

基于颜色直方图的快速镜头分割方法

俞 璐^{1,2}, 乔瑞萍², 胡宇平², 赵 静^{1,2}

(1 宇航动力学国家重点实验室, 陕西 西安 710043; 2 西安交通大学 电子与信息工程学院, 陕西 西安 710049)

摘 要: 提出了一种基于视频帧核心区域直方图的快速视频分割方法, 该方法包括预处理、切变检测和渐变检测三个阶段。预处理提出基于人类视觉特性的视频帧核心区域分块, 并利用自适应阈值滤除视频中不包含镜头转换的视频序列; 切变检测利用直方图距离判定边界位置; 渐变检测针对渐变过程颜色直方图的变化规律, 提出一种结合直方图熵差和熵差异常点的检测方法。实验结果表明, 本文方法在大幅度降低运算复杂度的同时, 提高了切变和渐变镜头边界的检测精度。

关键词: 镜头边界检测; 预处理; 切变; 渐变

中图分类号: TP391

文献标识码: A

文章编号: 1000-7180(2014)02-0072-05

A Fast Shot Boundary Detection Method Based on Color Histogram

YU Lu^{1,2}, QIAO Rui-ping², HU Yu-ping², ZHAO Jing^{1,2}

(1 State Key Laboratory of Astronautic Dynamics, Xi'an 710043, China;

2 School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: A fast shot boundary detection method based on visual kernel region histogram is proposed in this paper. The method consists of pre-processing, cut detection and gradual transition detection. In the pre-processing round, the even blocking of kernel region based on human visual features is presented and the adaptive local threshold is adopted to filter out non-boundary segments. Cut detection is carried out by the use of histogram distance. According to the color histogram change rule of gradual transition, an approach combining the histogram entropy difference with its outliers is used in gradual transition detection. The experiments indicate that the computational complexity can be reduced significantly, and the detection results can be improved on both abrupt and gradual shot boundary by our method.

Key words: shot boundary detection; pre-processing; cut; gradual transition

1 引言

随着媒体和网络技术的迅猛发展导致视频的数量每天都在大幅增加, 庞大的信息量使得视频的管理和分析变得十分困难, 视频检索技术成为当前的研究热点。视频镜头分割作为视频检索的基础, 直接决定了视频后续处理的效果。视频的镜头变换可分为切变和渐变, 切变是指在两个相邻帧之间完成镜头的变换; 渐变则持续多帧, 通过淡入淡出、溶解和扫换等多种形式完成镜头的变换。目前, 很多研究者

提出了一些经典的算法进行镜头边界的检测, 但是这些算法都具有一定的局限性, 大部分算法只单一针对切变或者渐变镜头^[1], 能够同时检测切变和渐变镜头的算法复杂度又比较高^[2], 目前缺少一种能够同时针对切变和渐变镜头的快速镜头边界检测方法。本文正是在这个背景下, 提出了一种基于颜色直方图的快速有效的镜头分割方法, 以解决镜头边界检测存在的问题。

收稿日期: 2013-04-19; **修回日期:** 2013-06-23

基金项目: 国家自然科学基金资助项目(60902045); 宇航动力学国家重点实验室开发基金项目(2013ADL-DW0302); 西安交通大学实践基金(2013300)

2 镜头检测系统框架

本文提出的镜头分割方法将检测过程分为预处理模块,切变检测模块和渐变检测模块三个部分.

2.1 预处理模块

2.1.1 核心区域分块直方图

基于人类视觉的居中性和主体选择性,本文将视频帧以 1:6:1 的比例划分为两个区域,如图 1 所示,图中灰色区域处于视频帧的中心位置,该区域是人类视觉的焦点,因此一般包含图像要表达的主体内容,信息量较大;图中白色区域一般包含图像背景,或者字幕,信息量较小,若镜头的变换发生在同一场景内,反而会因为相似的背景对检测造成干扰.因此本文只针对图 1 中的灰色区域进行特征提取,这样不仅能够很好地体现空间位置的重要性,降低背景对边界检测的影响,而且可以提高算法的效率.

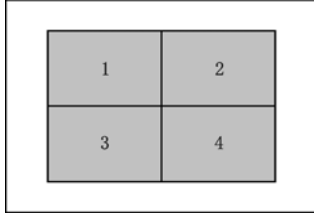


图 1 核心区域分块示意图

2.1.2 视频分段查找

大部分的边界检测方法需要对视频的每一帧进行扫描,提取视频帧的特征,再对特征进行比较,从而判断镜头边界.然而,视频中存在镜头边界的帧数远远小于视频中总的帧的数量,导致大量的运算时间浪费在非边界帧上.本文为了提高算法效率,在进行镜头边界检测前,加入预处理阶段^[3],筛选出包含边界的视频片段.

根据人眼的视觉暂留性质,一个镜头至少需要包含 40~50 帧^[4],为了保证一个子段中至多包含一个镜头变换,本文选用 11 帧为单位将视频均匀分段,每个子段的第一帧和前一子段的最后一帧重叠,如图 2 所示.

预处理阶段,本文选择符合人眼视觉的 HSV 色彩空间的色度分量直方图作为检测特征,为了弥补直方图忽略像素位置信息的缺陷,对核心区域进行了均匀分块,如图 1 所示.计算每个子段首尾两帧之间的核心区域色度分量分块直方图距离:

$$D_i = \sum_{m=1}^4 \sum_{k=1}^N |b_{k,m,j} - b_{k,m,j+10}| \quad (1)$$

式中, D_i 为第 i 个子段的直方图距离; N 为色度分量

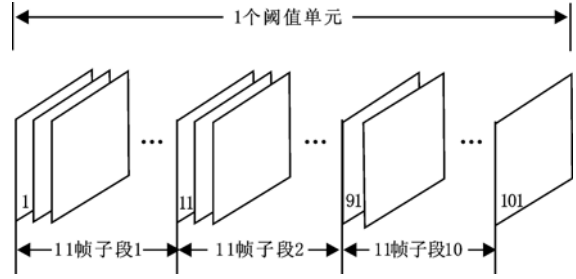


图 2 分段查找机制示意图

直方图的总级数; $b_{k,m,j}$ 为第 j 帧的第 m 个子块的色度分量归一化直方图中第 k 级的值, $j = 10 \times (i-1) + 1$. 如果 D_i 大于阈值,则认为该 11 帧子段中可能存在镜头边界,将该段作为候选段送入切变检测模块,否则将该子段滤除.

本文利用陷波原理自适应的计算各子段的判决阈值. 首先将每 10 个子段作为一个阈值单元,如图 2 所示,再利用下式计算各子段的阈值:

$$T_i = 1.1 \times \mu_i + 0.2 \times \frac{\mu_c}{\mu_i} \times \sigma_i \quad (2)$$

式中, T_i 为第 i 个子段的阈值, μ_c 为所有子段首尾帧直方图距离的均值, μ_i 和 σ_i 分别为阈值单元内除第 i 个子段外其余子段首尾帧直方图距离的均值和方差. 采用陷波原理的优点是对于含镜头边界的子段,其直方图距离较大,因此在计算 μ_i 和 σ_i 时剔除了一个较大的值,这样降低了阈值;对于不含镜头边界的子段,由于剔除的值较小,这样提高了阈值.

在实际应用中,可能存在少数包含镜头变换的子段,其首尾帧直方图的距离小于阈值,但是与相邻子段相比,其首尾帧直方图距离还是会有明显的增加,因此,若第 i 个子段满足

$$(D_i > 2D_{i-1} \text{ 或 } D_i > 2D_{i+1}) \text{ 且 } D_i > 0.8\mu_g \quad (3)$$

则将该子段也作为候选段,送入切变检测模块.

2.2 切变检测模块

镜头的切变发生在连续的两帧,由于前后两个镜头的内容具有较大的变化,因此切变的前后两帧图像的直方图也会产生较大的变化. 由于光照或者视频质量问题引起的相邻帧颜色变化会导致切变的误检,因此本文选用对颜色变化不敏感的亮度分量直方图作为切变检测特征,以此提高切变检测的准确率,同时可以缩短计算的时间,检测步骤如下所示.

step1 计算候选段内相邻两帧的直方图距离:

$$d_l = \sum_{m=1}^4 \sum_{k=1}^N |v_{k,m,l} - v_{k,m,l+1}|,$$

$$l = 1, 2, \dots, 10 \quad (4)$$

式中, d_l 为候选子段内第 l 帧和第 $l+1$ 帧的核心区域亮度分块直方图距离, $v_{k,m,l}$ 为候选子段内第 l 帧的第 m 个子块的亮度分量归一化直方图中第 k 级的值;

step2 若 d_n 满足:

$$(1) d_n = \max(d_l) \quad (5)$$

$$(2) d_n > \text{th} \times \frac{d_{n-2} + d_{n-1} + d_{n+1} + d_{n+2}}{4} \quad (6)$$

则认为该子段的第 n 帧和第 $n+1$ 帧之间发生了镜头的切变. 本文中, 取 $\text{th} = 2.75$.

由于渐变过程持续时间较长, 一般连续的候选子段中更有可能存在渐变, 因此首先将连续的候选子段送入渐变检测模块. 但是实际中不排除因为渐变过程短导致独立候选子段中也存在渐变, 但是这类独立子段具有首尾帧直方图距离较大的特点, 因此若独立候选子段的首尾帧直方图距离满足 $D_i > 2 \times T_i$ 时, 则将该独立候选子段也送入渐变检测模块.

2.3 渐变检测模块

经过切变检测后, 进入渐变检测模块的候选段主要是含有渐变或者是含有镜头或物体的移动. 由于绝大部分的渐变是溶解和淡入淡出, 因此本文主要考虑这两种渐变的检测.

溶解过程是指前一镜头的画面逐渐减弱的同时后一镜头的画面逐渐增强, 整个过渡过程在颜色直方图分布上表现为直方图的分布区域由窄变宽再变窄, 即完成从上一镜头结束帧的颜色直方图逐渐过渡到下一镜头起始帧的颜色直方图. 淡出过程是指前一镜头的画面逐渐变淡消失最终变为全黑帧, 整个过渡过程在颜色直方图分布上表现为直方图的分布区域逐渐变窄, 同时向零点移动, 最终只在零点处的取值不为零; 淡入过程直方图的分布变化与淡出过程相反. 通过以上分析可知, 渐变过程直方图分布的变化都具有一定的规律, 由于熵可以用来体现分布的特性, 一般当直方图分布区域较大时, 直方图的熵也较大, 反之, 直方图的熵则较小, 因此本文利用熵来体现颜色直方图的分布情况. 溶解过程的颜色直方图的熵会出现一个由小变大再变小的过程; 淡出淡入过程颜色直方图的熵会出现一个由大变小再变大的过程. 但是, 镜头或物体的移动过程的颜色直方图的分布变化则比较混乱, 即移动过程的颜色直方图的熵波动性较大, 不具有规律.

为了使渐变过程的颜色直方图分布变化规律更

为显著, 并且提高检测速度, 本文首先将连续的候选子段合并为一个候选片段, 再以一定步长对候选片段进行采样, 计算采样后帧的色度分量直方图的熵, 不同过程的采样帧的直方图熵变化曲线以及其对应的相邻帧直方图熵差情况如图 3 所示. 从图中可以看出, 溶解和淡出淡入过程的直方图熵变化情况与上文的分析一致, 并且变化幅度较大, 变化曲线较为光滑; 镜头或物体的移动过程的直方图熵变化则较混乱, 不规律, 由此会导致直方图的熵差出现如图 3 (f) 和 (h) 中标记的异常点, 即该点的直方图熵的增减性和相邻两个点的增减性都不一致. 因此, 本文提出了一种结合直方图熵差和异常点的检测方法, 用来辨别渐变和镜头或物体的移动. 在实际情况中, 不排除渐变过程的直方图熵的变化也会出现抖动, 导致存在异常点, 但是相对于直方图熵的变化量, 其抖动较小.

为了进一步的提高渐变检测的准确性, 本文在进行渐变镜头转换的判定条件中加入了对于候选片段首尾帧直方图距离的检测. 具体的渐变检测步骤如下:

step1 合并连续的候选子段为一个候选片段;

step2 以步长 L 对候选片段进行采样, L 等于合并的子段个数;

step3 计算采样后视频帧色度分量直方图熵差的均值:

$$H_r = \frac{\sum_{g=1}^{10} |h_{g+1,r} - h_{g,r}|}{10} \quad (7)$$

式中, H_r 为合并后的第 r 个候选片段的色度分量直方图熵差绝对值的均值, $h_{g,r}$ 为第 r 个候选片段中第 g 个采样帧的色度分量直方图熵;

step4 计算候选片段中直方图熵差的最大异常点:

$$U_r = \max(|u_g|) \quad (8)$$

$$\text{s. t. } u_g u_{g-1} < 0 \text{ 且 } u_g u_{g+1} < 0$$

式中, U_r 为第 r 个候选片段中的最大异常点, u_g 为候选片段中第 g 个直方图熵差, $g = 1, 2, \dots, 10$. 若候选片段中不存在异常点, 则 $U_r = 0$;

step5 计算候选片段首尾帧的 H 分量直方图距离:

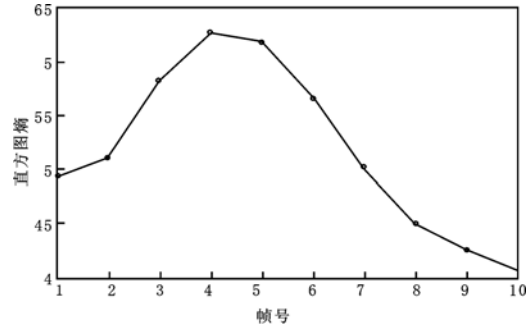
$$S_r = \sum_{k=1}^N |b_{k,10 \times L+1,r} - b_{k,1,r}| \quad (9)$$

式中, S_r 为第 r 个候选片段首尾帧直方图距离, $b_{k,1,r}$ 和 $b_{k,10 \times L+1,r}$ 分别为第 r 个候选片段的第一帧和最后

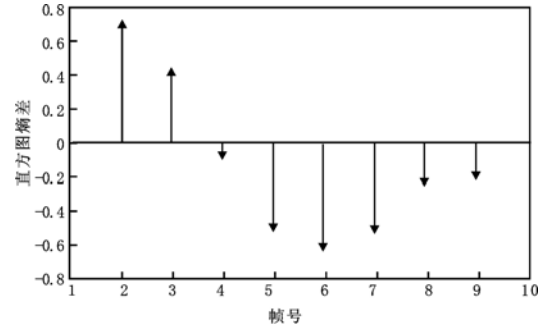
一帧的色度分量归一化直方图第 k 级的值;

step6 若满足 $U_r < \alpha H_r$ 且 $S_r > \beta$, 则判断该候

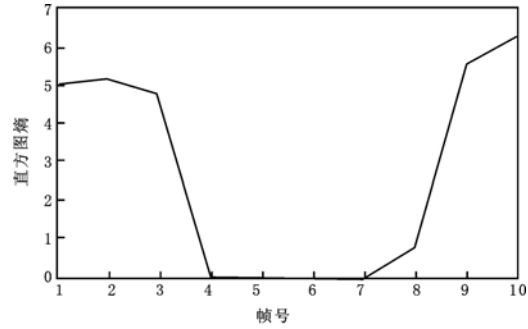
选片段中存在渐变. 本文中, 取 $\alpha = 0.8, \beta = 0.6$.



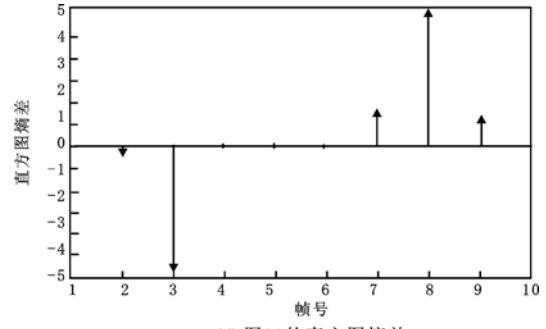
(a) 溶解过程直方图熵变化曲线



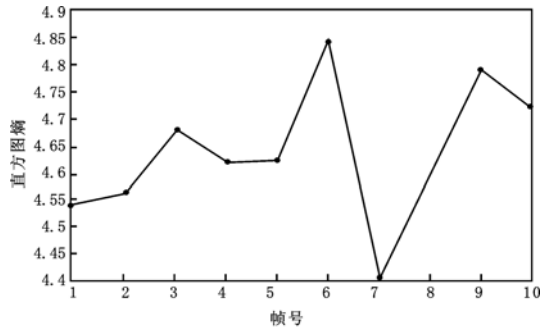
(b) 图(a)的直方图熵差



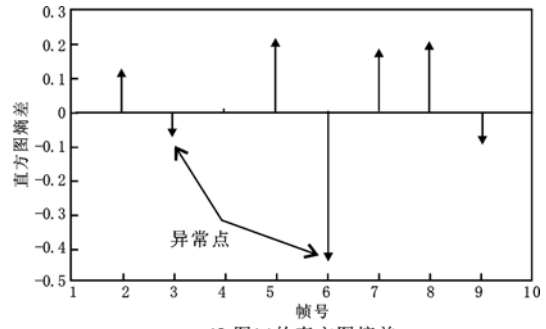
(c) 淡出淡入过程直方图熵变化曲线



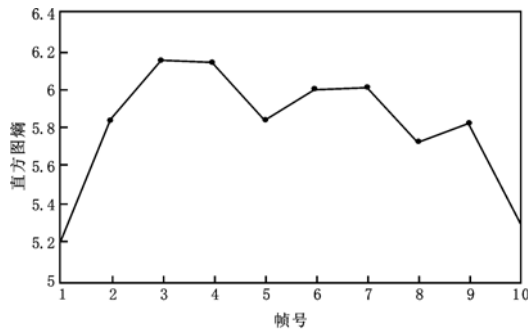
(d) 图(c)的直方图熵差



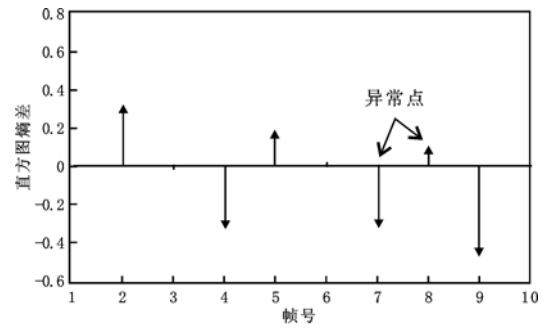
(e) 镜头移动过程直方图熵变化曲线



(f) 图(e)的直方图熵差



(g) 物体大幅度移动过程直方图熵变化曲线



(h) 图(g)的直方图熵差

图3 不同过程直方图熵及直方图熵差示意图

3 实验结果和分析

本文所有实验在 mablab2012 中通过编程实现, 实验所用的测试视频来自 TRECVID 视频库. 选用

查全率和查准率^[5]作为实验结果的评估标准, 它们的定义如下:

$$R_r = \frac{N_c}{N_c + N_u} \times 100\% \quad (10)$$

$$R_p = \frac{N_c}{N_c + N_f} \times 100\% \quad (11)$$

式中, N_c 为正确检测到的镜头边界数目, N_u 为漏检的镜头边界数目, N_f 为误检的镜头边界数目。

实验结果表明,通过预处理模块后,过滤掉的不包含镜头边界的视频帧占原视频长度的 76.31%,即由于预处理的引入,过滤掉了大量不包含镜头边界的视频片段,大幅度提高了镜头边界检测速度。

表 1 所示为本文提出的算法的检测结果。由表 1 可知,本文对镜头切变和渐变都具有良好的检测效果,尤其是渐变检测就目前该领域的研究成果而言,在查全率和查准率方面的效果都是令人满意的。

由于大部分的视频边界检测算法没有将视频的

切变和渐变分开检测,为了检验本文算法分别在切变和渐变检测上的有效性,本文编码实现了文献[6]中提出的基于双阈值的镜头分割算法,利用本文的测试视频进行测试,对比结果如表 1 所示。从实验结果可以看出,本文算法在切变检测的查全率和查准率上分别有 9.09% 和 6.38% 的提高,在渐变检测的查全率和查准率上分别有 15.29% 和 42.68% 的提高。

本文算法相对于文献[6]的渐变镜头查准率有显著的提高,是因为经过预处理模块后,绝大部分的非渐变片段被滤除,减小了其对渐变检测的干扰,同时结合直方图熵差和熵差异异常点的算法能有效的体现渐变的特性,提高准确率。

表 1 仿真及对比实验结果

检测结果	本文方法			双阈值法		
	切变	渐变	整体	切变	渐变	整体
镜头总数	231	85	316	231	85	316
N_c	226	79	305	205	66	271
N_u	5	6	11	26	19	45
N_f	1	3	4	15	57	72
$R_r/\%$	97.84	92.94	96.52	88.75	77.65	85.76
$R_p/\%$	99.56	96.34	98.71	93.18	53.66	79.01

4 结束语

本文利用切变和渐变过程在直方图变化上的不同规律,提出了一个包含预处理模块,切变检测模块和渐变检测模块的视频边界检测方法。仿真实验证明,由于预处理筛选过滤掉了不含有边界的视频帧,并且本文采用的直方图特征具有提取简单、快速的优点,因此该算法在保证检测准确率的同时,有效地降低了算法的计算复杂度,提高计算速度。

参考文献:

- [1] Li S, Lee M C. Effective detection of various wipe transitions[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2007, 17(6): 663-673.
- [2] Mohanta P P, Saha S K, Chanda B. A model-based shot boundary detection technique using frame transition parameters[J]. IEEE Transactions on Multimedia, 2012, 14(1):223-233.
- [3] Li Y N, Lu Z M, Niu X M. Fast video shot boundary

detection framework employing pre-processing techniques[J]. IEEE Image Process, 2009, 3(3):121-134.

- [4] Gao Guangyu, Ma Huadong. Accelerating shot boundary detection by reducing spatial and temporal redundant information[C]// IEEE International Conference on Multimedia and Expo. Piscataway, NJ, USA: IEEE, 2011:1-6.
- [5] Boreczky J S, Row L A. Comparison of video shot boundary detection techniques[J]. Journal of Electronic Imaging, 1996, 5(2):122-128.
- [6] Zhang Hongjiang, Wu Jianhua, Zhong Di, et al. An integrated system for content-based video retrieval and browsing[J]. Pattern Recognition, 1997, 30(4): 643-658.

作者简介:

俞璐 女,(1987—),硕士研究生.研究方向为视频检索。
乔瑞萍 女,(1966—),副教授.研究方向为图像视频处理及其 DSP 实现。
胡宇平 男,(1987—),硕士研究生.研究方向为图像识别。
赵静 女,(1980—),硕士研究生.研究方向为图像复原。