基于窗帧差的镜头边界系数模型及其应用

方 勇, 戚飞虎, 冉 鑫
(上海交通大学计算机科学与工程系, 上海 200030)

摘 要: 针对视频结构分析提出了一种新的镜头边界系数模型. 首先, 按照特定的计算方法计算当前帧的窗帧差, 使得窗帧差在镜头内与镜头边界呈现不同的特征; 然后, 采用相应的镜头边界检测算子计算镜头边界相似系数, 镜头边界相似系数越大, 则该帧越可能是镜头边界; 最后, 为了提高镜头边界相似系数的检测能力, 根据镜头边界相似系数定义了镜头边界系数. 镜头边界系数具有良好的特性, 可以单独检测镜头边界, 也可以与传统的镜头边界检测方法相结合, 为镜头边界检测提供了一种鲁棒的框架. 实验结果表明, 基于镜头边界系数模型的镜头边界检测方法能明显改善镜头边界检测结果.

关键词: 视频分割; 镜头边界检测; 镜头边界系数; 视频检索
中图分类号: TP391 文献标识码: A 文章编号: 0372-2112 (2006) 05-0810-07

Shot Boundary Coefficient Model and Its Application Based on Windowed Frame Differences

FANG Yong, QI Fei-hu, RAN Xin
(Department of Computer Science and Engineering, Shanghai Jiaotong University, Shanghai 200030, China)

Abstract To aim at video structure analysis, a new shot boundary coefficient model is proposed. Firstly, windowed frame differences are calculated using the special approaches, which have different characteristics for intra-shot frames and shot transition frames respectively. Secondly, shot boundary detection operators are used to calculate shot boundary similarity coefficients. If shot boundary similarity coefficients are bigger, corresponding frames are more probably shot boundaries. Thirdly, to improve shot boundary similarity coefficients, shot boundary coefficients are defined based on shot boundary similarity coefficients. Shot boundary coefficients have many good characteristics, and are used to not only detect shot boundaries independently, but also combine traditional shot boundary detection methods for shot boundary detection. Therefore, a robust shot boundary detection framework is provided. Experimental results show that shot boundary detection methods, based on the shot boundary coefficient model, can improve the performance of shot boundary detection.

Keywords video segmentation; shot boundary detection; shot boundary coefficient; video retrieval

1 引言

镜头边界检测, 也称之为镜头分割、镜头变换检测或场景变换检测等, 是基于内容的视频检索中的关键技术之一, 是当前的研究热点. 目前, 大量的镜头边界检测方法已被提出, 并有大量的研究者对这些方法进行了分类、综述与比较[1,2]. 镜头边界检测可以从区域与边界两个方面进行考虑: 基于区域的方法通常使用聚类技术实现, 基于边界的方法通常可以分为两类: 一类是基于相邻帧差的方法, 主要用于切变镜头检测, 也可以用于渐变镜头检测, 根据所使用的特征信息可分别在未压缩视频或压缩视频中进行[3~5]; 另一类是基于模型的方法, 主要用于渐变镜头检测[6,7]. 一般情况下, 视频中的镜头绝大部分为切变镜头, 因此第一类方法受到了更多的重视.

基于相邻帧差的方法一般是根据所使用的特征信息计算相邻两帧的帧差, 然后将帧差跟预设阈值进行比较, 如大于阈值, 则认为当前帧为镜头边界. 但是, 直接使用帧差进行镜头边界检测是不可靠的. 由于摄像条件的变化, 如照明变化、摄像机运动、景物运动和场景变化等, 使得镜头内的帧差与镜头边界的帧差并不必然分开, 有可能是重

收稿日期: 2004-06-07; 修回日期: 2006-01-12
基金项目: 国家自然科学基金 (No 60072029, 60271033)

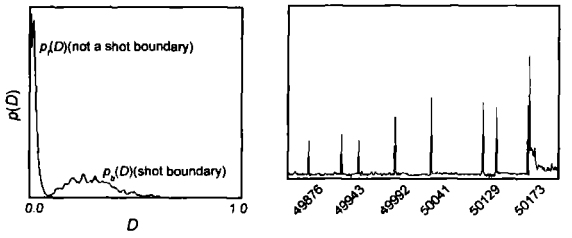
© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

叠的, 因此基于相邻帧差的方法的总体检测性能是有限的, 有待进一步深入的研究. 一些综述文献的评价结果也说明了这一点[1].

为此, 本文提出了一种新的镜头边界系数模型, 用于镜头边界检测, 明显地提高了切变镜头边界检测的能力.

2　镜头边界检测中的基本假设

基于相邻帧差的方法通常有两个假设: 第一, 镜头内帧差的分布与镜头边界帧差的分布是分离的, 没有重叠; 第二, 存在一个常数 T, 使得镜头内帧差小于 T, 而镜头边界帧差大于 T. 如果这两个假设得到满足, 基于相邻帧差的方法就能获得满意的结果. 但是在实际应用中, 这两个假设通常都难以得到满足. 由于摄像条件的变化, 镜头内帧差的分布与镜头边界帧差的分布通常是重叠的, 如图 1所示, 因此不存在假设二所要求的常数. 其次, 即使镜头内帧差的分布与镜头边界帧差的分布是分离的, 对于不同的视频, 假设二所要求的常数通常是不一样的, 难以确定, 限制了基于相邻帧差的方法的实际应用效果. 由于非镜头边界帧概率远大于镜头边界帧概率, 只要两者分布有微小的重叠, 就有可能给镜头边界检测带来严重的不利影响. 因此, 从理论上说, 基于相邻帧差的方法只能在准确率与查全率上进行折衷, 来取得可以接受的结果. 目前, 大部分的基于相邻帧差的方法都采用自适应阈值, 在一定程度上可以改善视频镜头边界检测结果.



(a) 归一化帧差分布图　　(b) 帧差图

图 1　镜头内帧差与镜头边界帧差的分布图

从假设条件的角度出发, 设计良好的镜头边界检测方法有两种途径: 第一, 放宽假设条件; 第二, 减少镜头内帧差的分布与镜头边界帧差的分布之间的重叠范围. 因此, 在此提出本文所要求的基本假设: 设镜头边界帧差为 D_b, 镜头内帧差为 D_f, 在镜头边界的邻域内,

$$D_b \gg D_f \tag{1}$$

基本假设比基于相邻帧差的方法的假设更接近实际情况, 通常情况下, 在镜头边界的邻域内, 镜头内相邻帧变化比较小, 而镜头间相邻帧变化比较大, 因此 $D_b \gg D_f$ 一般能得到满足, 没有假设一严格; 其次, 基本假设是一种相对关系, 不需要确定常数 T, 不存在假设二. 基本假设的有效性在后面的实验中会得到验证.

3　镜头边界系数模型

3.1　窗帧差

目前, 大多数镜头边界检测方法是计算相邻两帧的帧差, 这种帧差计算简单直接, 但是不具有任何特征, 进行镜头边界检测时必需和阈值进行直接比较, 容易受到摄像条件的影响. 为此, 本文提出了一种窗帧差计算方法, 具体来讲有四种实现方式: 0-对称加权窗帧差 (0-SW)、0-滑动平均窗帧差 (0-MA)、1/2-对称加权窗帧差 (1/2-SW) 和 1/2-滑动平均窗帧差 (1/2-MA), 前两种窗帧差在镜头边界处形成双极大值, 而后两种窗帧差在镜头边界处形成单极大值.

3.1.1　0-对称加权窗帧差

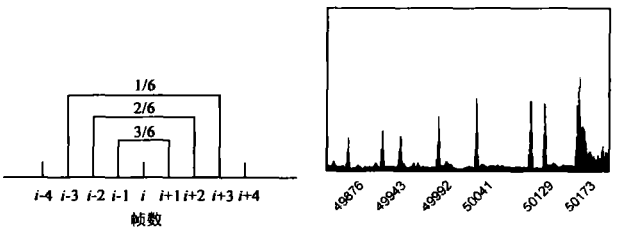
设窗的宽度为 $2N+1$, 则第 i 帧的 0-对称加权窗帧差的计算公式为

$$D_{0SW}(i) = \sum_{j=1}^{N} \frac{N-j+1}{K} D(i-j, i+j) \tag{2}$$

式中 $K = \sum_{k=1}^{N} k$, $D(m, n)$ 为第 m 帧与 n 帧的帧差. $D(m, n)$ 可以采用任何已有的帧差计算方法, 例如直方图交 (Histogram Intersection) 的帧差计算公式为

$$D(m, n) = 1 - \frac{1}{XY} \sum_{l=0}^{L-1} \min(H(m, l), H(n, l)) \tag{3}$$

其中, X、Y 分别为视频帧的水平与垂直像素数, L 为灰度级数, $H(m, l)$ 为第 m 帧直方图.



(a) 计算示意图　　(b) 窗帧差 (柱状图)

图 2　0-对称加权窗帧差示意图 (N=3)

0-对称加权窗帧差的计算示意图如图 2(a)所示, 其特征是计算窗帧差时不包含当前帧. 下面分析 0-对称加权窗帧差的特点, 以 $N=3$ 为例. 一般情况下, 镜头内的场景变化是比较缓慢的, 为了简化分析, 假设在计算窗内镜头内任意两帧之间的帧差相等, 帧差为 D_f; 镜头间的场景变化是剧烈的, 为了简化分析, 假设在计算窗内镜头间任意两帧之间的帧差相等, 帧差为 D_b. 根据式 (2), 在镜头边界邻域 $2N$ 内, 设在第 $i-1$ 帧与第 i 帧之间发生镜头切变, 则 $i-3, i-2, \cdots, i+2$ 连续 6 帧的窗帧差为 $\frac{1}{6}(5D_f + D_b)$, $\frac{1}{6}(3D_f + 3D_b)$, $\frac{1}{6}(6D_b)$, $\frac{1}{6}(6D_b)$, $\frac{1}{6}(3D_f + 3D_b)$, $\frac{1}{6}(5D_f + D_b)$, 根据基本假设式 (1), 则窗帧差可简化为 $\frac{1}{6}(D_b)$, $\frac{1}{6}(3D_b)$, $\frac{1}{6}(6D_b)$, $\frac{1}{6}(3D_b)$, $\frac{1}{6}(D_b)$, 化简取其相对比例关系为 1∶3∶6∶6∶3∶1, 可见在镜头边界处的窗帧差之间的比例关系是确定的, 如图 2(b)所示. 根据这个特点, 可以设计专门的检测算子来检测镜头边界.

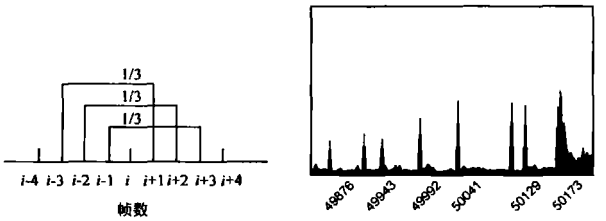
© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

3.1.2 0-滑动平均窗帧差

设窗的宽度为 $2N+1$, 则第 i帧的 0-滑动平均窗帧差的计算公式为

$$D_{0\text{-}MA}(i)=\frac{1}{N}\sum_{j=1}^{N}D(i-(N-j+1),\ i+j) \qquad (4)$$

其中, $D(m,n)$为第 m 帧与 n帧的帧差.



(a) 计算示意图 (b) 窗帧差（柱状图）

图 3 0-滑动平均窗帧差示意图 (N=3)

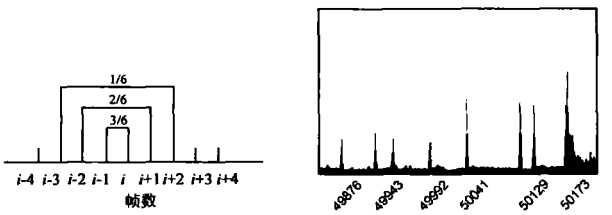
0-滑动平均窗帧差的计算示意图如图 3(a)所示, 其特征也是计算窗帧差时不包含当前帧. 下面分析 0-滑动平均窗帧差的特点, 以 $N=3$为例. 根据式(4), 在镜头边界邻域 $2N$ 内, 设在第 $i-1$帧与第 i帧之间发生镜头切变, 则 $i-3$, $i-2$, …, $i+2$连续 6帧的窗帧差为$\frac{1}{3}(2D_f+D_b)$, $\frac{1}{3}(D_f+2D_b)$, $\frac{1}{3}(3D_b)$, $\frac{1}{3}(3D_b)$, $\frac{1}{3}(D_f+2D_b)$, $\frac{1}{3}(2D_f+D_b)$, 根据基本假设式(1), 则窗帧差可简化为$\frac{1}{3}(D_b)$, $\frac{1}{3}(2D_b)$, $\frac{1}{3}(3D_b)$, $\frac{1}{3}(3D_b)$, $\frac{1}{3}(2D_b)$, $\frac{1}{3}(D_b)$, 化简取其相对比例关系为 1∶2∶3∶3∶2∶1, 可见在镜头边界处的窗帧差之间的比例关系也是确定的, 如图 3(b)所示. 根据这个特点, 可以设计专门的检测算子来检测镜头边界.

3.1.3 1/2-对称加权窗帧差

设窗的宽度为 $2N$, 则第 i帧的 1/2-对称加权窗帧差的计算公式为

$$D_{1/2SA}(i)=\sum_{j=1}^{N}\frac{N-j+1}{K}D(i-j,\ i+j-1) \qquad (5)$$

其中, $K=\sum_{k=1}^{N}k$, $D(m,n)$为第 m 帧与 n帧的帧差.



(a) 计算示意图 (b) 窗帧差（柱状图）

图 4 1/2-对称加权窗帧差示意图 (N=3)

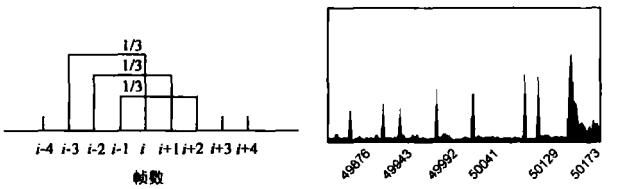
1/2-对称加权窗帧差的计算示意图如图 4(a)所示, 其特征是计算窗帧差时包含当前帧. 下面分析 1/2-对称加权窗帧差的特点, 以 $N=3$为例. 根据式(5), 在镜头边界邻域 $2N-1$内, 设在第 $i-1$帧与第 i帧之间发生镜头切变, 则 $i-2$, $i-1$, …, $i+2$连续 5帧的窗帧差为$\frac{1}{6}(5D_f+D_b)$, $\frac{1}{6}(3D_f+3D_b)$, $\frac{1}{6}(6D_b)$, $\frac{1}{6}(3D_f+3D_b)$, $\frac{1}{6}(6D_f+D_b)$, 根据基本假设式(1), 则窗帧差可简化为$\frac{1}{6}(D_b)$, $\frac{1}{6}(3D_b)$, $\frac{1}{6}(6D_b)$, $\frac{1}{6}(3D_b)$, $\frac{1}{6}(D_b)$, 化简取其相对比例关系为 1∶3∶6∶3∶1, 可见在镜头边界窗帧差之间的比例关系也是确定的, 如图 4(b)所示. 根据这个特点, 可以设计专门的检测算子来检测镜头边界.

3.1.4 1/2-滑动平均窗帧差

设窗的宽度为 $2N$, 则第 i帧的 1/2-滑动平均窗帧差的计算公式为

$$D_{1/2MA}(i)=\frac{1}{N}\sum_{j=1}^{N}D(i-(N-j+1),\ i+j-1) \qquad (6)$$

其中, $D(m,n)$为第 m 帧与 n帧的帧差.



(a) 计算示意图 (b) 窗帧差（柱状图）

图 5 1/2-滑动平均窗帧差示意图 (N=3)

1/2-滑动平均窗帧差的计算示意图如图 5(a)所示, 其特征是计算窗帧差时包含当前帧. 下面分析 0-滑动平均窗帧差的特点, 以 $N=3$为例. 根据式(6), 在镜头边界邻域 $2N-1$内, 设在第 $i-1$帧与第 i帧之间发生镜头切变, 则 $i-2$, $i-1$, …, $i+2$连续 5帧的窗帧差为$\frac{1}{3}(2D_f+D_b)$, $\frac{1}{3}(D_f+2D_b)$, $\frac{1}{3}(3D_b)$, $\frac{1}{3}(D_f+2D_b)$, $\frac{1}{3}(2D_f+D_b)$, 根据基本假设式(1), 窗帧差可简化为$\frac{1}{3}(D_b)$, $\frac{1}{3}(2D_b)$, $\frac{1}{3}(3D_b)$, $\frac{1}{3}(2D_b)$, $\frac{1}{3}(D_b)$, 化简取其相对比例关系为 1∶2∶3∶2∶1, 可见在镜头边界处窗帧差之间的比例关系是确定的, 如图 5(b)所示. 根据这个特点, 可以设计专门的检测算子检测镜头边界.

3.2 镜头边界相似系数与镜头边界系数

针对四种类型的窗帧差, 设计了四种镜头边界检测算子: 0-对称加权检测算子(SW_0算子)、0-滑动平均检测算子(MA_0算子)、1/2-对称加权检测算子($SW_{1/2}$算子)和 1/2-滑动平均检测算子($MA_{1/2}$算子). 镜头边界检测是通过计算视频的窗帧差序列与对应的检测算子的归一化相关系数来实现的.

3.2.1 镜头边界相似系数

定义第 i帧的镜头边界相似系数(BSC, Boundary Sim i-

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

larity Coefficient)为

$$BSC_{0\text{-}SW}(i)=\frac{\sum_{j=-N}^{N-1}(D_{0\text{-}SW}(i+j)\cdot SW_0(N+j))}{\sqrt{\sum_{j=-N}^{N-1}(D_{0\text{-}SW}(i+j))^2}\cdot\sqrt{\sum_{j=-N}^{N-1}(SW_0(N+j))^2}} \tag{7}$$

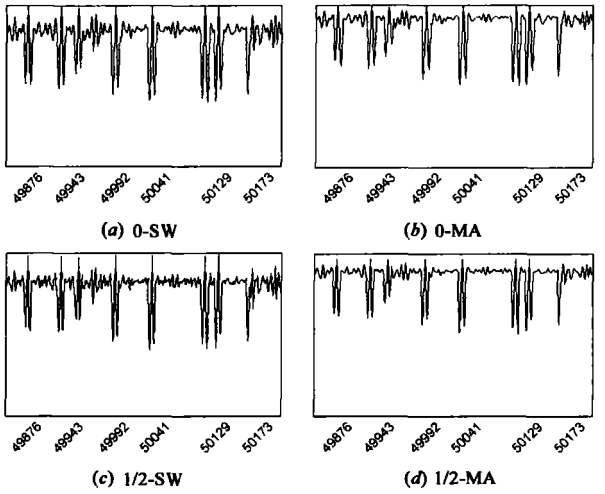
$$BSC_{0\text{-}MA}(i)=\frac{\sum_{j=-N}^{N-1}(D_{0\text{-}MA}(i+j)\cdot MA_0(N+j))}{\sqrt{\sum_{j=-N}^{N-1}(D_{0\text{-}MA}(i+j))^2}\cdot\sqrt{\sum_{j=-N}^{N-1}(MA_0(N+j))^2}} \tag{8}$$

$$BSC_{1/2\text{-}SW}(i)=\frac{\sum_{j=-(N-1)}^{N-1}(D_{1/2\text{-}SW}(i+j)\cdot SW_{1/2}(N+j-1))}{\sqrt{\sum_{j=-(N-1)}^{N-1}(D_{1/2\text{-}SW}(i+j))^2}\cdot\sqrt{\sum_{j=-(N-1)}^{N-1}(SW_{1/2}(N+j-1))^2}} \tag{9}$$

$$BSC_{1/2\text{-}MA}(i)=\frac{\sum_{j=-(N-1)}^{N-1}(D_{1/2\text{-}MA}(i+j)\cdot MA_{1/2}(N+j-1))}{\sqrt{\sum_{j=-(N-1)}^{N-1}(D_{1/2\text{-}MA}(i+j))^2}\cdot\sqrt{\sum_{j=-(N-1)}^{N-1}(MA_{1/2}(N+j-1))^2}} \tag{10}$$

其中, $D(i)$为窗帧差序列, $SW_0(j)$、$MA_0(j)$、$SW_{1/2}$、$MA_{1/2}(j)$分别为 0-对称加权检测算子、0-滑动平均检测算子、1/2-对称加权检测算子、1/2-滑动平均检测算子.

镜头边界检测算子是一维的, 其系数为理想情况下的镜头边界邻域内窗帧差的相对比例关系. 以 $N=3$为例, 对于 SW_0算子, 定义其为一个一维矢量: (1, 3, 6, 6, 3, 1); 对于 MA_0算子, 定义其为 (1, 2, 3, 3, 2, 1); 对于 $SW_{1/2}$算子, 定义其为: (1, 3, 6, 3, 1); 对于 $MA_{1/2}$算子, 定义其为 (1, 2, 3, 2, 1). 因此, 如果第 i帧为切变镜头边界, 根据式 (7)、式 (8)、式 (9)、式 (10), 则其镜头边界相似系数应接近于 1, 如图 6所示.



(a) 0-SW　(b) 0-MA　(c) 1/2-SW　(d) 1/2-MA

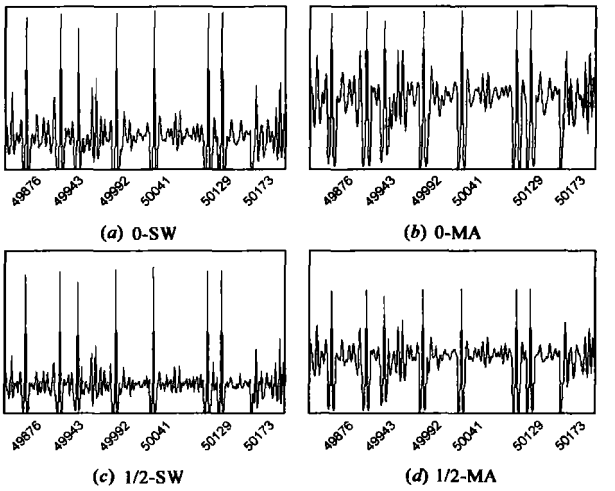
图 6　镜头边界相似系数示意图

3.2.2 镜头边界系数

由图 6可以看出, 虽然镜头边界相似系数在镜头边界处接近于 1, 但整个镜头边界相似系数分布比较集中, 基本上都大于 0.5, 为了增强系数的区别能力, 我们定义第 i帧的镜头边界系数 (BC, Boundary Coefficient)为

$$BC(i)=\exp(-\alpha(1-BSC(i))) \tag{11}$$

其中, $BSC(i)$为镜头边界相似系数, α为常数, 通常为 [5, 15]之间的一个常数, 典型值为 10.



(a) 0-SW　(b) 0-MA　(c) 1/2-SW　(d) 1/2-MA

图 7　镜头边界系数示意图 (α=10)

与图 6对应的镜头边界系数曲线如图 7所示. 由图 7可以看出, 镜头边界系数放大了镜头边界与镜头内帧的差异性, 便于进行镜头边界检测或与其他的镜头边界检测方法相结合. 镜头边界系数具有许多优良的特点: (1), 具有良好的抗噪声帧的能力; (2)对运动不敏感, 具有良好的鲁棒性; (3)不需要特别的帧特征提取, 与传统的帧差计算所需的特征一样; (4)镜头边界系数在不同的镜头内其均值比较一致; (5)只有当前帧的两边邻域帧之间发生明显改变, 且每一边的邻域帧一致时, 镜头边界系数才比较大, 对于理想的切变镜头边界, 其值为 1 这是一个非常好的特点, 是镜头边界检测与关键帧提取的基础. 因此, 镜头边界系数是描述视频帧性质的一个非常优良的参数, 相对于帧差来说, 不易受到摄像条件的影响, 在基于内容的视频信息处理中有着广泛的用途.

4 镜头边界检测

获得镜头边界系数后, 可以单独使用镜头边界系数来检测镜头边界; 也可以与其他的镜头边界检测方法相结合, 提高检测性能.

4.1 使用镜头边界系数的镜头边界检测方法

从图 7可以看出, 镜头边界系数在每个镜头内的均值非常一致, 不会因为不同的镜头而发生明显的改变, 这与

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

一般的相邻帧差很不相同. 基于相邻帧差的方法一般使用自适应阈值能明显地改善检测结果, 而基于镜头边界系数的检测方法不需要自适应阈值, 而采用固定阈值, 该阈值通常设定为一个接近于 1的常数.

基于镜头边界系数的镜头边界检测方法在实际应用中也有一定的缺点, 如果在切变镜头边界邻域内帧与帧之间发生剧烈的不规则变化, 则镜头边界处的镜头边界系数会降低, 而在镜头内也有可能因为视频帧的抖动而使镜头边界系数增加, 这样有可能降低算法的检测性能. 但是, 这种镜头内的视频帧抖动的镜头边界系数比正常的镜头边界处的镜头边界系数要小得多.

4.2 结合相邻帧差的镜头边界检测方法

从上一小节的分析得知, 单独使用镜头边界系数检测镜头边界存在一定的不足, 因此, 为了改善方法的检测性能, 提出了一种鲁棒的镜头边界检测方法, 将镜头边界系数与相邻帧差结合起来, 取得了较好的效果. 结合的方式就是用镜头边界系数去调整相邻帧差.

设第 i帧的帧差为

$$D_{BC}(i) = BC(i) \cdot D(i) \tag{12}$$

其中, $BC(i)$为镜头边界系数, $D(i)$为相邻帧差, 可以为任何一种帧差描述方式, 如直方图交. 获得视频序列的帧差后, 就可使用已有的镜头边界检测方法来检测镜头边界, 如自适应阈值方法等. 新的帧差曲线和帧差分布曲线如图 8、图 9所示. 从图 8帧差曲线图可以看出, 曲线相对于图 1(b)都得到了明显改善, 非镜头边界的帧差得到了明显的减小; 从图 9帧差分布曲线图可以看出, 非镜头边界帧差分布与镜头边界帧差分布的重叠范围大大减小, 非常有利于降低镜头边界误检率和漏检率, 提高总体检测结果.

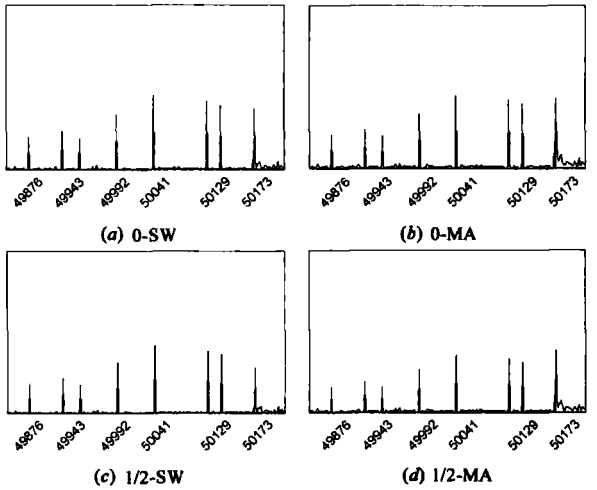


图 8 调整的帧差曲线示意图 (α=10)

4.3 自适应阈值设定方法

基于阈值的帧差镜头边界检测方法使用的都是自适应阈值, 有多种计算方法[3,8], 通常使用如下的阈值设定公式

$$T(i) = \mu(i) + t \cdot \sigma(i) \tag{13}$$

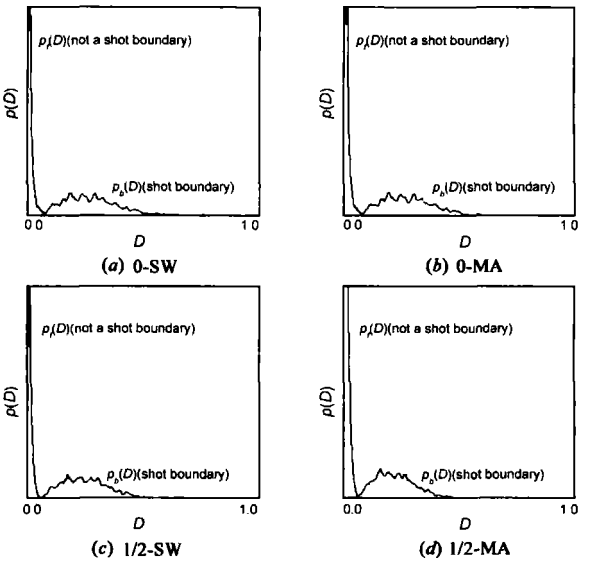


图 9 调整的帧差分布曲线示意图 (α=10)

式 (13)中, $u(i)$为均值, $\sigma(i)$为标准差, t为常数, 通常取值 5或 6均值与标准差的计算方式也有多种, 可参见文献[4].

5 实验结果与讨论

5.1 实验数据集

本实验采用了两个数据集: 电影视频数据集和电视视频数据集. 每个数据集都被转换成为分辨率为 320×240的标准视频. 每个数据集都比较复杂, 由两种不同类型的、摘自于著名电影、电视的真实视频片段组成, 视频中包含了快速运动、战争场面、打斗场面等, 对算法的性能测试比较全面. 各个数据集的参数如表 1所示.

表 1 视频数据集参数

数据集	帧率	帧数	时间	切变镜头数
Movie	23.976	140566	1:37:44	705
Sitcom	25.000	96890	1:04:35	632

每个数据集都借助作者设计的工具软件和人机交互的方式确定标准镜头, 标出镜头边界. 为了克服人的主观因素的影响, 对于有争议的镜头分割, 采用小组表决的方式确定. 根据这种方式产生标准镜头边界数据, 用以评价算法性能.

5.2 评价标准

实验过程中采用了通用的评价标准: 准确率 (Precision)和查全率 (Recall), 其计算公式如下:

$$\text{准确率} = \frac{\text{正确检测数}}{\text{正确检测数} + \text{误检数}}$$

$$\text{查全率} = \frac{\text{正确检测数}}{\text{正确检测数} + \text{漏检数}}$$

$$\text{误检率} = 1 - \text{准确率}$$

$$\text{漏检率} = 1 - \text{查全率}$$

实验过程中, 采用了 ROC曲线作为性能评价方法. 由于基于直方图交和自适应阈值的镜头边界检测方法得到

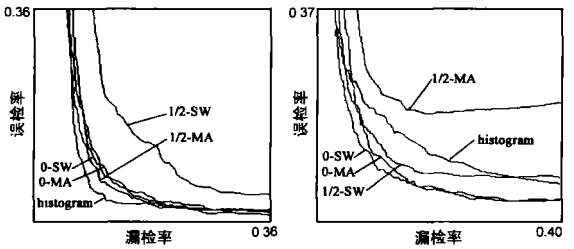
© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

了广泛的应用,并且因为其较好的检测性能[1]和较快的运算速度[2],我们采用其作对比实验,详细地进行了比较. 此外,为了说明镜头边界系数模型作为一种鲁棒的镜头边界检测框架的有效性,还采用了文献[9]中基于信息论的最新镜头边界检测方法做对比. 由于基于信息论的镜头边界检测采用互信息描述帧间变化,相应的镜头边界系数计算时需要调整检测算子,并且调整的互信息为

$$I_{BC}(i) = (1 - BC(i)) \cdot I(i) \tag{14}$$

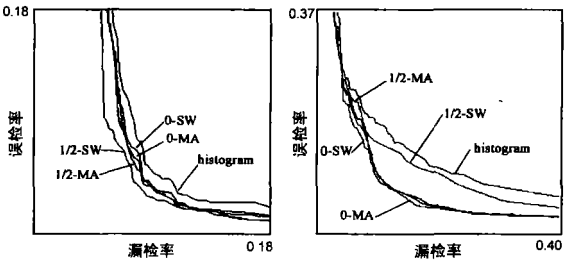
5.3 实验结果

图 10(a)为单独使用镜头边界系数检测电影视频数据集镜头边界的 ROC曲线图,图 10(b)为电视视频数据集 ROC曲线图. 图 11(a)为镜头边界系数调整的帧差并结合自适应阈值检测电影视频数据集镜头边界的 ROC曲线图,图 11(b)为电视视频数据集 ROC曲线图. 表 2、表 3为单独使用镜头边界系数检测镜头边界的结果; 表 4、表 5为使用镜头边界系数调整的帧差并结合自适应阈值进行镜头边界检测的结果. 以上实验中,使用的帧差计算方法为直方图交,计算窗帧差的窗宽为 7或 6,α为 10,计算均值和标准差的窗宽为 25,表 2、表 3中参数 t为最优固定阈值,表 4、表 5中为式(13)中的最优参数. 表 6、表 7为采用文献[9]中的互信息作为帧差计算方法的比较结果.



(a) 电影视频数据集　　(b) 电视视频数据集

图 10　使用镜头边界系数检测镜头边界的 ROC 曲线图



(a) 电影视频数据集　　(b) 电视视频数据集

图 11　使用鲁棒的镜头边界检测方法检测镜头边界的 ROC 曲线图

表 2　使用镜头边界系数检测镜头边界的结果 1(%)

镜头检测方法	电影视频数据集			电视视频数据集		
	t	准确率	查全率	t	准确率	查全率
0-SW	0.888	98.00	73.01	0.908	95.00	78.45
0-MA	0.918	98.00	75.43	0.938	95.00	78.45
1/2-SW	0.888	98.00	33.10	0.947	95.00	42.31
1/2-MA	0.717	98.00	75.71	0.947	95.00	42.31
标准直方图交		98.00	73.86		95.00	45.80

表 3　使用镜头边界系数检测镜头边界的结果 2(%)

镜头检测方法	电影视频数据集			电视视频数据集		
	t	准确率	查全率	t	准确率	查全率
0-SW	0.767	92.33	89.06	0.815	90.97	89.86
0-MA	0.851	93.61	87.78	0.863	88.27	90.65
1/2-SW	0.797	92.47	77.98	0.830	88.75	88.27
1/2-MA	0.656	93.18	89.77	0.737	77.81	89.38
标准直方图交		94.12	91.34		84.63	87.16

表 4　镜头边界系数结合相邻帧差检测镜头边界的结果 1(%)

镜头检测方法	电影视频数据集			电视视频数据集		
	t	准确率	查全率	t	准确率	查全率
0-SW	4.61	95.00	91.48	4.78	95.00	83.52
0-MA	4.45	95.00	91.62	4.74	95.00	81.14
1/2-SW	4.61	95.00	92.33	4.78	95.00	69.57
1/2-MA	4.46	95.00	91.62	4.75	95.00	80.20
标准直方图交	4.27	95.00	90.34	4.79	95.00	45.80

表 5　镜头边界系数结合相邻帧差检测镜头边界的结果 2(%)

镜头检测方法	电影视频数据集			电视视频数据集		
	t	准确率	查全率	t	准确率	查全率
0-SW	4.67	96.45	90.48	4.76	90.65	89.07
0-MA	4.47	95.45	91.48	4.69	91.28	89.22
1/2-SW	4.67	96.73	91.19	4.69	83.52	91.13
1/2-MA	4.51	95.74	91.48	4.71	92.08	88.11
标准直方图交	4.23	94.12	91.34	4.63	84.63	87.16

表 6　镜头边界系数结合相邻互信息检测镜头边界的结果 1(%)

镜头检测方法	电影视频数据集			电视视频数据集		
	ε_c	准确率	查全率	ε_c	准确率	查全率
0-SW	3.11	95.00	98.54	3.21	95.00	96.89
0-MA	2.97	95.00	98.69	3.19	95.00	96.75
1/2-SW	3.01	95.00	98.37	3.18	95.00	95.04
1/2-MA	2.98	95.00	98.16	3.07	95.00	95.21
互信息	3.02	95.00	97.28	3.15	95.00	94.31

表 7　镜头边界系数结合相邻互信息检测镜头边界的结果 2(%)

镜头检测方法	电影视频数据集			电视视频数据集		
	ε_c	准确率	查全率	ε_c	准确率	查全率
0-SW	3.13	97.15	98.02	3.24	96.45	96.57
0-MA	3.08	97.45	98.14	3.27	96.67	96.32
1/2-SW	3.05	96.83	97.59	3.21	96.22	94.98
1/2-MA	3.01	96.64	97.98	3.16	96.08	95.09
互信息	3.05	95.12	97.04	3.11	94.43	95.17

使用镜头边界系数的镜头边界检测方法与本文提出的鲁棒的镜头边界检测方法可以说是一种框架,可与任何一种传统的基于相邻帧差的镜头边界检测方法相结合,提高检测结果. 其次,使用镜头边界系数的镜头边界检测方法与鲁棒的镜头边界检测方法没有增加提取特征信息,只是增加邻域帧的特征信息的差异计算时间,而特征差异计算相对于特征提取时间来说,是微不足道的,如本文采用的直方图交,每帧还是只需计算一次直方图. 从算法的实际运行效率来看,与标准的直方图交方法作比较,没有明显的差别,都可以实现实时镜头边界检测. 当使用互信息

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

作为帧差计算方式时,两种滑动平均镜头边界系数与文献[9]中的计算时间基本一致,每帧只需计算一次联合直方图;而两种对称加权镜头边界系数,每帧需要计算 N 次联合直方图,计算量会增加.因此,从整体性能上看,应该优先使用 0-滑动平均镜头边界系数.

6 结论

本文提出了一种新颖的镜头边界系数模型,并将其应用于视频镜头边界检测.该方法首先计算当前帧的窗帧差,并用相应的检测算子计算镜头边界相似系数,通过镜头边界相似系数可以计算镜头边界系数;通过镜头边界系数,可以采用固定阈值检测镜头边界,也可以与传统的基于相邻帧差的镜头边界检测方法相结合,改善检测结果,提高方法的稳定性和可靠性.

参考文献:

[1] U Gargi, R Kasturi, S H Strayer. Performance characterization of video-shot-change detection methods[J]. IEEE Trans on Circuits and Systems for Video Technology, 2000, 10(1): 1-13.

[2] S Lefevre, J Holler, N Vicent. A review of real-time segmentation of uncompressed video sequences for content-based search and retrieval[J]. Real-Time Imaging, 2003, 9(1): 73-98.

[3] Y Yusoff, W J Christmas, J Kittler. Video shot cut detection using adaptive thresholding[A]. In: Electronic Proceedings of the Eleventh British Machine Vision Conference[C]. Bristol, UK, 2000. http://www.bmva.ac.uk/bmvc/2000/papers/p37.pdf

[4] R Dugad, K Ratakonda, N Ahuja. Robust video shot change detection[A]. In: Proceedings of the IEEE Workshop on Multimedia Signal Processing[C]. Redondo Beach, CA, 1998: 376-381.

[5] H J Zhang, A Kankanhalli, S W Smoliar. Automatic partitioning of full-motion video[J]. Multimedia Systems, 1993, 1(1): 10-28.

[6] P Aigrain, P Joly. Automatic real-time analysis of film editing and transition effects and its application[J]. Computer and Graphics, 1994, (1): 93-103.

[7] A Hampapur, R Jain, T Weymouth. Digital video segmentation[A]. In: Proceedings of the Second ACM International Conference on Multimedia[C]. San Francisco, USA, 1994: 357-364.

[8] 成勇,须德.一种自动选取阈值的视频镜头边界检测算法[J].电子学报, 2004, 32(3): 508-511.
Cheng Yong, Xu De. A method for shot boundary detection using adaptive threshold[J]. Acta Electronica Sinica, 2004, 32(3): 508-511. (in Chinese)

[9] Z Cernekova, I Pitas, C Nikou. Information theory-based shot cut/fade detection and video summarization[J]. IEEE Transaction on Circuits and System for Video Technology, 2006, 16(1): 82-91.

作者简介:

方 勇 男,1974年生于湖南武冈,博士研究生,主要研究方向为图像与视频处理、计算机视觉、信息检索、机器学习、帧率变换和视频增强与恢复等. E-mail: zhyfang@sjtu.org

戚飞虎 男,1938年出生于浙江余姚,教授,博士生导师,主要研究方向为图像处理、视频处理、模式识别和多媒体通信等.

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net