

논문-06-11-2-07

멀티미디어 이동형 단말을 위한 축구경기 비디오의 지능적 디스플레이 방법

서기원^{a)}, 김창익^{a)†}

An Intelligent Display Scheme of Soccer Video for Multimedia Mobile Devices

Keewon Seo^{a)} and Changick Kim^{a)†}

요약

이 논문에서 우리는 작은 휴대형 단말기에서 축구영상을 화면에 보여주는 자동적이면서도 효율적인 방법을 제안하고자 한다. 정보통신 기술의 급속한 발전으로 작은 LCD 패널을 가진 멀티미디어 장치의 사용이 대단히 빠른 속도로 증가하고 있다. 이러한 작은 장치들의 증가 추세에도 불구하고 일반 TV나 HDTV 용으로 녹화되는 대부분의 동영상들은 작은 이동형 기기 사용자들이 화면 내의 상황을 인식하는데 있어서 많은 불편한 경험을 갖게 한다. 예를 들어, 원거리 샷 카메라 기법으로 찍힌 축구 경기 동영상의 경우, 운동장 내의 공과 선수들의 모습이 매우 작아서 작은 화면으로 알아보기가 힘든 경우가 발생한다. 따라서 소형 디스플레이 시청자들의 원활한 이해를 위해 지능형 디스플레이 기술을 제안하고자 한다. 이를 위한 핵심기술의 하나가 관심영역을 자동으로 결정하는 일이다. 여기서 관심영역이란 시청자들이 화면 내에서 다른 부분에 비해 더욱 관심을 갖게 되는 부분을 말한다. 이 논문에서, 우리는 소형 단말기를 위한 축구 경기 비디오의 지능적 디스플레이에 초점을 맞춘다. 화면 내에서 시각적으로 현저한 부분의 검출에 관심을 갖는 방법 대신, 축구 경기 비디오 고유의 특징을 이용하는 도메인 한정적인 접근법을 이용한다. 제안된 알고리즘은 크게 세 단계 - 그라운드 색 학습과 샷 분류, 관심영역 결정으로 구성된다. 실험 결과를 통해 제안된 알고리즘이 이동형 단말기 상에서 지능형 디스플레이를 위한 좋은 해결책을 보이고자 한다.

Abstract

A fully automatic and computationally efficient method is proposed for intelligent display of soccer video on small multimedia mobile devices. The rapid progress of the multimedia signal processing has contributed to the extensive use of multimedia devices with a small LCD panel. With these emerging small mobile devices, the video sequences captured for standard- or HDTV broadcasting may give the small-display-viewers uncomfortable experiences in understanding what is happening in a scene. For instance, in a soccer video sequence taken by a long-shot camera technique, the tiny objects (e.g., soccer ball and players) may not be clearly viewed on the small LCD panel. Thus, an intelligent display technique is needed for small-display-viewers. To this end, one of the key technologies is to determine region of interest (ROI), which is a part of the scene that viewers pay more attention to than other regions. In this paper, the focus is on soccer video display for mobile devices. Instead of taking visual saliency into account, we take domain-specific approach to exploit the characteristics of the soccer video. The proposed scheme includes three modules; ground color learning, shot classification, and ROI determination. The experimental results show the propose scheme is capable of intelligent video display on mobile devices.

Key Words : 축구 비디오 분석, 관심영역, 지능형 디스플레이, 이동형 기기

a) 한국정보통신대학교 (ICU) 시각정보처리연구실

Information and Communications University (ICU) Visual Information Processing Lab.

† 교신저자 : 김창익(ckim@icu.ac.kr)

※ 본 연구는 정보통신부 및 정보통신연구원(ETRI)의 디지털미디어연구소 지원사업의 연구결과로 수행되었음.

I. 서론

최근 멀티미디어 신호처리 및 전송기술의 급속한 발전과, DVB H (Digital Video Broadcasting Handheld) 및 DMB (Digital Multimedia Broadcasting) 등과 같은 새로운 형태의 이동형 TV 서비스의 등장으로 인해 작은 LCD 패널을 통해 비디오를 시청하는 경우가 잦아지고 있다. 하지만 경제적인 이유 등으로 인하여 대부분의 서비스들은 기존의 화면 크기를 단지 축소하여 이동형 서비스에 제공하고 있는 실정이다. Knoche 등에 의해 실시된 실험^[1]에 의하면 해상도가 큰 영상의 직접적인 축소는 시각적으로 상세한 부분에 대한 손실을 가져옴을 알 수 있었다. [1]에서 제시된 실험 결과에 따르면, 이러한 손실은 축구 비디오와 같은 필드스포츠인 경우 더욱 크고 뮤직비디오, 뉴스, 애니메이션의 순으로 조금씩 줄어들었다. 특정 내용을 담은 비디오의 경우, 이러한 문제를 해결하기 위해서는 소형 이동 단말기 시청자들이 관심 영역 (Region of interest 또는 ROI)만이 확대된 화면을 시청할 수 있도록 해주는 지능형 디스플레이 기술의 개발이 필요하다. 여기서 관심영역이란 사용자가 화면상에서 가장 관심을 갖게 되는 영역 또는 사용자가 다른 영역에 비해 더욱 주의를 기울이는 영역으로 정의될 수 있다. 관심 영역은 상황인지 콘텐츠 적응, 트랜스 코딩, 지능형 정보 관리 등에 활용 된다^[2]. 또한 관심영역의 지정은 비디오 장면(scene)을 의미적으로 분석하는 첫 단계가 될 수 있으므로 영상 분석을 위해서도 매우 중요한 기술이라 할 수 있다.

그동안 관심영역을 결정하기 위한 여러 방법들이 연구되어 왔다. Itti^[3] 등은 국부적으로 존재하는 현저함(saliency)을 측정하기 위해 여러 개의 공간적 시각적 특징들을 하나의 특징점 맵(saliency map)으로 나타낼 수 있는 시각적 주의 모델을 제안하였으나 동영상으로의 확장은 제대로 고려하지 않았다. Cheng^[2] 등은 그들의 논문에서 명암, 색상 및 모션 정보에 기초하여 비디오에서 관심영역을 결정하는 방법을 연구하였다. 그 논문에서, 중요한 물체는 높은 색상 또는 명암의 대비를 갖는다는 가정 하에 비디오의 각 영상 프레임에서 가장 눈에 띄는

특징점을 가리키는 방법을 제안하였다. 하지만 중요한 물체의 물체가 어둡거나 낮은 대비를 가질 수 있으므로 이 가정은 항상 옳은 것은 아니다. 비디오 스킴밍(Skimming)과 요약을 위해 Ma 등이 제안한 또 다른 비디오 분석 방법에서는 모션, 스피치, 카메라 동작 및 비디오 편집 정보 등을 이용하여 사용자의 관심 모델이 제시되었다^[4]. 이상과 같이 비디오 화면내의 현저한 특징들을 이용하여 관심영역들을 추출하는 방법들은 효과적인 해결책을 제시하고는 있으나, 축구 경기 비디오와 같은 필드 스포츠 비디오의 경우에는 일단 관심영역의 추출이 필요한 샷과 필요하지 않은 샷들이 혼재하며, 전자의 경우 화면안에 현저한 특징을 갖는 작은 물체들이 동시에 여러 개가 존재 할 수 있다. 이 경우 복수개의 관심영역이 존재할 수 있으므로 화면내의 일부만을 뽑아내어 확대해 보여주고자 하는 이동형 단말을 위한 지능형 디스플레이에는 적합하지가 않다. 따라서, 축구 비디오를 다루는 경우에는 특징점에 기반한 접근 방법^{[2][3][4]} 대신 도메인 한정적인(domain-specific) 접근 방법을 이용하는 것이 바람직하다.

본 논문에서는 이동형 단말기를 위한 축구 경기 비디오의 디스플레이에 초점을 맞추고자 한다. 축구경기 방송의 경우, 많은 사람들이 실외에서 이동 중에도 시청하기를 원하는 콘텐츠의 대표적인 예이며 그림 1에서 보듯이 원거리 샷 영상의 경우 지나치게 축소된 물체들로 인해 쾌적한 시청이 쉽지가 않다. 축구 경기의 화면은 크게 원거리 샷과 비원거리 샷으로 구성된다. 원거리 샷 프레임은 원거리

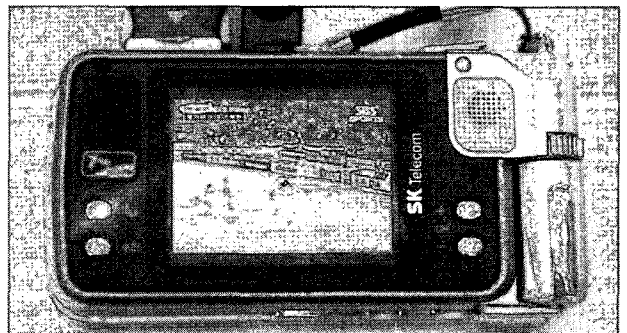


그림 1. DMB 폰에서 방송중인 축구 경기.

Fig. 1. A DMB phone displaying a soccer video.



(a) (b) (c)

그림 2. (a) 원거리 샷 영상. 관심영역을 추출하여 확대 재생할 필요가 있음. (b) 중거리 샷 영상. (c) 근거리 샷 영상.

Fig 2. Three kinds of frame in soccer video. (a) Long-shot frame, which requires a magnified view of ROI. (b) Mid-shot frame. (c) Close-up shot frame.

서 촬영되었으므로 작은 화면으로 시청할 때 공과 선수의 크기가 작아 알아보기 어렵다. (그림 2 (a) 참조). 이 경우 관심영역을 추출하여 확대 재생하여 시청하면 사물을 인식하기가 보다 쉽다. 미디엄 샷이나 클로즈 업 샷 같은 비원거리 샷의 경우 원 영상을 그대로 보여주어도 비디오 시청에는 문제가 없다. (그림 2 (b), (c)). 따라서 작은 화면에서 보다 시청하기 쉬운 디스플레이를 위해 프레임은 두 가지 종류의 프레임으로 구분되어야 하며 원거리 샷의 경우에는 전체 화면의 일부를 관심영역으로 지정하여 확대 재생한다. 본 논문이 제안하는 축구 경기 비디오의 디스플레이 방법은 크게 다음 세 단계로 나눌 수 있다: 1) 그라운드 색상 학습 2) 샷 분류 3) 관심영역 결정. 축구 비디오에서는 그라운드 색상이 샷 구분이나 그라운드 내의 사물과 외부의 사물을 구분하는데 있어서 중요한 역할을 한다. 하지만 경기 마다 그라운드 색상은 조금씩 다르기 때문에, 시스템 내에서 그라운드 색상에 대한 고정된 파라미터를 사용하기가 곤란하다. 따라서, 축구 비디오 분석의 첫 단계에서는 주어진 축구경기 비디오에서 그라운드 색상을 학습하는 과정이 필요하다. 그라운드 색상의 범위가 결정되고 나면 이를 바탕으로 샷 분류와 관심영역 결정 단계를 거치게 된다.

이 논문은 다음과 같이 구성되어 있다. 2장에서는 본 연구의 타당성을 검증하기 위하여 작은 화면 상에서 축구 경기 비디오를 위한 지능형 디스플레이의 필요성을 사전 조사한 결과를 소개하며 3장에서는 본 논문에서 제안하는 방법에 대해 기술한다. 4장에서는 실험 결과를 제시하며 이에

대한 결론을 5장에 기술한다.

II. 연구의 타당성 검증: 지능형 디스플레이의 필요성

[1]에서 보고된 조사결과에 따르면, 이동형 소형 TV 시청을 위한 해상도에는 최저 한계가 존재한다. 만일 TV 콘텐츠가 이동형 단말에서의 시청을 위해 특별히 제작되거나 편집되지 않은 경우, 해상도가 168x126 보다 낮게 축소된 경우 시청자의 시청 만족도에 있어서 급격한 저하가 나타나며, 콘텐츠가 축구와 같은 야외 스포츠인 경우 시청 만족도는 더욱 악화됨을 알 수 있었다. 본 논문에서 우리는 축구 경기 비디오를 위한 지능형 디스플레이 방법을 소개하고자 하며, 이러한 연구의 필요성을 검증하기 위하여 [1]에서 검증된 결과를 바탕으로 하여 축구 경기 비디오에 국한된 더욱 자세한 조사를 실시하였다. 즉, [1]에서 사용된 디스플레이 크기보다 더욱 많은 크기들을 가지고 디스플레이 크기와 시청자 만족도에 관한 관련성을 조사하였으며 또한 관심영역의 크기에 관한 시청자 평가도 실시하였다. 시청자 평가는 65536 색상과 최대 해상도 480x640를 갖는 4-inch TFT LCD 스크린 상에서 이루어졌으며 비디오는 TCPMP^[14]를 이용하여 복호화하였다. 시청자 평가를 위해 스물두명의 대학원생들이 참여하였으며 유로2004 게임 중 네덜란드와 체코간의 경기에서 뽑아낸 30초 분량의 비디오가 사용되었다.

1. 해상도와 시청자 만족도와의 관계

실험 참가자들은 MPEG-1으로 부호화 및 복호화된, 원본 포함 여섯 개의 서로 다른 해상도(640x480, 480x360, 320x240, 240x180, 160x120, 120x90)를 갖는 축구 경기 비디오를 보면서 시청 만족도를 0 (최소)에서 1(최고)까지의 범위내에서 답하도록 요구되었다. 부호화시 사용된 양자화 파라미터의 크기는 부호화로 인한 화질의 열화를 느끼지 않을 만큼 충분히 작은 것으로 사용되었다. 비디오 클립은 큰 것에서 작은 것의 순서로 제시되었으며 그림 3에서 보듯이 시청자의 만족도는 해상도가 480x360보다 낮아지면서 급격히 떨어지기 시작하였으며 [1]에서 보고된 바와 같이 해상도가 160x120 이하가 될 때, 만족도는 0.5 이하로 내려감을 알 수 있었다. 특히, 화면 내 물체들의 구체적인 특징 (detail)을 이해하는데 있어서는 해상도가 320x240 이하로 내려감에 따라 어려움이 발생하기 시작하였다.

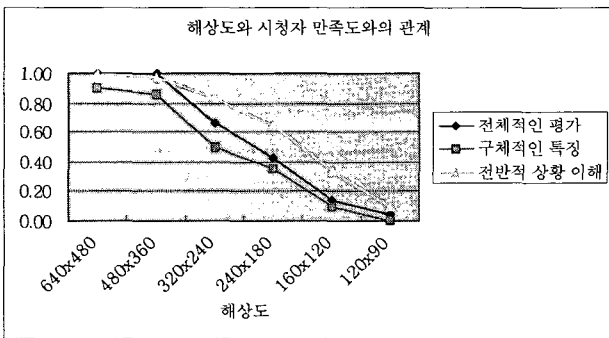


그림 3. 해상도와 시청자 만족도와의 관계
Fig. 3. User satisfaction of the soccer video as resolution changes

2. 지능형 디스플레이를 위한 최적의 관심영역 크기 평가

이전 절에서 요약되었던, 해상도 (또는 영상의 크기)가 낮아짐에 따라 시청자의 만족도가 저하되는 현상을, 화면 내에서 관심영역을 추출하여 확대해서 보여줌으로써 해결할 수 있는지를 확인하기 위해 다양한 크기의 관심영역만을 보여 주면서 시청자들의 만족도를 알아보는 실험을 실시하였다. 예를 들어, 화면 내에서 추출해내는 관심영역의

크기가 작을 경우, 확대되는 비율이 커짐으로 인해 관심영역에 대한 이해가 쉬워지는 반면 화면 전체를 통해 알 수 있는 전반적인 이해는 감소할 수 있을 것이다. 우리는 이러한 교환관계 (trade-off)를 갖는 시청 상황 속에서 관심영역만을 제시할 때의 시청자 만족도가 변화하는 양상을 그래프로 나타내었다. 먼저 640x480 크기를 갖는 원본 영상에 대해, 640x480, 480x360, 320x240 and 240x160과 같이 네 종류의 서로 다른 크기를 갖는 관심영역을 수동으로 추출하였으며 이들을 각각 원본 영상의 크기로 확대하여 보여주는 경우에, 시청자의 시청만족도가 변화하는 모습을 측정하였다. 앞에서 제시된 네가지의 관심영역 크기는 영상의 폭 (또는 높이)의 비율로 따지면 원본 영상에 대해 각각 1:1, 1:0.75, 1:0.5 및 1:0.375와 같다. 이렇게 네가지의 서로 다른 크기를 갖는 관심영역들은 320x240으로 크기가 조절되어 피실험자들에게 제시된다. 참석자들에게는 각 관심영역 크기별로 0(최저)에서 1(최고)까지의 만족도로 답하기를 요구하였으며 그림 4에서 보듯이 1:0.75의 경우, 전반적으로 가장 만족도가 높게 나타남을 알 수 있었다. 즉, 원본 영상에 대해 폭 (또는 높이)의 비율이 4분의 3 가량 되는 관심영역에 대해 시청자들은 화면의 전반적인 이해를 놓치는 것을 최소화하면서 관심영역을 더욱 확대하여 볼 수 있음을 만족하는 것으로 나타났다. 원본과 관심영역의 폭의 비율이 1:0.5 이하로 내려감에 따라 예상되는 바와 같이 세부 특징에 대한 이해는 증가하는 반면 화면의 전반적인 이해는 급격히 떨어짐을 알 수 있었다. 이러한 실험결과를 토대로, 우리는 다음 장들에서 제시되는 제안된 알고리

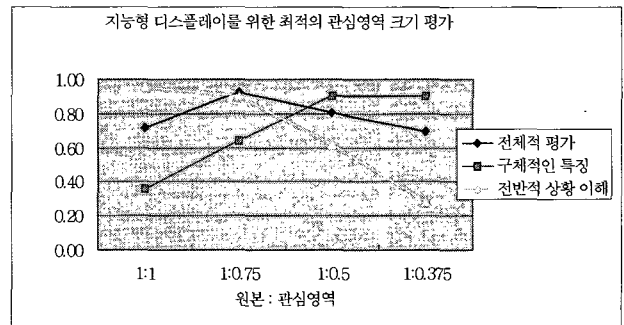


그림 4. 지능형 디스플레이를 위한 최적의 관심 영역 크기 평가
Fig. 4. Evaluation of ROI size for intelligent display.

즘의 구현과 실험에 있어서 1:0.75 에 가까운 비율을 적용하기로 하였다.

III. 제안 알고리즘

우선, 비디오는 다음과 같이 샷의 연속 (sequence of shots)으로 정의할 수 있다. 여기서 샷이란 한번의 카메라 녹화 과정으로 취득되는 영상 열 (image sequence)을 뜻한다.

$$Video = \langle Shot_0, \dots, Shot_{K-1} \rangle, \quad (K \text{ 는 비디오의 프레임 갯수}) \quad (1)$$

샷은 다시 일련의 프레임으로 나타낼 수 있다.

$$Shot_k = \langle f_k^0, \dots, f_k^{N-1} \rangle \quad (2)$$

여기서 f_k^i 는 $Shot_k$ 의 i 번째 프레임을 말하며, N 은 샷의 프레임 갯수이다. 또한 앞으로 비디오 전체에서 i 번째 프레임은 f^i 으로 표시하겠다. k 번째 샷의 i 번째 프레임의 관심

영역은 $ROI(f_k^i)$ 로 표시할 수 있다. 샷의 종류에 따라 관심 영역은 영상 프레임의 전체가 될 수도 있고 일부가 될 수도 있다. 즉

$$\begin{cases} ROI(f_k^i) \subset f_k^i, & \text{if } Class(Shot_k) \text{ is a long-shot} \\ ROI(f_k^i) = f_k^i, & \text{if } Class(Shot_k) \text{ is not a long-shot} \end{cases} \quad (3)$$

이는 영상 프레임이 원거리 샷에 속하지 않을 때의 ROI는 원 영상과 동일하게 정하며, 원거리 샷에 속할 때의 관심영역은 영상 프레임 내의 한 부분으로 정하는 것을 뜻한다. 소기의 목적을 달성하기 위한 과정을 그림 5에 나타내었다.

1. 그라운드 컬러 학습

축구 비디오 분석에서 그라운드의 색상은 매우 중요한 역할을 한다. 이는 특히 샷의 종류를 판별하는 데에 중요하다. 우리가 제안하는 알고리즘에서는 그라운드 색을 띄는 화소의 양이 샷의 종류를 판별하는 중요한 단서가 된다. 비디오 시퀀스에 따라 그라운드의 색상은 조금씩 다르기 때

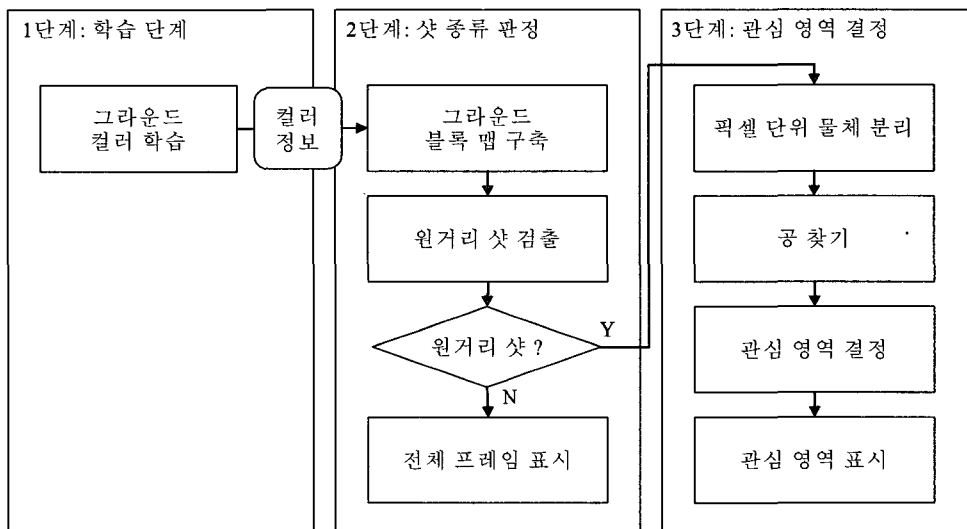


그림 5. 전체 과정

Fig. 5. Block diagram of the proposed algorithm

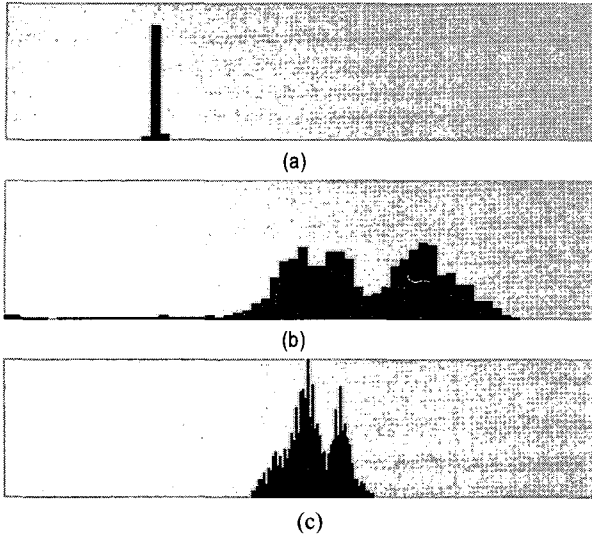


그림 6. HSV 컬러 공간에서의 히스토그램. (a) 색조 (b) 채도 (c) 명암
Fig. 6. Histograms for HSV color channels: (a) Hue. (b) Saturation. (c) Value.

문에, 시스템의 첫 단계에서 비디오 시퀀스의 그라운드 색상을 학습하는 일이 필요하다. 그림 6은 일반적인 축구 비디오의 원거리 샷에서 취한 HSV 히스토그램이다. 그림에서 알 수 있듯이 색조(Hue) 값의 분포는 노랑 녹색의 영역에 특히 집중되어 있다. 반면에 채도(Saturation)와 명암(Value) 값들의 분포는 비디오에 따라 각기 다르다. 이에 따라 우리는 다음과 같은 그라운드 색 학습 방법을 제안한다.

우선 우리는 색조 채널 및 채도 채널에 64개의 빈(bin), 명암에는 256개의 빈을 배정하였다. 이 빈의 개수를 각각 NH , NS , NV 라고 하겠다. 따라서 i 번째 프레임에 대한 HSV 히스토그램은 다음과 같이 정의된다.

$$\begin{aligned} Hue_i[k] &: 0 \leq k < NH \\ Saturation_i[k] &: 0 \leq k < NS \\ Value_i[k] &: 0 \leq k < NV \end{aligned} \quad (4)$$

또한 HSV 히스토그램을 비디오의 첫 30 프레임 동안 축적하여 사용하는데, 이를 각각 HUE , $SATURATION$,

$VALUE$ 라 하고 다음과 같이 정의한다.

$$\begin{aligned} HUE[k] &= \sum_{i=0}^{29} Hue_i[k] \\ SATURATION[k] &= \sum_{i=0}^{29} Saturation_i[k] \\ VALUE[k] &= \sum_{i=0}^{29} Value_i[k] \end{aligned} \quad (5)$$

위의 정의를 이용하여 우리는 다음의 두가지 값을 정의한다.

$$ValuePeakIndex = i, \quad (6)$$

where $VALUE[i] \geq VALUE[p]$ for all $0 \leq p < NV$, and

$$SaturationMean = \frac{\sum_{i=0}^{NS-1} i \cdot SATURATION[i]}{\sum_{i=0}^{NS-1} SATURATION[i]} \quad (7)$$

$ValuePeakIndex$ 는 Value 히스토그램의 피크 값이 있는 빈의 인덱스를 뜻한다. 우리는 다음 조건이 만족되는 경우 그라운드 색의 학습이 성공적으로 이루어졌다고 판정한다.

$$0.7 \times frame.width \times frame.height \times 30 < \sum_{i=9}^{21} HUE[i] \quad (8)$$

이는 $9 \leq i \leq 21$ 의 범위에서 $HUE[i]$ 를 합한 결과가 전체 프레임에 해당하는 픽셀 수의 70%를 넘으면 학습이 제대로 이루어진 것으로 보는 것이다. 색조 값 분포 내 그라운드 색상의 범위는 여러 축구 비디오를 관찰한 결과 9~21로 정하였다. 만약 이 조건이 만족하지 않으면 학습에 실패한 것으로 간주하여 다음 30 프레임 동안 같은 작업을 다시 수행한다.

많은 축구 경기를 관찰한 결과 우리는 그라운드의 색에는 대체적으로 $g > r > b$ 라는 관계가 있음을 발견할 수 있었다. 여기에 공이나 회색의 픽셀이 그라운드로 잘못 판정

되는 것을 막기 위해 몇 가지 조건을 추가하였다. 어떤 색의 RGB 값을 HSV 값으로 바꾸어 채도 S 와 명암 V 를 구하는 식은 다음과 같다^[5].

$$S = \frac{Max(R, G, B) - Min(R, G, B)}{Max(R, G, B)} \quad (0.0 \leq S \leq 1.0) \quad (9)$$

$$V = Max(R, G, B) \quad (0.0 \leq V \leq 1.0)$$

여기서 R, G, B 값은 0에서 1사이의 값으로 정규화 된 상태이다.

식 (9)와 $g > r > b$ 라는 그라운드 영역의 상관 관계를 이용할 때, $G \approx Max(R, G, B)$ 이고 $NV=256$ 이므로 V 는

$$g = G \cdot NV \text{로 표시할 수 있다. 또한 } S = \frac{g-b}{g} \cdot NS = \frac{G-B}{G} \cdot NS$$

인데 이는 $G \approx Max(R, G, B)$ 이고 $B \approx Min(R, G, B)$ 이기 때문이다. 정리하면, rgb 컬러 공간에서 각각의 픽셀이 그라운드에 속하는지 아닌지를 판단하는 식은 아래와 같이 나타낼 수 있다.

$$Ground(x, y) = \begin{cases} 1 & \text{if } \begin{cases} g > 0.95 \cdot r \text{ and} \\ r > 0.95 \cdot b \text{ and} \\ g < ValuePeakIndex + \theta_1 \text{ and} \\ \frac{g-b}{g} \cdot NS > SaturationMean - \theta_2 \end{cases} \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

여기서 r, g, b 는 (x, y) 에서의 RGB 값으로 0 255의 값을 갖는다. θ_1 과 θ_2 은 각각 $NV/6$ 과 $NS/4$ 로 설정하였다. 픽셀의 Value 값이 ValuePeakIndex 보다 너무 높거나 Saturation 이 SaturationMean 보다 너무 낮은 경우 그 픽셀은 그라운드에 속하지 않는 것으로 판정한다. 이 방법은 Ekin^[6]이나 Wan^[7]이 제안한 방법에 비해 빠르며 그림 7 처럼 운동장의 일부에 그림자가 드리운 경우에도 제대로 동작한다.

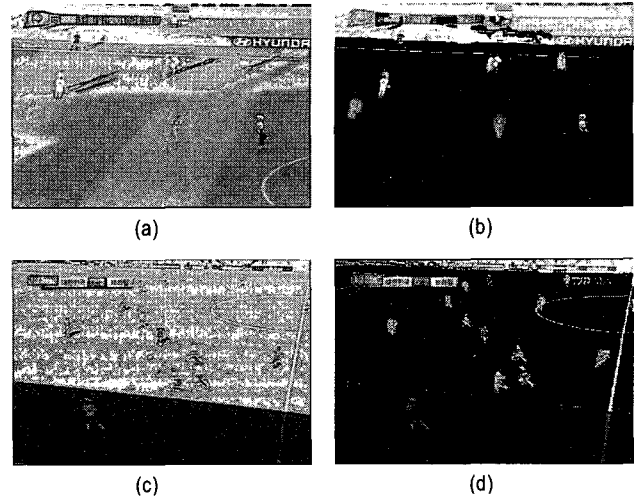


그림 7. (a), (c) 일광이 일정하지 않은 경기. (b), (d) 제안한 방법에 의한 그라운드 색 구분 (검정색으로 표시된 부분)
Fig. 7. (a), (c) Shadowed ground. (b), (d) Detected ground by using the proposed algorithm.

2. 그라운드 블록 맵 구축

이전 절에서 우리는 각 픽셀이 그라운드에 포함되는지의 여부를 결정할 수 있었다. 빠른 샷 경계 검출과 샷 구분을 위해, 전체 프레임을 16×16 블록으로 나누어 그라운드 블록 맵 $GB(i, j)$ 를 만들었다. $GB(i, j)$ 을 정의하기 위해 우선 한 블록 내 픽셀의 집합인 B_{ij} 를 다음과 같이 정의한다.

$$B_{ij} = \{(x, y) | i \times 16 \leq x < (i+1) \times 16, j \times 16 \leq y < (j+1) \times 16\} \quad (11)$$

이를 이용해 우리는 그라운드 블록 맵 $GB(i, j)$ 를 다음과 같이 정의한다.

$$GB(i, j) = \begin{cases} 1(Ground) & \text{if } \frac{\sum_{(x, y) \in B_{ij}} Ground(x, y)}{16 \times 16} \geq 0.2 \\ 0(Object) & \text{otherwise} \end{cases} \quad (12)$$

이는 한 블록 내 픽셀의 20% 이상이 그라운드 컬러를 띠면 그 블록을 그라운드로 본다는 것을 의미한다. 우리는

이후 과정에서 프레임 f_i 의 그라운드 블록 맵을 $GB_i(x, y)$ 로 표시할 것이다.

3. 샷 경계 검출

샷 경계 검출은 비디오 분석의 기본이 되는 과정이다. 샷 경계에서만 샷을 분류함으로써 계산량을 줄이고, 일시적인 화면 상태 변화로 인한 판정 오류를 줄임으로써 샷 종류 판정의 정확성을 높일 수 있다. 샷 경계를 검출하는 방법으로는 여러 가지 방법이 제시되고 있는데^[8], 우리는 그라운드 블록 맵의 시간적 불연속성(Temporal Block Difference, TBD)을 체크하여 효율적이고 빠르게 샷 경계를 검출하였다.

$$TBD_i = \sum_x \sum_y \{GB_{i-3}(x, y) \otimes GB_i(x, y)\} \quad (13)$$

여기서 \otimes 는 XOR연산을 뜻한다. 우리는 $TBD_{i-1} < \theta_{ShotChange}$ 이고 $TBD_i \geq \theta_{ShotChange}$ 일 때 i 를 샷의 경계로 지정하였다 ($\theta_{ShotChange} = 30$ 사용). 빠르게 변화하는 장면의 경우 $TBD_i \geq \theta_{ShotChange}$ 인 상황이 많이 발생하기 때문에 $TBD_{i-1} < \theta_{ShotChange}$ 를 함께 두어 경계 검출 오류를 줄였다. 또한 우리는 현재 그라운드 블록 맵을 세 프레임 이전의 것과 비교하였는데, 이는 페이드 인/아웃이나 시각적 효과 (visual effect)

로 인해 샷이 점진적으로 변하는 경우에 대응하기 위해서이다.

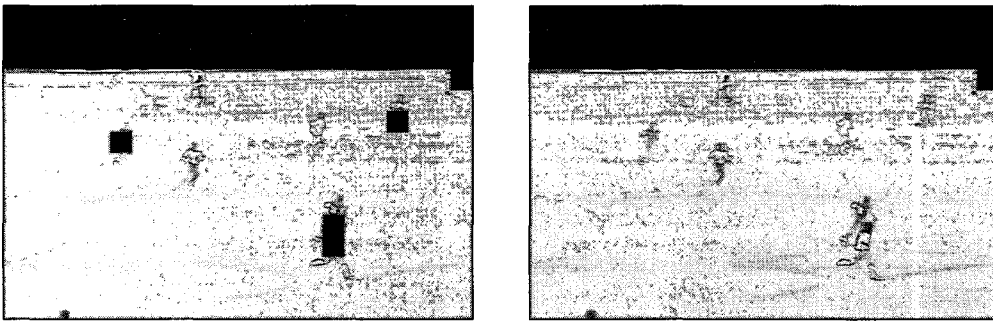
4. 샷 종류 판정

샷 경계가 검출되면 그 지점에서 샷 종류 판정이 수행되어야 한다. 원거리 샷과 그렇지 않은 샷을 구분해내기 위해 우리는 그라운드 블록 맵을 이용하였다. 첫번째 단계로 그림 8처럼 선수들로 인해 그라운드 내에 생긴 홀을 채우는 작업이 필요하다. 보다 자세히 설명하면 다음과 같다.

$$GB(i, j) = 1 \text{ if } \begin{cases} [GB(i-1, j)=1 \text{ and } \{GB(i+1, j)=1 \text{ or } GB(i+2, j)=1\}] \text{ or} \\ \{[GB(i-2, j)=1 \text{ or } GB(i-1, j)=1] \text{ and } GB(i+1, j)=1\} \end{cases} \quad (14)$$

이는 세로 방향으로 한 개나 두 블록 높이의 구멍이 있는 경우 이를 채우는 것을 의미한다.

홀을 채우고 나면 그라운드 블록 맵의 각 컬럼 k 에 대해 우리는 가장 긴 녹색 세그먼트 LGS_k (Longest Green Segment)를 찾아낸다. LGS_k 는 그림 9에 잘 나타나있다. 우리는 각 영상 프레임 내에서 황금 영역 내 LGS의 길이를 측정하여 샷의 종류를 알아낼 수 있다^[6]. 여기서 황금영역이란 영상을 가로 및 세로 방향으로 3:5:3으로 나눌 때 5의 비율에 해당하는 중앙 영역을 말하며 본 논문에서는 가로 방향에 대해서만 황금 영역을 구분하기로 하였다. 그림 9에



(a) (b)
 그림 8. (a) 그라운드 블록 맵 (b) 홀을 채운 후의 모습
 Fig. 8. (a) Ground block map (GBM). (b) GBM with holes filled.

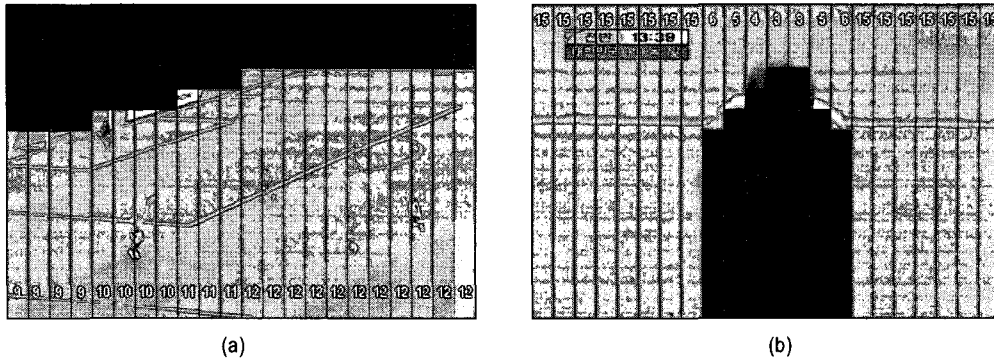


그림 9. LGS를 합친 결과. LGS의 길이를 각 컬럼에 표시하였다. (a) 원거리 샷 (b) 비 원거리 샷
 Fig. 9. LGS for each column. The values are shown in white on each column. (a) Long-shot. (b) Non-long-shot.

서 볼 수 있듯이 θ_L 보다 작은 LGS가 하나라도 존재하면 그 샷을 원거리 샷이 아닌 것으로 판정한다.

$$Class(f) = \begin{cases} \text{Non-long-shot, if } |LGS_k| < \theta_L \text{ for } GSLeft \leq k \leq GSRight \\ \text{Long-shot, otherwise} \end{cases} \quad (15)$$

여기서 $\theta_L = BlocksInColumn/3$, $GSLeft = BlocksInRow \times 3/11$, $GSRight = BlocksInRow \times 8/11$ 로 값을 정하였다.

5. 픽셀 단위 영상 분할

샷 종류 판정이 끝나면 각각의 프레임은 식 (15) 에 의해 원거리 샷과 그렇지 않은 샷으로 구분된다. 원거리 샷이 아닌 경우, 전체 영상 프레임보다 작은 관심영역을 굳이 설정할 필요가 없다. 사용자에게 프레임 전체를 보여줘도 사용

자에게 충분하기 때문이다. 하지만 원거리 샷에 대해서는 관심영역을 지정하여 확대 재생할 필요가 있다. 우리는 시청자들이 공의 주변 상황에 대해 보다 관심을 갖는다는 가정을 하였으며, 이어 따라 관심영역을 지정하는 데에 있어 공의 현재 위치 및 이전 영상에서 설정된 관심영역의 위치 정보를 가장 중요하게 취급할 것이다. 공을 찾기 위해서 우리는 식 (10)의 $Ground(x, y)$ 를 이용하여 픽셀 단위 이진화 영상을 구축하였다.

$$BinaryMap(x, y) = \begin{cases} 0 & \text{if } Ground(x, y) = 1 \\ 1 & \text{otherwise} \end{cases} \quad (16)$$

이를 통해 우리는 그림 10-(a)와 같은 영상을 얻을 수 있다. 이 영상에 대해 우리는 8방향의 영역 채색 기법 (Connected Component Labeling)을 수행하여 그라운드

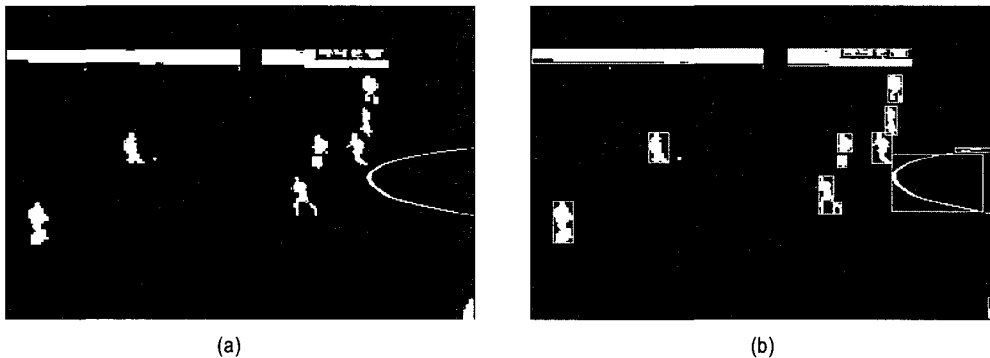


그림 10. (a) 이진화 영상 (b) 영역 채색 기법을 통해 검출한 물체의 최소 경계 사각형을 표시하였음.
 Fig. 10. (a) Binarized image. (b) Objects bounded by their MBR.

내 물체들의 목록을 얻을 수 있다. 그림 10-(b)에 얻어낸 물체들의 최소 영역 사각형(Minimum Bounding Rectangle, MBR)을 노란색으로 표시하였다. 각 물체들의 가로/세로 비, 평균 밝기, 구성 픽셀 수 등의 속성에 따라 우리는 각각의 물체를 공, 선수, 공과 비슷한 작은 물체, 기타 물체로 나누었다.

6. 공 찾기

물체들의 목록을 얻어냈지만, 한 프레임의 정보만을 이용하여 어느 물체가 진짜 공인지 판단하는 것은 쉽지 않은 일이다. 왜냐하면 운동장 내에는 공으로 오인할 수 있는 선수들의 양말, 신발이나 라인들이 있기 때문이다. 게다가 공은 경기 중에 선수들에 의해 가려지거나 빠른 속도로 인해 모양이 왜곡되기도 한다. 몇몇 프레임에서는 공이 프레임 내에 존재하지 않기도 한다. 그 동안 공을 추적하는 여러 가지 연구들이 있었다^{[9][10][11]}. 하지만 [9]에서 제시한 방법은 공의 처음 위치를 수동으로 지정해야 하는 단점이 있었다. [10]에서는 고정된 카메라로 촬영한 축구 영상을 이용하였기 때문에 우리의 경우에는 맞지 않다. [11]은 공의 궤적을 얻기 위해 이후 프레임의 정보를 필요로 하므로 실시간으로 동작할 수 없다는 단점이 있다. 우리는 간단하고 이전 값에만 의존하여 방송 영상을 실시간으로 처리할 수 있는 방법을 제안한다.

우리는 “가장 오랫동안 그라운드 내에서 추적된 공 후보가 공일 확률이 가장 많다”라는 가정을 하였다. 이는 일시적인 노이즈의 영향을 줄일 수 있다. 우선 공 후보 목록을 유지하면서 새로 찾아낸 물체 중 공으로 분류된 물체를 공 후보 목록에 추가한다. 공 후보는 다음 프레임에서 위치적으로도 가깝고 속성이 가까운 물체로 계승된다. 계승된 물체가 공인 경우 연차를 증가시키지만, 반대의 경우 그 물체의 연차를 깬다. 그 결과 가장 연차가 높은 물체가 공일 확률이 가장 높다. 연차가 0보다 작은 물체는 더 이상 추적하지 않는다. 이 방법을 통해 우리는 높은 확률로 실제 공을 찾아낼 수 있다. 실제 축구 경기 비디오 내에 공이 존재하지 않거나, 공의 모양이 제대로 분리되지 않는 등의 이유로 공을 찾지 못하는 경우, 우리는 가장 최근에 찾은 공의

위치를 현재 공의 위치로 지정하였다. 또한 세 프레임 연속으로 공을 찾지 못하는 경우 우리는 화면의 중앙을 임의로 공의 위치로 지정하였다.

7. 관심영역 창의 지정

관심영역 창을 배치하는 가장 간단한 방법은 관심영역 창의 중심을 공의 위치에 놓는 것이다. 하지만 여러 프레임에 걸쳐 관심영역의 중심이 공을 따라가는 경우 화면이 흔들리거나 사용자가 보기에 너무 빠르게 움직일 가능성이 있다. 다시 말해 작은 디스플레이를 사용하는 사용자들이 보기에 편하게 하려면 관심영역 창이 부드럽게 움직여야 한다. 또한 공이 빠르게 움직이는 경우를 감안하여 디스플레이 창을 움직이는 방법에는 가속도의 개념이 포함되어야 한다. 이에 대한 해결 방안을 이 절에 제안하였다.

어떤 프레임이 샷 종류 결정 단계에서 원거리 샷으로 판정된 경우 관심영역 창의 중심은 다음과 같이 그대로 공의 위치로 초기화 된다.

$$window_i = lob_i \quad (17)$$

$$dispi = 0 \quad (18)$$

여기서 $window_i$ 는 프레임 f^i 에서 디스플레이 창의 중심을 뜻하며, lob_i 는 f^i 에서의 공의 위치(Location of Ball)를 뜻한다. 또한 $dispi$ 는 f^i 에서 관심영역 창이 실제로 움직여야 할 거리(displacement)를 나타낸다. 계속해서 원거리 샷이 이어지는 경우에는 우선 공의 위치와 이전 디스플레이 창 사이의 거리를 계산하여 이에 따라 다르게 접근한다.

$$diff = lob_i - window_{i-1} \quad (19)$$

간단한 설명을 위해 가로 방향과 $diff \geq 0$ 인 상황만 고

려해도 일반성을 잃지 않는다. 우리는 다음 의 세가지로 경우를 나누었다. (*frame.width* 는 프레임의 너비를 뜻 한다.)

1) $0 \leq diff \leq disp_i$

$$window_i = window_{i-1} \quad (20)$$

$$disp_{i+1} = \alpha \cdot disp_i \quad (0 < \alpha < 1) \quad (21)$$

2) $disp_i < diff \leq frame.width / 2$

$$window_i = window_{i-1} + disp_i \quad (22)$$

$$disp_{i+1} = disp_i + a_1 \cdot sign(diff) \quad (23)$$

3) $frame.width / 2 < diff$

$$window_i = window_{i-1} + disp_i \quad (24)$$

$$disp_{i+1} = disp_i + a_2 \cdot sign(diff) \quad (25)$$

경우 1) 공의 위치와 관심영역 창의 중심 사이의 거리가 이전 이동 거리보다 짧은 경우이다. 이 경우, 관심영역 창의 위치는 변경되지 않는다. 그 대신 *dispi+1* 의 값에 α 를 곱하여 감소시킨다.

경우 2) 관심영역 창을 *dispi*만큼 이동한 후, *dispi+1* 를 식(23)에 따라 변경한다. $|disp_i| > 0$ 이므로 관심영역 창은 쉽게 기존 방향의 반대 방향으로 이동 하지 않는다.

경우 3) 2와 동일하나 공이 빠르게 움직이는 경우에 대응 하기 위해 더 큰 가속 파라미터를 갖는 경우이다. ($0 < a_1 < a_2 < 1$)

이렇게 디스플레이 창의 위치를 결정하고 나면 마지막으로 그림 11과 같은 경우를 처리하여 사용자에게 보여줘야 한다. 디스플레이 창이 (a)로 결정된 경우 사용자에게는 (b)를 보여

줘야 한다. 단 위의 과정에서 사용된 *window_i*는 바꾸지 않는다.

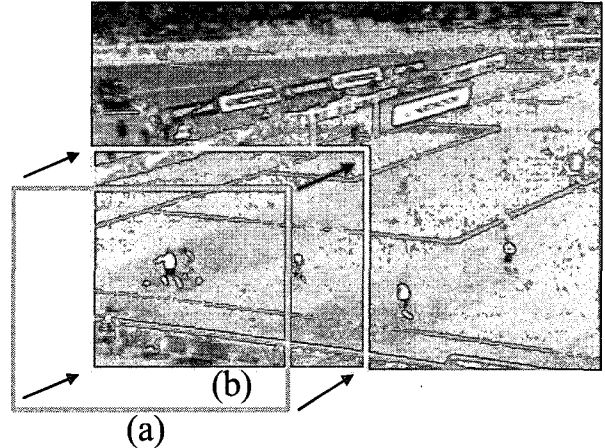


그림 11. (a) 계산된 관심영역의 위치가 프레임을 벗어나면 (b)처럼 관심영역의 위치를 프레임 안으로 이동시켜 디스플레이 한다. Fig. 11. (a) The calculated location of the ROI window. (b) The bounded window inside the frame.

IV. 실험 결과

이 시스템은 Win32 환경에서 Visual Studio 2003 (C++) 를 이용하여 개발하였으며 MPEG 복호화를 위해 FFMpeg

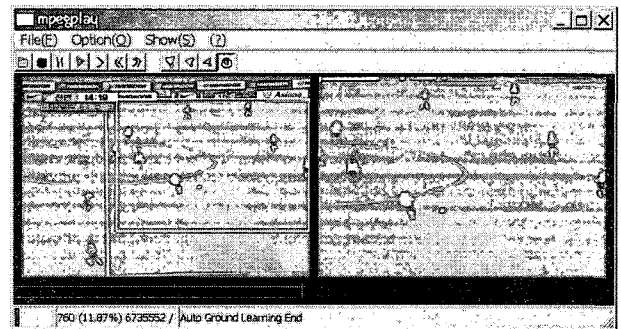


그림 12. 실험 결과. 모든 과정은 자동적이며 실시간으로 수행된다. (왼쪽) 원 영상 위에 노란색의 관심영역 창이 놓여있다. (오른쪽) 관심영역이 디스플레이 크기에 맞게 확대 되어있다. Fig. 12. Screen shot of our system. All processes are fully automatic and run in real-time. (Left) Original frame and ROI window indicated by yellow box. (Right) ROI is magnified and displayed.

라이브러리를 이용하였다. 우리는 5분짜리 축구 방송 세가지를 352x240, 29.97fps의 MPEG-1 포맷으로 부호화하여 실험에 이용하였다. (대한민국 대 독일: 2004년 친선 경기, 대한민국 대 쿠웨이트: 2006 월드컵 최종 예선, 일본 대 이탈리아: 2004년 아테네 올림픽.) 모든 과정은 Pentium-4 2.4GHz PC에서 실시간으로 수행되었다. 관심영역 결정을 위한 파라미터는 각각 $\alpha=0.85$, $\alpha_1=0.05$, $\alpha_2=0.15$ 로 설정하였다. 또한 352x240의 크기를 갖는 원래 영상에 대해 각 변에 대해 1:0.68의 비율로 축소된 240x164 크기의 관심영역을 추출하였다.

1. 그라운드 검출 성능의 평가

제 3장에서 제안된 그라운드 검출 성능을 검증하기 위하여 100개의 서로 다른 축구경기 비디오에서 그라운드에 해당하는 부분을 160x80 크기로 잘라 낸 후 식 (10) 및 (12)를 이용하여 검출된 그라운드 블록의 갯수를 세었다. 즉, 잘라낸 160x80의 안에 존재하는 16x16 크기의 블록 50개 중 몇 개가 그라운드 블록으로 인식되는 가를 알아 보았다. 다양한 잔디 색상과 일광 조건의 변화로 인한 그림자 간섭 등에도 불구하고 5000 (=50 x 100)개의 그라운드 블록 중 98.74%에 해당하는 4937개의 블록이 그라운드 블록으로 올바르게 인식됨을 알 수 있었다.

2. 샷 분류 과정에 대한 평가

샷 분류 판정에 대한 결과를 표 1에 나타내었다. 표에서 볼 수 있듯이 90% 이상의 높은 정확도를 보였다.

표 1. 샷 분류 과정의 성능
Table 1. Accuracy of the proposed shot classification technique.

	원거리 샷	기타	전체
대한민국 vs. 독일	93.84 %	99.26 %	95.18 %
대한민국 vs. 쿠웨이트	98.08 %	82.97 %	93.75 %
일본 vs. 이탈리아	94.81 %	94.71 %	94.78 %

3. 관심영역 결정 과정에 대한 평가

결정된 관심영역을 평가하는 객관적인 기준에 대해서는 말하기 힘들다. 하지만 우리는 비디오 프레임 내에 공이 오랫동안 존재하지 않는 경우 축구 경기 시청이 만족스럽지 않다는 것을 경험적으로 알고 있다. 이에 따라 우리는 관심영역 평가를 위해 관심영역 내에 실제로 공이 존재하는 가를 측정해보았다. 관심영역의 크기로는 240x164 크기의 창을 이용하였으며 그 결과를 표 2에 나타내었다. 80~90%의 높은 확률로 공이 관심영역 내에 존재함을 알 수 있었다.

표 2. 관심영역 결정 과정의 성능
Table 2. Evaluation of ROI determination

	관심영역 내 공의 존재 여부
대한민국 vs. 독일	91.08 %
대한민국 vs. 쿠웨이트	79.20 %
일본 vs. 이탈리아	84.17 %

또한 우리는 얼마나 장시간 동안 공이 관심영역을 벗어나는지에 대해 관찰하기 위해, 공 찾기 과정의 결과를 그림 13처럼 시간에 따른 그래프로 나타내었다. 각각의 그래프는 5분짜리 비디오 클립을 뜻하며, 회색 영역은 공을 제대로 찾은 경우, 검정색 영역은 공을 찾지 못한 경우를 의미한다. 추가적으로 진한 회색 영역은 공이 원 영상에 존재하지 않는 경우, 연한 회색 영역은 원거리 샷으로 판정되지 않아 공 찾기 과정이 수행되지 않은 시간대를 의미한다. 그 그래프에서 볼 수 있듯이 대부분의 시간 동안 공을 제대로 찾아내고 있으며 공을 놓치는 시간이 길지 않음을 알 수 있다. 또한 3장 7절에서 설명된 바와 같이 시간에 따른 관심영역 창의 흔들림은 최소화되어 있으며 이는 [15]에서 동영상 데모의 구동을 통해 직접 확인할 수 있다.

4. PDA에서의 성능 평가

우리는 실제 모바일 기기에서 이 시스템의 성능을 측정하기 위해 PDA용으로도 개발하였다. 개발에는 Microsoft

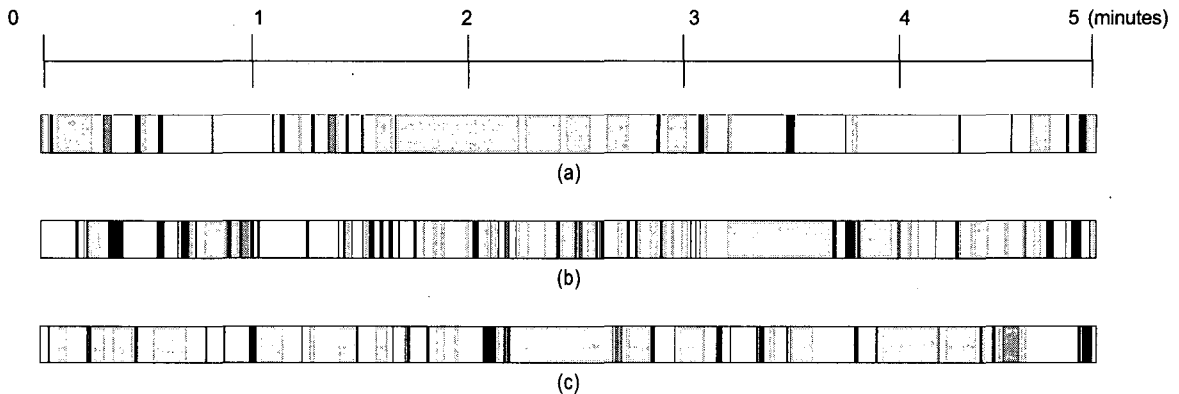


그림 13. 시간 축에 따른 공과 ROI 위치의 관계. 흰색: 공이 ROI 내에 존재함. 검정색: 공이 ROI 내에 존재하지 않음. 진한 회색: 공이 원 영상 내에 존재하지 않음. 연한 회색: 비 원거리 샷인 관계로 공 추적 과정이 수행되지 않음. (a) 대한민국 vs. 독일 (b) 대한민국 vs. 쿠웨이트 (c) 일본 vs. 이탈리아.
 Fig. 13. The relationship of locations between the ball and ROI window along the time axis. white: the ball exists inside ROI window, black: the ball out of ROI window, dark gray: the ball out of original frame, light gray: ball detection not conducted due to non-long-shot frame. (a) Korea Vs. Germany. (b) Korea Vs. Kuwait. (c) Japan Vs. Italy.

Embedded Visual Studio 4.0 을 이용하였다. PDA의 스펙은 표 3에 나타내었다. 성능 측정을 위해 우리는 MPEG1 포맷으로 97초의 재생 시간을 갖는 축구 비디오 하나를 이용하였다. 각 함수의 수행 시간을 측정하기 위해 세 번 반복 실험하여 걸리는 시간의 평균을 측정하였으며 그 결

과는 표 4에 나타내었다. 관심영역 분석에 걸린 평균 시간은 다음의 식을 이용하여 계산하였으며 34.14fps를 나타내었다.

$$\text{Average processing rate for analysis} = 1000000 / \left(\frac{LST + NLT}{N} \right) \text{ (frames/sec)} \quad (26)$$

표 3. 테스트에 이용한 PDA의 사양

Table 3. Test environment for PDA porting (Hardware)

하드웨어	
모델명	HP iPAQ hx4700
CPU	Intel PXA270 624MHz
RAM	64 MB
화면	VGA (640x480), 4.0 inch

여기서 LST 는 원거리 샷 전체를 처리하는 데 걸린 평균 시간이며, NLT 는 비 원거리 샷에 소요된 시간이다 (단위: μs). N 은 전체 프레임의 개수이다.

또한 MPEG 디코딩 시간과 화면 출력이 포함된 평균 재생 속도는 다음과 같이 계산되며 15.33fps의 속도를 보였다.

표 4. PDA에서 이 시스템을 돌렸을 때 각 함수별 걸린 평균 시간
 Table 4. Elapsed time for intelligent displaying on PDA.

	프레임의 갯수 N	걸린 시간 (μs)			평균 소요 시간(μs) T	T/N (μs)	
		1	2	3			
원거리 샷	전체	2174	80113470	79954110	80116630	80061403(=LST)	36827
	공 검출	2174	61968150	61978940	61970870	61972653	28506
	2차 공 검출	796					
	샷 분류	2174	18145320	17975170	18145760	18088750	8320
비 원거리 샷	전체	708	5292070	5215670	5215100	5240947(=NLT)	7402
	샷 분류	708	5292070	5215670	5215100	5240947	7402
전체 소요 시간 (MPEG 복호화와 화면 그리기 시간 포함)		2912	189930280	189859220	190053940	189947813	65229

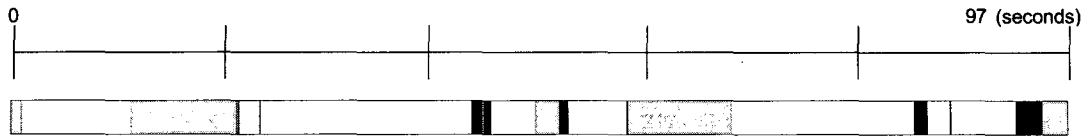


그림 14. PDA에서 수행된 관심영역 결정 과정의 평가.
Fig. 14. Evaluations of ROI determination on PDA.

T 는 재생에 걸린 총 시간을 뜻한다. 그림 14에서 볼 수 있듯이 PDA에서도 샷 분류 정확도는 떨어지지 않았다.

$$\text{Average playing rate} = \frac{1000000}{N} \frac{T}{N} \text{ (frames/sec)} \quad (27)$$

V. 결론 및 토의

우리는 축구 비디오의 원거리 샷을 검출하고 관심영역을 지정하여 확대 재생하는 지능형 디스플레이 시스템을 제안하였다. 제안한 알고리즘이 제대로 동작하는 것을 보이기 위해 우리는 여러 개의 국가간 축구 경기 비디오 클립을 이용하여 테스트 하였다. 이전 장에서 보인 것처럼 제안한 시스템은 일반 PC 상에서 실시간으로 동작하였으며 PDA 에서 관심영역 결정 과정은 34.14fps, 전체 디스플레이에는 15.33 fps의 속도를 보였다.

현재 우리는 제안된 알고리즘이 그라운드 색상이나 일광 조건 등, 그라운드 조건의 변화에 보다 강인하도록 더욱 많은 실험 영상들에 대해 강건성 실험을 진행 중에 있다. 또한 축구 경기 비디오에 있어서 원거리 샷의 시청은 때로는 양측 선수들의 전반적인 배치 상황을 파악하는데 유용하게 이용될 수도 있다. 따라서, 이동형 단말기에 관심영역 디스플레이 기능을 on/off 할 수 있는 버튼이 제공된다면 시청자의 선호도에 따라 자연스럽게 디스플레이 모드를 변경하는 일이 가능할 것이다. 또한 관심영역이 추출되고 확대된 디스플레이에서 점수 상자 (score box)가 사라지는 문제를 해결하기 위하여 원 영상에서 점수 상자를 자동으로 검출하여 관심영역의 임의의 위치에 표시해 주는 알고리즘의 개발을 현재 진행 중에 있다. 이 또한 시청자의 선호도에

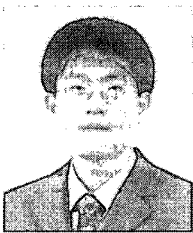
따라 쉽게 on/off 시켜주는 일이 가능하다. 이러한 기술들은 이동형 콘텐츠 개인화^[13]의 관점에서 매우 중요하다. 즉, 스크린의 크기가 제한된 이동형 멀티미디어 단말기의 경우, 위와 같이 다양한 선택권을 사용자에게 제공함으로써 더욱 향상된 시청 경험을 보장할 수 있게 될 것이다. 제안된 시스템을 인코더와 디코더 중 어느 쪽에 설치할 것인가에 관한 문제도, [13]에서 논의되고 있는 콘텐츠 개인화의 구현 위치 및 수준에 관련된 것으로서, 단말기의 처리능력, 송신측에서의 신호처리 능력, 응용 분야 등을 감안하여 결정되어야 할 것이다.

참고 문헌

- [1] H. Knoche, J. D. McCarthy, and M. A. Sasse, "Can small be beautiful?: assessing image resolution requirements for mobile tv," in *proc. of the 13th annual ACM international conference on Multimedia*, pp. 829 - 838, ACM Press, (New York, NY, USA), 2005.
- [2] W. H. Cheng, W. T. Chu, and J. L. Wu, "A visual attention based region of interest determination framework for video sequences," *IEICE Transactions on Information and Systems*, E 88D, pp. 1578 - 1586, 2005.
- [3] L. Itti, C. Koch, and E. Niebur, "A model of saliency based visual attention for rapid scene analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no.11, pp. 1254 - 1259, Nov. 1998.
- [4] Y. F. Ma and H. J. Zhang, "A model of motion attention for video skimming," in *Proc. ICIP*, pp. 129 - 132, 2002.
- [5] J. D. Foley, A. van Dam, S. K. Feiner, J. F. Hughes, *Computer Graphics Principles and Practice*, pp. 592,593, Addison Wesley, 1996.
- [6] A. Ekin, A. M. Tekalp, and R. Mehrotra, "Automatic soccer video analysis and summarization," *IEEE Transactions on Image Processing*, vol. 12, pp. 796 - 807, 2003.
- [7] K. Wan, X. Yan, X. Yu, and C. Xu, "Real time goal mouth detection in mpeg soccer video," in *proc. of the eleventh ACM inter-*

- national conference on Multimedia, pp. 311 - 314, ACM Press, (New York, NY, USA), 2003.
- [8] R. W. Lienhart, "Comparison of automatic shot boundary detection algorithms," in proc. of Storage and Retrieval for Image and Video Databases VII 3656(1), pp. 290 - 301, SPIE, 1998.
- [9] Y. Seo, S. Choi, H. Kim and K. Hong, "Where Are the Ball and Players? Soccer Game Analysis with Color Based Tracking and Image Mosaick," in proc. of ICIAP '97: Proceedings of the 9th International Conference on Image Analysis and Processing Volume II, pp. 196 - 203, Springer Verlag, London, UK, 1997.
- [10] Y. Ohno, J. Miura and Y. Shirai, "Tracking Players and a Ball in Soccer Games," in proc. of Int. Conf. on Multisensor Fusion and Integration for Intelligent Sys., Taipei, Taiwan, 1999.
- [11] X. Yu, C. Xu, H. W. Leong, Q. Tian, Q. Tang, K. Wan, "Trajectory Based Ball Detection and Tracking with Applications to Semantic Analysis of Broadcast Soccer Video," in proc. of MULTIMEDIA '03: Proceedings of the eleventh ACM international conference on Multimedia, pp. 11 - 20, ACM Press, (New York, NY, USA), 2003.
- [12] X. Fan and X. Xie and H. Q. Zhou and W. Y. Ma, "Looking into Video Frames on Small Displays," in proc. of the eleventh ACM international conference on Multimedia, pp. 247 - 250, ACM Press, (New York, NY, USA), 2003.
- [13] S.R. Subramanya and B.K. Yi, "Mobile Content Customization," IEEE Multimedia, vol. 12, Issue 4, pp. 104-105, 2005.
- [14] TCPMP, <http://tcpmp.corecodec.org>
- [15] <http://vega.icu.ac.kr/~ckim/projects>

저 자 소 개



서 기 원

- 2004년 2월 : 한국과학기술원(KAIST) 전자전산학과 전산학전공(학사)
- 2006년 2월 : 한국정보통신대학교(ICU) 공학부 디지털미디어전공 (석사)
- 2006년 1월~현재 : (주) WRG 연구원



김 창 익

- 1989년 2월 : 연세대학교 전기공학과 (학사)
- 1991년 2월 : 포항공과대학교(POSTECH) 전기전자공학과 (석사)
- 1991년 1월~1997년 7월 : SKC Ltd. R&D 센터 선임 연구원
- 2000년 12월 : 워싱턴주립대학교 전기공학과 (박사)
- 2000년 12월~2005년 1월 : Senior Member of Technical Staff, Epson Palo Alto Laboratory, Epson R&D Inc.
- 2005년 2월~현재 : 한국정보통신대학교(ICU) 공학부 조교수