

视频镜头分割方法综述^①

魏 玮, 刘 静, 王丹丹

(河北工业大学 计算机科学与软件学院, 天津 300130)

摘 要: 视频序列的镜头分割亦称镜头变化检测是视频检索中的关键技术之一. 对五种常用的视频分割算法作了综述, 包括像素法、直方图法、 X^2 直方图法、 X^2 直方图分块法、边缘轮廓变化率法, 并详细介绍了各种算法中的帧差异值的计算以及介绍了他们的优缺点, 并且通过实验分析对各种算法进行了比较, 五种算法的优缺点实验中得到了很好的体现.

关键词: 镜头分割; 镜头检测; 视频检索; 视频分割; 直方图

Survey of Methods for Partitioning Video into Shots in Video

WEI Wei, LIU Jing, WANG Dan-Dan

(School of Computer Science and Engineering, Hebei University of Technology, Tianjin 300130, China)

Abstract: Partitioning a video sequence into shots or detecting shot change is one of the key techniques in video indexing. This article summarizes some commonly used methods for partitioning video into shots which are pixel level-based methods, histogram methods, X^2 histogram methods, X^2 block histogram methods, edge silhouette rate of changing methods, and detailedly introduces various algorithms of frames difference value calculation, and introduces the advantages and disadvantages of them, and the results of these algorithms are compared through the experiments, the advantages and disadvantages of Five algorithm at the experiments are very well reflected.

Key words: shot segmentation; shot detection; video indexing; video segmentation; histogram

面对庞大的视频信息, 如何建立一个快速、简便、精确的视频检索系统成为近年来一个热门课题. 视频镜头分割是视频检索的第一步, 是随后的关键帧提取的基础, 镜头分割的准确性将直接影响到后续处理的效果. 关键帧是用来描述一个镜头内部主要内容的某帧或某几帧图像, 通过镜头分割后对每个镜头提取关键帧就可在此基础上对视频建立索引, 为视频检索提供了快捷简便的手段, 并且极大降低了视频检索系统的处理时间, 使得视频检索系统的实时性得到很大提高. 因此, 视频镜头的分割技术是视频检索的关键技术, 从一开始就得到广泛的研究, 并取得了丰富的研究成果^[1]. 然而直到今天, 仍然没有在各种情况下都能表现出良好性能的算法.

镜头是指由一个摄像机镜头连续拍摄的一组内在相关的连续帧. 镜头的分割是指正确的判断镜头的边界, 一般使用的镜头分割的方法是计算各视频帧之间的差值, 通过比较选择符合条件(一定大)的帧差值, 这样, 确定镜头边界. 由于视频内容的变化、噪声、快速运动以及光线剧烈变化等极端情况都有可能造成镜头的错误分割, 这样阈值的选取成了关键技术. 一般来说, 镜头转变曲线会发生明显变化, 但是由于光照影响可能会误检为切变(图 1 Abrupt Change), 而有些情况下, 例如渐变(淡入、淡出等)曲线发生的变化缓慢而不明显(图 1 Gradual Change). 汇总综述文献[2]及相关文献, 视频检索中镜头分割方法如下.

① 收稿时间:2012-05-30;收到修改稿时间:2012-06-28

1 基于像素的镜头分割算法

基于像素^[3]的镜头分割主要是对视频帧的图像底层处理过程, 包括亮度、灰度或者色彩值, 其计算简单, 原理是计算两帧之间的每个对应的像素的灰度(亮度)的变化。

相邻两帧对应点的灰度(亮度)差为

$$f_d(i, j) = |f_{n+1}(i, j) - f_n(i, j)|$$

式中, $f_n(i, j)$ 、 $f_{n+1}(i, j)$ 分别代表第 n 帧和第 $n+1$ 帧像素 (i, j) 的灰度(亮度), 则相邻两帧之间的总帧差为

$$F_d(f_n, f_{n+1}) = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N f_d(i, j)$$

式中, M 、 N 为图像的尺寸. 若总帧差大于某一设定阈值, 则判断镜头内容发生变化. 该方法原理简单、便于实现. 缺点是对于摄像机及镜头内运动物体、光线条件的剧烈变化敏感, 容易误检。

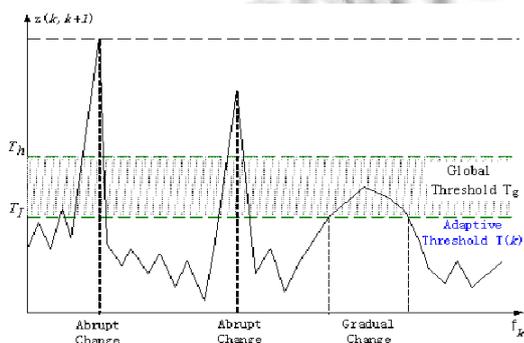


图 1 镜头分割帧差异值的分布和阈值

2 基于直方图的镜头分割算法

基于直方图^[4]的算法是最普遍的分割方法, 它实现简单方便, 而且对于大多数视频能得到较好的效果. 基于直方图的算法是在基于像素的比较上发展来的, 基于直方图的算法通常是将相邻帧之间的各个像素的灰度、亮度等分为 N 个等级, 再针对每个等级统计像素数做成直方图比较, 给出两个图像的直方图, 则直方图帧差计算公式如下所示:

$$D = \frac{1}{2N} \sum_i |h_m(i) - h_n(i)|$$

其中, N 为图像帧像素的总数. $h_m(i)$ 、 $h_n(i)$ 表示的是两个视频帧在 i 这个直方图单位上面的距离。

基于直方图法不考虑像素的位置信息, 而使用其亮度和色彩的统计值, 缺点是对结构不同而直方图却很相近的两帧造成漏检而且对于光线变化比较剧烈的情况下, 帧差值会受到很大的干扰。

3 基于 X^2 直方图的镜头分割算法

X^2 直方图^[5]的算法因能放大最大帧差及算法比较稳定而得到广泛应用, 为了使直方图帧差更好反映两帧间的差别, 可把直方图帧差按下式归一化为:

$$X^2 = \begin{cases} \sum_{i=1}^k \frac{(h_m(i) - h_n(i))^2}{\max(h_m(i), h_n(i))}, & (h_m(i) \neq 0 \vee h_n(i) \neq 0) \\ 0, & \text{else} \end{cases}$$

式中, k 为图像帧像素的总数, $h_m(i)$ 、 $h_n(i)$ 表示的是两个视频帧在 i 这个直方图单位上面的距离. 可见, X^2 值越大, 则两帧间差异越大; 反之则越小. 直方图的归一化也叫直方图均衡化, 是通过使用累积函数对灰度值进行“调整”以实现对比度的增强. 直方图均衡化处理的“中心思想”是把原始图像的灰度直方图从比较集中的某个灰度区间变成在全部灰度范围内的均匀分布. 直方图均衡化^[6]就是对图像进行非线性拉伸, 重新分配图像像素值, 使一定灰度范围内的像素数量大致相同. 简单说就是把给定图像的直方图分布改变成“均匀”分布直方图分布. 其缺点: ① 变换后图像的灰度级减少, 某些细节消失; ② 某些图像, 如直方图有高峰, 经处理后对比度不自然的过分增强. 算法对于摄像机及镜头内物体运动具有良好的容忍程度, 但是实现复杂, 计算复杂度高。

4 基于 X^2 直方图分块的镜头分割算法

X^2 直方图分块法^[7]顾名思义是在 X^2 基础上改进的, 近几年对 X^2 直方图分块的研究比较少, 为了减少运动、光照等引起的帧差值的变化, 本文将各帧分块处理, 比较每个块的直方图, 将差值最大的块剔除, 剩下的块比较帧差异值, 其计算公式如下:

$$X^2 = \begin{cases} \sum_{i=1}^k \sum_{j=1}^p \frac{(h_m(i, j) - h_n(i, j))^2}{\max(h_m(i, j), h_n(i, j))}, & (h_m(i, j) \neq 0 \vee h_n(i, j) \neq 0) \\ 0, & \text{else} \end{cases}$$

式中, k 为图像帧像素的总数, p 为块数, $h_m(i)$ 、 $h_n(i)$ 表示的是两个视频帧在 i 这个直方图单位上面的距离. 缺点是计算时间较 X^2 直方图时间长, 计算复杂度高, 优点对于运动的物体有良好的容忍程度。

5 基于边缘轮廓变化率的镜头分割

此方法^[8]的主要思想是通过计算边界的变化程度来确定镜头的边界. 首先利用 canny 算子将图像边缘化, 然后计算出帧间的总体位移, 以此进行配准, 然

后计算边缘的数量和位置. 帧差由边缘变化的比例表示, 即边缘从一帧到另一帧移进和移出的比例. 但是由于该方法是先边缘化再进行配准最后才进行边缘比较, 因此此方法对于运动时稳健的但计算比较复杂. 设 Q_k 为 k 帧中与 $k+1$ 帧中最近边缘的距离大于给定阈值 T 的边缘像素数目的百分比; 同样设 Q_{k+1} 为 $k+1$ 帧中与 k 帧中最近边缘的距离大于给定阈值 T 的边缘像素数目的百分比, 则帧差为:

$$D = \text{Max}(Q_k, Q_{k+1})$$

6 实验结果与分析

首先, 选择 2 个视频剪辑作为实验数据, 它们的格式都是 avi 格式. 第一个视频时长 18 秒, 共有 465 帧; 另一个时长 17 秒, 共有 448 帧, 播放速度都是 25 帧/秒, 每个视频都有三个人走动, 本文就是将每个人走入镜头和走出镜头的镜头边界检测出来.

本文比较了两个算法的运行速度和准确率, 实验在 WindowXP 系统下利用 matlab 运行五个算法, 对上述的视频段进行镜头检测. 五种算法的运行速度和准确率如表所示.

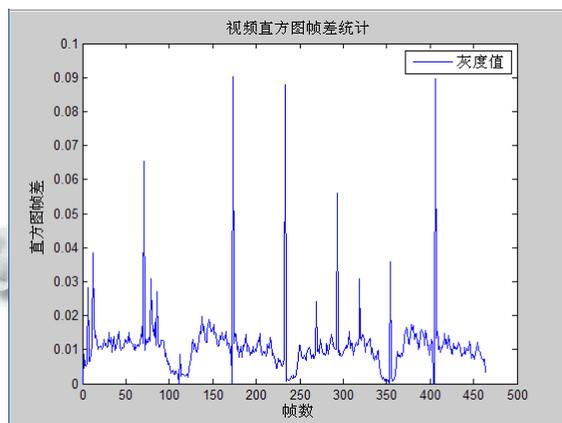
表 1 五种算法性能的比较

视频	方法	实际检测镜头数	正确检测镜头数	计算时间(秒)	准确率
视频 1	像素法	5	4	13.3	80%
	直方图法	12	6	12.1	50%
	X ² 直方图法	5	4	12.5	80%
	X ² 直方图分块法	6	5	15.2	83%
	边缘轮廓变化率法	7	4	55.6	57%
视频 2	像素法	5	3	12.4	60%
	直方图法	8	4	12.8	50%
	X ² 直方图法	6	4	12.4	67%
	X ² 直方图分块法	7	5	14.6	71%
	边缘轮廓变化率法	6	3	55.0	50%

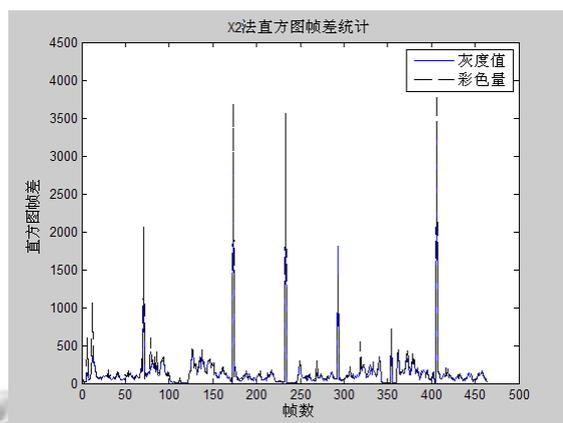
由表 1 可知像素法、X²直方图法和 X²直方图分块法准确率都比较高, 其中 X²直方图分块法准确率最高

但是时间比像素法和 X²直方图长. 而直方图和边缘轮廓变化率法准确率相对较低, 但是边缘轮廓变化率法时间比直方图法时间长 4 倍多. 由于光照变化影响, 以上五种算法性能都没有很达到最优.

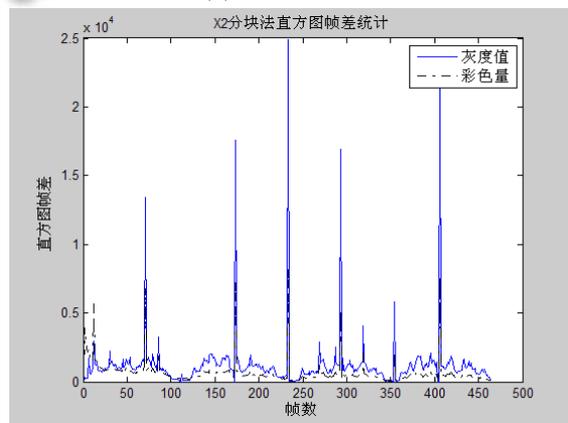
如图 2 展示了第一段视频的其中四种镜头检测的结果比较:



(a) 直方图法



(b) X²直方图法



(c) X²直方图分块法

(下转第 69 页)

选修人数超出限选人数的选课课程,之后由系统随机进行筛选,选该门课的学生会得到一个随机数,按随机数大小进行排序,较小的保留.如果某一门课程的学生选课名额未满足,这些有效的选课记录也不参加下一轮筛选,只有后面报名的学生才参加筛选.

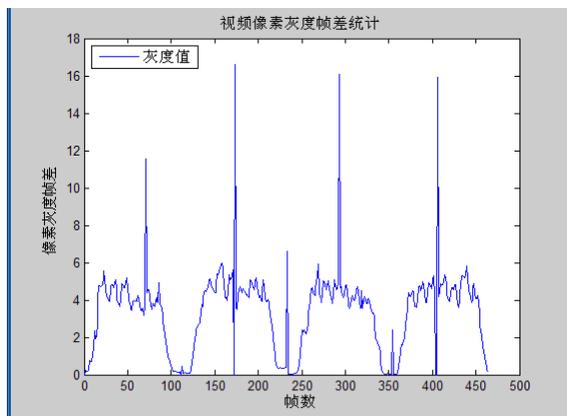
8 结语

校园网上选课系统是一个基于 Web Servers 的三层架构的管理信息系统,采用 B/S 架构,保证开发过程的逻辑结构更加清晰,开发内容的模块化,并且提高开发速度,使整个系统的开发完全是并行的,这样,不同人员之间的开发是完全可以独立的,不会因为其他人的功能模块没有实现,从而影响到自己本身模块的开发,这样就大大的加快了开发速度,缩短了开发周期,并且,应为该系统完全采用了模块化的开发机制,对整个系统的维护和操作更加方便,不需要重复的去修改程序,只修改需要修改的部分,而不影响其它功能模块的正常运行.

参考文献

- 1 吴洁明.软件工程应用实践教程.北京:清华大学出版社,2003.58-60.
- 2 黄敬仁.系统分析.北京:清华大学出版社,2003.9-13.
- 3 Kingston JH.电子商务.北京:机械工业出版社,2003.
- 4 Marshal B. XML 示例程序导学.北京:清华大学出版社,2003.55-60.
- 5 Long L. C++ and Object-Oriented Programming. Prentice Hall, Inc. 2005, 6-9.
- 6 Comer DE. Computer Networks and Internet. Prentice Hall, Inc.2003, 23-26.
- 7 赵强.J2EE 应用开发.北京:电子工业出版社,2003.33-36.
- 8 杨学瑜.ASP 入门与提高.北京:清华大学出版社,2003.56-63.
- 9 罗超里.管理信息系统.北京:清华大学出版社,2004.13-16.
- 10 赵广辉,钟洛.基于WEB的选课系统的设计与实现.武汉理工大学学报(信息与工程管理版),2005,27(2):50-51.
- 11 周震.基于 ASP 技术开发的网络选课管理系统.洛阳师范学院,2005,20(2):71.

(上接第 7 页)



(d) 像素法

图 2 四种镜头检测比较

由图 2 曲线图可知镜头内帧差异值变化不大,但是由于光照和快速运动也会发生很大变化,镜头和镜头的边界也发生了很大的变化,怎么样区分是否镜头边界这几种算法都没有得到很好地诠释.而且选取的阈值也不是唯一的,如直方图选择的阈值是 0.025 渐变的阈值是 0.013, X2 直方图法选择的是 800 渐变的阈值是 175, 这些值都是根据经验得到的,没有一种通用的算法来计算阈值,这样得到的结果也不是最完美的.

参考文献

- 1 Hanjalic A. Shot-boundary detection:unraveled and resolved? IEEE Circuits and Systems for Video Technology, 2002,12(2): 90-105.
- 2 Lienhart R. Comparison of automatic shot boundary detection algorithms. Proc. of IS&T/SPIE Storage and Retrieval for Image and Video Databases VII,1999,1:36-56.
- 3 段富,朱伟.视频存储中一种改进的镜头分割算法.微计算机信息,2009,25(73):176-179.
- 4 孙季丰,李颖雄.视频镜头自动分割的若干方法.华南理工大学学报(自然科学版),2003,31(8):10-14.
- 5 侯冠华,史萍.视频分割与场景聚类算法研究.中国传媒大学学报自然科学报,2006,13(2):32-37.
- 6 陈文飞,许雪峰,苗作华,李康顺.亮度自适应的保熵直方图均衡化方法.计算机工程与应用,2012:1-6.
- 7 Tsamura E, Mezaris V, Kompatsiaris I. Gradual transition detection using color coherence and other criteria in a video shot meta-segmentation framework. IEEE, 2008:45-49.
- 8 张思俊.一种基于 Canny 算子边缘检测的车牌图像增强方法.重庆交通大学学报(自然科学报),2012:1-4.