

MPEG-2 압축 영역의 TV 스포츠 뉴스 색인을 위한 효율적인 장면전환 및 기사검출

김 성 국[†] · 박 영 규^{††} · 유 원 영[†] · 김 준 철^{†††} · 이 준 환^{††††}

요 약

본 논문에서는 MPEG-2로 압축된 TV 스포츠 뉴스 비디오의 색인을 위한 효율적인 뉴스 기사(Article)와 장면(Shot) 전환 검출 알고리즘을 제안한다. 제안된 알고리즘은 MPEG-2 비디오 스트림을 디코딩하지 않고 이들이 포함하고 있는 정보들만 사용하기 때문에 디코딩에 따른 시간을 절약할 수 있다. 특별히 장면 전환 검출에 있어서는 계층적인 검출 방법을 이용하기 때문에 시간을 보다 절약할 수 있으며, 압축도메인의 휘도 및 색차 성분을 동시에 이용하여 갑작스러운 조명변화에 대처할 수 있다.

또한 디졸브나 와이프 등의 특수효과에 의한 장면변화도 압축도메인에서 검출할 수 있도록 설계하였다. 기사검출에 있어서는 앵커 프레임의 CCV(Color Coherent Vector)의 개념을 이용하여 강건한 검출이 이루어 질 수 있도록 구성하였다.

Efficient Article and Scene Change Detections for TV Sports News Indexing in MPEG-2 Compressed-Domain

Sung-Guk Kim[†] · Young-Kyu Park^{††} · Won-Young Yoo[†] · Joon-Cheol Kim^{†††} · Joon-Whoan Lee^{††††}

ABSTRACT

In the paper, we propose efficient article and scene change detection algorithms to make the index of sports news compressed in MPEG-2 domain. In the proposed algorithm, the information in MPEG-2 compressed domain is directly used without decoding to save the computation time. The scene change detection algorithm is constructed in an hierarchical method so that the time for detection can be greatly reduced. Also, the algorithm can provide the robust detection against abrupt illuminance change because the luminance and chrominance components are simultaneously considered. Also, the scene change caused by special effect such as dissolve and wipe can be detected in the compressed domain. In the article detection, the algorithm is constructed for robust detection of the anchor frame using the concept of CCV(Color Coherent Vector).

1. 서 론

최근 디지털 비디오의 사용이 급격히 증가함에 따라

※ 본 연구는 과학재단 특정 기초 연구과제(97-01-00-01-01-3)에 의거 수행되었음.

† 준 회 원 : 전북대학교 대학원 전자공학과

†† 정 회 원 : 오성INC 연구소 연구원

††† 정 회 원 : 서남대학교 전기전자통신공학부 교수

†††† 정 회 원 : 전북대학교 전자공학과 교수

논문접수 : 1998년 12월 29일, 심사완료 : 1999년 4월 2일

비디오 색인에 의한 다양한 기능을 제공할 수 있는 멀티미디어 서비스 시스템의 개발이 이루어지고 있다. 이러한 멀티미디어 서비스를 위한 기반 기술중 하나로서 여러 가지 장면 전환(컷, 디졸브, 와이프)을 검출할 수 있는 알고리즘의 개발은 내용 기반 비디오 검색 시스템에 반드시 필요한 기술이다.

한편 디지털 비디오는 아주 많은 저장공간을 필요로

하기에 저장과 전송을 위해서는 이들 비디오 데이터의 압축은 필수적이다. 이러한 압축 표준에는 MPEG-1, MPEG-2, 또는 Motion JPEG 등이 보편화되어 있으며, 현재 MPEG-4와 객체 지향 부호화의 표준이 완성단계에 있다. 이들 중 MPEG-2 비디오 압축방법은 HDTV 등의 방송 표준화와 관련이 있기 때문에 MPEG-4가 보편화되기 이전까지는 당분간 멀티미디어 서비스 시스템 개발에 주종을 이룰 것으로 예상된다.[1,12,13]

이러한 MPEG이나 Motion JPEG 등으로 압축된 영상에서 장면 전환을 검출하기 위해 사용되는 알고리즘은 크게 두 가지로 나눌 수 있다. 한가지 방법은 처리 시간을 줄이기 위해서 압축된 상태의 데이터를 디코딩을 최소화하면서 장면 전환을 검출하는 방법과 다른 한가지는 압축된 영상 데이터를 디코딩 한 후에 장면 전환을 검출하는 방법이다.[2,3,4,5,6] 후자의 방법은 정확한 장면 전환 검출은 가능할지 모르지만 디코딩 과정(움직임 보상, IDCT 등)을 거치게 되므로 처리 속도가 느린 단점이 있다. 따라서 본 논문에서는 처리속도 향상을 위해서 전자의 방법을 이용하는 알고리즘을 제안한다.

이미 MPEG 기반으로 압축되어진 비디오 데이터에서 장면 전환을 검출하는 많은 알고리즘이 개발되었지만, 스포츠 뉴스와 같이 빠른 움직임을 영상내에 많이 가지고 있는 경우와 조명등의 변화가 많은 비디오에 있어서는 보다 정확한 장면 전환 지점을 검출하기 위해서는 개선된 알고리즘이 필요하다.

본 논문에서는 비교적 스포츠 뉴스와 같이 빠른 움직임을 영상내에 많이 가지고 있는 경우와 조명등의 변화가 많은 스포츠 뉴스와 같은 비디오를 대상으로 효율적인 장면 전환 검출 알고리즘을 제안한다. 본 논문에서 제안하는 장면전환 검출은 화면이 급작스럽게 변화하는 컷과 점진적으로 변화하는 디졸브 및 와이프의 검출을 포함한다. 이들 장면은 MPEG-2 비디오를 디코딩하지 않고 압축된 비디오 스트림이 가지고 있는 정보만을 이용하며, 계층적으로 컷 검출을 하고 있기 때문에 검출 속도를 개선할 수 있다.

또한 앵커프레임이 갖는 특성을 이용하여 압축도메인에서 강건하게 앵커프레임을 검출하는 알고리즘을 제안하며, 뉴스의 기사가 같은 특성 즉 앵커 프레임에서 시작하고 다음 앵커 프레임 이전까지가 기사인 특징을 이용하여 뉴스기사를 검출한다. 앵커 프레임의 검출에는 압축도메인의 DCT(Discrete Cosine Trans-

form) 계수와 CCV의 개념을 이용한다.

본 논문의 구성은 다음과 같다. 제2장에서는 제안된 계층적인 장면전환 알고리즘을 기술하였으며, 제3장에서는 기사검출을 위한 CCV를 이용하는 앵커프레임의 검출방법을 기술하며, 제4장에서는 실험결과를 제시하며, 제5장에서 결론을 맺는다.

2. 장면 전환 검출 알고리즘의 제안

장면 전환은 크게 두 가지로 분류할 수 있다. 컷(Cut)과 같이 갑작스럽게 장면이 바뀌는 경우와 장면 전환 지점 사이에 특수 편집 효과(디졸브, 와이프, 페이드 인/아웃 등)를 삽입하여 점진적으로 장면이 전환되는 경우가 있다. 본 논문에서는 컷과 디졸브 및 와이프등을 동시에 고려한 장면 전환 검출 알고리즘을 제안한다.

지금까지 MPEG으로 압축된 비디오 데이터에서 장면 전환을 검출하기 위해서 사용된 방법들을 살펴보면 압축도메인에서 디코딩 없이 장면전환을 검출하는 방법과 디코딩 후 공간 도메인에서의 프레임을 이용하는 방법으로 대별할 수 있다. 그러나, 장면 전환 검출시 속도 개선을 위해서 전자의 방법으로 나아가고 있는 추세이며, 이 방법은 압축 도메인에서 장면 전환을 검출하는 알고리즘들도 정확도 면에서는 조금도 뒤떨어지지 않는다고 보고되고 있다.[4,5,6] 첫 번째 방법은 I-프레임에서 DC 영상(DCT DC 계수로 이루어진 영상)을 구하고 P-프레임과 B-프레임에서는 움직임 보상을 통하여 DC 영상을 구한 후, 장면 전환을 바로 이 재구성된 DC 영상들을 서로 비교해서 검출하는 방법이다.[4] 이 방법은 IDCT 과정을 거치지 않고 모든 프레임의 정확한 장면 전환을 검출할 수 있다는 장점이 있으나 제안된 다른 방법에 비해 속도가 늦는다는 단점이 있다. 두 번째 방법은 P-프레임과 B-프레임의 MB(Macro Block)의 종류(Type)를 이용하는 방법으로 움직임 보상과 IDCT 과정을 거치지 않으므로 앞에서 제시된 방법보다 훨씬 더 많은 처리 시간을 단축하였다.[5] 하지만 이 방법은 GOP(Group of Picture)의 구조가 Closed-GOP 인 경우에는 I-프레임에서 장면 전환을 검출하지 못할뿐만 아니라 점진적인 장면 전환을 찾지 못하는 단점이 있다. 세 번째 방법은 I-프레임간의 에지(Edge) 개수의 변화량을 이용해서 장면전환 지점을 검출하는 방법으로 8×8 블록 사이의 DC 계수의

작분이 급격히 변화하면 블록 사이에 에서가 존재하는 것으로 가정했다.[6] 이 방법은 움직임 보상과 IDCT 과정을 거치지 않음으로 앞의 두 방법보다 처리시간이 가장 짧다는 장점이 있지만 객체의 움직임이 많거나 카메라 오퍼레이션이 많은 데이터에서는 오인식률이 높으며 I프레임만을 검사하기 때문에 정확한 장면전환 위치를 알수없다는 단점이 있다.

본 논문에서는 (그림 1)과 같은 계층적인 컷 검출 알고리즘을 제안한다. 일반적으로 MPEG-2의 비디오 시퀀스는 랜덤한 데이터의 접근을 용이하게 하기 위해서 GOP 구조를 가지고 있으며, 각 GOP는 반드시 하나의 I-프레임을 포함한다. 따라서 각 GOP의 I-프레임들만을 비교하여 이들 사이에 갑작스러운 장면전환이 있으면 이들 사이의 프레임에서 장면전환 검출을 시도하며, 그렇지 않으면 탐색할 필요가 없다. 일반적으로 1개의 GOP가 0.5초 정도의 프레임을 담고 있다면 한 장면이 0.5초 이내에 제한된 경우는 없기 때문에 이러한 방법에 의해 장면을 놓치는 일은 없다고 할 수 있

다. 이러한 계층적인 방법은 모든 프레임을 검사하지 않기 때문에 GOP당 1개의 I-프레임을 포함하고 한 개의 GOP에 15개의 프레임이 있다고 가정할 경우 약 15배의 처리속도를 개선할 수 있다.

2.1 신경망을 이용한 컷 가능성 검출

MPEG으로 압축된 비디오에서는 Y·Cb·Cr 컬러 좌표계를 사용한다. 여기서 Y는 밝기 성분을 나타내고 Cb·Cr은 색차성분을 나타낸다. 화소의 밝기차(Y 성분의 차)는 영상에서 물체의 큰 움직임이나 카메라의 팬(fan)과 틸트(tilt)가 일어나면 이것에 민감하여 아주 큰 값을 가지게 된다. 따라서 장면전환이 일어나지 않은 경우에도 장면 전환으로 오인식하게 되는 경우가 발생한다. 색차 상관도(Cb, Cr 성분)는 위에서 언급한 물체의 큰 움직임과 카메라의 동작에 비교적 둔감하기 때문에 화소의 밝기차(Y 성분)와 색차 상관도(Cb, Cr 성분)를 함께 사용하면 장면 전환으로 오인식되는 부분을 보완할 수 있다. 따라서, 제안된 알고리즘에서는 I-프레임들에서 컷에 의한 장면 전환은 먼저 I-프레임 간의 DC 영상의 밝기차와 색차 상관도를 이용한다. 그러나 스포츠 뉴스와 같이 빠른 움직임이 많거나 카메라 오퍼레이션(Panning, Tilting, Zoom In/Out 등)이 많은 경우에는 장면전환을 위한 임계값 설정에 많은 어려움이 따르게 된다. 이 문제를 해결하기 위해서 의사 신경망을 도입하였다. 신경망은 I-프레임간의 밝기차(ΔY)와 색차 상관도(ρ_n)를 입력받아서 장면 전환이 유무를 판별해서 출력하게 된다. 여기서 사용된 신경망의 훈련 데이터로는 스포츠 뉴스에서 장면 전환이 일어났을 경우와 장면 전환이 일어나지 않았을 경우의 밝기차(ΔY)와 색차 상관도(ρ_n)를 사용했다.

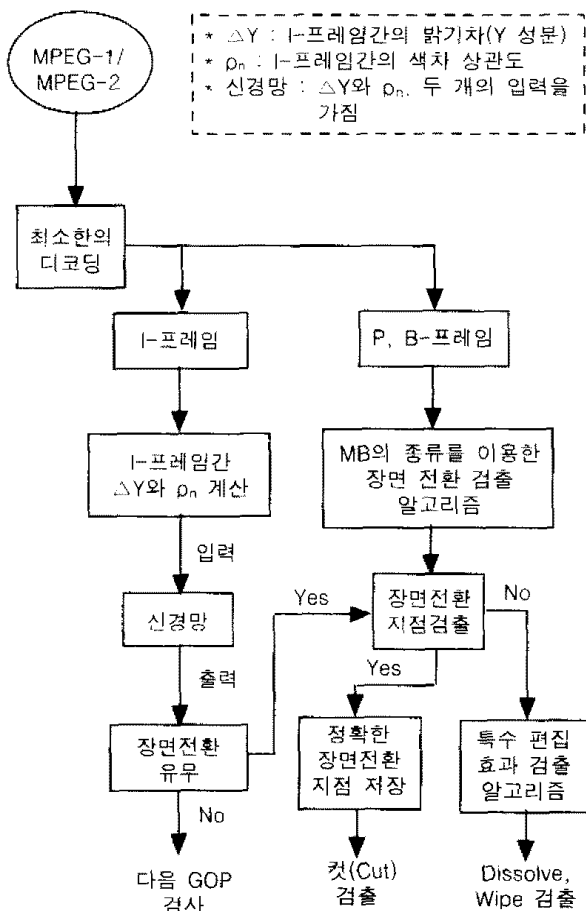
알고리즘에서 이용한 I-프레임간의 밝기차(ΔY)는

$$\Delta Y = \sum_{i=1}^M (Y_{n,i} - Y_{n-1,i}) \quad (1)$$

과 같고, 여기서 Y_n 과 Y_{n-1}은 현재와 이전의 I 프레임에서의 Y 성분의 DC영상의 히스토그램이다. 또한 색차 상관도(ρ_n)는

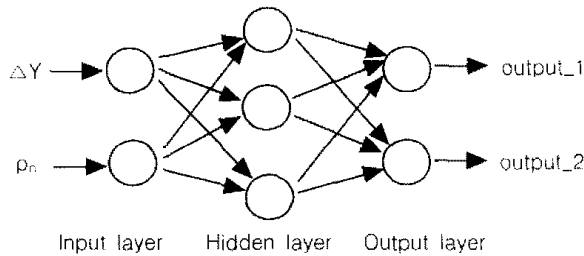
$$\rho_n = \frac{1 - (\sum_{k,l} H_{n,k,l} \times H_{n-1,k,l})^2}{(\sum_{k,l} H_{n,k,l}^2 \times \sum_{k,l} H_{n-1,k,l}^2)} \quad (2)$$

과 같이 정의되며 여기서 H_n과 H_{n-1}은 현재 I 프레임에서의 Cb, Cr 성분의 히스토그램을 나타낸다.



(그림 1) 제안한 계층적인 컷 검출 알고리즘의 흐름도

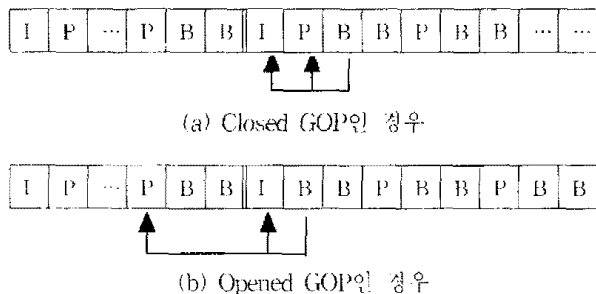
I-프레임간의 밝기차(ΔY)와 색차 상관도(ρ_n)를 이용한 장면전환 가능성 검출시 사용한 신경망은 (그림 2)와 같다. 사용된 신경망은 하나의 Hidden layer를 포함한 2개의 입력 노드와 2개의 출력 노드를 가지는 구조로 구성되어 있다. 신경망이 2개의 입력(밝기차(ΔY), 색차 상관도(ρ_n))을 받아서 Hidden layer를 거쳐 2개의 출력값을 내보내게 되면 이 값에 따라서 장면 전환 유무를 판단하게 된다. 신경망을 사용한 이유는 임계값 결정을 훈련에 의해 결정할 뿐 아니라 임계평면을 비선형적으로 설정함으로써 장면 전환 검출시 오인식을 피하기 위해서이다.



(그림 2) I-프레임 사이의 장면전환 가능성 검출을 위한 신경망 구조

2.2 컷 검출 알고리즘 제안

두 개의 GOP에 포함된 I-프레임 사이에서 컷 가능성이 검출된 후 장면 전환은 한 GOP내로 국한된다. 즉 한 GOP 내에서 I, P 또는 B-프레임에서 정확한 장면전환을 찾아내야 한다. 또한 GOP 구조는 GOP 내의 영상이 다른 GOP로부터의 독립 재생이 가능한 지의 여부에 따라서 Closed GOP와 Opened GOP로 (그림 3)과 같이 나누어지는데 이에 따른 고려도 필요하다.



(그림 3) GOP의 구조

2.2.1 컷 검출을 위한 속성

정확한 컷 검출을 위해서 <표 1>과 같은 MB의 종류를 이용하였다. I-프레임은 영상 전체가 MB_{intra} 로

인코딩이 되어 있어서 참조 영상 없이도 자체적으로 디코딩이 될 수 있지만, P-프레임과 B-프레임에서는 움직임 벡터를 이용하여 MB이 인코딩이 되어 있기 때문에 MB의 종류에 따라서 과거나 미래의 참조 영상이 있어야만 디코딩이 되어 진다. 장면이 전환되거나 영상내에 움직임이 많은 경우에는 과거의 영상으로부터 참조할 부분이 없어서므로 P-프레임의 경우 MB_{intra} 의 개수가 증가하게 되고, B-프레임에서는 과거의 영상을 참조하기 보다는 미래의 영상을 많이 참조하여 인코딩이 된다. 이러한 특징을 이용하여 P-프레임과 B-프레임에서 컷이 일어났는가를 검출할 수 있다.

<표 1> MB의 종류

사용된 MB	의 미
MB_{intra}	참조 영상 없이 MB 자체적으로 인코딩된 MB(움직임 벡터 없음)
$MB_{forward}$	과거의 영상을 참조하여 움직임 벡터를 이용하여 인코딩된 MB
$MB_{backward}$	미래의 영상을 참조하여 움직임 벡터를 이용하여 인코딩된 MB

GOP내의 I-프레임을 제외한 모든 프레임에서 장면 전환의 유무를 검사하는 방법은 다음과 같다. 먼저 GOP내의 P-프레임에서 R_p 를 구하고, B-프레임에서 R_b 를 구한다. 여기서 R_p 와 R_b 는 다음과 같이 정의하였다.

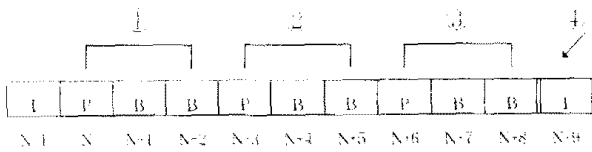
$$R_p = \# \text{ of } MB_{intra} / \# \text{ of } MB_{forward} \text{ in P-프레임} \quad (3)$$

$$R_b = \# \text{ of } MB_{backward} / \# \text{ of } MB_{forward} \text{ in B-프레임} \quad (4)$$

장면 전환이 있거나 움직임이 많은 경우에 P-프레임에서 R_p 값은 커지게 되고, B-프레임에서 R_b 값은 아주 큰 값을 가지는 특징을 이용하여 장면 전환의 유무를 판단할 수 있다.

2.2.2 Closed GOP의 경우 P-프레임 및 B-프레임에서의 장면전환 검출

MPEG으로 압축된 비디오 데이터에서 장면 전환 지점을 검출할 때에 유의해야 할 점은 프레임들이 화면에 출력되어지는 순서와 인코딩된 순서가 서로 다르다는 것이다. GOP 구조가 Closed GOP이고, P-프레임 사이에 B-프레임이 2개가 들어가 있는 경우로 (그림 4)에 나타나 있다.



(그림 4) MB의 종류를 이용한 장면 전환 검출 (Closed GOP인 경우)

제안된 알고리즘에서는 (그림 4)와 같은 Closed GOP인 경우 GOP내의 I-프레임을 제외한 ①, ②, ③에서 장면 전환 유무를 검사하고 장면 전환이 발견되지 않으면 다음 GOP를 검사하게 된다. GOP내의 ①, ②, ③에서 장면 전환 검출에 사용된 알고리즘은 아래와 같다.

```

< MB의 종류를 이용한 장면 전환 검출 알고리즘 >
IF (  $R_{n,n} > T_1$  ) {
    IF (  $R_{n,n+1} > T_2$  )
        N+1 번째(B 프레임)에서 장면전환 있음;
    ELSE IF (  $R_{n,n+2} > T_2$  )
        N+2 번째(B 프레임)에서 장면전환 있음;
    ELSE
        N 번째(P 프레임)에서 장면전환 있음;
}
ELSE
    장면 전환 없음;

N, N+1, N+2 : MPEG 2 Video Sequence에서 부호화되는 순서의 프레임 번호
T1, T2 : 장면 전환 검출에 사용된 임계값
    
```

즉 P-프레임에서 $MB_{n,n}$ 개수가 많더라도 계속되는 (실제로 디스플레이 순시상에서는 앞서 나오는) B 프레임에서 장면전환이 있을 수 있기 때문에 판단을 보류하고, 이들 프레임에서의 장면전환이 없을 때에만 P-프레임에서 장면전환이 있음을 판단하는 방식을 이용한다.

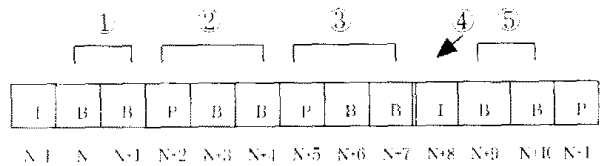
2.2.3 Closed GOP의 경우 I-프레임에서의 장면 전환 검출

(그림 4)와 같이 GOP 구조가 Closed GOP인 경우에는 GOP내에 있는 P, B-프레임들(①, ②, ③)에서 장면 전환 지점을 찾지 못하면 마지막으로 다음 GOP의 I-프레임(④)에서 장면전환이 일어났는지를 검사한다. 다음 GOP의 I-프레임에서 장면전환이 일어났는지는 이전 GOP의 P-프레임들에서 $MB_{n,n}$ 개수가 모두 설정된 임계값 이하이면 장면전환이 일어난 것으로 판단한

다. 제안된 것 검증 알고리즘은 I-프레임간에 컷이 삽입되어있을 가능성이 있는 경우에만 GOP내에 있는 P, B 프레임들을 검사한다. 따라서 I-프레임을 제외한(GOP내의 모든 프레임들을 검사한 결과, 장면전환 지점이 발견되지 않으면 다음 I-프레임에서 장면전환이 일어났을 확률이 아주 크다. 하지만 움직임이 영상내에 많은 경우나 특수편집효과가 삽입되어진 경우에도 위와 같은 현상이 일어나게 되므로 오인식을 줄이기 위해서 GOP내에 있는 P-프레임들의 $MB_{n,n}$ 개수가 모두 임계값 이하인지를 검사한 후, 모두 임계값 이하이면 I-프레임에서 장면전환이 일어난 것으로 판단을 하고 그렇지 않은 경우에는 특수편집효과 검출 알고리즘을 이용하게 된다.

2.2.4 Opened GOP의 경우 장면전환 검출

GOP 구조가 (그림 5)와 같이 Opened GOP인 경우는 앞에서 언급한 GOP 구조가 Closed GOP인 경우에 적용했던 것 검증 알고리즘을 그대로 적용할 수 없다.



(그림 5) MB의 종류를 이용한 장면 전환 검출 (Opened GOP인 경우)

GOP 구조가 (그림 5)와 같이 Opened GOP인 경우에도 I-프레임간에 컷이 삽입되어있는 가능성이 있는 경우에만 GOP 내의 P, B프레임들(②, ③)을 MB의 종류를 이용해서 장면전환 지점을 검출하게 되고, 장면 전환 지점이 검출된 경우에는 다음 GOP로 넘어가게 된다. 하지만 장면전환 지점이 검출되지 않은 경우에는 GOP내에 있는 P-프레임들의 $MB_{n,n}$ 개수가 모두 설정된 임계값 이하인지를 검사하게 되고, 설정된 임계값 이하이면 다음 GOP의 I-프레임 바로 다음에 위치한 B-프레임들(⑤)을 검사하게 된다. 하지만 그렇지 않은 경우에는 특수편집효과 검출 알고리즘을 이용해서 특수편집효과 삽입 유무를 판단하게 된다. 다음 GOP의 B-프레임들(N+9번째, N+10번째)에서 $MB_{backward}$ 가 많은 경우에는 I 프레임(N+8번째)을 참조해서 부호화된 경우이고, $MB_{forward}$ 가 많은 경우에는 이전 GOP의 P-프레임(N+5번째)을 참조해서 부호화된 경우이다.

Opened GOP 구조를 가진 경우에는 Closed GOP 구조를 가진 경우와는 다르게, I 프레임(④) 다음에 위치하는 B-프레임들(⑤)을 검사하는 이유는 B-프레임들이 시간상으로 I-프레임보다 과거의 영상이기 때문이다. 따라서 ⑤에 있는 B-프레임들의 매크로블록 종류를 이용해서 정확한 장면 전환지점을 찾는 방법은 아래와 같다.

< ⑤ 위치에서 MB의 종류를 이용한 장면 전환 검출 알고리즘 >

```

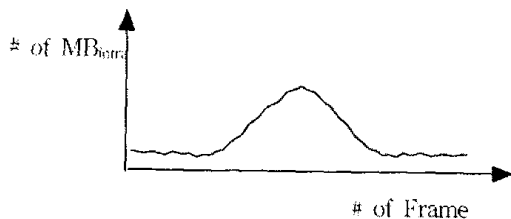
IF (  $R_{b,n-9} > T_1$  )
    N+9 번째(B 프레임)에서 장면전환 있음;
ELSE IF (  $R_{b,n-10} > T_1$  )
    N+10 번째(B 프레임)에서 장면전환 있음;
ELSE
    N+8 번째(I 프레임)에서 장면전환 있음;
N+8, N+9, N+10 : MPEG-2 Video Sequence에서 부호화된 순서의 프레임 번호
T1 : 장면 전환 검출에 사용된 임계값
    
```

2.3 특수 편집 효과 검출 알고리즘의 제안

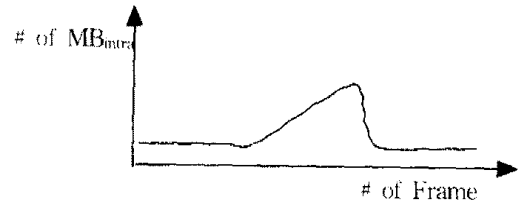
전 절에서 제시한 장면 전환 검출 알고리즘은 점진적인 장면변화(디졸브, 와이프, 페이드 인/아웃 등)를 검출하지 못한다. 따라서 이러한 특수 편집 효과가 삽입된 부분을 정확히 검출하기 위한 추가적인 알고리즘을 필요로 한다. 제안된 알고리즘에서는 이를 고려하여 많은 카메라 움직임이 포함되어진 경우에 장면 전환으로 오인식되는 점을 보완하였다.

2.3.1 디졸브 및 와이프 가능성 검출

특수 편집 효과(디졸브, 와이프, 페이드 인/아웃 등)의 지속 시간은 1초 정도이다. 따라서 2개의 GOP 내에서 P-프레임들의 인트라로 코딩된 매크로 블록의 개수의 변화량을 관찰함으로써 특수 편집 효과가 삽입되는지의 여부를 알 수 있다. 특수 편집 효과 중에서 디졸브나 와이프와 같은 경우는 P-프레임들에서 MB_{intra} 개수는 (그림 6)과 같은 변화를 나타낸다.



(a) Dissolve



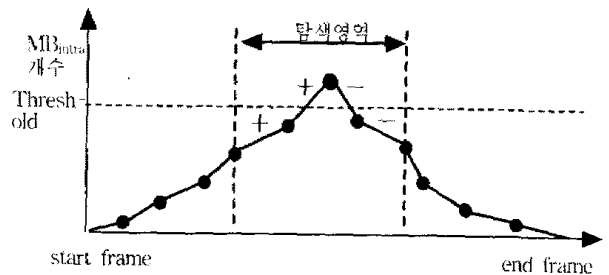
(b) Wipe

of MB_{intra} : P-프레임에서 인트라 부호화된 MB 개수
 # of Frame : 프레임 개수

(그림 6) 디졸브와 와이프의 특징

2개의 GOP내의 P-프레임들에서 나타난 MB_{intra} 개수의 변화량과 (그림 6)에 나타난 디졸브와 와이프의 특징곡선과의 매칭 방법이 (그림 7)에 나타나 있다.

(그림 6)에 나타나 있는 디졸브와 와이프의 특징곡선과의 비교는 (그림 7)에서 볼수 있듯이 특수편집효과가 나타나기 직전의 P-프레임(start frame)부터 시작해서 특수편집효과가 끝난 직후의 P-프레임(end frame)까지의 MB_{intra} 개수를 이용한다. (그림 7)의 탐색영역 내에서 MB_{intra} 개수가 설정된 임계값보다 크고 MB_{intra} 개수 변화량의 기울기가 급격히 증가했다가 급격히 감소하는 부분이 발견되면 이 부분에 디졸브나 와이프 삽입 가능성이 있다고 판단하고, 제안된 디졸브와 와이프 검출 알고리즘을 적용한다. 2개의 GOP내의 P-프레임들에서 나타난 MB_{intra} 개수의 변화량이 (그림 6)에 나타난 것과 다르다면 이 부분에는 카메라 오퍼레이션이나 상당히 큰 객체의 움직임이 있는 부분으로 장면전환 지점으로 인식하지 않고 다음 GOP를 검사하게 된다.



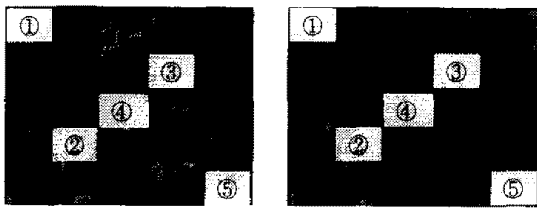
(그림 7) 특수편집효과 삽입 가능성 검출방법

2.3.2 디졸브 및 와이프의 분별

일단 디졸브 또는 와이프에 의한 장면전환이 있다고 판단되면 이러한 특수 편집 효과가 일어나기 전영상에서 부터 끝난 후의 영상들에 대한 DC 영상들을 구한

다. 구해진 DC 영상들에서 편집 효과가 일어나기 전 영상과 끝난 후의 영상을 이용하여 이것이 디졸브인지 와이프인지 아니면 단지 장면전환 없이 카메라의 움직임에 의한 것인지를 판단하게 된다.

이러한 판별은 디졸브와 와이프의 진행시 시간에 따른 영상의 변화 모습을 이용한다. 디졸브와 같은 경우에는 영상의 전반적인 변화가 일어나게 되고 디졸브가 시작되어 나타나는 영상들은 점진적으로 디졸브가 끝난 후의 영상에 가까워 지게 된다. 반면에 와이프의 경우에는 변화가 일부분에서 시작하여 영상전체로 퍼져 나가게 된다. 특수편집효과 중 디졸브의 검출은 DC 영상들 차이의 분산만으로도 구할 수 있지만 와이프와 같은 경우는 카메라의 오퍼레이션(Panning, Tilting)에 의해 나타나는 영상들과 구분이 힘들기 때문에 (그림 8)에서와 같이 영상의 부분에 레이블링을 하여 와이프 검출을 가능하게 하였다.



(그림 8) 특수 편집 효과 검출 방법

일반적으로 디졸브의 경우에는 디졸브가 시작되어 나타나는 영상들을 서로 비교했을 때 변화가 영상전체 (①번~⑤번 모두)에서 일어나게 된다. 하지만 디졸브는 장면이 급작스럽게 바뀌는 것이 아니기 때문에 디졸브가 시작되어 나타나는 영상들을 서로 비교했을 때 각 레이블링된 지역에서의 변화량은 설정된 임계값보다 작아야 하고 이런 현상이 나타나면 디졸브가 삼입된 부분으로 판단한다.

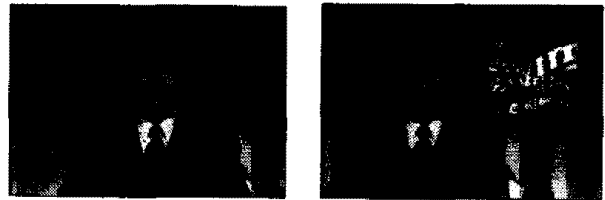
와이프의 경우에는 영상의 변화가 일부분(①번에서 ⑤번중 한 곳)에서 시작되어 영상 전체로 퍼져가게 되기 때문에 와이프가 시작되어 나타나는 영상들과 와이프가 시작되기 전 영상, 때로는 와이프가 끝난 후의 영상과의 비교를 통해서 와이프를 검출할수 있다. 와이프가 일어나기 전 영상과 와이프가 시작되어 나타나는 영상들을 서로 비교했을 때, 변화된 지역의 개수가 점차적으로 증가하면 와이프로 판단한다.

디졸브나 와이프가 아닌 경우에는 영상내에 카메라의 움직임(Tilting, Panning, Zoom in/out 등)이나 움직

임이 많이 포함된 경우이다. 세안된 장면전환 검출 알고리즘에서는 이러한 부분을 장면전환으로 인식하지 않음으로써 기존의 장면전환 검출방법에서 장면전환으로 오인식했던 부분을 보완하였다.

3. TV 뉴스 색인(Indexing)을 위한 뉴스 기사 검출

TV 뉴스는 여러개의 기사(Article)로 구성되고 하나의 기사는 여러개의 장면(Shot)으로 이루어져 있다. 따라서 TV 뉴스 색인을 위해서 가장 중요한 것은 기사의 시작 지점과 끝 지점을 정확하게 검출해 내는 것이다.^[1] TV 뉴스의 특성상 기사는 앵커가 나오는 영상(앵커 프레임)에서부터 기사가 시작되어서 내용이 나오고 다시 앵커가 나오면 다른 기사가 시작되어 진다. 이러한 특징을 이용하면 TV 뉴스를 이루고 있는 여러개의 기사를 정확하게 검출해 낼 수 있다. 다시 말하면 앵커 프레임을 정확하게 찾으면 기사의 시작 지점과 끝 지점을 찾아 낼 수가 있다.



(a) 아이콘이 없는 경우 (b) 아이콘이 있는 경우

(그림 9) 기사 검출에 쓰인 앵커 프레임의 종류

TV 스포츠 뉴스에서는 앵커가 대부분 한 명만이 나오고 앵커의 위치가 고정되어서 있고, 앵커의 조그만 움직임을 제외하면 영상내에 움직임이 없다고 할 수 있다. 이런 점을 이용해서 먼저 앵커의 얼굴이 나오는 부분을 지정하였고, 압축 효율을 향상시키기 위해서 사용된 움직임 벡터의 값도 아주 작은 점을 앵커 프레임 검출에 적용하였다. 앵커 프레임은 뉴스 아이콘을 가진 경우와 그렇지 않은 경우로 두 가지로 나누었고 뉴스 아이콘이 있는 경우의 앵커 얼굴의 위치와 뉴스 아이콘이 없는 경우의 앵커 얼굴의 위치를 지정하였다. 앵커 프레임의 검출은 장면 전환이 있는 I-프레임에서만 앵커 프레임 검출 알고리즘을 통하여 검사하였다. 앵커 프레임이 검출이 되어지면 영상내에 움직임의 유무를 판단하기 위해서 I-프레임 다음에 나오는

I-프레임을 검사하여 움직임이 없으면 최종적으로 앵커 프레임이라고 결정하였다.

제안된 알고리즘에서는 지정된 앵커 얼굴의 위치(1, 2)에서 얼굴색이 임계값 이상으로 나오면 일단 앵커 프레임일 가능성이 있지만 영상에 따라서 지정된 위치에 앵커가 아닌 다른 사람의 얼굴이 위치하게 되면 앵커 프레임으로 오인식하게 되는 경우가 생기게 된다. 이런 점을 보완하기 위해서 CCV를 사용하여 앵커 얼굴 부분외에 배경까지도 함께 고려함으로써 인터뷰 장면 등에서 앵커 프레임으로 오인식 될 수 있는 부분을 제거했다.

DCT로 압축되어진 영상에서 앵커 얼굴 부분과 배경부분의 비교를 위해서 사용한 텍스처(Texture)는 DCT의 에너지 분포를 사용하였고 다음과 같이 정의한다.[8,9,10,11]

$$h(i, j) = \frac{[DCT(i, j)]^2}{\sum_{i=0}^q \sum_{j=0}^q [DCT(i, j)]^2} \quad (5)$$

(그림 9)에 나타난 것처럼 앵커 프레임의 종류에 따라서 앵커의 얼굴이 있는 부분과 배경 부분으로 나누고, 각 부분에 포함된 8×8 블록들에서 식 (5)를 사용하여 DCT의 에너지 분포를 구해서 앵커 프레임간의 비교에 사용하였다. 8×8 블록들에 대하여 DC 계수와 AC 계수들의 DCT 에너지 분포를 먼저 구하고, 얼굴 부분과 배경부분에서 구한 각 블록들의 DCT 에너지 분포에 대해서 평균과 분산을 각각 구한 후에 검출되어질 앵커 프레임과 비교를 하는 텍스처로 사용하였다. MPEG 비디오가 DCT 기반으로 압축되어져 있기 때문에 앵커 프레임과의 유사도를 측정하기 위해서 DCT 에너지 분포를 사용함으로써 IDCT 과정을 거치지 않고서도 압축 영역에서 분별력이 뛰어난 특징 벡터의 추출이 가능하였다.

4. 실험 및 결과

특수 편집 효과(디졸브, 와이프 등)를 포함한 장면 전환 검출 실험을 위해서 MPEG-2로 인코딩된 704×480 해상도를 가진 11분 정도의 KBS 스포츠 뉴스를 대상으로 하였고, 하나의 GOP는 19개의 프레임으로 구성되어 있으며 GOP 구조는 Closed GOP인 비디오 데이터를 사용하였다. 먼저 실험 데이터에서 임의로

30개의 장면 전환과 30개의 장면 전환이 없는 부분의 I-프레임들을 선택하여 제안된 신경망으로 훈련을 시켜 장면전환 가능성이 있는 부분을 찾았는데 약 1000장의 I-프레임중 250장 정도의 장면 전환 가능성이 높은 부분을 찾았으며 여기엔 장면전환이 있는 모든 부분을 찾았으며, 찾은 부분중 장면 전환이 아닌 부분은 제안된 알고리즘을 통해서 보안하였다. KBS 스포츠 뉴스를 대상으로 제안된 알고리즘을 통해서 얻은 기사 및 장면 전환 검출 결과는 <표 2>에 나타나 있다. 실험 결과를 보면 기사는 100 % 정확하게 검출이 되었으며 컷 검출에 있어서도 97 % 이상의 정확한 결과를 보였다. 특수 편집 효과 검출에 있어서는 조금 낮은 인식률(디졸브 75%, 와이프 77.5%)을 보였는데 특수 편집 효과가 너무나 다양하기 때문에 제안된 특수 편집 효과 검출 알고리즘으로는 모든 편집 효과를 검출하지는 못했다.

실험은 SPARC-10 환경에서 실시하였는데 초당 약 30 프레임의 디코딩 및 디스플레이 시간 보다는 작게 걸리는 것을 확인할 수 있었다.

<표 2> KBS 스포츠 뉴스를 분석한 결과

(a) 기사(Article) 검출 결과

종 류	KBS 스포츠 뉴스	검출 결과	정확도(%)
기사(Article)	7	7	100

(b) 장면 전환 검출 결과

종 류	KBS 스포츠 뉴스	검출 결과	정확도(%)
컷(Cut)	185	180	97.2
디졸브(Dissolve)	8	6	75.0
와이프(Wipe)	9	7	77.7
합 계	202	192	95.0

5. 결 론

본 논문에서는 MPEG-2 기반으로 압축되어진 TV 스포츠 뉴스에 대하여 앵커의 얼굴색과 CCV를 사용하여 뉴스 기사(앵커프레임)를 검출하는 알고리즘을 제안하였다. 또한 I-프레임간의 DC 영상의 차이(밝기차, 색차상관도)속성을 이용한 신경망 분별기로 GOP사이에 장면전환이 있는지의 여부를 알아낸 후, GOP 내에

서 정확한 것을 검출하기 위해 I 프레임과 B 프레임에서의 MB의 종류를 이용하는 방법을 제안하였으며, 디졸브 또는 와이프등에 의한 장면전환도 검출하는 계층적인 장면전환 알고리즘을 제안하였다. 제안된 장면 전환 알고리즘에서는 모든 프레임 쌍에 대한 검출을 피하였기 때문에 장면전환의 검출속도를 개선할 수 있으며, MPEG-2의 비디오를 복호화하지 않은 상태에서 MPEG-2 비디오 시퀀스가 가지는 부가정보들을 최대한 이용하였다.

제안된 알고리즘은 일반적으로 시간에 따른 변화가 심한 KBS 스포츠 뉴스에 적용한 결과 실험 데이터에 대해 100%의 기사검출 능력을 나타내었고, 장면 전환 검출의 경우 97.5%, 특수효과(디졸브, 와이프) 검출의 경우는 76.25%의 검출 성능을 보였다. 속도에 있어서는 MPEG-2 비디오를 SPARC 10에서 실험시 소프트웨어 디코딩 및 디스플레이속도(약 30프레임/초)보다 빠른 속도의 검출을 실현하였다.

앞으로 앵커 프레임의 종류의 다양화 및 압축영역에서 추출할 수 있는 움직임 정보와 객체의 윤곽선 추출 및 카메라 오퍼레이션 추출등 그 밖의 다른 정보들의 첨가 및 알고리즘의 개선을 통해서 비디오 색인시스템을 구축할 예정이며, 특히 NoD(News on Demand)서비스에서 이용할 수 있는 DB의 내용을 추출함으로써 이를 이용한 검색 및 서버 구축에 많은 영향을 미치리라 기대된다.

참 고 문 헌

- [1] ISO/IEC 13818-2 : "Information Technology- Generic Coding of Moving Pictures and Associated Audio : Video," International Standard, March 1996.
- [2] Omer N. Gerek and Yucel Altunbasak, "Key Frame Selection from MPEG Video Data," SPIE Vol.3024, pp.920-925, 1997.
- [3] Jianhao Meng and Shih-Fu Chang, "Tools for Compressed Domain Video Indexing and Editing," SPIE Vol.2670, 1996.
- [4] Yasuyuki Nakajima, Kiyono Ujihara, Akio Yoneyama, "Universal Scene Change Detection on MPEG-Coded Data Domain," SPIE Vol.3024, pp.992-1003, 1997.
- [5] Jianhao Meng, Yujen Juan, Shin Fu Chang,

"Scene Change Detection in a MPEG Compressed Video Sequence," SPIE Vol.2419, 1996.

- [6] 나인엽, 우석훈, 원지선, 한희일, "MPEG-2 비트열에서 데이터 발생량 추정에 의한 장면전환 검출," 신호 처리합동학술대회, Vol.10, pp.69-72, 1997.
- [7] Y. Ariki and Y. Saito, "Extraction of TV News Articles Based on Scene Cut Detection Using DCT Clustering," ICIP 96, Vol.3, pp.847-850, 1996.
- [8] Hualu Wang, "Compressed-domain Image Search and Applications," Technical Report No.414, Center for Telecommunications Research, Columbia University, 1995.
- [9] Bo Shen and Ishwar K. Sethi, "Direct Feature Extraction from Compressed Images," SPIE Vol.2670, 1996.
- [10] Bo Tao and Bradley Dickinson, "Texture Classification on Block-Transformed Data," SPIE Vol.3024, pp.964-970, 1997.
- [11] Rob Reeves, Kurt Kubik and Wilfried Osberger, "Texture Characterization of Compressed Aerial Images Using DCT Coefficients," SPIE Vol.3022, pp.398-407, 1997.
- [12] Keith Jack, Video Demystified, HARRIS, 1996.
- [13] 정제창, 그림으로 보는 최신 MPEG, 교보문고, 1996.



김 성 국

e-mail : sgkim@ailab.chonbuk.ac.kr

1997년 전북대학교 전자공학과 졸업(학사)

1999년 전북대학교 대학원 전자공학과 졸업(공학석사)

관심분야 : 비디오 분석, 영상처리



박 영 규

e-mail : os3993@chollian.net

1996년 전북대학교 전자공학과 졸업(학사)

1999년 전북대학교 대학원 전자공학과 졸업(공학석사)

1999년~현재 오성INC 연구소 연구원

관심분야 : 영상처리 및 문자인식



유 원 영

e-mail : zero2@iceng.chonbuk.ac.kr
 1996년 전북대학교 전자공학과 졸업(학사)
 1998년 전북대학교 대학원 영상정보공학과 졸업(공학석사)
 1998년~현재 전북대학교 대학원 전자공학과 박사과정

관심분야 : 비디오 분석 및 편집, 영상처리



이 준 환

e-mail : chlee@moak.chonbuk.ac.kr
 1980년 한양대학교 전자공학과 졸업(학사)
 1982년 한국과학기술원 전자공학과 졸업(공학석사)
 1982년~1985년 전북대학교 전자공학과 조교

1985년~1987년 전북대학교 전자공학과 전임강사

1990년 미주리대 전산학과 졸업(공학박사)

1990년~1994년 전북대학교 전자공학과 조교수

1994년~현재 전북대학교 전자공학과 부교수

관심분야 : 영상처리 및 분석, 인공지능



김 준 철

e-mail : jkkin@tiger.seonam.ac.kr
 1986년 전북대학교 전자공학과 졸업(석사)
 1988년 전북대학교 대학원 전자공학과 졸업(공학석사)
 1988년~1991년 LG 산전 연구소 연구원

1991년~1995년 전북대학교 대학원 전자공학과 졸업(공학박사)

1993년~현재 서남대학교 전기전자통신공학부 조교수

관심분야 : 영상처리 및 컴퓨터비전