

압축된 MPEG 비디오에서 최적화된 검색간격을 이용한 빠른 장면전환 검출에 관한 연구

김중헌, 김신형, 박두영, 장종환
배재대학교 정보통신공학과

e-mail : jhkim@icelab.paichai.ac.kr, zeros@videolab.pcu.ac.kr,
dypark@mail.paichai.ac.kr, jangjw@mail.paichai.ac.kr

A Study of Fast Scene Change Detection Using Optimized Temporally Sampling from Compressed MPEG Videos.

Joong-Heon Kim, Shin-Hyoung Kim, Doo-Yeong Park, Jong-Whan Jang
Dept. Information & Communication of Engineering, PaiChai University

요 약

비디오 데이터의 효율적인 저장, 관리를 위해서는 장면전환 검출을 통한 비디오 분할 기술에 대한 연구가 필요하다. 기존의 장면전환 검출 알고리즘은 복호화에 의한 연산량 및 프레임들간의 비교에 의한 연산량이 많고 잡음에 의한 오검출 및 미검출이 발생하게 된다. 본 논문에서는 MPEG 압축 비디오에서 압축 영역에서의 빠른 장면전환 검출을 위한 최적의 검색 간격을 유도하였고, 검색간격 내의 B 프레임의 방향성을 직접 추출하여 장면전환검출에 이용하므로 빠르고 정확한 장면전환 검출 알고리즘을 제안한다.

1. 서론

현대 사회의 가장 큰 특징 중에 하나는 다양한 멀티미디어 데이터를 사용하고 있다. 멀티미디어 데이터를 복수의 표현 미디어라고 할 때, 여러가지 구성요소 가운데서 동영상은 멀티미디어 데이터의 핵심이라고 말할 수 있으며, 이러한 동영상에 대한 압축/복원 기술의 국제적인 표준 가운데서 우리에게 가장 잘 알려진 것이 바로 MPEG 일 것이다.

이러한 동영상 압축 기술의 표준화로 많은 동영상들이 MPEG 기술을 이용하여 압축/저장되고 있으며, 그 수가 기하급수적으로 늘어나고 있다. 디지털 저장 매체에 저장된 동영상들의 수가 늘어남에 따라서 이 정보들을 좀더 효율적으로 이용하기 위하여 체계적으로 분류, 저장할 필요성이 생겨나게 되었다. 이런 필요하에 나타나게 된 것이 근래에 대두되기 시작한 MPEG 7 이다.

MPEG 7은 content-based storage and retrieval 을 위한 국제 표준으로써, 이를 위하여 동영상 데이터 색인화

를 위한 기술이 필요하게 될 것이다. 동영상 데이터 색인화를 위해서는 이를 위해서 우선적으로 필요한 기술이 바로 장면전환 검출 알고리즘이라고 할 수 있다. 현재 장면전환 검출 알고리즘에 대한 연구가 많이 진행되고 있으며, 많은 연구 결과들이 발표되고 있다. 기존의 장면전환 검출 알고리즘을 살펴보면 [1]는 두 영상의 픽셀 및 히스토그램 차를 이용한 장면전환 방법을 제안하였는데 이 방법은 구현하기 간단하고 널리 사용되는 방법으로 급진적으로 변화는 장면은 정확히 검출하는 편이나 움직이는 물체가 존재하거나 빛의 변화에 따른 오검출이 많은 단점이 있다. [2] 및 [3]등은 모델기반 장면전환 검출 기법을 제안하였다. 그러나 이 방법들은 압축된 영상에 적용할 수 없기 때문에 복호화에 의한 시간이 많이 소비 되는 문제점이 있다. 최근 이러한 단점들을 해결하기 위해 압축된 비트스트림이 갖고 있는 내용을 직접 분석하여 특정 정보를 가지고 동영상의 장면전환 검출에 활용하려고 노력하고 있다[4][5][6]. 위와 같은 방법들은 비 압축영역이든 압축영역이든 방대한 영상자료를 매 프레임마

다 비교해야 하므로 계산량이 매우 많으며 검색속도 또한 아주 느리게 된다.

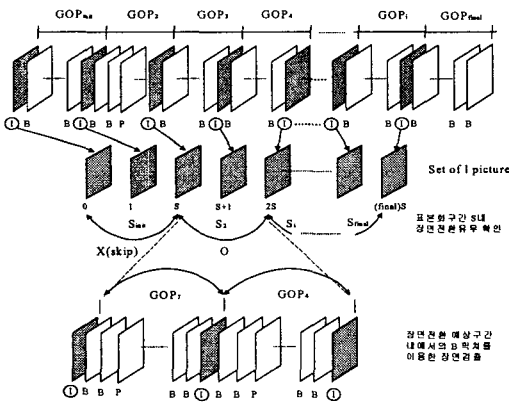
본 논문에서는 압축 영역에서 비 순차검색 방법을 이용하여 프레임 비교횟수를 줄이므로 검출 속도를 높이고 검색 영역내의 B 프레임의 특징을 직접 추출하여 정확한 장면전환 프레임을 검출해 낸다.

본 논문의 구성은 2장과 3장에서 순차 검색방법의 문제점을 해결하기 위한 비순차 검색방법과 검색 영역내에 장면전환이 존재 할 경우 영역안의 B 프레임의 특징을 직접 추출하여 장면전환 프레임을 검출하는 알고리즘을 제안하였다. 4 장에서는 제안한 알고리즘을 적용한 실험치 결과를 기술하며 5 장에서는 결론 및 향후 과제를 제시하였다.

2. 비순차 검색 방법

비 순차검색 방법은 동영상을 일정한 시간간격으로 표본화 하여 장면전환 여부를 검색하는 방법으로 장면전환이 존재하는 구간에서만 프레임을 비교하여 검색하기 때문에 동영상 전체 프레임의 각각의 프레임마다 연속적으로 비교하는 순차검색 방법보다 비교 횟수를 줄일 수 있어 검색시간 단축에 유리하다. [그림 1]은 본 논문에서 제안한 압축영역에서의 비 순차 검색방법으로 검색속도를 최소화 하기위해 B, P 픽처는 고려하지 않고 GOP 단위를 기준으로 I 픽처만을 묶어 I 픽처의 DC 영상을 이용한 히스토그램 비교를 통해 표본화 검색구간 S 내에 장면전환 유무를 파악하고 장면전환이 존재하는 구간에서만 B 픽처를 이용한 장면전환 지점을 찾는 방법을 보여주고 있다.

NTSC 용 MPEG 은 하나의 GOP 에 보통 15 개 이하의 프레임 을 포함하고 한 GOP 에서 I 프레임 이 적어도 한번은 출현하며 각 GOP 의 대표 프레임이라 할 수 있기 때문에 I 프레임 만으로 장면전환이 존재하는 구간을 검출하여도 크게 무리는 없다[7].



[그림 1] 제안한 비순차 검색 방법

2.1 최적 장면검색간격(S_b)

본 논문에서 최적 장면검색 간격을 구하기 위한 알고

리즘이 잘 성립하기위해 다음 조건을 가정한다.

1. 장면전환이 한곳에 집중적으로 분포하지 않는다.
2. 표본화 검색구간 S 내에는 장면전환 지점이 없거나 최대 1 개 존재한다.

전체 GOP 개수를 N, GOP 내의 B 프레임의 개수를 B, 표본화 검색 간격을 S, 총 프레임 비교 횟수를 T, 평균 장면전환 거리를 d, 장면전환 수를 k라 한다면, 먼저 평균 장면전환 거리 d는 다음과 같이 정의할 수 있다.

$$d = \frac{N}{k} \tag{1}$$

총 검색 비교 횟수 T는

$$T = \frac{N}{S} + k \cdot S + B \tag{2}$$

여기서 $\frac{N}{S}$ 은 전체 동영상의 표본화 검색구간의 개수로 그 구간 안에 장면전환이 있는지를 검사하기위해 각각의 표본화 구간의 처음 I 픽처와 마지막 I 픽처를 비교하여 장면전환의 유무를 판별한다.

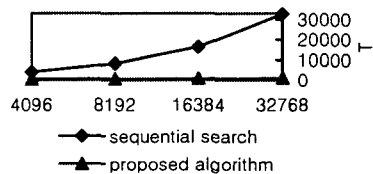
$k \cdot S$ 는 장면전환이 존재하는 구간내의 I 픽처 비교 횟수를 나타낸다.

총 검색 비교 횟수 T를 최소로 만드는 S를 S_b라 할 때 S_b를 구하기 위해 식(2)의 양변에 미분을 취하고 미분값이 0을 만족하는 S_b를 다음과 같이 얻는다.

$$\frac{d}{dS}(T) = -\frac{N}{S^2} + k \tag{3}$$

$$-\frac{N}{S_b^2} + k = 0 \tag{4}$$

$$S_b = \sqrt{\frac{N}{k}} = \sqrt{d} \quad \text{단, } S_b \text{는 양의 정수} \tag{5}$$



[그림 2] 방식별 프레임 비교횟수

[그림 2]에서 [■]는 순차검색방법의 프레임 비교 횟수를 나타내고 [□]는 제안한 알고리즘의 프레임 비교횟수를 나타낸다. [그림 2]에서 보듯이 순차검색과는 비교횟수의 차이가 크게 나타나고, 기존의 비 순차검색 방법과도 비교횟수를 줄일 수 있다. 검색대상이 되는 동영상의 프레임수가 많을 경우 제안한 알고리즘은 기존의 방법보다 비교횟수를 크게 줄일 수 있다.

3. 검색 영역내에서의 B 프레임의 특징을 이용한

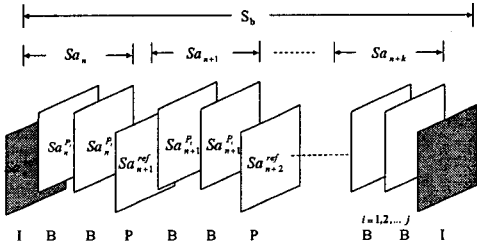
장면전환 검출

표본화 구간내에 장면전환이 존재 할 경우 정확한 프레임을 검출하기 위해 검색 영역내의 B 프레임의 특징을 압축 영역에서 직접 추출하여 장면전환 프레임임을 검출한다.

일반적으로 MPEG 시퀀스에서는 기준이 되는 프레임 사이에 압축 효율을 높이기 위해 두 개 이상의 B 프레임이 열을 이루고 있다. 각각의 B 프레임들의 매크로 블록들은 forward 또는 backward 의 방향성을 가지게 되고 방향성에 의해 B 프레임 안에서 매크로 블록의 종류는 크게 4 종류로 나눌 수 있다. Forward, backward, Intra, 그리고 양방향 모두 존재하는 Interpolation 이 있다. 인코딩 시 각 매크로 블록의 방향성은 식(6)에 의해 결정된다.

$$Var = Vard - Varc \begin{cases} Intra, & \text{if } Var > 0 \\ MV, & \text{otherwise} \end{cases} \quad (6)$$

식(6)에서 Vard 는 참조 영상과 픽셀 값들의 차이의 절대값이 최소인 매크로 블록과 현 영상의 매크로 블록과의 차의 분산값이고 Varc 는 현 영상의 매크로 블록의 분산값이다. 이때 Vard 는 forward , backward,, interpolation 방향중 최소가 되는 값으로 결정된다. 식(6)에서 보는것과 같이 참조영상과 차의 분산값이 적으면, 즉 참조 영상과 유사성이 높이면 그 매크로 블록은 방향성을 가지게 되고 그렇지 않으면 방향성이 없는 Intra 타입이 된다.



[그림 3] 정확한 장면전환 검출을 위한 Sa 영역 정의

[그림 3]은 기준 프레임 사이에 B 프레임이 2 개가 존재하는 일반적인 시퀀스에서 식(5)의 최적 장면전환 검색 간격인 S_b 영역내에 기준프레임부터 다음 기준 프레임 전까지를 정확한 장면전환 프레임을 검출하기 위해 Sa 로 정의 하였다. [그림 3]에서 Sa_n^{ref} 는 Sa_n 의 기준 프레임이고 Sa_nⁱ 는 Sa_n 내의 B 프레임의 개수이며 i 는 0 < i ≤ j 의 크기를 가진다.

Sa_n 의 B 프레임들인 Sa_nⁱ 에서 forward 타입의 매크로 블록 개수를 Fw(Sa_nⁱ), backward 타입의 매크로 블록 개수를 Bw(Sa_nⁱ)라 정의 할 때 각각의 B 프레

임의 방향성을 다음과 같이 결정할 수 있다.

$$Df(Sa_n^i) = Fw(Sa_n^i) - Bw(Sa_n^i), i=1,2..j \quad (7)$$

$$Df(Sa_n^i) = \begin{cases} 1, & \text{if } D(Sa_n^i) \geq Th_f \\ 0, & \text{if } -D(Sa_n^i) \geq Th_b \\ -1, & \text{otherwise} \end{cases}, i=1,2,...j \quad (8)$$

식(8)의 Th_f 는 양의 정수로서 forward 방향성을 결정하는 임계값이고 Th_b 역시 양의 정수로 backward 방향성을 결정하는 임계값이다. Df(Sa_nⁱ) 값이 1 이면 Sa_nⁱ 프레임의 방향성은 forward 가 되고 0 이면 backward, -1 일 경우는 한쪽 방향으로의 방향성이 없는 B 프레임이 된다. 만약 Df(Sa_nⁱ)가 1 의 값을 가진다면 Sa_nⁱ 프레임은 forward 방향의 기준 프레임을 참조한 매크로블록의 수가 많고 backward 방향의 영상과는 연관성이 적으므로 backward 방향의 영상은 전혀 다른 영상일 경우가 많게 된다.

본 논문에서는 Sa_n 영역 내에서 장면전환이 발생하는 경우를 2 가지로 정의 하였다.

1. 기준 프레임에서 장면전환이 발생
 2. Sa_n 영역의 임의의 k 번째 B 프레임에서 장면전환이 발생
- 위에서 정의한 각각의 장면전환 검출의 경우를 식(8)에 의해 다음과 같이 정의 하고 검출할 수 있다.

$$Cdf(Sa_n^k) = \begin{cases} Sa_{n+1}^{ref}, & \text{if } Df(Sa_n^i) = 1, i=1,2..j \\ Sa_n^i, & \text{if } \begin{cases} Df(Sa_n^i) = 0, (k \leq i \leq j) \\ Df(Sa_n^i) = 1, (0 < i < k) \end{cases} \end{cases}, k \leq j \quad (9)$$

식(9)에서 Df(Sa_nⁱ)가 1 이면 영역내의 모든 B 프레임들이 forward 방향성을 가지므로 backward 방향에 존재하는 다음 영역의 기준 프레임인 Sa_{N+1}^{ref} 프레임에서 장면전환이 검출된다. 또한 영역내의 임의의 B 프레임인 Sa_n^k 에서의 장면전환 검출은 Sa_n^k 이전 B 프레임들은 forward 방향성을 Sa_n^k 를 포함한 이후의 B 프레임들은 backward 방향성을 가진다.

2.1 장면전환 검출 개선

뚜렷한 장면전환이 존재할 경우 B 프레임의 열은 forward 또는 backward 중 한쪽으로 치우치는 방향성을 가지게 된다. 그러나 B 프레임의 방향성에는 양쪽 방향을 동시에 가지는 interpolation 타입과 아무런 방향성을 가지지 않는 intra 타입이 존재하게 된다. 이 두 타입은 한쪽의 영상을 참조하지 않기 때문에 장면

전환 검출에 있어서 예러로 정의 할 수 있다.

Interpolation 타입의 개수를 $Inter(Sa_n^i)$, intra 타입의 개수를 $Intra(Sa_n^i)$ 라 할 때 식(8),(9)에 의해 검출된 장면전환 중 오검출을 식(10)에 의해 판단할 수 있다.

$$\frac{Inter(Sa_n^i) + Intra(Sa_n^i)}{Fw(Sa_n^i) + Bw(Sa_n^i)} \geq Th_e, i = 1, 2, \dots, j \quad (10)$$

Th_e 는 오검출을 판단하는 임계값이다.

3. 실험 및 결과

동영상의 장면전환 검출에 있어서 순차검색과 본 논문에서 제안한 압축 영역에서의 비 순차 장면전환 검출 알고리즘의 성능을 분석하였다. 장면전환 검출에서의 성능을 비교하기 위해 윈도우 2000 환경에서 Visual C++ 6.0 을 이용하여 구현하였으며 실험 대상 동영상은 MPEG 로 압축된 드라마, 뉴스영상, 광고 각 1 편식을 선정하였고 영상길이는 각 영상의 종류별 반영시간 특성을 고려하여 가장 보편적인 값으로 선정했다.

<표 1>은 드라마, 광고, 동영상에 대한 각각의 특징을 보인 것이다. 평균장면전환 거리는 식(1)을 이용하여 얻었고 평균 장면전환 거리는 식(5)를 이용하여 얻었다.

	드라마	광고	뉴스
총 프레임수 (총 GOP 수)	9932 (663)	440 (30)	5158 (344)
장면전환 개수(k)	60	8	31
평균 장면전환 거리(d)	12	4	12
표본화 검색 간격(S_b)	3	2	3

<표 1> 실험 대상 동영상의 특성

<표 2>는 순차검색과 제안한 검색방법과의 장면전환 검출 실험결과를 보인 것이다. 특히 드라마 및 뉴스에서 물체의 움직임이나 후레쉬와 같은 잡음에 의한 오검출이 감소되는 것을 볼 수 있다. 식(8)의 Th_f 와 Th_b , 식(10)의 Th_e 는 실험에 의하여 Th_f 와 Th_b 는 매크로 블록 개수의 70%로 Th_e 는 30%로 설정하고 실험하였다.

	드라마		광고		뉴스	
	순차	제안한 방법	순차	제안한 방법	순차	제안한 방법
검출	68	60	8	8	64	34
미검출	0	0	0	0	1	1
오검출	8	0	0	0	34	4
비교 회수	9932	1263	440	120	5158	684

<표 2> 실험 영상에 대한 장면전환 검출 결과 및 프레임 비교회수

<표 2>에서 보는 것 처럼 후레쉬나 물체의 움직임에 의한 오검출이 기존 순차검색 보다 약 70%의 감소를 보였고 장면전환 검출을 위한 비교회수 또한 70% 이상 감소되는 것을 관찰 할 수 있었다.

4. 결론 및 향후 과제

본 논문에서는 장면전환 검출에 있어서 프레임 비교회수를 최소화 하여 계산량 및 검출속도를 향상시키는 최적 검색간격을 유도하였고 장면전환 후보 구간에서 복호화 과정없이 B 프레임의 특징만을 이용하여 비교하는 방법으로 장면전환 검출을 시행하여 기존의 알고리즘 보다 처리 시간의 단축과 정확한 검출 및 복잡하지 않은 검출 과정으로 우수한 결과를 도출하였다. 다양한 MPEG 비디오 스트림을 적용시킨 실험을 통해 정확한 장면전환 프레임 검출 및 수행 시간 측면에서 효율성을 확인하였다.

향후 연구 방향으로서는 압축 데이터 상에서 점진적인 장면전환을 효과적으로 검출해 내는 방법과 B 프레임이 존재하지 않는 시퀀스에서도 적용할 수 있는 방법에 관하여 연구할 계획이다.

참고문헌

- [1] A. Nagasaka and Y. Tanaka, "Automatic Video Indexing and Full Motion Search for Object Appearance," Proc. Of IFIP on Visual Database System, pp. 113-127, Sep. 1991.
- [2] A. Hampapur, R. Jain, and T. Weymouth, "Production Model Based Digital Video Sequention," Multimedia Tools and Applications, Vol. 1, No. 1, pp. 9-46, Mar. 1995.
- [3] B. Shahary, "Scene Change Detection and Content-Based Sampling of Video Sequences," Proc. Of SPIE, Vol. 2419, pp. 2-13, Feb. 1995.
- [4] B.L. Yeo and B. Lie, "Rapid scene analysis on compressed video," IEEE Trans. On Circuits and Systems for Video Technology, vol. 5, no. 6. pp. 533-544, Dec. 1995
- [5] R. Lienhart, S. Preiffer and W. Effelsberg, "Video Abstracting". ACM Communication, Vol. 40, No. 12, pp. 55-62, 1997.
- [6] G. Medioni, R. Nevatia and I. Cohen, "Event detection and Analysis from Video Stream", DARPA98, pp. 63-72, 1998.
- [7] Nilesh V. Patel and Ishwar K. Sethi, "Compressed Video Processing for Cut Detection" IEEE Proc, Video, Image and Signal Processing 1996.