

압축된 영상 시퀀스에서 내용 기반 색인을 위한 장면 분석 및 키 프레임 추출

오 상 현(吳相憲), 김 상 렬(金相烈), 김 주 도(金柱到), 이 근 영(李根泳)

성균관대학교 전기전자컴퓨터공학부

전화 : (0331) 290-7193 / 팩스 : (0331) 290-7180

The Scene Analysis and Keyframe Extraction for Content-Based Indexing on Compressed Image Sequence

Sangheun Oh, Sanglyeol Kim, Judo Kim, Keunyoung Lee

School of Electrical and Computer Engineering, Sungkyunkwan University

E-mail : sam8080@ece.skku.ac.kr

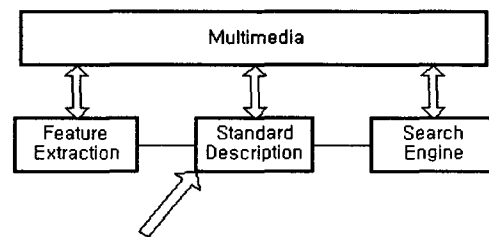
Abstract

In this paper, we proposed several scene analysis algorithms. These algorithms using image difference and histogram operate on the sequence of DC coefficient which is extracted from Motion JPEG or MPEG without full-frame decompression. Since DC sequence has the most information of full frame while it has reduced data. Experimental results show less than 1/64 of full frame analysing complexity and exactly analyze scene changes and extract key frames.

1. 서론

최근 멀티미디어 기술이 발전하면서 데이터의 저장에 대한 기술 뿐 만 아니라 정보내용에 기반한 색인의 요구가 증가하는 추세이다. 특히, 동영상과 같은 멀티미디어 데이터의 내용기반 검색을 위해 내용을 표현하는 방식에 관한 국제표준화작업을 MPEG-7: Multimedia Content Description Interface라는 이름으로 시작하였다[1]. 그림1은 MPEG-7과 관련된 정보처리 과정과 MPEG이 표준화 하고자 하는 범위를 개략적으로

보여준다. 이 중 특징 추출(Feature Extraction) 및 검색엔진(Search Engine)은 MPEG-7의 전처리 과정으로 신속하고도 정확한 처리 방식이 요구된다.



Scope of MPEG-7 standardization

그림 1. MPEG-7 표준화의 범위

본 논문에서는 계산시간을 단축하면서도 fade-in, fade-out 등과 같은 선형적 장면 전환 뿐만 아니라 panning과 같은 비선형적 장면전환을 검출하는 방법을 제안한다. 또한 검출된 장면전환에 대해 객체의 움직임에 무관한 키프레임 검출 방법을 제안한다.

2장에서는 기존의 비디오 데이터로부터 장면 전환을 추출하는 방법을 설명하며 3장에서는 이를 개선하기 위한 효율적인 장면전환 및 키프레임 추출 방법을 제

안한다. 4장에서는 실험 결과를 보이고 5장에서 결론을 맺는다.

2. 기존의 장면전환 추출방법

2.1 복원된 영상 시퀀스 장면전환 추출 방법

복원된 영상 시퀀스로부터 장면전환을 추출하는 대표적인 방법으로는 주로 화소단위 비교법, 유사율 측정법과 히스토그램 비교법 등이 있다. 화소단위 비교법은 연속된 프레임에서 각 프레임의 대응하는 화소의 SAD(Sum of Absolute Difference)를 비교하여 이 값이 임계값을 넘으면 장면전환으로 판단하는 방법이다. 히스토그램 비교법은 화소의 세기를 히스토그램으로 표현하여 히스토그램의 차로 유사도를 측정한다.

2.2 압축된 영상 시퀀스의 장면전환 추출 방법

기존의 MPEG으로 압축된 영상은 I프레임과 차영상의 DCT 계수값과 움직임 벡터를 가지고 있다. DCT 계수값은 복원 영상의 정보를 그대로 가지고 있으므로 신뢰할 수 있다. 압축된 영상 시퀀스의 장면전환 추출 방법으로는 움직임 벡터와 I와 P 프레임의 DC 계수값의 분산값을 비교하는 방법등이 있다.

Yeo 와 Liu의 방법[2]에서는 DC 계수 값만으로 시퀀스를 만들어 각 해당블록의 DC값의 SAD를 비교하여 프레임간의 유사성을 비교하는 방법을 사용하였다. 이 알고리즘은 영상복원의 필요가 없고 계산량을 64분의 1로 줄일 수 있는 장점이 있지만 급격한 변화와 점진적 변화를 각각 따로 검출해야 하는 단점이 있으며 점진적 변화의 경우 fade-in, fade-out과 같은 선형적 변화의 검출에는 유용하지만 panning과 같은 비선형적 변화의 검출에는 효과적이지 못하며 검출된 장면전환을 분석하지 못하는 약점을 가지고 있다.

3. 제안한 장면전환 추출 알고리즘

논문에서는 Yeo와 Liu의 방법[3]에서 제안한 일차 근사법을 이용하여 DC영상을 검출해 내고 이 중 B프레임을 뺀 I와 P프레임만을 사용하여 장면 변화의 검출에 사용하였다.

그림2와 그림3은 복원된 Suzie에서 각 프레임간의 SAD와 I와 P만의 DC 시퀀스를 이용하여 SAD를 구한 그래프이다. M=3인 경우 I와 P 프레임의 DC 시퀀스는 복원과정을 거치지 않고 프레임 수를 3분의 1로 줄이면서도 복원영상의 SAD의 성질을 그대로 가지고 있음을 알 수 있다.

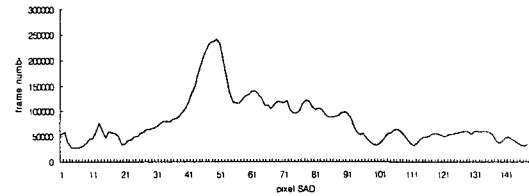


그림 2. 프레임간 pixel SAD(Suzie 150 프레임)

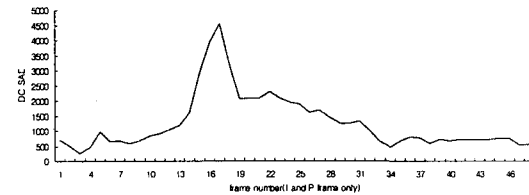


그림 3. I와 P 프레임간 DC SAD(Suzie 150 프레임)

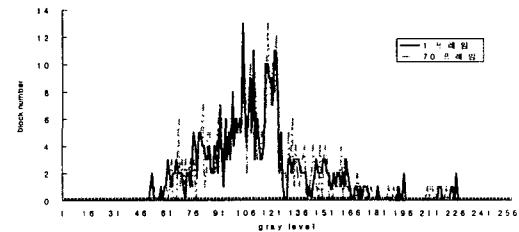


그림 4. Suzie 영상에서 첫 번째프레임과 70번째 프레임의 히스토그램(DC only)

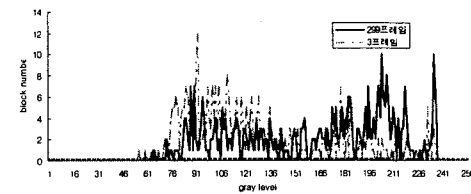


그림 5. Foreman 3번째 프레임과 299번째 프레임의 히스토그램(DC only)

장면변화의 검출에는 이러한 DC값의 SAD값의 합을 이용하여 유사성을 비교하여 검출해내고, 검출된 변화의 분석에는 DC값에서의 밝기 히스토그램을 이용하여 비교하였다.

그림4의 suzie 시퀀스의 경우 새로운 객체의 등장인 없는 경우로 처음과 70번째 프레임의 DC히스토그램이 유사한 것을 알 수 있고 그림5의 foreman 시퀀스의 경우 panning의 결과로 인한 다른 객체의 등장으로 인해 첫 프레임과 마지막 프레임의 DC히스토그램의 차이가 큰 것을 알 수 있다.

본 논문에서는 DC영상의 이러한 성질을 이용하여 압축된 영상의 복원과정을 거치지 않고 급격한 변화와

점진적 변화를 동시에 검출 및 분석하여 키프레임을 추출해 내는 방법을 제안한다.

3.1 장면 전환 및 키프레임 추출

급격한 변화는 장면전환에 의한 변화와 사진기의 플래시와 같은 섬광에 의한 변화로 크게 나뉘어진다.

X_i , ($i=1,2 \dots ,N$)을 DC 영상의 시퀀스라고 하고 그 차를 $D_i=d(X_i, X_{i+1})$ 라고 하면

$$d(X, Y) = f \sum_{k,l} |x_{k,l} - y_{k,l}|^\alpha \quad (1)$$

로 표현된다. 여기서 k, l 는 각 프레임의 블록을 나타내며 α 는 1을 쓴다.

H_i 를 DC 영상의 히스토그램이라고 할 때 그 차를 $HD_i=d(H_i, H_{i+1})$ 라고 하며 α 는 2를 쓴다. 이 때,

단계1. 크기가 $2m-1$ 의 양면 대칭의 sliding window 내에서 최대값을 가진 경우($D_i > D_j$, $j=i-m+1, \dots, i-1, i+1, \dots, i+m-1$)

단계2. 최대값이 두 번째로 큰 값의 n 배 이상의 값을 가지는 경우 조건을 만족하면 장면전환으로 정의하고 그렇지 않을 경우 단계3으로 간다.

단계3. $HD_{i-1}=(H_{i-1}, H_{i+1})$ 가 프레임 DC 개수의 h 배 이상이 되는 경우 장면전환으로 정의한다.

단계2는 플래시에 의한 급격한 변화를 장면전환으로의 검출을 배제하기 위한 조건이고 단계3은 점진적인 장면 변환 중에 생기는 급격한 장면 전환에 의해 두 번째로 큰 값이 최대값과 유사한 경우가 발생해서 장면전환의 검출에 실패하는 경우를 위한 조건으로 대부분의 장면전환의 경우 단계2에서 검출된다.

키프레임은 장면전환에 의한 다음장면의 시작 프레임을 키프레임으로 추출하고 다시 장면전환 발생시 장면 전환 바로 이전의 프레임과의 히스토그램을 비교하여 그 차이가 h 보다 큰 경우 키프레임으로 추출한다. 이는 다음의 점진적 장면변환에서 검출하지 못 할 정도의 객체의 느린 등장을 검출하기 위한 조건이다.

3.2 점진적 변화의 검출

DC 시퀀스에 의한 점진적 변화는 객체의 움직임에 의한 변화와 프레임내 객체들이 나타나고 사라지는 변화와 같은 경우에 발생한다.

점진적 장면 변환 검출에서는 급격한 변화에서 사용된 대칭적 sliding window 대신 i 번째 프레임을 중심으로 그 이전 $i-m+1$ 까지의 이전 시퀀스를 pre_sliding

window로, i 프레임 이후 $i+m-1$ 까지를 next_sliding window로 규정한 후 검출한다.

단계1. D_i 가 pre_sliding window에서의 최대값으로 pre_sliding window내의 최소값의 s 배 보다 큰 값을 가지며,

단계2. 최대값과 최소값의 차이가 t 보다 큰 경우의 프레임 i 를 점진적 변화의 시작점으로 잡는다.

단계3. 시작점이 검출된 상태에서 D_j 가 next_sliding window에서 최대값으로 next_sliding window 내의 최소값의 s 배 보다 작은 값을 가지며

단계4. 최대값과 최소값의 차이가 t 보다 작은 경우의 프레임 j 를 점진적 변화의 종결점으로 잡는다.

단계1, 3의 경우는 fade-in, fade-out과 같은 프레임간의 차이가 선형으로 변화하는 경우와 panning에서 보는 바와 같이 프레임간의 차이가 무질서하게 증가하는 비선형적 변화 등 두 가지의 변화를 모두 검출해 내기 위한 조건이다. 단계2, 4의 조건은 최소값에 비례한 최대값만으로 점진적 변화의 시작점을 검출할 경우 프레임간의 차이가 적은 window에서 작은 변화만으로 점진적 변화의 시작점으로 잡는 오류를 방지하기 위한 것이다.

3.3 점진적 장면 변화의 분석 및 키프레임 추출

객체의 움직임에 영향을 받지 않는 히스토그램 비교의 특징을 이용하여 점진적 장면 변화를 분석하여 키프레임을 추출해 낸다.

단계1. 점진적 장면변화의 시작 프레임을 i , 마지막 프레임을 j 라고 할 때 그 이전에 추출된 키 프레임과 점진적 변화의 j 프레임의 히스토그램을 비교하여 SAD가 h 미만이면 객체의 움직임에 의한 변화로 정의하고 종료한다.

단계2. h 이상인 경우 새로운 객체의 등장에 의한 점진적 장면 변화로 정의한다.

단계3. 그 이전에 추출된 키프레임의 히스토그램과 점진적 장면 변화로 검출된 프레임 내의 히스토그램을 비교하여 그 차값이 h 미만인 프레임은 키프레임 추출에서 제외한다.

단계4. j 프레임의 히스토그램을 단계3에서 제외되지 않고 남은 프레임의 히스토그램의 SAD를 비교, 값이 h 이상인 프레임을 key 프레임으로 결정 후, 이 프레임을 j 로 갱신하고 i 프레임까지 단계4를 반복

단계3, 4는 새로운 객체의 등장에 의한 점진적 장면 변화 내에서 중간에 나타났다가 사라지는 객체의 출현을 키프레임으로 추출해 내기 위한 조건이다.

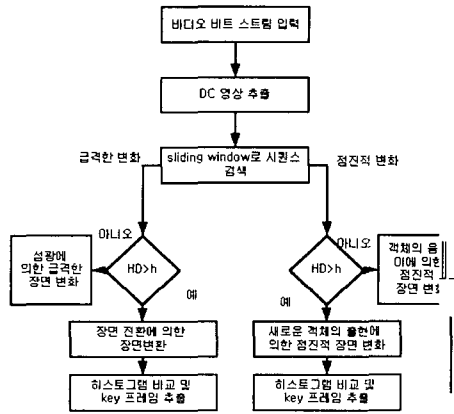


그림6. 제안한 방법의 블록 다이어그램

4. 실험결과

실험영상으로 176×144 Trevor, Container, News, Suzie, Salesman, Carphone, Foreman, Coast의 QCIF 시퀀스를 연결하여 총 2322 프레임 중 I와 P 프레임을 테스트 영상으로 사용하였다. 또한, 실험에 쓰인 파라미터로는 sliding window의 범위 $m=5$, 급격한 장면 전환 검출의 임계값 $n=3$, 점진적 장면 변화 검출의 임계값 $s=4$, $t=DC$ 개수의 n 배, 히스토그램 임계값 $h=DC$ 개수의 4배를 사용하였다.

실험결과를 살펴보면 전체 시퀀스에서 급격한 장면 전환으로는 영상시퀀스의 연결부분에서 7개, Trevor에서 1개, News에서 3개씩 모두 11개를 검출하였다. 점진적 변환으로는 Suzie(5프레임), Carphone(14프레임), Foreman(21프레임), Coast(6프레임) 시퀀스에서 각 하나씩 검출하였고 이 중 Trevor는 객체에 등장에 의한 장면전환으로 판명되어 3개의 키프레임을 추출하였고 Coast에서는 마지막 프레임에서 새로운 객체를 인식하여 총 15개의 키프레임을 검출하였다.

그림8의 (a)에서 (d)까지는 News 시퀀스에서 검출된 장면전환에서 추출된 키프레임이며 (e)부터 (h)까지는 실제로 panning을 하는 Foreman 시퀀스에서 검출된 점진적 변화 중 객체의 출현으로 판정되어 추출된 키프레임이다. 나머지 점진적 변화는 모두 객체의 움직임에 의한 변화로 판정되어 키프레임이 추출되지 않았

다.

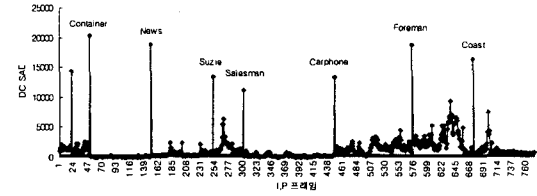


그림 7. 혼합영상에서의 DC SAD

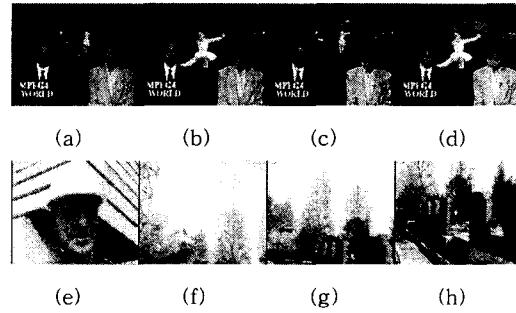


그림 8. news(a-d),foreman(e-f)

5. 결론

본 논문에서는 압축된 영상으로부터 DC값을 추출하여 복잡한 복원 과정 없이 장면 분석 및 객체에 근거한 키프레임을 추출하기 위한 방법을 제시하였고 실험 결과 DC값의 SAD와 히스토그램만으로 객체의 움직임에는 영향받지 않는 키프레임을 추출해 낼 수 있었다.

참고문헌

- [1] ISO/IEC JTC1/SC29/WG11, "MPEG-7 Context and Objective," MPEG98/N2207, Tokyo, Mar. 1998.
- [2] B.L.Yeo and B.Liu, "Rapid Scene Analysis on Compressed Video," IEEE Trans on Circuit and Systems for Video Technolgy, vol.5, no.6, pp. 533-544, Dec. 1995.
- [3] B.L.Yeo and B.Liu, "On the Extraction of DC Sequence from MPEG Compressed Video," Proceedings of ICIP, vol. 2, pp260-264, 1995.