

# 필드와 모션벡터의 특징정보를 이용한 스포츠 뉴스 비디오의 장르 분류

송 미 영<sup>\*</sup> · 장 상 현<sup>\*\*</sup> · 조 형 제<sup>\*\*\*</sup>

## 요 약

비디오의 브라우징, 검색, 조작을 위해서 비디오 내용을 기술하는 색인이 요구된다. 지금까지, 색인의 구성은 대부분 비디오 내용에 제한된 키워드를 수작업으로 할당하는 전문가에 의해 수행되었는데 이는 비용과 시간을 소비하는 작업이므로 비디오 내용을 자동으로 분류하는 것이 필요하다.

이 연구는 축구, 골프, 야구, 농구, 배구 등 5종의 스포츠 뉴스 비디오의 분석과 요약을 위해서 자동적이고 효율적인 방법을 제안한다.

우선, 스포츠 뉴스 비디오를 앵커 장면과 스포츠 기사 장면으로 분류한다. 장면 분류는 앵커 장면의 영상 전처리와 색상 특징을 기반으로 한다. 그리고 필드의 우세색상과 모션 방향을 특징으로 이용하여 스포츠 장면을 5개의 장르로 분류한다.

241개의 스포츠 뉴스 장면에 대한 실험에서 75%의 정확도를 얻었다. 따라서 제안된 기법은 향후 개별 스포츠 뉴스와 스포츠 하이라이트를 위한 뉴스 비디오를 검색하는데 이용될 수 있을 것이다.

키워드 : 장르 분류, 비디오 요약, 스포츠 뉴스 비디오, 우세색상, 모션방향

## Automatic Genre Classification of Sports News Video Using Features of Playfield and Motion Vector

MiYoung Song<sup>\*</sup> · SangHyun Jang<sup>\*\*</sup> · HyungJe Cho<sup>\*\*\*</sup>

### ABSTRACT

For browsing, searching, and manipulating video documents, an indexing technique to describe the video contents is required. Until now, the indexing process is mostly carried out by specialists who manually assign a few keywords to the video contents and thereby this work becomes an expensive and time consuming task. Therefore, automatic classification of video content is necessary.

We propose a fully automatic and computationally efficient method for analysis and summarization of sports news videos for 5 sports news video such as soccer, golf, baseball, basketball and volleyball.

First of all, sports news videos are classified as anchor-person shots, and the other shots are classified as news reports shots. Shot classification is based on image preprocessing and color features of the anchor-person shots. We then use the dominant color of the field and motion features for analysis of sports shots. Finally, sports shots are classified into five genre type.

We achieved an overall average classification accuracy of 75% on sports news videos with 241 scenes. Therefore, the proposed method can be further used to search news video for individual sports news and sports highlights.

Key Words : Genre Classification, Video Summarization, Sports News Video, Dominant Color, Motion Vector

### 1. 서 론

디지털 동영상 압축과 네트워크 기술의 발전으로 동영상에 대한 수요가 폭발적으로 늘어나고 있으며, 사용자는 동영상의 내용 정보를 직접 검색하려는 요구가 증가하고 있다. 즉, 인기 영화에서 “특정 배우가 연기하는 부분”만 보고

싶다거나, 축구 경기에서 “골을 넣는 장면”만을 찾고자하는 욕망이 있을 것이다. 이처럼 동영상이 포함하고 있는 내용에 대한 검색을 위해서 동영상에 포함된 소리, 영상의 이해(Understanding)와 분석(Analysis)을 통해 그 내용을 대체(Substitution)하는 기술이 필요하게 되었다[1,2].

최근 인터넷 포털에서도 이미지와 동영상에 대한 검색을 강화하고 있으나, 그 방법이 파일 단위의 속성을 수작업으로 입력한 메타 데이터나 방송 대본과 같이 텍스트를 기반으로 검색하는 수준이다. 그러나 오래전부터 동영상의 색상,

\* 정 회 원 : 수원여대 멀티미디어통신과 전임강사

\*\* 정 회 원 : 한국교육학술정보원 책임연구원(기획혁신팀장)

\*\*\* 정 회 원 : 동국대학교 컴퓨터·멀티미디어공학과 교수

논문접수 : 2006년 9월 14일, 심사완료 : 2006년 12월 2일

모션, 질감, 오디오 등을 활용하여 동영상 내용을 자동으로 색인하는 분야에 대해 많은 연구를 하고 있다. 이는 메타데이터를 작성하는 수고를 덜 수 있을 뿐만 아니라 내용 기반으로 동영상을 검색할 수 있기 때문이다. 우리가 쉽게 접하는 동영상은 영화, 드라마, 스포츠, 다큐멘터리, 뉴스, 교양, 오락 등 다양한 장르가 있다. 이 중 대부분이 비정형화된 형태의 동영상으로써 자동 색인이 매우 어렵지만, 뉴스와 같이 정형화된 동영상은 자동 색인과 내용 검색이 가능하다[3]. 그래서 뉴스 비디오는 자동화된 비디오 색인을 통해 원하는 기사를 찾는 주문형 뉴스(News-On-Demand) 시스템까지 발전하고 있다[3].

따라서, 이 연구는 동영상의 내용에 대한 의미적 검색을 위해서 첫째는 스포츠 뉴스 비디오를 의미 있는 장면으로 분리하고, 둘째는 선정된 주요 키프레임에서 특징 정보를 추출하고 내용을 자동으로 분석하여 색인을 구성함으로써 궁극적으로 최근 검색 분야의 연구 핵심인 의미적 검색(Semantic Retrieval)이 가능하도록 한다.

이 연구의 2장에서는 스포츠 비디오에서의 특징정보와 이를 활용한 분야 등에 관련된 연구들을 살펴보고, 3장은 스포츠 뉴스 비디오의 처리과정을 설명하고, 4장에서는 앵커와 스포츠 장면의 분리를 위한 앵커 장면 추출방법을 알아본다. 5장에서는 추출된 스포츠 장면에 대해서 필드 색상과 모션 방향의 분석 방법을 살펴보고, 6장에서는 분석된 특징을 기반으로 스포츠 장면의 장르를 결정한다. 그리고 7장에서는 제안한 연구의 실험결과를 확인하고, 마지막으로 8장에서는 결론을 맺는다.

## 2. 관련연구

현재 스포츠 비디오의 연구는 고수준 의미적 질의(High-Level Semantic Queries)를 기반으로 원하는 스포츠 장면을 검출하는 색인 및 검색 시스템(Indexing and Retrieval Systems), 스포츠 이벤트 장면내에 부가적인 정보를 삽입하여 관찰자에게 사실적이고 새로운 시각적 체험을 위한 증강현실표현(Augmented Reality Presentation) 등으로 분류할 수 있다.

### 2.1 스포츠 비디오의 일반적인 연구

대부분 특정 스포츠 비디오내에서 스포츠 경기의 이벤트와 하이라이트 검출, 경기 진행 절차상의 구조적 분석, 선수들의 형태나 윤곽을 추출하는 객체의 검출과 분할, 경기장내의 공과 선수의 위치를 파악하기 위한 공과 선수의 추적, 카메라 조정 등의 연구가 진행되고 있으며, 최근에는 스포츠 경기의 모습을 3차원으로 재구성하는 연구가 이루어지고 있다.

이벤트 검출(Event Detection)은 골 슈팅, 파울, 득점 등과 같은 사건(Event) 장면을 추출하는 것으로 비디오의 구조, 이벤트, 내용 등의 여러 속성들을 결합하여 분석한다. 이벤트의 내용 분석은 특징정보(Features)와 의미적 개념

(Semantic Concepts)간의 관계성을 명시하거나 특정 정보와 의미적 개념간의 관계성을 획득하는 것이다. 하이라이트 검출(Highlights Detection)은 특별히 관심있는 사건 장면을 추출하는 것으로 예를 들어, 테니스 경기에서의 득점 장면이나 코치가 선수들의 능력을 개선을 돕기 위해서 선수의 실수 장면을 추출하는 것이다. 이는 하이라이트의 발생 시점, 발생 시작 시점부터 끝나는 시점을 찾는다. 동영상내의 이벤트와 하이라이트 검출의 특징 정보로는 비디오, 오디오, 텍스트 채널 등을 활용한다.

구조 분석(Structure Analysis)은 스포츠의 시간적인 그룹화를 의미한다. 예를 들어, 테니스 경기는 시간상의 구조가 명확한 스포츠로 몇 개의 세트로 구성되어 있고, 다시 몇 게임으로 나누어져있고, 포인트 부분으로 이루어진다. 축구인 경우는 구조적인 특징이 적다. 스포츠 비디오의 구조 분석은 단순히 경계(Cut) 검출이 아니고, 게임의 규칙에 대한 지식과 결합된다. 이는 경기 부분과 휴식 부분으로 구별하는 것으로 불필요한 휴식 장면 부분을 삭제함으로써 비디오의 시간길이를 단축할 수 있기 때문이다[23].

객체 검출과 분할(Object Detection and Segmentation)은 스포츠에서 관심이 집중되는 객체를 추출하는 것이다. 즉, 스포츠 비디오에서는 객체의 중요한 동작에 초점을 맞춰서 보여준다. 예를 들면, 축구나 테니스 스포츠에서는 오직 선수의 위치에 관심이 집중하고 있고, 체조나 다이빙 스포츠는 선수의 자세와 움직임에 주요 관심사이다. 이러한 스포츠 분석은 정확한 객체 형태를 검출할 수 있는 자동화 분할 알고리즘과 객체 형태의 자세를 인식할 수 있는 알고리즘을 요구한다[17,18]. 이러한 정확한 객체 영역은 3차원으로 재구성하여 시각화하는데 필요한 요소가 된다[24].

공과 선수 추적(Ball and Player Tracking)은 비디오 분석에서 공과 선수의 위치는 중요한 요소이므로 이 위치를 추적하기 위한 방법이다. 특히 작은 공의 추적은 어려운 문제이므로[25], 여러 개의 프레임들 통해서 공을 추적하는 강력한 추적 알고리즘을 필요로 한다[19,20].

카메라 조정(Camera Calibration)은 선수 추적시 영상 좌표상에서는 선수의 위치를 판별하기 쉽지만, 실제 경기장에서의 선수의 위치에 대한 정보를 얻기 힘들기 때문에 영상 좌표와 실제 위치상간의 기하학적인 매핑을 위해서 요구된다. 일반적으로 경기장의 실제 위치는 선으로 표시되어 있다는 기반 지식을 통해서 카메라 조정에 따라 선을 식별하여 이를 활용한다[17-18, 26].

3차원 재구성(3D Reconstruction)은 여러 개의 카메라로 여러 위치에서 얻은 영상을 통해서 경기장을 3차원으로 재구성하는 방법이다[15]. 일반적으로 여러 대의 고정된 카메라의 사용은 카메라의 조정(Calibration)이 단순하고, 계산 시간을 감축하고, 카메라의 조정의 정확성을 개선할 수 있다. 더욱 중요한 것은 현재는 방송 스포츠 비디오는 축구 경기장과 같은 커다란 경기장이나 선수 자세의 일부분만을 표현하지만, 앞으로는 3차원 재구성을 통해서 경기장의 전체 표현이 가능할 것이다.

2.2 의미적 스포츠 분류의 연구

스포츠의 다양한 특징을 검색할 수 있도록 스포츠의 의미적 장르 분류와 요약(Summarization)에 대한 연구들이 진행되고 있다. 논문 [27-30]에서는 정의된 스포츠 장르로 분류하고, [31-33]는 스포츠내의 의미적 요약을 검출한다.

Wang은 스포츠 비디오 샷의 의미적 장르 분류를 위한 방법을 제안하였다. 특징 정보로는 모션과 색상을 기반으로 경기장의 색상비, 배경의 모션비와 일관성비 등을 이용하여 결정트리(Decision Tree)를 통해서 의미적 샷으로 분류하였다. 샷의 의미적 장르 분류는 아이스 하키, 골프, 야구, 농구의 스포츠로 제한하였고, 또한 스포츠 뉴스의 시작과 끝부분, 텍스트, 그리고 기타 등으로 구분하였다. 제안된 방법을 통해서 정의한 4개의 장르를 포함한 1시간이상의 15개 스포츠 뉴스 비디오에 대한 실험 결과는 88%의 정확도를 얻었다[27].

Gibert는 아이스하키, 미식축구, 축구 등의 스포츠 장르를 분류하는 방법을 제안하였고, 분류의 정확도는 93%를 얻었다. 스포츠 비디오는 아이스하키(39분), 야구(66분), 미식축구(67분) 그리고 축구(47분) 등을 포함한 220분의 비디오로 각 스포츠 종목의 비디오에서 10분은 학습데이터로 사용하고 나머지 부분은 실험 데이터로 사용하였다. 여기서 스포츠 비디오의 샷 분류는 100초(3000프레임)씩 경계를 나누어 스포츠 장르를 판별하였다[28]. 이 방법은 한 스포츠 경기 내용만을 가지고 학습 데이터와 실험 데이터로 분할하여 스포츠 장르를 분석하였기 때문에 스포츠 내용의 특징 정보들이 일관성을 유지할 수 있어서 정확도를 높게 얻을 수 있었다. 하지만 경기장이나 선수의 특징 정보가 다른 동일 종목의 스포츠 비디오에 대해서 정확도를 보장할 수 없다.

Sahouri는 주성분 분석법(Principal Component Analysis)를 활용하여 농구, 아이스하키, 축구, 배구 등의 스포츠 종목에 대한 모션 패턴을 정의하였고 스포츠내에서 경기와는 무관한 불필요한 프레임 정보들을 제거하여 전체 비디오 프레임율을 줄이기 위한 방법을 제안하였다[29]. 이 방법에서의 특징 정보인 모션 패턴은 짧은 길이의 비디오인 경우는 모션 패턴의 일관성을 유지하기 어렵기 때문에 시간이 긴 길이의 스포츠 비디오 데이터를 사용하였다.

Hong은 스포츠 비디오 샷을 클러스터링하기 위해 넓은 각도의 샷을 검출하여 테니스, 농구, 배구, 미식축구 등의 구조를 정의하였고[30], Assfalg은 다양한 스포츠 종목의 비디오에서 경기장, 선수, 관람객의 의미적 요약을 위한 방법을 제안하였고[31,32], Ekin은 우세색상을 통해서 카메라의 거리가 먼 샷, 중간 샷, 관람객 샷 등으로 구분하였다[33].

이와 같은 대부분 스포츠의 의미적 장르 분류나 요약에 대한 연구에서 사용되는 스포츠 비디오의 데이터는 한가지 스포츠 종목만을 다루고 있거나, 시간상으로 짧은 길이의 스포츠 뉴스 비디오보다는 긴 길이의 스포츠 비디오에서 이루어진다. 이는 시간적으로 짧은 길이의 스포츠 뉴스 비디오는 전체 스포츠 경기내에서 중요한 하이라이트 부분만을 압축하여 편집하므로 장시간의 전체 스포츠 경기보다도 일

관성 있는 특징 정보를 검출하기가 어렵기 때문에 스포츠내에서 특징 정보 추출의 용이성과 일관성을 위해서 긴 길이의 스포츠 비디오를 선택한다. 논문[27]에서는 스포츠 뉴스 비디오를 가지고 제안하였지만 스포츠 뉴스의 시간상 길이는 1시간이상으로 비교적 긴 길이의 스포츠 뉴스 비디오를 다루고 있다.

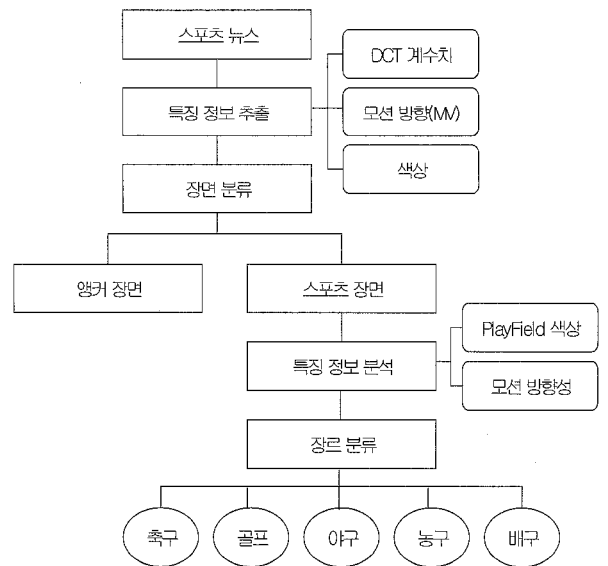
따라서 이 연구는 하이라이트 중심인 5분이내의 짧은 스포츠 뉴스 비디오내에서 자동적으로 앵커와 스포츠 장면을 분할하고, 추출된 스포츠 장면의 프레임에서 우세색과 모션 방향의 검출을 통해서 스포츠의 장르를 분류하는 방법을 제안한다.

3. 스포츠 장르 분류의 처리 과정

이 연구는 스포츠 뉴스 비디오내에서 스포츠 기사 장면을 검출하고 분석하여 스포츠 장르로 분류하는 것이다.

입력 영상은 스포츠 뉴스 비디오를 대상으로 영상의 크기는 320 X 240이다. 비디오 영상을 MPEG4 영상 압축 과정시 전체 크기 영상을 사용하는 것이 아니라 외각 부분은 제외한다. 이는 시간적인 비용을 줄일 수 있고, 영상의 상단 부분과 하단 부분에는 방송사명이나 제목, 자막의 표시로 인해 동일한 장면인데도 불구하고 DCT계수치가 급변하여 잘못된 장면(Scene)이나 샷(Shot)의 검출을 막기 위해서이다.

그리고 앵커 장면 검출과 스포츠 기사 장면의 스포츠 필드 색상을 분석하기 위해서 영상 처리(Image Processing) 과정이 이루어진다. 영상 처리의 목적은 영상에서의 불균일한 색상을 부드럽게 만들어서 색상의 분포도를 균일화하고, 또한 색상수의 분포도를 줄여서 주변의 유사 색상들을 동일한 색상으로 그룹화 효과를 얻기 위해서 이다. 즉, 색상의 분포도를 균일하게 하기 위해서 키프레임 영상에 대해 5X5 미디언 필터링(Median Filtering)의 영상 처리를 과정을



(그림 1) 스포츠 장르 분류의 처리 과정

거친다. 그리고 색상수의 분포를 줄이고 주변의 유사 색상들과의 그룹화를 위해서 16수준의 양자화(16 Level Quantization) 처리 과정이 이루어진다.

따라서 MPEG4 영상 압축시 영상의 크기는 256 X 128로 하여 DCT 계수치와 모션 벡터를 산출하고 영상 처리를 거쳐 영상의 화소차를 비교하여 앵커 기사 장면과 스포츠 기사 장면으로 분리한 후, 검출된 스포츠 기사 장면의 키프레임들에 대해서 영상 처리를 통해서 필드 색상을 분석하고 모션 방향성의 특징정보를 활용하여 축구·골프·야구·농구·배구 등의 스포츠 장르로 분류한다.

#### 4. 앵커와 스포츠 장면 검출

스포츠 뉴스 비디오내의 앵커 장면들은 모두 유사하고 스포츠 기사 장면간의 경계이므로 앵커 장면을 검출하는 것은 동시에 스포츠 기사 장면을 얻을 수 있으므로 앵커 장면과 스포츠 장면을 분리할 수 있다.

앵커 장면의 검출은 DCT 계수치와 시간적 특징을 이용하여 후보 앵커 장면을 추출하고, 이를 첫 앵커 프레임(참조 프레임)과의 화소 색상 비교를 통해서 유사도(similarity)가 높으면 앵커 장면으로 확정한다.

스포츠 뉴스 비디오의 MPEG4 영상 압축시 I 프레임의 매크로 블럭에 대해서 Y, Cr, Cb의 DCT 계수치의 합을 산출하고, 이웃하는 프레임간의 DCT 계수차에 대한 평균을 구한다(식 1).

$$DCT_k = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (B_{k+1}[x][y] - B_k[x][y]) \quad (식 1)$$

여기서 B는 Y, Cr, Cb의 DCT 계수치의 합에 대한 매크로 블럭이고, M과 N은 매크로 블럭의 가로축과 세로축의 총 개수이다.

앵커 장면 추출에 사용되는 임계치( $Th_{shot}$ )는 (식 2)와 같이 전체 영상의 통계적 특성을 이용하여 (식1)에서 산출된  $DCT_k$ 에 대해 평균값(AVG)과 표준편차(STD) 값에 가중치  $\alpha$ 를 적용하여 임계치 이상인 경우를 후보 앵커 장면으로 간주한다.

$$Th_{shot} = AVG + \alpha * STV \quad (식 2)$$

추출된 후보 앵커 장면의 시작과 마지막 프레임에 대해서 참조 프레임과의 화소값을 비교하여 정확한 앵커 장면의 범위를 검출한다. 이 연구에서 참조 프레임이란 스포츠 뉴스 비디오의 시작 프레임은 항상 앵커로부터 시작한다는 가정하에 스포츠 뉴스 비디오의 첫 번째 프레임을 말한다.

이와 같이 (그림2(a))의 참조 프레임과 후보 앵커 장면으로 검출된 프레임인 (그림3(a))과의 화소값을 비교하기 위해서 (그림3(b)(c))과 같이 5X5 미디언 필터링과 16수준 양자



(a) 원 영상 (b) 영상처리후 결과

(그림 2) 참조 프레임



(a) 원 영상 프레임 (b) 5X5 미디언 필터링 (c) 16 level 양자화

(그림 3) 후보 앵커 프레임

화의 영상 처리 과정을 거친 후, 후보 앵커 장면의 프레임인 (그림3(c))와 참조 프레임인 (그림2(b))간의 화소차의 평균이 임계치 이상인 경우의 프레임을 최종 앵커 장면으로 검출한다.

#### 5. 스포츠 장면 분석

검출된 앵커 장면을 제외한 스포츠별 기사 장면에 대해서 (식 1)과 (식 2)를 적용하여 스포츠 장면 분석을 위한 키프레임을 추출한다.

##### 5.1 필드(Playfield) 분석

스포츠 필드는 스포츠마다 서로 다른 형태나 색상의 특성을 가지고 있다. 스포츠별 필드를 살펴보면, (그림 4(a))와 (그림 4(b))는 각각 축구와 골프의 필드로 주로 초록색이고, 야구 필드는 초록색, 흙색(주황색)이거나 (그림 4(c))와 같이 두 가지 색상을 포함하기도 한다. (그림 4(d))는 배구로 연분홍색의 필드를 가지고, (그림 4(e))의 농구 필드는 노란색(주황색) 등의 특징을 가진다. 따라서 스포츠를 분류하는 데 있어서 필드의 색상은 유용한 특징정보라고 할 수 있다.

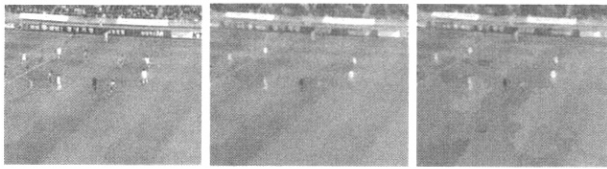


(a) 축구 (b) 골프 (c) 야구



(d) 배구 (e) 농구

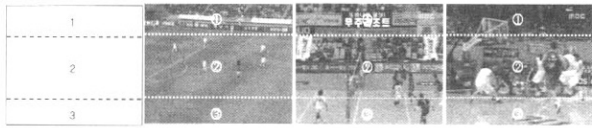
(그림 4) 스포츠 필드



(a) 키프레임 영상 (b) 미디언 필터링 (c) 16수준 양자화  
(그림 5) 영상처리 결과



(a) 축구 (b) 배구 (c) 농구  
(그림 8) 우세색 검출 결과



(그림 6) 영상의 영역 분할  
(a) 축구 (b) 배구 (c) 농구  
(그림 7) 필드의 영역 분할 예

스포츠 필드의 색상을 분석하기 위해서, 우선 영상 화소치의 분포도를 균일화하고 유사 색상간의 그룹화를 위해서 영상 처리 과정을 거친다. 영상 처리 과정을 살펴보면, (그림 5)와 같이 스포츠 장면에서의 키프레임 원영상에 대해서 5X5 미디언 필터링을 적용한 후, 이를 16수준 양자화 처리 과정을 거친다. 이러한 영상 처리 결과는 각각 (그림 5(b))와 (그림 5(c))로 유사 색상들이 그룹화된 것을 볼 수 있다.

키프레임의 영상 처리 과정을 거친 후, 스포츠별 필드를 식별하기 위해서 (그림 6)과 같영상을 3부분의 영역으로 분할하여 영역별 우세 색상(Dominant Color)의 분포도를 분석한다. 3영역의 분할 목적은 (그림 7)의 예와 같이 스포츠의 영상을 살펴보면, 1영역은 대부분 필드와 상관없는 관중이나 광고판 등의 스포츠 필드 색상을 식별하는데 불필요한 요소들을 많이 포함하고 있으며, 3영역은 자막으로 인해 필드 영역의 색상을 파악하기 어렵게 한다. 또한 (그림 7)의 예에서 축구나 골프인 경우는 (그림 7(a))와 같이 1,2,3영역에서 거의 동일한 색상이므로 우세색을 식별하기 쉽지만, (그림 7(b))와 (그림 7(c))의 경우는 1,2,3영역에 대해서 동일 색상의 분포가 아니므로 우세색을 검출하기 어렵다. 따라서 필드의 우세 색상을 검출하기 위해서, 추출된 키프레임의 영상에 대해 (그림 6)처럼 1,2,3영역을 종합적으로 분석한다.

스포츠별로 필드 색상을 구별하기 위해서 키프레임의 영상 처리 과정 후에 HSV상의 색상으로 변환하여 [표 1]과 같이 정의된 Hue값의 범위에 따라 1,2,3영역의 색상 분포가 임계치 이상인 경우를 스포츠 필드의 우세색으로 검출한다. 이와 같이 (그림 6)처럼 1,2,3영역별로 <표 1>에 정의된 Hue값의 범위에 따라 필드의 우세색 검출 결과가 (그림 8)이다.

<표 1> 필드 색상 분포

우세색 (Dominant Color)	초록색	노랑색 (주황색/홍색)	연분홍색
Hue 값의 범위	$65 \leq H \leq 110$	$40 \leq H \leq 55$	$20 \leq H \leq 35$

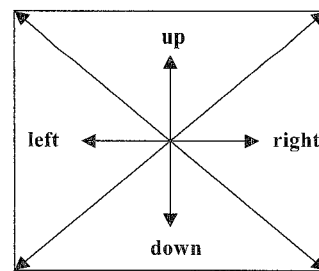
### 5.2 모션 방향성 분석

스포츠 비디오 영상에서 카메라나 객체의 움직임 방향은 스포츠를 구별하는데 유용한 특징 정보이다. 스포츠별 카메라의 모션 방향성을 살펴보면, 일반적으로 필드 색상이 비슷한 축구와 골프에서 축구는 한쪽 방향인 왼쪽이나 오른쪽 방향으로 움직임을 보이고, 골프는 거의 움직임 없다. 또한 배구의 경우는 네트를 중심으로 왼쪽과 오른쪽 방향의 움직임을 동시에 보이거나 공의 방향에 따라 왼쪽 또는 오른쪽 방향의 움직임을 보이고, 농구의 경우는 공의 방향에 따라 위쪽 방향성을 보이거나 불규칙하게 나타나고, 야구는 불규칙한 방향성을 가진다.

스포츠 영상에서 모션 방향성의 분석은 MPEG4 영상 압축 처리 과정에서 프레임의 매크로 블록에 대한 모션 벡터 값을 추출하고 모션 방향의 크기값을 산출하기 위해서 (식 3)과 같이 매크로 블록의 x축 방향과 y축 방향의 평균값을 구한다. 여기서 (식 3)의 M과 N은 각각 X, Y방향에 대한 모션 벡터(Motion Vectors)의 총 수이다.

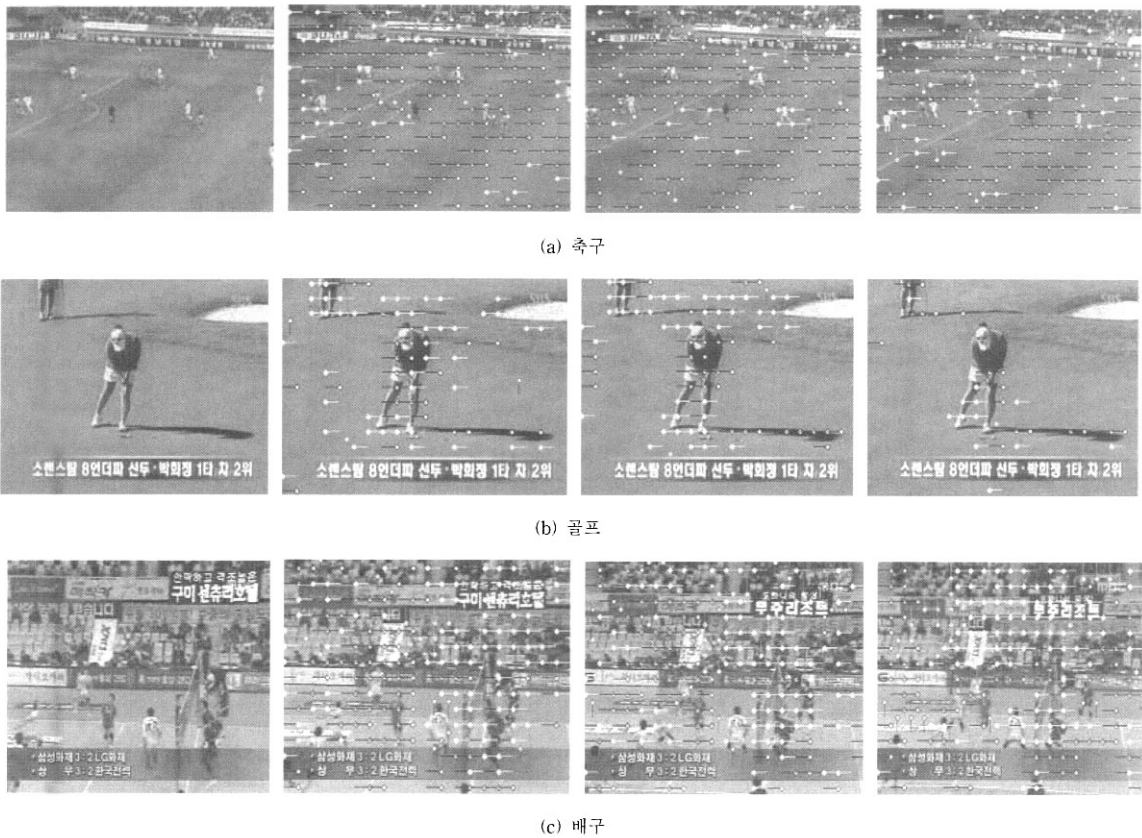
따라서 해당 프레임에 대한 모션 방향성은 모션 방향의 크기값에 따라 (그림 9)와 같이 4방향성으로 결정하고, 키프레임에 대한 최종 모션 방향성은 이웃한 연속 프레임들이 동일한 방향성을 가지고 있는지를 통계적 분석을 통해서 결정한다.

$$x_{ave}(i, j) = \frac{1}{M} \sum_{i=0}^{M-1} x_i \quad y_{ave}(i, j) = \frac{1}{N} \sum_{j=0}^{N-1} y_j \quad (\text{식 3})$$



(그림 9) 모션 방향

(그림 10)은 키프레임과 이웃한 연속 프레임에 대해서 모션의 방향을 추출한 결과를 표현한 것으로, (그림 10(a))는 공의 흐름에 따라 여러 프레임이 동일하게 오른쪽 방향을 보이고, (그림 10(b))는 거의 움직임이 없고, (그림 10(c))는 네트를 중심으로 양쪽 방향이 동시에 보인다.

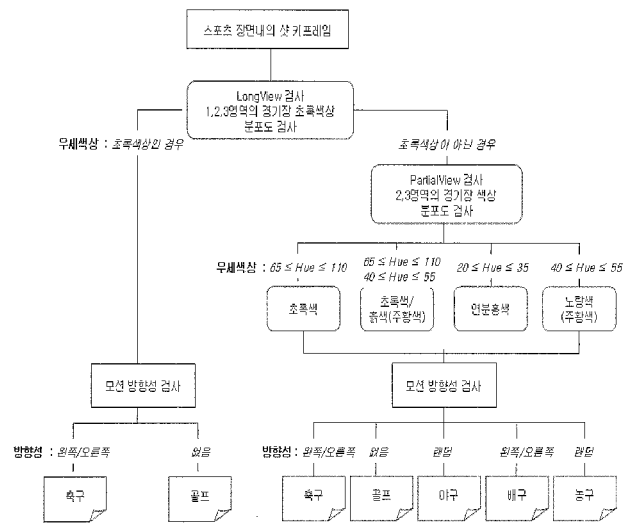


(그림 10) 모션 움직임 방향

### 6. 스포츠 장르 분류

스포츠 뉴스 비디오에서 추출된 스포츠 기사 장면의 장르는 <표 2>에 정의된 스포츠별 필드의 우세색, 모션 방향성의 특징 정보를 가지고 (그림 11)의 분류 과정을 통해서 결정한다.

우세색의 검출은 (그림 6)과 같이 3영역으로 나누어 전체 보기(LongView)와 부분 보기(PartialView)로 분석한다. 전체 보기(LongView) 분석이란 카메라가 경기장 필드를 최대한 포함할 수 있도록 먼 거리에서 영상을 포착하는 경우로 1,2,3영역에 대해서 모두 동일한 우세색을 나타내는 것을 말한다. 부분 보기(PartialView) 분석은 농구나 배구와 같이 필드의 색상이 전체적으로 나타나지 않는 경우에 대해서 필드 색상을 분석할 수 있도록 키프레임의 일부 영역인 2영역이나 3영역의 영역별 우세색을 검출하는 것이다.



(그림 11) 스포츠 장르 분류

<표 2> 스포츠별 특성

스포츠	우세색(Dominant Color)	모션방향
축구	초록색	좌 or 우
골프	초록색	변화 거의 없음
야구	초록색, 황색(주황색)	일정하기 않음
농구	노랑색(주황색)	일정하기 않음
배구	연분홍색	좌 and 우

스포츠 필드의 색상을 분석하기 위해서 전체 보기의 분석을 통해서 우세색이 초록색인 경우와 초록색이 아닌 경우로 분류한다. 전체 보기 분석에서 우세색이 초록색인 경우에 대해서 모션 방향성에 따라 축구와 골프로 분류한다. 전체 보기 분석에서 우세색으로 초록색이 아닌 경우는 부분 보기의 분석을 통해서 각 스포츠별 필드의 색상을 검출하고, 이에 대한 모션 방향성에 따라 스포츠 장르를 분류한다.

### 7. 실험 결과

실험 환경은 Pentium급 PC에서 Windows의 운영체제, Microsoft Visual C++ 6.0에서 이루어졌다.

입력 동영상은 축구·골프·야구·농구·배구 등을 기사 내용으로 포함한 스포츠 뉴스 비디오로 영상의 해상도는 320 X 240을 표준 크기로 한다. 이 연구에서 입력 동영상의 제약 조건으로는 반드시 스포츠 기사 장면사이에는 앵커 장면이

포함되어야 있어야 하고, 동영상의 시작 프레임은 앵커 프레임이어야 한다. 이는 처음의 앵커 프레임을 기반으로 스포츠 기사 장면간의 앵커 장면을 검출하기 때문이다.

앞의 4장에서 제안된 방법을 통해서 앵커 장면과 스포츠 장면을 검출한 결과가 (그림 12)이다. (그림 12)에서 스포츠 뉴스 비디오의 기사별로 검출된 결과로서 기사별로 앵커 장면의 시작 프레임과 끝 프레임의 결과와 스포츠 장면의 키 프레임들을 추출한 결과를 나타내고 있다.

기사 번호	아나운서 화면 시작 / 끝 프레임	스포츠 장면 추출된 키 프레임
No1	 00000      00060	
No2	 00344      00384	
No3	 00684      00728	

기사 번호	아나운서 화면 시작/끝 프레임	스포츠 장면 추출된 키 프레임
No1	 00000      00080	
No2	 00408      00460	
No3	 00740      00748	

(그림 12) 앵커와 스포츠 장면 검출

<표 3> 학습 데이터(Training Mode)

	축구	골프	야구	농구	배구	기타	장면수
축구	19	4			1	2	26
골프	5	3	3			1	12
야구	3		24				27
농구				15		1	16
배구				1	5		6
정확도	19	3	24	15	6		66/87(77%)

<표 4> 테스트 데이터(Test Mode)

	축구	골프	야구	농구	배구	기타	장면수
축구	50	4	2				56
골프	12	4	3				19
야구	3	4	25			1	33
농구			1	17	5	2	25
배구				3	18		21
정확도	50	4	25	17	18		114/154(74%)

본 실험은 스포츠 장르를 구별하기 위해서 <표 3>과 같이 스포츠 뉴스 비디오의 총 25개, 총 기사 장면수는 87개를 기본 학습 데이터(Training Mode)로 하여 스포츠별 필드 우세색상과 모션 방향성의 기준치를 찾고, 이를 바탕으로 <표 4>와 같이 테스트 데이터(Test Mode)는 45개의 스포츠 뉴스 비디오와 기사별 154개 장면수를 가지고 스포츠 장르의 분류 정확도를 측정하였다.

<표 3>과 <표 4>에서 스포츠별 장르를 분석한 결과를 살펴보면, 축구, 야구는 정확하게 검출되는 것을 볼 수 있고, 배구와 농구도 필드 색상과 모션의 방향성에 따라 정확하게 검출 결과를 얻었다. 골프인 경우는 우세색 검출을 바탕으로 방향성 검사에서 방향의 변화가 많이 나타나므로 정확도가 떨어짐을 알 수 있다.

이 연구에서 스포츠 뉴스 비디오의 전체 정확도는 학습 데이터와 테스트 데이터에 대한 전체 장면수는 241개이고 정확하게 분류한 개수는 180개로 75%의 정확도를 얻었다.

### 8. 결론 및 향후 과제

이 연구는 스포츠 뉴스 비디오의 내용에 대한 의미 검색을 위해 첫째, 앵커 장면과 스포츠 장면으로 분리(Segmentation)하고, 둘째, 선정된 스포츠의 주요 장면(Keyframe)의 필드 색상과 모션 방향의 특징 정보를 통해서 자동으로 스포츠 장르를 분류하여 의미 검색(Semantic Retrieval)이 가능하도록 하기 위한 방법을 제안하였다.

이는 하이라이트 중심인 5분이내의 짧은 스포츠 뉴스 비디오내에서 스포츠 장면을 검출하고 분석하여 스포츠 장르로 분류하는 것이다. 입력 동영상은 스포츠 뉴스 비디오를 대상으로 비디오 영상의 크기는 320 X 240이다. MPEG4 영상 압축시 영상의 크기는 256 X 128로 하여 DCT 계수치와 모션 벡터를 산출하여 키프레임을 추출하고, 이 키프레임에 대해서 영상 전처리를 거친 후 영상의 화소를 비교하여 앵커 장면과 스포츠 장면으로 분리하고, 스포츠 장면의 키프

레이들에 대해서 필드 색상과 카메라 모션 방향 분석을 통해서 축구·골프·야구·농구·배구 등의 스포츠 장르로 분류하였다. 이러한 제안된 방법을 통해서 스포츠 뉴스 비디오의 스포츠 장르 분류는 75%의 정확도를 얻었다.

앞으로 보다 정확도를 높이기 위해서는 영상이 가지고 있는 객체(Object)의 이동(Motion), 객체의 모양(Shape), 질감(Texture)등의 추가적인 특징 요소를 활용한 연구가 이루어져야 한다.

### 참고 문헌

- [1] J. Calic and E. Izquierdo, "Efficient Key-Frame Extraction and Video Analysis", Proc. IEEE Int. Conf. on Image Processing (ICIP2002), Sep. 2002.
- [2] Yueting Zhuang, Ruogui Xiao and Fei Wu, "Key Issues in Video Summarization and It's Application", Proc. IEEE Int. Conf. Information, Communications and Signal Processing, 2003.
- [3] By H. J. Zhang, C. Y. Low, S. W. Smoliar and J. H. Wu, "Video Parsing, Retrieval and Browsing : An Integrated and Content-Based Solution", International Multimedia Conference, 1995, ACM.
- [4] N. Babaguchi, Y. Kawai and T. Kitahashi, "Event Based Indexing of Broadcasted Sports Video by Intermodal Collaboration", IEEE Trans. on Multimedia, 4(1) pp.68-75, 2002.
- [5] L. Duan, M. Xu, T. S. Chua, Q. Tian and C. Xu, "A Mid-Level Representation Framework for Semantic sports Video Analysis", ACM Multimedia 2003, pp.33-44.
- [6] A. Ekin, A.M. Tekalp and R. Mehrotra, "Automatic Soccer Video Analysis and Summarization", IEEE Trans. on Image Processing, Vol. 12:7(2003), pp.796-807.
- [7] W.-N. Lie and S.-H. Shia. "Combining Caption and Visual Features for Semantic Event Classification of Baseball Video", ICME 2005.
- [8] S. Nepal, U. Srinivasan and G. Reynolds, "Automatic Detection of 'Goal' Segments in Basketball Videos", ACM MM01, pp.261-269, 2001.
- [9] Z. Xiong, R. Radhakrishnan, A. Divakaran and T. S. Huang, "Highlights Extraction from Sports Video Based on an Audio-Visual Marker Detection Framework", ICME 2005.
- [10] H. Xu, T.-H. Fong and T.-S. Chua, "Fusion of Multiple Asynchronous Information Sources for Event Detection in Team Sports Video", ICME 2005.



- [11] Sportvision Inc., "System for Enhancing a Video Presentation of a Live Event", U.S. Patent 6,597,406, January 26, 2001,
- [12] J. Assfalg, M. Bertini, C. Colombo and A. Del Bimbo, "Semantic Annotation of Sports Videos", IEEE Multimedia, vol.9, p.52-60, April 2002
- [13] Y. Rui, A. Gupta and A. Acero, "Automatically Extracting Highlights for TV Baseball Programs", ACM MM00, pp.105-115, Oct. 2000.
- [14] K. W. Wan, X. Yan and C. Xu, "Automatic Mobile Sports Highlights", ICME 2005
- [15] M. Xu, N. C. Maddage, C. Xu, M. Kankanhalli and Q. Tian, "Creating Audio Keywords for Event Detection in Soccer Video", ICME 2003, Vol II, 281-284.
- [16] T. D'Orazio, C. Guaragnella, M. Leo and A. Distanti, "A New Algorithm for Ball Recognition Using Circle Hough Transform and Neural Classifier, Pattern Recognition", vol. 37, pp.393-408, 2004.
- [17] D. Farin, J. Han and P.H.N. de With, "Fast Camera Calibration for the Analysis of Sport Sequence", ICME 2005.
- [18] D. Farin, S. Krabbe, W. Effelsberg and P.H.N. de With, "Robust Camera Calibration for Sport Videos Using Court Models", SPIE, vol. 5307, pp.80-91, 2004.
- [19] X. Yu, C. Xu, H. W. Leong, Q. Tian, Q. Tang and K. W. Wan, "Trajectory-Based Ball Detection and Tracking with Applications to Semantic Analysis of Broadcast Soccer Video", ACM MM03, pp.11-20, 2003.
- [20] X. Yu, H. W. Leong, J. H. Lim, Q. Tian and Z. Jiang, "Team Possession Analysis for Broadcast Soccer Video Based on Ball Trajectory", PCM 03, pp.1811-1815, 2003.
- [21] X. Yu, T. S. Hay, Xin Yan and E. Chng, "A Player-Possession Acquisition System for Broadcast Soccer Video", ICME 2005.
- [22] Y. Gong, T. S. Lim and H. C. Chua, H. J. Zhang, and M. Sakauchi, "Automatic Parsing of TV Soccer Programs", 2nd Int. C. on Multimedia Comp. and Sys., pp.167-174, 1995.
- [23] L. Xie, P. Xu and S. F. Chang, A. Divakaran, and H. Sun, "Structure Analysis of Soccer Video with Domain Knowledge and Hidden Markov Models", P.R. Letter, vol 25, pp.767-775, 2004.
- [24] T. Koyama, I. Kitahara and Y. Ohta, "Live 3D Video in Soccer Stadium", 2nd IEEE & ACM Int. Symp. on Mixed and Aug. Reality, pp.178-186, 7-10 Oct. 2003, Japan.
- [25] L. Duan, M. Xu and Q. Tian, "Semantic Shot Classification in Sports Video", SPIE: (SRMD) 03, pp.300-313, 2003.
- [26] I. Reid and A. Zisserman, "Goal-Directed Video Metrology", ECCV 96 (vol II), LNCS 1065, pp.647-658, 1996.
- [27] Wang, D-H. tain, Q. Gao, S. Sung and W.-K., "News Sports Video Shot Classification with Sports Play Field and Motion Features", ICIP'04 International Conference, Oct. 2004, Vol.4, pp.2247-2250.
- [28] X. Gibert, H. Li and D. Doermann, "Sports Video Classification using HMMs", Multimedia and Expo, 2003. ICME'03. Proceedings. 2003 International Conference on, Volume: 2, 6-9 July 2003, pp.345-348.
- [29] E. Sahouria and A. Zakhori, "Content Analysis of Video Using Principal Components", IEEE Trans. on Circuits and Systems for Video Technology, Vol. 9, No. 8, Dec 1999.
- [30] Hong Lu, Yap-Peng and Tan, "Content-Based Sports Video Analysis and Modeling", in Proceedings of 7th International Conference on Control, Automation, Robotics and Vision (ICARCV '02), pp.1198 - 1203.
- [31] Assfalg, J. Bertini, M. Colombo and C. Bimbo, "Semantic Annotation of Sports Videos", Multimedia, IEEE Publication, Apr-Jun 2002, Volume: 9, pp.52-60.
- [32] Assfalg, J. Bertini, M. Colombo, and C. Bimbo, Extracting Semantic Information from News and Sport Video", Image and Signal Processing and Analysis, 2001. ISPA 2001. pp.4-11, 2001.
- [33] Ekin, A. Tekalp and A.M., "Shot Type Classification by Dominant Color for Sports Video Segmentation and Summarization", Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03). April 2003, Volume 3, pp.173-176.



### 송 미 영

e-mail : songmy@swc.ac.kr

1995년 안양대학교 전자계산학과(학사)

1998년 동국대학교 컴퓨터공학과 대학원  
(공학석사)

2004년 동국대학교 컴퓨터공학과 대학원  
(공학박사)

2004년~현재 수원여대 멀티미디어통신과 전임강사

관심분야 : 컴퓨터 그래픽스, 게임, 컴퓨터 비전



장 상 현

e-mail : shjang@keris.or.kr  
1996년 동국대학교 컴퓨터공학과(학사)  
1998년 동국대학교 컴퓨터공학과 대학원  
(공학석사)  
2007년 동국대학교 컴퓨터공학과 대학원  
(공학박사)

1997년~1999년 한국학술진흥재단 부설 첨단학술정보센터  
연구원  
1999년~현재 한국교육학술정보원 책임연구원(기획혁신팀장)  
관심분야: 유비쿼터스/상황인식 기술, 이러닝 시스템, 교육게임



조 형 제

e-mail : chohj@dgu.edu  
1973년 부산대학교 전자공학과(학사)  
1975년 한국과학기술원 전기·전자공학과  
대학원(공학석사)  
1986년 한국과학기술원 전기·전자공학과  
대학원(공학박사)

1986년~현재 동국대학교 컴퓨터·멀티미디어공학과 교수  
관심분야: 멀티미디어 정보처리, 컴퓨터비전, 컴퓨터 그래픽스,  
게임