心理学、社会学、神经科学以及计算机科学等多个相关研究领域中的观点

和模型来阐述当前的研究进展。但是,当前仍缺少专门从“个性化”的角度来概括和总结当前研究方法的论文。相对于传统的分析方法,个性化分析能更准确地反映不同个体在视频内容上的不同情感感知和兴趣,从而提供基于个人兴趣或需求的人性化视频服务。基于此,本文综述了当前个性化视频情感内容分析的最新研究成果,并对未来的研究提出了一些思路 and 方向。

2 个性化视频情感内容分析简介

关于情感的研究可追溯到 19 世纪,达尔文^[3]在《人与动物的情感》一书中率先提出情感是人类生存和进化的结果,并暗示某些情感的普遍性;随后 Ekman^[4]证实共存在不同文化和种族的 6 种基本情感。随着人工智能的发展,Picard^[5]于 1997 年正式提出“情感计算”,其主要目的是赋予计算机感知、识别和理解人类情感的能力。但直到 2006 年, Hanjalic^[6]才首次对视频中的情感内容进行分析。随着个性化视频服务需求的日益增长,研究者^[7-8]提出个性化视频情感内容分析方

收到日期:2016-12-28 返修日期:2017-03-21 本文受国家自然科学基金(61402362),陕西省教育厅基金(16JK1553),西安市碑林区科技局项目(GX1616)资助。

张立刚(1982—),男,博士后,副教授,主要研究方向为图像处理、表情识别和情感语义分析,E-mail:lg.zhang@xaut.edu.cn;张九龙(1974—),男,博士后,副教授,主要研究方向为图像处理、模式识别,E-mail:chinajiu-long@hotmail.com(通信作者)。

法,旨在实现能满足个体用户情感需求的动态视频输出,但当前针对该方法的研究仍处于起步阶段。

个性化视频情感内容分析^[7-8]旨在找出能引起观看者情感反应的视频内容,并动态更新系统输出来满足个体用户的实时情感需求。该分析方法不仅是解决当前视频数据增长、实现个性化视频服务的有效途径,也是下一代情感智能计算机的发展方向。情感是最高层的视频语义特征,直接影响观看者的感知、决策及对视频服务的满意度。在很多现实情况中,人们喜欢观看能够引起他们情感共鸣或相似情感体验的视频内容。情感也是用户表达个人兴趣和需求的重要渠道,人们常本能地通过行为特征(表情等)表达对事物的爱好和满意度。Pantic^[9]指出情感是构建下一代以人为中心的人类计算的核心因素,能帮助计算机智能感知用户的情感状态,并进行自然、友好、高效、可信的反馈和交互。

3 当前研究综述

如图 1 所示,本文从视频特征、用户特征及个性化 3 个方面对国内外视频情感内容分析的研究现状进行综述。

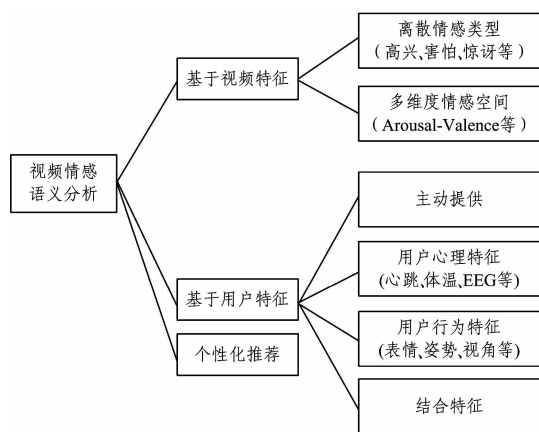


图 1 基于视频情感内容分析技术的分类

Fig. 1 Classification of video based on affective content analysis techniques

3.1 基于视频特征的情感内容分析的研究现状

基于视频特征的情感内容分析通过视频内部的听、视觉和文本等特征反映视频的情感内容。根据不同的情感表达方式,当前研究大致可分为离散情感类型和多维度情感空间两类^[10]。

3.1.1 离散情感类型

离散情感类型方法基于相关领域(美学、艺术学等)的先验知识,提取视频创作过程中产生的视觉(颜色、亮度、纹理、形状、运动等)、音频、文本等特征,并利用机器学习算法把视频划分到预选取的一组离散情感类型中,如 6 种基本情感(生气、高兴、沮丧、害怕、悲伤和惊讶)。代表性研究包括:Iric 等^[11]提取音高、能量、MFCC、颜色、亮度、运动强度、镜头长度等 7 种特征,并采用 LTDM 模型将电影视频分为 8 类情感;Jana^[12]提取颜色、纹理、和谐度、人脸等 4 类特征辨别图片的情感类型;Canini 等^[13]结合视觉特征(MPEG-7 颜色特征、颜色能量、亮度、色饱和度、运动)、听觉特征(音频强度、音色、节奏)和影片语法特征(视频长度、光谱构成、视频类型变化率)

来分析电影视频的情感内容。类似地,中国科学院 Wang 等^[14]通过提取听觉、视觉和颜色情感特征识别恐怖镜头。然而,当前研究对不同听视觉特征的类型选取未给出统一标准。

3.1.2 多维度情感空间

多维度情感空间方法运用心理学中的多维度空间表达情感信息,它利用相关领域(心理学、电影创作学、美学、艺术学等)的先验知识或训练机器学习算法,把视频特征映射到多维度空间中的点或区域,从而实现视频情感内容分析或状态分类。该方法最早由 Hanjalic 提出^[6],他通过提取了两组不同的听视觉特征分别线性模拟 Arousal-Valence(A-V)空间中的 A,V 维度值,实现对视频在情感层次的内容特征的分析。Zhang 等^[15]也提取两组听视觉特征,但采用 Affinity Propagation 算法分别非线性模拟 A,V 维度值。近期较有代表性的研究有: Xu 等^[16]首先应用与 A 维度关联的听视觉特征和 C-mean 聚类方法识别视频中情感内容的强度,然后再利用与 V 维度关联的听视觉特征和 CRF 方法辨别视频的情感类型; Niu 等^[17]基于两组分别对应于 A,V 维度的听视觉特征来计算视频间的相似度。该方法的应用受到工业界的高度重视,如英国国家广播公司(BBC)采用该方法开发的 Mood GUI 系统^[18],利用视频在“快、慢”和“严肃、幽默”二维空间中的情感内容帮助用户浏览电视节目。Baveye 等^[19]将 Convolutional Neural Networks (CNNs)用于预测 30 个视频的连续 A 和 V 维度值。实验结果发现, CNNs 的参数选取和学习需要较大的视频量,且 CNNs 所提取的视频特征和 Support Vector Machines for Regression (SVR)算法相结合的方法获得了最高的预测 Pearson 关联系数。然而,当前基于多维度空间的分析方法仅分别提取与每个维度关联的特征,忽视了多维度之间在时间和空间上的关联。

综上所述,基于视频特征的情感内容分析侧重于对可能影响观看者情感的视频内容的分析。由于缺少低层视频特征到高层用户情感感知的精确映射,且用户对视频内容的情感感知具有很强的主观性和实时性,因此该方法不能反映个体用户的动态情感变化,不能满足个性化视频情感内容分析的要求。

3.2 基于用户特征的情感内容分析的研究现状

针对视频低层特征应用于高层用户情感感知分析时存在的不足,国内外学者开始将目光转向基于用户特征的情感内容分析。该方法通过分析终端用户观看视频过程中的心理和行为反应信号得出用户的情感状态、兴趣、喜好、个性等信息,用于视频语义分析,如确定视频中哪些片段能使用户产生不同情感感知(高兴、悲伤等)。根据不同的用户特征获取方式,当前研究可分为以下 4 类。

3.2.1 情感信息的主动提供

该方法要求观看者主动提供个体真实的情感感知给情感内容分析系统。这是创建个性化情感语义分析的理想途径,但需要用户投入大量的时间和精力,且不能实现系统输出的自动实时更新。Zhang 等^[20]用 SVR 算法建立视频特征到 A,V 维度的映射,并利用个体观看者主动反馈的 A,V 值实现个性化视频检索。Canini 等^[13]将视频中的听视觉特征映射到一个三维内涵空间,并根据该空间中视频间的距离进行视频

推荐。该系统要求用户预先提供检索词更新推荐内容,因此,其需要用户主动提供信息以对视频情感内容进行分析,且未反映用户情感的动态变化。

3.2.2 用户心理特征

该方法通过监测和分析用户观看视频时的心理反应信号(包括脑电波(EEG)^[21]、体温、血压、心率、呼吸等),实现视频情感语义分析。该方法无需用户主动提供情感感知信息,但需要用户佩戴特定仪器,且对心理信号分析要求较高,即需要医学、心理学等领域的相关专业知识把心理信号转化为情感语义。Arthur等^[22]利用观看者的皮肤电反应、呼吸幅度、呼吸次数、血容量脉冲及心率信号实现视频摘要,而Katti等^[23]则利用瞳孔膨胀度和眼睛注意力来实现。Yazdani等^[24]发现EEG比神经末梢信号能更有效地区分视频情感语义(积极/消极、高/低、喜欢/不喜欢)。需要指出的是,心理特征已被广泛应用于人的情感分析^[25]。Mehmood等^[26]结合用户的神经元信号和视频中的听视觉特征来反映视频的情感内容,从而实现视频内容摘要的提取。

3.2.3 用户行为特征

该方法通过监控用户观看视频时的行为反应,包括表情、姿势、眨眼、注视及控制动作^[17]等,分析视频情感内容。该方法无需用户主动提供信息且无需佩戴特定仪器,但要求对用户的行为反应进行精确定位、跟踪、特征提取和情感分类。Soleymani等^[27]概括了当前情感语义标记中采用的用户行为和心理特征。在用户行为特征中,表情是情感最主要的表达方式^[28],具有直接、自然、接近人类感知、无需用户佩戴仪器、易操作等优点,因此目前对该方法的研究主要集中于用户表情。

Joho等^[10,29]是最早将表情用于视频情感内容分析的研究者之一,他们利用观看者的表情状态实现视频摘要和重要片段的提取。Jiao等^[30]成功地将用户表情状态用于验证图片标记结果的正确性。Tkalcic等^[31]发现基于表情的图片推荐虽然达不到手动标注结果的精确率,但却优于低层特征。Peng等^[32]通过检测眼睛运动、眨眼、头部移动特征获取用户的注意力,并结合表情状态预测该用户对视频内容的兴趣。在国内,Zhao等^[33]利用基于Haar的时间和空间特征识别观看者的表情状态,从而进行电影视频分类和视频推荐。Wang等^[34]采用Bayesian网络模拟多类表情之间以及表情与心理情感状态之间的关系,从而进行视频情感标记。以上研究围绕普通视频用户展开,但未考虑用户情感感知的动态变化和个体差异。

3.2.4 结合特征

该方法结合不同类型的用户特征或(和)视频特征来克服单一信息源的缺点,用于用户情感感知识别和视频情感内容分析。Arapakis等^[35]结合用户表情状态和神经心理信号预测话题与用户的关联度。Soleymani等^[36-37]结合EEG、瞳孔反应和注视距离标记视频情感内容,并发现心理信号和视频听视觉特征在反映视频情感内容上具有很强的相关性和一致性。在国内,中国科学技术大学Wang等^[38]结合用户的EEG和视频听视觉特征标记视频情感内容,并通过实验证实结合两类特征可获得更高的精确率。

综上所述,基于心理特征的情感内容分析通过用户主动提供或佩戴仪器获取情感感知信息,这种信息获取方式限制了该方法的应用范围。基于用户行为特征,特别是表情的分析方法通过监控用户的行为反应克服了该缺点,是实现不干扰用户的情感内容分析的有效方法,但该方法对用户行为特征的精确定位、跟踪等技术要求较高。当前对该方法的研究很少考虑用户情感的动态变化和个体差异。同时,研究表明用户特征和视频特征的结合有助于提高情感内容分析的精确率。

表1对当前研究文献中所使用的特征分类进行了简要的概述。

表1 基于视频情感内容分析的特征分类

Table 1 Feature classification of video based on affective content analysis

| 特征分类 | 子类型 | 特征内容 |
|------|-----|-------------------------|
| 视频特征 | 视觉 | 颜色、亮度、纹理、能量、形状、色饱和度和运动等 |
| | 听觉 | 音频强度、音色、节奏等 |
| | 语法 | 镜头长度、光谱构成、视频类型变化率等 |
| 用户特征 | 心理 | 脑电波(EEG)、体温、血压、心率、呼吸等 |
| | 行为 | 表情、姿势、眨眼、注视、控制动作等 |

3.3 个性化情感内容分析的研究现状

针对用户情感感知的个体差异,研究者近几年提出了个性化情感内容分析方法^[7-8],旨在找出能引起用户情感反应的视频内容,并通过动态更新系统输出来满足个体用户的实时情感需求,解决用户情感感知的动态变化和个体差异问题。

当前仅有少数文献对个性化情感内容分析方法进行研究。Arapakis等^[39]用表情反映个体用户对视频搜索结果的满意度,并据此判断视频与该用户的关联度。Arapakis等^[7]通过将其与普通预测模型进行对比,证实了基于用户表情的个性化视频情感内容分析能提高话题关联度的预测精度,但该研究未考虑用户情感的实时变化。Zhang等^[8]利用个体用户反馈的情感信息实现个性化视频检索,但该系统需要用户主动提供感知信息。Benini等^[40]建立一个三维内涵空间实现个性化视频情感分类,但该空间仍基于低层视频特征,因此不能直接反映用户情感感知的个体差异。Niu等^[17]把视频听视觉特征映射到A-V空间,并在该空间中进行视频归类,然后通过收集用户在浏览视频中的行为信息(如视频浏览时间)实现个性化视频推荐。由于用户的浏览行为信息与情感感知之间存在复杂的关系,因此该系统在表达用户情感感知的直观性上仍有待改进。

综上所述,当前对个性化视频情感内容分析的研究仍存在以下3方面的不足。1)未考虑用户情感感知的动态变化。当前研究对视频情感内容的分析和用户情感的获取是分开进行的,即预先获取用户观看所有视频的表情状态,然后再分析视频的情感内容。因此,其分析结果不能满足实时变化的用户情感需求。2)未考虑用户特征与视频特征的结合。当前研究仅从用户角度分析,未考虑视频内容特征对用户感知的影响。由于视频特征是引起用户情感感知的直接原因,不同的视频类型和内容对用户感知可能有很大影响。文献^[35]已证实用户特征和视频特征的结合有助于提高情感内容分析的精确率。3)未考虑用户情感感知的个体差异。大多数研究仅以普通用户为研究对象,而未考虑用户在情感上的个体差异。

因此,其分析结果不能反映用户个性化的情感感知。

4 当前挑战和未来研究方向

基于以上文献的综述,当前对个性化视频情感内容分析技术的主要挑战和未来研究方向如下。

1) 基于视频特征的情感内容分析方法尚缺少公认的高效的视频听视觉类型。由于缺少低层视频特征到高层用户情感感知的精确映射,该方法只能反映视频的情感内容,并不能有效地反映高层个人用户的情感和兴趣。因此,对不同类型的听视觉特征进行对比和分析,有助于找到更有效的特征组合。另外,将视频特征和用户行为特征相结合的方法将是未来的研究热点之一。

2) 基于用户行为特征的方法所面临的主要挑战是缺乏检测、跟踪和识别个人用户行为特征的高效且健壮的方法。虽然当前对用户表情、姿势、注视等行为的研究很多,但学术界仍缺乏能应用于现实环境的用户行为特征识别方法。因此,对该方法的研究也是个性化视频情感内容分析技术未来的研究方向之一。

3) 用户心理特征在反映用户情感和兴趣上具有较高的精确率,但需要用户佩戴特别的仪器设备来收集特征信号,这在很大程度上限制了该方法的广泛应用。因此,研究和开发更适合用户日常佩戴的仪器也是未来的一个研究重点。

4) 近些年,一些基于用户对视频内容的反应特征的数据库被创建,包括 MAHNOB-HCI^[41], DEAP^[42] 和 LIRIS-ACCEDE^[43]等。然而仅有极少数数据库包含同时对视频内容和用户反应特征的情感标注,且所包含的视频数量相对较少,如 DECAF^[44] 仅有 36 个视频。当前该领域仍然缺少公认的同时具有视频情感内容标注和相应的视频参看者情感状态标注的大型公开数据库用于测试和对比各种不同方法的识别性能。

5) 人体用户的情感和兴趣本身就是一个研究学科,个体之间会存在很大的差异,相关的心理学等学科^[2,45] 对该领域的研究成果将有助于个性化视频情感内容分析技术的发展,特别是对人的个性特征(如年龄、性别、职业、专业、地理位置等)和兴趣之间联系的研究。

6) 基于近几年深度学习在计算机视觉领域所获取的高性能,如何将深度学习技术有效地应用于情感内容分析也是未来的研究课题。目前,已有文献^[46-47] 进行了尝试,且取得了具有竞争力的结果。然而,深度学习一般要求较大的训练数据集,涉及大量的参数学习和较大的计算量,如何解决这些问题都是未来的挑战。

结束语 针对目前尚缺少对个性化视频情感内容分析进展的综述的问题,从视频内容、用户特征以及个性化推荐 3 个方面对当前的研究方法和成果进行概述和讨论,为本领域未来的研究提供一些研究思路和建议,更进一步促进本领域的发展。

参考文献

[1] WANG S,JI Q. Video Affective Content Analysis: A Survey of State-of-the-Art Methods[J]. IEEE Transactions on Affective

Computing,2015,6:410-430.

- [2] BAVEYE Y,CHAMARET C,DELLANDREA E,et al. Affective Video Content Analysis: A Multidisciplinary Insight[J]. IEEE Transactions on Affective Computing,2017,PP(99):1.
- [3] DARWIN C. The expression of the emotions in man and animals[M]. Oxford England:Philosophical Library,1955.
- [4] EKMAN P. Strong Evidence for Universals in Facial Expressions- A Reply to Russells Mistaken Critique[J]. Psychological Bulletin,1994,115:268-287.
- [5] PICARD R W. Affective Computing[M]. Cambridge: The MIT Press,1997.
- [6] HANJALIC A,LIQUN X. Affective video content representation and modeling[J]. IEEE Transactions on Multimedia,2005,7:143-154.
- [7] ARAPAKIS I,ATHANASAKOS K,JOSE J M. A comparison of general vs personalised affective models for the prediction of topical relevance[C]//The 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval. 2010:371-378.
- [8] ZHANG S,HUANG Q,TIAN Q,et al. Personalized MTV Affective Analysis Using User Profile[C]//Advances in Multimedia Information Processing-PCM 2008:327-337.
- [9] PANTIC M,PENTLAND A,NIJHOLT A,et al. Human Computing and Machine Understanding of Human Behavior: A Survey[C]//Artificial Intelligence for Human Computing. 2007:47-71.
- [10] JOHO H,STAIANO J,SEBE N,et al. Looking at the viewer: analysing facial activity to detect personal highlights of multimedia contents[J]. Multimedia Tools and Applications,2011,51:505-523.
- [11] IRIE G,SATOU T,KOJIMA A,et al. Affective Audio-Visual Words and Latent Topic Driving Model for Realizing Movie Affective Scene Classification[J]. IEEE Transactions on Multimedia,2010,12:523-535.
- [12] JANA M,ALLAN H. Affective image classification using features inspired by psychology and art theory[C]//Proceedings of the International Conference on Multimedia. 2010:83-92.
- [13] CANINI L,BENINI S,LEONARDI R. Affective Recommendation of Movies Based on Selected Connotative Features[J]. IEEE Transactions on Circuits and Systems for Video Technology,2013,23:636-647.
- [14] JIANCHAO W,BING L,WEIMING H,et al. Horror movie scene recognition based on emotional perception [C] // 17th IEEE International Conference on Image Processing (ICIP). 2010:1489-1492.
- [15] SHILIANG Z,QI T,SHUQIANG J,et al. Affective MTV analysis based on arousal and valence features[C]//IEEE International Conference on Multimedia and Expo. 2008:1369-1372.
- [16] XU M,XU C,HE X,et al. Hierarchical affective content analysis in arousal and valence dimensions[J]. Signal Processing,2013,93:2140-2150.
- [17] NIU J,ZHAO X,ZHU L,et al. Affivir: An affect-based Internet video recommendation system[J]. Neurocomputing,2013,120:422-433.

- [18] EGGINK J. Evaluation of a mood-based graphical user interface for accessing TV archives[C]//IEEE International Conference on Multimedia and Expo Workshops (ICMEW). 2013:1-6.
- [19] BAVEYE Y, DELANDRÉA E, CHAMARET C, et al. Deep learning vs. kernel methods: Performance for emotion prediction in videos[C]//International Conference on Affective Computing and Intelligent Interaction (ACII). 2015:77-83.
- [20] SHILIANG Z, QINGMING H, SHUQIANG J, et al. Affective Visualization and Retrieval for Music Video[J]. IEEE Transactions on Multimedia, 2010, 12: 510-522.
- [21] SOLEYMANI M, PANTIC M. Multimedia implicit tagging using EEG signals[C]//IEEE International Conference on Multimedia and Expo (ICME). 2013:1-6.
- [22] MONEY A G, AGIUS H. Analysing user physiological responses for affective video summarisation[J]. Displays, 2009, 30: 59-70.
- [23] KATTI H, YADATI K, KANKANHALLI M, et al. Affective Video Summarization and Story Board Generation Using Pupillary Dilation and Eye Gaze[C]//IEEE International Symposium on Multimedia (ISM). 2011: 319-326.
- [24] YAZDANI A, LEE J S, VESIN J M, et al. Affect recognition based on physiological changes during the watching of music videos[J]. ACM Trans. Interact. Intell. Syst., 2012, 2: 1-26.
- [25] NASOZ F, ALVAEREZ K, LISETTI C, et al. Emotion recognition from physiological signals using wireless sensors for presence technologies[J]. Cognition, Technology & Work, 2004, 6(1): 4-14.
- [26] MEHMOOD I, SAJJAD M, RHO S, et al. Divide-and-conquer based summarization framework for extracting affective video content[J]. Neurocomputing, 2016, 174: 393-403.
- [27] SOLEYMANI M, PANTIC M. Human-centered implicit tagging: Overview and perspectives[C]//IEEE International Conference on Systems, Man, and Cybernetics (SMC). 2012: 3304-3309.
- [28] MEHRABIAN A. Communication without words[J]. Psychology Today, 1968, 2(9): 52-55.
- [29] JOHO H, JOSE J M, VALENTI R, et al. Exploiting facial expressions for affective video summarisation[C]//Proceedings of the ACM International Conference on Image and Video Retrieval. 2009: 1-8.
- [30] JIAO J, PANTIC M. Implicit image tagging via facial information[C]//Proceedings of the 2nd International Workshop on Social Signal Processing. 2010: 59-64.
- [31] TKALČIČ M, ODIĆA, KOŠIR A, et al. Impact of Implicit and Explicit Affective Labeling on a Recommender System's Performance[C]//International Conference on Advances in User Modeling. Springer-Verlag, 2012: 342-354.
- [32] PENG W T, CHU W T, CHANG C H, et al. Editing by Viewing: Automatic Home Video Summarization by Viewing Behavior Analysis[J]. IEEE Transactions on Multimedia, 2011, 13(3): 539-550.
- [33] ZHAO S, YAO H, SUN X. Video classification and recommendation based on affective analysis of viewers[J]. Neurocomputing, 2013, 119: 101-110.
- [34] WANG S, LIU Z, ZHU Y, et al. Implicit video emotion tagging from audiences' facial expression[C]//Multimedia Tools and Applications. 2014: 1-28.
- [35] ARAPAKIS I, KONSTAS I, JOSE J M. Using facial expressions and peripheral physiological signals as implicit indicators of topical relevance[C]//Proceedings of the 17th ACM International Conference on Multimedia. 2009: 461-470.
- [36] SOLEYMANI M, PANTIC M, PUN T. Multimodal Emotion Recognition in Response to Videos[J]. IEEE Transactions on Affective Computing, 2012, 3: 211-223.
- [37] SOLEYMANI M, CHANEL G, KIERKELS J J M, et al. Affective ranking of movie scenes using physiological signals and content analysis[C]//Proceedings of the 2nd ACM Workshop on Multimedia Semantics. 2008: 32-39.
- [38] WANG S, ZHU Y, WU G, et al. Hybrid video emotional tagging using users' EEG and video content[C]//Multimedia Tools and Applications. 2013: 1-27.
- [39] ARAPAKIS I, MOSHFEGHI Y, JOHO H, et al. Integrating facial expressions into user profiling for the improvement of a multimodal recommender system[C]//IEEE International Conference on Multimedia and Expo. 2009: 1440-1443.
- [40] BENINI S, CANINI L, LEONARDI R. A Connotative Space for Supporting Movie Affective Recommendation[J]. IEEE Transactions on Multimedia, 2011, 1: 1356-1370.
- [41] SOLEYMANI M, LICHTENAUER J, PUN T, et al. A Multimodal Database for Affect Recognition and Implicit Tagging[J]. IEEE Transactions on Affective Computing, 2012, 3: 42-55.
- [42] KOELSTRA S, MUHL C, SOLEYMANI M, et al. DEAP: A Database for Emotion Analysis Using Physiological Signals[J]. IEEE Transactions on Affective Computing, 2012, 3: 18-31.
- [43] BAVEYE Y, DELLANDREA E, CHAMARET C, et al. LIRIS-ACCEDE: A Video Database for Affective Content Analysis[J]. IEEE Transactions on Affective Computing, 2015, 6: 43-55.
- [44] ABADI M K, SUBRAMANIAN R, KIA S M, et al. DECAF: MEG-Based Multimodal Database for Decoding Affective Physiological Responses[J]. IEEE Transactions on Affective Computing, 2015, 6: 209-222.
- [45] PORIA S, CAMBRIA E, BAJPAI R, et al. A review of affective computing: From unimodal analysis to multimodal fusion[J]. Information Fusion, 2017, 37: 98-125.
- [46] ZHU Y, JIANG Z, PENG J, et al. Video Affective Content Analysis Based on Protagonist via Convolutional Neural Network[C]//Pacific Rim Conference on Multimedia: Advances in Multimedia Information Processing. 2016: 170-180.
- [47] YU G, LI X, SONG D, et al. Encoding physiological signals as images for affective state recognition using convolutional neural networks[C]//38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2016: 812-815.