

中图分类号: TP391.41 文献标志码: A 文章编号: 1006-8961(2011)06-0995-07

论文索引信息: 倪煜, 赵耀, 朱振峰. 结合标签传递的镜头边界检测与分类 [J]. 中国图象图形学报, 2011, 16(6): 995-1001

结合标签传递的镜头边界检测与分类

倪煜, 赵耀, 朱振峰

(北京交通大学信息科学研究所, 北京 100044)

摘要: 镜头是视频的基本组成单元, 其自动检测与分类是视频分析的重要任务。为了有效利用视频流视觉上的感知特性, 提出一种基于标签传递的镜头边界检测与分类算法。该算法利用半监督学习的标签传递机制, 通过视频流中连续多帧之间的相关性, 将预先构造的初始状态标签通过相关图不断传递, 以揭示不同镜头变化类型的视觉感知特征。然后利用多类 SVM 分类器进行镜头类型分类。实验结果表明, 本文算法能有效识别多种镜头类型, 对视频分析、检索等具有一定实用价值。

关键词: 镜头检测; 标签传递; 镜头分类; 支持向量机

Label propagation for shot boundary detection and classification

Ni Yu, Zhao Yao, Zhu Zhenfeng

(Institute of Information Science, Beijing Jiaotong University, Beijing 100044 China)

Abstract: As a fundamental unit in video analysis, automatic shot detection and classification plays a significant role. To keep consistent with the characteristics of human visual perception, the semi-supervised label propagation based shot boundary detection and classification technique is proposed in this paper. Taking the correlations among consecutive frames in video stream into consideration, the pre-constructed initial state of label for each shot category is propagated continuously via correlation graph, of which the final convergent state can be exploited to reveal the intrinsic description of various shot categories. Furthermore, we apply a multi-class SVM to fulfill the shot classification. The experimental results show the effectiveness of the proposed algorithm, from which the performance of video analysis and retrieval can be expected to benefit.

Keywords: shot detection; label propagation; shot classification; support vector machine

0 引言

随着多媒体技术的发展, 对海量视频数据进行检索已成为目前的一个重要研究课题。镜头检测与分类是视频检索技术中一项基本的、重要的部分, 它是整个视频检索中需要解决的首要问题, 其检测、分类效果的好坏直接影响视频检索的性能。

视频流通常被划分为视频、场景、镜头、关键帧。镜头是指摄像机从打开到关闭的过程中记录下的一组连续图像帧。目前视频一般都经过一定的效果处理, 因此我们所指的镜头也包括由于视频编辑而产生的类似摄像机开关的效果下产生的视频序列。视频检索的首要任务是区分不同的镜头, 即找出视频镜头边界, 并能识别出不同镜头边界类型。镜头边界主要分为切边 (cut) 和渐变, 渐变又包括消隐

收稿日期: 2009-07-21; 修回日期: 2010-05-27

基金项目: 国家杰出青年科学基金项目 (61025013); 中新联合研究计划项目 (2010DFA11010); 中央高校基本科研业务费专项资金项目 (2009JBZ006); 北京市自然科学基金项目 (4112043)。

第一作者简介: 倪煜 (1980—), 男。北京交通大学信息科学研究所人机交互工程专业硕士研究生, 主要研究方向为基于内容的视频检索。E-mail: nu007a@126.com。

(dissolve)、划变(wipe)、淡入淡出(fade in/out)^[1]。在视频编辑中,往往运用多种变换过程将不同镜头结合起来,以表达一定语义或视频特点。因此通过镜头变化方式可获取一定的视频高层语义。

根据视频数据流存在方式,镜头边界检测算法可分为面向非压缩域和面向压缩域的方法。面向非压缩域^[2]主要有基于像素的算法、基于直方图的算法、基于块直方图的算法、基于运动矢量的算法、基于边缘变化率的算法,这类算法需要设定一定阈值。面向压缩域^[2]的算法主要有基于 DCT 系数的算法、基于直流(DC)系数的算法等。但这些算法未能利用视频在时间上的连续特性,往往只通过对临近两帧之间的视觉底层特征做比较,未能考虑连续多帧的相互关系,无法体现镜头变化在视觉上的直观特性。

提出一种利用标签传递算法对镜头连续帧进行时-空分析,从而挖掘镜头视觉本征特性。本质上,其属于基于图理论(graph-based)的方法;图理论技术主要研究节点和边组成的图形的数学理论和方法,目前在视频镜头分割、场景分析中得到广泛的应用。Zhao 等人^[3]基于 Normalized cut (Ncut) 技术实现了视频场景的分析与分类;Yang 等人^[4-5]曾提出基于非监督模式的拉普拉斯特征映射(Laplacian eigenmap),来获取镜头流形特征。本文算法与 Yang 等人方法类似,均属于基于图的方法。但本文是通过一种监督方式,根据人们对镜头变换方式的视觉直观感知,针对不同镜头变换设置不同初始标签。通过帧间关系图逐步传递,其最终状态则揭示了镜头的视觉感知特征。将标签传递的算法引入视频的特征提取中,这种特征具有很好的稳定性,通过判断是否符合某种预设镜头类型,从而实现镜头边界检测与分类。

1 标签传递算法

标签传递(LP)算法^[6-9]属于一种基于图的半监督学习算法。其利用样本间关系建立关系图模型。在这个图模型中,各个已标记样本和未标记样本均被表示成图中的一个节点,样本点之间的距离为图上各边权值。特征空间上距离相近的两个样本点具有相似标记的概率较高。这种学习问题通常称做半监督学习。半监督学习的关键是两种一致假设前提:1) 临近的样本点往往拥有相同的标签;2) 相同

的流形结构上的点往往拥有相同的标签。

标签传递的目的是用已标记节点的标签信息去预测未标记节点的标签信息。采用文献[7]中提出的标签传递技术,来构造镜头变化的视觉本征特性。给定一组数据集 $\mathbf{x} = \{x_1, \dots, x_l, x_{l+1}, \dots, x_n\}$ 和一组标签集 $\mathbf{L} = \{1, \dots, c\}$, $\{y_1, \dots, y_l\} \in \mathbf{L}$ 为前 l 个数据点标签,其余为未标记的数据点。标签传递(LP)算法^[6]过程如下(详细论述参见文献[7]):

1) 计算相似性邻接矩阵

$$W_{ij} = \begin{cases} \exp\left(-\frac{d(x_i, x_j)}{2\sigma^2}\right) & i \neq j \\ 0 & i = j \end{cases} \quad (1)$$

式中 $d(x_i, x_j)$ 为某种距离度量。

2) 归一化 \mathbf{W} 得到相似性矩阵

$$\mathbf{S} = \mathbf{D}^{-1/2} \mathbf{W} \mathbf{D}^{-1/2} \quad (2)$$

式中, \mathbf{D} 为对角矩阵,且有 $D_{ii} = \sum_j W_{ij}$ 。

3) 迭代计算 $F(t+1) = \alpha \mathbf{S} F(t) + (1-\alpha) Y$ 直到收敛。其最终收敛状态为

$$F^* = (1-\alpha) (\mathbf{I} - \alpha \mathbf{S})^{-1} Y$$

式中,参数 $\alpha \in [0, 1]$, F 是 t 时刻的标签状态, Y 为标签初始值 $F(0) = Y$ 。

图 1 为 two moons 合成数据在两个标记样本的情况下,经过多轮迭代后的标签传递结果^[7]。

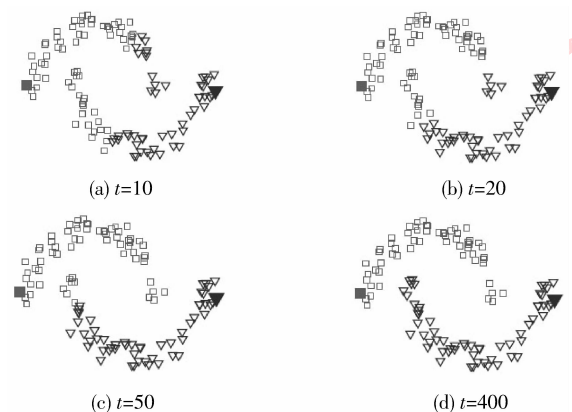


图 1 Two moon 合成数据分类结果

Fig. 1 Classification results on two moon synthetic data

2 镜头边界本征特征提取

视频流可以看做一个高维欧氏空间 \mathbf{R}^{N+1} , N 表示每帧像素个数,另一维是时间。视频中的每帧可以看做该空间中的一点,镜头边界中的每帧对应的

高维空间点有着不同的相关性。从视觉上观察,镜头的这种相关性呈现某种变化趋势。如果能获取这种视觉本征特征,则可以用低维特征表示高维视频。本文算法采用标签传递方法挖掘镜头在视觉上的感知特性,用低维特征表示镜头内在的视觉属性,其算法流程如图 2 所示。

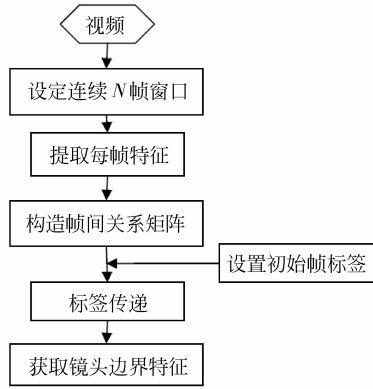


图 2 镜头边界本征特征提取算法流程图

Fig. 2 Flow chart of LDP algorithm

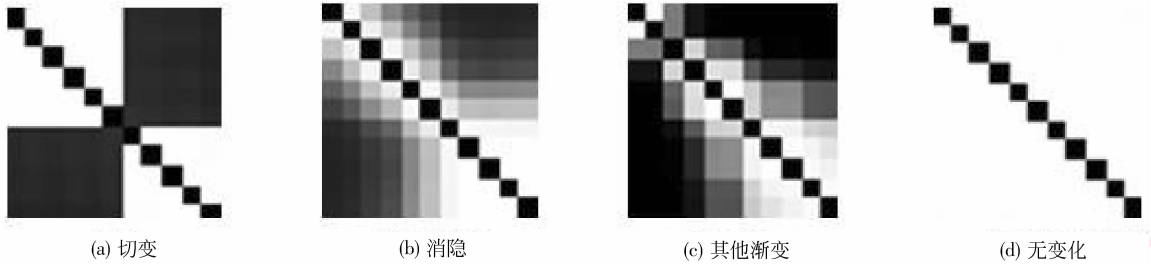


图 3 帧间关系图谱

Fig. 3 Correlation graph between frames

2.2 获取视觉感知特征

根据镜头在视觉上给人的直观感知,对窗口中心位置两侧的帧赋以不同的标签值。在镜头分割点两侧的视频帧所包含的内容往往属于不同的类别,因而该标签值具有所属类别的概率特性。式(3)为标签初始值的通用表达式。

$$Y = \{L_1, \dots, L_{(N-1)/2}, L_k, L_{(N+1)/2}, \dots, L_N\} \quad (3)$$

式中 $k = ((N-1)/2) + 1$, L_i 表示标签值,取 $-1 \sim 1$ 的值。如果通过相关矩阵传递后的标签仍然能保持有其初始特性,说明该边界的视觉特性和初始假设的情况是一致的。因此本文针对上述几种镜头变化种类分别构造 4 种初始化标签值,具体设置值如式(4)。其中, $i = 1, 2, 3, 4$ 分别表示切变、消隐渐变、

2.1 构造帧间关系图 W

首先提取视频底层特征。本文算法核心并非在图像特征提取上,因此选用较简单的 HSV 颜色空间下色调 H 和饱和度 S 的联合颜色直方图表示帧的特征向量 x_i 。其中 H 量化为 $1 \sim 30$, S 量化为 $1 \sim 32$ 。此外,忽略亮度分量 V ,避免因亮度而引起的变化。

然后,以某帧为中心位置,选择时间轴上临近的 N 帧,形成一个长为 N 的窗口,将窗口中的每一帧当做一个节点,应用式(1)构造一个 $N \times N$ 的帧间关系图 W 来表示连续 N 帧的帧间相似度。 $d(x_i, x_j)$ 表示帧间特征向量距离,其中, x_i, x_j 分别表示窗口中的第 i, j 帧的底层特征向量。鉴于卡方距离良好的度量特性^[10],本文中采用卡方距离 $d(x_i, x_j) = \sum_k 2(x_i^k - x_j^k)^2 / (x_i^k + x_j^k)$ 。图 3 显示了 $N = 11$ (综合考虑复杂度,效率)时不同镜头类型的帧间关系图,其中颜色越浅表示相似性越高。最后利用式(2)对矩阵 W 进行归一化处理,得到矩阵 S 。

其他渐变和无镜头变化。对于切变类型,

$$Y_i = \begin{cases} \{1, 1, 1, 1, 1, 0, -1, -1, -1, -1, -1\} & i = 1 \\ \{1, 1, 0.5, 0.5, 0.5, 0, -0.5, -0.5, -1, -1\} & i = 2 \\ \{1, 1, 0, 0, 0, 0, 0, 0, -1, -1\} & i = 3 \\ \{1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1\} & i = 4 \end{cases} \quad (4)$$

可以假设切变点两侧视频帧内容完全属于不同类,因此将其设置为 Y_1 。对于消隐变化,其变化过程给人一种阶梯状渐变过程,因此设其初始标签为 Y_2 。此外视频中包含一些无法确定的渐变类型,将其归结为其他类渐变。由于其变化过程较为复杂,对于

中间部分的视频帧无法确定类别,因此初始标签设为 Y_3 。对于无变化镜头,其视觉特性始终保持一

致,因此设为 Y_4 。基于感知观察的各种镜头类别的初始标签如图 4 所示。

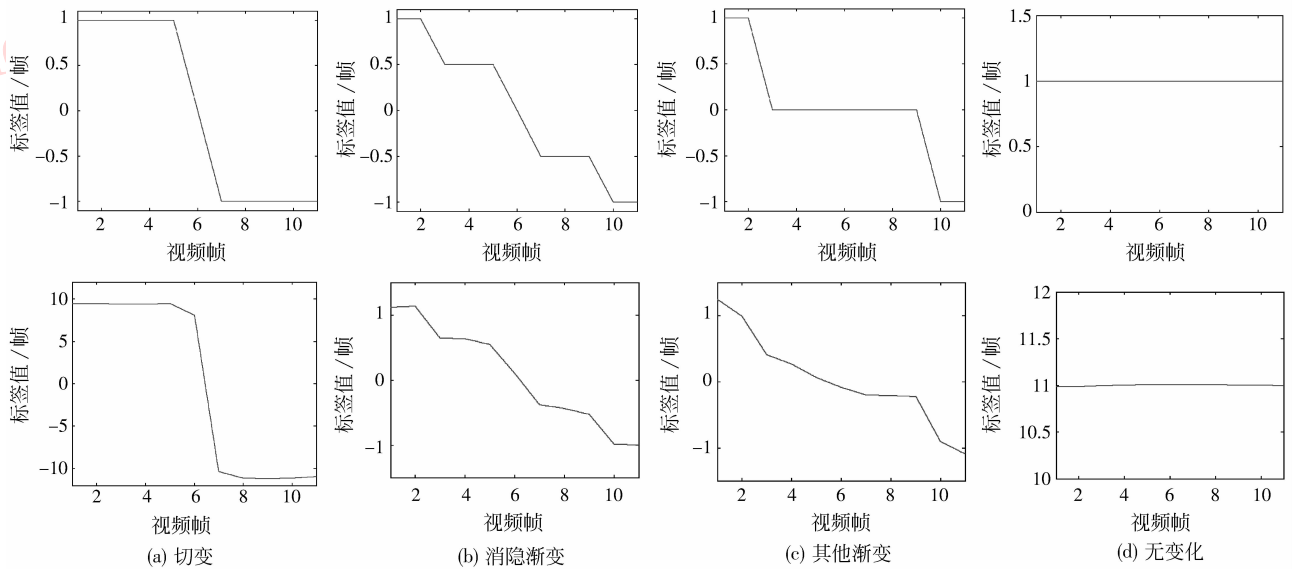


图 4 不同镜头边界经过传递后的内在特性

Fig. 4 Intrinsic characteristics of different shot category via LP

根据 $F^* = (1 - \alpha)(I - \alpha S)^{-1}Y$,用初始标签值 Y 计算出经过帧间相关图 W 传递之后的视频帧标签特性 F 。对于不同的镜头边界, F 呈现出不同的特性,因而具有更好的区分度。如图 4 第 2 行所示为不同镜头类型的初始状态标签经过标签传递后的收敛结果。从图 4 中可以看出,经过标签传递后的镜头标签特征揭示了其边界本质特性,不但具有很好的区分度,而且维数很低。此外,通过实验可以发现,不同的标签初始值,其适应的镜头边界类型不同。图 4 给出了不同类型的初始标签在与之对应类型的镜头变化位置,经过标签传递后,其收敛状态仍然能够被较好地保持初始状态。因此,可以为每种镜头变化类型的收敛状态训练出一种模型,镜头边界通过不同的模型所得预测概率不同,与镜头边界变化一致的模型将获得较大的概率输出。

3 基于 SVM 的镜头特征分类算法

在第 2 节得到的镜头视觉感知特征的基础上,采用多类支持向量机完成不同镜头类型的分类。

在检测模块中,利用典型镜头检测算法颜色直

方图帧差法进行预检测,且对短镜头进行自动合并。通过对视频数据的观察,对于跨度小于 L (本文 $L=6$) 的视频帧具有较强的一致性,所以这种镜头分割点不予计算在内。这种算法漏检率低,而误检率高,可以检测出所有可能的镜头分割点。并且排除了大量非镜头分割点,可以很好地达到预检测效果。

针对预检测得到的候选边界位置按照第 3 节算法提取镜头标签特征值,本文实验中选择连续 $N=11$ 帧作为窗口,该值的选取主要是综合考虑了计算效率与镜头变化感知特征获取的复杂度两方面因素。分别采用式(4)中 4 种初始标签值,针对每种镜头边界特征训练一个 SVM 分类器。

最后,提取测试样本的 4 种边界特征,分别输入与其相应的分类器,并判断该镜头边界是否和初始特征保持一致,即其分类结果和初始标签类型是否一致。针对每一个候选镜头边界,均可得到 4 种结果,分别表示其所属各个类别的概率。最终分类结果为最大概率输出值所对应的类别,分类及训练流程如图 5、图 6 所示。从图 5 可以看出,与文献[4]一致,本文算法采用了一个由粗到精的策略,来实现镜头分割。

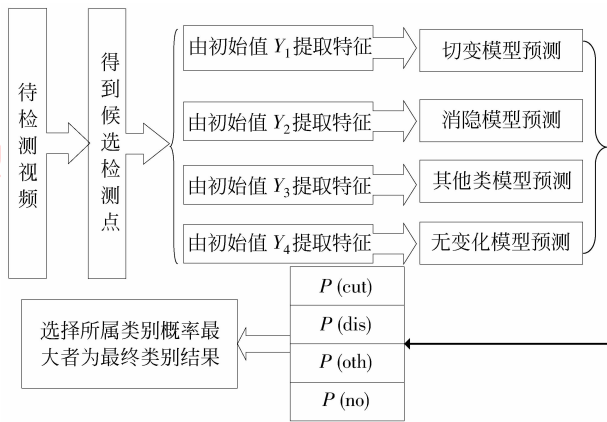


图 5 分类模块框架图

Fig. 5 Framework of classification module

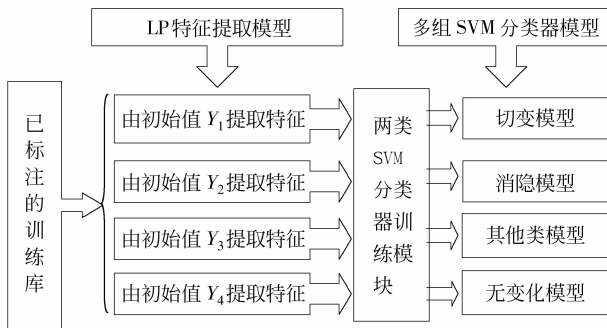


图 6 训练模块框架图

Fig. 6 Framework of training module

4 实验结果与分析

为了验证算法的有效性,我们采用目前国际流行的标准视频库 TRECVID06 的测试集作为实验数据,数据包括 13 段视频,553 678 帧,其中镜头边界有 3 079 个。这些数据涵盖各种类型镜头边界和内容形式,具体数据信息如表 1 所示。由于需要判断无变化类别镜头,因此从视频库中采样 1 000 例镜头内连续 N 帧视频流,作为无变化镜头样本。

表 1 训练与测试数据

Tab. 1 Training and testing dataset

TRECVID 视频库	视频帧数	镜头边界数		
		切变	消隐	其他
训练	451 738	1 074	985	301
测试	101 940	424	142	63

实验使用 OpenCV 开源工具包与 LibSVM 开源工具包,结合 VC++ 6.0 作为开发工具。在训练模块

中,使用 RBF 核函数^[11],并利用交叉验证来确定 SVM 中的 C 和 γ 值。

为评价本文算法,采用常用的性能评价标准查全率、准确率以及 F 评估。

- 1) 查全率 $R = N_c / (N_c + N_m)$;
- 2) 准确率 $P = N_c / (N_c + N_f)$;
- 3) F 评估:

$$F = \frac{(1 + \beta^2) \times R \times P}{\beta^2 \times R + P}$$

式中, N_c 表示正确检测到的镜头边界数; N_m 表示漏检的镜头边界数; N_f 表示误检的镜头边界数; β 表示查全率和准确率之间的相关性,实验均取为 1。

表 2 为 9 组实验的参数设定,表 3 为 Run4 的实验结果。从表 3 中可以看出,切边和消隐渐变的检测效果较好,但其他类渐变的效果较差,主要是不同种类的渐变之间区分度较差,使用一种初始标签较难判断结果。从所做实验的结果数据中看出其他渐变常常会被误分为消隐渐变。主要由于视频帧特征的提取仅使用了颜色特征,较难区分不同种渐变之间的相似性,对镜头边界相关性矩阵的计算会造成误差。而且由于受一些变化过程较长的镜头边界影响,有些镜头的变换跨度长达 30 帧,小范围内很难区分其变化。

此外,还对其他组织的实验结果与本文基于 Label Propagation 的镜头检测算法(LP 算法)进行比较。表 4 和表 5 分别是 CUHK^[12] 和 JDL^[13] 在 TRECVID06 库上的镜头边界检测结果,表 6 为本文算法实验结果。从表 4~6 的对比结果可以看出,本文算法在切变和渐变中的多种指标均取得较理想结果。但渐变的准确率稍低于 JDL 方法,主要由于渐变中包含了多种较难确定类型的渐变过程。如前所述,与文献[4]的非监督方式的镜头变化感知特征获取方法不同,本文采用一种基于标签传递的有监督方式。在感知特征获取的复杂度方面,本文算法涉及 $N \times N$ 逆矩阵的求解,而文献[4]中将涉及 $N \times N$ 矩阵的特征值分解,其复杂度均为 $O(N^3)$ 。本文取 $N = 11$,在效率和复杂度方面效果较优。此外,本文采用一个由粗到精的分层检测策略,在“粗”检测层,采用复杂度低但具有较高虚警率的算法来获得具有镜头变化的候选集,然后在“精”检测层,通过标签传递获得镜头变化的感知特征,来降低虚警率,从而可以从整体保证算法效率和检测性能。

表 2 各组实验所设参数

Tab. 2 Parameter setting for σ and α

Run	Run0	Run1	Run2	Run3	Run4	Run5	Run6	Run7	Run8
σ	0.25	0.50	0.75	0.25	0.50	0.75	0.25	0.50	0.75
α	0.99	0.99	0.99	0.90	0.90	0.90	0.83	0.83	0.83

表 3 镜头分类结果 ($\sigma=0.5, \alpha=0.9$)Tab. 3 Results of shot classification ($\sigma=0.5, \alpha=0.9$)

镜头类型	实际镜头数	正确检出数	漏检数	误检数	检出数	R	P
全部变化	629	550	79	322	872	0.874	0.630
切变	424	399	25	142	541	0.941	0.738
淡入淡出	142	109	33	123	168	0.768	0.649
其他渐变	63	42	21	57	99	0.667	0.424

表 4 CUHK 在 TRECVID06 库上的镜头边界检测结果^[12]Tab. 4 Results of shot boundary detection by CUHK on TRECVID06^[12]

Run	全部变化			切变			渐变		
	R	P	F	R	P	F	R	P	F
tv6cuhk1	0.648	0.700	0.673	0.659	0.767	0.709	0.620	0.560	0.588
tv6cuhk2	0.627	0.684	0.654	0.695	0.686	0.690	0.444	0.677	0.536
tv6cuhk3	0.627	0.715	0.679	0.667	0.778	0.718	0.591	0.572	0.581
tv6cuhk4	0.649	0.659	0.654	0.713	0.657	0.684	0.476	0.668	0.556
tv6cuhk5	0.655	0.693	0.673	0.660	0.760	0.706	0.643	0.557	0.597
tv6cuhk6	0.634	0.678	0.655	0.700	0.677	0.688	0.455	0.679	0.545
tv6cuhk7	0.657	0.709	0.682	0.675	0.772	0.720	0.608	0.569	0.588
tv6cuhk8	0.631	0.694	0.661	0.705	0.697	0.701	0.434	0.677	0.529
tv6cuhk9	0.660	0.685	0.672	0.668	0.755	0.709	0.640	0.544	0.588
tv6cuhk10	0.638	0.672	0.655	0.708	0.671	0.689	0.451	0.676	0.541

表 5 JDL 在 TRECVID06 库上的镜头边界检测结果^[13]Tab. 5 Results of shot boundary detection by JDL on TRECVID06^[13]

Run	全部变化			切变			渐变		
	R	P	F	R	P	F	R	P	F
1	0.626	0.721	0.67	0.619	0.716	0.664	0.646	0.736	0.688
2	0.635	0.718	0.674	0.64	0.709	0.673	0.622	0.744	0.678
3	0.623	0.721	0.668	0.61	0.72	0.660	0.656	0.726	0.689
4	0.628	0.73	0.675	0.613	0.724	0.664	0.67	0.746	0.706
5	0.654	0.711	0.681	0.636	0.703	0.668	0.703	0.729	0.716
6	0.624	0.728	0.672	0.635	0.716	0.673	0.594	0.766	0.669
7	0.624	0.726	0.671	0.621	0.724	0.669	0.631	0.731	0.677
8	0.627	0.715	0.668	0.613	0.719	0.662	0.667	0.705	0.685
9	0.638	0.712	0.673	0.62	0.708	0.661	0.687	0.721	0.704

表6 LP算法在TRCVID06库上的镜头边界检测结果

Tab.6 Results of our proposed LP based shot boundary detection on TRCVID06

Run	全部变化			切变			渐变		
	R	P	F	R	P	F	R	P	F
1	0.828	0.643	0.724	0.941	0.725	0.819	0.595	0.469	0.525
2	0.847	0.652	0.737	0.941	0.694	0.799	0.654	0.554	0.600
3	0.820	0.684	0.746	0.953	0.733	0.829	0.546	0.552	0.550
4	0.839	0.680	0.751	0.943	0.738	0.828	0.624	0.545	0.582
5	0.874	0.631	0.733	0.941	0.738	0.827	0.737	0.566	0.640
6	0.835	0.683	0.751	0.936	0.768	0.844	0.624	0.508	0.560
7	0.835	0.589	0.690	0.948	0.730	0.825	0.600	0.361	0.451
8	0.841	0.659	0.739	0.929	0.752	0.831	0.659	0.484	0.558
9	0.795	0.763	0.779	0.920	0.821	0.868	0.537	0.611	0.571

5 结论

提出一种利用半监督学习的标签传递方法挖掘镜头本征特性的算法,该算法以人的视觉感知为监督条件,通过标签传递获取镜头边界视觉本质特征。同时该方法有很多可扩展性,为该领域的研究提供了一个新颖的思路,实验结果也验证了本文算法的有效性。

参考文献 (References)

- [1] Zhang Yujin. Contented-Based Visual Information Retrieval [M]. Beijing: Science Press, 2003: 50-210. [章毓晋. 基于内容的视觉信息检索[M]. 北京: 科学出版社, 2003: 50-210.]
- [2] Liu Zhengkai, Tang Xiaoou. A survey of methods for partitioning video into shots in video indexing [J]. Computer Engineering and Applications, 2002, 38(23): 84-87. [刘政凯, 汤晓鸥. 视频检索中镜头分割方法综述[J]. 计算机工程与应用, 2002, 38(23): 84-87.]
- [3] Zhao Yanjun, Wang Tao, Wang Peng. Scene segmentation and categorization using ncuts [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR2007). Minneapolis, Minnesota, USA: IEEE Computer Society, 2007: 1-7.
- [4] Yang Linjun, Lu Hong, Wang Bei, et al. Shot boundary classification by temporal pattern discovery from Laplacian eigenmap [J]. IEEE Electronics Letters, 2005, 41(17): 958-960.
- [5] Wang Bei, Yang Linjun, Lu Hong, et al. Post-refinement of shot boundary detection based on manifold feature [J]. Journal of Computer Research and Development, 2006, 43(11): 1993-1998. [王贝, 杨林军, 路红, 等. 基于流形特征的镜头边界检测后处理算法[J]. 计算机研究与发展, 2006, 43(11): 1993-1998.]
- [6] Zhu Xiaojin, Ghahramani Zoubin. Learning From Labeled and Unlabeled Data With Label Propagation: Technical Report CMU-CALD-02-107 [R]. Pittsburghers, USA: Carnegie Mellon University, 2002.
- [7] Zhou Dengyong, Bousquet O, Lal T N, et al. Learning with local and global consistency [C]//Advances in Neural Information Processing Systems 16. Cambridge, MA, USA: MIT Press, 2003: 321-328.
- [8] Wang Fei, Zhang Changshui. Label propagation through linear neighborhoods [J]. IEEE Transactions on Knowledge and Data Engineering, 2008, 20(1): 55-67.
- [9] Blum A, Lafferty J, Rwebangira M R, et al. Semi-supervised learning using randomized mincuts [C]//Proceedings of the Twenty-First International Conference on Machine Learning. Banff, Alberta, Canada: ACM, 2004: 13-21.
- [10] Rubner Y, Tomasi C, Guibas L J. The earth mover's distance as a metric for image retrieval [J]. Int. J. Computer Vision, 2000, 40(2): 99-121.
- [11] Bian Zhaoqi, Zhang Xuegong. Pattern Recognition [M]. 2nd ed. Beijing: Tsinghua Press, 1999: 284-304. [边肇祺, 张学工. 模式识别[M]. 2版. 北京: 清华大学出版社, 1999: 284-304.]
- [12] Hoi SCH, Wong LLS, Lyu A. Chinese University of Hong Kong at TRCVID 2006: Shot Boundary Detection [R/OL]. [2009-07-03]. http://www-nlpir.nist.gov/projects/tpubs/tv6.papers/chinese_uhk.pdf.
- [13] Liu Chunxi, Liu Huiying, Jiang Shuqiang. JDL at TRCVID 2006 Shot Boundary Detection [R/OL]. [2009-07-03]. http://www-nlpir.nist.gov/projects/tpubs/tv6.papers/cas_jdl.pdf.