

압축 영역에서 intra mode와 에지 방향성을 이용한 H.264 비디오 장면 전환 검출

Scene change detection using intra prediction mode and edge direction in H.264/AVC compression domain

홍보현*, 엄민영**, 최윤식***

Bohyun Hong and Minyoung Eom and Yoonsik Choe

Abstract - This paper presents a novel scene change detection method using intra prediction mode and edge direction in H.264/AVC. When scene change occurs, there are less temporal correlation between frames, most of macro-blocks encoded in intra mode. Using this property, the method calculates the percentage of intra mode blocks in each predictive frame in order to get candidates of scene change frame. To further find scene change, we obtain edge histogram of each candidates by using eight prediction direction of intra prediction mode in H.264/AVC. We detect scene change frames with l^1 -norm of edge histograms. The experimental results show that the method is efficient and robust.

Key Words : scene change detection, H.264/AVC, Intra prediction, Edge Histogram

1. 서론

최근 네트워크와 멀티미디어 기술의 발달에 따라 많은 양의 멀티미디어 정보가 생산되고 있다. 특히 효율적인 전송을 위한 압축 기술이 발달됨에 따라 압축된 멀티미디어 정보를 분석하고 분류할 필요가 있다. 이 중 압축 영역에서 장면 전환 검출은 가장 중요한 분야 중의 하나로 인식되고 있다.

MPEG-2 기반의 압축 영역에서 연구되어진 장면 전환 검출 방법들은 DCT 계수를 이용하여 DC-영상을 사용하는 방법[1], 순방향·역방향 MV를 사용하는 방법[2], Intra 프레임의 비트율을 사용하는 방법[3] 등이 있다. 그러나 이러한 방법들은 H.264에 적용하기에 적합하지 않다. H.264에서 DC 영상을 만들기 위해서는 거의 full decoding이 필요하며, 다중 참조 프레임을 이용하고 매크로블록(MB)의 크기도 다양하여 MV를 직접적으로 사용하기 어렵다. 마지막으로 Intra 프레임의 경우 Intra prediction mode 등으로 인하여 적은 비트율을 가지므로 장면 전환 검출에 사용되기 어렵다.

따라서 H.264의 특징을 이용한 장면 전환 검출 방법이 연구되어지고 있다. 이 중 인접한 intra 프레임의 intra prediction mode의 분포 차이를 척도로 압축 영역에서 장면 전환 검출하는 방법[4]이 있는데 이 방법은 intra 프레임만을 고려하기 때문에 정확한 장면 전환 검출이 어렵다. 따라서 본 논문에서는 intra frame 뿐 아니라 inter frame에서도 intra mode가 가지는 에지 방향성을 이용하여 H.264에 대해 압축 영역에서 고속 장면 전환 검출 알고리즘을 제안한다.

2절에서는 기존에 연구되어진 Intra prediction mode의 분포를 척도로 한 장면 전환 검출 방법[4]과 그에 대한 문제점을 기술하고, 3절에서는 각 intra prediction mode와 그 블록의 에지 방향성의 관계를 이용하여 새로운 장면 전환 검출 방법 소개한다. 마지막 4절에서는 제안한 장면 전환 검출 방법을 실험하고 기존의 장면 전환 검출 방법과 비교하여 그 효율성과 성능을 확인한다.

2. 기존의 장면 전환 검출 방법

Intra prediction은 H.264에 적용된 새로운 방법 중의 하나이다. Intra 프레임을 위한 예측은 16×16 과 4×4 로 나뉘어진다. 연속적인 Intra 프레임 사이에서 장면 전환이 일어날 경우 prediction mode의 분포에 큰 변화가 있을 것이다. 즉, intra prediction mode 분포의 차이는 영상 자체의 차이를 뜻하므로 prediction mode의 분포에 큰 변화가 있을 경우 장면 전환이 일어나는 것이다.[4]

i 번째 프레임의 j 번째 MB의 intra prediction mode를 나타내기 위하여 변수 $Mode_i^j$ 를 아래와 같이 정의한다.

$$Mode_i^j = \begin{cases} 0, & \text{for } 16 \times 16 \text{ prediction mode} \\ 1, & \text{for } 4 \times 4 \text{ prediction mode} \end{cases} \quad (1)$$

두 intra 프레임의 유사도 측정을 위하여 각 sub-block(한 프레임을 중복되지 않는 20개의 sub-block으로 나뉜다.)당 i 번째와 $i+1$ 번째 intra 프레임 사이의 정규화된 prediction mode의 분포를 나타내는 Δf_i^k 를 정의한다.

$$\Delta f_i^k = \frac{1}{|S^k|} \left| \sum_{j \in S^k} Mode_i^j - \sum_{j \in S^k} Mode_{i+1}^j \right| \quad (2)$$

위 식에서 S^k 는 k 번째 sub-block에 포함되는 MB의 집합

* 연세大學 전기전자學科 碩士課程

** 연세大學 전기전자學科 博士課程

*** 연세大學 전기전자學科 正教授·工博

※ 본 연구는 한국 과학재단 목적 기초 연구

(R01-2003-000-10412-0) 지원으로 수행되었음.

을 의미하고, $|S^k|$ 는 sub-block에 포함되는 MB의 개수를 의미한다.

장면 전환 결정 함수($d_{i,i+1}$)는 Δf_i^k 를 임의의 임계값(λ)과 비교하여 다음과 같이 정의한다.

$$d_{i,i+1} = \frac{1}{N_{sub}} \sum_k q_i^k, \quad q_i^k = \begin{cases} 1, & \Delta f_i^k \geq \lambda \\ 0, & \Delta f_i^k < \lambda \end{cases} \quad (3)$$

N_{sub} 는 한 프레임안의 sub-block의 개수이다. 여기서 $d_{i,i+1}$ 이 임의의 임계값(T')보다 클 경우 i 번째와 $i+1$ 번째 intra 프레임 사이에서 장면 전환이 일어났다고 결정된다.

3. 제안하는 장면 전환 검출 방법

기존의 H.264 에서의 장면 전환 검출 방법의 경우 intra 프레임만을 고려하기 때문에 정확한 위치의 장면 전환 검출이 어렵다. 본 논문에서는 intra 프레임 뿐 아니라 inter 프레임까지 고려하는 방법을 사용하여 좀 더 정확한 장면 전환 검출 방법을 제안하고자 한다. 이를 위해 본 논문에서는 한 프레임 당 intra mode 로 부호화 되는 MB의 백분율을 구하여 장면 전환 프레임의 후보를 추출한 후, intra prediction mode 의 8가지 방향성을 이용하여 각 후보 프레임의 에지 히스토그램을 구한다. 최종적으로 히스토그램의 l^1 -norm 을 이용하여 장면 전환 검출을 수행한다.

3.1 장면 전환 후보 검출

장면 전환이 일어날 경우 시간적 상관관계가 적기 때문에 다른 프레임을 참조하는 inter mode 보다 intra mode 로 부호화되는 MB의 수가 더 많다. 이러한 특성을 바탕으로 각 프레임당 intra mode 로 부호화 되는 MB의 백분율을 구하여 일정한 임계값보다 클 경우 장면 전환 후보 프레임으로 검출한다. 아래 표 1. 은 3415장의 프레임 중 48 번의 장면 전환이 일어나는 영상의 intra 와 inter 로 부호화 되는 MB의 평균 백분율을 나타낸 것이다.

	intra mode (%)	inter mode (%)
장면 전환 프레임	97.59	2.41
그 외	4.49	95.51

표 1. intra mode 와 inter mode 로 부호화 된 MB의 백분율

3.2 장면 전환 검출

3.2.1 intra prediction mode 를 이용한 에지 특성 추출

후보로 추출된 프레임의 에지 특성을 알아보기 위하여 intra mode로 부호화된 MB의 prediction mode를 살펴보고 그에 따라 에지 방향을 정의한다. H.264 의 intra prediction mode 는 그림1, 2와 같이 4가지의 16x16 prediction mode 와 9가지의 4x4 prediction mode로 나뉘어진다.[5] 16x16 prediction mode 로 선택되는 경우는 4x4 블록 단위로 나누어 동일한 mode를 지정해준다. 각 mode에 따른 에지 방향은 arctan 를 이용하여 표2 와 같이 정의한다.

위에서 정의된 에지 방향을 이용하여 8x8 블록 단위의 평균 에지 방향($e_{8 \times 8}$)을 구한다. $e_{8 \times 8}$ 을 구할 때 4x4 블록의 mode가 1일 경우 8x8 블록 안에 둔각의 에지 특성을 나타내

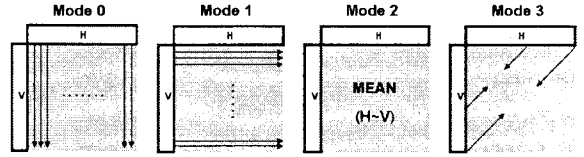


그림 1. 16x16 intra prediction mode

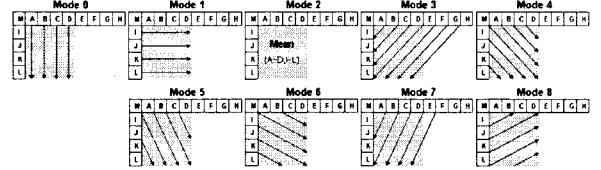


그림 2. 4x4 intra prediction mode

mode	0	1	2	3	4	5	6	7	8
에지 방향 (각도°)	90	0	non-edge	45	135	116	153	63	26

표 2. intra prediction mode의 에지 방향

는 4x4 블록이 많은 경우 180°로 에지 방향을 정의하고, 각각의 에지 특성을 나타내는 4x4 블록이 많은 경우 0°로 에지 방향을 정의한다. 또한 모든 4x4 블록이 non-edge의 특성을 가지는 경우에만 $e_{8 \times 8}$ 을 non-edge 로 결정하고, 그 외의 경우 평균을 취할 때 non-edge 의 개수는 고려하지 않는다. $e_{8 \times 8}$ 를 구하는 식은 다음과 같다.

$$e_{8 \times 8} = \begin{cases} \frac{1}{N} \sum_{i=1}^4 e_{4 \times 4 i}, & e_{4 \times 4 i} \neq non_edge \\ non_edge, & e_{4 \times 4 i, 2, 3, 4} = non_edge \end{cases} \quad (4)$$

여기서 $e_{4 \times 4 i}$ 는 8x8 블록에서 i 번째 4x4 블록의 에지 방향을 뜻하며, N 은 non_edge 가 아닌 4x4 블록의 개수를 의미한다.

이렇게 구해진 $e_{8 \times 8}$ 을 이용하여 그 블록이 가지고 있는 에지 영역을 정의한다. 그림 3에서는 블록을 네 가지 에지 영역으로 분류하고 8x8 블록의 평균 가중치가 1번 영역에 포함되면 vertical, 2번 영역에 포함되면 horizontal, 3번 영역에 포함되면 45° diagonal, 4번 영역에 포함되면 135° diagonal 이라고 판단한다.

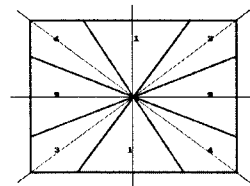


그림 3. 에지 영역

3.2.2 에지 히스토그램

MPEG-7 에지 히스토그램 기술자(EHD) 는 영상을 16개의 sub-image로 나누고, 각 sub-image 당 로컬 에지 분포를 나타내는 다섯 가지 종류의 에지 bin을 정의하여 에지 bin 80 개로 하나의 영상을 대표한다.[6] 이 논문에서는 위에서 설명

한 에지 추출 방법을 이용하여 80개의 bin을 추출한다.

3.2.3. 장면 전환 검출

정확한 장면 전환 프레임의 검출을 위하여 후보로 추출된 프레임들의 에지 히스토그램 차이 정보를 사용한다. 에지 히스토그램 bin의 차이 정보, EHBD (Edge Histogram Bin Difference)는 아래와 같이 정의된다.

$$EHBD(K) = \sum_{i=1}^{80} |EHB_i(K) - EHB_i(K-1)| \quad (5)$$

위 식에서 $EHB_i(K)$ 는 K번째 프레임에서 i번째 에지 히스토그램 bin을 나타낸다. $EHBD$ 가 임계값보다 크면 두 프레임간의 에지 변화가 큰 것이므로 장면 전환이 발생한 것으로 정의한다.

4. 실험 방법 및 결과

실험에 사용된 비디오 영상은 352×240 크기의 30프레임 당 1장의 intra 프레임으로 부호화된 H.264 비디오 영상(실험1)과 첫 번째 프레임만 intra 프레임이고 나머지는 inter 프레임으로 부호화된 H.264 비디오 영상(실험2)으로 구성되어 있다. 실험에 사용된 후보 추출의 임계값은 77%이며, 장면 전환 검출의 임계값은 740이다.

장면 전환 검출 성능의 평가 기준으로는 Recall 과 Precision 이 사용되었으며 해당 식은 아래와 같이 정의된다.

$$Recall = \frac{N_c}{N_c + N_m} \times 100$$

$$Precision = \frac{N_c}{N_c + N_f} \times 100 \quad (6)$$

여기서 N_c, N_m 그리고 N_f 는 각각 옳게 검출된 개수, 검출하지 못한 개수, 잘못 검출된 개수를 뜻한다. 식 (6)에 따라 N_m 와 N_f 이 작을수록 Recall 과 Precision 은 100%에 가까워지며 좋은 성능을 나타낸다고 할 수 있다.

표 3.4는 실험1에 대한 기존의 방법과 제안한 방법의 Precision 과 Recall 을 보여준다. 대부분의 실험 영상에 대해서 기존의 방법보다 제안한 방법의 Recall 이 월등하게 높은 반면 Precision 은 약간 떨어짐을 알 수 있다. 이것은 제안하는 방법이 장면 전환을 검출하지 못하는 경우는 적으나 장면 전환이 일어나지 않는 곳에서 장면 전환으로 검출 되는 경우가 있음을 보여준다. 이것은 후보를 추출하는 과정에서 intra 프레임이 모두 후보로 간주되어 잘못 검출된 경우가 생기기 때문이다.

실험2의 결과는 표5와 같다. 이와 같은 실험 영상은 주기적으로 intra 프레임이 존재하지 않기 때문에 기존의 방법으로 실험이 불가능하다. 제안한 방법의 Recall과 Precision을 살펴보면 첫 번째 실험에 비하여 Precision 값이 월등히 좋아짐을 알 수 있다.

5. 결론

본 논문에서는 각 프레임의 MB의 intra 부호화의 백분율을 최도로 장면 전환의 후보를 추출하고, 후보 프레임의 intra prediction mode의 방향성을 이용하여 얻은 에지 히스

	scene	Correct	Miss	False	Recall(%)	Precision(%)
1	26	17	9	3	65.38	85
2	23	5	17	1	22.73	83.33
3	14	9	5	11	69.23	45
4	10	3	7	3	33.33	50
5	7	2	5	0	33.33	100
6	6	5	1	1	83.33	83.33
평균					51.22	74.44

표 3. 기존의 방법의 Recall 과 Precision (실험1)

	scene	Correct	Miss	False	Recall(%)	Precision(%)
1	26	24	2	17	92.31	58.54
2	23	20	3	3	86.96	86.96
3	14	12	2	14	85.71	46.15
4	10	7	3	5	70	58.33
5	7	6	1	0	85.71	100
6	6	5	1	2	83.33	71.42
평균					84	70.23

표 4. 제안한 방법의 Recall 과 Precision (실험1)

	scene	Correct	Miss	False	Recall(%)	Precision(%)
1	26	23	3	13	88.46	63.89
2	23	20	3	1	86.96	95.24
3	14	10	3	4	71.43	76.92
4	10	10	0	0	100	100
5	7	7	0	0	100	100
6	6	3	3	1	66.67	80
평균					85.59	86.01

표 5. 제안한 방법의 Recall 과 Precision (실험2)

토그램의 차이 정보를 사용하여 장면 전환을 검출하는 방법을 제안하였다. 실험 결과 기존의 방법과 비교하여 제안한 방법이 정확한 검출을 가능케 함을 알 수 있다.

참 고 문 헌

- [1] Jungrim Kim, Sungjoo Suh, and Sanghoon Sull, "Fast Scene Change Detection for Personal Video Recorder", IEEE Transaction on, vol. 49, pp. 683-688, 2003
- [2] Dawood A.M., and Ghanbari M., "Scene Cut Detection from MPEG Video Stream Coded without B Pictures", IEEE International Conference on, vol.3, pp.1645-1648, 2001
- [3] Li H., Liu G., Zhang Z., and Li Y. "Adaptive Scene Detection Algorithm for VBR Video Stream", IEEE Transactions on Multimedia, vol. 6, pp. 624-633, 2004
- [4] Sung Min Kim, Ju Wan Byun, and Chee Sun Won, "A Scene Change Detection in H.264/AVC", LNCS 3786, pp. 1072-1082, 2005
- [5] Iain E. G. Richardson, "H.264 and MPEG-4 Video Compression", Wiley
- [6] B. S. Manjunath, Philippe Salembier, Thomas Sikora, "Introduction to MPEG-7 Multimedia Content Description Interface", Wiley