

基于分块图像小波相对能量和熵的视频镜头边界检测算法*

戴小文^{1a,2}, 蔡志平², 钟桂英^{1b}

(1 西南交通大学 a. 电气工程学院, b. 建校指挥部, 成都 610031)

(2 厦门大学电子工程系, 福建 厦门 361000)

摘要: 提出了对检测窗内图像帧进行分块并进行小波变换, 计算尺度 j 上分块图像的低频、水平、垂直和对角信号的小波相对能量和小波熵的方法, 给出了分块图像小波熵的计算公式. 采用尺度 j 上分块图像低频信号小波熵作为特征向量检测视频镜头是否发生转换, 即在检测窗内, 帧间低频信号小波熵值只发生一次大突变(与设定的阈值相比)就是剪切转换, 若发生多次比较大的变化就是渐变转换. 采用检测窗内分块图像在尺度 j 上的垂直、水平和对角高频信号的小波熵作为另一特征量可以区分渐变镜头的类型: 淡进/出(fade in/out)或扫换(wipe). 采用大量的高清和标清码流进行实验并与基于概率的方法进行比较. 结果证明: 本文的视频镜头检测方法具有强的鲁棒性, 能检测出镜头转换的类型: 剪切、淡进/出和扫换, 同时有较高的查准率和查全率.

关键词: 视频镜头; 检测; 图像分块; 小波变换; 小波熵

中图分类号: TP391.41

文献标识码: A

文章编号: 1004-4213(2008)07-1492-5

0 引言

镜头作为视频内容的基本结构单元, 由摄像机一次连续拍摄的帧序列组成. 镜头边界检测的任务就是有效地识别镜头边界类型及其位置, 即通过比较视频序列的帧间差异来寻找镜头边界变化的规律. 镜头检测比较复杂, 主要表现在: 1) 需要找到受光线及其他参数影响小的视频帧的特征量; 2) 对渐变镜头的检测比较复杂. 镜头边界检测的方法有多种, 如抽取一帧或一组视频帧的一个或多个特性, 并对这些特性采用不同的方法来实现检测镜头转换^[1-3]. 文献[3]中提出了分层的多分辨率镜头边界检测方法, 该方法是先对每帧图像进行 Harr 小波分解, 然后使用低频分量的颜色直方图差来区分突变, 再使用高频分量的边缘特征(Edge Spectrum Average, ESA)来检测淡入淡出, 并用双色差(Double Chromatic Difference, DCD)来检验溶解(dissolve), 最后使用空间平移运动矢量(Spatial Translation Motion Vector, STMV)将候选边界分为渐变、突变和错误检测点. 此算法不仅计算量大, 而且对渐变的分类没有给出具体的方法. Hanjalic^[4]主要针对剪切转换(cut)和溶解的渐变转换镜头提出基于概率的算法, 把每一帧视频图像

分割成没有交迭的块, 从每一块中抽取出 YUV 颜色元素. 相邻帧间的块之间通过颜色元素来匹配, 此方法采用的特征向量维数多, 对光照等变化比较大. Liehart^[5]对 dissolve 图像提出采用训练分类器的算法. Cernekova^[6]提出, 为了降低特征矢量维数, 对每一帧的颜色直方图进行单值分解. 通过比较每一帧的特征矢量和当前分割的特征矢量的平均值来进行视频分割. 此算法的缺点是对内容变化大和内容变化小的镜头检测错误多. Lelescu and Schonfeld^[7]提出针对压缩图像的统计算法, 如 MPEG 图像, 此算法以每一个 I 帧和 P 帧的每个块的颜色和亮度值作为特征矢量, 对这些特征矢量采用主成分分析(PCA). 此算法的难点是在图像开始 M 帧的镜头特征的抽取.

熵是一个随机事件的不确定性或信息量的度量, 小波熵能反映图像小波变换频率空间的能量分布信息, 因为小波函数在频域与时域上均不具有脉冲选择性质, 而具有一定的支撑空间, 因此, 图像小波熵能反映图像的能量分布特征. 本文提出把图像划分成不重叠的块, 对每一块图像进行小波分解, 在尺度 j 上, 计算各分块图像的小波相对能量谱, 此能量谱可以作为图像的特征向量. 但为降低特征向量的维数, 采用各分块图像的低频信号小波熵以及垂直、水平和对角高频信号的小波熵作为分块图像的特征向量. 本文的镜头检测算法包括两步骤: 第一, 顺序捕捉一组视频帧作为检测窗, 对检测窗内的每帧图像进行分块并进行小波变换, 在尺度 j 上计算

* 国家自然科学基金(50407009)和信息产业部电子发展基金(555)资助

Tel: 13959286946 Email: xwdai678@yahoo.com.cn

收稿日期: 2008-01-14

低频信号小波熵来检测镜头是否发生转换. 同时, 以检测窗内低频信号小波熵是否发生多次突变来判别镜头转换是剪切还是渐变; 第二, 利用分块图像在尺度 j 上小波变换的高频信号小波熵对渐变镜头再进行分类, 判别是淡进/出还是扫换.

1 图像小波分解

镜头边界类型一般分为两种: 剪切突变(图 1 所示的电视镜头切换)和渐变(图 2 所示的淡进淡出、图 3 所示的扫换). 镜头的剪切变换是突然间完成的, 但每一个镜头的内容也由多帧构成, 而渐变转换一般需要几十帧甚至几百帧才能完成.

小波变换在许多领域得到很好的应用, 如图像压缩、故障检测等. 由图像小波变换的实现过程知道: 图像数据的每一级分解总是将上级低频数据划分为更精细的频带. 其中 LH_j 频带是先将上级低频

图像数据 LL_{j-1} 在水平方向低通滤波后(行方向), 再经垂直方向高通滤波(列方向)而得到的, 因此, LH_j 频带中包含了更多垂直方向的高频信号. 相应地, HL_j 频带主要是原图像水平方向的高频成分, 而 HH_j 频带是图像中对角线方向高频信息的体现, 尤其是以 45° 或 135° 角的高频信息为主, LL_j 是图像在经过 j 级分解后得到的图像低频信号. 采用这些图像的低频信号和图像在各个尺度各个方向上的高频信息可以实现视频镜头的检测. 从多分辨率分析的角度考虑小波图像的各个频带时, 这些频带之间并不是完全无关的, 特别是各个高频带, 它们是图像同一边缘、轮廓和纹理信息在不同方向、不同分辨率下由细到粗的描述, 它们之间必然存在着一定的关系. 这些频带中所对应的边缘、轮廓和相对位置是相同的.



图 1 SBS 电视节目中的剪切镜头切换

Fig. 1 Transition process of cut shot

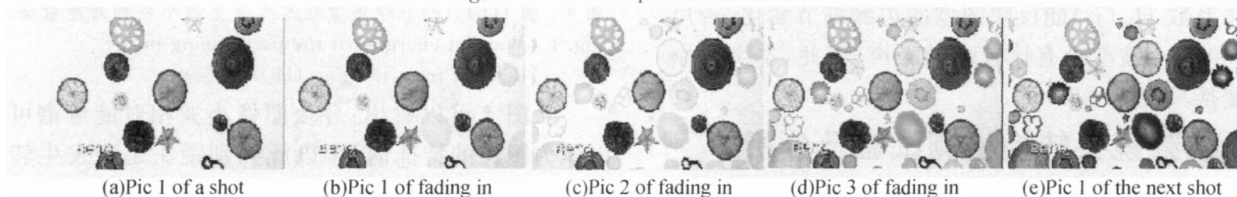


图 2 Fade in 的渐变镜头图

Fig. 2 Transition process of fade in shot

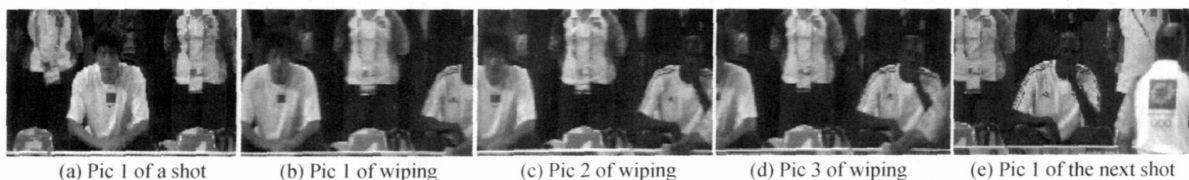


图 3 Wipe 的渐变镜头

Fig. 3 Transition process of wipe shot

2 小波相对能量和小波熵

当小波基函数是一组正交基函数时, 小波变换具有能量守恒的性质, 这样可以定义单一尺度下的小波能量为该尺度下小波系数的平方和

$$E_j = \sum_k |C_j(k)|^2 \quad (1)$$

式中, $(j = 1, 2, \dots, J)$ 、 $(k = 1, 2, \dots, N)$, J 为最大分解尺度, N 为所分析信号的采样点数. 依此类推, 信号总能量的表达式为

$$E_{\text{total}} = \sum_j \sum_k |C_j(k)|^2 = \sum_j E_j \quad (2)$$

式中, j, k 的意义同式(1), 因此, 小波相对能量定义为

$$p_j = E_j / E_{\text{total}} \quad (3)$$

p_j 之和为 1, 如果把小波系数矩阵处理成一个概率分布序列, 由这序列计算得到的熵值就反映了这个系数矩阵的稀疏程度, 也就是信号概率分布的有序程度, 这种熵就叫小波熵^[8,9], 也称小波能量谱, 定义为

$$W_E = - \sum_j p_j \ln p_j \quad (4)$$

3 图像分块及分块图像的小波熵

把图像分为不重叠的子块, 设分块图像的大小为 $w \times v, w \in M, v \in N, M \times N$ 为图像的大小, M 表示图像的最大行数, N 为图像最大列数, 用 W_l 来表示分块图像, l 是分块图像的标记号. 把图像划分成 L 个互不重叠的分块图像(若图像边界部分不能恰好划分, 做适当地添零延拓, L 为最多分割的子块图像数量).

用 $E_{mj}(W_l)$ 表示分块图像 l 小波变换后的小波能量, 按照小波相对能量的定义, 用 $p_{mj}(W_l)$ 表示图像在 j 尺度下分块图像 l 经小波变换后小波能量与总能量的比值, 与式(3)定义的相对能量类似. 因此, 定义在 j 尺度下分块图像 l 的小波熵为

$$H_m(j) = - \sum_j (p_{mj}(W_l) \log(p_{mj}(W_l))) \quad (5)$$

式中 m 是指分块图像经小波变换后在 j 尺度下低频、垂直、水平和对角信号中的一种, $l = 1, 2, \dots, L$ 表示分块图像的标记. 在每一尺度下, 分块图像经小波变换后的低频、垂直、水平和对角小波系数矩阵, 都可以计算 $H_m(j)$, 因此可以得到分块图像小波熵的变化曲线.

采用分块图像小波变换低频系数的 $H_m(j)$ 能反映图像的基本轮廓变化; 对分块图像小波变换的高频系数 $H_m(j)$ 能反映图像的边缘细节变化. 若尺度 $j = 1$, 小波变换有比较多的噪声, 因此, 本文取小波变换尺度 $j = 2$.

4 视频镜头转换检测试验及分析

采用不同的小波基得到的小波相对能量和分块图像小波熵是不一样的. Daubechies 已经证明, 既具有紧支集, 又具有对称性的正交小波是不存在的, 在图像处理应用中, 为了减少处理后图像的相位延迟, 通常要求小波具有对称性, 使处理后图像的边缘失真最小, 为此, 放松对正交性的要求, 而采用双正交小波^[10]. 通过理论分析和实验验证, 本文采用紧支集双正交小波 Bior3.7 作为小波熵计算的小波基^[11].

4.1 视频镜头剪切转换的检测试验

图1是镜头剪切转换的高清视频图像(1920 × 1080), 对图1(b)和图1(c)先进行图像分块, 分块图像大小为 32 × 32. 对分块图像进行小波变换, 小波基函数为 bior3.7, 尺度 $j = 2$, 计算每个分块图像的

小波相对能量, 得到图1(b)和(c)的分块图像的小波相对能量谱分别如图4. 横坐标为按顺序取出的分块图像标记号, 纵坐标为分块图像的小波能量与整幅图像能量的比值.

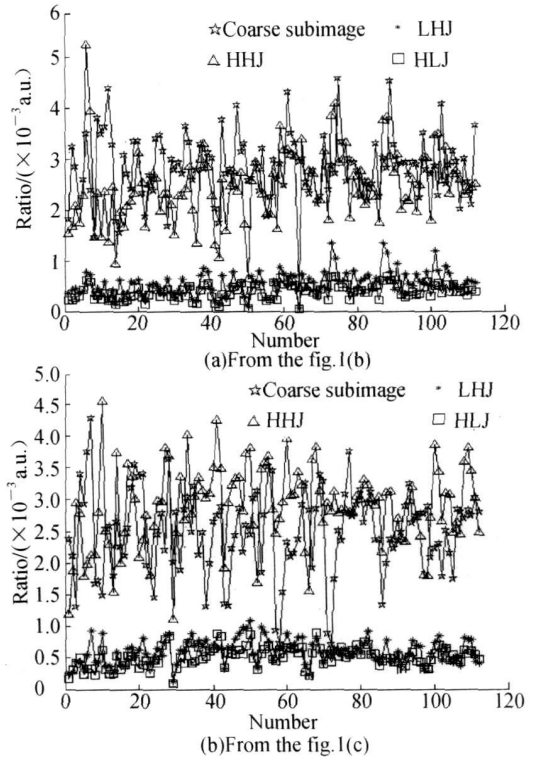


图4 图1(b)、(c)分块图像在尺度2上的小波相对能量谱
Fig. 4 Wavelet energies of the partitioning image divided from the Fig. 1(b) and (c)

从图4可以看出: 分块图像小波相对能量谱可以作为图像的特征向量, 以此判别镜头是否发生转换. 为降低特征向量的维数, 本文在分块图像小波相对能量的基础上, 采用分块图像的低频信号小波熵作为图像的特征向量来判断镜头是否发生转换. 假定顺序捕捉若干视频帧作为一个检测窗, 计算检测窗内各帧的分块图像在尺度 $j = 2$ 时的小波熵. 表1是图1~3镜头转换过程中检测窗内帧间分块图像在 $j = 2$ 时低频信号的小波熵. 由表1可得知图1(b)与(c)之间、图2(a)与(b)、图2(d)与(e)之间、图3(a)与(b)、图3(b)与(c)、图3(c)与(d)之间的小波熵变化比较大. 因此, 在检测窗内帧间小波熵只发生一次大的突变(设定一个合适的判别阈值)可以判别为剪切转换, 若发生多次比较大的变化就是渐变转换. 因此, 可以检测出图1是发生了镜头的剪切转换, 而图2和3可能发生了渐变镜头的转换.

表1 图1~3分块图像 $j = 2$ 时低频信号小波熵

图号	1(a)	1(b)	1(c)	1(d)	1(e)	2(a)	2(b)	2(c)	2(d)	2(e)	3(a)	3(b)	3(c)	3(d)	3(e)
小波熵	1.84	1.86	1.66	1.57	1.60	1.65	1.53	1.57	1.59	1.41	1.91	1.82	1.67	1.77	1.75

4.2 渐变镜头的类型的检测

由4.1节可知, 采用分块图像的小波变换低频

信号的小波熵作为视频帧的特征向量可以检测镜头是否发生了转换以及判别是剪切还是渐变转换. 本

小节对渐变镜头的类型进行检测判别, 即是淡进淡出还是扫换。

图 1~3 检测窗内帧间分块图像在尺度 $j = 2$ 的水平、垂直和对角高频信号的小波熵的欧式距离(用 E_{ab} 表示帧 a, b 之间的欧式距离) 如表 2。

从表 2 可以看出, 检测窗内淡进/出与扫换镜头转换的高频信号的小波熵间的欧氏距离是不同的, 淡进/出镜头转换的帧间分块图像的高频信号的小波熵的欧氏距离比较大, 而扫换的镜头转换帧间分块图像的高频信号的小波熵的欧氏距离要小。因此, 采用检测窗内分块图像高频信号的小波熵间的欧氏距离可以分辨出渐变镜头转换的类型, 即淡进/出还是扫换。

表 3 实验结果比较图

镜头类型	突然切换			查准率	查全率	淡进/出转			查准率	查全率
检测方法	镜头数	误检	漏检	/%	/%	换镜头	误检	漏检	/%	/%
本文方法	100	3	1	97	99	170	18	21	89.41	87.64
A. Hanjalic	100	2	1	98	99	170	27	24	84.11	85.88

对于扫换的镜头转换实验表 3 中没有列出, 实验结果查准率和查全率分别达 84.4% 和 81.5%。值得注意的是: 查准率和查全率与检测算法中的阈值选取方法有一些关系。从表 3 可以清楚看出采用本文的两步检测方法, 对于淡入/出和扫换的镜头转换, 其查准率和查全率都比文献[4]的检测方法要高。

不同视频镜头转换的检测算法计算复杂度是不同的, 经实验证明采用窗内帧间图像的颜色直方图差来检测的算法计算复杂度最低, 但效果较差。A. Hanjalic 方法的计算复杂度要比本文的低 10% 左右。本算法的复杂度与分块图像的大小有关系, 分块图像越小, 计算小波变换的次数越多, 复杂度越高。总体上本算法计算复杂度要比文献[3]的低许多。

6 结论

顺序捕捉一组视频帧作为检测窗, 对检测窗内的图像进行分块并小波变换, 计算其小波相对能量谱, 此能量谱可以作为图像检测的特征向量。为降低特征向量的维数, 本文采用在 j 尺度上, 分块图像的低频信号小波熵作为图像的特征向量, 检测窗内分块图像的低频信号小波熵, 判别是否发生镜头的转换, 并根据检测窗内帧间分块图像的低频信号小波熵是否出现多次大的变化(与设定的阈值相比), 判别是剪切变换还是渐变转换。在此基础上, 采用检测窗内分块图像在 j 尺度上的水平、垂直和对角高频信号的小波熵值作为图像的另一特征量, 还可以对渐变转换镜头的类型进行区分, 即是淡进/出还是扫换。大量的实验及相关对比表明, 本文的检测方法有

表 2 图 2、图 3 检测窗内帧间分块图像镜头变换检测窗内帧间小波熵的欧式距离

图号	E_{2a-2b}	E_{2b-2c}	E_{2c-2d}	E_{2d-2e}	E_{3a-3b}	E_{3b-3c}	E_{3c-3d}	E_{3d-3e}
欧氏距离	0.187	0.063	0.034	0.073	0.046	0.009	0.011	0.016

5 镜头检测实验

利用仿真软件 Matlab 编程, 实验测试码流有: LG 高清测试 TS 码流(2 Gbytes, 高清 1 920 × 1 080)、SONY 高清测试 TS 码流(2 Gbytes)、Benq 高清测试码流(1.6 Gbytes)、2004 年雅典奥运会比赛节目高清 TS 流(2 Gbytes), 韩国 SBS 电视台高清节目源, 也有部分标清的 TS 码流。得到实验检测结果并同文献 4 A. Hanjalic 方法相比较, 如表 3。

较强的鲁棒性并且可以检测出镜头转换的类型, 同时有较高的查准率和查全率。

参考文献

- [1] CITSACES C, NIKOLAIDIS N, PLTAS I. Video shot detection and condensed representation [J]. *IEEE trans Signal Processing Magazine*, 2006, 23(2): 28-37.
- [2] HOU Yurqing, Ge Bao, PENG Jirye. A novel approach to the detection of shot boundary[J]. *Acta Photonica Sinica*, 2005, 34(10): 1597-1600.
侯青育, 葛宝, 彭进业. 视频分镜头边界检测的一种新方法[J]. *光子学报*, 2005, 34(10): 1597-1600.
- [3] YU H H, WOLF W. A hierarchical multiresolution video shot transition detection scheme [J]. *Computer Vision and Image Understanding*, 1999, 75(1/2): 196-213.
- [4] HANJALIC A. Shot boundary detection: Unraveled and resolved [J]. *IEEE Trans Circuits and Systems for Video Technology*, 2002, 12(2): 90-105.
- [5] LIENHART R. Reliable dissolve detection [C]. *SPIE*, 2001, 4315(1): 219-230.
- [6] CERNEKOVA Z, PITAS I, NIKOU C. Information theory based shot cut/fade detection and video summarization [J]. *IEEE Trans Circuits and Systems for Video Technology*, 2006, 16(1): 82-91.
- [7] LELESCU D, SCHONFELD D. Statistical sequential analysis for real-time video scene change detection on compressed multimedia bitstream [J]. *IEEE Trans Multimedia*, 2003, 5(1): 106-117.
- [8] LEMIRE D, PHARAND C, RAJAONAH J C, et al. Wavelet time entropy, T wave morphology and myocardial ischemia [J]. *IEEE Transactions in Biomedical Engineering*, 2000, 47(7): 967-970.
- [9] HE Zhengyou, CAI Yumei, QIAN Qingqian. A study of wavelet entropy theory and its application in electric power

system fault detection[J]. *Proceedings of the CSEE*, 2005, 25(5): 38-43.

何正友, 蔡玉梅, 钱清泉. 小波熵理论及其在电力系统故障检测中的应用研究[J]. *中国电机工程学报*, 2005, 25(5): 38-43.

- [10] ZHE Xi-an. Wavelet analysis and its application in SST images emulation compression[D]. Beijing: China University of Mining & Technology, 2003: 35-39.

朱希安. 小波分析及在 SST 图像仿真压缩中的应用[D]. 北京: 中国矿业大学, 2003, 35-39.

- [11] DONG Wei, JUN Tian, WELLS R O. A new class of biorthogonal wavelet systems for image transform coding[J]. *IEEE Transaction on Image Processing*, 1998, 7(7): 1000-1013.

An Algorithm of Video Shot Detection Based on Partitioning Image Wavelet Entropy

DAI Xiaowen^{1a,2}, CAI Zhiping², ZHONG Guirong^{1b}

(1 a College of Electrical Engineering, b The Headquarters of Building New School, Southwest Jiaotong University, Chengdu 610031, China)

(2 Department of Electronic Engineering, Xiamen University, Xiamen, Fujian 361000, China)

Received date: 2008-01-14

Abstract: Shot boundary detection is considered to be the primitives for higher level content analysis, indexing, and classification. A new video shot detection algorithm based on wavelet energy and wavelet entropy of the partitioning image is presented. The proposed detector includes twofold. The solution to shot-boundary detection or distinguish the cut and gradual shot using the low frequency of partitioning image wavelet entropy is provided. The formula of partitioning image wavelet entropy is given out. A solution is given to detect the fade in/out or wipe shot boundary using the high frequency of partitioning image wavelet entropy. Major advantages of the detector are its robust and good performance, while there is also the possibility to detect different types of shot boundaries simultaneously. The experiment results indicate that the new algorithm gives better performance than A. Hanjalic.

Key words: Shot-boundary; Detection; Partitioning image; Wavelet transfer; Wavelet entropy



DAI Xiaowen was born in 1967, and received the Ph. D. degree in vehicle engineering from Southwest Jiaotong University in 2001. He is currently an associate professor at College of Electrical Engineering Southwest Jiaotong University, Chengdu, Sichuan. His research interests focus on signal processing for digital video technologies and multimedia applications.