Video2Cartoon: Generating 3D Cartoon from Broadcast Soccer Video

Dawei Liang[1]; Yang Liu[1]; Qingming Huang[3]; Guangyu Zhu[1];
Shuqiang Jiang[2]; Zhebin Zhang[2]; Wen Gao[1, 2, 3]

[1](School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China)

[2](Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100080, China)

[3](Graduate School of Chinese Academy of Sciences, Beijing 100039, China)

{dwliang, yliu, qmhuang, gyzhu, sqjiang, zbzhang, wgao}@jdl.ac.cn

ABSTRACT

In this demonstration, a prototype system for generating 3D cartoon from broadcast soccer video is proposed. This system takes advantage of computer vision (CV) and computer graphics (CG) techniques to provide users new experience that can not be obtained from original video. Firstly, it uses CV techniques to obtain 3D positions of the players and ball. Then, CG techniques are applied to model the playfield, players, and ball. Finally, 3D cartoon is generated. Our system allows users to watch the game at any point of view using a 3D viewer based on OpenGL.

Categories and Subject Descriptors

I.4.5 [**Image Processing and Computer Vision**]: Reconstruction – Transform methods; I.3.7 [**Computer Graphics**]: Three-Dimensional Graphics and Realism – Animation

General Terms

Algorithms, Experimentation

Keywords

Soccer Video, Camera Calibration, 3D Cartoon Generation

1. INTRODUCTION

Soccer has been one of the most popular sports games all over the world. It has driven howling success of TV broadcasting service. At each game many cameras are deployed around the playfield to capture every movement of the players, referees and ball. However, on television, only a single viewpoint is available to the viewers, at any time. Some viewers may wish to see how the ball streaks towards the goalmouth from the goalie's point of view. Others may want to observe the progress of a goal from a bird's eye point of view. By using CV and CG techniques, such desire can be satisfied.

There is some related work in the literature. Bebie and Bieri [1] presented the SoccerMan system, which can generate an animated 3D scene from a soccer video sequence. In their system, player is modeled as the so-called animated texture object (quadrilateral holding the player's texture). Hence, the locations of point of view are limited. Recently, Yu et al. [5] presented a 3D

reconstruction and enrichment system, which can reconstruct not only the goalmouth but also the midfield scene. However, it only provides the main camera's point of view.

In this demonstration, we present a system – Video2Cartoon, which can generate 3D cartoon from broadcast soccer video, and allows users to watch the game from arbitrary point of view. The most related work to us is that of Matsui et al. [4]. They also generate animation from broadcast soccer video. However, our work differs from theirs in the following aspects: first, we adopt a new camera calibration method [3], which can obtain camera calibration parameters even when the corresponding points are insufficient; second, a sophisticated method [6] is employed to track soccer players and ball, which is robust to background clutter and occlusion; last, our system can obtain the ball's 3D positions under certain assumptions.

2. SYSTEM OVERVIEW

Figure 1 provides an overview of our proposed system. It consists of three main modules. The first one is 3D information extraction, which consists of camera calibration, players and ball detection and tracking, and 3D position estimation of players and ball. The second one is playfield and player modeling. The last one is 3D cartoon generation. In what follows, we will introduce these modules in detail.
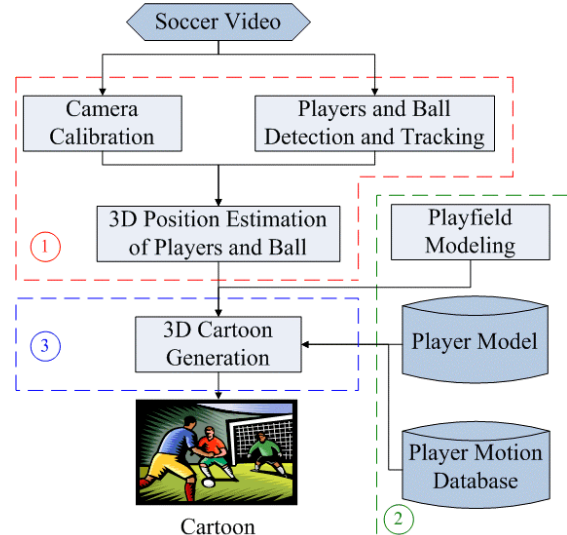

Figure 1. An overview of the proposed system.

3. 3D INFORMATION EXTRACTION

3.1 Camera Calibration

In order to obtain the 3D positions of players and ball, camera must be calibrated for each frame. Since playfield is a plane, calibration is reduced to estimate the transformation between the playfield and its image, which is a 3 by 3 matrix called homography. Here we adopt our newly developed method [3]. If there are more than four corresponding points, with no three of them collinear, the homography can be estimated directly. For the frames with insufficient corresponding points, we use global motion estimation (GME) and the calibrated frames to compute their homographies. To robustly estimate the GME, two strategies are exploited. One is removing the moving objects (i.e. players) based on adaptive Gaussian Mixture Model playfield detection, and the other is using Kanade-Lucas-Tomasi tracker to determine horizontal and vertical translation.

3.2 Objects Detection and Tracking

Multiple objects detection and tracking approach based on support vector regression (SVR) particle filter [6] is adopted. SVR particle filter is an improved particle filter that can achieve better performance with small particle set and enhance the efficiency of tracking system. Object detection based on playfield detection is also combined into the tracking framework, which makes the tracking algorithm fully automatic.

3.3 3D Position Estimation

It is reasonable to assume that the players move on the playfield in most cases, so the 3D position can be obtained by homography. The bottom line midpoint of the minimal bounding box of the player is regarded as his image position. Motivated by [2], we propose an improved version of their algorithm. Under the assumption that the ball follows a parabolic trajectory in the air, the 3D position can be obtained by seeking the intersection point of a line and a plane. The line is determined by the camera position (obtained by self-calibration) and the shadow position (the corresponding world position of the ball's image) of the ball on the playfield. The plane is determined by the starting and ending points of the ball on the playfield. Even if the ending point is not obtained in some conditions, e.g. heading the ball, the plane can be identified by searching a plane in which the ball's trajectory is most close to a parabola.

4. PLAYFIELD AND PLAYER MODELING

Playfield modeling is done according to the Laws of the Game of FIFA. To enhance the viewing experience, playfield is not only texture-mapped using the playfield detection result, but also decorated with some auxiliary materials, such as billboards, racetrack, auditoria and so on. Player model is built according to H-anim1.1, which is the specification for a standard humanoid, and hence it can do lots of complex actions.

5. 3D CARTOON GENERATION

Our system is implemented based on OpenGL in Visual C++ 6.0. Figure 2 shows a snapshot of the prototype. The key issue of 3D cartoon generation is the players' motion. We build player motion database in advance, including some simple actions such as

running, walking, etc. More complex actions will be developed in future work. A preprocessing procedure is carried out to smooth the players' trajectories, and after that, the players' velocities are calculated to control the players' motion direction and motion type (e.g. running, walking, etc.). However, the players' poses are hardly to be recovered, since lots of information is lost during imaging.

Our system is not fully automatic, mainly due to the confusion before a goal that makes the tracking failure, and the difficulty to automatically determine the ball's starting and ending points on the playfield. These cases can be resolved by human intervention. In the prototype, the users can control a virtual camera to pan, tilt, zoom and change viewpoint, and can also roam around the playfield using the direction key and the mouse.
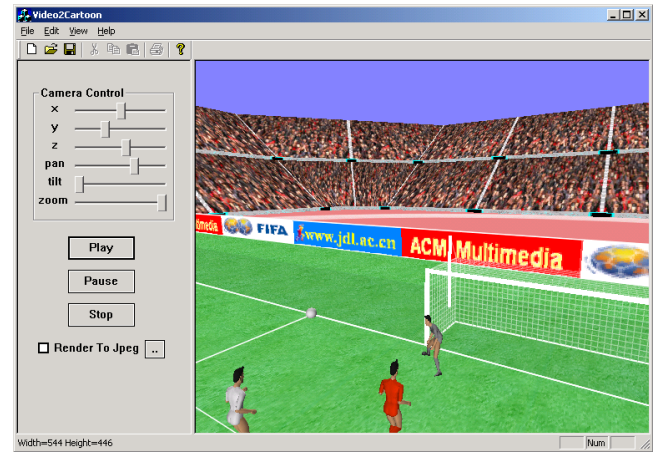

Figure 2. A snapshot of the prototype.

6. ACKNOWLEDGEMENTS

7. REFERENCES

[1] T. Bebie and H. Bieri, SoccerMan-reconstructing soccer games from video sequences, Proc. of ICIP, 898 – 902, 1998.

[2] T. Kim, Y. Seo, and K. Hong, Physics-based 3D position analysis of a soccer ball from monocular image sequences, Proc. of ICCV, 721 – 726, 1998.

[3] Y. Liu, Q. Huang, Q. Ye, and W. Gao, a new method to calculate the camera focusing area and player position on playfield in soccer video, Proc. of VCIP, 2005.

[4] K. Matsui, M. Iwase, M. Agata, T. Tanaka, and N. Ohnishi, Soccer image sequence computed by a virtual camera, Proc. of CVPR, 860 – 865, 1998.

[5] X. Yu, X. Yan, T. S. Hay, and H. W. Leong, 3D reconstruction and enrichment of broadcast soccer video, Proc. of ACM Multimedia, 2004.

[6] G. Zhu, D. Liang, Y. Liu, Q. Huang, and W. Gao, Improving Particle Filter with Support Vector Regression for Efficient Visual Tracking, to appear in Proc. of IEEE ICIP, 2005.