

文章编号: 1001 - 9081 (2009) 12 - 3422 - 05

电影视频景别音阶的识别与检测

于俊清, 黄岳

(华中科技大学 计算机科学与技术学院, 武汉 430074)

(yjqing@hust.edu.cn)

摘要:通过对电影景别音阶的识别检测,可以有效地分析和检索电影视频情感变化的片段。在系统分析以往研究成果的基础上,利用电影领域知识构建了局部运动占有率、摄像机运动和镜头间相似度等新的特征,结合常用的视频特征,采用贝叶斯分类器来识别电影视频的镜头景别,并根据景别变化同观众情感之间的关系,设计了5种能够激发观众情感的景别音阶,在景别识别的基础上实现了对景别音阶的检测。实验结果表明,选取的特征能够得到较好的检测结果,与其他方法相比,远景与近景的识别在准确率和查全率上均有不同程度的提高。

关键词:景别音阶;特征提取;分类;识别

中图分类号: TP391 **文献标志码:** A

Scene tune recognition and detection in film videos

YU Jun-qing, HUANG Yue

(School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan Hubei 430074, China)

Abstract: It has been a research hotspot to retrieve quickly and accurately the film clips that reflect the fluctuation of audiences' mood according to individual user's needs. Detection of the film scene tune is an effective way to detect these clips. New features such as local motion share ratio, camera motion and shots similarity were proposed with the combination of film domain knowledge. Using the above-mentioned features and common features mentioned in other papers, a Bayesian classifier was utilized to classify shot types. According to the relationship between scene and emotion of audience, five scene tunes that could stimulate emotion of audiences were defined and detected based on the recognition of shot types. The experimental results show that the chosen features achieve good results. Compared with the results of other studies, both precision and recall rate of long shot and close-up shot are improved.

Key words: scene tune; feature extraction; classification; recognition

0 引言

近年来,由于视频数据呈现出爆炸性的增长,如何对这些数据进行有效地组织与管理,让用户能够从海量的视频信息中检索到自己需要的片段,已成为当今多媒体领域的研究热点。在这种背景下,基于内容的视频检索(Content-based Video Retrieval, CBVR)便应运而生。电影视频是用户最为喜闻乐见的一种视频类型,随着科技的发展,人们不仅仅满足于电影视频的单纯播放,而提出了许多更高的要求。例如,自动检测到引起观众情绪变化的电影视频片段,或者观众希望浏览特殊景别组合的视频片段,有的观众也希望浏览引起特定情绪的片段。为满足上述需求,通过研究电影拍摄艺术和视觉心理等相关领域知识,可以发现电影视频的景别音阶能够有效地诱发观众的情感变化。景别是指由于摄像机相对静止物体的运动或者物体本身的运动,造成画面上被摄物体形象的改变^[1],一般分为近景、中景和远景。不同景别的镜头可以带给观众不同的感受,景别之间的组合又能产生单个镜头所不具有的含义,更能带给观众巧妙的视觉变化,从而诱发不同的心理感受。例如近景、中景、远景这样的镜头组合使节奏逐渐加快,让观众的视线逐渐从整体引向局部,注意力也

逐渐趋于集中。我们将具有特定意义、能够激发观众情感变化的镜头组合称为景别音阶。检测和识别视频片段中的景别音阶,能够辅助在电影视频中寻找到观众情感变化的片段,可以应用在视频检索和视频点播等领域,也是电影摘要生成的关键技术之一,具有很好的理论研究和实际应用意义,同时也具有一定的挑战性。

本文在系统分析镜头景别分类和识别现有研究成果的基础上,针对存在问题定义并提取了能够反映镜头景别类型特征的特征,提出并实现了电影视频镜头景别的识别方法和特定景别音阶的检测方法。

1 相关工作

当前通过检测景别音阶来检索反映观众情绪变化片段的相关研究较少。单纯对镜头景别的识别,有不少研究者涉及,但大多基于体育视频或者新闻视频,由于电影视频的限定规则和领域知识相对其他视频类型较少,因此专门针对电影视频的很少。文献[2]利用足球视频中草地颜色占整个画面中的比例,将镜头类型分为近景、缩放和全景。文献[3]利用球场的形状,在网球视频和斯诺克比赛视频中检测是否出现球场的矩形形状,利用出现矩形面积的大小判断镜头是否属于

收稿日期: 2009 - 07 - 22。

基金项目: 国家自然科学基金资助项目(60703049); 中国博士后基金资助项目(20060400847); 华为基金资助项目(YJCB2006047N)。

作者简介: 于俊清(1975 -),男,内蒙赤峰人,副教授,博士,CCF高级会员,主要研究方向:数字媒体处理与检索; 黄岳(1985 -),男,福建福州人,硕士研究生,主要研究方向:数字视频处理。

球场全景。在球类运动中,利用边缘点的比例和分布规律也已经成为一种镜头分类的方法^[4]。无论是电影视频还是体育视频,最重要的拍摄焦点总是人物,其中人脸所在的位置和比例直接反映了镜头景别的不同。文献[5]利用人脸中心到视频帧画面中心的距离,结合人脸的大小、镜头长度、摄像机运动和某些音频特征通过贝叶斯分类器,将视频镜头分为全景、远景、中景、特写和大特写。相对于远景镜头,在近景镜头中局部运动一般比远景镜头要激烈,所以局部运动活跃度作为反映局部运动激烈程度的一种特征也常受到研究者的青睐,已被应用到体育视频的镜头分类中^[6]。多种视频特征相结合或者视频特征与音频特征相结合的方法也是一种常用的镜头分类方法。文献[7]结合足球视频中的重放镜头的检测和比赛中断过程中的音-视频特征来区分精彩镜头。文献[8]通过提取运动向量域、文本、颜色和镜头长度这些底层特征,根据领域知识模型映射为相应的中层特征,然后利用决策树进行镜头的分类。

通过对以往研究成果的分析可以看出,研究者所使用的特征对体育视频的镜头景别有较好的分类效果,但并不能应用在电影视频中。本文结合电影领域知识,构建了反映电影镜头景别特点的一些特征,并利用这些特征来对镜头的景别进行分类识别。

2 特征的设计、提取和选择

在景别音阶的提取过程中,对每一个镜头提取特征向量并进行分类识别是一个重要的过程。如何使提取的特征向量中的每一维都能够有效地反映电影镜头景别的特点是景别识别的关键。

2.1 局部运动占有率

利用观众对电影画面中不同区域关注度不同的原理,将视频画面划分成不同的区域^[9],并设定不同的权值,如图 1 所示。局部运动占有率能在一定程度上区分电影镜头的景别。在许多电影视频中,近景镜头的局部运动面积一般较大,局部运动活跃度高。在远景镜头中,通常没有局部运动,或者局部运动面积占整个画面的范围小,局部运动活跃度较低。

1/3	1/3	1/3	
1	2	1	3/12
4	8	4	4/12
1	2	1	5/12

图 1 画面关注区域的权值设定

局部运动占有率特征的计算公式为:

$$M_{in} = \frac{1}{k} \sum_{R} \left(\frac{M_k}{A_{area}} \times \frac{w_{pos}^k}{8} \right) \quad (1)$$

其中: k 代表区域的个数, M_k 表示每一个区域的局部运动的块的数目, A_{area} 表示一帧中块的总数, w_{pos}^k 表示第 k 个区域所占的权值, R 代表帧的整个范围。

为了验证局部运动占有率特征的有效性,实验中截取了影片《国家宝藏》中的一个场景片段,如图 2 所示。每一帧的局部运动占有率如图 3 所示,其中垂直线表示镜头的边界。由图可见,第 4~6 个镜头的局部运动占有率比较高,都是近

景;第 8 个镜头局部运动占有率较低,属于远景;第 2 和 3 个镜头局部运动占有率在两者之间,则是中景。可见,局部运动占有率特征在反映镜头景别上是有效的。从镜头 1 来看,是近景镜头,但其局部运动并不明显。第 7 个镜头虽然是一个中景,但局部运动都集中在中央部分,权值较大,所以局部运动占有率也呈现出较高的数值。由以上分析可知,局部运动占有率是一个有效的特征,但是并不适用于所有的情况,如局部运动较少的近景就难以被区分。



图 2 电影《国家宝藏》中一个场景的镜头关键帧序列

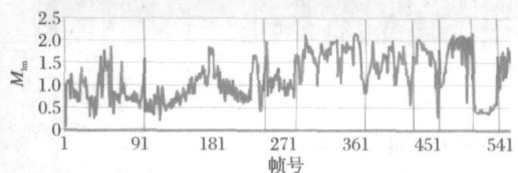


图 3 电影《国家宝藏》中一个场景的局部运动占有率曲线

2.2 人脸特征

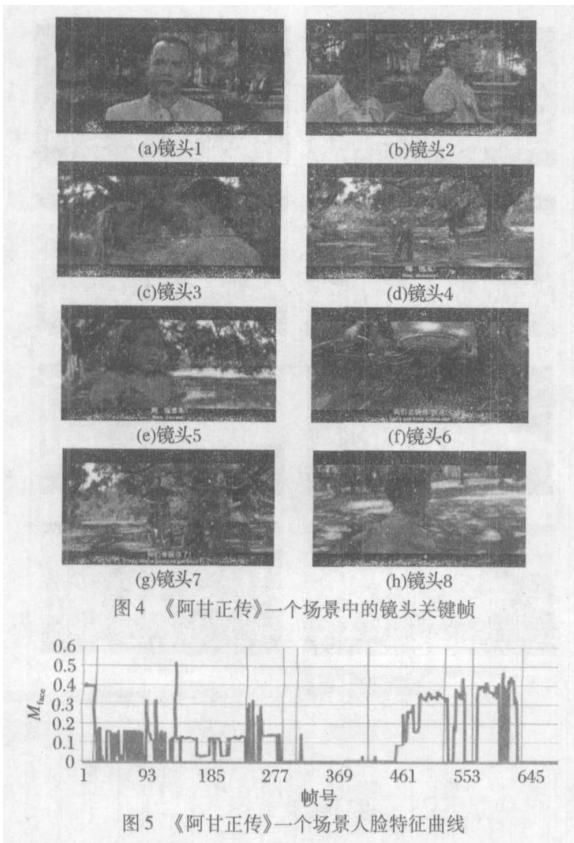
人是在电影视频中出现最多的元素。通常在近景的对话镜头中,人脸所占的比例较大,较容易检测。对中景镜头来说,通常在一个画面中出现多个人脸,人脸占有的比例较小,而且位置也不一定在画面的中心部分。远景镜头一般并不表现人脸。综上所述,为了识别出镜头的景别,需要在一个画面中得到一些人脸的信息,具体有人脸出现的位置、大小以及人脸的个数。人脸检测在具体的实现中采用了 Adaboost 算法^[10]。

人脸特征的计算如下:

$$M_{face} = \max \left(\frac{A_1}{A_{frame}}, \frac{A_2}{A_{frame}}, \dots, \frac{A_k}{A_{frame}} \right) \times \frac{w_{pos}^i}{8} \quad (2)$$

其中: A_k 表示第 k 个人脸所占有的面积,用像素数目来表示; A_{frame} 表示该帧的像素总数; w_{pos}^i 表示第 i 个人脸所在画面位置的权值,权值的设定与画面关注区域的权值相同。公式表示求出最大的那个人脸占整个帧画面的比,再乘以它所处位置的权值。

本文随机抽取了电影《阿甘正传》中的一个场景,如图 4 所示。计算该场景每一帧的人脸特征值,得到的结果如图 5 所示。从图 4 中可以看出,第 1 和第 8 个镜头都属于近景,在第 6 个镜头中没有人脸的出现,其较高的 M_{face} 值应为人脸检测算法的误检所致。第 2 个镜头和第 3 个镜头的平均 M_{face} 值较上述近景镜头低,但又不为 0,为中景镜头。实验中每一个镜头的 M_{face} 值是对该镜头中检测出人脸的帧的 M_{face} 求均值得到的。



2.3 摄像机运镜模式特征

电影视频的摄像机运镜模式主要分为摄像机静止、水平转动、垂直转动、变焦和其他模式。求取摄像机运镜模式的算法可以参考文献 [11]。

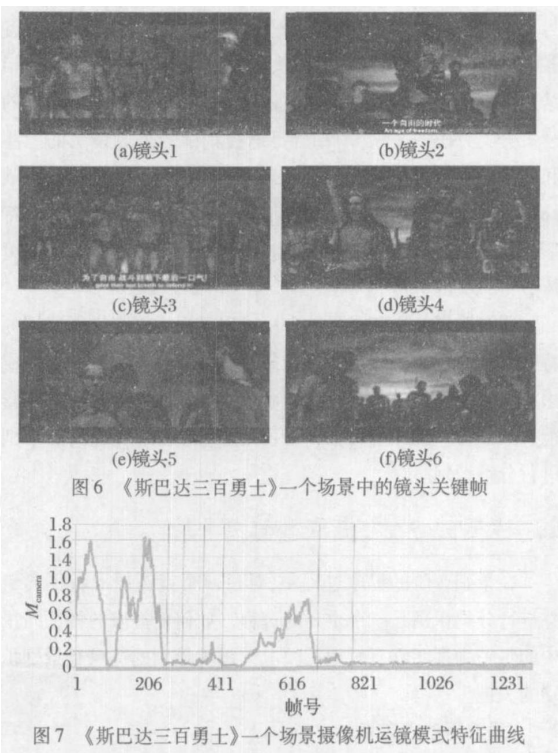
将摄像机运镜模式作为一个放大因子,对镜头活跃度进行放大,如式 (3) 所示:

$$M_{camera} = (A_{shot})^{am} \quad (3)$$

其中: A_{shot} 是该镜头的运动活跃度; am 作为一个放大因子,体现了摄像机的运镜模式, $am \in [0, 2]$ 。

由于摄像机的推进是一个由中景到近景或者远景到中景的渐进过程,景别的变化最为强烈,所以 am 设置为 2,同理摄像机拉远的 am 设置为 0。摄像机静止运镜手法和其他运镜手法在三种不同的景别中都被用到, am 设为 1。摄像机水平运动分为有跟踪和无跟踪拍摄手法两种。跟踪拍摄手法多体现于近景,因此其放大系数被设置为 1.5。至于跟踪拍摄手法的检测只需要判断摄像机运动和局部运动是否具有一致性即可,无跟踪拍摄手法的水平运动和摄像机垂直运动一般出现在中景或者远景, am 设置为 0.5。针对电影《斯巴达三百勇士》随机提取了如图 6 所示一个场景中的几个镜头。对该场景逐帧求 M_{camera} 值,见图 7。第 1 个镜头是一个跟踪拍摄的

镜头,又表现出摄像机的水平运动,所以其 M_{camera} 值较高,而第 5 个镜头是先静止后跟踪的镜头,这些都是近景镜头, M_{camera} 在图中也呈现出较高的数值;在第 2~4 个镜头中,虽然其中有近景,但由于摄像机的运镜手法是静止,基本没有全局运动,所以其 M_{camera} 曲线较低。由以上分析容易知道, M_{camera} 值可以区分有摄像机运动情况(摄像机非静止状态)下的近景和远景,但是对没有摄像机运动(摄像机静止)的镜头景别区分无能为力。



2.4 其他特征

除了以上特征外,还可以提取一些常被用来检测镜头景别的视频特征,如镜头间相似度、镜头持续时间和边缘点占有率等。

镜头间相似度被归一化到 0~1,表示当前镜头与参考镜头从视觉上的相似程度。在电影视频中,远景镜头常被用来作为一个过渡镜头,所以远景镜头与其下一个镜头的镜头间相似度通常低于近景镜头和中景镜头。利用这样的领域知识,提取镜头相似度特征用来分类镜头的景别是可行的。我们把当前一个镜头与下一个镜头的颜色族直方图 [12] 的相似度作为当前镜头的相似度。

镜头的持续时间也能在某种程度上反映镜头的景别。远景镜头的持续时间相对较短;远景镜头包含的内容较多,必须要给观众足够的时间来观察镜头内的事物,因此平均持续时间应该较长,可以简单地利用镜头包含的帧数来代替镜头持续时间。

边缘点占有率表示整个帧是属于边缘点的像素占整个帧像素的比例,它是识别远近景中比较有效的特征。在近景镜头中,导演着力表现事物的细节部分,摄像机镜头离被摄物体较近,画面中呈现出相同纹理的区域较大,低纹理区域较多,呈现的边缘点就较少。远景镜头表现的事物较多,边缘特征表现较为明显。求取边缘的算法可参考文献 [13]。

3 景别分类识别的实验结果分析

3.1 视频镜头测试库的建立

景别分类识别需要大量典型的视频训练和测试数据,视频镜头测试库可以为镜头景别的分类提供大量的数据,其质量的好坏直接影响到分类结果。为了表现客观性,依照互联网电影数据库(Internet Movie Database, MDB),实验选取了具有一定代表性的 34 部影片作为视频测试中镜头片段的片源,类型涵盖了动作片 9 部、剧情片 11 部、科幻片 7 部、恐怖片 2 部和喜剧片 5 部,选取镜头片段 5 874 个,其中近景镜头 3 384 个、远景镜头 275 个、中景镜头 2 215 个。

3.2 镜头景别的识别与实验结果

分析在分类模型上选择朴素贝叶斯模型,它有着坚实的数学基础,所需要的参数不多,对缺失的数据不敏感,并且许多文献在解决类似的问题上也采用了这种模型,这样使本文的实验结果更具可比性。

实验将使用包含 3 384 个近景镜头、2 215 个中景镜头和 275 个远景镜头的训练集,对每一个样本提取总数为 6 维的特征向量,特征向量由摄像机运动模式特征、人脸特征、局部运动占有率、镜头间相似度、镜头持续时间和边缘点占有率组成。在训练集和测试集的分配上,使用 10 折(10-fold)交叉验证,即把训练集分成相等的 10 个子集,每一次取 1 个子集作为测试集,另外 9 个子集作为训练集,最后得到 10 次结果的均值作为对算法识别精度的估计。

为了解决远景镜头的数目相对近景镜头和中景镜头来说相对较少的问题,在近景镜头和中景镜头的样本中,选择最为典型的一部分,即减少近景镜头和中景镜头的数量,使它们和远景镜头的数目基本达到平衡。依照以上思路,选取了典型的近景镜头 402 个、远景镜头 275 个和中景镜头 316 个,利用朴素贝叶斯分类模型的识别精度如表 1 所示。混乱矩阵如表 2 所示,混乱矩阵表明了每一类镜头被错分到其他类型的镜头的数目,有比较大的参考价值,混乱矩阵中有下划线部分表明某一类型的镜头被正确识别的数目。

表 1 景别类型识别结果

类型	镜头数	检测数	正确数	正确率/%	查全率/%
近景	402	393	302	76.8	75.1
远景	275	236	220	93.2	80.0
中景	316	361	226	62.6	71.5

表 2 识别结果的混乱矩阵

近景	远景	中景	识别结果
<u>302</u>	8	92	近景
12	<u>220</u>	43	远景
79	8	<u>226</u>	中景

由表 1~2 可见,本文提出的特征容易将远景从其他镜头类型中区别开来,但对中景镜头的识别并不尽如人意。这主要是由于有些中景镜头具有远景的特点,有些却具有近景镜头的特点,所以中景镜头的准确率和查全率比远景和近景镜头低。人脸特征容易把近景和其他景别的镜头区分开来,但近景镜头并不一定都出现人脸画面,所以妨碍了其识别精度的进一步提升。从识别错误的结果上可以看出,在被错误分

类的样本中,近景镜头多被错分到中景,较少被错分到远景。远景镜头多被错分的中景,较少被错分到近景,其中一个原因可能是由于中景与近景、中景与远景之间的边界较为模糊所致。

3.3 与现有其他方法的比较

以下附上文献[5]的识别结果并与本文的结果进行比较,如表 3 所示。本文与文献[5]同样使用朴素贝叶斯分类器和 10 折交叉验证,仅在特征的选择上有所差异,因此两种方法具有一定的可比性。值得说明的是,文献[5]把景别分为大远景、远景、中景、中近景、特写和大特写 6 类,但最后识别时,也整合成 3 类:将大远景和远景合并为远景,近景和中近景合并为中景,特写和大特写合并为近景。从表 3 分析可知,除了中景的准确率低于文献[5]的识别结果外,本文的识别方法都要高很多。

表 3 文献[5]的识别结果与本文识别结果的比较

类型	准确率/%		查全率/%	
	文献[5]方法	本文方法	文献[5]方法	本文方法
近景	39.0	76.8	53.0	75.1
远景	38.0	93.2	89.0	80.0
中景	89.0	62.6	48.0	71.5

4 景别音阶的识别

景别的变化造成了被摄物体在画面中内容的改变,是导演表达其拍摄意图的一个重要手段。不同的景别表现出不同的时空内容,近景表现微观的细节部分,远景体现宏大的整体状况。电影镜头景别对观众情绪的影响如图 8 所示,不同镜头的组合可以带给观众不同的感受。不同镜头景别的组合,形成一种特别的节奏,带给观众的情感更加丰富。例如两个连续的远景镜头接上一个中景镜头,给观众一种舒缓和轻松的氛围,在这种情况下,观众并不关注电影画面,注意力较分散。如果是一个全景连接上一个特写镜头,这样的镜头序列给人以强烈的对比,节奏更加地紧凑和急切,情绪的变化也更加激烈。因此,如何对镜头景别进行搭配和组合,使场景更具层次感,事物的表达更有序,诱发观众深刻的情感,也是导演或者拍摄者的一项很重要的创作活动。

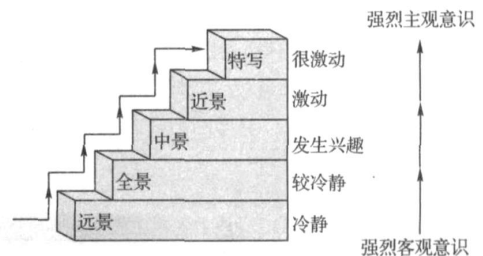


图 8 电影镜头景别对观众情绪的影响

4.1 典型的景别音阶

在参考了电影蒙太奇手法中镜头的组接和蒙太奇句子的有关特点后,定义了以下几种具有特殊意义的镜头景别组合,即景别音阶。

1) 远景 中景 近景:由远及近的方式,使导演要叙述的内容逐步地暴露出来。由于远景镜头的持续时间一般比近景镜头和中景镜头长,这种景别音阶的节奏逐渐加快,使观众的视线逐渐从整体引向局部,注意力也逐渐趋于集中。

2) 近景 中景 远景:由近及远的方式,与之前由远而近的方式相反,这种镜头的组合把观众的注意力从细节引向整体,作用是将节奏降低,把观众的情感从高昂转向低沉。

3) 全景 中景 近景 中景 全景 (或者近景 中景 远景 中景 近景):这是一种景别变化反复循环的组合,这种类型的景别音阶将上述两种镜头组合类型相结合,使观众的情感跌宕起伏。导演常用这种手段控制观众情感的走向,把握场景的节奏。

4) 近景 远景 (或者远景 近景):这是跳跃式的组合,往往令观众产生一个情绪上的突变,就如景别的突变一般。

5) 近景 近景 近景 (远景 远景 远景):该类型的镜头组合中各个镜头的景别保持不变,这样的镜头景别组合具有累积、对比的意思,给予观众一种情绪上的积累。在电影视频中,这种场景多见于对话场景或者是打斗场景。前者令观众的关注度持续保持,后者则使观众的激动程度逐步累加。

6) 景别随意变化的形式:不像上述的几种景别组合那样具有一定的规律。每一个镜头的景别可能是按照导演自身对剧情的理解设置的,没有一定的规律。该种类型极为复杂,且无特定的形式,因此这里不对其进行检测。

4.2 景别音阶的检测

由于输入的数据是视频片段,要对视频片段进行景别的识别,必须先进行镜头分割,然后对每一个视频镜头提取特征,再通过贝叶斯模型进行训练。测试的镜头利用训练后得到的模型参数预测景别,最终在景别识别的基础上实现景别音阶的检测。以下简要介绍下整个景别音阶的检测步骤。

1) 镜头分割。将视频片段按镜头进行分割。

2) 特征提取。对每一个镜头提取特征,包括摄像机运动模式特征、人脸特征、局部运动占有率、镜头间相似度、镜头持续时间和边缘点占有率,总共 6 维特征。

3) 利用贝叶斯模型进行训练。

4) 将需要识别的镜头特征向量,通过训练得到的模型,可以对其景别进行预测。

5) 将一种景别用一个字符代替,一个输入的视频片段最终生成一个字符串序列,在这个序列上查找景别音阶就是一个查找子串的过程,可以利用 KMP 算法查找是否存在给定的子串 (即某些典型的景别音阶) 景别音阶检测原型系统如图 9 所示,其中第 1、4、5、6 个镜头表示识别出的镜头的景别音阶。图中表现的是一个远景 近景的景别音阶和一个近景 近景 近景镜头组成的景别音阶。

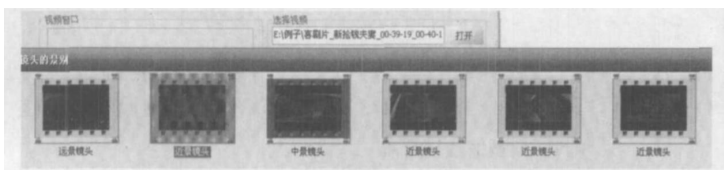


图 9 景别音阶识别与检测原型系统实例

5 结语

本文结合电影视频的领域知识,在镜头景别识别的基础上,通过检测电影视频中的景别音阶来提取电影视频中诱发观众情感变化的片段,这为基于观众情感的视频检索提供了一种切实可行的研究思路。但是仍然存在一些问题:

1) 中景镜头的识别结果不尽如人意,如何解决中景镜头与其他景别之间的模糊边界?

2) 人脸特征,局部运动占有率等特征的计算量相对较大,是否可选取计算量小且有效的特征替代?

3) 训练集中远景镜头相对不足。

今后我们将继续深入研究以下几个方面:

1) 设计更加有效的特征向量,如加入一些音频特征,进一步提高镜头景别的识别率;

2) 不同景别音阶能够在一定程度上诱发观众产生不同的情感变化,但对量化情感变化的强度还需要做进一步的研究工作。

参考文献:

- [1] 陈明, 黄心渊. 电影视听语言中的景别变化 [J]. 电视字幕·特技与动画, 2006, 12(8): 70 - 73.
- [2] XU P, XIE L, CHANG S F, *et al*. Algorithms and system for segmentation and structure analysis in soccer video [C]// Proceedings of IEEE International Conference on Multimedia & Expo. Dhaka: IEEE, 2001: 721 - 724.
- [3] DAHYOT R, REA N, KOKARAM A C. Sport video shot segmentation and classification [J]. Proceedings of the SPIE, 2003, 5150: 404 - 413.
- [4] YUAN YUMING, WAN CHUNRU. The application of edge feature in automatic sports genre classification [C]// Proceedings of IEEE Conference on Cybernetics and Intelligent Systems. Singapore: IEEE, 2004: 1132 - 1135.
- [5] FERRER M Z, BARBERIM, WEDA H. Automatic classification of field of view in video [C]// Proceedings of IEEE International Conference on Multimedia and Expo. Toronto: IEEE, 2006: 1609 - 1612.
- [6] YU XIAODONG, DUAN LINGYU, TIAN QI. Shot classification of sports video based on features in motion vector field [C]// Proceedings of Advances in Multimedia Information Processing, LNCS 2532. Berlin: Springer, 2002: 253 - 260.
- [7] TJONDRONEGORO D W, CHEN Y-P P, PHAM B. Classification of self-consumable highlights for soccer video summaries [C]// Proceedings of IEEE International Conference on Multimedia and Expo. Piscataway: IEEE, 2004: 579 - 582.
- [8] DUAN LINGYU, XU MIN, TIAN QI. Semantic shot classification in sports video [C]// Proceedings of The International Society for Optical Engineering. Santa Clara: SPIE, 2003, 5021: 300 - 313.
- [9] MA YU-FEI, LU LIE, ZHANG HONG-JIANG, *et al*. A user attention model for video summarization [C]// Proceedings of the 10th ACM International Conference on Multimedia. New York: ACM, 2002: 533 - 542.
- [10] 艾海舟, 肖习攀, 徐光祐. 人脸检测与检索 [J]. 计算机学报, 2003, 26(7): 874 - 881.
- [11] XUE XIANG-YANG, ZHU XING-QUAN, XIAO YOU-NENG, *et al*. Using mutual relationship between motion vectors for qualitative camera motion classification in MPEG video [C]// Proceedings of Second International Conference on Image and Graphics [S 1]: SPIE, 2002, 4875: 853 - 860.
- [12] NEVENKA D, JACQUELYN M, LALITHA A. Color superhistograms for video representation [C]// Proceedings of IEEE International Conference on Image Processing. Kobe: IEEE, 1999: 314 - 318.
- [13] 王娜, 李霞. 一种新的改进 Canny 边缘检测算法 [J]. 深圳大学学报, 2005, 4(2): 149 - 152.