

# Multi-Pass 구조를 가지는 Shot 경계 검출기법

## Shot Boundary Detection Algorithm using Multi-Pass Mechanism

성창우, 강대성

Changwoo Seong and Dae-Seong Kang

### 요 약

본 논문에서는 MPEG으로 압축된 영상에 대해서 급격한 장면 변화에 의한 shot 경계(cut)와 점진적 장면 변화에 의한 shot 경계(dissolve)를 검출하여 비디오 shot을 분할하는 기법을 제안한다. 제안하는 shot 경계 검출 기법은 기존에 잘 검출하지 못하였던 점진적 장면 변화에 의한 shot의 경계 검출을 위한 알고리즘이다. 먼저 압축영역의 기법 중 DCT DC 값을 비교하는 방법을 이용하여 급격한 장면 변화에 의한 shot 경계를 검출한다. 그리고 움직임 벡터(MV)의 비를 비교하는 방법을 사용하여 점진적 장면 변화에 의한 shot 경계의 후보지들을 얻어내고, 선택된 후보지들 중  $n$ 번째와  $n+2$ 번째 후보지 영상으로 faded image를 만들어  $n+1$ 번째 후보지의 영상과 유사도를 비교하여 점진적 장면 변화에 의한 shot 경계를 검출한다. 이와 같이 압축영역의 기법으로 급격한 장면 변화에 의한 shot 경계와 점진적 장면 변화에 의한 shot 경계 후보지를 검출해 내고, 검출된 shot 경계 후보지들에서 점진적 장면 변화에 의한 shot 경계를 검출하는 방법을 함으로서 MPEG 비디오 영상의 복원량을 최소화하여 수행 속도를 높이면서도 급격한 장면 변화에 의한 shot 경계와 점진적 장면 변화에 의한 shot 경계, 두 가지 모두를 효과적으로 검출할 수 있었다.

### ABSTRACT

This paper describes an efficient algorithm for shot boundary detection in MPEG video stream. There are two types of shot boundary: abrupt and gradual. The proposed algorithm for detecting the abrupt shot boundaries used DCT DC value in compressed domain. The proposed algorithm of the gradual change detection consists of two-pass mechanism. In the first pass, the expected positions of shot boundaries are extracted using ratio value of motion vectors. After decoding frames that are extracted in the first pass, we will make the dissolving image using  $(n)$ th and  $(n+2)$ th image of expected position. The gradual shot boundaries are selected by similarity of the dissolving image and the image of  $(n+1)$ th expected position. As applying the algorithm for detecting shot boundaries, the gradual changes as well as the abrupt changes are detected efficiently. Experimental results indicate that the proposed method is computationally fast for detecting shot boundaries and robust to the variation of the video characteristic that is different for the kind of videos.

### I. 서 론

컴퓨터와 초고속 통신의 발달로 인해 비디오 영상 정보의 디지털화가 급속히 진행되고 있다. 그

리고 이들 디지털 비디오 정보가 증가함에 따라 이를 보다 효율적으로 이용하기 위해 내용기반 검색을 바탕으로 한 비디오 정보의 데이터베이스화에 대한 연구들이 진행되고 있다. 이들 연구 중 비디오 색인에 대한 연구는 비디오 정보 데이터베이스화를 위한 기초가 되는 작업이다. 비디오의 색인에 대한 연구는 크게 두 가지로 나누어

분 수 있다. 먼저 한 가지는 디지털 비디오의 분할이며, 다른 한 가지는 분할된 정보를 바탕으로 특징을 추출하는 것이다. 디지털 비디오 분할의 경우 형태론적 기본 단위 중 하나인 shot으로 분할하는 방법에 대한 연구가 많이 진행되고 있다. 일반적으로 shot의 경계는 크게 두 가지로 나누어진다. Cut에 의한 급격한 장면 변화(abrupt scene change)와 fade in / fade out / dissolve 등과 같은 영상효과에 의한 점진적 장면 변화(gradual scene change)가 있다. Shot 경계를 검출하는 방법은 압축되지 않은 비디오 영상에서 각 프레임들의 특징을 이용하여 shot을 분할하는 비압축 영역에서의 작업과 MPEG 등으로 압축된 영상을 복원하지 않고 압축된 stream 내의 여러 가지 parameter의 특징을 이용하거나, 일부 정보만을 복원하여 shot을 분할하는 압축 영역에서의 작업으로 크게 나누어 볼 수 있다. 일반적으로 비압축 영역에서의 작업은 원 영상을 그대로 사용함으로써 다양한 형태의 영상 효과를 이용하여 압축 영역의 작업에 비해 정확한 결과를 얻을 수 있으나 압축되어진 영상을 복원하여야 한다는 결점이 있고, 압축 영역의 경우 수행 속도는 빠르지만 비압축 영역에 비해 정보의 양이 한정되므로 보다 빠르고 정확하게 shot의 경계를 검출하기 위해서는 두 가지 영역의 기법을 병행하여 사용할 필요성이 있다. 본 연구에서는 압축 영역의 기법 중 DCT DC 계수를 이용하여 급격한 장면 변화를 검출하고, 움직임 벡터를 이용하여 점진적 장면 변화에 의한 shot 경계의 후보지를 선정한다. 선정된 후보지들에 대해 영상을 복원하여 복원된 후보지의 영상과 후보지의 영상에서 만들어진 faded image로 유사도를 비교하여 점진적 장면 변화에 의한 shot 경계를 검출하였다. 본 논문의 구성은 다음과 같다. 먼저 2장에서는 기존의 연구 동향에 대해 알아보고, 이들의 장단점과 특징에 대해서 기술할 것이다. 3장에서는 세안한 알고리즘에 대해 설명할 것이며, 4장에서는 실험 결과를 논할 것이며, 5장에서는 결론과 향후 과제들에 대해서 기술할 것이다.

## II. 기존 연구 동향

비디오 shot 분할 기법은 크게 비압축 영역에서의 기법과 압축 영역에서의 기법 두 가지로 나누어 질 수 있다. 일반적으로 비디오 영상을 저장

하기 위해서는 영상 압축을 행하므로 비압축 영역에서의 사용되는 기법들은 압축된 영상을 복원한 후 복원 영상을 이용하는 방법이며, pixel-by-pixel difference나 히스토그램에 의한 frame difference를 이용하는 방법이 있다[1][2]. 압축 영역에서 사용되는 기법들은 영상을 복원하지 않고 압축된 비디오 스트림의 파라미터들을 이용하여 shot 경계를 검출한다[3][4][5]. 대표적인 압축 영역의 기법들은 DCT 계수를 이용하는 방법과 움직임 벡터를 이용하는 방법을 들 수 있다. DCT 계수를 이용하는 방법의 경우 DC 계수를 이용하는 방법이 많이 이용되고 있는데 이는 비압축 영역의 frame difference를 이용하는 것과 유사한 성질을 가지고 있다. DCT DC 계수를 이용하여 shot 경계를 검출할 경우 일반적으로 카메라의 움직임이나 화면 내 물체들의 움직임 등에는 영향을 받지 않으나, 명암이나 색상의 변화에 민감하고 명암이나 색상 변화가 적은 shot의 변화나 dissolve와 같은 점진적 shot 경계를 검출하기 어렵다는 문제점을 가지고 있다. 움직임 벡터를 이용하는 방법에는 P와 B 프레임에서의 부호화된 매크로블럭(MB)과 움직임 벡터 비나 전방향 예측 움직임 벡터(FMV)와 역 방향 예측 움직임 벡터(BMV)의 비율을 이용하는 방법과 움직임 벡터의 움직임의 크기를 이용하는 방법이 있다[6]. 움직임 벡터를 이용하는 경우 색상의 변화나 명암의 변화에는 강하나 카메라의 움직임이 있다면, 물체의 움직임이 심한 경우 shot 경계를 잘못 인식하는 경우가 많이 있다. 점진적 장면의 변화가 일어나는 shot의 경계에 대한 연구는 비압축영역에서 영상의 edge count를 이용하는 방법이 있다[7][8]. 이 방법은 급격한 장면 변화에 의한 shot 경계를 찾은 다음 검출된 shot 경계들 사이에서 edge image를 얻어 edge image의 pixel 수의 변화를 이용하여 점진적 장면 변화에 의한 shot 경계를 검출한다. 이 밖에도 물체의 움직임을 추적하여 shot의 경계 검출을 시도하는 경우도 있다[9].

## III. 제안된 shot 경계검출 방법

본 논문에서 제안하는 Shot 경계를 검출하는 방법은 크게 두 개의 과정으로 나누어 볼 수 있다. 첫 번째 과정은 압축 영역의 기법인 DCT DC 계수와 움직임 벡터의 비율 이용하여 급격한

장면 변화에 의한 shot 경계와 점진적 장면 변화에 의한 shot 경계가 있는 후보를 검출하는 과정을 검출하는 것이고, 두 번째 과정은 fading 영상의 생성 원리를 이용하여 검출된 후보들 중 점진적 장면 변화에 의한 shot 경계를 검출하는 과정이다. 그림 1은 전체 알고리즘을 도식화한 것이다.

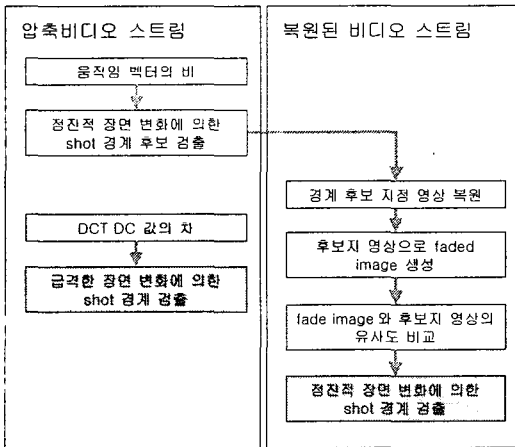


그림 1. 전체 shot 경계 검출 알고리즘

1. 압축 영역에서의 작업

MPEG[10]으로 압축된 영상에서 DCT 변환된 값들 중 DC 값은 DCT 변환된 8 x 8 영상의 평균을 나타낸다. 따라서 DCT 변환된 프레임에서 DC 값들로 영상을 만들고, 만들어진 영상의 히스토그램을 구하면 복원 영상의 히스토그램과 유사한 특성을 갖는다. 이러한 DCT 변환의 성질을 이용하여 영상을 복원하지 않고 DC 영상의 히스토그램 값의 변화만으로도 shot이 변화하는 지점을 검출할 수 있다. DC 영상의 히스토그램을 이용할 경우 장면이 서서히 변화하는 shot 경계는 검출하지 못하므로 이를 보완하기 위해 움직임 예측 벡터의 비를 이용한다.

MPEG으로 압축된 영상의 B 프레임에는 전방향 움직임 예측 벡터(FMV)와 역방향 움직임 예측 벡터(BMV)가 있다. 이들 움직임 예측 벡터들은 프레임 간의 유사도를 이용하여 이전 혹은 이후의 프레임으로 현재 프레임을 표현하는 것들로 현재 프레임에서 장면 변화가 일어나는 경우 프레임 간 유사도가 적어져 유사도가 적은 방향의

움직임 벡터가 급격히 줄어들게 된다. 이러한 움직임 벡터의 성질을 이용하여 shot 경계를 검출할 경우 화면의 색상이나 명암의 급격한 변화에는 강인하나 물체의 움직임이 크거나 한 경우 잘못된 shot 경계를 검출한다. 하지만 장면이 서서히 변화하는 shot 경계도 검출해 내는 장점이 있어 점진적 장면 변화에 의한 shot 경계 후보지를 선정하는데 사용하였다. 급격한 장면 변화에 의한 shot 경계 지점 검출과 점진적 장면 변화에 의한 shot 경계 후보지 검출 과정은 다음과 같다.

- Step 1. N-프레임에서의 DCT DC 값의 평균 ( $DC_N$ )을 구한다.
- Step 2.  $DC_N$ 과  $DC_{N-1}$  값의 차가 경계값 ( $Threshold_{DC}$ )보다 크면 급격한 장면 변화에 의한 shot 경계로 한다.
- Step 3. 전체 비디오 스트림에서 FMV와 BMV의 비를 구한다.
- Step 4. FMV와 BMV의 비의 구간 평균값을 계산할 window를 설정한다.
- Step 5. Window의 중앙에 위치하는 프레임의 FMV/BMV 비와 window 내의 FMV/BMV 비의 평균값을 비교한다.
- Step 6. 중앙 프레임의 FMV/BMV 비가 크면 점진적 장면 변화에 의한 shot 경계 후보로 설정한다.

2. 점진적 장면 변화에 의한 shot 경계의 검출

움직임 벡터의 비를 이용하여 검출한 shot 경계 후보지의 경우 물체의 움직임이 많은 장면이나 카메라의 panning이 일어나는 장면, 점진적 장면 변화에 의해 shot이 변화하는 경우를 모두 포함하고 있다. 이들 중 점진적 장면 변화에 의해 shot이 변화하는 지점을 찾기 위해서 영상 편집시 fade 효과를 생성하는 원리를 이용하였다. Fade 효과는 아래와 같은 수식으로 표현할 수 있다.

$$E(t) = O(t_S)\alpha(t) + I(t_E)[1 - \alpha(t)] + C, \quad (1)$$

$$\forall t \in (t_S, t_E)$$

여기서,  $E(t)$ 는 시간  $t$ 에서 화면에 나타나는 영상이며,  $O(t_S)$ 는 fade 효과가 시작되는 시간  $t_S$

의 비디오 스트림이며,  $I(t_E)$ 는 fade 효과가 끝나는 시간  $t_E$ 의 비디오 스트림이다.

이때  $O(t_S)$ 가 0인 경우를 fade in,  $I(t_E)$ 가 0인 경우를 fade out,  $O(t_S)$ 와  $I(t_E)$  모두 0이 아닌 경우를 dissolve라 한다.  $\alpha(t)$ 는 시간  $t_S$ 에서  $t_E$  사이의 구간에서 1에서 0의 값으로 단조 감소하는 일차원 함수이다.

식(1)과 같은 fade 효과의 특징을 이용하여 n번째 후보지의 영상을 out going 영상으로, n+2번째 후보지의 영상을 in coming 영상으로 하여 임의의 fade image를 만들어낼 수 있다. 이렇게 만들어진 fade image와 n+1 번째 후보지의 영상의 유사도를 비교하여 두 영상의 유사도가 높은 경우 n+1 번째 후보지는 점진적 장면 변화에 의한 shot이 변한 곳으로 검출해 낸다. 다음은 후보지들에서 점진적 장면 변화에 의한 shot 경계를 검출 과정을 나타낸 것이다.

- Step 1. 점진적 장면 변화에 의한 shot 경계 후보지를 복원한다.
- Step 2. 수식 (1)을 이용하여 n번째 후보지와 n+2번째 후보지로부터 fade image를 만든다.  
이때  $\alpha(t)$ 는 0.9에서 0.1까지 0.1 step으로 하여 총 9개의 fade image를 만든다.
- Step 3. Step 2에서 만들어진 fade image와 n+1 번째 후보지 영상의 유사도를 비교한다.
- Step 4. 유사도가 높으면 점진적 장면 변화로 검출한다.
- Step 5.  $\alpha(t)$ 는 0에서 1까지 0.1씩 증가시켜가면서 Step 2에서 Step 4까지의 과정을 반복한다.
- Step 6. 전체 후보지들에 대해 Step 1에서 Step 5까지의 과정을 반복한다.

#### IV. 실험 및 결과

본 논문에서 제안한 알고리즘의 성능을 검사하기 위해 MPEG 스트림에서 직접 shot 경계의 위치를 찾아내었다. 그림 2는 실험에 사용한 MPEG 스트림에서 찾아낸 전체 shot 경계 지점을 나타낸 것이다. 그래프에서 높게 보이는 부분은 급격한 장면 변화에 의한 shot 경계이고, 낮게 보이는 부분은 점진적 장면 변화에 의한 shot

경계이다. 그림 3은 전체 MPEG 스트림에서 FMV와 BMV의 비를 구하여 그래프로 그린 것이다.

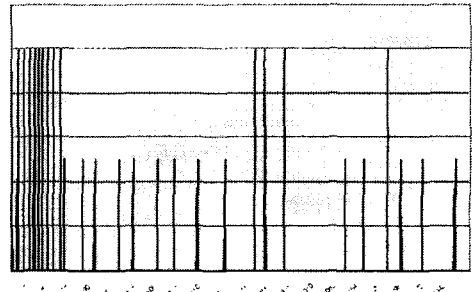


그림 2. 실험에 사용된 MPEG 영상의 shot 경계 위치

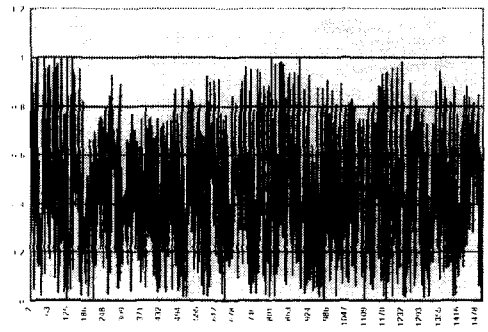


그림 3. 실험에 사용된 영상의 FMV/BMV 비

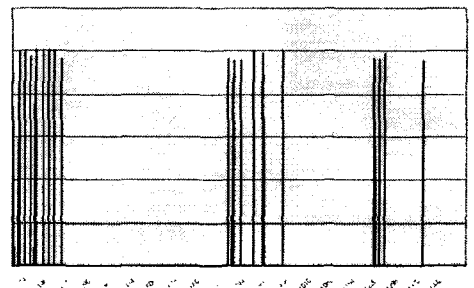


그림 4. 급격한 장면 변화에 의한 Shot 경계

그림 4는 DCT DC 값을 이용하여 급격한 장면 변화에 의한 shot 경계 지점을 검출한 결과를 보여주고 있으며, 그림 5는 움직임 벡터의 비를 이용하여 검출한 shot 경계 후보지들을 보여주고 있다.

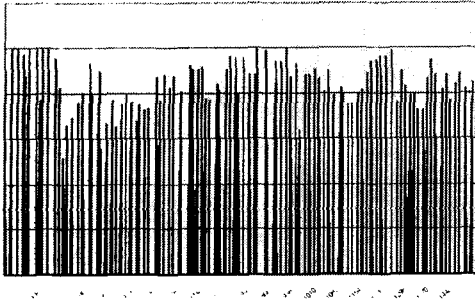


그림 5. 점진적 장면 변화에 의한 shot 경계 후보

이렇게 shot 경계 후보지와 cut이 검출되어지면 shot 경계 후보지들의 영상을 복원한다. 그리고 복원된 shot 경계 후보지들의 영상(n번째 후보지와 n+2번째 후보지)으로 fade image를 만들어 낸다. 그림 6은 fade image를 만드는데 사용한 후보지 영상의 예를 보여주고 있다. Fade image가 만들어지면 n+1번째 후보지의 영상을 유사도를 비교하여 점진적 장면 변화에 의한 shot 경계 지점을 검출한다. 그림 7의 왼편 그림은 그림 6의 후보지 영상들로부터 만들어진 fade image의 예이고, 오른편 그림은 n+1번째 후보지 영상의 예를 보여주고 있다. 그림 6, 7과 같은 결과가 나오면 n+1번째 후보지는 점진적 장면 변화에 의한 shot 경계 지점이 된다. 제안한 알고리즘을 이용하여 shot 경계를 검출한 최종 결과는 표 1과 같다.



그림 6. n번째와 n+2 번째 후보지 영상



그림 7. dissolve 영상과 n+1번째 후보지 영상

표 1. 제안한 알고리즘을 이용하여 Shot 경계를 검출한 결과

검출항목	전체 shot 경계	급격한 장면 변화 검출 후	후보지 검출 후	최종 결과
후보지	X	X	97	X
급격한 장면 변화	13	16	X	16
점진적 장면 변화	11	X	X	9
오검출	X	3	73	3
미검출	X	0	0	2

실험에 사용된 MPEG 스트림은 1700개의 프레임으로 구성이 되어 있고 13개의 급격한 장면 변화에 의한 shot 경계와 11개의 점진적 장면 변화에 의한 shot 경계를 가지고 있다. FMV와 BMV 비율로서 shot 경계 후보지를 검출한 결과를 보면 물체의 움직임이 심하거나 카메라의 움직임이 있는 경우 오인식한 것은 많이 있으나, 급격한 장면 변화에 의한 shot 경계와 점진적 장면 변화에 의한 shot 경계를 다 포함하고 있으므로 shot 경계 후보지 선정에 적합하다. 급격한 장면 변화에 의한 shot 경계의 검출 결과 중 오인식된 것들은 fade in/fade out 시간이 짧은 dissolve를 급격한 장면 변화에 의한 shot 경계로 인식한 것과 카메라의 panning을 서로 다른 shot 경계로 인식한 경우가 있었다.

### V. 결 론

본 논문에서는 압축 영역의 기법 중 DCT DC를 이용하여 급격한 장면 변화가 일어난 곳을 검출 하였으며, 움직임 벡터의 비율을 이용하여 점진적 장면 변화에 의한 shot 경계 후보지를 검출하였다. 그리고, 검출된 경계 후보지들에 대해 fade 효과의 특징을 이용하여 점진적 장면 변화에 의한 shot 경계 지점을 검출하는 방법을 제안하였다. 제안한 알고리즘의 특징은 압축된 비디오 stream의 복원량을 최소화하면서도 급격한 장면의 변화에 의한 shot 경계뿐만 아니라 점진적 장면 변화에 의한 shot 경계로 효과적으로 검출할 수 있는 것이다. 앞으로의 과제는 오인식율을 줄이고 카메라의 움직임에 의한 영상의 변화를 감

지할 수 있는 방법에 대한 연구와 함께 보다 다양한 종류의 비디오 영상에 대해 shot 경계 검출 성능을 시험하여 비디오 영상 특징에 영향을 받지 않는 shot 경계 검출 기법을 연구하여야 할 것이다.

접수일자 : 2000. 9. 9. 수정완료 : 2000. 10. 25  
 본 논문은 정보통신부의 2000년도 대학기초연구 지원사업의 지원으로 수행되었음.

### 참고문헌

[1] Y. Tonomura, K. Otsuji, A. Akutsu, and Y. Ohba, "Stored video handling techniques", NTT Review, Vol. 5, pp. 82-90, 1993.

[2] S. W. Smith and H. J. Zhang, "Content-Based Video Indexing and Retrieval", IEEE Multimedia, summer 1994, pp 62-72, 1994.

[3] Boon-Lock Yeo and Bed Liu, "Rapid Scene Analysis on Compressed Video", IEEE Transaction on Circuits and Systems for Video Technology, Vol. 5, No. 6, pp 533-544, 1995.

[4] Shih-Fu Chang, "Compressed-Domain Techniques for Image/Video Indexing and Manipulation", IEEE ICIP-95, Vol. 1, pp 314-317, 1995.

[5] Jianhao Meng and Chih-Fu Chang, "Tools for Compressed Domain Video Indexing and Editing", SPIE Conference on Storage and Retrieval for Image and Video Database, Vol. 2670, 1996.

[6] Jianhao Meng, Yujen Juan, Shih-Fu Chang, "Scene Change Detection in MPEG Compressed Video Sequence", IS&T/SPEG Symposium Proceeding Vol. 2419, 1995.

[7] H. Yu, G. Bozdagi, S. Harrington, "A Feature-based Hierarchical Video Segmentation", IEEE ICIP-97, Vol. 2, pp 498-501, 1997.

[8] Ramin Zabih, Justin Miller, Kevin Mai, "A feature-based algorithm for detecting and classifying production effects", Multimedia

Systems, Vol. 7, pp 119-128, 1999.

[9] Wei Jyh Heng and King N. Ngan, "Integrated Shot Boundary Detection using Object-based Technique", IEEE ICIP-99, Vol. 2, pp 289-293, 1999.

[10] K. R. Rao and J. J. Hwang, "Techniques & Standards For Image · Video & Audio Coding", PTR, pp 201-272, 1996.



성창우(Changwoo Seong)  
 準會員

1996년 동아대학교 전기  
 공학과 학사.

2000년 동아대학교 전자  
 학과 석사과정 재학

관심분야 영상처리, 멀티미디어 색인 및  
 검색, 영상통신 등.



강대성(Dae-Seong Kang)  
 正會員

1984년 경북대학교  
 전자공학과 학사.

1991년 Texas A&M Un  
 iv., Electric Eng. 석사.

1994년 5월 Texas A&M  
 Univ., Electric Eng. 박사.

1984년-1989년 국방과학연구소 연구원.

1994년-1995년 한국전자통신연구소  
 선임연구원.

1995년-현재 동아대학교 전기전자컴퓨터 공  
 학부 조교수.

관심분야 영상처리, 패턴인식, 영상코딩,  
 통신시스템 등.