

방송 축구 영상으로부터 3차원 애니메이션 변환을 위한 축구 선수 동작 인식

장원철, 남시욱, 김재희

연세대학교 전기·전자공학과 컴퓨터 비전 연구실

Pose Recognition of Soccer Players for Three Dimensional Animation

Wonchurl Jang, Siwook Nam, Jaihie Kim

Computer Vision Lab., Dept. of Electrical and Electronic Eng., Yonsei Univ.

E-mail: jwchappy@nownuri.net

Abstract

To create a more realistic soccer game derived from TV images, we are developing an image synthesis system that generates 3D image sequence from TV images. We propose the method for the team and the pose recognition of players in TV images. The representation includes camera calibration method, team recognition method and pose recognition method. To find the location of a player on the field, a field model is constructed and a player's field position is transformed by 4-feature points. To recognize the team information of players, we compute RGB mean values and standard deviations of a player in TV images. Finally, to recognize pose of a player, this system computes the velocity and the ratio of player(height/width). Experimental results are included to evaluate the performance of the team and the pose recognition.

I. 서론

최근 축구경기의 자동분석에 대한 관심이 고조됨에 따라 그에 대한 연구가 활발히 수행되고 있다[3][5][6][7]. 축구 경기는 제한된 수의 카메라로 촬영하기 때문에 카메라가 미리 배치되지 않은 각도에서의 경기 장면은 볼 수 없다. 선수의 위치 정보와 동작정보를 이용하여 축구 경기의 주요장면을 3차원 애니메이션으로 재연하게 되면 자유로운 각도로 경기장면을 다시 볼 수 있게된다[1][2]. 그러나 실제 3차원 애니메이션을 하기 위해서는

축구 영상으로부터 모션 캡처에서와 같은 선수 관절의 동작에 대한 자세한 정보를 추출해야 한다. 하지만 축구 영상에서 선수영역이 작고 카메라와 선수의 움직임에 의한 번짐(blur)이 발생하기 때문에 선수에 대한 자세한 정보를 추출하는 것이 어렵다. 본 논문에서는 3차원 애니메이션을 위한 선수의 동작인식 과 팀 인식 방법에 대해 제안한다.

II. 시스템 개요

제안한 시스템에서는 방송 축구 영상이 입력으로 들어오면 경기장에서 선수 위치를 추정하고, 각 선수의 동작을 인식한다. 선수 위치와 동작 정보를 3차원 그래픽 합성 시스템에 입력하면 원하는 각도에서 경기장면을 재연하게 된다. 그림 1은 축구 영상의 3차원 그래픽 애니메이션 시스템의 개요를 나타낸 것이다.

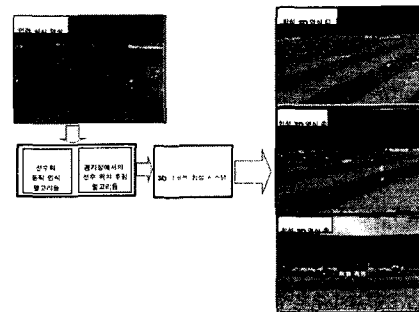


그림 1. 축구영상의 3D그래픽 애니메이션 시스템

III. 경기장에서 선수 위치 추정

축구 영상으로부터 경기장에서의 물체 위치를 추정하기 위해서는 영상과 경기장간의 위치 변환 관계를 파악해야 하고, 이를 위해서는 기준이 되는 특징점을 정의해야 한다. 본 논문에서는 그림 2에서와 같이 경기장 전체에서 43개의 특징점을 정의한다. 축구 영상에 4개 이상의 특징점이 존재하고 그 중에서 3개 이상이 한 직선위에 존재하지 않으면 투영 변위필드모델에 의해 경기장에서의 위치를 직접 구할 수 있다[4].

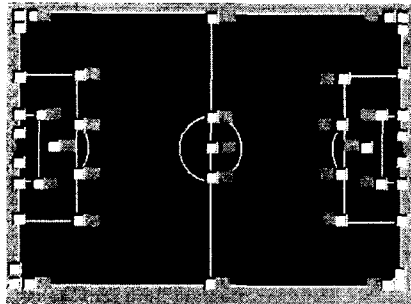


그림 2. 경기장에서의 43개의 특징

IV. 선수 동작 인식 및 팀 인식

4.1. 선수 영역 추출

축구영상에서는 선수의 크기가 작고, 선수와 카메라의 움직임에 의한 번짐이 존재하기 때문에 단순한 이진화(threshold) 방법으로 선수를 추출하는 것은 어렵다. 본 논문에서는 경계 검출(sobel)을 한 후 클로징(closing)을 사용해 분할하여 선수 영역의 왜곡을 줄였다[1]. 선수 추출 결과는 아래와 같다.



그림 3. 물체 분할(Segmentation) 순서

- <단계 1> 축구 영상에 대해 흑백 영상으로 변환한다.(그림 4.(a))
- <단계 2> 소벨 에지 검출을 한다. 컬러 정보를 이용해 이진화하는 경우 정확한 임계치 설정이 어렵기 때문에 배경과의 차이를 찾아내는 에지 검출 방법을 사용했다.(그림 4.(b))
- <단계 3> 클로징 처리를 해준다. 에지 검출시 선수의 다리와 같은 관절 부분이 몸통에서 분리되는 현상이 종종 생기는데 이를 해결하기 위해 클로징을 행한다.(그림 4.(c))
- <단계 4> 크기 필터링(size filtering)을 통해 선수 이외의 것들을 제거 해준다.(그림 4.(d))

$$S = \max_{i \in N} \{G(i)\}$$

G(i): 색인별 사이즈, N:전체 색인 값, S: 선수 영역.

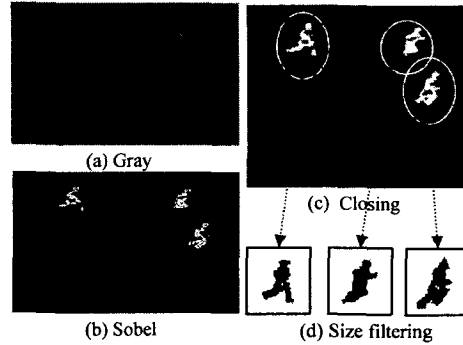
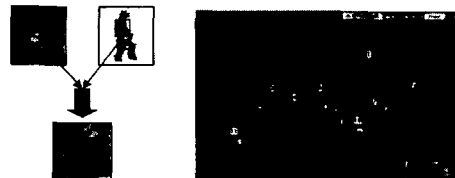


그림 4. 선수 추출 결과

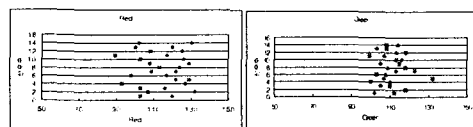
4.2 팀 인식

축구 영상의 선수들로부터 팀의 분류는 그림 5.(a) 와 같이 영상에서 분할된 선수들의 상체 부분의 RGB 컬러 각각의 평균과 편차를 사용한다. 선수의 상체 영역을 사용하는 이유는 두 팀의 유니폼 부분이 팀을 구분하는데 적합한 특징이기 때문이다. 경기장에 있는 모든 선수(그림 5.(b))에 대해 그림5.(a)처럼 선수의 몸통에서 15×15 마스크(mask)를 사용해 RGB 값을 추출하고 평균을 구한다. 그림 5.(c)와 같이 RGB 각각의 평균값들의 분포들이 생긴다. 여기서 분산이 가장 큰 컬러를 선택하여 팀 인식을 위한 기준 컬러로 선택한다. 선택된 컬러의 평균을 팀 인식의 임계치로 정의하여 두 팀을 구분한다.



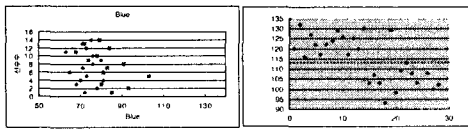
(a) 선수 상체 영역

(b) 경기 장면

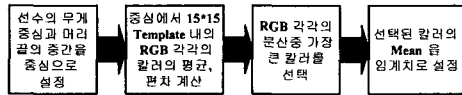


<Red>

<Green>



(C) 선수들의 RGB 평균값 분포 와 기준 칼러



(d) 팀 인식 방법
그림 5. 팀 인식 방법 및 과정

4.3 선수 동작 인식

영상을 3차원 애니메이션하기 위해서는 선수의 동작을 인식해야 한다. 선수 동작인식 범위를 그림 6과 같이 1단계에서 4단계로 구분하였다.

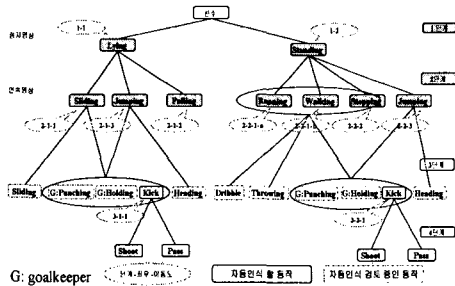


그림 6. 선수 동작 인식 범위

단계 1은 정지 영상에서 얻는 동작으로 ‘서있는 자세’, ‘누워있는 자세’로 정의한다. 단계 2는 연속 영상에서 얻는 동작으로 “뛰다”, “걸는다”, “멈춘다”로 정의한다. 단계 3은 공과 관련된 자세로 “공을 차는 자세”, “공을 가지고 달리는 자세”로 정의한다. 본 논문에서는 단계 1, 단계 2 동작 인식에 대한 방법을 제안한다.

4.3.1 단계 1동작 인식

단계 1동작(standing, lying) 인식은 정지 영상에서 추출된 선수의 수직, 수평 투영(projection) 결과를 이용한다. 그림 7과 같이 추출된 선수에 대해 수직, 수평 투영하여 선수의 가로, 세로축 평균 길이를 계산한다.

$$ratio = \frac{\frac{1}{N} \sum_{i=0}^N V(i)}{\frac{1}{M} \sum_{j=0}^M H(j)} \quad (1)$$

H(j): 수평축 투영값. V(i): 수직축 투영값.
N: 상위 5%, 하위 5%의 값을 제외한 수평축 길이.

M: 상위 5%, 하위 5%의 값을 제외한 수직축 길이.

식(1)을 사용해 축구 영상에서 선수의 자세 비율을 계산하여 임계치에 따라서 ‘서있는 자세’와 ‘누워있는 자세’를 인식하게 된다. 임계치는 0.96을 사용한다. 이 임계치는 120명의 선수를 비교하여 얻은 결과이다.

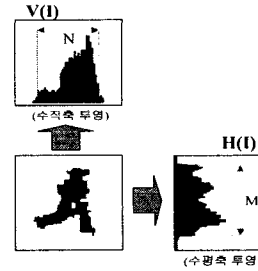


그림 7. 선수 투영도

4.3.2 단계 2동작 인식

단계 2동작(run, walk, stop)은 연속 영상에서 얻어진 선수의 경기장에서 위치를 이용하여 속력을 계산해 인식하게 된다. 동작인식을 위한 속력 검출 과정은 그림 8과 같다.



그림 8. 단계 2동작 인식과정

화면에서 선수를 추적할 때 다른 선수와 겹침, 선수의 비선형 이동 등으로 선수의 위치가 갑자기 크게 변하는 경우가 생긴다. (그림 10)

이를 해결하기 위해 미디언(median) 필터링과 평균(mean) 필터링을 연속으로 사용한다. 필터 처리된 선수의 경기장에서 위치를 이용해 선형 근사화를 한다. 선형 근사화를 통해 선수의 이동 방향이 크게 변하는 프레임을 계산해 낸다. 그리고 이를 이용해 선형적으로 이동한 소구간을 정의한다. 소구간에서 선수의 평균 이동 속력을 계산하여 단계 2동작을 판단한다.

V. 실험 결과

4.2 절에서 기술한 선수의 상체 부분의 RGB 평균과 편차를 이용해 20경기 207명에 대해 실험한 결과 팀 인식률은 82.6%였다. 17.4%의 에러가 발생한 이유는 선수 수가 적은 경기(4~6명)에서는 샘플수가 부족하여 임계치의 정확성이 떨어지기 때문이다.

단계 1동작을 인식하기 위해서 식(1)을 사용해 선수의 자세 비율을 계산한 결과 그림 9의 결과를 얻었다. 그림 9의 결과는 선수자세 비율(선수의 세로/가로)이 누워있는 자세(0.5~0.7)와 서있는 자세(1.2~2.2)로 크게 두 그룹으로 분포한다. 그림 9의 결과를 이용하여 '서있는 자세'와 '누워있는 자세'를 구분하기 위한 임계치를 식(2)에 의해 0.96으로 정했다. 이때 인식 어려움은 5%(120개중 6개)였다.

$$P(A) \frac{1}{\sigma_A \sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x - \mu_A}{\sigma_A})^2} = P(B) \frac{1}{\sigma_B \sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x - \mu_B}{\sigma_B})^2} \quad (2)$$

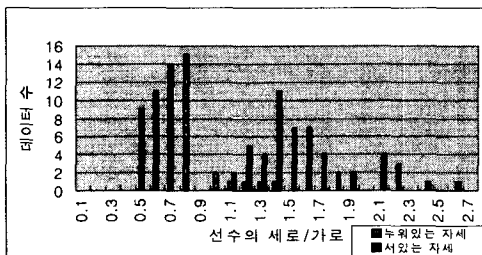


그림 9. 선수 자세 비율 분포

단계 2동작을 인식하기 위해서 선수의 이동 속력을 이용한다. 그림 10에 표시된 잡음(noise) 때문에 선수의 이동 속력을 정확히 계산하기가 힘들다. 단계 2동작 인식에 적합한 이동 궤적을 얻기 위해 4.3.2절에 제안한 필터링 방법을 사용한 결과 그림 10과 같이 잡음의 영향을 제거할 수 있었다.

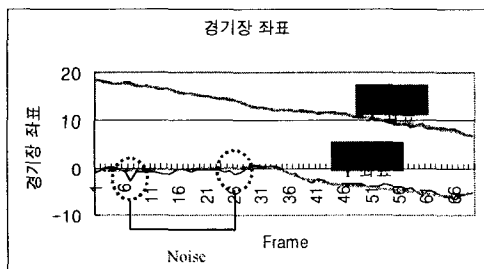


그림 10. 선수의 이동 궤적

VI. 결론

본 논문에서는 경기장에서 선수의 위치를 추정하고, 정지 영상과 연속 영상에서 선수의 팀과 동작을 인식하는 것에 대해 연구했다. 동작 인식의 실험 결과를 보면 팀 인식은 82.6%였고, 단계 1동작은 95%의 인식율을 보였다. 이 인식율의 성능을 향상시키기 위해 입력 영상에서 선수를 분

할하는 방법의 개선이 필요하며, 더 정확한 인식율 계산을 위해 다양한 축구 영상에 대한 실험이 필요하다. 이번 연구를 바탕으로 앞에서 제안한 선수의 '공을 차는 자세', '점프 자세' 등 연속 영상에서 얻은 정보를 이용해 선수 동작을 인식하는 방법에 대한 연구가 필요하다. 마지막으로 단계 2동작의 인식 결과에 대해 정량적 검증 방법이 필요하다.

VII. 참고 문헌

- [1] Bebie, T. Bieri H. "SoccerMan Reconstruction Soccer Games from Video Sequences" Image Processing, 1988. ICIP 98. Proceedings. 1988 International Conference on Volume: 1, pp. 898-902, 1988
- [2] Matsui, K. Iwase, M. Agata, M. Tanaka, T.T. Ohnishi, N. "Soccer Image Sequence Computed by a Virtual Camera" Computer Vision and Pattern Recognition, 1998. Proceedings. 1998 IEEE Computer Society Conference on, 1998, pp. 860-865.
- [3] Yihong Gong, Lim Teck Sin, Chua Hock Chuan, Hongjiang Zhang, Masao Sakauchi. "An Automatic Video Parser for TV soccer Games" Multimedia Computing and Systems, 1995., Proceedings of the International Conference on, 1995, pp. 167-174
- [4] 남시욱, 이재철, 이종현, 이재호, 이상길, 김재희, "투영 변위 필드 모델을 이용한 물체 이동거리 추정 및 추정 정확도의 정량적 평가", 1999년 인공지능, 신경망 및 퍼지시스템 종합학술대회 논문집, pp.275-278, 1999년 10월.
- [5] S.S Intille, A.F. Bobick, "Tracking using a local closed-world assumption: Tracking in the football domain", MIT Media Lab Perceptual Computing Group, TR No.296, MIT, August, 1994
- [6] 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, Florence, Italy, 1997
- [7] T. Kim and K.-S. Hong, "Physics-based 3D position analysis of a soccer ball from monocular image sequences" International Conference on Computer Vision, Bombay, India, pp.721-726, Jan. 1998