

# 스포츠 장르 분석을 위한 스포츠 뉴스 비디오의 의미적 장면 분류

송 미 영<sup>\*</sup>

## 요 약

앵커 장면 검출은 내용기반 뉴스 비디오 색인과 검색 시스템에서 비디오 장면의 의미적 파싱과 색인을 추출하는데 중요한 역할을 한다. 이 논문은 스포츠 뉴스의 단위 구조화를 위해서 뉴스 동영상에 존재하는 앵커 구간을 구분해내는 효율적인 알고리즘을 제안한다. 앵커 장면을 검출하기 위해서, 우선 MPEG4 압축 비디오에서 DCT 계수치와 모션 방향성 정보를 이용하여 앵커 후보 장면을 결정한다. 그리고 검출된 후보 앵커 장면으로부터 영상처리 방법을 활용하여 뉴스 비디오를 앵커 장면과 비앵커(스포츠) 장면으로 분류한다. 제안된 방법은 앵커 장면 검출 실험에서 평균적으로 98%의 정확도와 재현율을 얻었다.

## Semantic Scenes Classification of Sports News Video for Sports Genre Analysis

Mi-Young Song<sup>\*</sup>

## ABSTRACT

Anchor-person scene detection is of significance for video shot semantic parsing and indexing clues extraction in content-based news video indexing and retrieval system. This paper proposes an efficient algorithm extracting anchor ranges that exist in sports news video for unit structuring of sports news. To detect anchor person scenes, first, anchor person candidate scene is decided by DCT coefficients and motion vector information in the MPEG4 compressed video. Then, from the candidate anchor scenes, image processing method is utilized to classify the news video into anchor-person scenes and non-anchor(sports) scenes. The proposed scheme achieves a mean precision and recall of 98% in the anchor-person scenes detection experiment.

**Key words:** Scene Classification(장면분류), Anchor-Person Scene Detection(앵커장면검출), DCT Coefficients(DCT 계수), Motion Vector(모션 벡터), Image Processing(영상처리)

## 1. 서 론

최근의 압축 기술과 통신 기술의 급격한 발달에 따라 디지털 비디오 정보의 활용이 폭발적으로 증가하고 있으며 이를 효과적으로 관리하고 검색하기 위한 방법이 연구되고 있다.

특히 비디오 영상내의 내용을 기반으로 장면 분할을 수행한 후, 각 분할된 영역을 대표 프레임으로 표현함으로써 전체 비디오의 시퀀스를 효과적으로 이해(Understanding)·분석(Analysis)· 요약(Summarization)하고, 이를 바탕으로 색인 및 검색에 이용하는 기술이 활발히 대두되고 있다[1-6]. 비

※ 교신저자(Corresponding Author) : 송미영, 주소 : 경기도 수원시 권선구 수원여대길 62(441-748), 경기도 화성시 봉담읍 상기리 336-27(445-895), 전화 : 031)290-8957, FAX : 031)290-8945, E-mail : songmy@swc.ac.kr

접수일 : 2007년 1월 11일, 완료일 : 2007년 4월 4일

<sup>\*</sup> 수원여자대학 멀티미디어통신과

※ 본 연구는 2006년도 수원여자대학 연구과제 지원에 의해 수행되었음

디오의 색인은 영상의 내용에 따라 다른 접근을 하게 되는데, 특히 영화나 드라마, 또는 스포츠 영상과 달리 뉴스는 영상 자체가 시간 순으로 앵커 구간과 비앵커(기사자료) 구간을 포함하고 있는 구조적인 구성으로 되어 있다. 이러한 뉴스 비디오의 장면 단위의 구조화를 위해서 우선적으로 앵커 장면의 구간 검출이 요구된다.

앵커 장면의 검출은 주로 칼라 영상을 이용하여 얼굴 영역의 추출이 이루어진다. 컬러 영상은 밝기 정보 이외에도 색상 정보를 포함하고 있으므로 얼굴과 같이 고유한 색상을 지닌 영역을 분할하는데 유리하다. 이러한 점을 기반으로 얼굴의 피부색을 활용한 얼굴 영역을 추출하는 방법이 많이 이용되고 있다 [7-18].

[7]에서는 누적 히스토그램 기법을 활용하여 샷 경계를 검출하고 샷 경계에서 얼굴 영역들을 찾기 위해서 피부 컬러 정보를 이용하였다. 앵커 객체를 추출하기 위해서 사람의 상체 부분의 컬러 정보를 이용하여 앵커 후보 프레임을 검출하며, 후보 프레임들에 대해서 그래프 이론을 활용한 클러스터 분석 알고리즘을 적용하여 앵커 프레임과 비앵커 프레임을 분류하였다[8]. 이 방법은 객체 컬러와 배경 컬러 그리고 움직임 정보의 모든 조건이 앵커 프레임과 유사한 특징을 갖고 있는 프레임은 앵커 프레임으로 오검출된다. 또한 뉴스 비디오가 야외를 배경으로 한 경우 다양하게 변화는 배경으로 인해서 앵커 프레임을 검출할 수 없다.

[9]에서는 앵커 장면의 공통된 특징인 한 뉴스 프로그램을 진행하는 앵커는 동일하다는 점이고, 동일한 스튜디오라는 점 등의 지식을 기반으로 앵커 얼굴의 검출 방법과 옷 색상의 히스토그램 비교 방법을 활용하여 검출하는 방법을 제안하였다. 제안된 알고리즘을 통해서 여러 개의 KBS 9시 뉴스 비디오에 적용하여 실험한 결과로 정확도는 모두 96% 이상을 얻었다. 그러나 이 방법은 뉴스 기사내에서 앵커와 비슷한 옷의 색상이 나오는 경우나 처음 앵커의 특징을 추출하는데 있어서 특징이 잘 나타나지 않는 경우에 오검출이 발생된다.

[10]에서는 앵커 얼굴 위치의 프레임 변화량을 계산하는 FRFD(Face Region Frame Difference)를 활용하여 앵커 구간의 마지막을 찾았고, 비앵커 구간에서는 앵커 얼굴 추출하여 앵커 구간의 시작점을 찾았

다. 처리 속도를 빠르게 하기 위해서 MPEG 동영상의 복호화를 통한 대략적 분석과 전체 복호화를 통해 얻어진 앵커 얼굴 후보들을 비교 검증하여 앵커 구간을 검출하였다. 이 방법은 화면의 사람 얼굴이 앵커 얼굴의 제약 조건을 만족하거나 앵커 구간과 비슷한 구간인 경우에 오인식된다.

[11]에서는 앵커 장면을 배경 영역과 얼굴 영역으로 나누어서 배경 영역과 얼굴 영역에 대한 모델을 결정하고[12], 앵커 모델에 만족하는 장면을 앵커 장면으로 검출하였다. 이 방법은 고정된 배경 색상을 가진 뉴스에서는 좋은 성능을 보이지만 배경이 다른 뉴스이거나 배경이 바뀌는 뉴스에서는 만족스럽지 못한 결과를 얻는 단점이 있다.

[13]에서는 장면 전환점을 이용하여 앵커 장면을 검출하는 방법으로 뉴스 배경의 변화와 뉴스 캡션으로 인한 오류를 줄이기 위해서 각각 16개의 블록으로 분할한 후 동일 위치의 블록들 간의 유사도를 측정하여 가장 차이가 많은 8개의 블록을 제외한 나머지 8개의 블록들에 대해서만 유사도를 측정하고 유사도를 특정 임계치의 조건에 따라 MST로 구성하여 클러스터링된 결과에서 두 개 이상의 프레임이 존재하는 집합들에 대해서만 앵커 장면을 검출하였다. 그러나 이 방법은 조건에 의해서 앵커는 아니지만 앵커와 유사한 장면이 여러 번 나올 경우 오검출하는 문제점이 있다. 또한 앵커가 기사 중간에서 나타나는 장면이 아니고 단 한번 나오는 장면은 검출할 수 없다.

[14]에서는 컬러 히스토그램과  $x^2$  히스토그램을 합성한 방법을 활용하여 뉴스 비디오의 장면 전환의 키프레임을 추출하였고, 추출된 키프레임을 대상으로 앵커 프레임의 공간적 구성과 얼굴의 특징 정보에 대한 사전 지식을 바탕으로 한 유사성 측정을 통하여 앵커를 인식하였다. 이 방법은 앵커를 추출하기 위해서 평활화, 에지추출, 이진화 등의 전처리로 인해 중요한 정보를 잃을 수 있을 뿐만 아니라, 앵커에 대한 사전 지식을 바탕으로 앵커 프레임을 검출하므로 비디오 내에서 사람의 인터뷰 장면이나 기사 장면 등의 앵커 프레임과 유사한 경우 오인식률이 발생한다.

[15]에서는 뉴스 비디오에서 앵커 프레임의 효과적인 검색을 위해서 DC 성분 추출과 이진 검색 기법 그리고 컬러 클러스터링을 이용하였다. 제안된 방법을 검증하기 위해서 47분 10초 분량의 84,836개 프레임으로 구성하여 MPEG2로 압축된 뉴스 비디오에

이터에 적용한 결과는 91.3%의 정확도를 보였다. 이 방법은 각 샷에 대한 키프레임의 검출에 있어서 밝기에 민감한 반응을 보이며, 또한 기사 장면에서의 자막에서도 다른 샷으로 분류해내거나 앵커 장면과 인접한 기사 장면이 앵커 장면과 유사할 경우 키프레임을 정확히 분류해내지 못하는 문제점이 있다. 또한 앵커 프레임 중에서도 앵커의 오른쪽에 나오는 슈퍼포지션 캡션이 없거나 컬러가 지나치게 두드러지는 경우 앵커 프레임 그룹에서 제외되는 결과를 보인다.

[16]에서는 MPEG 압축 도메인에서 최소 복호화를 수행하여 I 프레임의 DC 이미지를 추출함으로써 실행시간을 단축하고[17], 밝기 정보와 색상 정보를 함께 고려한 변형된 히스토그램 비교법을 도입하여 장면 전환을 검출하고, 신경망 기법을 통해서 추출된 장면 전환에 대한 앵커 프레임과 비앵커 프레임을 식별하지만 시간적인 비용이 소요된다.

위의 논문들에서 앵커 장면의 검출은 주로 앵커의 사전 지식을 기반으로 공간적 위치나 색상을 이용한 얼굴 영역의 추출을 통해 이루어진다. 이러한 방법들은 앵커 프레임 장면과 유사하거나 장면내의 사람이 나타나는 경우 앵커 프레임으로 인식하여 앵커 검출의 정확도를 저하시킨다.

따라서 이 논문은 기존의 방법인 앵커의 얼굴 정보나 공간적 위치 등의 여러 기반 지식과는 무관하게 자동으로 뉴스 비디오에서 앵커 기사 장면과 비앵커 기사 장면 검출을 위한 방법을 제안한다. 최근에는 스포츠 뉴스 비디오를 대상으로 스포츠 경기내의 다양한 특징을 검색할 수 있도록 스포츠의 의미적 장르 분류나 요약(Summarization) 등에 대한 연구들이 활발히 진행되고 있으며, 이를 위해 선행되어야 할 기술이 앵커 장면과 스포츠 장면을 자동적으로 분할하는 방법이므로 이 논문에서는 MPEG4 영상 압축 과정에서 DCT 계수치, 모션 방향성 등의 특징 정보로 후보 앵커 장면을 검출한 후 영상 처리 방법을 통해서 앵커 기사 장면과 비앵커(스포츠) 기사 장면으로 분할한다.

이 논문의 구성을 살펴보면, 2장은 스포츠 뉴스 비디오에서 장면 전환 분류의 처리과정을 설명하고, 3장에서는 앵커와 스포츠 장면의 분리를 위한 후보 앵커 장면 추출 방법을 알아본다. 4장에서는 추출된 후보 앵커 장면에서 정확한 앵커 장면과 스포츠 장면을 검출하는 방법을 살펴보고, 5장에서는 제안한 연

구의 실험 및 성능 평가 결과를 확인하고, 마지막으로 6장에서는 결론을 맺는다.

## 2. 장면 분류의 처리 과정

이 논문은 스포츠 뉴스 비디오에서 각 스포츠 기사별 장르를 분류하기 위한 전처리 과정으로 그림 1과 같은 처리과정을 통해서 앵커 기사 장면과 비앵커 장면인 스포츠 기사 장면으로 분할한다.

스포츠 뉴스 비디오의 영상을 입력으로 받아 MPEG4 영상 압축 과정서 DCT 계수치와 모션 방향성의 특징 정보를 추출한다. 추출된 DCT 계수치와 모션 방향성을 분석하여 후보 앵커 장면의 구간을 검출한다. 그리고 각 후보 앵커 장면 구간의 경계 프레임과 참조 앵커 프레임에 대해서 영상 처리 과정을 거친 후 영상의 유사성을 통해서 앵커 장면의 구간을 결정하여 앵커 기사 장면과 비앵커(스포츠) 기사 장면으로 분류한다.

입력 영상은 스포츠 뉴스 비디오를 대상으로 영상의 크기는 320 × 240이다. 비디오 영상을 MPEG4 영상 압축 과정서 전체 크기 영상을 사용하는 것이 아니라, 그림 2와 같이 외각부분을 제외한 256 × 128의 크기로 한다. 이는 시간적인 비용을 줄일 수 있고,

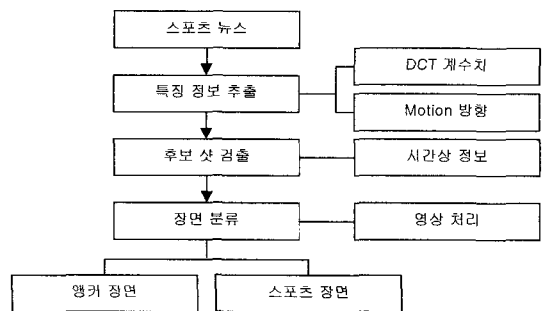


그림 1. 장면 분류의 처리 과정



그림 2. 입력 영상

영상의 상단 부분과 하단 부분의 방송사명이나 제목, 자막의 표시로 인해 동일한 장면임에도 불구하고 DCT 계수치나 모션 방향성의 급변으로 잘못된 장면이 검출되는 것을 막기 위해서이다.

### 3. 후보 앵커 장면 검출

후보 앵커 장면은 MPEG4 영상 압축 과정에서 DCT 계수치와 모션 벡터를 추출하고, 이를 분석하여 검출한다.

#### 3.1 DCT 계수 분석

스포츠 뉴스 비디오의 영상을 MPEG4 영상 압축 과정시 GOP는 IPPP로 하고, I 프레임의 매크로 블록의 수는 16 X 9개로 총 144개이다. 이 매크로 블록에 대해 Y, Cr, Cb의 DCT 계수치인  $DY$ ,  $DCb$ ,  $DCr$ 를 추출한다.

추출된  $DY$ ,  $DCb$ ,  $DCr$ 의 DCT 계수치를 이용하여 (1)과 같이 이웃하는 프레임간의 DCT 계수치의 평균을 산출한다.

$$\begin{aligned}
 DiffY_f &= \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (DY_{f+1}(x,y) - DY_f(x,y)) \\
 DiffCb_f &= \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (DCb_{f+1}(x,y) - DCb_f(x,y)) \\
 DiffCr_f &= \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (DCr_{f+1}(x,y) - DCr_f(x,y))
 \end{aligned} \tag{1}$$

여기서,  $M$ 과  $N$ 은 각각 매크로 블록  $x$ ,  $y$ 축의 총 개수이다.

(식1)에서 산출된 각 DCT 블록차를 합하여 (2)와 같이  $DCT_f$ 를 계산하고

$$DCT_f = |DiffY_f + DiffCb_f + DiffCr_f| \tag{2}$$

산출된  $DCT_f$ 에 대한 전체 평균값( $AVG$ )과 표준편차( $STV$ ) 값을 이용하여 (3)과 같이 임계치( $Th_{det}$ )를 통해서 장면의 전환 경계를 추출한다.

$$Th_{det} = AVG + \alpha * STV \tag{3}$$

여기서,  $\alpha$ 는 실험치에 의해 0.5이고, 후보 앵커 프레임은  $DCT_f > Th_{det}$  인 경우로 이를 만족하는 프

레이들의 연속 구간을 검출하여 DCT 계수치에 의한 후보 앵커 장면의 구간인  $AnchorDCT_k$ 을 결정한다.

#### 3.2 모션 방향성 분석

뉴스 비디오 영상에서 카메라나 객체의 움직임 방향은 유용한 특징 정보로 앵커 장면과 스포츠 기사 장면을 구별하는데 활용한다. 즉, 앵커 장면에서의 모션 방향성은 거의 변화가 없지만 스포츠 장면인 경우 모션 방향의 변화가 앵커 장면에 비해서 움직임의 변화가 많다.

스포츠 영상에서 모션 방향성의 분석은 MPEG4 영상 압축 처리 과정에서 프레임의 매크로 블록에 대한  $x$ 축 방향과  $y$ 축 방향의 모션 벡터값을 추출하여 모션 방향의 크기값( $dir$ )을 구한다. 각 매크로 블록의 모션 방향성은 산출된 모션 방향의 크기값( $dir$ )이 0이면 무방향으로, 그 외에는 그림 3과 같이 4개의 방향성으로 결정한다.

후보 앵커 프레임을 검출하기 위해서 (4)와 같이 산출된 매크로 블록들에 대한 모션 방향의 크기값( $dir$ )에서 무방향성인  $MV_d$ 를 산출한다.

$$MV_d = M - \sum_{i=0}^{M-1} D_i \tag{4}$$

$M$ : 매크로 블록의 총 수

$$D = \begin{cases} 1 & \text{if } dir = 4\text{방향} (Up, Down, Left, Right) \\ 0 & \text{otherwise} \end{cases}$$

그림 4는 키프레임과 이웃한 연속 프레임에 대해서 모션의 방향을 추출하여 4방향만을 표현한 것으로, 그림 4(a)와 그림 4(b)는 앵커 프레임으로 앵커의 모습에만 변화가 있을 뿐 배경영상에서의 변화는 거의 없다. 또한 그림 4(c)와 그림 4(d)는 스포츠 프레임으로 앵커 프레임에 비해 움직임의 변화가 많은 것을 볼 수 있다.

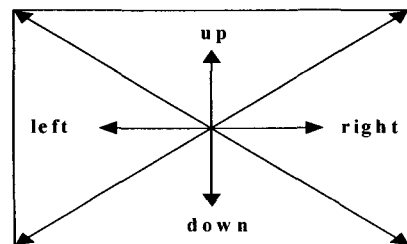


그림 3. 모션 방향

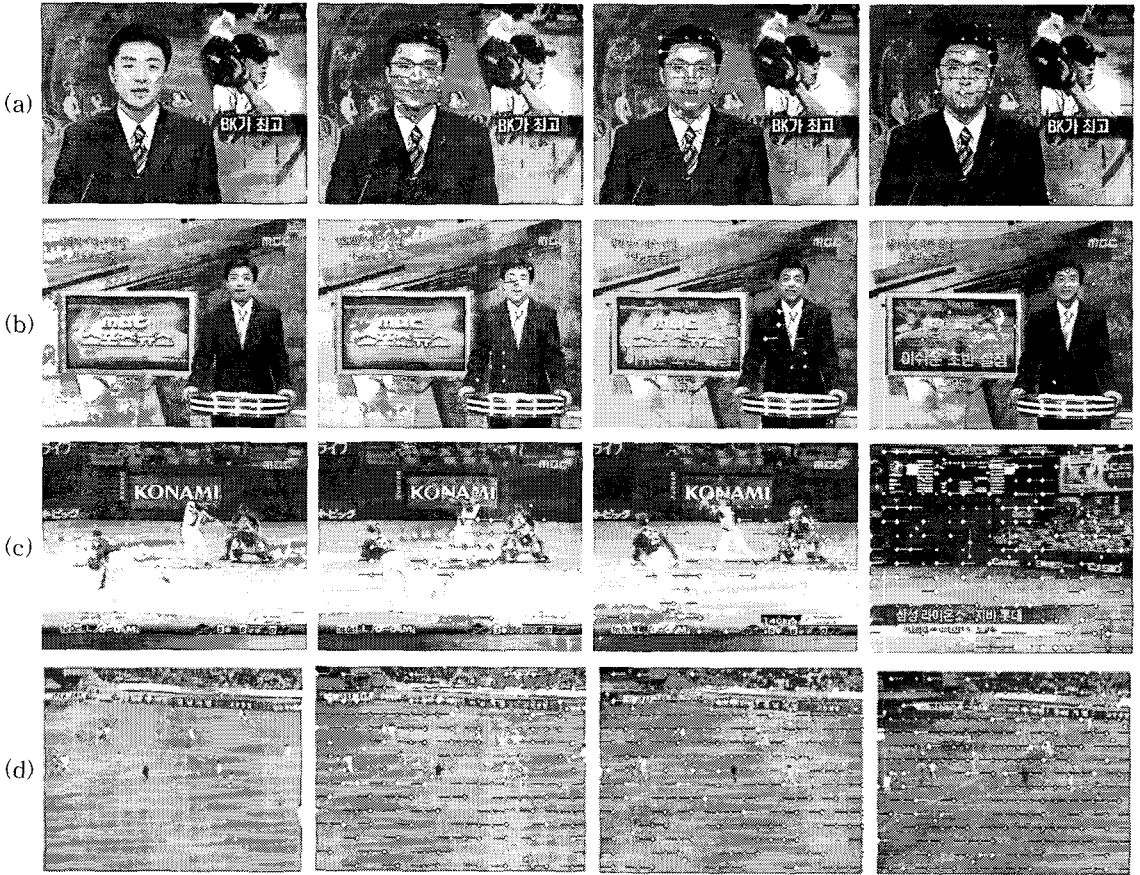


그림 4. 연속 프레임에 대한 모션의 방향성

산출된  $MV_d$ 에 대한 전체 평균값( $AVG$ )과 표준편차( $STV$ ) 값을 이용하여 (5)와 같이 임계치( $Th_{mv}$ )를 통해서 장면의 전환 경계를 추출한다.

$$Th_{mv} = AVG + \alpha * STV \quad (5)$$

여기서,  $\alpha$ 는 실험치에 의해 0.5이고, 후보 앵커 프레임은  $MV_d > Th_{mv}$ 인 경우로 이에 대한 연속 프레임들의 구간을 검출하여 모션 방향성에 따른 후보 앵커 장면의 구간인  $AnchorMV_d$ 을 결정한다.

최종 후보 앵커 장면 구간인  $AnchorScene_s$ 은 (6)과 같이 DCT 계수치를 이용하여 검출된 후보 앵커 장면 구간인  $AnchorDCT_k$ 과 모션 방향성에 의해 검출된 후보 앵커 장면 구간인  $AnchorMV_d$ 에 대해서 각 구간들을 통합하여 구한다. 또한 불필요한 후보 앵커 장면 구간들을 줄이기 위해서 후보 앵커 장면의

평균 구간 길이보다 짧은 구간은 제외시킨다. 이는 스포츠 뉴스가 경기내에서 하이라이트 부분만을 편집하므로 DCT 계수치나 모션 방향성이 급변하여 후보 앵커 장면으로 오검출되기 때문이다.

$$AnchorScene_s = AnchorDCT_k \text{ OR } AnchorMV_d \quad (6)$$

$$s = 1, 2 \dots l$$

$$k = 1, 2 \dots m$$

$$d = 1, 2 \dots n$$

여기서,  $m$ 은 DCT에서 추출된 후보 앵커 장면 구간의 총 수,  $n$ 은 모션 방향성에서 추출된 후보 앵커 장면 구간의 총 수,  $l$ 은 최종 추출된 후보 앵커 장면 구간의 총 개수이다.

#### 4. 앵커 장면 검출

앵커 장면 검출하기 위해서 그림 5와 같이 영상

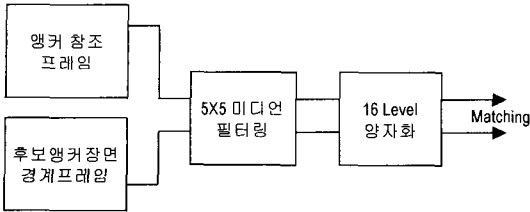


그림 5. 앵커 검출 과정

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 검출된 후보 장면의 경계 프레임 영상에 대해 5X5 미디언 필터링(Median Filtering)의 영상 처리를 과정을 거친다. 그리고 색상수의 분포를 줄이고 주변의 유사 색상들과의 그룹화를 위해서 16수준의 양자화(16 Level Quantization) 처리 과정이 이루어진다.

정확한 앵커 장면 구간을 검출하기 위해서 후보 앵커 장면 구간(Anchor Scene<sub>s</sub>)의 시작과 마지막 경

계 프레임을 기준으로 참조 프레임과 화소값을 비교한다. 여기서 참조 프레임은 뉴스 비디오의 시작 프레임은 항상 앵커로부터 시작한다는 가정하에 스포츠 뉴스 비디오의 첫 프레임을 말한다.

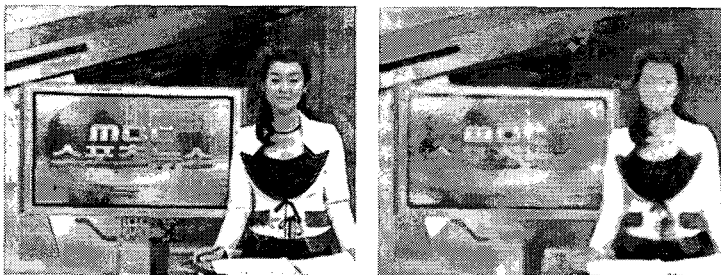
참조 프레임과 후보 앵커 장면의 프레임간의 화소값을 비교하기 위해서 그림 6과 같이 5X5 미디언 필터링(median filtering)과 16 레벨 양자화의 영상 처리 과정을 거친 후 후보 앵커 장면의 프레임과 참조 프레임은 (7)과 같이 화소차의 합이 임계치 ( $Th_{DP} = 0.5$ ) 이하인 경우의 프레임을 앵커 장면의 프레임으로 간주한다.

$$DP(k, ref) = \frac{\sum_{xy} (P_k(x, y) - P_{ref}(x, y))}{XY} \quad (7)$$

$P_k(x, y)$  : k프레임의 16레벨 양자화된 (x,y)좌표의 화소값

$P_{ref}(x, y)$ : 참조프레임의 16레벨 양자화된 (x,y)좌표의 화소값

정확한 앵커 장면을 검출하기 위해서 후보 앵커 장면 구간의 경계 프레임이 앵커 프레임이 아닌 경우



참조 프레임



(a) 원 영상 프레임

(b) 5X5 미디언 필터링

(c) 16level 양자화

그림 6. 앵커 프레임의 영상 처리 결과

는 후보 앵커 장면 구간내의 프레임들과 비교하여 앵커 프레임의 경계를 찾고, 또한 경계 프레임이 앵커 프레임과 동일하다 하더라도 경계 프레임의 바깥 쪽으로 비교 검색하여 앵커 장면 구간을 검출한다. 이와 같이 앵커 장면 구간을 검출함으로써 스포츠 장면의 구간이 자동으로 결정된다.

### 5. 실험 결과 및 분석

실험환경은 PentiumIV PC, Windows 2000의 운영체제, 구현언어는 Microsoft Visual C++ 6.0을 사용하여 행해졌다.

입력 동영상은 스포츠 기사 내용을 포함한 스포츠 뉴스 비디오로 영상의 해상도는 320 × 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.

본 실험은 앵커 기사 장면과 스포츠 기사 장면을 구별하기 위해서 스포츠 뉴스 비디오의 총 76개, 총 기사 장면수는 325개로 하여 (8)에서와 같이 재현율(Recall)과 정확도(Precision)의 척도법을 활용하여 성능을 평가하였다.

$$Recall = \frac{C_n}{C_n + M_n} \tag{8}$$

$$Precision = \frac{C_n}{C_n + F_n}$$

여기서,  $C_n$ 은 정확히 검출된 장면의 수를 나타내고,  $M_n$ 은 찾지 못한 장면의 수,  $F_n$ 은 잘못된 장면의 수를 나타낸다.

표 1은 제안된 방법의 성능 평가를 나타낸 결과로, 기존의 압축 동영상에서 활용되는 DCT 계수치와 모션 방향성만을 이용한 장면의 분류와 이에 영상처리

방법을 추가하여 통합한 장면의 분류를 비교한 결과치이다.

표 1(a)에서 DCT 계수치와 모션 방향성만을 통해서 장면 분류를 하지 못한 이유는 기존의 연구방법과 마찬가지로 기사 장면내에서 스포츠 장면이 아닌 인터뷰 장면이나 앵커와 유사한 사람이 나타나는 경우로 모션 벡터의 움직임이 적기 때문에 앵커 장면으로 검출되는 오류가 발생한다. 이러한 오검출을 해결하기 위해서 영상처리를 통해서 참조 앵커 프레임과의 유사성 비교하여 인터뷰 장면이나 기사 장면을 제거함으로써 재현율과 정확도가 모두 98%의 결과를 보였다.

제안된 방법을 통해서 앵커 장면 구간과 스포츠 장면 구간을 검출한 결과가 그림 7이다. 그림 7에서 스포츠 뉴스 비디오의 기사별로 검출된 결과로서 기사별로 앵커 장면 구간과 스포츠 장면 구간의 시작 프레임과 마지막 프레임을 나타낸다.

이 논문은 스포츠 뉴스 비디오에서 스포츠의 의미적 장르 분류나 요약을 위해서 선행되어야 하는 전처리 과정으로 기존의 방법인 앵커의 얼굴 정보나 공간적 위치 등의 여러 기반 지식과는 무관하게 DCT 계수치, 모션 방향성 그리고 영상처리를 통해서 앵커 기사 장면과 스포츠 기사 장면을 자동적으로 분할하는 방법을 제안하였다.

### 6. 결론 및 향후 과제

입력 동영상은 스포츠 뉴스 비디오를 대상으로 320 × 240의 표준 크기로 하였다. 그리고 시간적인 비용을 줄이고 영상의 상단 부분과 하단 부분의 방송 사명이나 제목, 자막의 표시로 인해 동일한 장면임에도 불구하고 DCT 계수치나 모션 방향성의 급변으로 잘못된 장면이 검출되는 것을 막기 위해서 MPEG4 영상 압축 과정시 영상의 크기는 256 × 128로 하여 DCT 계수치와 모션 벡터를 산출하였다. 이를 이용하여 후보 앵커 장면의 구간을 검출하였고, 후보 앵

표 1. 장면 분류 결과

방법	$n$	$C_n$	$M_n$	$F_n$	Recall	Precision
(a) DCT계수+MV	325	279	46	49	0.86	0.85
(b) DCT계수+MV+영상처리 (제안된 방법)		320	5	7	0.98	0.98



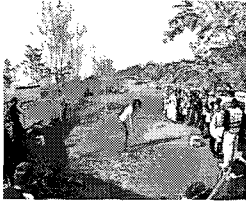











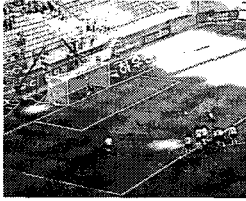





기사	앵커 장면		스포츠 장면	
	시작 프레임	마지막 프레임	시작 프레임	마지막 프레임
기사1				
	F0000	F0215	F0216	F1203
기사2				
	F1204	F1385	F1386	F2436
기사3				
	F2437	F2669	F2670	F3494
기사4				
	F3495	F3549	F3550	F3851
기사5				
	F3852	F4046	F4047	F4991

그림 7. 앵커와 스포츠 장면 분류 결과



커 장면에 대해서 영상 처리 방법을 활용하여 앵커 기사 장면을 검출하였다.

제안된 방법을 통해서 앵커 장면과 스포츠 장면의 분류 결과는 스포츠 기사 장면수로 325개를 포함한 76개의 스포츠 뉴스 비디오를 대상으로 98%의 정확도를 얻었다.

앞으로 분류된 스포츠 기사 장면을 통해서 스포츠의 장르를 분석할 수 있는 연구가 이루어져야 한다.

### 참 고 문 헌

- [1] Y. Alp Aslandogan and C.T. Yu, "Techniques And Systems For Image And Video Retrieval," *IEEE Transactions on Knowledge and Data Engineering*, Vol. 11, No. 1, pp. 56-63, 1999.
- [2] S. Khan and M. Shah, "Object Based Segmentation Of Video Using Color Motion And Spatial Information," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 2, pp. 746-751, 2001.
- [3] Z. Lei, W. Chou, J. Zhong, and Chin-Hui Lee, "Video Segmentation Using Spatial And Temporal Statistical Analysis Method," *IEEE International Conference on Multimedia and Expo*, Vol. 3, pp. 1527-1530, 2000.
- [4] J. Fan, A. K. Elmagarmid, X. Zhu, W. G. Aref, and L. Wu, "Hierarchical Video Shot Classification, Indexing, and Accessing," *IEEE Transactions on Multimedia*, Vol. 6, pp. 70-86, 2004.
- [5] X. Yang, Q. Sun, and Q. Tian, "Content-Based Video Identification : A Survey," *International Conference on Information Technology: Research and Education*, pp. 50-54, 2003.
- [6] Y. Zhuang, R. Xiao, and F. Wu, "Key Issues In Video Summarization And Its Application," *Information, Communications and Signal Processing and the Fourth Pacific Rim Conference on Multimedia*, Vol. 1, pp. 448-452, 2003.
- [7] 박기태, 황두선, 문영식, "앵커 객체 추출을 이용한 앵커 프레임 검출," *대한전자공학회 논문지*, pp. 17-24, 2006.
- [8] 박기태, 황두선, 이종철, 이석필, 문영식, "뉴스 동영상의 앵커 객체 추출을 이용한 앵커 장면 검출," *한국정보과학회 학술대회*, pp. 478-480, 2003.
- [9] 정진국, 이태연, 낭종호, 김경수, 하명환, 정병희, "뉴스 동영상 요약을 위한 앵커 장면 자동 추출 알고리즘," *한국정보과학회 학술대회*, pp. 274-276, 2001.
- [10] 전승철, 박성환, "뉴스 정보의 단위 구조화를 위한 효율적인 앵커구간 추출 알고리즘," *한국방 송공학회 논문지*, pp. 260-269, 2001.
- [11] W. Wang and W. Gao, "A Fast Anchor Shot Detection Algorithm On Compressed Video," *Proceedings of the Second IEEE Pacific Rim Conference on Multimedia: Advances in Multimedia Information Processing*, pp. 873-878, 2001.
- [12] A. Hanjalic, R.L. Lagendijk, and J. Biemond, "Template-Based Detection Of Anchorperson Shots In News Programs," *IEEE International Conference on Image Proceeding*, Vol. 3, pp. 148-152, 1998.
- [13] X. Gao and C. Tang, "Unsupervised Video Shot Segmentation And Model Free Anchorperson Detection For News Video Story Parsing," *IEEE Transactions on Circuits and Systems for Video Technology*, pp. 765-776, 2002.
- [14] 신성운, 임정훈, 이양원, 표성배, "뉴스 비디오에서 키프레임 추출과 앵커 인식," *한국멀티미디어학회 학술대회*, pp. 286-289, 2001.
- [15] 정정훈, 이근섭, 오화중, 최병욱, "MPEG으로 압축된 뉴스 데이터에서의 DC성분 추출과 컬러 클러스터링을 이용한 앵커 프레임 검색 기법," *대한전자공학회 학술대회*, pp. 729-732, 2000.
- [16] 정진국의 4, "MEPG 비디오 스트림의 칼라 히스토그램 정보와 매크로블록 정보를 이용한 새로운 샷 경계 검출 방법," *한국정보과학회 학술 발표*, Vol. 28, No. 1, pp. 418-420, 2001.
- [17] I. Koprinska and S. Carrato, "Detecting And Classifying Video Shot Boundaries In MPEG

Compressed Sequences,” *EUSIPCO Conf, Rhodes*, pp. 1729-1732, 1998.

- [18] 강현철, 이진성, 이완주, “뉴스 비디오에서의 효율적인 장면 전환과 앵커 화면 검출,” *한국정보과학회 논문지*, Vol. 32(12), pp. 1157-1163, 2005.



송 미 영

1995년 안양대학교 컴퓨터공학  
졸업(학사)

1998년 동국대학교 컴퓨터공학과  
졸업(공학석사)

2004년 동국대학교 컴퓨터공학과  
졸업(공학박사)

2004년~현재 수원여자대학  
멀티미디어통신과 전임강사

관심분야 : 게임, 컴퓨터 그래픽스, 컴퓨터 비전