

축구 동영상에서의 장면 구조 분석에 기반한 자동적인 하이라이트 장면 검출

박 기 태[†] · 문 영 식^{**}

요 약

본 논문에서는 축구 동영상으로부터 자동적으로 하이라이트 장면들을 검출하는 방법을 제안한다. 축구 동영상에서 하이라이트 장면들은 슈팅 장면들이나 골 장면들로 정의 될 수 있다. 우리는 축구 동영상에 대한 구조적 분석을 통해서 일반적으로 골 포스트(goal posts) 영역 주위에서 하이라이트 장면들이 나타나는 것과 하이라이트 장면 이후에는 TV 카메라가 축구 선수들이나 관중석을 확대해서 보여주는 것을 확인할 수 있었다. 본 논문에서 축구 동영상으로부터 하이라이트 장면들을 검출하기 위한 과정은 세 단계로 구성된다. 첫 번째 단계는 통계적인 문턱치(statistical threshold)를 이용한 그라운드(playing field) 영역을 추출한다. 두 번째 단계는 골 포스트를 찾기 위해서 그라운드 영역과 그라운드 가 아닌 영역들의 경계선 부분을 검출한다. 그리고 마지막 단계에서는 축구 선수나 관중들의 확대 장면을 검출하기 위해서 그라운드 가 아닌 영역들에 대해서 connected component labeling 기법을 적용하여 한 장면 내에서 그라운드 가 아닌 영역들의 비율을 계산한다. 본 논문에서는 하이라이트 장면 검출에 대한 성능을 평가하기 위하여 정확률(precision)과 재현율(recall)을 사용하고, 실험을 통하여 제안된 방법이 정확률 95.2%, 재현율 85.4%로 축구 동영상에서 하이라이트 장면을 효과적으로 검출할 수 있음을 확인하였다.

키워드 : 하이라이트 장면 검출, 그라운드 검출, 골 포스트 검출, 비 그라운드 확대 검출

Automatic Detection of Highlights in Soccer videos based on analysis of scene structure

Park, Ki Tae[†] · Moon, Young Shik^{**}

ABSTRACT

In this paper, we propose an efficient scheme for automatically detecting highlight scenes in soccer videos. Highlights are defined as shooting scenes and goal scenes. Through the analysis of soccer videos, we notice that most of highlight scenes are shown around the goal post area. It is also noticed that the TV camera zooms in a soccer player or spectators after the highlight scenes. Detection of highlight scenes for soccer videos consists of three steps. The first step is the extraction of the playing field using a statistical threshold. The second step is the detection of goal posts. In the final step, we detect a zooming of a soccer player or spectators by using connected component labeling of non-playing field. In order to evaluate the performance of our method, the precision and the recall are computed. Experimental results have shown the effectiveness of the proposed method, with 95.2% precision and 85.4% recall.

Key Words : Detection of Highlight Scene, Playing Field Detection, Goal Posts Detection, Zooming Of Non-Playing Field

1. 서 론

최근 멀티미디어 서비스와 정보통신의 발달로 과거 통신망 상에서 단순히 문자 정보만을 제공하던 것과는 달리 발전된 형태의 동영상들을 순차적으로 제공하는 멀티미디어 서비스에 대한 관심이 증대되어 왔다. 특히 동영상 멀티미디어 정보는 그 양이 매우 방대하기 때문에 데이터베이스에 저장하기 위한 색인 방법이 매우 중요하다. 효율적인 색인

방법이 아닐 경우에는 동영상에서 원하는 위치를 찾는 데 많은 시간을 소비해야 하기 때문에 동영상을 보다 빠르고 정확하게 사용자가 원하는 위치를 검색 할 수 있는 효율적인 색인 기법필요하며, 이를 위해 동영상에 대한 구조 분석 작업이 필수적으로 이루어져야 한다[1-3]. 예로써, 축구 경기 동영상은 카메라의 움직임과 선수들의 움직임이 역동적이고 구조면에 있어서도 뉴스 동영상처럼 정형화되어 있지 않아서 구조적 분석이 쉽지 않다. 하지만 일반적으로 슈트(shoot) 또는 골(goal) 장면과 같은 하이라이트 장면들은 그라운드 영역과 비 그라운드의 경계선이 일정한 기울기를 갖고, 그 경계선 위에 골 포스트가 있다는 것과 골 포스트를 포함하

* 본 연구는 한국과학재단 특정기초연구(R01-2006-000-10876-0)지원으로 수행되었음.

† 준 회 원 : 한양대학교 대학원 컴퓨터공학과 박사과정

** 정 회 원 : 한양대학교 컴퓨터공학과 교수(교신저자)

논문접수 : 2006년 9월 20일, 심사완료 : 2006년 12월 28일

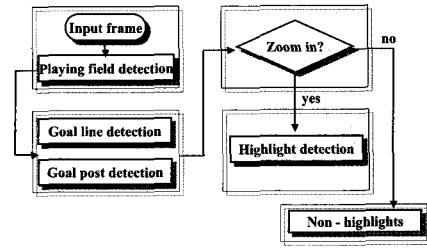
는 장면 이후 연속되는 장면들에서 선수나 관중석을 확대하는 것과 같은 구조를 축구 동영상에서 확인할 수 있다.

축구 동영상에서 하이라이트 장면을 검출하기 위해 선수의 움직임 추적, 공의 움직임 추적과 같은 움직임 벡터(motion vector) 정보를 이용하는 방법[4-6], 골 포스트를 검출하는 방법[7], 음성, 문자, 영상등의 다양한 특징값들의 정보를 이용한 방법[8], 장면 전환 검출 기법을 사용하여 축구 동영상의 구조적 특성을 이용한 방법[9]등이 제안되었다. 하지만, 움직임 벡터를 사용하는 방법은 축구 선수들의 유니폼 색상에 대한 사전 정보가 있어야 하고, 객체를 추적하기 위해 필요한 많은 특징을 얻어야 하기 때문에 계산이 복잡하고 시간이 많이 소요된다는 단점이 있다. [7]은 골 포스트를 찾기 위해 장면의 전체 영역에서 골 포스트 색상에 해당하는 화소들을 수평적으로 투영한 후 히스토그램을 구성하여 일정한 거리의 서로 다른 두 개의 극치(peak)를 검출한다. 그러나 이 방법은 모든 장면에 대해서 골 포스트를 찾는 방법이 적용되므로 연산량이 너무 많고, 장면 전체 영역에서 수평적으로 투영하기 때문에 잡음(noises)들로 인해 히스토그램에서 서로 다른 두 개의 극치를 추출하는데 어려움이 있다. [8]은 하이라이트 장면을 검출하기 위해서 선수의 확대 장면을 이용하는데, 축구 경기의 많은 확대 장면중에서 어떤 장면 이후에 확대 장면이라는 것이 명확하게 구분되어 있지 않다. 그러므로, 하이라이트가 아닌 장면인 경우에도 카메라가 응원중인 관중석을 비추게 되면 하이라이트 장면으로 오검출하는 문제점이 있다. [9]는 wipe 장면 검출, replay 검출 및 zoom-in/zoom-out 장면 전환 검출을 이용하여 하이라이트를 추출하는 방법을 제안하였다. 그러나 wipe 장면을 전체 동영상에 대해서 검색한 후 이웃한 wipe 장면을 묶어서 하나의 replay 장면으로 검출하기 때문에, 실시간적으로 하이라이트 장면을 검출할 수 가 없고, 또한 wipe 장면 패턴이 축구 동영상 마다 달라질 수 있기 때문에, 정확한 하이라이트 장면을 추출하기 힘들다. 그러므로, 본 논문에서는 축구 동영상에 대한 구조적인 분석을 통하여 일반적으로 축구 동영상에 나타나는 구조를 정형화하여 하이라이트 장면을 검출하는 방법을 제안한다.

본 논문의 구성은 2장에서 그라운드 영역 검출, 골 포스트 검출, 그리고 비 그라운드 영역을 확대한 장면을 검출하는 방법에 대해 기술하고, 3장에서는 다양한 환경을 가진 그라운드들을 포함하는 축구 동영상에 대해 제안한 방법을 이용한 실험 결과를 평가하고, 마지막으로 4장에서는 본 논문에 대한 결론을 맺는다.

2. 제안하는 방법

일반적으로 축구 동영상에서 하이라이트 장면이라고 하는 것은 슈팅 장면이나 골 장면으로 정의될 수 있다. 우리는 축구 동영상에 대한 구조적 분석을 통해서 일반적으로 골 포스트(goal posts) 영역 주위에서 하이라이트 장면들이 나타나는 것과 하이라이트 장면 이후에는 TV 카메라가 축구 선수들이나 관중석을 확대해서 보여주는 것을 확인할 수 있었다. 본 논문에서는 이러한 구조적인 정보를 이용하여 하이라이트 장면을 검출하고자 한다. 축구 동영상으로부터 하이라이트 장면을



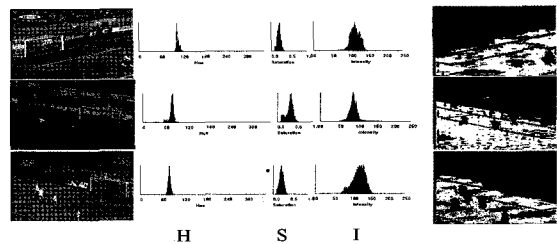
(그림 1) 제안한 방법에 대한 전체적인 흐름도

검출하기 위한 과정은 세 단계로 구성된다. 통계적인 문턱치(statistical threshold)를 이용한 그라운드(playing field) 영역을 추출한다. 두 번째 단계는 골 포스트를 찾기 위해서 그라운드 영역과 그라운드가 아닌 영역들의 경계선 부분을 검출한다. 그리고 마지막 단계에서는 축구 선수나 관중석의 확대 장면을 검출하기 위해서 비 그라운드 영역들에 대해서 connected component labeling 기법을 적용하여 한 장면 내에서 비 그라운드 영역들의 비율을 계산한다. (그림 1)은 제안한 방법에 대한 전체적인 흐름도를 보여준다.

2.1 그라운드 영역 검출(Detection of playing field)

일반적으로 축구 동영상에서의 그라운드는 다양한 색깔의 잔디, 날씨, 그리고 조명 상태와 같은 환경적인 변화들로 인해 한 가지 색이 아닌 다양한 색으로 보인다. 그러므로 기존 방법들에서 사용하고 있는 고정된 색상 정보를 이용하면 그라운드를 효과적으로 검출하기 힘들다. 따라서 본 논문에서는 고정된 색상의 문턱치 정보를 사용하지 않고, 환경적인 변화에 적응적인 그라운드 검출을 위해서 통계적인 문턱치를 계산한다. 이를 위해, HSI 색상 모델을 사용하여 각각의 채널에 대해서 히스토그램을 구성한 후에 각각의 히스토그램에서 극치(peak)값을 가지는 빈(bin)을 조사하여 일정한 문턱치 이상의 화소수를 갖는 극치 주변의 빈들에 대해서 통계적인 문턱치로 사용되는 평균(mean)과 표준편차(standard deviation)를 계산함으로써, 여러 환경에 적응적인 그라운드 영역을 검출한다. 각각의 채널에 대한 통계적인 문턱치는 식(1)에 의해 계산된다. (그림 2)는 식(1)을 사용하여 축구 동영상의 한 장면에서 그라운드 영역을 검출한 결과를 이진 영상으로 보여준다. 식(1)에서, μ , σ , 그리고 α 는 각각 평균, 표준편차, 그리고 가중치를 의미한다.

$$\begin{aligned}
 \mu_H - \alpha \cdot \sigma_H &\leq Hue \leq \mu_H + \alpha \cdot \sigma_H \\
 \mu_S - \alpha \cdot \sigma_S &\leq Saturation \leq \mu_S + \alpha \cdot \sigma_S \\
 \mu_I - \alpha \cdot \sigma_I &\leq Intensity \leq \mu_I + \alpha \cdot \sigma_I
 \end{aligned} \tag{1}$$

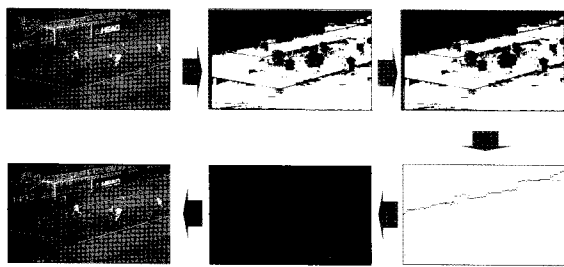


(그림 2) H, S, 그리고 I 채널에 대한 히스토그램 분포와 그라운드 영역 검출 결과

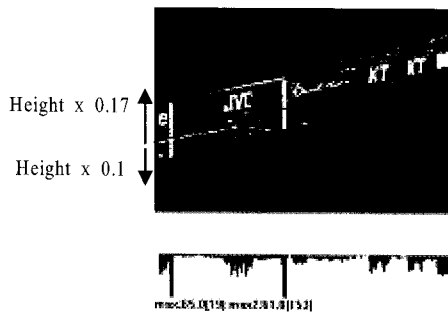
2.2 골 포스트 검출(Detection of goal posts)

하이라이트 장면 검출을 위한 두 번째 단계는 골 포스트를 검출하는 과정이다. 본 논문에서는 정확한 골 포스트를 검출하기 위해서 (그림 2)와 같이 첫 번째 단계에서 구성된 이진 영상으로부터 그라운드 영역과 비 그라운드 영역을 분할하는 경계선(goal line)을 먼저 찾는다. 이를 위해, 그라운드 영역과 배경 영역이 이웃한 화소들(pixels)을 찾고, 그런 다음, 해당 화소들에 대해 최소 제곱 평균 오차 접근법(least mean square error fitting)을 적용하여 하나의 직선(line)을 검출한다. 이 때, 직선이 수평선과 이루는 기울기(slop)가 0.15에서 0.5 사이의 값을 갖게 되면, 해당 직선을 goal line으로 결정한다. 그리고 goal line을 포함하는 장면을 골 포스트가 존재할 가능성이 있는 후보 장면으로 선택한다. (그림 3)은 최소 평균 제곱 오차 접근법을 이용하여 goal line을 검출하는 과정과 결과를 보여준다.

그런 다음, 골 포스트를 포함할 수 있는 후보 장면들로부터 실제 골 포스트를 포함하는 장면들을 추출해야한다. 일반적으로 축구 동영상을 분석하면, 골 포스트는 goal line 주위에서 흰색 화소들(white pixels)이 수직으로 두 개가 나타나는 것을 확인할 수 있다. 그러므로, 본 논문에서는 골 포스트를 찾기 위해서, 장면의 전체 영역이 아닌 검출된 goal line의 주변 영역들에 대해서만 흰색 화소들만을 수평축 상으로 투영한다. 그러면, 일정 이상의 간격을 갖는 서로 다른 극치들이 나타나는 것을 확인할 수 있고, 이럴 경우에 실제 골 포스트를 포함하는 장면으로 결정한다. (그림 4)는 실제 골 포스트를 포함하는 장면을 검출하기 위해 goal line의 주변 영역들에 대해서 흰색 화소들을 수평축으로 투영한 결과를 보여주고 있다. (그림 4)에서 볼 수 있듯이, 일정 이상의 간격을 갖는 서로 다른 두 개의 극치를 확인할 수 있다.



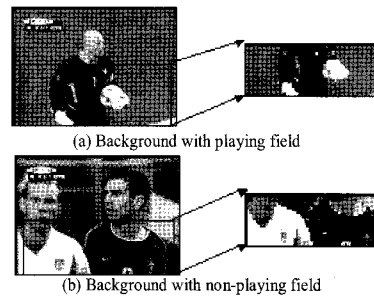
(그림 3) goal line의 검출 과정 및 결과



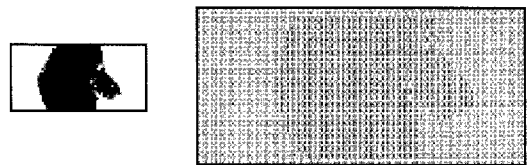
(그림 4) 골 포스트 검출 결과

2.3 비 그라운드 영역에 대한 확대 장면 검출

일반적으로 골 포스트를 포함하는 장면들이 하이라이트 장면으로 고려될 수 있지만, 모든 장면이 하이라이트라고 할 수 없다. 그래서 축구 동영상의 구조를 분석한 결과 하이라이트 장면 이후에 선수나 관중석을 확대하는 장면들과 같은 비 그라운드 영역이 나타나는 것을 확인할 수 있었다. 그러므로, 본 논문에서는 하이라이트 장면 이후의 비 그라운드 장면을 검출하기 위해서, 축구 동영상의 한 장면에서 비 그라운드 영역에 해당하는 화소들에 대해 connected component labeling 기법을 사용하여 비 그라운드 영역의 비율을 계산하다[10]. 만약 비 그라운드 영역의 비율이 일정한 문턱치 이상이면 비 그라운드를 확대한 장면으로 검출한다. (그림 5)는 선수들을 확대한 비 그라운드 장면 검출 예제를 보여주고 있다. 본 논문에서는 비 그라운드 장면을 검출하는 과정에서 전체 영역을 고려하지 않고, 화면의 절반 이하의 영역에 대해서만 고려하였다. 왜냐하면, TV 카메라가 (그림 5)-(a)에서 보는 바와 같이 선수들의 허리의 윗부분을 주로 촬영하기 때문에, 화면 상단은 그라운드 영역이 포함될 수 있기 때문이다. 또한 연산량을 줄일 수 있는



(그림 5) 축구 선수들을 확대한 장면 검출 예제



(그림 6) connected component labeling 기법 적용 영상

장점이 있다. 그런 다음, 해당 영역에 대해서 그라운드 영역과 비 그라운드 영역을 이진 영상으로 변환한 후 비 그라운드 영역들에 대해서 connected component labeling 기법을 사용하여 화면의 크기에 대해 일정한 비율 이상이면 비 그라운드 영역을 확대한 장면으로 결정한다. (그림 6)은 connected component labeling한 결과 영상을 보여주고 있다.

3. 실험 결과

본 논문에 대한 실험 평가를 위하여, 2002 한국/일본 월드컵 대회에서 방송된 축구 동영상과 유럽 챔피언스 대회에서 방송된 축구 동영상을 테스트 영상으로 사용하였다. 또한, 본 논문에서 사용된 축구 동영상들은 그림자가 지거나 햇빛이 강하게

내리찍는 것과 같은 다양한 환경을 포함하는 동영상 9가지를 사용하였다. 실험 환경은 펜티엄 4 3.0GHz 컴퓨터 시스템을 사용하였고, Visual C++ 6.0 컴파일러를 사용하여 구현하였다.

성능 평가를 위해서는 정확률(precision)과 재현율(recall)을 사용하였다. 정확률은 하이라이트 장면으로 검출한 전체 장면 수에 대한 실제 하이라이트 장면을 검출한 수를 정의한 것이고, 재현율은 전체 하이라이트 장면 수에 대한 실제 하이라이트 장면을 검출한 수를 정의한 것이다. 식(2)와 식(3)은 정확률과 재현율을 나타낸다.

$$precision = \frac{n_c}{n_c + n_f} * 100 \tag{2}$$

$$recall = \frac{n_c}{n_c + n_m} * 100 \tag{3}$$

식(2)와 식(3)에서 n_c 는 실제 검출된 하이라이트 장면 수이고, n_f 는 하이라이트로 잘 못 검출된 장면의 수이다. n_m 은 검출하지 못한 하이라이트 장면의 수이다.

<표 1>과 <표 2>는 제안한 방법에 대한 실험 결과를 보여주고 있으며, 정확률이 95.16%이고, 재현율이 85.41%로써 효과적으로 하이라이트 장면을 검출하는 것을 확인할 수 있다. 그러나, <표 1>에서 나타나는 결과 보면, goal line 주변에서 경기가 벌어지면서 슈트나 골이 아닌 장면 이후에 연속된 장면들이 선수수를 확대하는 장면이 나타나면서 하이라이트 장면으로 오검출하는 결과를 볼 수 있으며, 또한 하이라이트 장면이지만 축구장 전체를 TV 카메라가 너무 멀리(zoom out) 잡아서 goal line 검출은 성공하지만, 골 포스트 검출하지 못하여 검출하지 못하는 결과도 볼 수 있다.

<표 1> 실험 성능 평가

Soccer video clips	# of highlight scenes	extracted highlight scenes	false	miss
Video 1	19	18	1	2
Video 2	23	23	2	2
Video 3	12	13	1	0
Video 4	19	18	2	3
Video 5	11	9	0	2
Video 6	10	6	0	4
Video 7	15	13	0	2
Video 8	7	6	0	1
Video 9	19	19	2	2

<표 2> 정확률과 재현율 결과

	precision	recall
Average	95.2 %	85.4 %

4. 결 론

본 논문에서는 축구 동영상의 구조적 분석을 통한 하이라이트 장면을 검출하기 위한 방법을 제안하였다. 축구 동영상의 구조에 대한 분석을 통하여 그라운드 영역을 검출한 후 정확한 하이라이트 장면을 검출하기 위해서 골 라인과 골 포스트를 추출하고, 그 후에 연속되는 장면들이 선수나 관객들을 확대하는 장면들이면 나타나는지를 확인하였다. 그리고, 다양한 환경을 포함하는 축구 동영상에 대하여 실험한 결과 정확률이 95.2%, 재현율이 85.4%로써 효과적으로 하이라이트 장면을 검출하는 것을 확인할 수 있었다. 향후 과제로는 성능 개선을 위해 오디오 정보와 같은 다른 특징들을 추가로 고려해야 할 것이고, 또한 다른 알고리즘과의

비교 분석이 이루어져야 할 것이다.

참 고 문 헌

- [1] K. Wan, J. Lim, C. Xu, and X. Yu, "Real-time Camera Field View Tracking in Soccer Video," in *Proceedings of IEEE ICASSP*, Vol. 3, pp. 185-188, 2003.
- [2] A. Ekin and A. Tekalp, "Automatic Soccer Video Analysis and Summarization," *IEEE Trans. on Image Processing*, Vol. 12, No. 7, pp. 796-807, 2003.
- [3] K. Wan, X. Tan, X. Yu, and C. Xu, "Real-time Goal-mouth Detection in MPEG Soccer Video," in *Proceedings of ACM MM2003*, pp. 311-314, 2003.
- [4] Y. Gong, L. T. Sin, C. H. Chuan, H. Zhang, M. Sakauchi, "Automatic Parsing of TV Soccer Programs," in *Proceedings of the International Conference on Multimedia Computing and Systems*, pp. 167-174, 1995.
- [5] V. Kobla and D. S. Doermann, "Extraction of Features for Indexing MPEG Compression Video," in *Proceedings of IEEE First Workshop on Multimedia Signal Processing*, pp. 337-342, 1997.
- [6] 최송하, 이성환, "방송용 축구 경기 비디오의 자동 색인 및 분석 기술," *한국정보과학회 가을 학술발표논문집*, Vol. 25, No. 2, pp. 550-552, 1998.
- [7] 황본우, 방건, 이미숙, 이성환, "축구 경기 비디오 분석을 위한 하이라이트 자동 추출 방법," *한국정보과학회 가을 학술발표 논문집*, Vol. 24, No. 2, pp. 347-350, 1997.
- [8] S. Dagtas and M. Abdel-Mottaleb, "Extraction of TV Highlights Using Multimedia Features," *IEEE Fourth Workshop on Multimedia Signal Processing*, pp. 91-96, 2001.
- [9] 김재홍, 남종호, 하명환, 정병희, 김경수, "스포츠 중계 방송의 특성을 이용한 축구동영상 하이라이트 생성 알고리즘," *정보과학회 논문지: 소프트웨어 및 응용 제 30권* pp. 727-843, 2003.
- [10] G. Pass, R. Zabih, and J. Miller, "Comparing Image Using Color Coherence Vectors," in *Proceedings of ACM Multimedia 96*, pp. 65-73, 1996.

박 기 태

e-mail : parkkt@cse.hanyang.ac.kr

2000년 한양대학교 전자계산학과(공학사)
 2002년 한양대학교 대학원 컴퓨터공학과(공학석사)
 2002년~현재 한양대학교 대학원 컴퓨터공학과(박사과정)
 관심분야 : 내용기반 영상 검색, 패턴인식, Image/Video indexing, MPEG-7 등



문 영 식

e-mail : ysmoon@cse.hanyang.ac.kr

1980년 서울대학교 공과대학 전자공학과(공학사)
 1982년 한국과학기술원 전기 및 전자공학과(공학석사)
 1990년 University of California at Irvine Dept. of Electrical and Computer



Engineering(공학박사)

1982년~1985년 한국전자통신연구소 연구원
 1989년~1990년 Inno Vision Medical 선임연구원
 1992년~현재 한양대학교 컴퓨터공학과 교수
 관심분야 : 영상처리, 패턴인식, Image/Video Indexing, MPEG-7 등