

# 基于足球广播视讯的自动场景分类与球员侦测及追踪技术研究

张晓健

苏州工艺美术职业技术学院 江苏苏州 215104

**摘要:** 在本篇论文中我们提出一套足球广播录影视讯的自动场分析判断和球员侦测与追踪的法则。我们的方法包括二个阶段, 第一个是自动划分录影的场景分析工作。我们提出一套基于HSV颜色的高斯分类模型并设计了一个随视讯内容自动更新模型的方法。第二个是球员分类阶段。我们提出一个基于HSV直方图判断树的无监督分类方法进行球员的分类工作, 并且利用线性预估模型进行球员的追踪定位工作。实验结果证明我们提出的方法是简单并且有效的。

**关键词:** 足球视讯; 场地切割; 球员侦测与追踪; 渐进式高斯模型更新

## Research on automatic scene classification and player detection and tracking technology based on football broadcast video

Xiaojian Zhang

Suzhou Arts and Design technology Institute Suzhou Jiangsu 215104

**Abstract:** In this paper, we present a set of rules for automated scene analysis, player detection, and tracking in football broadcasting videos. Our approach consists of two stages. The first stage involves automatic scene segmentation in the recorded footage. We propose a Gaussian classification model based on the HSV color space and introduce a method for automatically updating the model according to video content changes. The second stage focuses on player classification. We propose an unsupervised classification method based on HSV histogram decision trees for player categorization. Additionally, we employ a linear prediction model for player tracking and positioning. Experimental results validate the simplicity and effectiveness of the proposed approach.

**Keywords:** Football video; field cutting; player detection and tracking; progressive Gaussian model updates

近年来, 体育视频的事件挖掘在理解、事件检索和内容搜索方面变得越来越重要。体育项目, 如足球、篮球、足球和网球, 都是最有趣的部分。其中一个有趣的应用是自动体育视频摘要和浏览。基于语义的分析, 如亮点检测、战术分析和玩家活动分析是应用程序的目标。由于视频媒体的爆炸式增长和庞大的受众数量, 自动视频内容分析已经成为一个非常有趣的研究课题。在这些视频中, 广播体育视频因其受众众多而备受研究者的关

注。体育视频数据的分析, 如亮点提取、视频摘要和战术分析, 可以为视频内容服务提供商和体育比赛观看者提供极大的便利。本文提出了一种足球视频广播的自动分析系统。尽管已经对足球视频进行了许多研究, 但由于视频中的内容复杂, 特征难以提取, 自动分析仍然是一个悬而未决的问题。由于许多困难, 这项任务相当具有挑战性, 例如不同的日光、阴影、遮挡、相似的玩家外观和低辨别力、不同的玩家数量、突然的相机运动、游戏。

视频中的镜头包含许多不同的条件, 摄像机总是跟随球员移动, 摄像机会根据场地和球员的比赛情况频繁放大或缩小。由于广播视频中的镜头很复杂, 因此广播视频的处理比个人制作的视频要困难得多。因此, 大多

---

**作者简介:** 张晓健, 苏州工艺美术职业技术学院, 江苏苏州, 本科, 苏州大学, 研究生, 苏州大学, 学历: 硕士, 职称: 副教授, 研究方向: 足球教学与训练。

数研究都集中在广播视频上，其主要目标是对视频进行自动分析。根据足球视频自动分析所需的流程，我们可以简单地将其分为三个主题，讨论如下：

1. 足球运动员场地的分割和场景的分类：视频序列由不同的镜头组成。每个镜头都包含一系列视频场景或帧。为了检测游戏中的玩家，需要首先检测每帧中游戏的游戏场地。通常，根据观看条件和镜头中每一帧的场景内容，足球视频可以分为四种镜头。它们有远射、中射、近射和远射。长镜头的场景包含游戏场地的全局视图。它是由摄像机在小视场下远距离拍摄的。它显示了游戏的全局游戏状态，如团队游戏的位置、团队组建和长传。从长远来看，游戏区域对整个图像的比例很高。场景通常包含许多玩家，而玩家的体型通常较小。

2. 球员的检测和跟踪：在足球比赛中，除了进球场景外，球员的状态或动作也是观看或分析战术分析、比赛抽象和总结的有趣部分。因此，正确的球员分割步骤对于在足球比赛中跟踪球员至关重要。大多数用于运动物体检测的算法都是基于背景减法的：通过从参考背景模型中减去当前图像来提取前景物体。关于这些方法的一些综述可以在中找到。运动应用的主要特征使研究人员使用了基于运动场与不同颜色（球员和球）的分割的运动检测方法。其他有趣的方法是基于数学形态学的：特别，提出了一种基于分水岭变换的有趣分割算法。

### 一、处理方法

本文提出了一种在足球直播视频中实现球场自动分割、射门分类、球队聚类和球员跟踪的方案。系统框架如图7所示。整个程序由五个块组成。第一个块是初始处理块，包括图像帧捕获、游戏场主色估计和每个帧的色调饱和度（HS）颜色直方图分析。下一个块是镜头分类块。在这个块中，我们首先基于颜色直方图进行射门边界检测，并通过使用决策树将每个射门分为长视图、近视图或外场视图中的一个。然后，估计图像序列中运动场颜色的HS多元高斯分布，并对每个长视角镜头序列进行增量更新，以适应在足球比赛期间改变运动场的颜色的光照条件，下一个区块是玩家分割区块。采用两步形态学算子来消除分割后的噪声、阴影和线迹的影响。此外，为了在礼堂边缘获得更好的玩家检测结果，应用了顶部边界检测算子。由于游戏场上有两个游戏团队，玩家的聚类方法是通过使用基于HS颜色直方图的简单2-均值聚类算法来实现的。使用无监督的训练过程，从而自动执行分类。为了解决球员制服为黑色或白色时的H-S直方图问题，在H-V直方图的基础上增加了第二决策树。最后，在玩家检测之后，采用了基于跟踪的玩家聚类方法。使用线性运动矢量和预测模型的玩家跟踪模型被应用于每个分割的玩家，以便处理多个被遮挡的玩

家跟踪。下面将详细说明整个过程。

我们描述了镜头分类的过程。该过程可分为四个步骤，初始处理、镜头边界检测、运动场分割和镜头分类。在初始处理期间，将捕获的图像从RGB转换为HSV颜色空间，以便计算用于镜头边界检测的H-S 2-D直方图，并用于生成用于运动场分割的初始H-S高斯模型。然后，使用卡方距离来定位镜头边界，并将增量高斯模型更新过程应用于草地运动场分割，以处理不同日光环境下的阴影效果。基于分割后使用决策树的类草比率结果，将镜头分为三类。

### 二、镜头边界检测与分类

在足球视频分析中，球场颜色起着非常重要的作用，尤其是在区分射门类别方面。由于不同的足球场、摄像机拍摄角度和阴影的存在，球场颜色可能会有所不同，因此必须对每一帧进行球场检测，并且草地颜色模型应适应捕捉到的帧。因此，选择合适的颜色域是获得良好分割结果的一个非常重要的关键点。使用RGB进行地面检测，分割结果当场景中包含阴影时，运动场的分割结果并不那么精确。与RGB、CMY或YIQ等面向硬件的颜色模型不同，HSV颜色模型涉及色调、饱和度和亮度的人类感知概念。此外，已经观察到HSV空间更适合计算颜色变化，因为将色调与饱和度分离，并且亮度在mos下增加了鲁棒性。本文提出了一种在线自动运动场颜色模型训练方案。给定一个视频序列，该方案自适应地选择训练场景进行初始训练。由于视频序列可以从上面定义的两个镜头中的任何一个开始，我们不能直接使用视频序列开始部分的场景的下半部分进行训练。提出了一种从测试视频序列开始检测第一个长视角的草地运动场检测方案。在初始步骤中，如果输入帧的均值落入这两个区间，则该帧将被接受并定义为初始训练帧。否则，检测器检查下一帧，直到检测到草模型为止。在找到满足的5个帧作为初始训练帧之后，可以估计高斯模型。该模型将在初始分割过程中使用。由于足球视频是由不同的视图或不同的剪辑组成的。为了建立一个更好、精确的运动场高斯分布模型，我们需要定位镜头边界并确定每个片段的观看条件。在镜头转换过程中，捕捉到的场景中的内容并不稳定，因此这些图像不能作为训练图像。对于不属于远景的镜头，镜头的图像也不能用作训练图像。因此，镜头边界检测在下面被使用和废弃。在确定镜头边界并对帧进行分割后，我们可以通过镜头分类来确定镜头边界后的剪辑类型。投篮类型一般可分为四种，即远视、中视、近景和外场。由于中景和近景的场景内容非常相似，而我们在论文中的目的是寻找远景场景，因此我们只考虑将视频分为远景、近景和外场三种镜头。

在决定帧的拍摄类型之后，为了决定两个拍摄边界

之间的剪辑的类型, 如果拍摄边界之后的五个以上帧属于相同类型的视图, 则将拍摄的剪辑分类为五个帧中的相同类型。如果检测到新的镜头边界, 则再次应用新的类型检测。

运动场高斯模型的精度是影响运动场分割精度的一个重要因素。由于足球比赛至少可以持续90分钟, 因此阳光条件、体育场等可能会对拍摄的画面产生不同的影响。由不同日光引起的操场或玩家阴影的示例。阴影可能出现在游戏过程中的任何地方, 因此我们提出了一种增量训练解决方案, 根据拍摄场景的情况生成自适应高斯游戏场模型。如上所述, 在初始阶段期间, 使用满足上面列出的平均值和色调范围约束的第一帧来估计第一高斯分布。然后, 对于每个帧, 使用上述拍摄类型检测方案来检查输入帧的类型。如果超过五个帧属于长视图类型, 则将这五个帧的所有运动场候选像素添加到用于upd的训练集中。

### 三、检测和跟踪

讨论了玩家检测和跟踪的过程。在球员检测阶段, 获得分割的足球场区域的顶部边界, 并去除外场区域。然后检测分割后与外场区域连接的球员斑点, 并重新检测球员斑点。估计用于识别玩家的特征, 并且应用玩家检测滤波器来检测所有真实的可能的玩家候选斑点。在球员训练阶段, 使用训练阶段每个球员球衣的颜色进行2均值聚类算法, 将所有可能的训练球员斑点自动分类为两个团队。最后, 构造了一个空间位置相似矩阵来确定连续帧之间玩家斑点的对应关系。然后, 线性运动和预测阶段被用于球员跟踪, 并且相应地执行每个球员的团队分类。

在运动场分割过程之后, 每个图像像素可能属于两类中的一类, 运动场像素或非运动场像素, 结果示例的右上角所示。黑色像素是可能的运动场像素, 而白色像素是非运动场像素。然后可以通过使用区域生长来获得运动场和非运动场区域。可以看出, 对于长视角下的足球比赛图像, 分割后的图像下半部分存在一个主要的运动场区域边界。该区域边界以上的那些区域或斑点属于外场区域, 而非赛场斑点的其余斑点是可能的球员斑点。为了检测所有可能的球员, 需要首先检测球场的顶部边界, 以便能够很好地隔离连接到外场的球员斑点。对于远景场景, 当摄像机视图聚焦在中场时, 运动场的顶部边界由一条直线组成, 当摄像机图像聚焦在角落时, 由两条直线组成。通常, 霍夫变换用于检测边界线。但是, 首先, 需要计算通过外场区域边界的所有可能的线, 并且需要投票过程来估计线方程。其次, 由于顶部边界可能由一个或两个线段组成, 因此需要后处理过程来决定霍夫变换后提取的线段数量。为了避免这种复杂性, 我们提出了一种简单快速的方法来检测分割后的二值图像的顶边界, 而不是使用霍夫变换。第三个序列包含彩色和

白色的球员球衣, 因此应用S-V直方图。在这种情况下, 尽管阴影与播放器结合在一起, 但分割结果是精确的, 因此阴影不会影响最终的跟踪性能, 并且它们也都被正确跟踪。从第9299帧到第9319帧, 播放器2和4变为合并状态, 然后在第9324帧中分离。从第9324帧到第9344帧, 尽管球员2和4与中场线合并, 并且在第9334帧的标记中丢失, 但是在跟踪过程的预测阶段仍然可以检测到它们。

第一个序列包含黑色或白色的球员球衣, 因此应用S-V直方图。我们可以从第4624帧中的播放器4、5和9、19中看到我们的跟踪方案的效果。播放器4和5处于合并状态, 播放器9和19从帧4599到4604变为合并状态, 然后从帧4619到4624变为分离状态。这两个斑点的初始标记是6, 并且在处理之后, 它们被正确地跟踪。此外, 通过这些帧可以容易地观察跟踪后的一致性标记的结果。使用H-S颜色模型的帧2262中每个斑点的概率结果如图46所示。总共有6名玩家被正确检测和追踪。尽管斑点1、2和5被合并在一起, 但每个斑点对团队的概率很容易区分。球员1和球员5分别有55%和62%的信心属于蓝白队, 球员2有96%的信心属于红队。球员正确归属蓝白队的平均概率为82.09%, 而正确归属红队的平均几率为98.05%。结果非常令人满意。

### 四、结论

本文提出了一种简单而有效的足球直播视频场景自动分析和球员检测与跟踪方案。我们的方法包括两个阶段, 第一阶段是基于色调和饱和度的二维高斯颜色模型的视频场景自动分析阶段。提出了一种增量颜色模型更新方法, 可以处理运动场中有阴影的场景。第二个阶段是玩家分析和跟踪阶段。提出了一种基于颜色直方图的方法来使用决策树对玩家进行分类, 并使用具有线性预测模型的空间相似性矩阵来跟踪玩家。实验结果表明, 该方法简单有效。

然而, 在本研究中, 重点是玩家对远景场景的分析, 其他类型的场景也应进行分析。建议对未来工作进行以下研究:

1. 在更多摄像机运动和更复杂场景下的玩家跟踪方案。
2. 镜头检测方案研究不仅针对切割型边界, 也针对其他类型的镜头边界。
3. 针对所有不同类型的场景的玩家跟踪。
4. 足球视频的亮点提取和视频自动摘要。

### 参考文献:

- [1] 大卫·罗. 体育、文化与媒介: 不羁的三位一体[M]. 吕鹏, 译. 北京: 清华大学出版社, 2013
- [2] 牛思远. 最后的狂欢: 一个足球撬动整个经济世界[N]. 南方日报, 2014-07-(11).
- [3] 梁维卿. 试论足球运动的全球化与个性化[J]. 体育文化导刊, 2003(6): 25-27.