

篮球赛事视频镜头分割

夏磊,张明

(上海海事大学 信息工程学院,上海 201306)

摘要:针对篮球赛事视频分割问题,提出一种基于场地三分线的新分割算法。首先提取出帧图像的 Canny 边界,再对提取到的结果进行优化去杂;接着在优化过的边界中搜寻三分线并记录下三分线上点的坐标;然后根据这些点的坐标来判断三分线的弯曲方向,将其作为后续分割的依据。基于以上算法,以 Microsoft Visual Studio 2010 与 OpenCV2.3 为开发平台,实现了一个篮球赛事视频镜头分割的原型系统,并通过该系统验证了新方法在正确检出和漏检方面的优势。虽然不是很明显,但是在误检方面的优势很显然。

关键词:镜头分割;Canny 边界;三分线;弯曲方向

中图分类号:TP37

文献标识码:A

文章编号:1673-629X(2015)02-0216-05

doi:10.3969/j.issn.1673-629X.2015.02.049

Shot Segmentation of Basketball Tournament Video

XIA Lei,ZHANG Ming

(College of Information Engineering,Shanghai Maritime University,
Shanghai 201306,China)

Abstract:In view of basketball tournament video segmentation,propose a new segmentation algorithm based on the three-point line. First,the Canny boundary of frame image is extracted and optimized,and then the three-point line has been searched with the coordinates of those points recorded which determine the bending direction of the three-point line and could be used as the basis for the subsequent split. By the development platform of Microsoft Visual Studio 2010 and OpenCV2.3,a prototype system of basketball tournament video shot segmentation has been achieved,which verifies that the new method's advantage is not obvious in terms of being correctly detected and undetected,but it's notable in terms of false detection.

Key words:shot segmentation;Canny border;three-point line;bending direction

0 引言

全世界每天都有各种各样的篮球赛事举行,录制的赛事视频也非常多,而其中美国的 NBA 篮球赛事更是牵动着全世界篮球爱好者的心的,每年的常规赛和季后赛,几乎都是看点频频,吸引着全世界的观众,甚至让有些球迷都达到了废寝忘食的程度。但众多球迷们也是要工作和生活学习的,没有那么多的宝贵时间去观看每一场球赛,所以这里就需要做一些筛选工作,尽量选取一些在大家看来都很精彩的片段以供球迷更好地欣赏。传统的人工筛选在一定程度上满足了球迷们的这种需求,例如各种视频集锦之类的,但是这种方式往往工作量会很大,而且由于各人的偏好不一样,存在很大程度上的主观性,每个人的理解和需求可能都会

不一样,所以人们在试图研究一种机制去自动实现篮球赛事视频的精彩镜头提取,摆脱对人工的高度依赖性,尽量为每一个球迷提供个性化的服务。而这所有的努力都必须从视频镜头分割开始,镜头分割的好坏直接影响着后续研究的成功与否^[1],这是至关重要的一步,所以文中着重研究篮球赛事视频的镜头分割技术,以期后续研究作一些铺垫,从而让更多的篮球视频研究工作者从中受益。

视频镜头分割技术的研究也有着比较早的历史,产生了一些比较经典的分割算法,大致分为以下两种:一是基于视频底层特征的方法,如 A. Ekin 等利用统计篮球比赛视频图像中球场主要颜色分布率,将比赛视频分割为比赛和暂停的视频片段^[2];二是视频基本特

收稿日期:2014-03-13

修回日期:2014-06-16

网络出版时间:2014-12-27

基金项目:上海市自然科学基金资助项目(11ZR1415200)

作者简介:夏磊(1990-),男,研究方向为信息处理与模式识别;张明,教授,博士,研究方向为多媒体应用、多媒体数据库、多媒体信息检索与处理、信息安全、物流、航运信息化。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20141227.1343.030.html>

征和辅助信息相结合的方法,如 Kyungsu 等提出基于篮球比赛中的音视频特征来抽取语义信息^[3]。而主流的篮球赛事视频的分割方法大多是基于颜色直方图的,利用像素的统计特征进行分析。在众多的已有分割方法中,有一种分块的颜色直方图方法^[4]比较适合于篮球视频的镜头分割,因为篮球场地上颜色比较统一和集中,禁区和场地其余部分就是两种不同的颜色,而且其他的杂色也比较少,比较适合于这种方法。但是仔细研究发现,假如按分块的颜色直方图方法来划分图像区域,其颜色直方图的统计结果还是很受相机抖动和闪光灯的影响,而这又是体育赛事视频不可避免的问题。所以不得不致力于开发一种新的分割方法来更好地实现篮球视频的镜头分割。

鉴于已有各种方法都不能很好地对篮球赛事视频进行镜头分割,文中在研究前人工作成果的基础之上,提出了一种新的基于三分线和禁区线的篮球赛事视频分割方法,文中主要讲述该方法的思路和系统实现。

1 对单个图像帧的处理

从视频流中提取帧图像之后,首先该方法对单个帧图像的处理大致分为以下几步:一是剪切源图像;二是提取 Canny 边界;三是该边界进行优化去杂;四是在优化的边界里搜寻曲线并记录曲线上点的坐标值;五是根据坐标值判断曲线的弯曲方向并返回曲线弯曲方向的类型。

1.1 剪切源图像

通过研究从视频流中提取的帧图像不难发现,每帧图像的顶端总有一些观众画面,而且往往由于观众衣服颜色或者肤色等因素的影响,提取的 Canny 边界会很杂乱,严重影响后续研究的进行。另外,这一部分又很难成为感兴趣区域,所以选择在进一步研究开始之前就对源图像进行必要的裁剪。如此一来,既可以大大减小计算量,又可以提高后续研究工作的精度。

首先,用 OTSU 算法(大津法)^[5]对提取的帧图像进行二值化;然后再用 Mean-Shift 算法^[6]对图像进行切割。实验结果表明,具体的裁剪尺度是保留源图像的底端 7/10 较为合适,如图 1 所示为剪切处理后的图像。



图 1 剪切后的图像

1.2 提取 Canny 边界

检测边缘的方法有很多,例如 Roberts 算子^[7]、Sobel 算子^[8]、Prewitt^[9]算子、Laplacian^[10]算子和 Canny 算子^[11]等。从本质上讲,Canny 算子不是只通过进行简单的梯度运算来决定像素是否为边缘点的,在决定一个像素是否为边缘点时,需要同时考虑其他像素的影响。它也不是进行简单的边界跟踪,在寻找边缘点时,需要根据当前和前面处理过的像素来进行判断。它把边缘检测问题转换成检测函数极大值的问题,其基本思想是:首先用高斯滤波器平滑图像;再用一阶偏导的有限差分来计算梯度的幅值和方向;接着对梯度幅值进行非极大值抑制;最后用双阈值算法^[12]检测和连接边缘。

如图 2 所示,就是从图 1 中提取出来的 Canny 边界。



图 2 提取到的 Canny 边界

1.3 对提取到的 Canny 边界优化去杂

很明显,图 2 中的 Canny 边界包含了目标边界:三分线和禁区线边界,这是满足要求的;但同时又带有很多的无关边界点,这对后续研究存在很大的干扰;所以这里要做的工作就是保留所需要的三分线和禁区线边界点,除去无关边界点。通过观察可以发现,所需要的三分线和禁区线边界都是双线形式的,而且双线之间有一个相对固定的距离,并且它们水平方向的前后都有较长一段距离是没有其他杂点的。基于这些特征,下面给出一种优化去杂的方法,根据图 2 的源数据来做分析,将新数据保存在一个新的结果图中。

从图 2 中第一个像素点开始,遍历到最后一个像素点结束。

Step1:判断当前点的像素值是否大于零,若大于则进入 Step2,否则左距离 d 加 1,并将结果图中对应点置为 0;

Step2:判断左距离 d 是否大于距离阈值 distance,若大于则进入 Step3,否则将左距离 d 重置为 0,并将结果图中对应点的值置为 0;

Step3:先将左距离 d 置为 0,再判断左边界非零点的个数是否小于 5,若小于则进入 Step4,否则将结果图中对应点到左边界最右点的值都置为 0,并将图 2

中的当前点设置为左边界最右点的下一点;

Step4:判断双线之间的零点个数是否小于 6 且大于 2,若是则进入 Step5,否则将结果图中对应点到零点序列最右点的值都置为 0,并将图 2 中的当前点设置为零点序列最右点的下一点;

Step5:判断右边界的非零点个数是否小于 5,若小于则进入 Step6,否则将结果图中对应点到右边界最右点的值都置为 0,并将图 2 中的当前点设置为右边界最右点的下一点;

Step6:判断后续零点的个数是否大于距离阈值 distance,若大于则将图 2 中当前点的值赋给结果图中当前点,再将当前点的下一点到后续零点序列最右点的值都置为 0,并将图 2 中当前点设置为后续零点序列最右点的下一点,最后将左距离 d 置为后续零点个数;否则将结果图中对应点到后续零点序列最右点的值都置为 0,并将图 2 中的当前点设置为后续零点序列最右点的下一点。

最后得到的结果图就是优化去杂后的 Canny 边界,记为 A 。

1.4 在优化的边界里搜寻曲线并记录曲线上点的坐标值

优化后的 Canny 边界是比较理想的,而且有两个特征:

- (1)三分线的整体造型类似抛物线;
- (2)三分线的边界点也都是相对集中而且比较长的,而其他的边界点大部分都是很零散的,即使有很少部分相对集中,它们的长度也比较短。

基于这两个特征,下文分别给出的两个方法都是力求准确地定位满足要求的边界曲线,并记录下边界曲线上点的坐标值以供后续使用。

1.4.1 改进型的抛物线 Hough 变换

根据 1.4 提出的特征(1),引进一种改进型的抛物线 Hough 变换^[13-15],用该方法来检测 A 中所包含的抛物线,如图 3 所示。

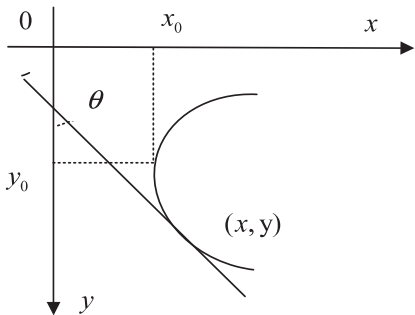


图 3 改进型的抛物线 Hough 变换

抛物线方程为

$$(x - x_0) = a(y - y_0)^2, a \neq 0 \quad (1)$$

求得得:

$$d_x/d_y = 2a(y - y_0) \quad (2)$$

取抛物线上任意一点 (x, y) ,则抛物线在该点的切线方向为 $d_x/d_y = 2a(y - y_0)$ 。设抛物线在该点的切线与 y 轴的夹角为 θ ,则有 $\tan\theta = d_x/d_y$,并由此可得:

$$\begin{cases} x_0 = x - \tan^2\theta/4a \\ y_0 = y - \tan\theta/2a \end{cases} \quad (3)$$

紧接着,改进型的抛物线 Hough 变换步骤如下:首先设置一个三维累加器数组 $A(a, x_0, y_0)$,对于图像中任一边缘点 (x, y) ,利用边缘梯度方向预测 θ 值,变化 a 值并由式(3)计算得到 (x_0, y_0) ,对累加器数组进行投票。如此遍历完所有边缘点后,寻找累加器的峰值点,即可获得抛物线的顶点和曲率。

1.4.2 扫描法

根据 1.4 提出的特征(2),采用一种扫描法:

Step1:先初始化一幅标记图像和一幅新的结果图像,大小都和 A 一样,再初始化一个三维零矩阵,用于记录边界曲线上点的坐标值,再从 A 的第一个像素点开始遍历整幅图像;

Step2:根据标记图像中对应点的值来判断 A 中当前点是否被记录过,如果是则遍历下一点,否则进入 Step3;

Step3:判断当前点的值是否大于 0,若是则进入 Step4,否则遍历下一点;

Step4:判断当前点的正下方 $n * m$ 区域内有无非零点,若有则纵向长度 length 加 1,并将当前点设置为该非零点,继续 Step5;否则退出 Step4,并将当前点恢复到最初的那个当前点位置,再进入 Step5;

Step5:判断从当前点开始的纵向长度是否大于长度阈值 lengththread,若大于则进入 Step6,否则遍历下一点;

Step6:判断当前点的正下方 $n * m$ 区域内有无非零点,若有则将当前点的值赋给新的结果图像中对应点,再用三维矩阵记录下该点的坐标值(第一维表示的是这幅图像中的第几条边界曲线,第二维表示的是当前边界曲线上的第几个点,第三维表示的是当前点的坐标),还要将标记图像中的对应点值重置为 1,最后将当前点设置为该非零点,继续 Step6;否则退出 Step6,并将当前点恢复到最初的那个当前点位置,并继续遍历下一点。

最后得到的结果图就是从 A 中提取出来的边界曲线图。

1.4.3 实验结果对比分析

两种方法在前述步骤提取的 Canny 边界效果比较好的情况下(例如,优化的 Canny 边界图里有比较完整

的抛物线形状的三分线边界),可以得到几乎相同的实验结果,没有明显的区别。但在前述步骤提取的 Canny 边界效果不是很好的情况下(例如,优化的 Canny 边界图里没有比较完整的抛物线形状的三分线边界,而只是一段弧线,或者连弧线都没有),扫描法在性能上要优胜许多。因为扫描法可以根据弧线给出相应的判断结果,而改进型的抛物线 Hough 变换法得不到所期待的实验结果,并且篮球赛事视频大部分图像帧提取的 Canny 边界都没有完整的抛物线形状三分线边界,所以扫描法更适用。

1.5 根据坐标值判断边界曲线弯曲类型

从 A 中提取出来的边界曲线结果图已经准确提取到了所需要的边界曲线,接下来就是根据三维矩阵所存的边界曲线上点序列的坐标值来判断曲线的弯曲方向。这里给出基于向量夹角余弦值的判定方法:

从每幅图的第一条边界曲线开始遍历,直到最后一条边界曲线结束遍历。

Step1:为了排除边界曲线上前后点可能不满足要求的影响,在取首尾点的时候都向内缩进了 s 个点,取定首点为 (x_1, y_1) , 尾点为 (x_2, y_2) , 中点为 (x_3, y_3) , 并取到 x 轴的单位向量 \vec{x} , 再初始化一个一维零矩阵, 用来标记每张帧图像上提取到的多条边界曲线所对应的弯曲类型, 1 表示向右, -1 表示向左, 0 表示该帧图像上没有提取到对应的边界曲线;

Step2:过中点 (x_3, y_3) 向首尾点连线作垂线,并由计算可得垂点坐标分别为:

$$x_0 = \frac{(x_2 - x_1) \times x_3}{y_2 - y_1} + y_3 - y_1 + \frac{(y_2 - y_1) \times x_1}{x_2 - x_1}$$
$$y_0 = \frac{(y_2 - y_1) \times x_0}{x_2 - x_1} - \frac{(y_2 - y_1) \times x_1}{x_2 - x_1} + y_1$$

Step3:求得由中点 (x_3, y_3) 指向垂点 (x_0, y_0) 的向量,并将其单位化,用 $\vec{t}(t_x, t_y)$ 表示:

$$t_x = \frac{x_0 - x_3}{\sqrt{(x_0 - x_3)^2 + (y_0 - y_3)^2}}$$
$$t_y = \frac{y_0 - y_3}{\sqrt{(x_0 - x_3)^2 + (y_0 - y_3)^2}}$$

Step4:根据两点间距离公式求出中点和垂点的距离 $dis = \sqrt{(x_0 - x_3)^2 + (y_0 - y_3)^2}$, 如果大于 2 则进入第五步,否则遍历下一条边界曲线;

Step5:计算从中点指向垂点的单位向量 $\vec{t}(t_x, t_y)$ 与 x 轴的单位向量 \vec{x} 之间夹角的余弦值 $value$, $value = \frac{\vec{t} \cdot \vec{x}}{|\vec{t}| |\vec{x}|} = \frac{t_x \times 1 + t_y \times 0}{\sqrt{t_x^2 + t_y^2} \times 1} = t_x$; 若 $value > 0$, 则将对应

的标记矩阵分量值赋为 1; 若 $value < 0$, 则将对应的标记矩阵分量值赋为 -1;

Step6(综合多条边界曲线的开口方向来确定该帧图像的类型值):查看标记矩阵的值,如果所有分量除了 0 就是 1, 则将该帧边界曲线的开口方向类型值 $type$ 赋为 1, 表示开口方向向右; 如果所有分量除了 0 就是 -1, 则将该帧边界曲线的开口方向类型值 $type$ 赋为 -1, 表示开口方向向左; 否则将该帧边界曲线的开口方向类型值 $type$ 赋为 0, 表示没有提取到满足要求的边界曲线。

2 分割方法

对每帧图像进行上述六步处理,可以得到每帧图像的一个边界曲线的开口方向类型值 $type$ 。这里提出的分割方法就是根据这个 $type$ 值来对整个赛事视频进行镜头分割,具体方法表述如下:

如果是第一帧,则直接作为镜头边界起点,记录下帧号和 $type$ 值,若 $type \neq 0$, 将 $type$ 值赋给 pre_type , 否则继续查看下一帧,将第一个不为 0 的 $type$ 值赋给 pre_type ;

如果是最后一帧,也直接作为镜头边界终点,记录下帧号和 $type$ 值;

否则,仅当 $type \neq 0$ 和 $pre_type \neq 0$, 并且 $pre_type \neq type$ 时,才将对应帧作为镜头分割点,然后记录下帧号,并将 pre_type 值更新为 $type$ 值。

按上述分割算法得到的还只是初步分割结果,接着还要对该分割结果进行必要的优化:要将除首尾分割点之外的所有分割点的帧号统一向前挪动 n 个点的距离,这是因为能检测到三分线的帧要比实际发生场地切换的帧慢 n 帧,所以在调整分割点位置的时候需要考虑到;要保证两个分割点之间的距离大于帧数阈值 $FrmThresd$ 。最后才能得到完善的分割结果。

3 系统性能评价

基于以上算法,文中以 Microsoft Visual Studio 2010 与 OpenCV2.3 为开发平台,初步实现了一个篮球赛事视频镜头分割原型系统。为了验证提出的新分割算法的优越性,分别采用两种分割方法(分块颜色直方图方法和文中提出的新方法)对 2013 年 NBA 总决赛“热火 vs 马刺”的第七场第一节录像前 3 000 帧和 2013 年 CBA 总决赛“山东 vs 广东”的第四场第一节录像前 3 000 帧作对比分析,视频分辨率为 1 280×720,实验结果数据如表 1 所示。

分析表中数据可知,文中提出的新方法虽然在正确检出和漏检方面的优势不是很明显,但是在误检方面的优势很显然。分块颜色直方图方法受相机抖动和

场地光照条件的影响比较大,造成的误检很频繁,而文中提出的新方法可以有效地避免这一干扰,取得了一定的进步。

表 1 对比实验数据

	实际镜头个数	分块颜色直方图方法			文中的新方法		
		正确检出个数	误检个数	漏检个数	正确检出个数	误检个数	漏检个数
NBA	7	6	4	1	7	1	0
CBA	6	5	3	1	6	0	0

4 结束语

文中在研究前人有关视频镜头分割技术的基础上,结合篮球赛事视频本身的特点,提出了一种不同于以往视频分割技术的新方法,以期为后续研究工作做一些铺垫,从而让更多的篮球视频研究者从中受益。虽然这个方法取得了一定的成功,但同时也存在一些顾虑:

- (1)分割的依据是篮球场地的三分线开口方向,这个准则是否符合视频镜头分割技术的有效范畴有待商榷;
- (2)没有充分考虑到特殊的镜头切换效果给分割带来的影响,还需继续完善;
- (3)该方法工程应用性较强,在理论支撑和理论创新方面略显不足,后续研究工作还是要在这方面做进一步努力。

参考文献:

[1] 刘佳兵. 视频检索中的视频镜头分割技术[J]. 福建电脑, 2007(1):66-67.

[2] Ekin A, Tekalp A M. Shot type classification by dominant color for sports video segmentation and summarization[C]//Proc of IEEE international conference on acoustics, speech, and signal processing. [s. l.]:IEEE,2003:173-176.

(上接第 215 页)

[7] Latiff N M A, Tsimenidis C C, Sharif B S. Energy-aware clustering for wireless sensor networks using particle swarm optimization[C]//Proc of 18th international symposium on personal, indoor and mobile radio communications. Athens: IEEE,2007:1-5.

[8] Kennedy J, Eberhart R C. Particle swarm optimization[C]//Proc of IEEE international conference on neural networks. Piscataway: IEEE, 1995:1942-1948.

[9] 杨伟伟, 刘润杰, 申金媛. 一种基于 LEACH 的高效节能协议[J]. 传感技术学报, 2010, 23(8):1153-1157.

[10] 武春涛, 胡艳军. 无线传感器网络 LEACH 算法的改进[J]. 计算机技术与发展, 2009, 19(3):80-83.

[3] Kim K, Choi J, Kim N, et al. Extracting semantic information from basketball video based on audio-visual feature [C]//Proc of CIVR2002. London: Springer, 2002:278-288.

[4] 庄越挺, 潘云鹤, 吴 飞. 网上多媒体信息分析与检索[M]. 北京:清华大学出版社, 2002:78-79.

[5] 肖超云, 朱伟兴. 基于 Otsu 准则及图像熵的阈值分割算法[J]. 计算机工程, 2007, 33(14):188-189.

[6] 杨 辉, 刘 军, 阮 松. 基于 Mean Shift 算法视频跟踪研究[J]. 计算机工程与设计, 2013, 34(6):2062-2066.

[7] Rosenfeld A. The max Roberts operator is a Hueckel-type edge detector[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, PAMI-3(1):101-103.

[8] Deng Caixia, Ma Weifeng, Yin Yin. An edge detection approach of image fusion based on improved Sobel operator [C]//Proc of 2011 4th international congress on image and signal processing. Shanghai: IEEE, 2011:1189-1193.

[9] Li Yunfeng, Guo Yonghao, Cao Yukun. Apple image classification method based on the prewitt operator[C]//Proc of 2009 1st international conference on information science and engineering. [s. l.]:[s. n.], 2009:1161-1163.

[10] Wang Xin. Laplacian operator-based edge detectors[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(5):886-890.

[11] 刘瑞祯, 于仕琪. OpenCV 教程-基础篇[M]. 第 2 版. 北京:北京航空航天大学出版社, 2007:225-226.

[12] 李 钰, 孟祥萍. 自适应双阈值 Canny 算子的图像边缘检测[J]. 长春工程学院学报:自然科学版, 2007, 8(3):44-46.

[13] 李亚利, 王生进, 胡 斌, 等. 基于改进型抛物线 Hough 变换的眼睛特征提取[J]. 清华大学学报:自然科学版, 2010, 50(1):100-103.

[14] 李 静, 张绍武, 杨 峰, 等. 改进的随机 Hough 变换抛物线运动目标航迹起始算法[J]. 火力与指挥控制, 2011, 36(2):42-45.

[15] 王国宏. Hough 变换及其在信息处理中的应用[M]. 北京:兵器工业出版社, 2005:15-19.

[11] Li Chengfa, Ye Mao, Chen Guihai, et al. An energy-efficient unequal clustering mechanism for wireless sensor networks [C]//Pro of IEEE international conference on mobile Adhoc and sensor systems. Washington, DC: IEEE, 2005.

[12] Manzoor B, Javaid N, Rehman O, et al. Q-LEACH: a new routing protocol for WSNs[J]. Procedia Computer Science, 2013, 19:1-6.

[13] Latif K, Ahmad A, Javaid N, et al. Divide-and-rule scheme for energy efficient routing in wireless sensor networks [J]. Procedia Computer Science, 2013, 19:340-347.

[14] 刘 明, 曹建农, 陈贵海, 等. EADEEG: 能量感知的无线传感器网络数据收集协议[J]. 软件学报, 2007, 18(5):1092-1109.

篮球赛事视频镜头分割

作者: [夏磊](#), [张明](#), [XIA Lei](#), [ZHANG Ming](#)
作者单位: [上海海事大学 信息工程学院, 上海, 201306](#)
刊名: [计算机技术与发展](#) 
英文刊名: [Computer Technology and Development](#)
年, 卷(期): 2015 (2)

引用本文格式: [夏磊](#). [张明](#). [XIA Lei](#). [ZHANG Ming](#) [篮球赛事视频镜头分割](#)[期刊论文]-[计算机技术与发展](#) 2015 (2)