

비디오 프레임 타입을 이용한 비디오 샷 검출

김영빈* · 류광렬* · 로버트스크라바시**

*목원대학교, **UPMC

Video Shot Detection Based on Video Frame Types

Young Bin Kim* · Kwang Ryol Ryu* · Robert J. Scلابassi**

*Mokwon University, **University of Pittsburgh Medical Center

conan9977@yahoo.com

요 약

논문은 비디오 픽처 타입을 이용한 비디오 샷 검출에 관한 연구이다. 제안한 방법은 압축된 비디오 프레임에 대하여 원 영상을 복원하지 않고, 압축 상태의 비디오 프레임을 이용한다. I픽처 프레임에서 DC영상을 복원하고, P픽처 프레임에서는 매크로블록의 개수를 이용하여 비디오 샷을 검출한다. 테스트 비디오를 이용하여 실험 결과 85~98%의 장면전환 검출이 가능 하였고, 압축비트스트림을 복원하여 장면전환의 샷을 검출 하는 기법에 비해 4배 빠른 검색이 가능하다

ABSTRACT

The video shot detection based on video picture type is presented in this paper. The detection algorithm is used MPEG compressed video frame directly, not reconstructed the original image. For shot detection, I and P frame of MPEG video bit stream are classified. The detecting scene cuts at I pictures are detected by reconstructed DC image. While scene cuts at P picture frame by monitoring the percentage of Intra-macroblocks per P picture. Experimental results on the test video bit stream is shown the detection rate of 85~98% and searching time is 4 times faster than the previously known video shot detection algorithm on the decompressed video shot.

Keywords

Changing scene shot detection, Video frame type, Intra-block, MPEG, moving compensation

1. 서론

인터넷을 통한 대용량의 멀티미디어 정보 중 압축된 비디오 데이터는 크기가 커지면서 특정 데이터를 순차적으로 검색하거나 원하는 장면을 보다 빠르고 정확하게 검색하기 위한 요구가 점차 증가하고 있다. 이러한 요구를 만족시키기 위해서는 비디오 스트림을 장면 단위로 분할하여 다른 내용의 장면으로 전환되는 위치들을 효율적으로 모두 검출해내는 장면 전환 검출 기법이 필요하다. 비디오 데이터의 장면들의 특성을 비교하여 서로 다른 내용의 장면을 변환되는 프레임을 검출하는 방식에는 크게 프레임 기반의 장면 전환 검출 방식과 압축 비디오 데이터 기반의 검출 방식으로 나눌 수 있다.[1~3] 프레임 기반의 검출 방식으로는 히스토그램과 같은 프레임의 특징을 이용하는 방식으로 프레임의 특징을 사용함으로써 압축 비디오 데이터 기반의 검출 방식 보다는 섬

세하게 장면 전환을 검출할 수 있다. 그러나 물체나 카메라의 빠른 움직임과 순간적인 밝기의 변화와 점진적인 장면 전환에는 민감하지 못하며, 장면 전환을 검출 하는데 소요 되는 시간이 많이 든다. 그리고 히스토그램의 임계값이 영상 값의 분포도에 따라 변경되는 단점이 있다. Miller와 Mai는 컷, 페이드, 디졸브, 와이프 등 여러 가지 장면 전환을 찾기 위해 에지를 사용했으며, Zabih는 복원 영상의 에지 정보를 픽셀 분포를 이용하여 디졸브, 페이드 및 와이프를 검출 하였다. 압축 비디오 데이터 기반의 장면 전환 검출 방식은 여러가지 장점을 갖는다. 첫째로, 압축 영역에서는 처리할 데이터 양이 화소 영역보다 많이 감소하므로 알고리즘 구현하는데 있어서 오버헤드가 감소하고, 둘째로, 압축 영역에서는 일반적으로 작은 양의 데이터를 처리하기 때문에 시간을 절약할 수 있다. 끝으로, 압축된 데이터를 완전히 복호화 할 필요가 없기 때문에 복호화 하는데 걸

리는 시간을 줄일 수 있다. 그러나 실질적인 프레임의 특징을 이용하지 않고 압축된 비디오 데이터의 특징을 이용하기 때문에 장면 전환의 정확성이 떨어지는 단점이 있다. Arman은 JPEG로 압축된 데이터의 선택된 영역에서 DCT 계수를 비교하여 셋 경계를 검출 하였고, Zhang등은 P픽처 또는 B 픽처의 프레임에서 타당한 움직임 벡터의 개수를 계산 하였다. 이러한 움직임 벡터의 개수가 임계값 보다 적어 영에 가까우면 카메라 움직임의 경계는 P 픽처 또는 B 픽처 형태의 프레임 전후에 존재한다고 제안하였다. Meng등은 급격한 셋 경계 검출을 위해 움직임 벡터의 분포를 이용하였고, 점진적인 셋 경계를 검출 하기 위해서는 DCT의 DC 계수의 편차를 이용 하였다. Nakajima는 I 픽처 프레임의 DCT 계수의 값을 이용하여 주파수 영역에서 빠른 비디오 셋 검출 방법을 제안하였다. 그러나, 비디오 셋이 P 또는 B 픽처 프레임에서 발생할 경우 비디오 셋 검출이 안되는 문제점이 있다.[4~9]

따라서, 본 논문에서는 압축된 MPEG 비트 스트림을 원이미지로 복원하지 않고, 모든 I, P 픽처 프레임에 대하여 비디오 셋 검출이 가능한 기법을 제안한다. 제안 알고리즘은 먼저, 픽처 프레임을 분류하고, I 픽처 프레임에 대해서 DC영상을 복원하여 비디오 셋을 검출하고, P 픽처 프레임에 대해서는 매크로블록의 개수를 연산하여 비디오 셋을 검출 하도록 한다.

II. 비디오 셋 검출

2.1. 셋 검출 과정

비디오 스트림의 픽처 타입은 압축되어 있는 데이터에 따라 I, P, B의 GOP로 구성되거나 I, P로 구성될 수 있다. 만일 B 픽처 타입이 없는 GOP의 경우 Zhang의 방법으로는 장면을 검출할 수 없게 된다. 따라서 I 픽처 타입에서 DCT 계수를 이용하여 DC 영상을 복원하여 I 픽처 전후의 I 픽처와 비교하여 장면 전환 발생을 검출하게 된다. 그러나 I 픽처만을 이용하여 장면 전환을 검출 하게 될 경우 I 픽처 사이의 P 픽처에서 발생하는 장면 전환은 검출을 할 수 없게 된다. P 픽처에서 발생하는 장면 전환은 P 픽처의 매크로블록의 개수를 이용하여 매크로블록의 개수가 임계값 이상인 경우 장면 전환 셋으로 검출하게 된다. 그림1에 제안하고 있는 알고리즘을 보이고 있다. 입력되는 압축 비디오 입력 스트림은 비디오 픽처 타입에 따라 I 픽처 프레임과 P 픽처 프레임으로 구분하여 장면 전환 셋을 검출한다. 압축된 데이터를 복호화 하지 않고 연산을 수행하기 때문에 신속한 검출이 가능하고 I 픽처와 P 픽처 모든 프레임에 대하여 장면 전환 검색을 수행기 때문에 입력 스트림에 대한 전반적인 검출이 가능하다.

2.2. 움직임 보상

MPEG 비디오 스트림은 계층화된 구조를 가지고 헤더에 의해 비디오 신호와 오디오 신호로 분류 된다. 비디오 신호는 서로 다른 여러 장의 정지영상들로 구성된다. 픽처는 정지 영상을 말하며 연속된 픽처를 GOP는 연속된 픽처이다.

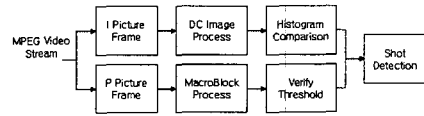


그림1. 비디오 셋 검출의 과정

헤더에 의해 각 계층으로 구분되며, MPEG에는 4종류의 픽처 프레임으로 구성된다. I 프레임은 움직임 보상을 하지 않고 해당 프레임만을 DCT 하여 부호화 한 것이고, P 프레임은 I 프레임이나 다른 P 프레임을 참조하여 움직임 보상을 적용하여 발생하는 차를 DCT 한다. B 프레임에서는 P 프레임 처럼 움직임 보상을 사용하지만 P 프레임과 달리 시간 축 상에서 앞과 뒤에 위치한 두 개의 참조 프레임(I 또는 P)으로부터 움직임 보상을 한다. GOP의 구성에서 I 픽처와 P 픽처의 구성과 I, P, B 픽처로 구성 될수 있다. GOP에서 I 픽처는 적어도 한 개 이상 필요로 하고 있다. P 픽처 프레임은 I 픽처 프레임을 참조하여 움직임 보상을 하기 때문에 I 프레임의 DCT 계수의 상관을 이용하여 빠른 장면 전환을 검출 할 수 있다. 그림2는 N=15이고 M=1인 GOP 구조를 갖는 비디오 시퀀스를 보이고 있다. 1나의 GOP에서 P 프레임은 모두 14개이고 P5~P12의 프레임은 점선으로 표시하였다. 희생의 I 프레임은 각 GOP의 시작 프레임이 된다. P1은 이전 프레임인 I 프레임을 참조하여 복원되는 두 번째 프레임이 된다.

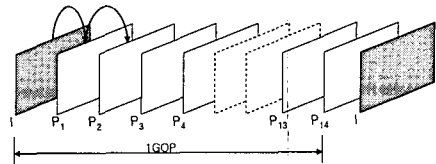


그림2. 비디오 프레임 움직임 보상

2.3. I, P픽처 프레임 장면 전환 셋 검출

I 픽처 프레임은 다른 프레임과 중복되는 정보를 제거하지 않고 인코딩 되어 있어 다른 프레임의 참조 없이 재구성이 가능하다. 압축되어있는비디오비트스트림에서원영상을복호화하지않고 DCT 계수의 DC 항을 이용하여 구성된 DC 영상은 크기가 원 영상의 1/64로 축소되어 작지만, 원 영상이 가지고 있는 중요한 특징들을 가지고 있어서 영상의 분석에 유용하게 사용할 수 있다. DC 영상을 사용하는데 있어서 장점은 원 영상에 비

해 빠른 처리가 가능한 것이다. 또한, I 픽처 프레임에서 DC 영상을 복원하기는 쉽지만, P, B 픽처 프레임에서 DC 영상의 복원은 쉽지 않다. 그러나, DC영상은 크기가 작기 때문에 중요한 특징들을 처리하는데 있어서 민감하게 반응하는 단점이 있다. DC영상을 복원해서 밝기 영상은 필터를 사용하여 단순화 한다. 인간이 밝기 영상을 인식할 때, 밝기 값이 0에서 50사이의 변화 정도나 200에서 250 사이의 변화 정도는 50에서 200 사이의 변화 정도와 많은 차이가 있기 때문에, 밝기 변화 단계를 변화가 심하게 느끼는 영역에서는 작게 분할하고, 밝기의 변화가 적게 느껴지는 영역에서는 크게 분할하도록 한다. 이렇게 정량화 하면 원영상의 급격한 밝기 변화가 단순화 되는 영상을 얻을 수 있다. 이러한 방법으로 I 픽처 프레임에서 복원된 영상은 히스토그램 변화를 이용하여 장면 전환 셋을 검출 한다. P픽처의 경우 I 픽처 프레임을 이용하여 장면 전환을 검출 하면 작은 사이즈의 DC영상을 복원하여 신속한 셋 검출이 가능하지만 P픽처 프레임에서 발생한 장면 전환은 검출이 어렵다. 따라서, P 픽처 프레임에서도 장면 전환 발생 여부를 확인 할 필요가 있다. P 픽처 프레임은 I 픽처 프레임 또는 최근의 다른 P 픽처 프레임을 기준으로 전환 예측, 즉 전환 이동 벡터를 검출하고 검출 된 이동 벡터를 이용하여 P 프레임의 구성한다. P 픽처 프레임에서 장면 전환 셋을 검출 하는 과정은 P 픽처 프레임 간의 매크로블록 변화를 이용한다.

MPEG 비디오 데이터는 매크로블록 단위로 부호화가 이루어지며, 각 프레임의 매크로블록을 부호화 한다. 즉, 비디오 셋이 발생하여 움직임 보상이 무효한 경우나 해당 프레임 내에서 움직임 보상이 더 좋은 경우엔 프레임 종류에 관계 없이 매크로블록을 인트라 모드로 부호화 한다. 인트라 모드로 부호화 된 매크로블록의 개수 θ 가 많은 프레임은 셋이 발생했을 가능성이 많으므로 해당 프레임을 장면 전환의 후보 셋으로 검출한다. 검출된 후보 셋의 θ 는 히스토그램 차의 임계값 선택에 중요한 정보가 된다. 식 (1)은 매크로블록의 임계값과 비디오 셋을 검출하기 위한 임계값에 대한 조건이다.

$$\begin{aligned} TH_{max} : & \text{if } MB_{max} \leq \theta \\ TH_{min} : & \text{if } MB_{min} \leq \theta \leq MB_{max} \end{aligned} \quad (1)$$

임계값 MB_{min} 은 인트라 모드로 부호화 된 매크로블록의 최소 개수로 총 매크로블록 개수의 40% 정도의 개수를 가지며 MB_{max} 는 후보 셋의 매크로블록의 70% 정도의 개수를 가진다. TH_{max} 는 후보 셋의 매크로블록 개수 θ 가 MB_{max} 이상일 때 적용되는 임계값이며 이전 장면 전환이 발생하였을 때를 검출 하기 위하여 작

은 임계값을 주었다. TH_{min} 은 이전 장면과의 어느 정도 유사성을 가지는 비디오 셋도 검출 할 수 있게 큰 임계값을 주었다. 그림3에는 이웃된 P2,P3프레임에서 매크로블록의 개수를 검출 하는 과정을 보이고 있다

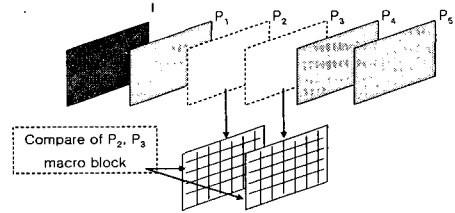


그림3. P프레임 매크로블록 검출

III. 실험 및 고찰

제안한 장면 전환 알고리즘의 성능을 평가하기 위하여 N=15 그리고 M=1의 GOP 구조로 압축된 3종류의 MPEG-4 비트스트림을 사용하였다. 테스트 비트스트림의 엔코딩 되기 전의 비디오 사이즈는 320x240이다. 알고리즘은 2단계 과정으로 구분 되어 처리 된다. 먼저, 프레임의 픽처 타입을 검출하고, 다음으로 장면 전환 검출의 기준이 되는 데이터를 추출하여 장면 전환 여부를 확인 한다. 즉, I 프레임의 경우에는 DC 영상을 복원하고, P 프레임의 경우에는 DCT 매크로블록의 개수를 검출한다. I 픽처 프레임은 GOP마다 1개씩 검출이 된다. 따라서, 첫번째 GOP에서 검출된 I 프레임과 두번째 GOP에서 I프레임을 검출하여 DC 영상을 복원한다. 각각의 GOP의 I픽처 프레임에서 복원 된 DC영상은 히스토그램 평활화를 거쳐 밝기 변화에 민감하지 않도록 처리를 한다. 복원된 DC 영상의 히스토그램 차를 연산하여 임계값이 0.6 이상인 경우 장면 전환 셋으로 검출 한다. P픽처 I 픽처 프레임의 사이에 연속으로 나타난다. I픽처 프레임에 이어서 입력되는 프레임을 시작으로 하여 모든 P픽처 프레임에 대하여 매크로블록의 개수를 연산한다

표1은 I픽처 프레임에 대하여, 표2는 P픽처 프레임에 대하여 실험영상을 적용하였을 때 장면 검출 결과를 보이고 있다. 표1에서 보는 바와 같이 I 픽처를 이용하여 비디오1, 비디오2, 비디오3에서 장면 전환 검출을 시도 하였을 때 98%, 92% 그리고 85%의 검출 비율을 보이고 있다. 비디오3 입력 영상에서 낮은 정확도를 보이고 있는 이유는 다른 두 개의 테스트 비디오에 비하여 빠르고 큰 움직임 특성이 발생하여 나타난 결과이다. 표2는 P 픽처 프레임에 매크로블록을 적용하여 장면 전환을 검출 하였을 때 결과를 보이고 있다. 실험 비디오에 대하여 비디오 셋 검출의 성공율은 98%, 96%, 93%의 결과를 보이고 있다. 이 때 임계값의 기준에 해당되는 매크로블록의 개수

가 전체 매크로블록의 70%인 경우이다. 임계값의 기준을 높일 경우 실제 장면 전환이 발생하지 않았지만 비디오 셋으로 검출 되는 정확도가 증가할 수 있다. 표3은 비디오 셋 검출 방법에 따른 검색 속도의 비교 결과 이다. 기법1은 압축된 비트스트림 데이터를 원이미지로 복원하고, 복원된 이미지의 히스토그램 차를 이용한 방법이다. 기법2는 I픽처 프레임의 DC 계수만을 이용한 방법이다. 기법3는 제안한 기법을 이용하였다. 비디오1을 입력 비트스트림으로 하여 1초 동안 비디오 셋의 검색성능을 측정하였다. 실험 결과 압축 데이터를 이용한 기법2, 3은 기법1보다 4~5배 빠른 결과를 보이고 있다

표 1. I 픽처 셋 검출

비디오/ 검출	검출	검출오류	전체 프레임수
비디오1	75	2	9000
비디오2	79	7	9000
비디오3	102	18	9000

표 2. P 픽처 셋 검출

비디오/ 검출	검출	검출오류	전체 프레임수
비디오1	156	3	9000
비디오2	180	6	9000
비디오3	264	17	9000

표 3. 검출 비교

비디오/ 기법	기법1	기법2	기법3
비디오	3.2(f/s)	15.6(f/s)	13.4(f/s)

IV. 결론

본 논문에서는 압축된 MPEG 비트 스트림을 디코딩 하지 않고 비디오 픽처 타입을 이용하여 빠른 비디오 셋 검출 기법을 제안하였다. 압축되어 있는 비디오 픽처 프레임을 사용하기 때문에 압축 비디오 데이터를 복원하기 위한 과정이 필요 없다. 이 방법은 I 픽처와 P 픽처로 구성된 GOP 구조의 비트 스트림에 적용이 가능하다. I 픽처 프레임에 대해서는 DCT 계수의 DC 항을 이용하여 DC영상을 복원하고 복원된 영상을 이웃된 GOP에서 복원된 I 픽처 프레임의 DC영상과 히스토그램 비교를 하여 셋 검출을 하였다. I 픽처 프레임을 참조하는 P 픽처 프레임에서 장면 전환이 발생하는 경우에는 P픽처 프레임의 매크로블록의 개수를 산출하여 임계값 이상인 경우 장면 전환 셋으로 검출 하였다. 이 결과 I 픽처 프레임에서는 셋 검출 비율이 85~98%의 검출 효과가 있었고, P픽처 프레임에서는 98~93% 장면

전환 비율을 보였다. 기존의 압축영상을 복원하여 비디오 셋을 검출하는 방법에 비해 제안한 기법은 4~5배 빠른 검색이 가능하였다. 그러나, 동일한 매크로블록으로 압축된 P 픽처 프레임의 경우 장면 전환으로 검출 되지 않는 경우가 발생될 수 있어 입력 영상에 대하여 체적화가 요구된다.

참고문헌

- [1] Rafael C. Gonzalez and Richard E. Woods, Digital Image Processing, Prentice-Hall, 2002.
- [2] Ricahrd S. Shernock Regis J. Crinon, Michael A. Dolan, John R. Mick Jr., Data Broadcasting, McGraw-Hill, 2001.
- [3] Ali M Dawood, Mohammad Ghanbari, "Scene Cut Detection form MPEG Video Stream Coded without B Pictures", ,pp.1645-1648,2001.
- [4] F. Arman, A. Hsu and M.-Y. Chiu, "Image processing on Compressed Data for Large Video Databases", ACMMultimedia'93,Jun.,pp.267-272,1993.
- [5] H. J. Zhang, C.Y. Low and S. W. Smolia, "Video Parsing and browsing using compressed data", Multimedia Tools and Applications, Vol. 1, No.1, Mar., pp.89-111, 1995.
- [6] Joung-Dae Jeon, Seong-II Cheong, Sang-Ryul Ryu, Sung-Ho Kim, "Cut Detection Using Macroblock and Frame Information", The Korea Information Science Society, vol28.No2, pp.271-273, 2001.10
- [7] In-Guk Jeong, Hwa-Sun Lee, Jong-Ho Nang, Gyung-Su Kim, Myung-Hwan Ha, Byung-Hui Jeong, "A New Shot Change Detection Scheme Using Color Histogram and Macroblock Information of MPEG Video Stream", The Korea Information Science Society, vol28. No 01, pp.418-420, 2001.4
- [8] Seong-Whan Lee, Young-Min Kim, Sung-Woo Choi, "Fast Scene Change Detection using Direct Feature Extraction from MPEG Compressed Videos", IEEE TRANSACTIONS ON MULTIMEDIA, VOL.2, NO.4, pp.240-254, 2000.12.
- [9] Young-Bin Kim, Kwang-Ryul Ryu, Robert J. Scabassi, "Changing Scene Detection using Histogram and Header Information of H.264 Video Stream", Proceedings of the Korean Institute of Maritime Information and Commucation Sciences Conference, pp.197-200, 2006. 06