

视频分镜头边界检测的一种新方法*

侯榆青¹ 葛宝² 彭进业¹

(1 西北大学电子科学系, 西安 710069)

(2 陕西师范大学物理学与信息技术学院, 西安 710062)

摘要 定义了用于渐变边界检测的累积帧差, 提出了一种基于高斯模型的镜头边界检测新方法, 对突变边界和渐变边界分别采用不同参数的高斯模型, 实现了阈值的自适应选择, 克服了现有的双阈值比较法中下边界阈值难以自适应设定的困难. 初步实验结果表明该方法能准确地检测突变和渐变边界.

关键词 分镜头边界检测; 高斯模型; 双阈值比较法

中图分类号 TP391 **文献标识码** A

0 引言

视频/图像分析一直是计算机视觉领域的热门课题, 其理论与应用研究受到广泛关注^[1,2]. 视频镜头边界检测是进行视频内容分析的首要步骤, 是基于内容的视频检索中更高级的语义、内容分析的基础. 镜头是一组连续的相互关联的帧, 是相机的一次连续拍摄, 代表时间或空间上连续的一组动作, 镜头的切换形成镜头边界. 镜头边界的类型一般可以分为突变型和渐变型两种. 渐变又可分为许多种, 主要包括溶解(dissolve)、淡入(fade in)、淡出(fade out)和扫换(wipe)等^[3]. 视频镜头边界检测的研究包括帧间不连续值(帧差)计算、阈值选择和检测性能评价三个方面. 在帧间不连续值计算的研究上, 已经提出像素比较法^[4]、颜色直方图法^[5,6]、MPEG-7 描述符方法^[7~9]等. 在阈值选择的研究上, 提出了高斯模型^[10]、视频编辑模型^[11]等方法. 检测性能评价用于衡量各种镜头检测算法优劣, 主要采用检出率和准确率、漏检率和错误率、ROC (Receiver Operating Characteristic) 曲线、覆盖系数、骰子(dice)系数等指标^[12,13]. 基于高斯模型的双阈值比较法^[10]既能用于突变边界的检测, 也能用于渐变边界的检测, 是一种有效的镜头边界检测方法. 但是在渐变边界的检测上, 由于渐变和镜头内部的帧间不连续值往往差距很小, 所以其中低阈值难以确定. 本文提出一种新的镜头边界检测方法, 对突变边界和渐变边界分别采用不同参数的高斯模型, 克服了双阈值比较法中下边界阈值选择的困难, 从而能更准确地实现渐变边界的检测.

1 基于高斯模型的阈值选择方法

1.1 高斯模型法

H. J. Zhang^[10]认为如果不考虑摄像机镜头切换和镜头移动, 那么帧与帧之间的差别由以下三种噪音引起: 第一, 将模拟信号转化为数字信号所带来的噪音; 第二, 由视频编辑设备所带来的噪音; 第三, 同一镜头内物体运动所带来的噪音. 假设这三种噪音服从高斯分布, 那么, 帧与帧之间的差别可分别认为是高斯噪音和相机移动(包括切换)所产生. 令 σ 为不连续值的标准差, μ 为均值, 那么

$$P(x) = \int_0^x \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \quad (1)$$

会在偏离均值的小范围内包含绝大多数不连续值, 也就是说, 不包括镜头转换的帧间不连续值会落在 0 到 $\mu + \alpha\sigma$ 内, α 为常量, 所以阈值 T_b 由式(2)计算

$$T_b = \mu + \alpha\sigma \quad (2)$$

当帧差大于 T_b 时, 则认为是镜头的边界. Hanjalic^[14]对其进行了改进, 将滑动窗口法和高斯模型相结合, 在局部范围内利用高斯模型求阈值.

1.2 双阈值比较法

为了检测渐变边界, 参考文献[10]中还提出了一种双阈值比较法. 该方法设置两个阈值 T_b 和 T_s ($T_b > T_s$), 当帧差大于 T_b 时, 为突变边界, 当帧差小于 T_b 而大于 T_s 时为渐变边界. 并且通过计算两种帧差来确定渐变边界的起始和结束位置: 一种帧差是连续两帧之间的帧差 $Z(k, k+1)$, 另一种帧差是相隔帧的帧差 $Z(k, k+l)$, 即相隔 l 帧的帧差. 当接续帧的帧差开始超过 T_s 时, 将这一帧判定为镜头渐变的起始帧, 当镜头渐变的起始帧检出后, 便开始计算 $Z(k, k+l)$, 即 k 逐渐增加的同时, 也逐渐增加 l , 显然, 相隔帧的帧差随着相隔帧数 l 的增加而增加, 因而相隔帧的帧差是一个累计帧差. 当相隔帧的帧差 $Z(k, k+l)$ 累计超过 T_b , 而连续帧的帧差低于 T_s 时, 这

* 西安市科技攻关计划(GG9907)项目资助

Tel: 029-88308119 Email: houyuqin@nwu.edu.cn

收稿日期: 2005-04-04

一帧便判定为镜头渐变的终止帧. 上述两种帧差是同时计算的, 在相隔帧差开始累计后, 同时观察连续帧的帧差 $Z(k, k+1)$, 如果 $Z(k, k+1)$ 小于 T_s , 则废除该起始帧, 接着重新寻找起始帧, 这样就确定了渐变边界起始帧和结束帧的位置. 显然, 这种方法要获得成功的关键之一是设置合适的 T_b 和 T_s , T_b 采用式(2)计算, 但 T_s 的合理取值比较困难, 这是因为一般来说渐变和镜头内部帧间差往往相差较小.

2 改进的方法

针对上述方法存在的问题, 提出了一种改进的阈值选择方法, 能同时检测突变和渐变边界. 以各帧颜色直方图为特征计算两种帧差, 一种为当前帧与后一帧的帧差

$$Z_k^{(1)} = Z(k, k+1) \tag{3}$$

另外, 定义另一种帧差为当前帧与前后各 N 帧的帧间差的累加, 即

$$Z_k^{(2)} = \frac{1}{2N+1} \sum_{i=-N}^N Z(k, k+i) \tag{4}$$

$Z_k^{(1)}$ 用于突变边界检测, 而 $Z_k^{(2)}$ 用于渐变边界检测. 与之对应, 用两个不同参数的滑动窗高斯模型分别进行检测, 阈值的选择仍然采用式(2), 其中

$$\mu = \frac{1}{W} \sum_{k=1}^W Z_k \tag{5}$$

$$\sigma = \left[\frac{1}{W} \sum_{k=1}^W (Z_k - \mu)^2 \right]^{1/2} \tag{6}$$

式中 W 为滑动窗的宽度, μ 和 σ 是当前滑动窗内 Z 值的均值和方差的估计. 式(5)和(6)是分别进行突变和渐变边界检测时的统一表达, 当用于突变边界的检测时, Z_k 用 $Z_k^{(1)}$ 代替; 用于渐变边界检测时, Z_k 用 $Z_k^{(2)}$ 代替. 也就是说, 利用式(4)定义的帧差, 渐变边界的检测可以采用与突变边界相同的检测模型, 避免了低阈值的设置. 需要注意的是, 由式(4)不难看出, 突变边界也会形成长度为 $2N+1$ 的渐变边界, 称其为虚假边界, 因此, 在综合两者的检测结果时需要将这些虚假边界剔除掉. 下面简单描述边界检测的过程:

1) 计算帧差. 统计各帧的 HSV 颜色直方图特征, 计算各帧位置的帧差 $Z_k^{(1)}$ 和 $Z_k^{(2)}$. 直方图间的不相似性测度采用城区距离.

2) 检测突变边界集. 根据 $Z_k^{(1)}$, 采用滑动窗高斯模型检测出突变边界集 $B_{cut} = \{B_1, B_2, \dots, B_n\}$.

3) 检测渐变边界候选集. 类似地, 利用 $Z_k^{(2)}$ 检测出渐变边界候选集 $B_{grad} = \{B_1, B_2, \dots, B_n\}$, 但与突变情形不同, 此时的 $B_n = \{k, k+1, \dots, k+m | m \geq 1\}$, 即检测到的边界是许多成簇的连续边界的集合.

4) 筛选渐变边界. B_{grad} 中长度为 $2N+1$ 的那些子集, 被认为是由于突变边界引起的虚假边界, 将它们剔除后就得到了渐变边界的集合 $B_{gradual}$.

5) 渐变边界的进一步定位. $B_{gradual}$ 中的子集只能粗略地反映渐变边界开始与结束位置, 由于 $Z_k^{(2)}$ 已经很平滑, 所以可通过从峰值位置开始分别向左右两边求二次差分, 根据二次差分的符号变化与连续的 p 个差分值小于预先设定的阈值来更精确地判定渐变开始和结束位置.

通过以上步骤, 可检测出突变的位置和渐变的开始、结束位置.

3 实验结果

有研究表明 HSV 颜色空间与人类对颜色的感知接近, 人眼可独立感知该空间各颜色分量的变化, 其中 H(Hue, 色调) 分量又较好地克服了一般光照变化对颜色直方图所带来的影响, 因此选择 HSV 颜色空间作为特征来计算帧差, 并且取 H、S、V 分别为 12、5、5 等份^[15].

图 1 中(a)、(b)是截取的同一段视频所得到的不连续值曲线(其中实线部分为不连续值曲线), (a)为采用式(3)直接求出的不连续值曲线, (b)为采用式(4)求出的不连续值曲线. 正如前面所分析的, 采用新的方法以后, 本来不太光滑的抛物线变得光滑起来, 从而, 检测出边界的峰值位置, 也就是渐变的中心以后, 渐变边界的起始和结束位置定位就变得容易得多. 虽然突变和渐变边界都采用滑动窗高斯模型法, 但它们的参数并不一样, 突变检测的参数应

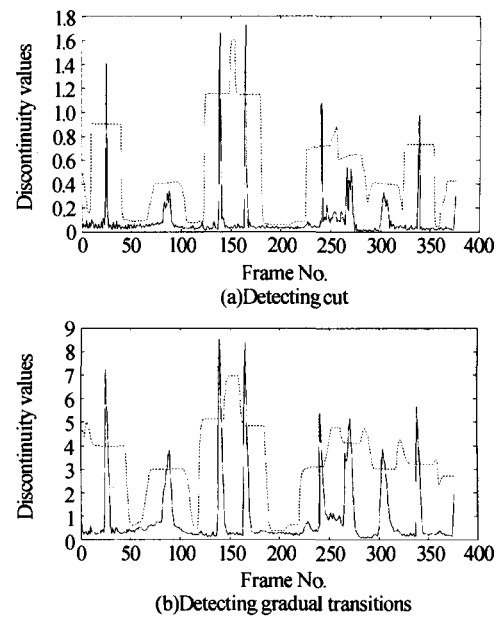


图 1 同一段视频所得到的不连续值曲线
Fig. 1 The discontinuity in a video sequence

使自适应的阈值变化更快一些,以便跟随变化的趋势.因为渐变区域在滑动窗内的标准差较小,所以为了检测出其峰值位置,它的阈值变化应该平坦一些.实验中检测突变的参数设置为:窗口宽度 $W=25$,高斯模型中的系数 $\alpha=3.7$;而在检测渐变边界时候,窗口宽度 $W=35$,系数 $\alpha=2.3$,算法中的 p 取值为 2, N 取值为 2.图中虚线部分分别为检测突变和检测渐变的自适应阈值曲线.可以看出,(a)图中的自适应曲线能很好地检测突变边界,而(b)图中的自适应曲线旨在检测渐变边界.综合两者的检测结果即可得到突变和渐变位置.

对 CNN 新闻的几个渐变较多的片断进行检测,总共有 62 个边界,其中含有 5 个突变和 57 个渐变边界.突变边界全部检测出,对于渐变,正确检测到 51 个,错检 4 个,漏检 6 个,检出率 89.47%,正确率 92.73%.在正确检测到的边界中,平均骰子系数^[13]达到 92.91%,和已有的检测性能^[16]比较,本文方法取得了更好的检测结果.

4 结论

对于渐变边界的检测,基于高斯模型的双阈值比较法简单、可行,但是其低阈值的设置较困难,且对于不同的视频有不同的最佳阈值.本文提出一种基于高斯模型的自适应阈值选择的新方法,通过定义的累积帧差,采用不同参数的高斯模型分别检测突变和渐变边界,再将两者综合就可以得到检测结果.该方法避免了双阈值比较法中下边界阈值的设置问题,从而取得了较理想的镜头边界检测效果.

参考文献

- 李明,吴艳,吴顺君.基于小波多通道特征级融合的彩色纹理图像分析.光子学报,2004,33(8):999~1003
Li M, Wu Y, Wu S J. *Acta Photonica Sinica*, 2004, 33(8): 999~1003
- 单越康,许昌.基于计算机视觉的光学目标自动定位系统.光子学报,2000,29(10):899~902
Shan Y K, Xu C. *Acta Photonica Sinica*, 2000, 29(10): 899~902
- Lienhart R. Comparison of automatic shot boundary detection algorithms. *SPIE*, 1999, 3656: 290~301
- Hanjalic A. Shot-boundary detection: Unraveled and resolved? *IEEE Transactions on Circuits and Systems for Video Technology*, 2002, 12(2): 90~105
- Yeo B L, Liu B. Rapid scene analysis on compressed video. *IEEE Transactions on Circuits and Systems for Video Technology*, 1995, 5(6): 533~544
- Nagasaka A, Tanaka Y. Auto video indexing and full-video search for object appearances. *In Visual Database Systems II*, 1992, A-7: 113~127
- Shim S H, Yang S J. Real-time shot boundary detection for digital video camera using the MPEG-7 descriptor. *SPIE*, 2002, 4666: 161~171
- Lee H K, Kim. Video contents summary using the combination of multiple MPEG-7 metadata. *SPIE*, 2002, 4664: 1~12
- Graves A, Lalmas M. Video retrieval using an MPEG-7 based inference network. *SIGIR'02*, 2002, 339~346
- Zhang H, Kankanhalli A, Smoliar S W, et al. Automatic partitioning of full-motion video. *Multimedia Systems*, 1993, 1: 10~28
- Hampapur A, Jain R C, Weymouth T. Production model based digital video segmentation. *Multimedia Tools and Applications*, 1995, 1(1): 9~46
- Ford R M. A quantitative comparison of shot boundary detection metrics. *SPIE*, 1999, 3656: 666~676
- Kobla V, Dementhon D, Doermann D, et al. Special effect edit detection using video trails: a comparison with existing techniques. *SPIE*, 1999, 3656: 302~313
- Hanjalic A, Lagendijk R L. A novel video parsing method with improved thresholding. *Proceedings of the Third Annual Conference of the Advanced School for Computing and Imaging*, Heijen, NL, Jun. 1997. 3~8
- Lienhart R, Zaccarin A. A system for reliable dissolve detection in videos. *Proceedings of ICIP*, 2001, 3: 406~409
- Wan X, Kuo C C J. Color distribution analysis and quantization for image retrieval. *SPIE*, 1996, 2670: 8~16

A Novel Approach to the Detection of Shot Boundary

Hou Yuqing¹, Ge Bao², Peng Jinye¹

¹ Department of Electronics, Northwest University, Xi'an 710069

² College of Physics & IT, Shaanxi Normal University, Xi'an 710062

Received date: 2005-04-04

Abstract For the detection of shot boundary in a video sequence, the selection of a proper threshold for computing discontinuity values is an important and difficult problem. A formula for calculating discontinuity values is defined, then a new Gaussian model based technique is proposed, which needs not setting low boundary threshold. It is a kind of improvement of twin-comparison method. Primary experimental results demonstrate that the proposed method enables to locate boundaries of gradual transitions and cuts exactly and quickly.

Keywords Shot boundary detection; Gaussian model; Twin-comparison



Hou Yuqing was born in 1963, She got her B. S.(Physics) and M. S. degrees (Electronics) from Northwest University in 1984 and in 1990, respectively. She is now an associate professor in Department of Electronics, Northwest University. Her current research interests are image and video processing.