

基于局部线性嵌入的视频镜头变换检测*

李家兵

(六安职业技术学院 信息工程系, 安徽 六安 237158)

摘要:介绍了一种新的视频镜头变换类型检测方法,利用局部线性嵌入算法提取镜头变换的一维流形特征,用K近邻分类器确定渐变镜头变换的类型。实验表明,本算法镜头变换类型的识别率为97.5%以上。

关键词:镜头边界检测;渐变;流形学习;局部线性嵌入

中图分类号:TP752

文献标识码:A

Shot transition detection based on locally linear embedding

LI Jia Bing

(Department of Information Engineering, Luan Vocational and Technical College, Luan 237158, China)

Abstract: This paper presents a novel method of shot transition detection based on locally linear embedding. Firstly, it extracts the manifold feature of shot transition using locally linear embedding, and then classified it by KNN. Experiments show that the recognition rate of shot transition are reached over 97.5%.

Key words: shot boundary detection; gradual transition; manifold learning; LLE(locally linear embedding)

随着日益丰富的视频数据的出现,人们迫切期望能在大量视频数据中方便、快速地检索到自己感兴趣的视频段。而要实现视频检索,通常是先将视频分割成独立的视频镜头。镜头是指一个摄像机镜头连续拍摄的一组内在相关的连续帧,用来表现在时空上连续的一组运动,是视频检索的基本单元。镜头之间的变换通常分为切变(Cut)和渐变(Gradual)两大类,其中渐变镜头变换又可分为溶解(Dissolve)、扫换(Wipe)、淡入(Fade in)与淡出(Fade out)^[1-3]等。检测出不同类型的镜头变换是对视频高层内容分析、分类、检索和其他应用的基础,因此,对镜头类型变换检测准确性的高低将直接影响到后续处理效果。

视频镜头类型检测算法,从对视频研究时起就得到广泛的重视,主要表现在阈值选取^[4]和特征度量^[5]等两个方面。其中阈值选取又分为固定阈值和自适应阈值两种,固定阈值选择常常存在选择阈值过高或过低的问题,造成查全率和误检率之间的矛盾;而自适应阈值,在视频处理过程中它的值要不断调整,这要花费大量时间

用于计算合适的阈值。因此对镜头变换检测算法的研究,特别是渐变镜头变换检测依然是视频分析与处理所研究的难点。

由于视频镜头变换的多样性以及视频帧高维性的特点,很难从高维中找到它们内在的特征,但通过选择恰当的算法在保持高维特征的情况下,把高维降低维,然后通过分析低维中的数据,得出高维视频变换的特征数据。根据流形学习的特性和视频数据的特点,把视频表示成一幅幅图像随着时间变化而变化的一个高维空间,图像之间在时间轴上是局部相关的,图像之间呈现出特有的分布,把这种分布,看成是一个低维的流形嵌入到高维空间中进行处理。

1 基于局部线性嵌入(LLE)的视频镜头变换学习

局部线性嵌入LLE(Locally Linear Embedding)^[6]是流形学习中的一种。流形学习是从观测到的现象中寻找事物的本质,找到产生数据的内在规律。本文用它来提取镜头之间的特征,以期得到对各种镜头分类的效果。

* 基金项目:安徽省教育厅自然科学基金资助重点项目(KJ2007A045,KJ2008A033)

图形图像与多媒体

Image Processing and Multimedia Technology

1.1 基于 LLE 特征提取

LLE 算法是从流形的局部入手,一个局部邻域上的流形可以近似看成是具有线性特征,在这个小的局部邻域上的一个点就可以用它的相邻点在最小二乘意义下最优的线性表示,通过连接的邻域就可以从局部逐步扩展到整体,处理过程如下。

给定视频 Video,把视频看成是由帧序列 $\{f_1, f_2, \dots, f_n\}$ 构成的。

(1)将视频帧序列 $f_i(i=1, \dots, n)$ 转化为 $M \times N(M, N$ 为视频的帧数、列数)的一维向量 F_i ,转换后视频帧变成 n 个 f 向量;

(2)用 LLE 算法提取镜头变换特征。

保存原流形中局部邻域间的相互关系,将高维数据映射到一维全局坐标系中,具体算法如下:

步骤 1:邻域点搜索。计算出每个向量 F_i 的邻域点,通常取欧式距离最小的 K 个点为邻域或者固定半径 ε 的球状邻域。

步骤 2:求重构误差最小。在 F_i 的邻域中,计算能重构每个 F_i 的权值 W_{ij} ,使重构误差最小,即式(1)的值最小。

$$\varepsilon(W) = \sum_i \|F_i - \sum_j W_{ij} F_j\|^2 \quad (1)$$

其中重构权值 W_{ij} 满足 2 个条件:当 F_i 不属于 F_j 的邻域时, $W_{ij}=0$; $\sum_i W_{ij}=1$ 。

步骤 3:映射到低维嵌入空间。这里以代价误差来衡量,对任何一数据点 i , W_{ij} 具有旋转、尺度和全局变换不变性,因此主要是优化低维坐标系下 Y_i ,使式(2)值最小。

$$\varepsilon_n(Y) = \sum_i \|Y_i - \sum_j W_{ij} Y_j\|^2 \quad (2)$$

(3)特征提取

根据上面算法的步骤,把视频中得到的 n 个 f 向量数据作为数据源,用到 LLE 算法中,经过局部嵌入学习,求出一维向量 Y ,根据 Y 向量的值画出如图 1 所示的特征图。

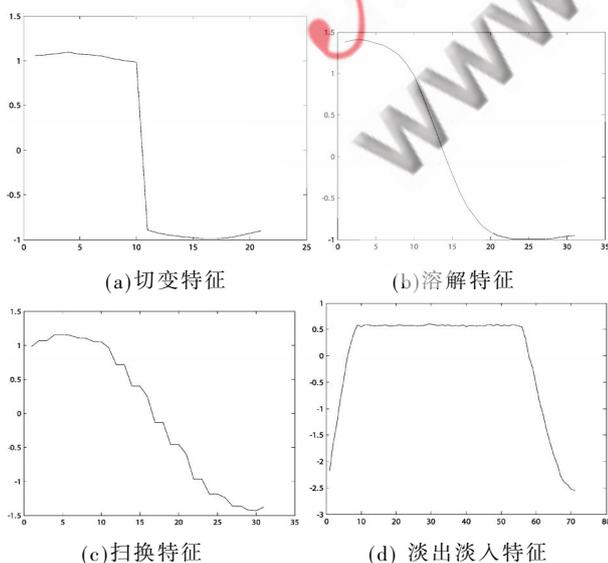


图 1 视频镜头变换特征曲线

《微型机与应用》2010 年第 2 期

因 LLE 对 K 值敏感,不同的 K 值,即使特征值不同也有可能得到相同的特征曲线,图 1 中算法的 K 值都取固定值 15。对于 2 个镜头之间是切变变换,两镜头变换处的 2 帧之间形成一个突然上升或下降的级跃,如图 1(a)所示;对于溶解型镜头变换,特征曲线表现为一个向下或向上的光滑曲线,如图 1(b)所示;对于扫换型镜头变换,特征曲线表现为一个不光滑的锯齿型曲线,如图 1(c)所示;对于淡入淡出型镜头变换,特征曲线成一个“n”形或类似于“n”形曲线,如图 1(d)所示。

为了便于特征的提取并能正确地反映镜头变化的过程,在提取特征时需选择合适的点数(即帧数),本文选择 21 个数据点作为镜头之间变换的过程进行实验。如果镜头变换过程中帧的数目多于 21 帧,则要对其进行抽样,采用等间距地抽取数据点,如淡入淡出镜头变换特征曲线,变换过程中共有 71 帧,如图 2 中的带星号线所示,经过抽样等距点后,变换成图 2 中左边的无星号的特征曲线,仍然保持原来渐变过程变换的特征。

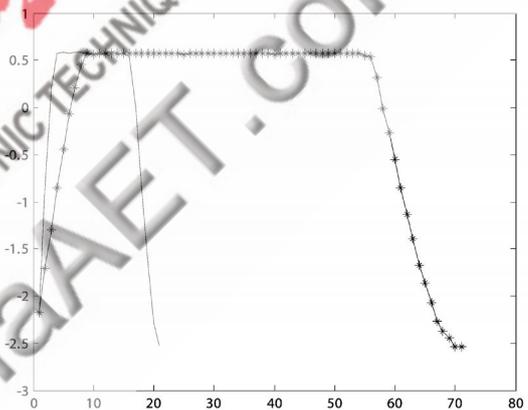


图 2 多帧渐变过程中特征抽样

1.2 K 近邻视频镜头学习

有了一维流形 Y 向量的数据,得到了视频帧序列的一维线性曲线。根据线性曲线的特点,用 K 近邻法对视频序列进行学习,找出镜头变换的类型。算法的流程如图 3 所示。

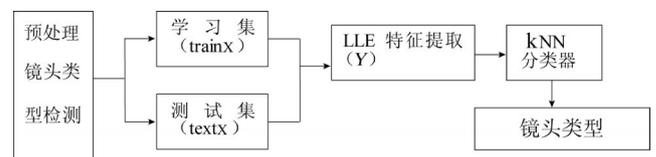


图 3 LLE 算法下视频镜头变换类型处理过程

- (1)预处理镜头类型检测;
- (2)将步骤(1)中处理的数据分为两部分,一部分作为学习集,另一部分作为测试集;
- (3)提取镜头类型一维流形的特征,并根据预处理的镜头类型作相应类型标签;
- (4)通过 K 近邻分类器测试测试集中镜头类型,并分出相应类别的识别率。

欢迎网上投稿 www.pcachina.com 53

图形图像与多媒体

2 实验与分析

2.1 选择数据源

为了验证本文所述方法的有效性,在开放式视频库(www.open-video.org)中下载几段视频,各段视频帧数、镜头类型数和镜头数量如表1所示。

表1 镜头情况

视频文件名	帧数	镜头类型及数量			
		切变	溶解	扫换	淡入淡出
chi97_08_m1	3 174	-	6	-	11
hcil2000_01	2 439	3	-	-	10
hcil2000_07	1 4047	-	72	4	-
BOR10_005	2 609	-	20	2	-
an001	919	18	3	-	1
donghuabaobei	9 360	-	-	14	-
20061118043813	11 047	11	60	3	3
BORSISTER2008	2 969	43	8	14	5
合计	46 564	75	169	37	30

2.2 实验比较与分析

从表1中抽取每种镜头数各30个,共120个,然后从120个镜头变换中随机抽取每种各20个作为学习数据,剩下作为测试数据。用KNN分类器进行分类,验证上述方法的实验结果及其受K值的影响情况。取相应的K值,实验结果数据如表2所示,不同K值实验结果如图4所示。

表2 学习测试数据及不同K值分类的结果

镜头变换类型	学习集镜头数	测试集镜头数	识别率(%)			
			K值			
			K=3	K=5	K=7	K=9
切变	20	10	100	100	100	100
溶解	20	10	90	80	70	60
扫换	20	10	100	90	90	80
淡入淡出	20	10	100	80	80	70
合计/平均	80	40	97.5	87.5	85	77.5

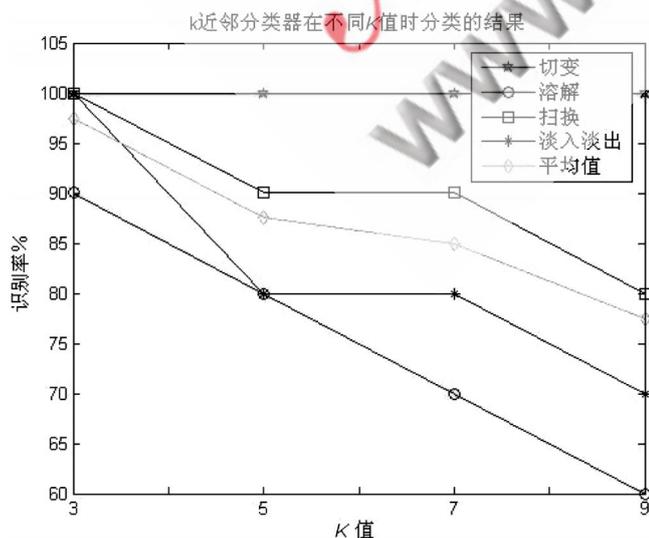


图4 不同K值分类的效果比较图

图4纵坐标数值表示镜头变换类型的识别率,横坐标表示K近邻分类器的不同K值。当K值为3时,除溶解的识别率为90%外,其他三种变换类型的识别率都达到了100%。从图4中可以看出,渐变识别率随着K值的增加而下降。这是因为在提取曲线特征,因帧数过多,间隔提取帧值时,造成渐变镜头特征曲线相互靠近,使渐变镜头类型特征曲线可能互串。从平均识别率看,在K值为3时,识别率达97.5%;在K值为9时,识别率为77.5%。

本文提出了一种基于流形学习和K近邻分类器相结合的镜头类型检测方法。通过流形学习获得了视频帧的原始特征,便于把高维变换到低维来分析、处理。

今后还要做的工作有:针对渐变或渐变中的一种采用流形实现镜头变换的检测;实现一个能自动处理视频镜头分割的系统,为关键帧提取、视频摘要做好坚实的基础。

参考文献

- [1] 黄茜,张海泉,杨文亮,等.基于灰度和直方图的阈值自适应镜头边界检测[J].科学技术与工程,2008,8(14):3787-3792.
- [2] 耿玉亮,须德,冯松鹤.一种快速有效的视频镜头边界检测方法[J].电子学报,2006,34(12):2272-2277.
- [3] 杨碧天,王煦法.一种渐变镜头检测方法[J].计算机工程与应用,2005(20):37-39.
- [4] SZE K W, LAM K M, QIU G. Scene cut detection using the colored pattern appearance model [A].IEEE ICIP-2003 [C]. Barcelona,Spain:IEEE Press,2003(2):1017-1020.
- [5] YOUM S, KIM W. Dynamic threshold method for scene change detection [A].IEEE ICME-2003 [C]. Baltimore, Maryland:IEEE Press,2003:337-340.
- [6] 尹峻松,肖健,周宗潭,等.非线性流形学习方法的分析与应用[J].自然科学进展,2007,17(8):1015-1025.

(收稿日期:2009-10-14)

作者简介:

李家兵,男,1973年生,硕士,讲师,主要研究方向:视频处理,数据挖掘。

(上接第51页)

- [5] 祁同林,左旭东.基于背景重构的运动目标检测算法研究与应用[J].计算机工程与应用,2007,43(16).

(收稿日期:2009-10-21)

作者简介:

魏嘉杰,男,1984年生,硕士研究生,主要研究方向:图像处理;

刘知贵,男,1966年生,教授,博士生导师,主要研究方向:信息工程及控制;

张活力,女,1970年生,助理会计师,主要研究方向:电子商务。