

体育视频中场地自动检测方法^①

殷伟良, 陈临强, 李 伟

(杭州电子科技大学 计算机学院, 杭州 310018)

摘 要: 主要针对虚拟广告系统设计了简单而有效的算法来实现体育视频中场地检测, 首先通过基于颜色空间直方图统计的方法实现了场地主区域的自动提取, 再通过 Top-Hat 变换和改进的最大类间方差法实现了场地边缘检测, 最后通过 Hough 直线检测和最小二乘拟合相结合的方法精确检测到所有的场地线。大量的实验表明, 本文提出的场地检测方法对于羽毛球场地, 网球场地, 乒乓球场地等都有很好的检测效果, 提取结果可用于摄像机定标和比赛场地的重建。

关键词: 场地检测; 颜色空间直方图; 顶帽变换; 大津法; 霍夫变换; 最小二乘法拟合

Automatic Detecting Field in Sport Video

YIN Wei-Liang, CHEN Lin-Qiang, LI Wei

(School of Computer, Hangzhou Dianzi University, Hangzhou 310018, China)

Abstract: In the virtual advertisement system, we use several simple and effective algorithm to solve the field detection problem in the sport video. Firstly, through a method based on color space histogram, we automatically detect the main area venues. Secondly, through Top-hat transform and the improved otsu algorithm to detect area edge. Finally, the Hough line detection and the least squares fitting method of combining accurate detection of all lines. Through A large number of experiments show that this method has a good effect on badminton field, tennis field and table tennis field. The results can be used for camera calibration and venues reconstruction.

Key words: field detection; color space histogram; Top-hat; otsu; Hough; least square method

体育比赛的组织者为了最大化挖掘广告的商业利润, 会布置尽可能多的广告牌在比赛场地周围。但是出于保护运动员等方面考虑, 实际上场地周围很多区域是不被允许摆放实体广告牌的, 如网球比赛的边线和底线附近。因此, 转播方可以利用视频融合技术, 分析实时转播的体育视频, 提取出视频中比赛场地框架, 然后把预先准备好的虚拟广告如广告图片或动画广告进行几何变换后融入到视频帧中指定的场地位置上, 这样得到了类似于实体广告效果的虚拟广告。如在羽毛球比赛、网球比赛或足球比赛中植入三维动画、虚拟广告牌、虚拟信息提示等。

虚拟广告是针对现实中的真实广告提出来概念, 虚拟广告是真实广告的三维模型, 可以是图片广

告、三维动画广告或视频广告等等。虚拟广告系统是通过视频采集卡实时采集到摄像机的视频源后, 分析采集到的视频, 然后通过本文提供的场地检测方法得到比赛场地, 进而通过场地线相交提取出场地线特征点, 可以通过建立特征点对待匹配, 计算出视频中场地与已知三维模型场地之间的透视变换关系, 通过 OSF 操作需要融合到视频中三维虚拟广告调整广告的位置, 通过计算得到透视变换把广告变换到二维, 最后把虚拟的广告融入视频帧进行实时播放。国外的公司已针对这一需求做出了相关的产品, 如: 以色列 Orad^[1]公司推出了 CyberSport 与 AD Vision 两个体育直播与视频融合解决方案。其中, 在足球比赛中使用的 TrackVision 系统, 已经是各个电视广播商在足球比赛

^① 基金项目:浙江省重大科技专项(2009C03015-4)

收稿时间:2011-08-21;收到修改稿时间:2011-10-03

直播中的首选产品；又如英国的 Hawk-Eye Innovations^[2]提供了视频分析系统，其鹰眼技术在网球、足球、斯诺克、板球比赛中已得到广泛应用。这些系统的实现过程中都会碰到体育视频中的场地检测问题，只有得到准确的场地信息才可以确定虚拟信息插入的位置，以及应用得到的场地信息去对一些特殊情况进行判断。本文主要针对虚拟广告系统中场地检测问题给出一种比较完善和自适应较强的方法。

本文提出的场地检测方法的主要流程为：场地主区域提取，边缘提取，场地线提取，线交点计算，计算出二维场地到三维场地的投影矩阵，场地检测流程如图 1 所示。本文中，对于主区域的提取，首先通过基于颜色空间直方图统计的方法提取出场地主颜色，然后找到面积最大的轮廓，最后通过腐蚀和膨胀操作得到一块相对比较完整的场地区域。接下来通过对主区域进行顶帽变换得到场地边缘，根据改进的最大类间方差法^[3]得到的自适应阈值将场地边缘图像进行二值化，再采用 Hough 直线检测和最小二乘拟合相结合的方法提取出直线^[4,5]，通过直线相交求得交点，进而与标准的场地建立对应关系。本文采用了几种实用型的改进技术，实现了视频主区域的自动提取，阈值自适应，可以精确的求得所有的场地线，取得较好的场地检测效果。

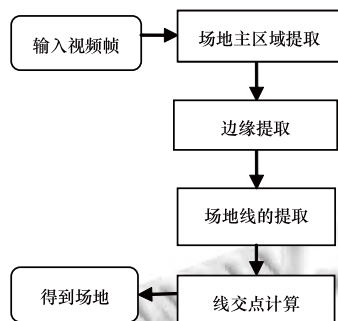


图 1 场地检测流程图

本文第 1 部分介绍研究背景,第 2,3,4 部分给出场地检测中所涉及的方法,第 5,6 部分分别为实验结果和结论。

1 场地主区域提取

在文献[6]中，刘国翌，杜威等采用了利用颜色分量差值的统计信息来分割球场主区域，可以得到比较精确的结果。但是其算法只能用在拥有先验知识的场

地中，无法实现场地的自适应提取。另一种方法是基于在场地类体育比赛节目中，一段视频图像的颜色主要分布在颜色空间中的几个比较密集区域，如图 2 所示。基于这一特点，研究者们提出了基于颜色空间直方图统计的提取方法。算法需要在颜色直方图中寻找一个数值普遍较大的密集区域，并将其作为球场颜色的大致分布区域。定义球场的主颜色为整个视频帧画面占 50% 以上的比例的颜色，如图 2 红土地网球场上的红色，实际提取主颜色是提取以红色为中心的领域范围，和红色比较接近的颜色也会被划分到主颜色范围中。比赛场地基本上位于关键帧的正中间，并且由于比赛场地本身大小的影响，它可能只占到整幅关键帧画面的约三分之一。那么，如果主场地以外的颜色也比较统一，那么比赛场地的主颜色就不一定能够成为颜色直方图中唯一的峰值，进一步说，颜色直方图中可能会出现两个以上的峰值，这样的情况很不利于提取比赛场地的主颜色。



图 2 原图

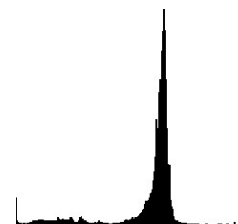


图 3 R 分量

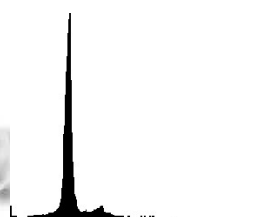


图 4 G 分量



图 5 B 分量

因此本文中提出了改进的场地类型的检测方法，方法如下：

首先，从体育视频中选出一个关键帧，关键帧代表存在比赛场地画面的视频帧，可以通过手动选取截图获得。

然后，在这视频帧的中心处选择一个宽度、高度均为关键帧的宽度、高度 0.5 倍的矩形区域。如图 6 所示黑色线包围区域：

$$\begin{aligned}
 H[i_{\min}] &\geq K * H[i_{\text{peak}}] \\
 H[i_{\min} - 1] &< K * H[i_{\text{peak}}] \\
 H[i_{\max}] &\geq K * H[i_{\text{peak}}] \\
 H[i_{\max} + 1] &< K * H[i_{\text{peak}}]
 \end{aligned} \tag{1}$$

其中, $H(i)$ 代表像素值为 i 时该像素值在视频帧中的数量, $H(i_{peak})$ 为颜色直方图的波峰值, K 为一常量, 取值为 0.2。



图 6 在矩形区域查找主颜色

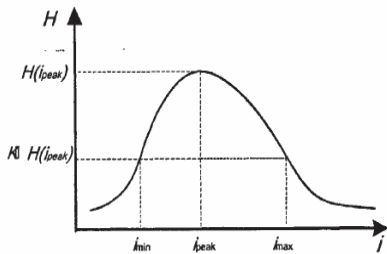


图 7 主颜色取值范围

最后, 利用通过下面介绍的步骤计算矩形区域内的主颜色范围:

第 1 步: 统计各个颜色分量 (RGB) 的直方图。

第 2 步: 在直方图上取某个区间 $[i_{min}, i_{max}]$, 如图 7 所示, 区间的左右端点在直方图上的取值要满足式(1), 使该区间包含了直方图的峰值 $H(i_{peak})$, 其中 K 为一常量, 在本文的实验中取值 0.2。

第 3 步: 扩展主颜色范围。修改颜色范围为 $[i_{min} - (i_{max} - i_{min})/2, i_{min} + (i_{max} - i_{min})/2]$, 这样就把整个波峰比较完整的包含在主颜色范围内, 纳入所有的相似颜色, 减少空洞。进而, 每个分量都得到一个颜色范围。

第 4 步: 取关键帧上各像素点的三个分量的灰度值逐一与上一步求得的相对应的颜色范围进行比较, 可以判断该像素点是否为主颜色像素, 根据主颜色范围得到二值化后的主颜色分布图, 即像素值落在主颜色范围内则把相应的灰度值置为 255, 其他置为 0。如图 8 所示。

第 5 步: 将若干较大的联通区域进行拼接, 以得到一块较完整的场地主区域; 然后对拼接结果进行膨

胀腐蚀等处理, 以去除噪声、填补空洞, 如图 9 所示; 最后与原图像进行掩模, 得到场地主区域, 如图 10 所示。

实验结果如下:



图 8 提取主颜色



图 9 去干扰和膨胀腐蚀

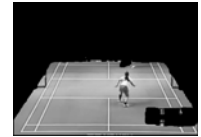


图 10 与原图掩膜

2 边缘提取和自适应阈值

通过上一步计算得到了场地的主区域, 然后需要进行场地的边缘提取。对主区域图像进行 Top-Hat 顶帽变换, 形态学中的 Top-Hat 算子是一种极好的高通滤波算子, 利用该算子通过选择合适的结构元, 就可以将需要的目标从复杂的背景中提取出来, 本文主要用来进行场地的边缘提取^[7]。消除了主区域图像的背景, 保留并突出了场地白线, 有利于提取场地边缘。根据比赛视频中场地白线的宽度, 本文选取结构元素 B 的尺寸为 3×3 , 如图 11 所示。

通过上一步计算得到场地边缘的灰度图, 然后采用改进的最大类间方差法对其进行二值化操作。步骤如下:

第 1 步: 灰度拉伸。首先找到最大和最小灰度 $MinGray, MaxGray$, 然后把 $(MinGray, MaxGray)$ 映射到 $(0, 255)$, 灰度拉伸的目的是让场地线和背景区分明显。因为经过对多种图像的具体实验发现, Otsu 算法致命的缺陷是当目标与背景灰度相差不明显时, 会出现无法忍受的大块黑色区域, 甚至会丢失整幅图像的信息, 因此通过灰度拉伸可以避免上述情况。映射过程如式(2)

$$f(i, j) = \frac{f(i, j) - MinGray + 1}{MaxGray - MinGray + 1} \quad (2)$$

其中 $f(i, j)$ 为图像上点 (i, j) 处的灰度值。

第 2 步: 通过高斯滤波去除噪声点, 以解决因为 Otsu 法对噪声敏感所带来的影响。如图 12 所示。

第 3 步: 更新滤波后产生的新的最大灰度值和最小灰度值 $(MinGray, MaxGray)$ 。

第四步: 在 $(MinGray, MaxGray)$ 范围内使用最大类间方差法得到的自适应阈值, 进而将场地边缘图像

进行二值化。具体方法如下:

对于图像, t 记为目标与背景的分割阈值, 目标像素数占图像比例为 ω_0 , 平均灰度 μ_0 ; 背景像素数占图像比例为 ω_1 , 平均灰度 μ_1 。图像的总平均灰度为:

$$\mu = \omega_0(t)\mu_0(t) + \omega_1(t)\mu_1(t) \quad (3)$$

在公式(3)中从最小灰度值到最大灰度值遍历 t , 当 t 使得方差公式(4)

$$g = \omega_0(\mu_0 - \mu)^2 + \omega_1(\mu_1 - \mu)^2 \quad (4)$$

最大时 t 即为分割的最佳阈值。其中 $\mu_0(t) \leq \mu \leq \mu_1(t)$ 。

此时 $\omega_0(t)(\mu_0(t) - \mu)^2 + \omega_1(t)(\mu_1(t) - \mu)^2$ 和 $(\mu_0(t) - \mu)^2(\mu_1(t) - \mu)^2$ 同时在处取得最大值, 因此两者是等价的。可以求 $g = (\mu_0(t) - \mu)^2(\mu_1(t) - \mu)^2$ 最大时的 t 为二值化的阈值。在这里方差公式(4)是灰度分布均匀性的一种度量, 方差值越大说明构成图像的两部分差别越大, 当部分目标点错分为背景或部分背景点错分为目标点都会导致两部分差别变小, 因此使类间方差最大的分割意味着错分概率最小。结果如图 13 所示:



图 11 Top-Hat 变换图 图 12 灰度拉伸和去噪 图 13 球场线图

3 场地线提取和线交点计算

通过上一步的计算得到场地边缘图, 然后采用 Hough 变换检测直线^[8,9], 但是直接使用 Hough 变换存在两个问题: 首先得到直线参数为离散值, 造成精度较差。其次较长的线附近的参数可能比短线段的计数多, 从而掩盖了较短线段; 针对以上存在的问题, 本文采用了一种改进的 Hough 变换, 结合最小二乘法提取图像中多条线段的方法^[10]。

整个直线检测算法步骤如下:

初始化 $\sin[v]$ 和 $\cos[v]$, 数组 ImageWhite, Param $[\rho][\theta]$ 。

对场地边缘图进行 hough 直线检测, 把满足条件的检测结果存入数组 Param。

把数组 Param 中最大值取出, 并找出所有距离这条直线 3 个像素的所有点, 利用上述得到的点通过最小二乘法拟合出新的直线。

把上面的直线邻近的点剔除, 同时把消除在数组 Param 中的计数。

计算提取出来直线的数量, 如果满足所需要数量则结束, 否则转到 3)。

其中参数空间为 ρ 和 θ , θ 为角度, 取值范围 $0 < \theta < 180$, ρ 取值范围从 $-\text{dist} \sim \text{dist}$, dist 是图像中对角线长度。 $\sin[v]$ 和 $\cos[v]$ 代表 $0 < \theta < 180$ 的正弦和余弦值, 数组 ImageWhite 中存储图像所有白色点的坐标。

在以上步骤中采用最小二乘拟合直线的方法来修正用 Hough 变换提取出的直线参数。用该方法求得的拟合直线满足所有输入点到直线的竖直距离的平方和最小, 从而提高了提取直线的精度。结果如图 14 所示:

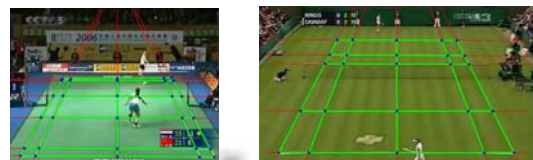


图 14 直线检测结果图

4 实验结果

本文采用多种类型比赛场地进行实验, 在羽毛球场地, 网球场地, 乒乓球场地上都可以取得比较好的实验结果。多种类型场地检测结果如表 1 所示。实验中主区域提取和阈值自适应取得很好的效果, 误差主要出现在直线提取部分, 容易发生漏检或误检问题。实验结果表明本文比赛场地检测效果成功率较高。

表 1 多种类型场地检测结果分析

| | 成功次数/主 区域提取次数 | 成功次数/阈值 自适应次数 | 成功次数/直 线提取次数 |
|-------|------------------|------------------|-----------------|
| 羽毛球场地 | 42/42 | 42/42 | 40/42 |
| 网球场地 | 37/37 | 37/37 | 36/37 |
| 乒乓球场地 | 32/34 | 32/34 | 31/34 |



图 15 网球场原图

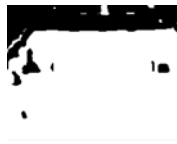


图 16 球场主区域图



图 17 去除空洞图



图 18 顶帽变换图



图 19 stu 二值化图



图 20 结果



图 21 羽毛球场原图



图 22 球场主区域图



图 23 去除空洞图

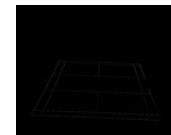


图 24 顶帽变换图

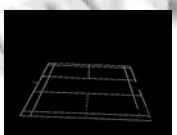


图 25 1ostu 二值化图



图 26 结果图

5 结语

本文主要解决了不同类型的体育视频中比赛场地检测问题,采用了几种简单有效的方法,就可以达到自动提取出羽毛球场地,网球场地和乒乓球场地等,减少人为的设定参数带来的不便。本文提出的场地检测方法实现了视频主区域的自动提取,自适应阈值,精确的求得所有的场地线,得到了比较好的效果。检测结果可以用于摄像机的定标和建立二维场地和三维

场地的关系。

参考文献

- 1 Orad homepage. <http://www.orad.co.il/en/index.asp>,2010.
- 2 Hawk-Eye Innovations homepage. <http://www.hawkeyeinnovations.co.uk>, 2010.
- 3 Ostu N. A threshold selection method from gray-level histogram. IEEE Trans, 1979,SMC29:62-66.
- 4 Li YQ, Wan KW, Yan X, Xu CS. Real Time Advertisement Insertion in Baseball Video Based on Advertisement Effect. MM'05, November 6-11, 2005, Singapore. 2005: 343-346.
- 5 Yan X, Yu XG, Chi TTP. A System for 3D Projected Virtual Content Insertion into Broadcast Tennis Video. MM'06. October 23-27, 2006, Santa Barbara, California, USA, 2006: 809-810.
- 6 刘国翌,杜威,李华.足球场地标志线的自动提取.计算机辅助设计与图形学学报,2003,15(7):870-874.
- 7 崔屹.图象处理与分析—数学形态学方法及应用.北京:科学出版社,2000.127-131.
- 8 Wan K, Yan X, Yu X, Xu CS. Real-time goal-mouth detection in MPEG soccer video. Proc. ACM Multimedia. 2003: 311-314.
- 9 Wan K, Yan X, Yu XG, Xu CS. Robust Goal-Mouth Detection for Virtual Content Insertion. MM'03. November 2-8, 2003, Berkeley, California, USA. 2003:468-469.
- 10 Walsh D, Raftery AE. Accurate and efficient curve detection in images: The importance sampling Hough transform. Pattern Recognition, 2002,35(7):1421-1431.