

MPEG 동영상에서 부호화된 블록의 개수를 이용한 점진적 장면 전환 영역 검출

남승필, 오화종, 최병욱
한양대학교 전자통신전파공학부

A Research on dissolve detection in MPEG video streams using coded block pattern

Seung-pil Nam, Haw-jong Oh, Byung-uk Choi
Dept. of Electrical and Computer Engineering, Hanyang Univ.
E-mail : nspil@hymail.hanyang.ac.kr

요약

멀티미디어 데이터베이스에서 장면전환 영역을 검출하는 것은 검색과 색인을 위해서 필수적이다. 동영상에서 장면전환 영역은 단순한 장면전환과 점진적인 장면전환으로 나눌 수 있다. 단순한 장면전환은 다음 장면과 구별이 쉬우나, 점진적인 장면전환은 그 구별이 쉽지 않다. 본 논문에서는 압축된 동영상에서 점진적인 장면전환 영역을 검출하는 효과적인 방법을 제시한다. 제안된 알고리즘은 MPEG-1으로 압축된 동영상에서 DC계수를 추출하고, 부호화된 휘도 블록의 개수를 추출하여 점진적 장면전환 영역을 검출한다. 제안된 알고리즘의 성능은 장면이 점진적으로 바뀌는 영역을 찾아내는 정확도를 기반으로 분석하였다.

1. 서론

동영상에서 장면전환을 검출하는 것은 멀티미디어 데이터베이스를 구축하기 위한 가장 기본적인 단계이다. 동영상은 대부분 압축된 형태로 존재하는데, 장면전환 영역을 검출하기 위해서 압축된 동영상을 전부 디코딩하여 장면 전환 영역을 추출하는 것은 시간적으로나 메모리 문제에 있어서 바람직한 방법이 아니며, 압축할

때 장면이 바뀌는 부분에서 얻어지는 정보를 잃게 된다. 이에 본 논문은 최소한의 디코딩으로 장면전환 영역을 검출한다.

장면전환은 단순한 장면전환과 점진적인 장면전환으로 크게 나눌 수 있다. 이것은 알고리즘에 많은 제한을 준다. 단순한 장면전환만 있는 동영상에서 성능이 좋은 알고리즘이라고 할 지라도 점진적인 장면전환이 있는 동영상에서는 그 성능을 기대할 수 없다. 그리고 위의 예의 역도 마찬가지이다. 따라서 본 논문은 이런 두 가지 상황을 모두 고려한 알고리즘을 제시한다.

본 논문의 구성은 2장에서 DCT를 이용하여 장면전환을 검출하는 기존의 기법을 살펴보고, 이 기법이 점진적 장면전환에서 초래될 수 있는 문제점과 이를 보완하는 기법을 기술하며, 3장에서는 점진적 장면전환 영역을 검출하기 위해 P-Frame의 부호화된 휘도 블록 개수를 이용하는 기법을 기술한다. 그리고 4장에서는 제안한 점진적 장면전환 영역 검출 알고리즘의 실험결과 및 분석을 기술하고 제안한 알고리즘의 결론을 기술한다.

2. DC성분을 이용한 장면전환 검출

2.1 점진적인 장면전환 영역

동영상의 구조는 의미적 단위인 scene과 이 scene을

구성하는 물리적인 단위인 shot이 있고, 또 shot을 구성하는 프레임으로 되어있다. 장면전환은 크게 shot과 shot이 이어진 단순한 장면전환과 shot과 shot이 겹쳐져있는 점진적 장면전환으로 나눌 수 있다. 여기서 장면전환점을 살펴보면, 단순한 장면전환의 경우는 앞 shot의 마지막 프레임과 뒤 shot의 첫 번째 프레임이 된다. 하지만 점진적인 장면전환은 앞 shot의 프레임과 뒤 shot의 프레임이 겹쳐져 있으므로 뒤 shot의 프레임이 겹치기 시작하는 앞 shot의 프레임으로부터 앞 shot의 프레임이 없어지는 뒤shot의 프레임까지가 된다. 이 구간에 속하는 프레임들은 앞 shot에 속하기도, 뒤 shot에 속하기도 하는 모호한 화면이 되므로 shot과 shot을 구분할 때 빠져야 할 부분이 된다. 따라서 점진적인 장면전환이 들어있는 동영상에서 장면전환 구분은 점진적인 장면전환이 시작하는 프레임까지 앞 shot이 되고, 점진적인 장면전환이 끝나는 프레임부터 뒤 shot이 된다.

2.2 DC 성분을 이용한 장면전환 검출

MPEG-1은 시간적인 중복을 제거하기 위해 움직임 보상 방법을 사용하고, 공간적인 중복을 제거하기 위해 DCT 압축 방식을 사용한다. DCT는 8X8 블록 단위로 이루어지며, I-프레임에서 DCT는 프레임 전체의 블록에 적용된다. DCT를 적용하면 블록은 평균값에 해당하는 DC성분과 나머지 AC성분으로 구분된다. 여기서 DC성분은 8X8블록의 특징을 대표적으로 나타낸다. I-프레임은 해당 프레임 정보만으로 부호화되는 프레임이다. 따라서 I-프레임에서 DC성분은 프레임 전체를 대표하는 특징으로 충분하다. 그리고 MPEG-1은 랜덤한 재생을 위해 GOP단위로 부호화되는데, GOP당 반드시 1개의 I-Frame이 있어야 한다. 여기서 GOP단위가 15라면 I-Frame은 15번째마다 나타난다. 따라서 I-Frame간 DC비교로 장면전환을 검출하는 것은 수행속도와 연산량에서 강력한 알고리즘이다. 본 논문도 이 알고리즘을 이용하여 장면전환 점을 찾을 것이다. 그런데 점진적인 장면전환 영역이 여러 I-프레임을 포함하고 있다면 단순한 장면전환보다 I-프레임간의 DC성분 차이는 크지 않을 것이다. 이것은 점진적으로 장면전환이 일어나고 있지만 DC성분 비교로는 검출하지 못하는 경우를 초래할 수 있다. 이를 보완하기 위해 3장에서 기술할 부호화된 휘도 블록의 개수와 인접한 I-프레임간의 비교와 이를 보완할 방법을 사용한다.

DC성분으로 I-프레임간의 비교를 하기 위해서 사용되는 방법은 여러 가지가 있다. 그 중에서 본 논문은 DC성분간의 차의 절대치를 정규화하여 그 값이 임계값보다 크면 장면전환이 일어난 것으로 검출하는 방법을 사용한다.

$$D(fm, fn) = \frac{1}{k} \sum_{i=1}^k \frac{|c(fm, i) - c(fn, i)|}{\max(c(fm, i), c(fn, i))}$$

여기서 m,n은 프레임 번호이고, I-프레임 전체 블록 수, c(fp,i)은 fp프레임에서 i블록의 DC성분이다.

2.3 I-프레임간 DC성분 비교 방법

점진적 장면전환 영역이 여러 I-프레임을 포함하고 있으면, 이 영역에서 I-프레임의 휘도 DC성분은 서서히 변하기 때문에 인접한 I-프레임 비교만으로는 동영상의 장면전환점을 놓치고 넘어가는 경우가 있다. 이를 보완하기 위해 3장에서 설명할 부호화된 휘도 블록 개수를 이용한다. 부호화된 휘도 블록의 개수로 점진적인 장면전환이 시작되는 GOP를 검출하면, 이 GOP의 앞 GOP내의 I-프레임을 고정시키고 I-프레임 비교를 수행한다. 이것은 점진적인 장면전환 구간을 명확히 하는 효과를 가진다. 그림(1)은 위의 방법을 나타내고 있다.

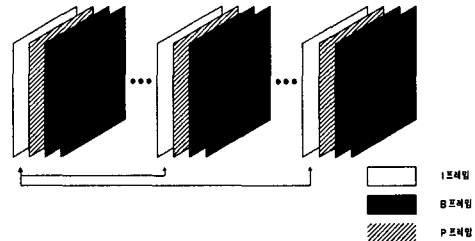


그림 1 I-프레임간 비교 방법

2.4 P-프레임내의 Intra매크로블록 이용

점진적 장면전환 영역이 아닌 단순한 장면전환을 인접한 I-프레임 비교로 검출하였다면 이 GOP내에서 장면이 바뀌고 있다. 더욱 장면전환점에 근접하기 위해서 P-프레임의 인트라 매크로블록 수를 사용한다. 인트라 매크로블록의 수가 많다는 것은 현재 P-프레임이 앞의 P-프레임 혹은 I-프레임과 다른 화면임을 나타낸다. 따라서 GOP내의 P-프레임 중 가장 많은 인트라 매크로블록 수를 가지는 P-프레임을 장면전환점으로 택한다.

3. 점진적 장면전환 검출 알고리즘

3.1 휘도 블록의 부호화 패턴

I-프레임은 모든 블록이 부호화된다. 그러나 P-프레임이나, B-프레임은 참조 프레임간의 예측을 통해 필요한 블록만 부호화한다. 부호화된 블록을 나타내기 위해서 P-프레임과 B-프레임은 CBP(부호화 블록 패턴)라는 정보를 가지고 있다. 이런 CBP 정보는 가변장 부호화되어 있으며, 한 매크로블록당 블록의 수는 MPEG-1인 경우 휘도 블록 수 4개와 색상 블록 수 2개로 6개이다. P-프

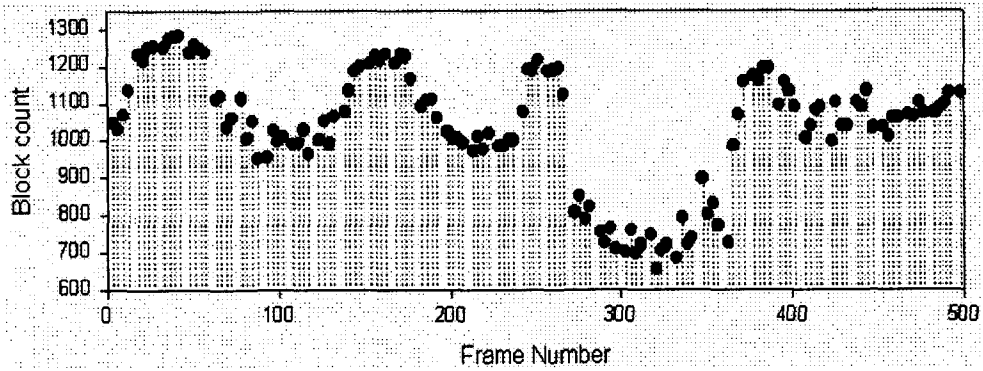


그림 2 P-프레임내의 부호화된 휘도 블록의 개수

레이프와 B-프레임에서 블록이 부호되어 있는지는 매크로블록타입의 매크로블록패턴 값이 있는가로 판단할 수 있다. 만약 이 매크로블록패턴 값이 없으면 매크로블록내의 모든 블록은 부호화되지 않는다. 여기서 부호화라는 것은 DCT 계수를 가지는 것을 의미한다.

블록이 6개이므로 부호화되는 종류는 모두 63개가 되는데, 다음 표(1)로 그 일부분만 나타낸다.

MB패턴의 VLC	CBP	Y	Y	Y	Y	Cb	Cr
111	60	1	1	1	1	0	0
1101	1	0	0	0	0	0	1
1100	8	0	0	1	0	0	0
1011	16	0	1	0	0	0	0
1010	32	1	0	0	0	0	0

표 1 MB 패턴(CBP)의 VLC

3.2 점진적 장면전환 영역 검출

GOP내에서 P-프레임이 나타나는 주기가 있다. 인코딩할 때 이 주기는 바꿀 수 있지만 보통 주기로 3을 사용한다. 이때 GOP가 15이면 P-프레임이 4개가 들어가는 것을 의미한다. 이것은 I-프레임만으로 비교했을 때 I-프레임 사이의 장면전환 점 검출에 P-프레임이 보완 역할을 할 수 있다 것이다. 그리고 I-프레임간 DC성분 비교는(식(1)) 휘도 차이를 구하고 그 차를 더하여 비교하는 것이다. 이것은 휘도가 바뀌는 화소의 수와는 무관한 값이다. 따라서 이를 보완하기 위해 화소별로 변화되는 개수를 구하는 것이 필요한데 이것은 P-프레임의 CBP를 통해 할 수 있다. 휘도 블록이 부호화 된다는 것은 부호화되는 화소에 휘도 성분이 바뀌고 있다는 것이다. 즉 부호화된 휘도 블록 개수를 조사하는 것은 변화된 화소를 조사하는 것과 같다.

제안한 알고리즘은 점진적 장면전환 영역에서 다음과

같은 특징을 반영하고 있다. 휘도 블록이 부호화 되는 경향을 보면 shot내에 비슷한 프레임이 지속되면 프레임간 예측을 통해서 부호화되는 블록 수가 점차 줄어드는 것을 알 수 있다. 그런데 점진적 장면전환(디졸브)이 나타나는 영역에서는 부호화된 휘도 블록의 개수를 조사해보면 앞 프레임 보다 많이 부호화되는 경향을 볼 수 있다. 이것은 이전 프레임에서 부호화되는 휘도 블록의 개수와 겹쳐지는 프레임에 의해서 부호화되는 휘도 블록이 더해지기 때문이다. 그림(2)을 통해 확인한다.

GOP단위로 부호화된 휘도 블록 개수를 조사해보면 점진적 장면전환이 일어나는 GOP에서 이전 GOP보다 훨씬 많은 것을 확인할 수 있다. 그리고 전 GOP에서 제일 적은 블록 수와 현재 GOP에서 제일 적은 블록 수와의 차이가 현저하게 차이가 난다. 또한 점진적 장면전환 영역에서 부호화되는 휘도 블록의 개수가 유지되는 것은 휘도 성분이 계속 변화하고 있기 때문이다. 이를 검출식으로 표현하면 다음과 같다.

$$D_{min}(G_m, G_n) = BC_{min}(G_m) - BC_{min}(G_n) \times 2$$

$$D_{max}(G_m, G_n) = BC_{max}(G_m) - BC_{max}(G_n) \times 2$$

$BC_{min}(G_m)$ 은 m번째 GOP에서 최소 부호화된 휘도 블록의 수의 전체 휘도 블록 수로의 정규화, $BC_{max}(G_m)$ 은 m번째 GOP에서 최대 부호화된 휘도 블록의 수의 전체 휘도 블록 수로의 정규화. $D_{min}(G_m, G_{m+1})$ 이 임계값(th1)보다 크면 m+1번째 GOP에서 점진적 장면전환 영역이 시작되고 $D_{max}(G_m, G_{m+1})$ 이 임계값(th2)보다 작으면 m번째 GOP에서 점진적 장면전환 영역이 끝난다. 검출한 GOP들 구간 중에서 점진적 장면전환 영역은 부호화된 휘도 블록 수가 장면전환 영역 전과 후의 GOP구간의 최대 휘도 블록 수보다 큰 프레임이 있는 영역이다.

3.3 알고리즘 구현 방법

GOP단위로 I-프레임의 DCT의 DC성분과 P-프레임의 부호화된 휘도 블록 수를 구한다. 그리고 인접 I-프레임 간의 DC성분비교를 하여 정해진 임계값보다 크면 장면 전환이 일어난 것이므로 이 GOP단위에서 P-프레임 중 인트라 매크로블록이 제일 많은 P-프레임이 장면전환점이다. 그런데 DC성분 비교에서 임계값보다 작다면 식(2)를 사용하여 점진적 장면전환 영역인지 검사한다. 그리고 점진적 장면전환이라면 이전 I-프레임을 고정시키고 다음 I-프레임과의 DC비교 과정을 해나간다. 수행 중 임계값보다 큰 I-Frame을 찾으면 점진적인 장면전환 영역을 벗어난 것이다. 그리고 식(3)으로 이를 검증한다.

"Scene change detection in a MPEG compressed video sequence" Digital Video Compression: Algorithm and Technologies, Pres. SPIE 2419, pp 14-25, 1995.

- [4] 정제창 역
"최신 MPEG" 교보문고.

4. 실험결과 및 결론

본 논문에서 실험에 사용한 동영상은 디졸브(dissolve)효과가 들어간 MPEG-1 동영상이다. 이 동영상은 인터넷에서 구했으며, 점진적 장면전환 점을 찾는 데 많이 사용되는 동영상이다. 실험은 3개의 동영상으로 하였고, 그 동영상은 다음과 같다. sunrise.mpg, beach.mpg, autumn.mpg는 모두 크기가 352x240 이며, 29.97 frames/s로 각각 총 프레임 수는 591프레임, 186프레임, 447프레임이다.

동영상	Dissolve	검출 프레임	FIND
sunrise	4	(18,57), (147,174) (246,264), (372,396)	100%
autumn	3	(33,72), (156,192) (333,366)	100%
beach	3	(54,72), (99,108) (147,162)	100%

표 2 실험 결과

본 논문이 제시한 점진적인 장면전환 영역 검출 알고리즘은 MPEG-1 동영상을 최소한 디코딩하여 필요한 특징을 추출해 내었다는 것을 실험결과가 검증하고 있다.

6. 참고문헌

- [1] N. Gamaz, X. Huang and S. Panchanathan
"Scene change detection in mpeg domain"
IEEE, 1998.
- [2] Ok-Bae Chang, Myung-Sup Yang, Jae-Hyun Lee
"Segmentation of Gradual Scene Transitions Motion Vector" Part of the SPIE Conference on Input/Output and Imaging Technologies. Vol. 3422 July 1998.
- [3] J. Meng, Y. Juan, and S.F. Chang