

基于颜色空间的自适应阈值镜头分割算法

汪翔^{1,2}, 罗斌^{1,2}, 翟素兰^{1,3}, 涂铮铮^{1,2}

(1. 安徽大学 计算智能与信号处理教育部重点实验室, 安徽 合肥 230039;

2. 安徽大学 计算机科学与技术学院, 安徽 合肥 230039;

3. 安徽大学 数学科学学院, 安徽 合肥 230039)

摘要:镜头分割是基于内容的视频检索的关键步骤,它会直接影响到视频检索的效果。文中介绍了几种常用的镜头分割方法,并且根据YUV颜色空间的各分量可分离的特点,提出了一种改进的基于YUV颜色空间的自适应阈值镜头分割方法。在突变镜头检测模块,为了消除由噪声带来的误检,加入了帧差值比法。在渐变镜头检测模块,使用了滑动窗口值方法。镜头分割的难点是渐变检测,算法在渐变检测上也取得了不错的成果。经过大量实验结果表明,改进的算法对镜头分割有很好的实验效果,算法计算复杂度低,易于实现。

关键词:镜头分割;YUV颜色空间;自适应阈值

中图分类号:TP301.6

文献标识码:A

文章编号:1673-629X(2012)09-0037-04

Self-threshold Shot Segmentation Based on Color Space

WANG Xiang^{1,2}, LUO Bin^{1,2}, ZHAI Su-lan^{1,3}, TU Zheng-zheng^{1,2}

(1. Key Laboratory of Intelligent Computing & Signal Processing of Ministry of Education, Hefei 230039, China;

2. School of Computer Science & Technology, Anhui University, Hefei 230039, China;

3. School of Mathematical Sciences, Anhui University, Hefei 230039, China)

Abstract: Video shot segmentation is the key procedure of content-based video retrieval (CBVR), it affects the results of video retrieval directly. It introduces some commonly used shot segmentation algorithms, presents a shot segmentation approach based on YUV color space while each component of YUV color space can separate. The ratio of interval frames is used to eliminate some misdetection brought by noise. The continuous window value method is used to detect the gradual change happen in the gradual change detection module. The difficult of shot segmentation is the gradual change, the approach has good results on it. As proved by the experimental results, the approach is brief and effective for shot boundary detection.

Key words: shot segmentation; YUV color space; adaptive threshold

0 引言

近几十年来,随着网络和计算机多媒体技术的发展,以及大量的手持摄像设备的普及,各种多媒体数据不断涌入人们日常的生活中,而视频数据,以其特有的数据结构特点,占据了多媒体数据的一大部分。人们在享受视频数据带来的方便、快捷的同时,也面临着如何有效管理组织、检索视频数据的问题。传统的基于文本标注的检索方法已经满足不了人们的需求,广大科研人员逐渐将精力转向了基于内容的视频检索算法

的研究当中,而镜头分割^[1]是视频结构化的第一步。

镜头分割算法^[2]目前主要有两大类,一类是基于非压缩域,另一类是基于压缩的。

非压缩域的镜头分割主要有:

(1)基于像素的方法^[3],它的缺点是受噪声,视频帧内大物体的运动,摄像设备的移动非常敏感,因为它完全取决于像素空间的位置,物体或拍摄镜头的一切运动都会使得相邻帧之间的差值突然增大,从而造成误检(False detection),它的优点是算法简单,容易实现。

(2)基于直方图的方法。它首先统计的是直方图的信息,不用考虑绝对的像素位置,所以可以容忍物体运动所造成的干扰。缺点是没有反应图像的整体内容,当镜头内容变化过快容易造成误检或漏检(Miss detection)。它的优点是降低了对拍摄镜头和镜头内大目标的物体运动以及噪声的敏感性。

收稿日期:2012-01-10;修回日期:2012-04-21

基金项目:国家自然科学基金资助项目(60772122);安徽省教育自然科学基金资助重大项目(KJ2011ZD10);博士研究生队伍建设计划(02203105)

作者简介:汪翔(1986-),男,硕士研究生,研究方向为图像处理和模式识别;罗斌,教授,博导,研究方向为图像处理与模式识别。

(3) 基于轮廓的方法, 对于视频帧内容比较简单的视频, 以及在渐变镜头的检测方面效果很好, 缺点是为了抵消运动所带来的影响, 需要使用图像配准 (Image registration) 技术来进行运动补偿, 因此使得计算量过大, 并且当帧内物体快速运动时容易造成误检。

基于压缩域的镜头检测方法主要有 Yeo 和 Liu 提出的一种利用 MPEG 压缩视频的 DC 序列进行镜头边界检测的方法^[4]。

文中提出一种基于 YUV 颜色空间的自适应镜头分割算法, 选用 Y 分量分块直方图来计算相邻帧的帧差值, 计算出自适应阈值, 根据阈值判断进入哪个模块, 在切变镜头的判定模块中, 加入了隔帧帧差值比法来鉴定是否受到闪光灯影响。在渐变模块中加入了连续窗口值来判断是否发生了渐变, 解决了镜头分割中的一大难点。

1 镜头分割算法

YUV 颜色空间^[4]是最常用的颜色空间, 它也可以直接通过压缩视频获取, 不需要经过解压缩。其中: Y 是指亮度 (Brightness), U 和 V 是指色度 (Chrominance)。

YUV 颜色空间可以通过下述公式由 RGB 空间转换得到:

$$\begin{pmatrix} Y \\ U \\ V \end{pmatrix} = \begin{pmatrix} 0.3 & 0.59 & 0.11 \\ -0.15 & -0.29 & 0.44 \\ 0.51 & -0.52 & -0.095 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

通常在同一视频场景内, 色度基本是一致的单一的, 即便是不在一个镜头内的相邻帧, 它们的色度也相差不大, 从而可以用 Y 分量作为判断镜头是否突变的主要依据, 在 YUV 颜色空间中, Y 和 U, V 分量是可以分离的, 这减少了大量的运算量。

1.1 相邻帧帧差值的计算

由于在统计直方图时, 失去了视频帧原有的像素点位置信息, 难以反应视频的空间信息, 所以采用分块直方图^[5]作为解决方案。分块示意如图 1 所示, 为了解决屏幕顶部经常有观众看台区域, 以及底部经常出

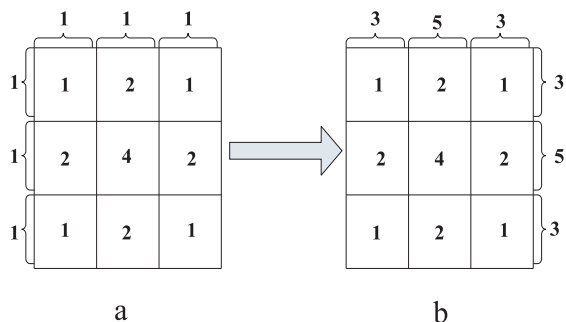


图 1 分块加权示意图

现的广告内容或字幕的变化对镜头边界检测的干扰, 通常视频帧的主要内容显示在图像的中间块位置, 所以改进了文献[5]中的均匀分块的方法, 将所有图像帧不均匀分成为 $M \times N$ (文中为 3×3) 大小的子块, 并修改了相应的加权矩阵。

计算相邻帧对应子块的直方图差值为:

$$\text{Minus_frame}_i = |Y_value_{k+1} - Y_value_k| \quad (1)$$

这里的 Y_value_k 为当前帧子块的 Y 分量直方图, Y_value_{k+1} 为下一帧对应子块的 Y 分量直方图。由图 1 我们可以得到每个子块的权值分别对应的加权矩阵 W 为:

$$W = \begin{pmatrix} w_1 & w_2 & w_3 \\ w_4 & w_5 & w_6 \\ w_7 & w_8 & w_9 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix}$$

由公式(1)可以计算得到相邻帧对应子块的差值, 因此相邻两帧之间的 Y 分量的直方图差值 $\text{FramDiff}_{k,k+1}$ 为公式(2)所示。

$$\text{FramDiff}_{k,k+1} = \frac{\sum_{i=1}^9 \text{Minus_frame}_i * w_i}{\sum_{i=1}^9 w_i} \quad (k \text{ 为帧号}) \quad (2)$$

1.2 自适应阈值的获取

由于同一镜头内相邻的两帧在表示的内容上是一致的, 所以它们之间的直方图差值是非常小的, 而发生镜头突变时, Y 分量直方图帧差值则明显要比同一镜头内平均帧差值大很多。如图 2 所示, (a) 中几个突变处显然比同一镜头内其他帧差值大很多, 可以利用这点作为镜头发生突变转换的依据。而镜头的渐变是通过亮度的不断变化来实现的, 出现在连续的多帧内, 且相邻帧差值也会发生较大的变化, 因此, 可以设立两个自适应阈值^[6]来分别进行判断。

计算当前检测镜头内帧差值总和 FramDiffAll 为

$$\text{FramDiffAll} = \sum_{k=1}^{n-2} \text{FramDiff}_{k,k+1} \quad (3)$$

这里的 i 的初始为当前镜头内的第一帧, n 为镜头首帧到当前检测帧的所有帧数。则当前检测镜头内的平均帧差值为

$$\text{AvgFramDiff} = \frac{\text{FramDiffAll}}{n-1} \quad (4)$$

突变镜头检测的自适应阈值 Th_a 为

$$Th_a = \alpha * \text{AvgFramDiff} \quad (5)$$

渐变镜头检测的自适应阈值为

$$Th_b = \beta * \text{AvgFramDiff}$$

这里的 α 和 β 是非常重要的, 对镜头检测准确度有着很大的影响, 经过大量的实验比对, 这里的 α 取

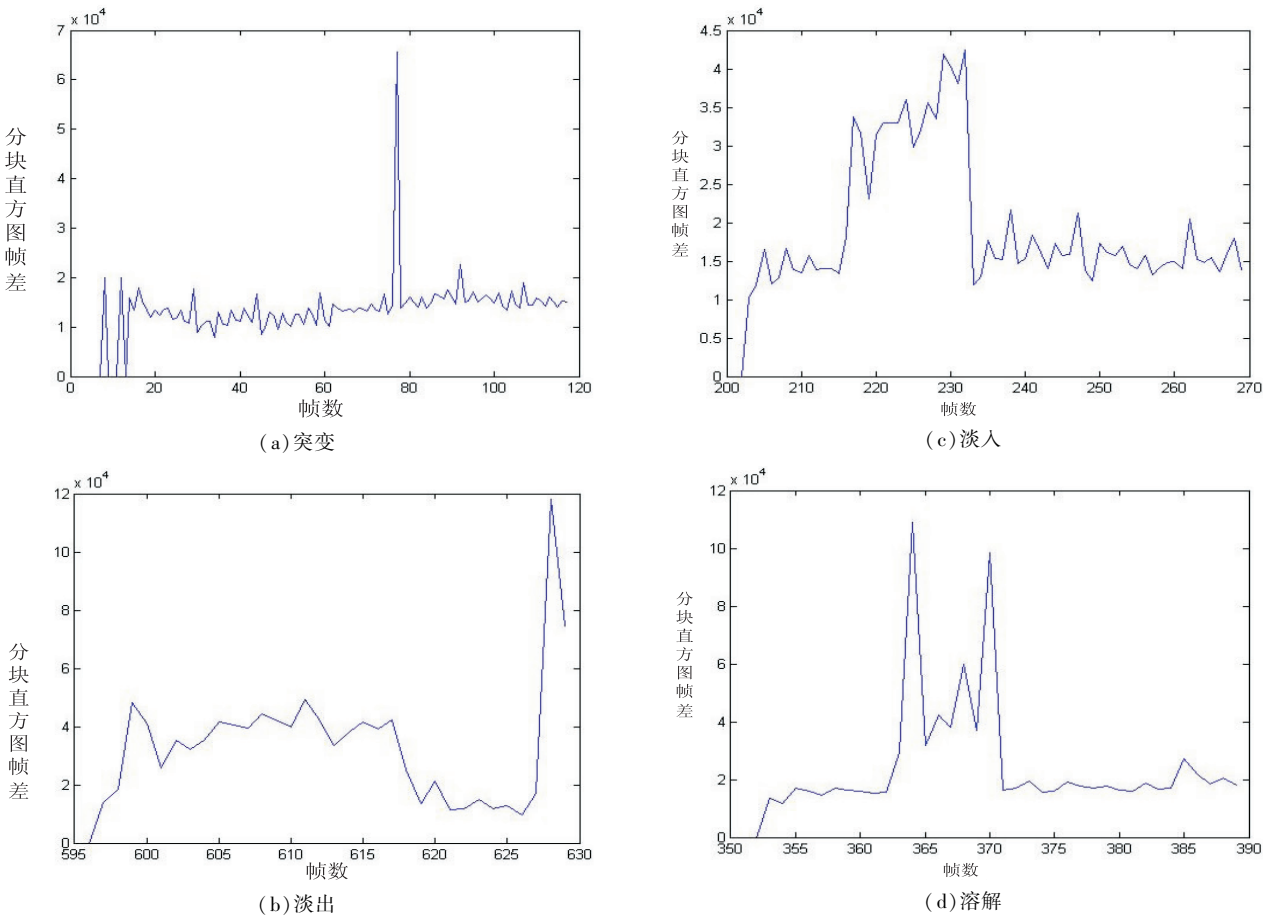


图2 相邻两帧Y分量分块直方图帧差值

2.5 至 4.0 之间, β 取 1.5 至 2 比较合适^[7]。

1.3 模型选择器

通过计算相邻帧的Y分量分块直方图帧差值和自适应阈值进行比较,判断是进入突变检测模块,还是进入渐变检测模块。

- (1) If() 则当前相邻的两帧属于同一镜头。
- (2) If() 则检测流程进入突变检测模块。
- (3) If() 则检测流程进入渐变检测模块。

1.4 突变检测模块 (闪光检测模块)

视频同一镜头内如果突然受到闪光影响会使相邻帧的Y分量分块直方图帧差值^[8]像检测到突变镜头一样,帧差值变的很大,因此,算法很容易受到光照的影响。统计得出,在发生光照影响时,前后的帧差值相差不大,而当发生镜头突然转换时,突变的前后帧差值差别是很大的,可以通过计算隔帧图像的帧差值来区分闪光灯类和突变类。

当检测到发生镜头发生突变时,计算当前位置 k 的前 j 帧(第 $k-j$ 帧) 和后 k 帧(第 $k+j$ 帧) 的帧差值 $FramDiff_{k-j,k+j}$,并将其与当前相邻帧的差值进行比较,二者得比值 R 为

$$R = \frac{FramDiff_{k-j,k+j+1}}{FramDiff_{k,k+1}}$$

(6)

如果 R 近似于 1 的话,那么认为当前发生了镜头突变,将当前帧加入镜头边界集合,如果不是那么认为当前受到了光照影响。由于闪光灯影响可能持续数帧,为此可以适当地增大步长 j 的值^[9],使得相隔 $2j+1$ 帧正好跳过闪光灯的影响。

1.5 镜头的渐变检测模块

由于镜头渐变^[10]是发生在连续的多帧内,因此算法设定一个渐变帧计数器 Fc ,初始值为零。当出现连续的 6 帧相邻帧差值^[11]在这个区间内,使得 Fc 的值等于 6 时,就认定此处发生了镜头的渐变,将当前帧加入镜头边界集合,在 Fc 小于 6 之前,检测帧差值出现了大于突变阈值和小于渐变阈值的情况下,将 Fc 清零,进入突变模块或下一帧检测。

1.6 算法具体步骤

假设一个长度为 N 的视频序列为 $V = \{f_1, f_2, \dots, f_n\}$, 输出镜头边界集合为 Ω 。

(1)输入视频序列 V 后,根据每一帧的 RGB 颜色空间获取 YUV 颜色空间,分离 YUV 颜色空间得到 Y 分量。

(2)计算相邻两帧的Y分量分块直方图帧差值 $FramDiff_{k,k+1}$ 。

(3)根据 $FramDiff_{k,k+1}$ 的值计算相应的自适应阈

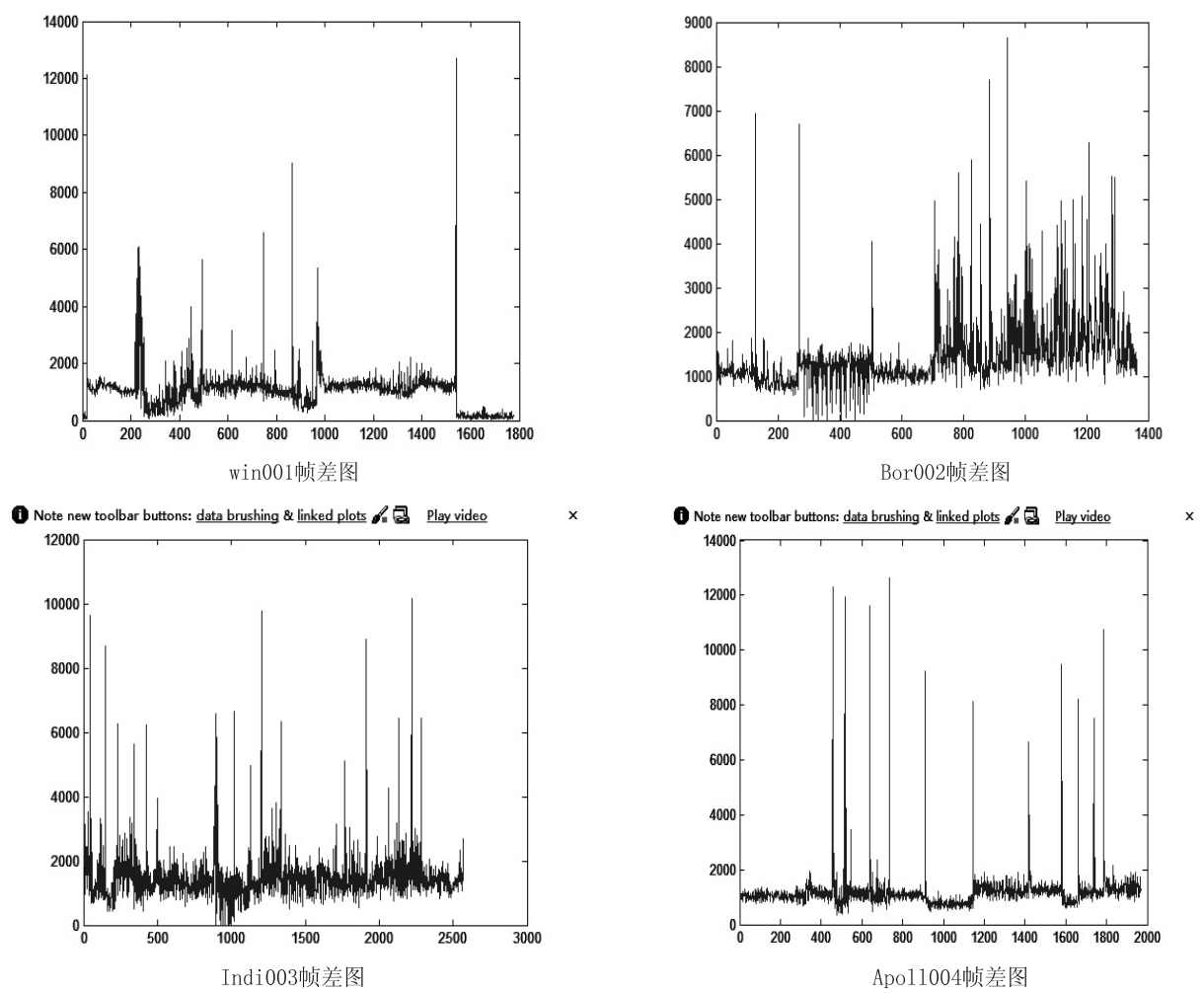


图 3 视频 1-4 在镜头分割中的帧差图

值。

(4)如果帧差值 $FramDiff_{k,k+1}$ 大于突变自适应阈值,转入突变检测模块,检测到没有受到闪光灯影响,那么将当前帧加入到 Ω ,否则转入步骤(6)。

(5)如果帧差值 $FramDiff_{k,k+1}$ 介于突变阈值和渐变阈值之间的话,转入渐变帧检测模块,是渐变镜头边界的则将当前帧加入 Ω ,否则根据帧差值转入步骤(4)或(6)。

(6)如果当前帧不是本段视频最后一帧的话,进入下一帧检测,直至结束。

2 实验结果分析

为了验证文中提出算法的有效性,选取了 4 个来自 OpenVideo 的视频片段作为实验素材,以 MatlabR2009a 为实验平台,进行了实验,并且和田玉民,吴自力提出的基于帧间联合直方图帧差的镜头边界检测算法^[12]做了对比。基于 YUV 颜色空间的自适应阈值突变检测过程中视频 1-4 的 Y 分量直方图帧差值如图 3 所示。

文中采用两个常用的指标查全率、查准率来评价

镜头的检测效果,具体的实验结果如表 1 所示。

$$\text{查全率} = \frac{\text{正确检出数}}{\text{正确检出数} + \text{漏检数}}$$
$$\text{查准率} = \frac{\text{正确检出数}}{\text{正确检出数} + \text{错误检出数}}$$

表 1 实验具体结果

视频名称	总帧数	镜头转换数	文中方法		对比方法	
			查全率	查准率	查全率	查准率
Win001	1778	10	97	88.9	87.5	84.5
Bor002	1365	14	92.3	91.6	81.8	75
Indi003	2572	17	89.5	87.8	80	66.7
Apo11004	1970	11	91.5	89.5	82.8	70

从实验结果来看,文中提出的算法检测效果要明显优于基于帧间联合直方图帧差的镜头边界检测算法,在查全率和查准率都优于对比算法,而且算法在时间复杂度上也是优于对比算法。

3 结束语

文中计算相邻帧的 Y 分量分块直方图帧差与双阈值的比较来判断突变、渐变以及闪光情况,其中分块

AgentTopDisc 进行拓扑发现消耗的能量为 1310.6mj, flooding 算法消耗的能量为 1924.9mj。因此,与 flooding 算法相比,AgentTopDisc 能够节省 30% 的能耗。因而 AgentTopDisc 算法在节能方面有较大优势。

3 结束语

为了解决无线传感器网络拓扑结构发现过程中,节点拓扑请求信息过多导致过多能量消耗问题,提出了 AgentTopDisc 算法。算法利用移动 Agent 的移动和计算特性,迭代进行网络拓扑发现:最开始以 Sink 节点为簇头,发现簇头一跳范围内的邻居中剩余能量最大的节点作为下一轮的簇头进行网络拓扑发现。最后由这些簇头连接形成网络拓扑。实验结果表明,相对于目前的 flooding 算法,AgentTopDisc 算法能够节约 30% 的能量,同时具有更好的网络连通性。

参考文献:

- [1] 孙利民,李建中,陈渝. 无线传感器网络[M]. 北京:清华大学出版社,2005.
- [2] Zhao Mingchen, Lei Jiayin, Wu Minyou, et al. Surface Coverage in Wireless Sensor Networks[C]//Proceedings of the IEEE INFOCOM 2009. [s. l.]: [s. n.], 2009:109-117.
- [3] 蒋鹏. 基于无线传感器网络的湿地水环境远程实时监测系统关键技术研究[J]. 传感技术学报, 2007, 20(1):183-186.
- [4] 杨树森,周小佳,阎斌. 无线传感器网络在环境监测中的

应用[J]. 计算机技术与发展, 2008, 18(9):170-172.

- [5] Paul J L. Smart sensor web; web-based exploitation of sensor fusion for visualization of the tactical battlefield[J]. IEEE Aerospace and Electronic Systems Magazine, 2001, 16(5):29-36.
- [6] Polastre J. Design and implementation of wireless sensor networks for habitat monitoring[D]. Berkeley: University of California, 2003.
- [7] Ni S, Tseng Y, Chen Y, et al. The broadcasting storm problem in a mobile Ad Hoc network[C]//ACM MOBICom'99. [s. l.]: [s. n.], 1999.
- [8] Deb B, Bhatnagar S, Nath B. A topology discovery algorithm for sensor networks with applications to network management[C]//Proceedings of the IEEE CAS Workshop on Wireless Communications and Networking. Pasadena, USA: [s. n.], 2002.
- [9] 陈晶. 基于 Agent 的无线传感器网 QoS 管理机制研究[J]. 计算机技术与发展, 2008, 18(11):230-232.
- [10] Li Yongzhong, Xu Jing, Zhao Bo, et al. A new mobile agent architecture for wireless sensor networks[C]//The 3rd IEEE Conference on Industrial Electronics and Applications. [s. l.]: [s. n.], 2008:1562-1565.
- [11] 张宇晴,郑小建,胡旦华. 无线传感器网络中基于 Agent 的高效路由算法的研究[J]. 计算机技术与发展, 2007, 17(9):120-122.
- [12] 卞永钊,于海斌,曾鹏. 无线传感器网络中的拓扑控制[J]. 计算机应用研究, 2008, 25(10):3128-3133.

(上接第 40 页)

上采用的是不均匀分块考虑了视频帧内容分布不均的情况,两个阈值是根据两个系数和镜头内平均帧差值的乘积,体现了自适应性,闪光检测模块用隔帧帧差法来判断,在渐变帧检测模块使用窗口计数器来检测镜头的渐变。文中的算法在镜头边界的检测方面取得了不错的效果,并且计算复杂度较低。

参考文献:

- [1] 钱刚,曾贵华. 典型视频镜头分割方法的比较[J]. 计算机工程与应用, 2004, 40(32):51-55.
- [2] 金红,周源华. 基于内容检索的视频处理技术[J]. 中国图象图形学报, 2000, 5(4):276-283.
- [3] Li S, Lee M. Effective detection of various wipe transitions[J]. IEEE Trans. on Circuits Syst. Video Technol, 2007, 17(6):663-673.
- [4] 章毓晋. 中国图像工程及当前的几个研究热点[J]. 计算机辅助设计与图形学学报, 2002, 14(6):489-500.
- [5] 印勇,侯海珍. 基于直方图帧差的自适应镜头分割算法[J]. 计算机工程与应用, 2010, 46(9):186-189.

- [6] 丁洪丽,陈怀新. 基于累积直方图的视频镜头边界检测方法[J]. 电讯技术, 2008, 48(3):65-69.
- [7] 潘磊,吴小俊,尤媛媛. 基于聚类的视频镜头分割和关键帧提取[J]. 红外与激光工程, 2005, 34(3):341-344.
- [8] Qu Zhongshui, Wang Jianwei. A Color YUV Image Edge Detection Method Based on Histogram Equalization Transformation[C]//2010 Sixth International Conference on Natural Computation. [s. l.]: [s. n.], 2010:3546-3549.
- [9] Yuan J, Wang H, Xiao L. A formal study of shot boundary detection[J]. IEEE Trans. on Circuits Syst. Video Technol, 2007, 17(2):168-186.
- [10] Cotsaces C, Nikolaidis N, Pitas I. Video shot detection and condensed representation: a review[J]. IEEE Signal Processing Magazine, 2006, 23(2):28-34.
- [11] Zabih R, Miller J, Mai K. A feature-based algorithm for detecting and classifying scene breaks[C]//Proc of ACM Multimedia. San Francisco, CA: [s. n.], 1995:189-200.
- [12] 田玉民,吴自力. 基于帧间联合直方图帧差的镜头边界检测算法[J]. 计算机应用, 2008, 21(10):49-50.