

基于改进模糊综合评价的电影情感分类

林新棋

(福建师范大学数学与计算机科学学院 福州 350007)

(福建师范大学网络安全与密码技术福建省高效重点实验室 福州 350007)

摘要 为了提高电影情感分类精确度,以模糊数学理论为基础,建立电影底层特征和高层认知情感之间的关系,提出一种基于改进模糊综合评价的电影情感分类新算法。首先,选择了反映情感变化的场景亮度、镜头切换率和颜色能量作为视频场景底层特征,介绍了相应的特征提取方法。其次,引入和改进模糊综合评价模型,并给出特征对情感的模糊隶属函数,建立了单因素评价矩阵。最后,对于不同情感,采用层次分析法确定底层特征之间的相对权重,根据改进的模糊综合评价模型计算出电影场景的情感模糊特征向量,用最大判决值和阈值原则确定待识别场景的情感类型。实验结果表明,所提出的算法能有效地提高电影场景的情感分类精确度。

关键词 情感模糊特征向量,模糊综合评价,单因素评价矩阵,视频情感内容

中图分类号 TP391.4 文献标识码 A

Film Affective Classification Based on Improved Fuzzy Comprehensive Evaluation

LIN Xin-qi

(Department of Mathematics and Computer Science, Fujian Normal University, Fuzhou 350007, China)

(Key Laboratory of Network Security and Cryptography, Fujian Normal University, Fuzhou 350007, China)

Abstract In order to improve the classification accuracy of the film scene emotion, a novel algorithm was proposed based on the improved fuzzy comprehensive evaluation in the fuzzy mathematics theory by establishing the relationship between the low-level features and high-level cognitive emotion. First, the scene luminance, shot cut rates and color energy were selected as the low-level features for their special characteristics that can be used to better distinguish different types of human emotional reaction. Further, the extractive methods were put forward. Secondly, after introducing and improving the fuzzy comprehensive evaluation model, fuzzy membership functions were formed to measure the fuzzy relationship between low-level features and emotion, and then the single factor evaluation matrix was built. Finally, the method of the analytic hierarchy process (AHP) was used to determine the relative weight matrix between the features, and the affective fuzzy feature vector was computed by the improved fuzzy comprehensive evaluation model. And the affective type of the film scene was obtained by the maximization value of the components of the affective fuzzy feature vector and threshold at last. The experimental results show that the proposed algorithm can effectively improve the accuracy of the film affective classification.

Keywords Affective fuzzy feature vector, Fuzzy comprehensive evaluation, Single factor evaluation matrix, Video affective content

1 引言

近年来,随着网络技术、多媒体技术和超大规模集成电路技术突飞猛进的发展,数字视频成为网络中一个重要的媒体资源,对人们日常生活产生越来越大的影响。如何有效地表示、组织、管理和传输这些数据,使人们在任何时间和地方从海量的数字视频中找到自己感兴趣的相关视频片段已成为一种迫切的需求。当前,部分学者、研究人员认为分析视频的情感层次内容——视频情感内容有助于解决上述问题^[1-4]。因此,视频情感内容已经成为信息领域具有前沿特征的研究课

题,对研究视频个性化浏览、检索、分类、高层语义识别等具有重要意义,并逐渐呈现出广阔的商业前景。

视频情感内容就是包含在视频中的期待用户在观看该视频片段过程中被诱发的情感强度和情感类型^[5]。近来,国内外不少学者对视频情感内容进行了研究,也取得若干进展。Hanjalic 等人利用底层特征,通过建立情感曲线模型来提取视频情感内容的强度信息^[5]。Hong-Bong Kang 等人通过隐马尔可夫模型建立底层特征和基本情感(高兴、悲伤、害怕)之间的映射关系^[6],从而识别视频片段诱发的情感类型。针对连续剧电影,Min Xu 等人先用隐马尔可夫模型识别几个音频

到稿日期:2013-05-20 返修日期:2013-08-06 本文受国家自然科学基金(61070062,11071041),福建省教育厅基金项目(JA12075,JA10064,JB11036),福建省高等学校科技创新团队(IRTSTFJ,N. J1917)资助。

林新棋(1972—),男,博士,副教授,硕士生导师,主要研究方向为视频、图像处理、编码理论, E-mail: xqlin@fjnu.edu.cn.

事件,然后根据音频事件和情感之间的蕴涵关系推导出情感描述^[7]。孙凯等人将遗传算法和隐马尔可夫模型相结合来实现基本情感的识别^[8]。闫乐林等人根据人类情感反应具有不确定属性提出基于未确知数学理论的视频情感内容识别算法^[9,10]。

人的情感反应属于心理学范畴,情感状态变化受到用户内部各种驱动因素和外部各种刺激因素的双重影响,其变化具有模糊性,刻画人的情感状态需要考虑多因素的影响和模糊性。上述方法均未考虑情感的模糊属性,识别结果不是很理想。本文基于模糊数学理论,建立底层特征对情感类型的模糊隶属函数,通过单因素评价矩阵刻画了底层特征和情感之间的模糊关系,通过相对权重矩阵刻画多因素对各类情感的影响因子,提出一种基于改进模糊综合评价的电影情感内容分类新算法。

本文第2节介绍了情感关联底层特征提取方法;第3节介绍了改进的模糊综合评价模型,通过该模型将视频底层特征向量转化为一个蕴涵情感信息的情感模糊特征向量;第4节介绍了电影情感分类方法;第5节给出实验和仿真结果;最后是总结。

2 底层特征

对于那些视频底层特征与情感反应之间存在关系和它们的提取模型这两个问题,很多研究者或机构试图尝试从心理学、电影摄影术和主观实验3个方面进行探究。迄今已经获得一些令人兴奋的结论,具体参见文献^[5-7,11-14]。在总结现有结论的基础上,本文假设亮度^[5]、镜头切换率^[5,6,13]和颜色能量^[13,14]与诱发的情感类型直接相关,能够有效地识别情感类型。与此同时,第 k 个场景的亮度、镜头切换率、颜色能量特征提取模型分别如式(1)一式(3)所示。

$$ScL(k) = \text{medium}_{t_1 \leq t \leq t_2} \left\{ \frac{\text{medium}_{i,j}(L_{Shot(k,t)}(i,j)) \cdot \sum_{j=1}^Y \sum_{i=1}^X \text{Sig}(L_{Shot(k,t)}(i,j))}{\max_t \left\{ \text{medium}_{i,j}(L_{Shot(k,t)}(i,j)) \cdot \sum_{j=1}^Y \sum_{i=1}^X \text{Sig}(L_{Shot(k,t)}(i,j)) \right\}} \right\} \quad (1)$$

$$ScS(k) = \text{medium}_{n_1 \leq n \leq n_2} \left\{ \frac{\min \{ r(k,n) - l(k,n) \}}{r(k,n) - l(k,n)} \right\} \quad (2)$$

$$ScR(k) = \frac{\text{mean}_{t_1 \leq t \leq t_2} \left\{ \sum_{j=1}^Y \sum_{i=1}^X f_{Shot(k,t)}(i,j) \cdot \sum_{j=1}^Y \sum_{i=1}^X g_{Shot(k,t)}(i,j) \right\}}{\max_k \left\{ \text{mean}_{t_1 \leq t \leq t_2} \left\{ \sum_{j=1}^Y \sum_{i=1}^X f_{Shot(k,t)}(i,j) \cdot \sum_{j=1}^Y \sum_{i=1}^X g_{Shot(k,t)}(i,j) \right\} \right\}} \quad (3)$$

其中, $Shot(k,t)$ 表示第 k 个场景的第 t 个镜头的关键帧, $r(k,n)$ (或 $l(k,n)$)表示第 k 个场景的第 n 个镜头的右(左)边界视频帧索引号。如果 $L_{Shot(k,t)}(i,j) > \alpha$,那么 $\text{Sig}(L_{Shot(k,t)}(i,j)) = 1$,否则等于0;如果 $S_{Shot(k,t)}(i,j) \geq 0.2$,那么 $f_{Shot(k,t)}(i,j) = 1$,否则等于0;如果 $0^\circ \leq H_{Shot(k,t)}(i,j) \leq 105^\circ$,或者 $315^\circ \leq H_{Shot(k,t)}(i,j) \leq 360^\circ$,那么 $g_{Shot(k,t)}(i,j) = 1$,否则等于0。 X 和 Y 表示视频帧的分辨率。 $L_{Shot(k,t)}(i,j)$ 、 $S_{Shot(k,t)}(i,j)$ 和 $H_{Shot(k,t)}(i,j)$ 分别表示对应视频帧在像素 (i,j) 处的 L 分量值、饱和度值和色调值。

对于亮度,15个实验者根据视频帧亮度将随机抽自3部电影的160个视频帧分为Hight-Key帧、Low-Key帧和其他的。如果一个视频帧至少有10个相同的标注,如Hight-Key,那么将该帧视为Hight-Key帧。最后,得到54帧为Hight-Key帧,91帧为Low-Key帧。 L 分量均值统计数据如表1所列。根据表1数据可知参数 $\alpha=65$ 可作为区分明亮和暗的阈值。

表1 Hight-Key视频帧和Low-Key视频帧的 L 分量均值统计

L 平均值	40以下	40~50	50~60	60~70	70以上
Low频率	0.247	0.302	0.297	0.135	0.019
平均值	70以下	70~80	80~90	90~100	100以上
Hight频率	0.093	0.354	0.186	0.245	0.122

对于色彩,当像素饱和度低于0.2时,人类视觉系统将该像素感知为灰色,故用阈值0.2来区分彩色和灰色。对于色调,红、黄和橙属于暖色,而蓝、绿和紫属于冷色。同时,色调差低于 15° 时,人类视觉系统无法区分像素颜色。最后,根据牛顿色圆,将色调介于 0° 与 105° 或者 315° 与 360° 之间的像素视为暖色像素,其余的像素视为冷色像素。

3 情感模糊特征向量

3.1 模糊综合评价

模糊综合评价定义为在模糊环境中,根据多个因素对一事物进行恰当的评价,然后按照某种方法对评价结果的全体进行排序或选优。这种方法是基于Zadeh的模糊数学理论为基础的。模型描述如下:

设 $X = \{x_1, x_2, \dots, x_m\}$ 为因素集, $Y = \{y_1, y_2, \dots, y_n\}$ 为评价集, $R_i = \{r_{i1}, r_{i2}, \dots, r_{in}\}$ 为因素 x_i 对评价集的单因素评价,也就是 r_{ij} 表示因素 x_i 对评价 y_j 的模糊评价价值, $W = [\omega_1, \omega_2, \dots, \omega_m]$ 为权重矩阵, ω_i 表示第 i 因素在评价中的影响因子或重要程度,或者贡献程度,如果多因素作用下的综合评价结果设为 $B = [b_1, b_2, \dots, b_n]$,则式(4)给出模糊综合评价计算模型。

$$B = [\omega_1, \omega_2, \dots, \omega_m] \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ r_{21} & r_{22} & \dots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & \dots & r_{mn} \end{bmatrix} \quad (4)$$

在实际中,应用该模型的关键是解决单因素评价矩阵和权重系数矩阵这两个问题。目前,前者常采用专家打分法、综合评定法、概率统计法、隶属函数法等给出各因素对评价的模糊评价价值。后者多数是专业领域的专家根据多年积累的经验进行确定,或采用层次分析法确定。

3.2 改进模糊综合评价

对于电影场景情感内容来说,本文的目标是识别用户在观看电影过程中所产生的各种情感类型。由于用户的情感反应受到多个因素影响,包括视频内容、用户本身的文化修养、外部的环境和观看电影时的基本情感状态等,因此电影情感内容具有模糊属性。为了简化,假设观众在观看电影过程中被诱发的情感变化仅与视频内容有关,与本身和外界因素无关。

一旦采用场景特征向量来表示视频场景内容,那么影响用户情感反应类型的因素就是场景的各个特征取值。一方面,由于不用特征诱发,不同的情感反应的影响程度不尽相

同;另一方面,模糊综合评价模型中,对于评价集中的不同评价,不同因素采用相同影响程度。显然直接采用模糊综合评价模型来识别视频情感内容存在缺陷。因此,对于不同的评价,各个因素权重矩阵采用不同的影响程度是比较合理的,于是,本文提出如下的改进模型描述。

设 $\mathfrak{M}=\{x_1, x_2, \dots, x_m\}$ 为因素集,其中 x_i 表示电影场景的第 i 个特征值取值; $\mathfrak{Y}=\{y_1, y_2, \dots, y_n\}$ 为评价集,其中 y_j 表示第 j 个情感评价类型; $R_i=\{r_{i1}, r_{i2}, r_{i3}, \dots, r_{in}\}$ 为因素 x_i 对评价集的单因素评价,也就是 r_{ij} 表示第 i 个特征 x_i 对第 j 个情感评价类型 y_j 的模糊评价价值, $R=[r_{ij}]$ 称为单因素评价矩阵; w_{ij} 表示第 j 因素在第 i 个情感评价类型中的影响因子或重要程度,或者贡献程度,称 $W=[w_{ij}]$ 为相对权重矩阵;那么矩阵 $[w_{ij}][r_{ij}]$ 的第 (i, j) 元素可理解为实际诱发用户产生第 i 个情感类型的电影场景被评价为第 j 个情感类型的模糊综合评价价值,称为模糊相对综合评价矩阵,如果多因素作用下的综合评价结果设为 $B=[b_1, b_2, \dots, b_n]$,那么式(5)为情感模糊综合评价计算模型。

$$B = \text{diag} \left\{ \begin{matrix} [w_{11} & w_{12} & \dots & w_{1m}] \\ [w_{21} & w_{22} & \dots & w_{2m}] \\ \vdots & \vdots & \ddots & \vdots \\ [w_{n1} & w_{n2} & \dots & w_{nm}] \end{matrix} \begin{matrix} [r_{11} & r_{12} & \dots & r_{1n}] \\ [r_{21} & r_{22} & \dots & r_{2n}] \\ \vdots & \vdots & \ddots & \vdots \\ [r_{m1} & r_{m2} & \dots & r_{mn}] \end{matrix} \right\} \quad (5)$$

其中,diag表示提取矩阵的对角线上的元素构成的行向量,称向量 B 为电影场景的模糊综合评价向量。

由于向量 B 中第 i 个元素表示电影场景诱发用户产生第 i 类基本情感的模糊综合评价价值,因此向量 B 蕴涵情感信息,本文利用该向量表示场景的内容,称之为情感模糊特征向量。

为了将情感模糊特征向量用于识别电影场景的情感内容,下面给出模型中单因素评价矩阵和相对权重矩阵的度量方法。

3.3 单因素评价矩阵

如何建立每个影响因子对评价集中每一个元素的模糊评价价值是应用模糊综合评价模型的一个关键技术。由于不同的应用中确定模糊评价价值的方法不尽相同,鉴于情感类型具有模糊属性,本文采用模糊隶属函数方法来确定模糊评价价值。

对于每一个影响因素 x_i ,事实上, x_i 就是场景的一个低层特征值。因此,不同场景的取值不一定相同。若将电影数据库中所有场景的影响因素的取值构成一个集合,记为:

$$U_i = \{u_i(k) \mid u_i(k) \text{ 表示第 } k \text{ 个场景的第 } i \text{ 个特征值}, 1 \leq k \leq N\}$$

其中, N 表示电影场景数据库中场景总数,那么 U_i 构成一个论域。

对于论域 $U_i (i=1, 2, 3, \dots, m)$ 以及模糊概念“ y_j ”($j=1, 2, \dots, n$),定义论域 U_i 到区间 $[0, 1]$ 的映射 μ_{ij} 如下:

$$\mu_{ij}: U_i \rightarrow [0, 1] \\ u_i(k) \mapsto \mu_{ij}(u_i(k))$$

这里 $\mu_{ij}(u_i(k))$ 表示第 k 个电影场景中的第 i 个特征对诱发用户产生第 j 类基本情感“ y_j ”的贡献程度,那么 μ_{ij} 是 U_i 上的一个隶属函数, $\mu_{ij}(u_i(k))$ 表示 $u_i(k)$ 对模糊概念“ y_j ”的隶属程度,也就是第 i 个因素空间(第 i 个特征空间)对第 j 个评价对象(第 j 类基本情感)的模糊隶属函数。因此在此定义下就产生一个模糊集合,记为 A_{ij} ,它与模糊概念“ y_j ”紧密相关,称

为情感模糊集合。实际上,该集合能够刻画电影场景中所选择的特征对诱发用户产生各种基本情感的贡献程度。

于是,对于给定的特征集和评价集,根据上述定义可获得如下形式的矩阵:

$$\begin{bmatrix} \mu_{11}(u_1(k)) & \mu_{12}(u_1(k)) & \mu_{11}(u_1(k)) & \dots & \mu_{1n}(u_1(k)) \\ \mu_{21}(u_2(k)) & \mu_{22}(u_2(k)) & \mu_{23}(u_2(k)) & \dots & \mu_{2n}(u_2(k)) \\ \mu_{31}(u_3(k)) & \mu_{32}(u_3(k)) & \mu_{33}(u_3(k)) & \dots & \mu_{3n}(u_3(k)) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu_{m1}(u_m(k)) & \mu_{m2}(u_m(k)) & \mu_{m3}(u_m(k)) & \dots & \mu_{mn}(u_m(k)) \end{bmatrix} \triangleq R$$

那么矩阵 R 就刻画了电影场景内各个特征与诱发用户产生各种基本情感之间的模糊关系,称 R 为低层特征和基本情感之间的 Fuzzy(模糊)关系矩阵,定义 R 为单因素评价矩阵。一旦给定第 i 个因素空间对第 j 个评价对象的模糊隶属函数表达式,对于给定的场景就可以得到一个对应的单因素评价矩阵。

文献[6]表明某类情感的电影场景对应的同一个特征值具有相似性,因此,下面的假设是合理的:特征值偏离相似区域越大,诱发该类情感的模糊隶属度就越小。在此基础上,利用样本数据,采用距离成反比的思想来构建模糊隶属函数。

假设 $\bar{U}_{ij} = \{u_{ij}(t) \mid t=1, 2, \dots, M_{ij}\} (i=1, 2, 3, \dots, M; j=1, 2, 3, \dots, N)$ 表示第 j 个参考情感的所有场景对应的第 i 个因素取值集合,或特征集合。对于任意一个测试场景,其因素向量为 $(u_1(k), u_2(k), u_3(k), \dots, u_m(k))$,定义 $u_i(k)$ 与 \bar{U}_{ij} 的距离为 d_{ij} ,见式(6):

$$d_{ij} = d(u_i(k), \bar{U}_{ij}) = \sqrt{\frac{\sum_{t=1}^{M_{ij}} (u_i(k) - u_{ij}(t))^2}{M_{ij}}} \quad (6)$$

根据假设,令:

$$f_{ij}(u_i(k)) = \frac{c_{ij}}{d_{ij}} = \frac{c_{ij}}{d(u_i(k), \bar{U}_{ij})} \quad (7)$$

其中, c_{ij} 是一个调节参数,具体根据对应样本的情感主观隶属度来计算。

假设对应样本的情感主观隶属度为 $g_{ij}(u_{ij}(t))$, $t=1, 2, \dots, M_{ij}$,为了获得 c_{ij} 的估计值,求最优化方程(8)的解:

$$\min \left\{ \sum_{t=1}^{M_{ij}} (f_{ij}(u_{ij}(t)) - g_{ij}(u_{ij}(t)))^2 \right\} \quad (8)$$

一旦得到上述最优化方程的最优解,就得到参数 c_{ij} 的取值,进一步可以计算特征因素对情感的初始模糊隶属函数 f_{ij} 。由于同一个因素对诱发用户产生的所有情感的模糊隶属程度之和应该为 1,因此所求的模糊隶属函数形如式(9)所示。

$$\mu_{ij}(u_i(k)) = \frac{f_{ij}}{\sum_{j=1}^n f_{ij}} \quad (9)$$

其中, $i=1, 2, \dots, m; j=1, 2, \dots, n$ 。

3.4 相对权重矩阵

对于给定的研究问题,如何确定合理的相对权重矩阵是成功地应用改进模糊综合评价模型解决该问题的一个重要因素。目前,比较成熟的方法是采用运筹学家萨蒂教授提出的层次分析法。下面结合电影场景情感类型详细描述层次分析法的具体步骤。

第 1 步 假设评价因素为第 t 个,邀请专家对因素集中的任意两个因素的相对重要性根据表 2(比例标度)给出相应的评价序数,并将结果按表 3 格式进行填写,其中 $a_{ij}^t = f_i(x_i,$

x_j),那么称矩阵 $A=(a_{ij}^t)$ 为因素集对第 t 个评价因素的判断矩阵。

表2 比例标度定义

因素 x,y 相比较	说明	$f_t(x, y)$	$f_t(y, x)$
x 与 y 同等重要	x, y 对第 t 个评价因素有相同的贡献	1	1
x 比 y 稍微重要	x 的贡献稍大于 y, 但不明显	3	1/3
x 比 y 明显重要	x 的贡献明显大于 y, 但不十分明显	5	1/5
x 比 y 十分重要	x 的贡献十分明显大于 y, 但不特别突出	7	1/7
x 比 y 极其重要	x 的贡献以压倒优势大于 y	9	1/9
x 比 y 处于上述两相邻判断之间	x 的贡献稍大于 y, 但不明显	2,4, 6,8	1/2,1/4, 1/6,1/8

表3 对评价因素 y_t 的判断矩阵

	x_1	x_2	...	x_m
x_1	a_{11}^t	a_{12}^t	...	a_{1m}^t
x_2	a_{21}^t	a_{22}^t	...	a_{2m}^t
...
x_m	a_{m1}^t	a_{m2}^t	...	a_{mm}^t

第2步 鉴于判断存在局限性,根据比例标度建立的判断矩阵未必与实际存在的关系完全一致,因此采用式(10)所定义的一致性指标对判断矩阵进行一致性判断。

$$CI = \frac{\lambda_{\max} - m}{m - 1} \quad (10)$$

其中, λ_{\max} 表示判断矩阵 A 的最大特征值。一般认为 $CI \leq 0.1$ 时,判断矩阵是可以接受的。当 $CI = 0$ 时, $\lambda_{\max} = m$, 这与理论相符; CI 值越大说明一致性判断越差。

第3步 根据判断矩阵求解权重。采用式(11)所定义的和法求之。

$$w_i = \frac{\sum_{j=1}^n a_{ij}^t}{\sum_{i=1}^n \sum_{k=1}^m a_{ik}^t} \quad (11)$$

于是因素集的 m 个因素对第 t 个评价因素 y_t 的综合评价贡献向量为 $W_t = (w_{t1}, w_{t2}, w_{t3}, \dots, w_{tm})$ 。进一步可以获得相对权重矩阵 $W = [W_1^T, W_2^T, \dots, W_n^T]^T$, 其中 $(\cdot)^T$ 表示矩阵的转置运算。

4 情感分类方法

对于给定的场景,设特征向量为 $(u_1(k), u_2(k), \dots, u_m(k))$, 根据改进模糊综合评价模型,则该场景的情感模糊特征向量 $B = [b_1, b_2, \dots, b_n]$ 可以用来描述场景内容。本文的目标就是根据场景的情感模糊特征向量给出场景激发的情感类型。显然,对于情感模糊特征向量 $B = [b_1, b_2, \dots, b_n]$, 选择最大判决值所对应的情感作为视频场景所诱发情感类型的判决结果是合理的。于是,判决函数定义形如式(12)。

$$d = \arg \max \{b_1, b_2, \dots, b_n\} \quad (12)$$

其中, d 表示最大值所对应的分量标号,即 $d = 1, 2, 3, \dots, n$ 中的一个或若干个。

但是,如果 b_1, b_2, \dots, b_n 都很小,那么在这种情况下根据判决函数做出的情感判断是很勉强的,有时甚至是不合理的。进一步,引入一个阈值 $\lambda \in [0, 1]$ (实验数据 $\lambda = 0.4$)。

如果 $\max \{b_1, b_2, \dots, b_n\} < \lambda$, 则认为电影场景所诱发的情感不是给定的这几类,或没有诱发用户产生情感,这时候做出“拒识”判决结果。

如果 $\max \{b_1, b_2, \dots, b_n\} \geq \lambda$, 则选择最大判决值所对应的

情感作为视频场景所诱发情感类型的判决结果。但是,当超过 λ 的最大值对应的类型不止一个时,随机输出一个值。

综上可知,对于给定的电影场景,根据底层特征提取模型和改进的模糊综合评价模型就能计算出一个情感模糊特征向量,再根据上述的情感类型识别思想就可给出该场景诱发的情感类型。

5 实验结果分析

为了验证上面提出的电影场景情感分类方法的可行性和性能,根据底层特征提取模型,验证用的因素集合为{亮度,镜头切换率,颜色能量},评价集合为{高兴,悲伤,害怕或担心},也就相当于3.2节提及的模型中 $m = n = 3$ 。

5.1 实验环境与实验数据

仿真相关参数:VS2005.NET 的 1.1 版本的 OpenCV, Microsoft WindowsXP 操作系统, P4-2.50GHz 的 CPU, 2GB 的内存, 200G 的硬盘容量。采用 17 英寸的电脑显示器、耳机等相关设备。

实验数据由 9 个长片电影:《宝贝计划》、《K 歌情人》、《活着》、《暖春》、《异灵第六感》、《男才女貌》、《我的兄弟姐妹》、《森冤》和《午夜凶铃》组成,片长总共约 973 分钟。首先利用文献[14]中算法将其分割成镜头;再利用文献[15]中算法识别场景边界,并进行适当的手工调整,最后共分成 654 个场景。

5.2 相对权重矩阵

根据 3.4 节,针对因素集合{亮度,镜头切换率,颜色能量}和评价集合{高兴,悲伤,害怕或担心},在专家给出因素集合中各因素对每个评价因素的相对重要性评价的基础上,获得的判断矩阵如表 4—表 6 所列。

表4 对评价因素 y_1 (高兴)的判断矩阵

	x_1	x_2	x_3	贡献
x_1	1	5	3	0.5378
x_2	1/5	1	1/5	0.0837
x_3	1/3	5	1	0.3785

表5 对评价因素 y_2 (悲伤)的判断矩阵

	x_1	x_2	x_3	贡献
x_1	1	1	3	0.4286
x_2	1	1	3	0.4286
x_3	1/3	1/3	1	0.1428

表6 对评价因素 y_3 (担心)的判断矩阵

	x_1	x_2	x_3	贡献
x_1	1	1/5	3	0.2510
x_2	5	1	5	0.6574
x_3	1/3	1/5	1	0.0916

根据表 4, $\lambda_{\max} = 3.1356$, $CI = 0.0678 < 0.1$, $W_1 = [0.5378, 0.0837, 0.3785]$; 根据表 5, $\lambda_{\max} = 3$, $CI = 0 < 0.1$, $W_2 = [0.4286, 0.4286, 0.1428]$; 根据表 6, $\lambda_{\max} = 3.1356$, $CI = 0.0678 < 0.1$, $W_3 = [0.2510, 0.6574, 0.0916]$ 。进一步, 3.4 节的相对权重矩阵具体如下:

$$W = \begin{bmatrix} W_1 \\ W_2 \\ W_3 \end{bmatrix} = \begin{bmatrix} 0.5378 & 0.0837 & 0.3785 \\ 0.4286 & 0.4286 & 0.1428 \\ 0.2510 & 0.6574 & 0.0916 \end{bmatrix}$$

5.3 主观评价

为了获得每个场景诱发观众的情感类型,对分割好的场

景进行主观评价实验。实验共有 10 名研究生参与,分成 3 个步骤。首先,每个实验者独立观看每个场景;其次,看完记录自己的情感反应类型;第三,休息 1 分 30 秒后,重复测试其他的场景。

实验结束后,对数据进行统计,如果某个场景的记录中有 n 个人标注同一个情感类型,当 $n \geq 7$ 时,就把该情感类型作为场景诱发情感类型的主观标注结果。最终,实验样本由 76 个“高兴”场景、55 个“悲伤”场景和 35 个“害怕”场景组成。

5.4 情感主观隶属度

为了获得 3.3 节中模糊隶属函数的参数,从样本中随机抽取 17 个“高兴”场景、17 个“悲伤”场景、17 个“害怕”场景。将这 51 个场景计算的底层特征值用于构建模糊隶属函数,并进行如下的实验。

被试者首先填写基础情感报告表,然后观看场景片段,接着,被试者根据自己观看过程的情感反应回答下面问题。举个例子,假设场景的情感类型主观标注为“高兴”,那么设计的问题具体为:“该电影场景能够诱发用户产生高兴类的情感吗?”,问题答案为:“真的”、“大致真的”、“似真似假”、“大致假的”和“假的”。每个被试者选择一个答案进行回答。其余的类似进行。

将上面获得的答案根据表 7 进行量化处理。对于每一个场景,计算 10 个答案的主观量化值的平均值,并把这个平均值作为该场景诱发对应的情感类型的主观隶属度。

表 7 主观量化表

答案	假	大致假	似真似假	大致真	真
量化值	0	0.25	0.5	0.75	1

根据 3.3 节的最优化问题,结合样本底层特征向量和主观隶属度,最终求出参数 C_{ij} 的值,用矩阵表示如下:

$$C = [c_{ij}] = \begin{bmatrix} 0.1311 & 0.2068 & 0.0964 \\ 0.1994 & 0.1389 & 0.1563 \\ 0.1490 & 0.2178 & 0.1345 \end{bmatrix}$$

5.5 性能评价

为了检测算法的性能,定义精确度和识全率两个性能评价指标,也就是精确度 = (正确检测数) / (正确检测数 + 虚假数),识全率 = (正确检测数) / (正确检测数 + 丢失数)^[16]。统计算法识别结果如表 8 所列。计算两个性能指标,并与现有的方法进行比较,获得的结果如表 9 所列。

表 8 情感类型检测统计结果

情感类型	正确检测	丢失	虚假
高兴	51	8	4
悲伤	31	7	4
害怕	16	2	2

表 9 与现有方法的对比(%)

	高兴	悲伤	害怕	
精确度	文献[6]	78.40	76.50	81.30
	文献[8]	93.10	86.20	78.40
	文献[17]	89.80	84.38	88.24
	本文	92.73	88.57	88.89
识全率	文献[17]	74.58	71.05	83.33
	本文	86.44	81.05	88.89
	提高	11.86	10.00	5.56

从表 9 数据可知:本文提出的方法大部分情感内容都能被正确识别,识别精确度大于 80%。与文献[6]相比,对应的

精确度至少高出 10%。与文献[8]相比,害怕类的情感识别精确度高出 10%,高兴类的精确度略低一点。与文献[17]相比,3 类基本情感的识别精确度都有一定的提高,尽管提高不大,但是对于高兴类和悲伤类,在识全率方面,本文提出的算法比文献[17]提高超过 10%。至于其他方法,文献[7]仅考虑连续剧的基本情感分类,能否用于其他类型电影尚未考证,故缺乏可比性。因此,本文提出的算法是有效的。

为了使新提出的算法更具客观性和通用性,主观评价和情感主观隶属度实验中的测试者由两组不同成员组成,并且用于建立模糊隶属函数的视频样本和性能评价用的视频样本也是不同的。另外,本文测试用的视频场景均带有明显的情感色彩,对于情感色彩较淡的场景视为其它类别。

结束语 本文使用场景亮度、镜头切换率和颜色能量 3 个底层视频特征作为电影情感类型识别的影响因素,考虑情感反应的模糊属性,充分利用模糊数学理论的优势,提出了一种基于改进模糊综合评价模型的电影场景情感分类新算法。实验结果表明,提出的算法分类精确度超过 80%,算法是有效的,达到预期的效果,与其他算法相比,平均检测精确度提高了 6.0278%。

尽管本文仅使用 3 个底层特征,使分类精确度和识全率有一定的提高,但是人类的情感反应是十分复杂的,要建立一个以电影作为刺激源、符合人类情感反应变化规律的情感模型,无疑将是一个非常有挑战性的课题。笔者认为,利用模糊理论的观点研究视频情感内容并予以数量化,为视频情感内容的研究提供了一条新的有益的探索途径。而且,随着对声音、运动、形状和纹理等视频底层特征和情感反应之间关系的研究的逐渐深入,将来基于模糊数学理论的视频情感分类精确度和识全率必定会有所提高。

参考文献

- [1] Hanjalic A. Video and image retrieval beyond the cognitive level: the needs and possibilities[C]//Proc. SPIE on Storage and Retrieval for Media Databases. San Jose, CA, USA, Jan. 2001: 130-140
- [2] Ong K-M, Kameyuma W. Classification of video shots based on human affect[J]. Information and Media Technologies, 2009, 4 (4): 903-912
- [3] 孙凯, 于俊清. 面向观众的个性化电影情感内容表示与识别[J]. 计算机辅助设计与图形学学报, 2010, 22(1): 136-144
- [4] Cui Yue, Jin Jesse S, Zhang Shi-liang, et al. Music Video affective understanding using feature importance analysis[C]//Proceedings of the ACM International Conference on Image and Video Retrieval. 2010: 213-319
- [5] Hanjalic A, Xu L Q. Affective video content representation and modeling[J]. IEEE Transactions on Multimedia, 2005, 7 (1): 143-154
- [6] Kang H B. Affective content detection using HMMs[C]//Proc. of 11th ACM International Conference on Multimedia. Berkeley, CA, USA, Nov. 2003: 259-262
- [7] Xu M, Chia L T, Yi H L, et al. Affective Content Detection in Sitcom Using Subtitle and Audio[C]//the Proceedings of 12th International Conference on Multimedia Modeling Processing. 2006: 129-134

(下转第 196 页)

97%，在规模为 64 和 128 的路由组中，占比都在 90% 以上，依然保持了较高的准确度。

进一步分析表 4 的数据还可发现，冲突路由平均分组数与路由组规模等级近似成正比关系，如图 2 所示。由此可推知，冲突路由平均分组数与路由组规模近似成对数函数关系，如图 3 所示。这说明，算法生成的路由分组数受路由组规模的影响加速递减，这使得即使在较大规模的路由组中，算法生成的路由分组数也是很有限制的。例如，在规模为 1024 的路由组中，实验所得冲突路由平均分组数仅为 5.56，最大分组数也只有 7。这体现出分批路由方法较强的规模适应性。

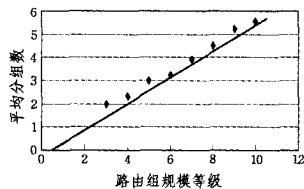


图 2 平均分组数与路由组规模间的关系

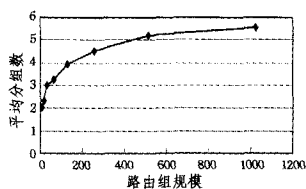


图 3 平均分组数与路由组规模之间的关系

结束语 混洗交换网络中，如何确定冲突路由的最小无冲突分组是分批路由的核心问题。基于布尔代数的最小无冲突分组生成方法将分组的确定转化为布尔运算，虽然运算的时间复杂度受路由组规模影响较大，但其计算方法简单，并且可计算所有最小无冲突分组，是混洗交换网络研究的又一重要理论成果。最小无冲突分组近似算法实现了冲突路由的快速分组，不仅时间效率高，而且具有较高的准确度，能满足大规模信息交换的需求，它与文献[15]中的路由冲突检测算法结合将构成 Omega 网络中一个理想的分批路由实施方案。

参考文献

[1] Stone H S. Parallel Processing with the Perfect shuffle[J]. IEEE Transactions on Computers, 1971, C-20(2): 153-161
 [2] Goke L R, Lipovski G J. Banyan networks for partitioning multi-processing systems[C]//Proc First Annual Computer Architecture Conf. 1973; 21-28
 [3] Wu C, Feng T. On a Distributed processor communication architecture[C]//Proc Compecon Fall, 1980; 599-605
 [4] Akbar S, Reza S N, Hamid Sarbazi A. The Shuffle-Exchange

Mesh Topology for 3D NoCs[C]//The International Symposium on Parallel Architectures, Algorithms, and Networks. IEEE, 2008, 23: 275-280

[5] Chou W Y, Chen R B, Chen Chiu-yuan. All-to-All Personalized Exchange Algorithms in Generalized Shuffle-exchange Networks [C]//Eighth International Conference on Networks, 2009; 185-190
 [6] Muller O, Baghdadi A, Jezequel M. From parallelism levels to amulti-asip architecture for turbo decoding[J]. IEEE Transactions on Very Large Scale Integration(VLSI) Systems, 2009, 17(1); 92-102
 [7] Ding N, Ma H L, Tan G Z. The Parallelization Design of Chaotic Encryption System Based on Multilevel Shuffle-Exchange Network[C]//Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, 2009; 698-671
 [8] Lawrire D H. Access and alignment of data in an array processor[J]. IEEE Transactions on Computers, 1975, C-24(12): 1145-1155
 [9] Clonan T J. Topological equivalence of optical crossover networks and modified data manipulator networks[J]. Appl. Opt., 1989, 28(13): 94-98
 [10] Raghavendra C S, Varma A. Rearrangeability of the 5-Stage Shuffle-Exchange Network for $N=8$ [A]//Proceedings of 1986 International Conference on Parallel Proceeding[C]. University Park, USA, 1987; 119-122
 [11] 戴浩, 沈孝钧. 在 7 级混洗交换网络中实现 16×16 的可重排性[J]. 电子学报, 2007, 35(10): 1875-1885
 [12] 葛方斌, 张涛, 宋金玉, 等. $3n-1$ 级混洗交换网络的重排性研究[J]. 通信学报, 2011, 32(10): 10-18
 [13] 李挥, 何伟, 伊鹏, 等. 排序集线器多级互连交换结构的多路径自路由模型[J]. 电子学报, 2008, 36(1): 1-8
 [14] Linial N, Tarsi M. Interpolation between bases and the shuffle exchange network[J]. European J Combin, 1989, 10(1): 29-39
 [15] Ge F B, Zhao M, Zhang T, et al. A new policy to solve routing conflict in shuffle-exchange network[J]. Science China Information Sciences, 2011, 54(7): 1512-1523
 [16] Raghavendra C S. On the Rearrangeability Conjecture of $(2\log 2N - 1)$ -stage Shuffle-Exchange Network [A]//IEEE Computer Society Technical Committee on Computer Architecture Newsletter[C]. 1994-95; 10-12

(上接第 165 页)

[8] Sun K, Yu J Q. Video affective content representation and recognition using video affective tree and hidden Markov models [C]//Proc. of the second International Conference on Affective Computing and Intelligent Interaction (ACII2007), Lisbon, Portugal, Sep. 2007; 594-605
 [9] 闫乐林, 温向明, 郑伟, 等. 基于未确知测度的视频情感内容识别[J]. 东南大学学报: 自然科学版, 2011, 41(3): 473-477
 [10] 闫乐林, 冯希叶. 一种基于内容的视频情感类型识别算法[J]. 计算机系统应用, 2011, 20(3): 102-105
 [11] Valdez P, Mehrabian A. Effects of color on emotions[J]. Journal of Experimental Psychology: General, 1994, 123(4): 394-409
 [12] Zettl H. 图像、声音、运动; 实用媒体学[M]. 赵森森, 译. 北京: 北京广播电影学院出版社, 2003

[13] Wang H L, Cheong L F. Affective understanding in film [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2006, 16(6): 689-704
 [14] Lefevre S, Vincent N. Efficient and robust shot change detection [J]. Journal of Real-Time Image Processing, 2007, 2: 23-34
 [15] Sundaram H, Chang S F. Computable scene and structures in films[J]. IEEE Transactions on Multimedia, 2002, 4(4): 482-491
 [16] Su J H, Huang Y T, Yeh H H, et al. Effective content-based video retrieval using pattern-indexing and matching techniques [J]. Expert systems with applications, 2010, 37(11): 5079-5085
 [17] 林新棋, 严晓明, 郑之. 基于模糊理论和三段论推理的电影情感分类[J]. 山东大学学报: 工学版, 2011, 41(4): 61-67