Soccer Image Sequences Mosaicing Using Reverse Affine Transform

Ho-Sub Yoon*, Jung Soh, Byung-Woo Min, and Young-kyu Yang
Computer Software Technology Laboratory
Electronics and Telecommunications Research Institute
161 Kajong-Dong Yusung-Gu, Taejeon, 305-350, KOREA
Tel + 82-42-860-5233, Fax+ 82-42-860-4844
E-mail : yoonhs@etri.re.kr

Abstract: In this paper, we develop an algorithm of soccer image sequences mosaicing using reverse affine transform. The continuous mosaic images of soccer ground field allows the user/viewer to view a "wide picture" of the player's actions

The first step of our algorithm is to automatic detection and tracking player, ball and some lines such as center circle, sideline, penalty line and so on. For this purpose, we use the ground field extraction algorithm using color information and player and line detection algorithm using four P-rules and two L-rules. The second step is Affine transform to map the points from image to model coordinate using pre-defined and pre-detected four points. General Affine transformation has many holes in target image. In order to delete these holes, we use reverse Affine transform. We tested our method in real image sequence and the experimental results are given.

1. Introduction

In the most televised sporting events, from soccer to football to hockey, it's not always possible to capture every moments of action from every perspective. After all, camera set at a fixed perspective can only see so much at one time, and they can easily be blocked or shielded by any assortment of interference. It's been a longtime desire of many producers, commentators and sports aficionados alike to be able to get more than that stiff 20D view of their favorite sports events and actually get "in the game," if only in visualized fantasy.

In this paper, we develop a algorithm of mosaicing and panorama view generation from soccer game image sequences. A panorama scene presents a larger view of soccer field and allows the user/viewer to view a "wider picture" of the actions. In addition, by overlaying the motion trajectories of the players and the soccer, we are able to convey a more unbiased feel of the field and see the relative paths of individual players and soccer in motion beyond the narration of the video.

The first important component of mosaicing and panorama view generation is to automatically detect and track player, ball and some lines such as center circle, sideline, penalty line and so on. The second important component is to find image to model transformation for automatically generate mosaicing and panorama view

Previous work has been done that the parsing of soccer game to determine the position of the field by Gong, et. al[1] and visual tracking of American football under "closed world" assumption by S. S. Intill and Bobick[2]. Gong's system can classified a sequence of soccer frames into various play categories, such as shut at the left goal, top left corner kick, play in right penalty area, in mid field, based on a prior model comprising four major components: a soccer court, a ball, the players and the motion vectors. Intill's approach is focused tracking method in the "football domain", but same techniques such as closed-world analysis and context-specific tracking can apply in "soccer domain".

The panorama view generation methods are proposed by L.Teodosio and W. Bender[3] by spatially and temporally the individual frames on top of each other, tacking into account the global camera actions(pan and zoom).

Some similar previous works are presented by D. Yow[4] and Y. Seo[5] in field of soccer game analysis. D.Yow's system present to detect and extract the soccer highlights by analyzing the image contents, and to present these shots of actions by the panoramic reconstruction of selected events. The main difference between proposed system and Yow is panoramic view generation algorithm for image to model transformation. Finally, Y. Seo and al are coworkers in our project during 1998.

The rest of the paper is organized as follows. In section 2, the ground segmentation algorithm is introduced by two steps; ground field extraction and player and line detection. In section 3, Affine transformation model is described using pre-defined four points. The section 4th topic is reverse Affine transform method to delete the holes area in target image. Finally, the conclusion is summarized in section 5.

2. Ground Segmentation

The soccer court provides us with the most significant and reliable clues for soccer scene analysis as we can get the position of play by recognizing the white line marks on the ground. Since the cameras usually focus on the players carrying the ball, most of the video frames captures only a portion of the court. Therefore, the first step for soccer game analysis is to extract the ground region and then, several tools have been developed and combined to automatically extract players and lines from ground region foe a soccer sequence. The algorithm proceeds as follows:

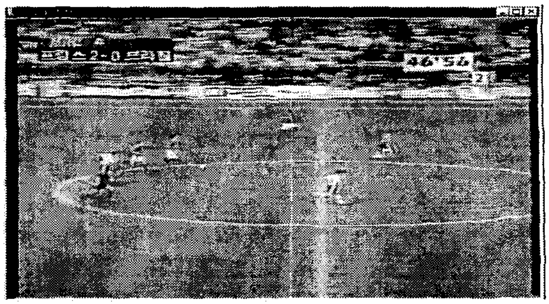
2.1 Ground fields extraction

The ground field has uniform color of strong green and occupies large area in the video sequence. From this knowledge, we make the R, G, B, color histogram for detecting ground region information. That means the peak value R_{peak}, G_{peak}, B_{peak} of color histogram are represented the ground color information because of assumption "the lager area from input image is ground area". Following equation is showed the ground field extraction method.

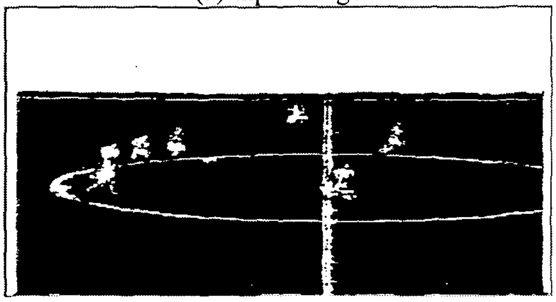
$$O(x,y)=\begin{cases} & |I_R(x,y)-R_{peak}|<Rt \\ & |I_G(x,y)-G_{peak}|<Gt \\ 1: & |I_B(x,y)-R_{peak}|<Bt \\ & I_G(x,y)>I_R(x,y) \\ & I_G(x,y)>I_B(x,y) \\ 0: & otherwise \\ & and\ G(x,y)<Gt \end{cases}$$

Where O(x,y) means binarized output image and I_R, I_G, I_B mean R, G, B value in a one pixel , and R_t, G_t, B_t mean threshold value for R, G, B. And G(x,y) means the gray level information and Gt means the threshold value for G(x,y). From 1th to 3th conditions are checked the ground color and condition 4th and 5th conditions are checked strong green color. The final condition discriminate between line and ground using gray level information.

Figure 1 (a) shows example of input image and (b) shows the result of this equation.



(a) Input image



(b) Output image

Figure 1. The result of ground extraction algorithm

2.2 Player and line detection

From the ground field extraction algorithm, white regions are divided into two segments with player and line. For separate the two segments, we initially process the 2-pass labeling algorithm and Hough transformation. The results of these algorithm have some information from isolated white segments such as centroid, MBR(minimum bounding rectangle), area(# of pixel), average gray level, average color level of R, G, B, Hough value about line segments and so on. From this information, we can distinguish the player and line segments.

Following player rules(P-Rule) are used to discriminate player regions.

P-Rule 1 (about size) : check size of MBR and area

P-Rule 2 (about shape) : check compactness

P-Rule 3 (about color) : check R, G, B distribution

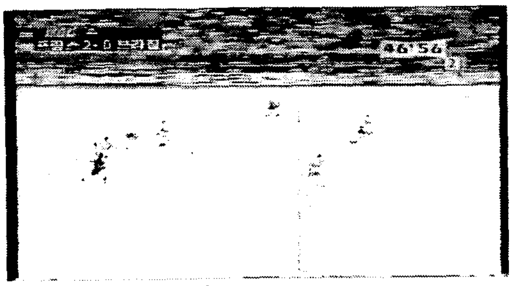
P-Rule 4 (about elongateness) : check versa vertical and horizontal length.

These rules present the characteristic information about player region which exist soccer game image sequence. In addition, following line rules(L-Rules) are used to separate line regions.

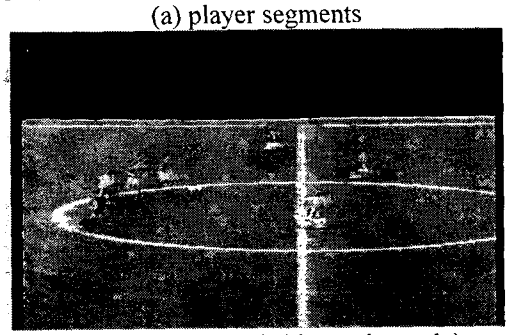
L-Rule 1 (about gray level) : check gray level

L-Rule 2 (about line segments) : check Hough value

These two type rules divide candidate region into player or line regions. Also, if one region is not satisfied condition of the some P-Rule then that region will be included in line region. Figure 2 shows the results of player and line detection.



(a) player segments



(b) Line segments (white region only)

Figure 2. The result of ground extraction algorithm

3. Affine Transform

For the affine transform, we define model coordinate and image coordinate

$$\text{Model Coordinate} \qquad \text{Image Coordinate}$$

$$\begin{bmatrix} x_{2i} \\ y_{2i} \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & 1 \end{bmatrix} \begin{bmatrix} x_{1i} \\ y_{1i} \\ 1 \end{bmatrix}$$

$$x_{2i} = \frac{a_{11}x_{1i} + a_{12}y_{1i} + a_{13}}{a_{31}x_{1i} + a_{32}y_{1i} + 1}, \; y_{2i} = \frac{a_{21}x_{1i} + a_{22}y_{1i} + a_{23}}{a_{31}x_{1i} + a_{32}y_{1i} + 1}$$

From this equation, the coordinate mapping problem is convert to the 8 variable searching problems from a_{11} to a_{32}. These 8 variable can be calculated by following equation using 4 point from (x_{11},y_{11}) to (x_{14}, y_{14}) which detected by image processing.

$$x_{1i}a_{11} + y_{1i}a_{12} + a_{13} - x_{1i}x_{2i}a_{31} - y_{1i}x_{2i}a_{32} = x_{2i}$$

$$x_{1i}a_{21} + y_{1i}a_{22} + a_{23} - x_{1i}y_{2i}a_{31} - y_{1i}x_{2i}a_{32} = y_{2i}$$

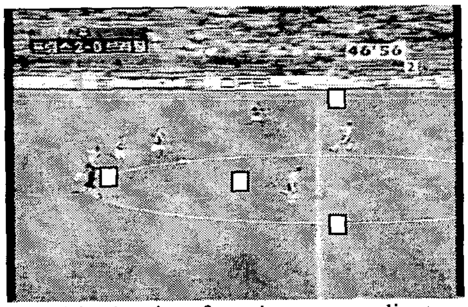
$$\begin{bmatrix} x_{11} & y_{11} & 1 & 0 & 0 & 0 & -x_{11}x_{21} & -y_{11}x_{21} \\ 0 & 0 & 0 & x_{11} & y_{11} & 1 & -x_{11}x_{21} & -y_{11}y_{21} \\ x_{12} & y_{12} & 1 & 0 & 0 & 0 & -x_{12}x_{22} & -y_{12}x_{22} \\ 0 & 0 & 0 & x_{12} & y_{12} & 1 & -x_{12}x_{22} & -y_{12}y_{22} \\ x_{13} & y_{13} & 1 & 0 & 0 & 0 & -x_{13}y_{23} & -y_{13}x_{23} \\ 0 & 0 & 0 & x_{13} & y_{13} & 1 & -x_{13}y_{23} & -y_{13}y_{23} \\ x_{14} & x_{14} & 1 & 0 & 0 & 0 & -x_{14}y_{24} & -y_{14}x_{24} \\ 0 & 0 & 0 & x_{14} & y_{14} & 1 & -x_{14}y_{24} & -y_{14}y_{24} \end{bmatrix} \begin{bmatrix} a_{11} \\ a_{12} \\ a_{13} \\ a_{21} \\ a_{22} \\ a_{23} \\ a_{31} \\ a_{32} \end{bmatrix} = \begin{bmatrix} x_{21} \\ y_{21} \\ x_{22} \\ y_{22} \\ x_{23} \\ y_{23} \\ x_{24} \\ y_{24} \end{bmatrix}$$

$$Aq = z, \quad q = A^{-1}z$$

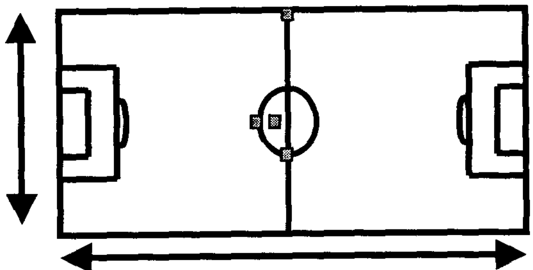
Using simple 4 points from (x_{11}, y_{11}) to (x_{14}, y_{14}) from image coordinate figure 3 (a), ground model coordinate are transformed to figure 3 (b) using Affine transformation, then all image points are transformed from ground image to model coordinate.

Therefore, the center circle, centerline and sideline detection algorithm is needed. The sideline and centerline detection algorithm can be easily developed using histogram projection because they have strong line segment. But center circle detection algorithm have complex processing because occlusion and its circumstance. Figure 3 (c) shows the results of detected side line and ceneterline and circle.

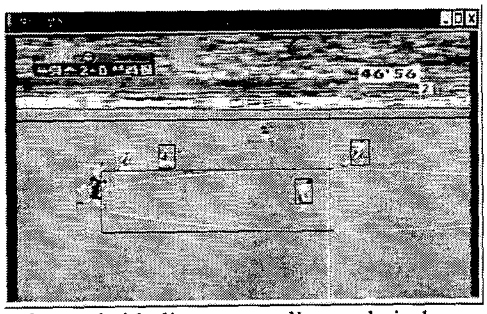
Additionaly, the basic 4 points are determined symatric location based on centerline position. If centerline exist on right side in the input image then left center circle detected and else right center circle deteced. Figure 3 (d) shows the result of Affine transfornation.



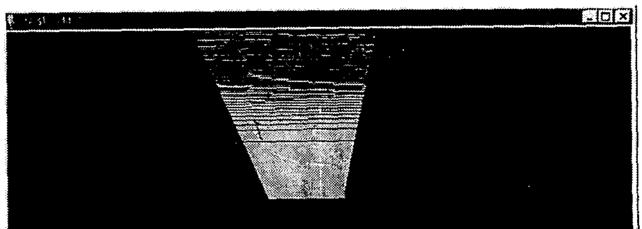
(a) 4 points from image coordinates



(b) 4 points from ground model coordinates



(c) deteced side line, center line and circle



(d) The result of Affine transform

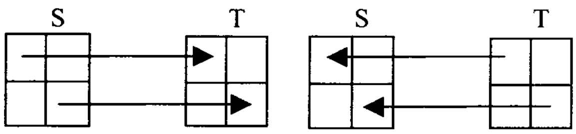
Figure 3. The step of the Affine transfornation.

4. Reverse Affine Transform

The results of affine transform as Figure 3 (d) have many holes in the model coordinate. There are two problems with forward mapping: holes and overlaps. Holes are pixels that are undefined, and the destination pixel has no corresponding source pixel. Overlaps occur when two input pixels get mapped to the same output pixel.

There are several interpolation algorithms to solve this problem such as nearest neighbor, bilinear, higher order, cubic convolution, B-spline, but more efficiency method is used to reverse transform.

Reverse mapping traverses the destination image and calculates via some inverse transform which pixel in source image will be used to produce the destination pixel. Computing destination image in this manner eliminates the problems of holes and overlaps. Figure 4 shows the mapping direction for coordinate transform and figure 5 shows the results of reverse affine transform.



(a) Forward mapping (b) Backward mapping

Figure 4. The mapping direction

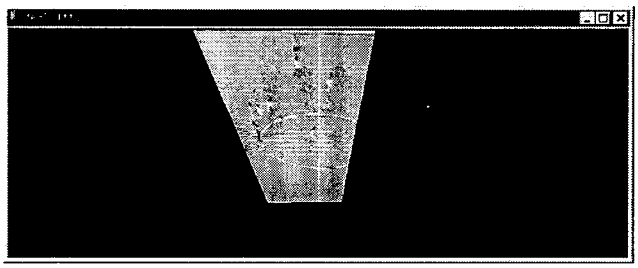


Figure 5. The result of reverse Affine transform

5. Conclusion

The proposed system was implemented on a personal computer with an image capture board (Matrox Meteor II). Also, an input image sequences were captured by a VHS tape from VTR with the resolution 640x480. The computing power is 7 ~ 8 frames per second using dual Pentium II 400 Mhz. Since our computer can process over 5 frames per second, it is possible to view the proposed system for real-time mosaic generation. The recognition software is implemented in Visual C++ 6.0 on Windows 98.

This paper describes an algorithm of soccer image sequences mosaicing using reverse affine transform. General Affine transfer's results are two problems: holes and overlaps. Holes are pixels that are undefined, and the destination pixel has no corresponding source pixel. Overlaps occur when two input pixels get mapped to the same output pixel. In order to solve these holes, we use reverse Affine transform. Reverse Affine transform is more time consuming processing because of twice matrix calculation than general processing, but it's results have good quality image than the other interpolation algorithm

There are some future works, we will develop algorithm which can deal with more complex problems such as occlusion reasoning of more complex case, tracking ball position, automatic scaling of player size, analysis of player motion(kick, dribble, run, walk, shoot), virtual 3D replay and so on.

References

[1] Y. Gong, T.S. Lim, H.C. Chua, H. Zhang, and M. Sakauchi, "Automatic Parsing of TV Soccer programs", Second International Conference in Multimedia Computing and Systems, pp. 167-174, May 1995

[2] S. Intelle, A. Bobick, "Visual Tracking Using Closed-Worlds", MIT Media Lab. Perceptual Computing Section Technical Reports No. 294, November 1994.

[3] L. Teodosio, W. Bender, "Salient Video Stills:Content and Context preserved", ACM Multimedia, pp.39-46, June 1993.

[4] Dennis Yow, Boon-Lock Yeo, Minerva Yeung and Bede Liu, "Analysis And Presentation of Soccer Highlights From Digital Video", The Second Asian Conference in Computer Vision(ACCV'95).

[5] Y. Seo, S. Choi, H. Kim and K.S. Hong, "Where are the ball and players ? Soccer game analysis with color-based tracking and image mosaic," International Conference on Image Analysis and Processing, Sept. 1997, Florence, Italy

[6] Randy Crane, *A simplified approach to Image Processing*, Prentice Hall PTR, pp.206-211.1997.

[7] S. Watanabe, T. Sato, M. Sengoku, and T. Abe, "Equalization scheme on two dimensional orthogonal coding multi-carrier CDMA," Proc. ITC-CSCC'97, Okinawa, Japan, Vol.2, pp.1121-1124, July 1997.