Automatic Tracking of Soccer Players

R. M. L. Barros\textsuperscript{1}, P. J. Figueroa\textsuperscript{2}, R. Anido\textsuperscript{2}, S. A. Cunha\textsuperscript{3}, M. S. Misuta\textsuperscript{1}, N. Leite\textsuperscript{2}, E. C. Lima Filho\textsuperscript{1}, R. Brenzikofer\textsuperscript{4}

\textsuperscript{1}Laboratório de Instrumentação para Biomecânica – UNICAMP - Brasil
\textsuperscript{2}Instituto de Computação – UNICAMP
\textsuperscript{3}Departamento de Educação Física – UNESP – Rio Claro
\textsuperscript{4}Instituto de Física “Gleb Wataghin” – Unicamp

Introduction

Soccer is a very popular sport in the world. In order to make possible an objective analysis of the game’s dynamics, it is important to develop methodologies able to detect and codify the great amount of information available during a match. Continuous data about position, velocity and acceleration of the players, in function of time, allow studies about biomechanical, physiological or strategic aspects. The automatic tracking of players from video records is a challenging problem due to the changes of the players’ features along the sequences of images, the players’ mutual occlusion and the long duration of the game.

The first studies about the player’s movement during the game was made by Reilly et al. (1976). They used audio recorder in order to register the estimated player position. Withers et al. (1982) already used a camera to track the movement of a unique player. Mayhew and Wenger (1985) also used a camera to track two players where each one was filmed alternately for 7 minutes. Then they computed the time spent in each activity, such as staying, walking, jogging, running and the number of times that occurred each activity.

In the recent years, there has been developed many methodologies in order to make possible the kinematics analysis of the players during game. For example, the system proposed by (Patridge, Mosher and Franks, 1991) and applied in the World Cup of 1990, uses the sequence of images captured by TV cameras for the annotations of the events and the player position. In the works presented by (Intille and Bobick, 1994; Taki, Hasegawa and Fukumura, 1996; Seo et al., 1997;Matsui et al., 1998) are used some image processing and computer vision techniques to automate the extraction of information about players position. The system presented by (Taki et al., 1996) is more similar to our system, because they also used many static cameras and is interested in getting the position of all players. However they can track automatically only isolated players, and for that they use just template matching technique.

The purpose of this research is the development of a computational environment for automatic soccer players tracking.

Methods

A first division Brazilian championship game was recorded using four stationary digital video cameras located in the highest place of the stadium. Each of them covered approx. a quarter of the pitch, including extra overlapped region. The sequences of images were analysed at 15 Hz frame rate. To make possible the players tracking in a complex scene, the following problems had to be dealt with: 1) Motion filtering: used to separate the moving objects (players) and the pitch, which consists of generating recursively a median image and subtracting of each frame. 2) Players’ segmentation: aiming to extract regions representing players in the image, the main tools are based on morphological operators. The segmentation algorithm consists of the following step: 2a) morphological filters (opening and closing), in order to eliminate noises; 2b) threshold the image in order to binarize; 2c) getting segmented regions as blobs; filtering region considering their size and form. 3) 2D Image-object transformations: in order to obtain the players’ representative co-ordinate of the segmented regions in the real game pitch. 4) Players’ trajectories: considering the existing redundancy due to overlapped regions, the smoothness of the trajectories and the limited velocity of the players, the trajectory of each player is identified.
Results and Discussion

The calibration, segmentation and tracking procedures have been applied in four sequences of 1800 frames each, which corresponds to 2 minutes of game. Each sequence represents one camera, so after tracking we reconstruct and merge the trajectories in order to have only one trajectory for each player.

The figure 1 illustrates the selected region and the segmentation algorithm applied to one sequence. On the right size, the figure shows the segmented blobs.

![Figure 1: (left side): One camera view and figure 1b: (right size) the segmentation using: 1) motion filtering and 2) players’ segmentation.](image)

The Figure 2 shows the trajectories of all players that we tracked without stopping (automatically), in two different cameras.

![Figure 2: Results of two minutes of tracking procedures applied to two sequences of images (2 cameras)](image)

The tracking was successful for all situation where there was a occlusion of one or two players, and some cases where were three players together. For more complex situation, when there are more than three players together in the same blob, we have to stop the tracking and start again after skipping some frames. Because the success of the tracking depends on the blobs splitting algorithm, the further improvement intend to treat this problem.

References


**Acknowledgements**

Research supported by FAPESP, CAPES, FAEP-UNICAMP and DAAD