

一种新的视频镜头边界检测 及关键帧提取方法

方 勇

计算机科学与工程系

上海交通大学

2004年11月12日

内 容 提 要

- 基于内容的视频检索简介
- 镜头边界系数模型
- 基于镜头边界系数的镜头边界检测
- 基于镜头边界系数的关键帧提取

1. 基于内容的视频检索简介

基于内容的视频检索：

- 对视频数据中蕴涵的视觉和语义内容进行计算机处理、分析与理解并根据内容进行检索，其本质是对无序的视频数据结构化，提取视觉与语义信息，保证视频内容能被快速检索。

基于内容的视频检索是与以往基于整个视频文件的检索相对应，是基于视频数据局部且与内容相关的检索。

关键问题是视频的组织与索引，包括视频结构分析、视频内容分析、组织与索引。

视频结构分析：

帧—镜头—情节—节目

视频内容分析：

低层内容建模（颜色、纹理、形状、空间关系、运动信息等）

中层内容建模（视频对象）

高层内容建模（语义概念等）

视频索引：

B-树、R-树、LIFT索引结构等，难以处理高维问题

CBVR与WWW:

基于内容的视频检索至少从两个方面与WWW发生紧密联系:

- 第一, 视频信息的获取;
- 第二, 视频数据的传输。

在WWW上实现基于内容的视频检索主要有两种形式:

- 第一, C/S或B/S结构;
- 第二, 视频搜索引擎。

2. 镜头边界系数模型

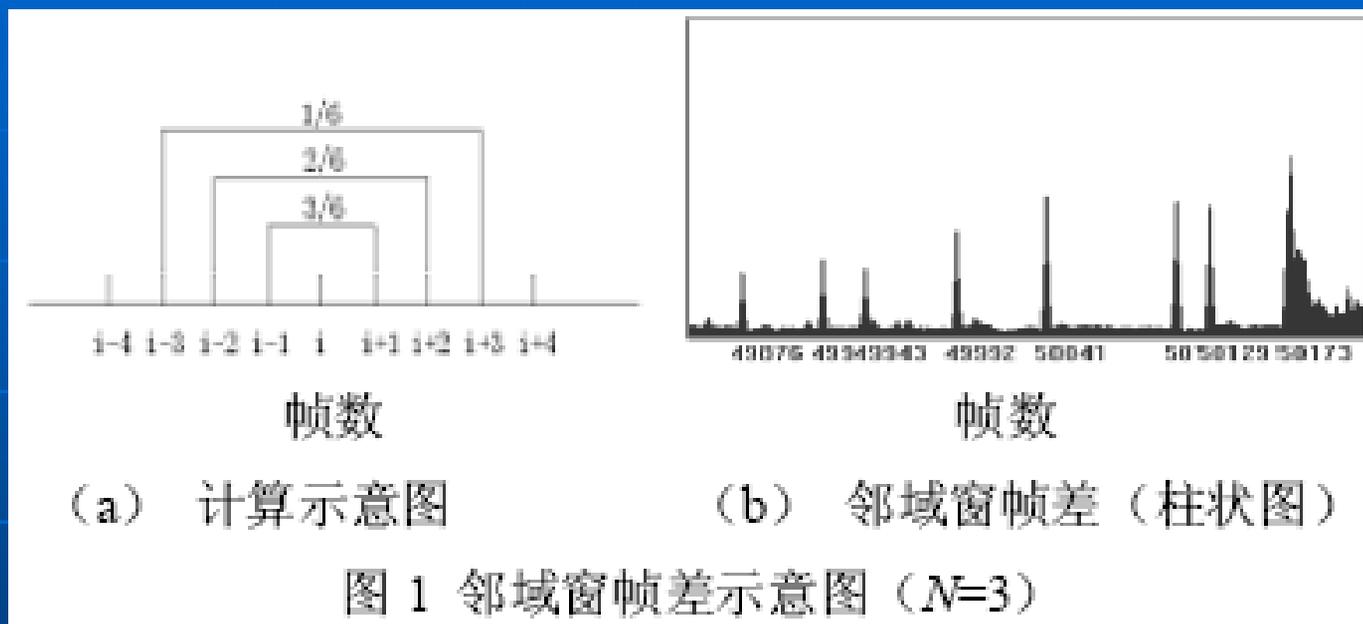
邻域窗帧差:

设邻域窗的宽度为 $2N+1$, 则第 i 帧的邻域窗帧差的计算公式为

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

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

$$D(m, n) = 1 - \frac{1}{XY} \sum_{l=0}^{L-1} \min(H(m, l), H(n, l)) \quad (2)$$



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 $i-3, i-2, \dots, i+2$

连续 6 帧的窗帧差为 $\frac{1}{6}(5D_f + D_d), \frac{1}{6}(3D_f + 3D_d), \frac{1}{6}(6D_d), \frac{1}{6}(6D_d), \frac{1}{6}(3D_f + 3D_d), \frac{1}{6}(5D_f + D_d)$, 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

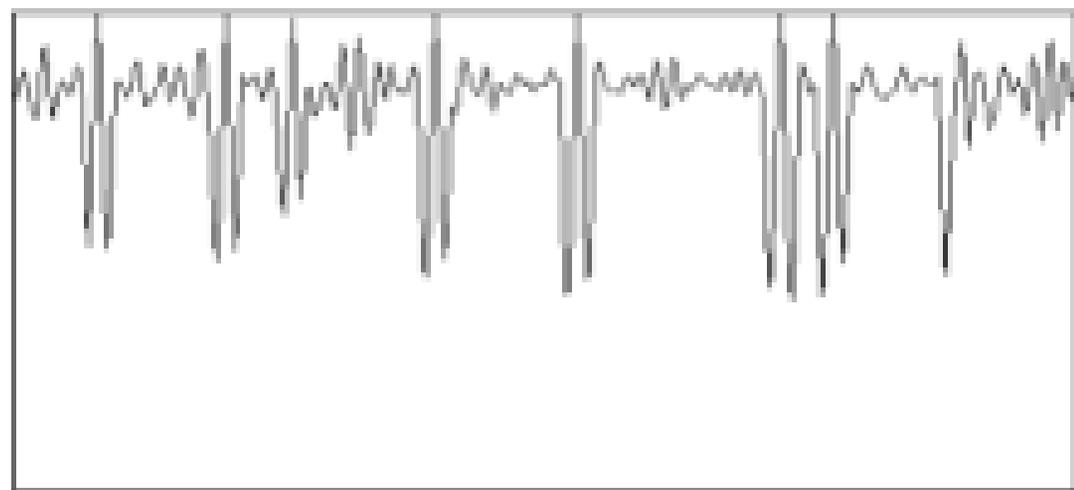
镜头边界相似系数:

定义第 i 帧的镜头边界相似系数 (C_{SBS}) 为

$$C_{SBS}(i) = \frac{\sum_{j=-N}^{N-1} (D_{SW}(i+j) \cdot O_{SW}(N+j))}{\sqrt{\sum_{j=-N}^{N-1} (D_{SW}(i+j))^2} \cdot \sqrt{\sum_{j=-N}^{N-1} (O_{SW}(N+j))^2}} \quad (3)$$

其中, $D_{SW}(i)$ 为邻域窗帧差序列, $O_{SW}(j)$ 为镜头边界检测算子。

对于 O_{SW} 算子, 我们定义其为一个一维向量: (1,3,6,6,3,1)。



49876 49949 49943 49992 50041 50501 50129 50173

帧数+

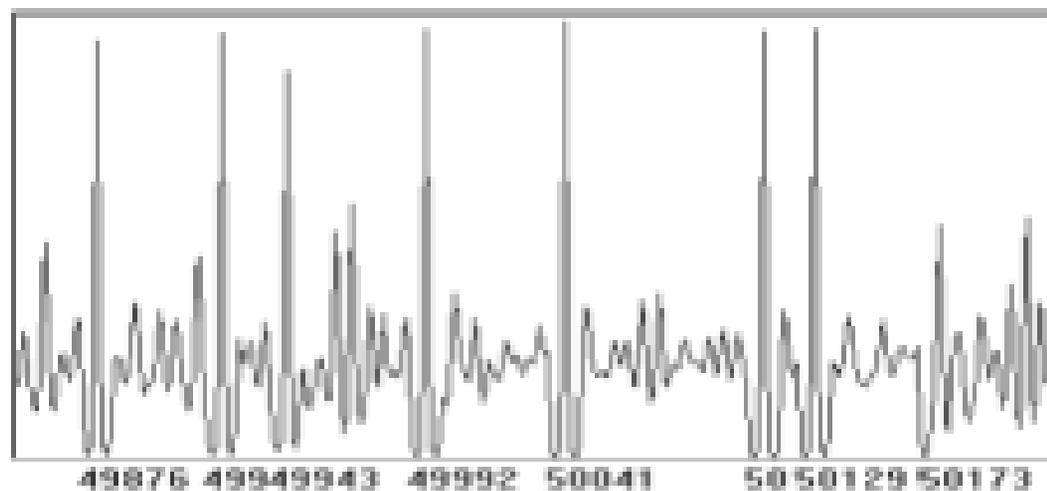
图 2 镜头边界相似系数示意图

镜头边界系数:

定义第 i 帧的镜头边界系数 (C_{SB}) 为

$$C_{SB}(i) = \exp(-\alpha(1 - C_{SBS}(i))) \quad (4)$$

其中, $C_{SBS}(i)$ 为镜头边界相似系数, α 为常数



帧数

图 3 镜头边界系数示意图 ($\alpha = 10$)

镜头边界系数的特点：

- 具有良好的抗噪声能力；
- 对运动不敏感，具有良好的鲁棒性；
- 额外计算量小；
- 两边邻域一致变化时，系数才比较大，理想情况下为1。

3. 基于镜头边界系数的镜头边界检测

大体上可以分为三类：

第一，基于帧差的镜头边界检测方法；

第二，基于模型的镜头边界检测方法；

第三，基于学习的镜头边界检测方法。

镜头边界检测方法:

第一，只使用镜头边界系数的镜头边界检测——固定阈值

第二，结合相邻帧差的镜头边界检测——自适应阈值

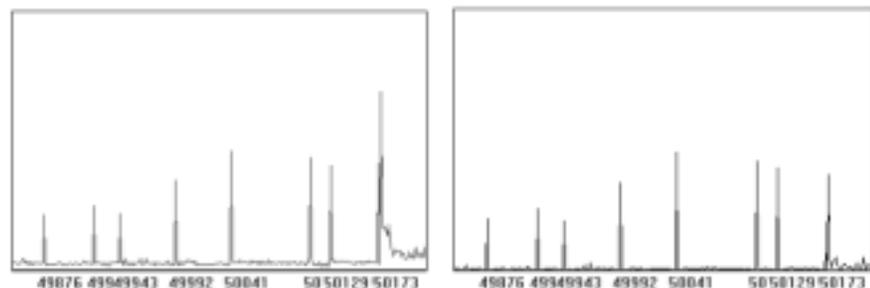
结合相邻帧差的镜头边界检测:

设第 i 帧的镜头边界系数帧差 (D_{SBC}) 为

$$D_{SBC}(i) = C_{SB}(i) \cdot D(i) \quad (5)$$

自适应阈值:

$$T(i) = \mu(i) + t\sigma(i) \quad (6)$$

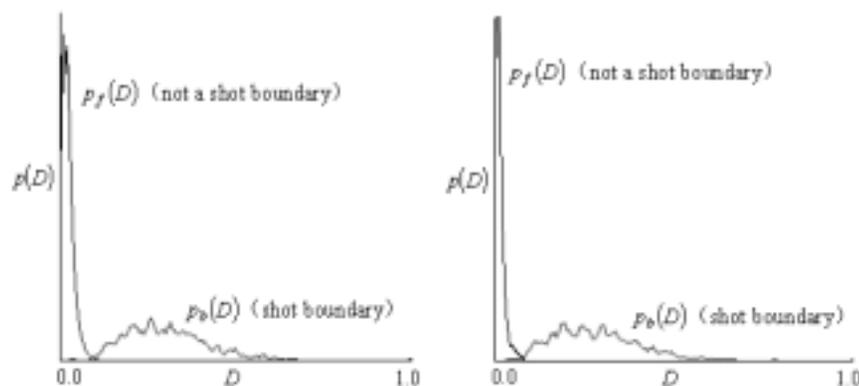


帧数

帧数

(a) 原始帧差曲线 (b) 镜头边界系数帧差曲线

图 4 帧差曲线示意图 ($\alpha = 10$)



(a) 原始帧差分布曲线 (b) 镜头边界系数帧差分布曲线

图 5 帧差分布曲线示意图 ($\alpha = 10$)

实验结果:

实验数据集:

表 1 视频数据集参数

Table 1 Parameters of video data sets

数据集	帧率 (fps)	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

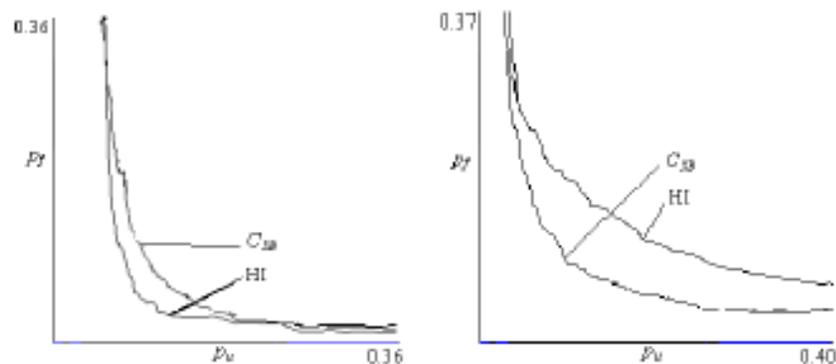
评价标准:

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

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

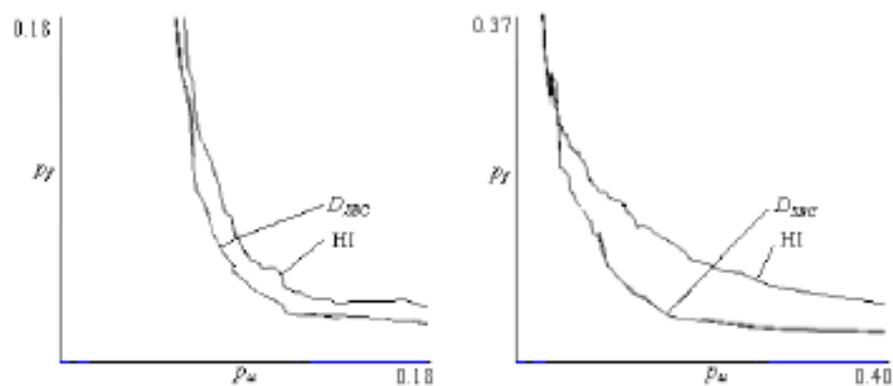
$$\text{误检率} = 1 - \text{准确率};$$

$$\text{漏检率} = 1 - \text{查全率}。$$



(a) 电影视频数据集 (b) 电视视频数据集

图 6 使用镜头边界系数检测镜头边界的 ROC 曲线图
Fig. 6 ROC curves of shot boundary detection using C_{SB}



(a) 电影视频数据集 (b) 电视视频数据集

图 7 使用镜头边界系数帧差检测镜头边界的 ROC 曲线图

表 2 使用镜头边界系数检测镜头边界的结果 1 (%)

Table 2 Experimental result 1 of shot boundary detection using shot boundary coefficients

镜头检测方法	电影视频数据集			电视视频数据集		
	固定阈值	准确率	查全率	固定阈值	准确率	查全率
镜头边界系数法	0.888	98.00	73.01	0.908	95.00	78.45
直方图交法		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 (%)

Table 3 Experimental result 2 of shot boundary detection using shot boundary coefficients

镜头检测方法	电影视频数据集			电视视频数据集		
	固定阈值	准确率	查全率	固定阈值	准确率	查全率
镜头边界系数法	0.767	92.33	89.06	0.815	90.97	89.86
直方图交法		94.12	91.34		84.63	87.16

表 4 镜头边界系数帧差检测镜头边界的结果 1 (%)

Table 4 Experimental result 1 of shot boundary detection using shot-boundary-coefficient frame differences

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
镜头边界系数帧差法	4.61	95.00	91.48	4.78	95.00	83.52
直方图交法	4.27	95.00	90.34	4.79	95.00	45.80

表 5 镜头边界系数帧差检测镜头边界的结果 2 (%)

Table 5 Experimental result 2 of shot boundary detection using shot-boundary-coefficient frame differences

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
镜头边界系数帧差法	4.67	96.45	90.48	4.76	90.65	89.07
直方图交法	4.23	94.12	91.34	4.63	84.63	87.16

4. 基于镜头边界系数的关键帧提取

有两个问题需要解决：

第一，关键帧数目；

第二，关键帧位置。

关键帧提取策略：

第一，设置一个最大关键帧数 M ；

第二，每个镜头的非边界过渡区的第一帧确定为关键帧；

第三，使用非极大值抑制法确定镜头边界系数极大值，并排序，以实现基于镜头边界系数的关键帧提取。



图 8 电影视频数据集关键帧提取示例



图 9 电视视频数据集关键帧提取示例

Thanks !