

# HDTV 인코더를 위한 적응적 율제어 방식

남재열\*, 강병호\*, 이호영\*\*, 하영호\*\*

계명대학교\*, 전자계산학과

대구시 달서구 신당동 1000 번지

jynam@kmucc.keimyung.ac.kr

경북대학교\*\*, 전자및 전기공학부

대구시 북구 산격동 1370 번지

## An adaptive rate control scheme for HDTV encoder

Jae Yeal Nam\*, Byung Ho Kang\*, Ho-Young Lee\*\*, Yeong-Ho Ha\*\*

Keimyung University\*, Taegu, Korea

Kyungpook National University\*\*, Taegu, Korea

### 요약

HDTV는 DTB에 비해 공간해상도가 4~6배 이상됨에 따라 처리할 샘플수가 증가하게 된다. 따라서 기존의 DTB 급으로 개발된 칩셋이나 코덱을 그대로 HDTV 코덱 개발에 적용하기가 어렵게 된다. 따라서 HDTV의 높은 샘플링 주파수 때문에 HDTV 코덱은 화면분할 방식과 같은 새로운 코덱 구조를 가지게 되며, 그 분할된 각각은 기존의 DTB 급 코덱으로 처리할 수 있게 된다. 이 때 각 대역마다 발생하는 부호화된 비트수가 다르기 때문에 일정한 전송율로 채널을 통하여 전송되기 위해서 버퍼 및 율제어 방식이 필요하게 된다. 본 논문에서는 각 대역별 영상의 특성과 버퍼 상황에 따라 적응적으로 율제어를 수행함으로써 화질의 균일성을 도모하고 부호화 효율을 증가시킬 수 있는 새로운 율제어 방식을 제안하였다. 또한 장면 전환이 발생했을 때에도 적응적으로 율제어를 수행할 수 있는 장면전환 검출 및 이를 이용한 적응적 율제어방식을 제안하였다.

### 1. 서론

디지털 비디오 신호의 전송은 매우 큰 데이터 전송폭을 요구하지만 실제로 사용 가능한 전송로는 매우 제한된 신호만이 전송되어질 수 있다. 따라서 비디오 신호의 압축이 절대적으로 필요하다. MPEG-2 비디오 부호화 방식[1]은 시간 중복성(Temporal redundancy)과 공간중복성(Spatial redundancy)을 제거함으로써 효과적으로 신호 대역폭을 줄일 수 있는데 여기에서 시간중복성은 움직임 보상된 예측 부호화 기법(Motion estimation and compensation)을 이용하여 제거하며, 공간중복성은 적응적 필드/프레임 DCT를 이용하여 제거할 수 있다. 또한 MPEG-2에서는 압축효율을 증대시키기 위하여 I, P, B-frame을 정의하고 전체 영상 시퀀스는 이러한 I, P, B-frame으로 구성되게 된다. 이렇

게 I, P, B-frame으로 차별화된 방식으로 부호화된 영상의 부호화 비트수는 각각 약 4:2:1 정도의 비율을 가지게 된다. 이 때 가변길이부호화(variable length coding)기법으로 코드가 할당됨으로 인해 생성되는 비트의 수가 시간에 따라 달라지게 된다. 발생하는 비트수는 가변비트율을 가지게 되는데 이러한 가변 비트율의 압축된 비트스트림을 고정 비트율의 채널로 전송하기 위해 비트스트림을 조절하는 방법이 필요하게 된다. 이러한 방법으로 채널로 전송되기 이전에 채널 버퍼를 두게 되며 이에 따른 율제어(Rate control)방법[2-12]이 필요하게 된다. 여기에 HDTV의 경우 1920x1024의 큰 해상도를 가짐으로 인해 프레임을 여러개의 밴드로 나누어서 각 서브밴드별로 DTB 급에 사용된 코덱을 이용하여 병렬처리를 하는 기법[14]이 제안되어지고 있는데 이러한 경우에는 율제어도 여러개의 서브밴드로 나누어진 프레임에 적합한 새로운 방법이 필요하게 된다. 이러한 각각의 서브밴드에 대해 적용되는 율제어 방법은 재구성된 영상의 화질에 커다란 영향을 미치게 되는데, MPEG-2 비디오 표준은 현재 압축된 비트스트림의 신택스와 복호화과정만을 정의하고 있기 때문에 인코더 설계시 어떠한 율제어 방식을 선택하느냐에 따라 코덱의 성능은 매우 달라질 수 있다. MPEG-2의 TM-6에서 이용된 율제어 방식은 3 단계로 수행되는데 장면전환이 발생되지 않는 동영상의 율제어에는 상당히 좋은 성능을 보이나, 장면전환이 발생할 경우에는 잘 처리하지 못하는 단점이 있다.[1]

본 논문에서는 Local variance를 이용한 장면전환 검출방안을 HDTV 영상에 적용하는 것을 제시하였다. 제안된 장면전환 검출방식은 모두 매크로블럭단위로 수행되기 때문에 한 프레임내에서의 국부적인 변화정보를 이용함으로써 보다 정확하게 장면전환을 검출할 수 있다. 또한 Local variance를 이용한 장면 전환 검출시 수행된 연산 결과를 율제어부에서 사용하며, 한 프레임을 여러개의 서

브랜드로 분할하여 각각의 서브밴드에 대해 공간 활동도에 따라 목표비트를 차등할당하여 단순히 서브밴드로 분할했을 경우에 비해 더욱 우수한 PSNR 값을 얻어내는 방법을 제시하고 있다.

## 2. 제안된 장면전환 검출 및 적응적 율제어 방식

### 2.1 장면전환 검출 방식

El-sakka 등은 영상구역분리(Segmentation)를 위한 새로운 척도로 국부분산을 제안하였는데[13], 본 연구에서는 이 국부분산을 동영상의 장면전환 검출에 이용하는 새로운 장면전환 검출방식을 제안하였다. 두 연속한 영상간의 밝기 신호성분의 국부분산은 아래와 같이 정의된다.

$$Lmean(i, j) = \frac{1}{4} \{x(i+1, j) + x(i-1, j) + x(i, j+1) + x(i, j-1)\} \quad (2-1)$$

$$Lvar = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \{x(i, j) - Lmean(i, j)\}^2 \quad (2-2)$$

일반적으로 분산은 영상의 활동도로 많이 이용하지만 국부영상의 변화도(활동도)를 제대로 검출하지 못한다. 국부분산이 영상의 국부적인 활동도를 더 정확하게 측정할 수 있는 성질을 이용한 새로운 장면전환검출 방식을 제안하였다.

이 방식의 경우 각 매크로블럭의 국부분산을 구하여 이웃한 영상간에 대응하는 매크로블럭간에 그 값의 차를 구한다. 장면전환이 일어날 경우 국부분산값이 크게 달라진다는 성질을 이용한 방식이다. 이 국부분산의 차가 어떤 정해진 문턱값을 넘어서는 매크로블럭의 갯수가 일정 갯수를 넘어설 경우 장면전환이 일어난 것으로 처리했다. 이때 매크로블럭의 크기인 16 x 16 전체에 대해서 국부분산값을 구하며 이렇게 구해진 값은 장면전환 검출에서도 사용하고 또한 율제어부의 적응적 양자화단계에서도 다시 이용함으로써 장면전환부와 율제어부에서 이중으로 계산하는 것을 피함으로써 계산량 및 H/W 복잡도를 줄일 수 있다. 이 방식에 대한 순서도는 (그림 1)에 나타내었다. (그림 1)에서 Local-var는 국부분산, NB\_Local\_var는 국부분산의 차가 임계치를 넘는 매크로블럭의 갯수를 나타내며, Th1 및 Th2는 임계치를 나타낸다.

### 2.2 적응적 율제어 방식

여기에서 제안하고 있는 율제어 방식은 MPEG-2 TM6 와 마찬가지로 3 단계로 구성되어 있으나 첫번째 단계와 세번째 단계를 변경하여 MPEG-2 TM6 방식을 개선하였다. 첫번째 단계에서 global complexity measure( $X_i, X_p, X_b$ )는 각 영상종류에 따라 다음과 같이 정의된다.

$$X_i = MS_i \cdot Q_i \cdot L_i \quad (2-3)$$

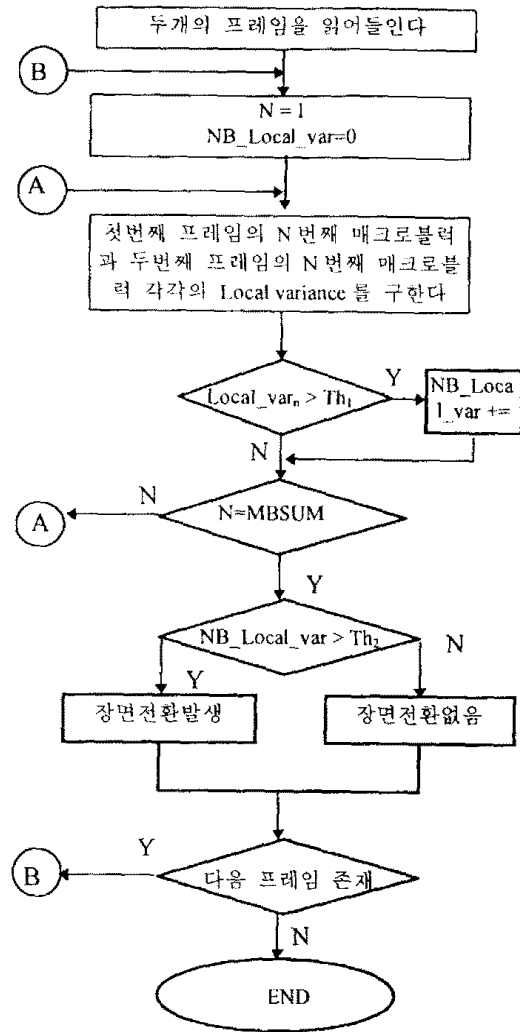
$$X_p = MS_p \cdot Q_p \cdot L_p \quad (2-4)$$

$$X_b = MS_b \cdot Q_b \cdot L_b \quad (2-5)$$

$$MS_i = S_i / MB\_SUM \quad (2-6)$$

$$MS_p = S_p / MB\_SUM \quad (2-7)$$

$$MS_b = S_b / MB\_SUM \quad (2-8)$$



(그림 1) 국부분산을 이용한 장면전환 검출방식의 순서도  
위에서 MB\_SUM은 한 영상내의 매크로블럭의 갯수를 의미하는데 1920 x 1024 에서 7680 이 된다.

$L_i$  = I-영상의 평균 국부분산값

$L_p$  = P-영상의 평균 국부분산값

$L_b$  = B-영상의 평균 국부분산값

초기값은 다음과 같다.

$$X_i = 310 * bit\_rate / 7680 \quad (2-9)$$

$$X_p = 120 * bit\_rate / 7680 \quad (2-10)$$

$$X_b = 80 * bit\_rate / 7680 \quad (2-11)$$

TM6 에서는 세번째 단계에서 매크로블럭을 4 개의 luminance frame-organised subblocks 과 4 개의 luminance field-organised subblocks 으로 나누어서 그 중 가장 작은 분산값을 공간활동도(Spatial activity)로 사용하여 2 단계에서 구해진 기준 양자화 파라미터를 정규화해서 mquant 값을 구하도록 되어있다. 본 논문에서는 세번째 단계에서 매크로블럭을 8 개의 서브블럭으로 나누는 대신 장면전환검출에서 16x16 매크로블럭에 대해 구했던 국부분산값을 공간활동도로 이용하였다.

HDTV 방식에서는 현재 기술수준으로는 1920 x 1024 의 해상도를 가진 영상을 실시간으로 부호화하는 것이 어렵기 때문에 하나의 프레임을 여러개의 서브밴드로 나누어서 압축부호화를 병렬로 수행되도록 고선명 TV 인코더 시스템이 구현될 필요

가 있다. 이때, 화면을 분할하여 처리하게 되면 기존의 율제어 방법을 그대로 적용하기가 어려우므로 기존의 율제어 방법을 개선하거나 새로운 율제어 방법을 사용하여야 하는데 본 논문에서는 각 서브밴드에 대해 각각의 버퍼를 따로 가지도록 하며, 각각의 버퍼에 대해 버퍼의 컨트롤을 따로 하도록 하고 있다. 이때 각 버퍼에 대한 목표비트는 식(2-12)-(2-14)를 이용하여 구한다.

$$T_{in} = T_i \times \frac{avg\_act_n + k}{\sum_{j=1}^m avg\_act_j + 4k} \quad (2-12)$$

$$T_{pn} = T_p \times \frac{avg\_act_n + k}{\sum_{j=1}^m avg\_act_j + 4k} \quad (2-13)$$

$$T_{bn} = T_b \times \frac{avg\_act_n + k}{\sum_{j=1}^m avg\_act_j + 4k} \quad (2-14)$$

where, n: n 번째 sub\_band  
 m: 한 프레임 내의 sub\_band 의 갯수  
 k: 80

장면전환이 일어날 경우 목표비트는 수정이 되는데 여기에서는 [12]에서 제안된 방식과 비슷하게 Sub GOP 단위로 수정하는 방식을 이용하였다. 다음 Sub GOP 에 장면전환이 있는 것으로 예측이 되면 현재 Sub GOP 에 할당되는 비트수를 줄여서 장면전환된 영상에 절약한 비트수 만큼 할당을 더해준다. 따라서 장면전환이 있는 영상의 화질저하를 줄여주게 된다.

### 3. 실험 결과 및 고찰

본 논문에서는 제안한 방식의 모의실험을 위하여 fashion show, locomotive, fruits 의 세가지 HDTV 영상을 사용하였다. 이러한 영상들을 임의로 섞어서 장면전환이 IGOP 당 1 회 정도가 있도록 시퀀스를 재구성하여 40 프레임짜리 시퀀스 2 가지를 만들어 사용하였다. 본 연구에서 이용한 실험환경은 (표 1)에 나타내었고, 장면전환검출 및 적응적 율제어 방식의 성능을 확인하기 위해 여러 종류의 시퀀스를 조합하여 hd\_temp1, hd\_temp2 의 두가지 재구성된 시퀀스를 만들었으며 그 구성도를 (표 2)와 (표 3)에 나타내었다. 성능의 객관적인 비교는 PSNR 및 주관적 화질 평가를 병행하였다.

#### 3.1 장면전환 검출

국부분산을 이용하여 hd\_temp1, hd\_temp2 의 2 가지 영상 시퀀스에 대하여 장면전환 검출 실험을 수행하였다. 지면 관계상 hd\_temp1 시퀀스와 hd\_temp2 시퀀스에 대한 결과만을 (그림 2)와 (그림 3)에 나타내었다. (그림 2)에서 장면전환이 발생한 6~7, 12~13, 15~16, 21~22, 24~25, 30~31, 36~37 영상들에서만 상당히 높은 수치를 기록하고 나머지 영상들에서는 낮은 수치를 기록하고 있으므로 정확하게 장면전환을 검출할 수 있음을 보여준다. 그리고, (그림 3)의 hd\_temp2 에서는 모든 프레임에서 장면전환이 발생하고 있는데, 그래프를 보면 한 프레임에 대해 한 프레임 내의 매크로블럭의 갯수인

7680 의 절반인 3840 을 훨씬 초과한 값이 모든 프레임에 대해서 나타나고 있으므로 정확하게 장면전환을 검출하고 있음을 보여준다.

표 1. 모의 실험 환경

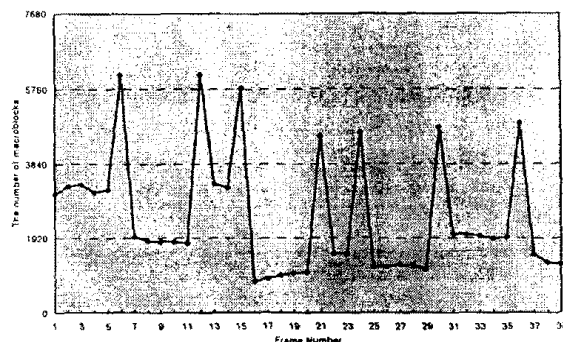
영상의 크기	1920 x 1024
기본 부호화 알고리즘	MPEG-2 MP@HL
GOP 구조	N=12, M=3
비트율	20Mbps

표 2. 재구성된 시퀀스 hd\_temp1 의 구성도

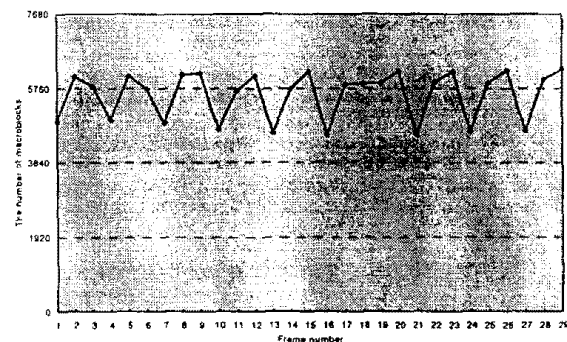
hd_temp1	
Frame Number	Sequence Name
1~6	loco
7~12	fshow
13~15	loco
16~21	fruits
22~24	fshow
25~30	fruits
31~36	fshow
37~40	fruits

표 3. 재구성된 시퀀스 hd\_temp2 의 구성도

Sequence Name	Frame Number
fruitsm	1, 4, 7, 11, 14, 17, 19, 22, 25, 28
fshow	2, 5, 8, 10, 13, 16, 21, 24, 27, 30
loco	3, 6, 9, 12, 15, 18, 20, 23, 26, 29



(그림 2) 국부분산을 이용한 hd\_temp1 시퀀스에 대한 장면전환 검출 실험결과

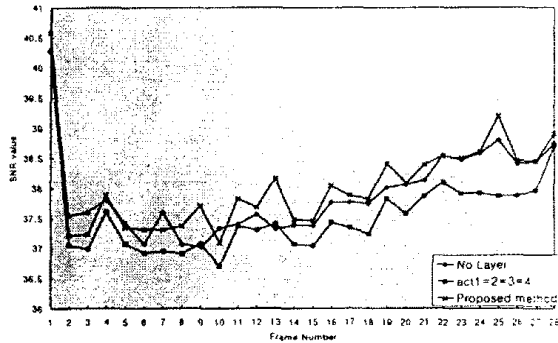


(그림 3) 국부분산을 이용한 hd\_temp2 시퀀스에 대한 장면전환 검출 실험결과

#### 3.2 율제어 방식

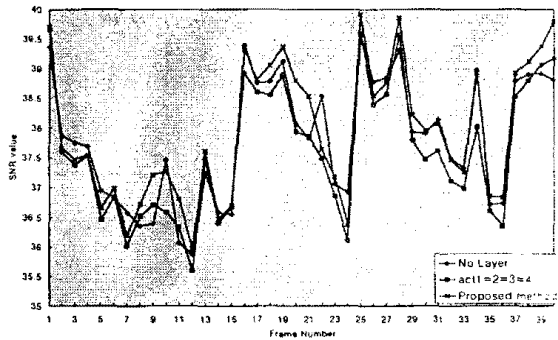
다음으로 본 논문에서 제시한 새로운 율제어 방식을 이용하여 나온 결과를 PSNR 그래프를 이용하여 나타내었다. 본 연구에서 사용한 S/W 코덱은 기본적으로 MPEG-2 MP@HL 을 기본 방식으로 사

용하였다. 예측방식은 적응적 필드/프레임 예측을 사용하였으며 프레임내 부호화 방식은 적응적 필드/프레임 DCT 및 양자화 방식을 사용하였다. 본 연구의 기본 방향이 프레임을 몇개의 서브밴드로 나누었을 때의 율제어 방법에 대해 제안하고 있기 때문에 네개의 서브밴드로 나누어서 각각의 서브밴드에 대해 같은 목표비트를 할당한 방식과 본 논문에서 제안하고 있는 방식인 각각의 서브밴드에 대해 차등적인 목표비트를 할당한 방식을 비교하여 그래프로 나타내었다. 부호화 비트율은 HDTV 영상을 부호화할 때 일반적으로 많이 이용하는 20Mbps를 이용하여 실험하였다. (그림 4)에서 보듯이 장면전환이 없는 경우도 제안한 방식이 보다 우수한 성능을 보여줌을 알 수 있다.



(그림 4) 장면전환이 없는 fashion show 시퀀스에 적용한 결과

(그림 5)는 장면전환이 있는 hd\_tmp1 동영상에 적용한 실험 결과이다. 실험 결과에서 볼 수 있듯이 제안한 방식은 장면전환이 일어났을 때 이를 정확히 검출하고 장면전환 발생이전 영상의 부호화시 절약한 비트를 장면 전환 영상에 재분배함으로써 PSNR 및 주관적 화질의 균일성을 도모할 수 있다.



(그림 5) 장면전환이 있는 hd\_tmp1 에 적용한결과

#### 4. 결론

본 논문에서는 국부분산을 이용한 HDTV 영상에서의 새로운 장면전환 검출 방식을 제안하였고, 이를 율제어부분에 적용한 비트 할당방식과 장면 전환시 처리 방법을 제안하였으며, HDTV 영상의 효율적인 인코딩과정을 위해 몇개의 서브밴드로 나누었을 경우의 율제어 방식을 제안하였다. 시뮬레이션 결과 제안된 장면전환검출방식은 동영상의 장면전환을 정확히 검출하였으며, 제안된 새로운 율제어방식은 기존의 MPEG-2 TM6 율제어 방식에 비하여 장면전환이 있을때와 없을때 모두 우수한 율제어 성능을 보여주었다. 또한 네개의 서브밴드

로 나누어서 부호화할 경우 각 밴드에 대해 영상의 국부활동도에 따라 비트를 할당함으로써 화질의 향상을 이루었다.

#### 5. 참고문헌

- [1] ISO/IEC/JTC1/SC29/WG11 MPEG93/457, 1993
- [2] Limin Wang, "Rate control for MPEG video coding," SPIE VCIP Vol.2501 pp.53-64, 1995
- [3] Limin wang, "Bit rate control for Hybrid DPCM/DCT video codec," IEEE. Trans. CSVT Vol.4, No.5, Pp509-517, Oct.1994
- [4] 이우용, 임경원, 조현덕, 최병선, 나종범, "MPEG 부호화기를 위한 효과적인 전송률 제어 알고리즘," 제 8 회 신호처리 합동 학술대회 논문집, pp.932-935, 1995
- [5] Seungkwon Paek, Jungsuk Kang, and Yangsuk Seo, "Rate control strategy based on human visual sensitivity for MPEG video coder," SPIE VCIP Vol. 2308 pp.322~330, 1994
- [6] Ishwar K.Sethi and Nilesh Patel, "A statistical approach to scene change detection," SPIE VCIP Vol. 2420, pp329~338, 1995
- [7] Seong Hwan Jang and Seop Hyeong Park, "An adaptive rate control algorithm for DPCM/DCT hybrid video codec adopting bi-directional prediction," SPIE VCIP Vol.2094, pp1237~1248, 1993
- [8] 이영선, 박상규, 장현식, 김진웅, "MPEG-2 인코더 시스템의 VBV 만족을 위한 실버퍼 제어 방식에 관한 연구," 제 8 회 신호처리합동학술대회 논문집 제 8 권 1 호, pp.936-939, 1995
- [9] 이영선, 호요성, 남재열, "장면 전환시 비트율 제어 방식 연구," 제 7 회 신호처리 합동 학술대회 논문집, pp.113-116, 1994
- [10] P.J. van der Meer, R.L. Lagandijk and J.Biamond, "A constant quality MPEG coder with bit stream shaping and peak rate control," SPIE VCIP Vol.2727, PP.145-153, 1996
- [11] C.Horne and A.Puri, "Video coding with adaptive quantization and rate control," SPIE VCIP vol.1818, pp.798-806, 1992
- [12] Jae Yeal Nam, Sang Gyu Park, Young Sun Lee, Sang Mi Lee, Chieteuk Ahn, and Sung Woong Ra, "An adaptive rate control scheme with scene change detection", Proceeding of the International Workshop on HDTV'94, pp.99-107, 1995
- [13] Mahmoud R. El-Sakka and Mohamed S. Kamel, "A segmentation criterion for digital image compression," Proc. Of IEEE ICIP'95, pp.2551-2554, 1995.
- [14] 안치득, 양재우, "고선명 TV 코덱 시스템," HCI'96 학술발표대회 논문집, pp.271-279,1996