摘　要

多媒体信息的爆炸式增长要求人们研究开发视频检索技术。近年来，以体育视频内容分析为主要目标的研究已经成为热门领域。这项研究主要为观众提供精彩摘要并以多种形式呈现给用户。虽然经过十余年的努力，这一领域已经取得了长足进步，但当前取得的成绩与人们的期望还有很大距离。

本文针对足球视频场景分析与三维重建中的若干关键技术进行了比较深入的研究，旨在提高它们的鲁棒性和可应用性。文中工作集中在以下三个方面：1）自动的球场检测算法，这项技术是场景分析与三维重建工作中共用的技术；2）运动对象的检测和跟踪，在体育视频中运动对象主要是指球员和球；3）从单视点广播视频中重建三维场景技术。本文的主要创新工作有：

1. 本文提出了一种基于本征图像的自动球场颜色检测算法。作为体育视频分析研究中一项关键技术，球场检测严重地受到阴影区域的影响，此算法可以有效地克服这一困难。算法首先自动地在色彩分布直方图中检测主颜色。因具有相似表面反射函数的像素值在正确的本征图像投影方向投影后，它们具有关于某一点对称分布。依据此特性，算法用已检测到的主颜色粗标定本征图像的投影方向。在此方向的指导下，算法根据本征图像的分布来判断第二主颜色是否是球场颜色的。此外，当确定了球场中阴影区域和非阴影区域后，本文提出通过最大化二者本征图像的分布相似度精细标定相机的本征图像的投影方向。实验表明此算法可以比较准确地检测球场颜色，同时精度的本征图像投影方向为在含有阴影区域的目标跟踪提供了一种光照不变的表观特征。
2. 本文提出了一种自适应的球场检测算法。针对广播视频流中的数据具有数量大、噪声多、同时来自多部相机的特点，在检测到的球场颜色的基础上，算法采用高斯混合模型对其建模，并且利用增量期望最大算法对模型参数进行在线估计。算法一方面提高了球场检测的效果，另一方面减少训练过程对数据存储的开销。实验表明，由于增量学习方法不断地强化模型的泛化能力，此算法取得了比单纯的直方图方法和一般的高斯混合模型方法更好的结果。
3. 本文提出了一种顺序数据关联策略应用于多目标的跟踪。多运动目标跟踪受到对象之间相互遮挡的困扰。因为比赛场地是平面，根据成像规律同一平面上物距大的目标的像位于物距小的目标的像的上部，此策略按照目标

的像的底边位置关系，从下到上依次关联各个目标的观测。与传统的基于联合概率数据关联方法相比，此策略大大减少了物距排列的搜索空间，提高了算法的效率，同时基本保证了基于此策略的数据关联是贝叶斯意义下的最优解。此外，本文提出一种最大化前景模板相似度同时最小化背景模板相似度的表观建模方法。这种方法在匹配目标时，同时考虑了目标背景的影响，从而有效地提高了跟踪器的跟踪性能。对于球的检测和跟踪，本文从实验角度上回答了只有跟踪和检测有机地结合才能取得更好的性能。

4. 本文提出了一种从单视广播视频中重建三维场景的方法。方法根据图像平面和场地平面的对应关系，并利用相机的自标定方法联合确定相机在场地中的位置；然后利用相机的位置、球的飞行平面、球的虚影之间的几何关系，本文推导出球的三维位置估计公式。另外，对于无法确定球的飞行平面的情况，本文提出一种基于物理限制的飞行平面预测算法。这种算法通过优化一个新设计的耗费函数，使其在最小值时对应于球的最合理的三维轨迹。此方法的优点是可以估计飞行中的球的三维位置而无需参考已知高度的物体；应用此法的限制条件少，其只要求相机位置是固定的。实验表明，此三维重建方法取得了令人满意的结果。

上述技术被应用到足球视频精彩检测和精彩片断的三维动画生成原型系统中去，取得了令人满意的效果。虽然以足球视频为背景，但球场检测以及目标检测与跟踪可以经过简单改进应用到其它体育视频分析中。

关键词　体育视频分析；球场检测；多目标跟踪；三维重建

Abstract

The explosive growth in the amount of multimedia information necessitates the development of content-based video indexing and retrieval techniques. Recently, sports video content analysis has become a hot research area. This research work's main objective is to automatically extract high-lights pieces from a video sequence for users quickly browsing and enriching them with different represent forms. Although great progress has been achieved in the past ten years, the successes of some systems can not satisfy users' requirement and is far from the research target.

This paper focuses on some key technologies in soccer video analysis and 3D scene reconstruction, in order to make them more robust and applicable. The main work includes the following three points: 1) automatically detecting playfield. This technology is commonly used in soccer video analysis and 3D scene reconstruction; 2) moving objects detection and tracking. In sports videos these objects are ball and players; 3) 3D scene reconstruction based on monocular broadcast video. The main novelties consist of the following points:

1. We propose a new algorithm for playfield color detection based on intrinsic image. As a key technology in sports video analysis, it is affected by shadow heavily. The proposed algorithm can overcome the problem. The algorithm first detects dominant color in the color histogram for a sequence. Since similar surface's intrinsic image has symmetrical distribution along the right projection direction, the camera sensor can be calibrated. As the projection direction directs, the algorithm justifies whether the second dominant is the playfield color by their intrinsic image distribution. Moreover, when the shadow regions and non-shadow regions are determined in playfield, the camera's sensor can be fined calibrated by maximize the similarity of the two regions' intrinsic image distribution. Experiment show that the algorithm can detect playfield colors even including shadow.
2. We propose an adaptive playfield detection algorithm. Based on the detected playfield color, the algorithm exploits GMM to model these colors distribution, whose parameters are updated by incremental expectation maximization algorithm. This method not only improves the detection results, but also reduces the memory requirement in training processing. Experiments show the method acquires better results than the methods using histogram or GMM.
3. We propose the named sequential data association strategy for multiple objects tracking. Multiple objects tracking is puzzled by occlusion. As the players are on a plane, a player's image is on lower part in the image if he has small scene depth. As the strategy suggests, the data association is performed in the sort ascending of the player images' vertical position. Compared with traditional joint probability data association, this method reduces the searching space. The method improves the tracking efficiency and almost acquires the optimal solution in the sense of Bayesian. More, we propose a new appearance model, which reinforces the foreground and strains the background. For ball detection and tracking, we

analyze the relationship between detection and tracking in the view of experiments.

4. We propose a new method for estimating the players' and ball's 3D position information from monocular broadcast video. For players, the homography between image and playfield is used to estimate their positions. For ball, by analyzing the geometry relation between ball, its "virtual" shadow and camera position, we derive the formulas for estimating flying ball's 3D position. Moreover, we propose a method to predict the flying plane if it can not be determined from images. This method designs a new cost function, which arrives at the minimum when the predicted flying plane is reasonable. This method has at least two merits. One is that it can estimate the flying ball's position without referring to other object with known height; the other is that only one assumption is made, the camera is fixed a position.Experimental results are satisfying.

The mentioned technologies have been used for exciting event detection and 3D cartoon generation for high-lights piece. The methods for playfield detection and objects detection and tracking can be applied to other sports video analysis.

Keyword sports video analysis, playfield detection, multiple objects tracking, 3D reconstruction