

HEVC의 계층적 부호화 구조를 고려한 LCU 단위의 비트율 제어 기법

박 동 일^{a)}, 김 재 곤^{a)†}, 임 성 창^{b)}, 김 종 호^{b)}, 김 휘 용^{b)}

LCU-Level Rate Control for HEVC Considering Hierarchical Coding Structure

Dong-Il Park^{a)}, Jae-Gon Kim^{a)†}, Sung-Chang Lim^{b)}, Jongho Kim^{b)}, and Hui-Yong Kim^{b)}

요 약

본 논문에서는 현재 표준화가 진행중인 HEVC(High Efficiency Video Coding)의 고정 비트율(CBR: Constant Bitrate) 부호화를 위한 비트율 제어(rate control) 기법을 다룬다. 기존의 H.264/AVC의 비트율 제어 기법에서는 계층적 부호화 구조의 특성을 고려하지 못한 제어 기법을 사용하여 정확한 비트율 제어에 어려움이 있다. 현재 HEVC의 참조 소프트웨어인 HM(HEVC test Model)에는 CBR 부호화를 위한 비트율 제어 알고리즘이 제공되지 않은 상태이며, 추후 비트율 제어 모듈이 추가될 것으로 예상된다. 이때 HM의 공통 실험조건에 정의된 임의접근(Random Access: RA) 모드의 계층적 부호화 구조에서의 보다 정확한 비트율 제어가 요구된다. 따라서, 본 논문에서는 계층적 부호화 구조에서의 효율적인 비트율 제어를 위해 2차 비트율-왜곡(Rate-Distortion: R-D) 모델 기반의 시간계층 및 프레임 타입에 따른 비트율 특성을 반영한 비트율 제어 기법을 제시한다. 또한, 제어 단위에 따른 비트 변동율과 화질 손실 간에 발생하는 상충관계를 고려하여 비트율 제어 단위를 프레임 단위와 HEVC의 기본 부호화 단위(Coding Unit: CU)로 설정하고 모의실험을 통해 계층적 부호화 구조를 위해 제시된 기법의 비트율 제어 성능을 확인한다.

Abstract

In this paper, a method of rate control for constant bitrate (CBR) coding of High Efficiency Video Coding (HEVC) is addressed. The existing rate control of H.264/AVC may not provide exact rate control in the case of hierarchical coding structure since it doesn't consider the characteristics of the hierarchical coding structure. It is expected that a rate control is added to the reference software called HM for CBR encoding in the near future. More accurate rate control may be required in a hierarchical structure of random access (RA) mode defined in the common test condition of HM. In this paper, we propose a method of rate control based on quadratic Rate-Distortion (R-D) model considering temporal layers and frame types in hierarchical coding structure for efficient rate control. In the consideration of the trade-off relationship between the bit fluctuation and the average PSNR, both of frame and coding unit (CU) are set as the basic unit of rate control. The performance of the proposed rate control method is verified by simulations along with the trade-off relationships for the both cases of basic unit.

Keywords: Rate Control, HEVC, Hierarchical Coding Structure, LCU

a) 한국항공대학교 항공전자 및 정보통신공학부

School of Electronics, Telecommunication & Computer Engineering, Korea Aerospace University

b) 한국전자통신연구원 방송융합미디어연구부 실감미디어연구팀

Realistic Media Research Team, ETRI

† 교신저자 : 김재곤(jgkim@kau.ac.kr)

※ 본 연구는 방송통신위원회의 ETRI 연구지원 사업[KCA-2011-(11921-02001), 무안경 다시점 3D 지원 UHDTV 방송 기술개발] 및 지식경제부 및 정보통신산업진흥원의 대학 IT 연구센터 지원사업[NIPA-2011-(C1090-1111-0001)]의 지원으로 수행되었음.

· 접수일(2011년7월18일), 수정일(2011년8월22일), 게재확정일(2011년8월22일)

1. 서론

ISO/IEC MPEG과 ITU-T VCEG의 공동 작업반인 JCT-VC(Joint Collaboration Team on Video Coding)에서는 초고해상도(UHD: Ultra HD) 비디오 부호화를 주 목표로 한 HEVC(High Efficiency Video Coding) 표준화^[1]를 진행하고 있다. 지난해 10월 광저우에서 개최된 3차 회의에서는 HEVC의 참조 소프트웨어인 HM(HEVC test Model) 1.0에 포함될 기술^[2]이 결정되었고, 올해 3월 제네바 5차 회의에 걸쳐 HM 3.0^[3]에 포함될 기술이 채택되었다. HEVC의 표준화 단계에서는 공통 실험조건^[4]을 명시하고 HM의 부호화 모드로 Intra-only, 저지연(Low-Delay: LD), 그리고 임의접근(Random Access: RA)의 3가지 모드를 정의하고 있다. 한편, HM에서는 아직 고정 비트율(CBR: Constant Bitrate) 부호화를 위한 비트율 제어를 제공하고 있지 않으며, 추후 비트율 제어 모듈이 추가될 것으로 예상된다.

본 논문에서는 HEVC의 공통 실험조건에서 명시하는 RA 부호화 모드에서의 비트율 제어 기법을 제안한다. RA 모드는 방송 및 저장 응용 등에서 사용될 수 있도록 주기적으로 I-프레임을 삽입하는 것을 특징으로 하며, LD 모드와 달리 역방향 예측이 허용되는 특징을 갖는다. 따라서 RA 모드에서는 양방향 예측을 포함한 계층적-B(Hierarchical-B) 구조를 사용할 수 있으며, 이를 통해 시간적 스케일러블리티를 지원할 수 있을 뿐만 아니라 더 높은 부호화 효율을 얻을 수 있다.

기존 H.264/AVC^[5]의 JM(Joint Model)^[6]에는 계층적 부호화 모드를 포함한 비트율 제어 기법^{[7][8]}이 제안 및 구현되어 있다. [7]은 IPPP 구조와 비참조 프레임을 사용하는 구조에 적합하도록 제안되어 계층적 부호화 구조를 고려하지 않고 버퍼 레벨과 프레임 복잡도만을 고려하여 비트할당을 하고 있다. [8]에서는 계층적-B 구조를 고려한 기법을 제시하였지만, P-프레임들의 부호화 결과와 현재 시간계층을 이용하여 I-와 B-프레임의 양자화 파라미터(Quantization Parameter: QP)를 결정하는 방식을 사용하므로 초기 QP에 따라 영향을 많이 받으며 정확한 비트율 제어가 어렵다. [9]는 H.264/SVC의 스케일러블 부호화 구조를 고려한

것으로 시간계층 별 비트할당 방식을 제시하였고, [10]은 계층적 부호화 구조를 고려한 개선된 비트율-양자화 모델을 제시하고 시간계층 및 프레임 타입을 반영하여 비트율 제어의 정확성을 높였다.

본 논문에서는 HEVC의 계층적-B 구조에서의 효율적인 비트율 제어를 위해 모든 시간계층과 프레임 타입에 대해 2차 R-D 모델을 적용함으로써 각 시간계층과 프레임 타입의 특성이 반영된 QP를 결정하는 비트율-제어 기법을 제시하고 그 결과를 확인한다.

또한, JM에서는 비트율 제어를 수행하는 기본적인 단위로 기본단위(Basic Unit: BU)를 정의하고 있으며, 이 기본단위는 한 프레임으로부터 매크로블록(MB)까지 설정이 가능하다. 이 기본단위는 움직임의 비트율 변동과 PSNR 화질간의 상충관계를 고려하여 프레임, MB 또는 그 사이의 적절한 단위로 설정할 수 있도록 제안^[7]되었다. 즉, BU를 하나의 프레임으로 설정할 경우, 높은 PSNR과 프레임 간 높은 비트율 변동을 보이는 반면, MB 단위의 제어는 낮은 비트율 변동과 약간의 PSNR 손실을 갖는다. 따라서, 효율적인 비트율 제어를 위해서는 적절한 BU의 설정이 요구된다.

현재 HEVC에서는 초고해상도(Ultra HD: UHD) 비디오의 효율적인 부호화를 위해 기존의 16x16 크기의 MB를 확장한 다양한 블록 크기를 갖는 CU(Coding Unit) 구조^[11]를 정의하고 있다. CU는 적용 범위에 따라 최대 블록 크기와 최대 깊이를 이용하여 16x16 이상의 블록 크기도 지원 가능하지만 복잡도와 성능에 따라 현재 최대 크기로 LCU(Largest Coding Unit)를 64x64 크기로 설정하여 표준화가 진행 중이다. 따라서, 본 논문에서는 비트율 제어를 위한 BU를 프레임 및 LCU로 제안 및 구현하여 그 결과를 확인하고자 한다.

본 논문의 구성은 다음과 같다. 제 2 장에서는 기존 H.264/AVC의 비트율 제어 기법에 대해 기술하고, 제 3 장에서는 HEVC의 계층적-B 부호화 구조를 위해 제안한 비트할당 및 QP 결정 기법을 기술하고, 제 4 장에서는 제안한 비트율 제어 실험결과를 보이고, 마지막으로 제 5 장에서 결론을 맺는다.

II. H.264/AVC 비트율 제어 기법

본 장에서는 H.264/AVC의 참조 소프트웨어인 JM에 구현되어 있는 비트율 제어 방식에 대해 기술한다. JM은 가변 비트율(VBR: Variable Bitrate)과 고정 비트율(CBR: Constant Bitrate) 방식을 모두 지원하지만, 본 논문에서는 CBR 방식에 대해서만 논한다.

1. 프레임 목표 비트량 할당

H.264/AVC의 비트율 제어는 크게 GOP(Group of Picture), 프레임, 기본단위 레벨 순으로 진행된다. 기본적으로, GOP에 할당되는 비트량은 GOP 내에 포함되는 모든 프레임들에 대한 할당량이며, 한 프레임이 부호화된 이후마다 갱신된다. 이 때 GOP에 할당된 남은 비트량에 대해 복잡도와 버퍼 레벨을 고려하여 P-프레임에 대한 비트할당을 한다. 이 때, 가상버퍼 점유량($B_c(n_{i,j+1})$)과 목표 버퍼레벨 ($Tbl(n_{i,j+1})$)은 다음 식 (1)와 (2)과 같이 유동적으로 변하는 유동 트래픽 모델(Fluid flow traffic model)^[7]을 따른다.

$$B_c(n_{i,j+1}) = B_c(n_{i,j}) + b(n_{i,j}) - \frac{u(n_{i,j})}{F_r} \quad (1)$$

$$Tbl(n_{i,j+1}) = Tbl(n_{i,j}) - \frac{Tbl(n_{i,2})}{N_p - 1} + \frac{\widetilde{W}_p(n_{i,j}) \times (N_B + 1) \times u(n_{i,j})}{F_r \times (\widetilde{W}_p(n_{i,j}) + \widetilde{W}_b(n_{i,j}) \times N_B)} - \frac{u(n_{i,j})}{F_r} \quad (2)$$

식 (1)에서 $b(n_{i,j})$ 는 i 번째 GOP 내의 j 번째 프레임에 대해 부호화된 비트량이고, $u(n_{i,j})$ 는 채널 대역폭이고, F_r 은 초당 부호화된 프레임 율을 나타낸다. 식 (2)의 N_p 는 GOP 내에 있는 P-프레임의 개수를 의미하고, N_B 는 GOP 내의 B-프레임의 수를 의미하며 $\widetilde{W}_p(n_{i,j})$ 와 $\widetilde{W}_b(n_{i,j})$ 는 P-프레임과 B-프레임의 평균 복잡도 비중을 의미한다. 현재 부호화하고자 하는 프레임의 목표 비트량을 예측하기 위해 가상버퍼 점유량 $Tbl(n_{i,j})$ 및 목표 버퍼레벨 $B_c(n_{i,j})$ 을 이용하여 다음 식 (3)과 같이 버퍼량 $\widetilde{T}(n_{i,j})$ 을 예측한다.

$$\widetilde{T}(n_{i,j}) = \frac{u(n_{i,j})}{F_r} + \gamma \times (Tbl(n_{i,j}) - B_c(n_{i,j})) \quad (3)$$

이와 같이, 현재 프레임의 버퍼량이 예측되면 다음 식 (4)에서 예측한 부호화 비트량 $\widehat{T}(n_{i,j})$ 과 함께 고려하여 식 (5)과 같이 최종적으로 목표 비트량 $T(n_{i,j})$ 을 결정한다.

$$\widehat{T}(n_{i,j}) = \frac{W_p(n_{i,j-1}) \times T_r(n_{i,j})}{W_p(n_{i,j-1}) \times N_{p,r}(j-1) + W_b(n_{i,j-1}) \times N_{b,r}(j-1)} \quad (4)$$

$$T(n_{i,j}) = \beta \times \widehat{T}(n_{i,j}) + (1 - \beta) \times \widetilde{T}(n_{i,j}) \quad (5)$$

식 (4)는 GOP 내에 존재하는 프레임들의 부호화에 할당할 수 있는 비트량 $T_r(n_{i,j})$ 과 현재 GOP 내의 부호화하지 않은 P-프레임과 B-프레임의 수 $N_{p,r}(j-1)$, $N_{b,r}(j-1)$ 와 복잡도 $W_p(n_{i,j-1})$, $W_b(n_{i,j-1})$ 를 고려한 전체 복잡도 중에 현재 P-프레임의 복잡도의 비중을 이용하여 계산한다.

2. MAD 예측 및 QP 결정

현재 프레임에 대한 복잡도를 예측하는데 사용되는 MAD(Mean Absolute Difference)는 R-D 최적화(RDO: Rate Distortion Optimization) 기반의 부호화 구조와 QP를 결정하는 비트율 제어 모듈 간의 chicken-and-egg 딜레마^[7]로 인해 식 (6)과 같이 선형 예측값을 계산하여 비트율 제어에 사용한다.

$$MAD_c = a_1 \times MAD_p + a_2 \quad (6)$$

이 때의 MAD_c 는 예측한 현재 프레임의 왜곡값, MAD_p 는 이전 프레임의 실제 왜곡값을 나타낸다. 현재 프레임의 복잡도를 반영하는 예측 왜곡값 MAD_c 와 목표 비트량 $T(n_{i,j})$ 을 이용하여 식 (7)와 같이 2차 R-D 모델^[7]에 적용하여 양자화 단계 Q_{step} 을 계산한다.

$$\frac{T(n_{i,j})}{MAD_c} = \frac{x_1}{Q_{step}} + \frac{x_2}{Q_{step}^2} \quad (7)$$

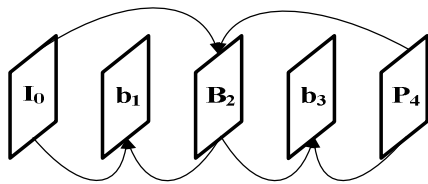
이 때의 x_1, x_2 는 이미 부호화된 이전 프레임으로부터 얻는 2차 R-D 모델의 파라미터 값이다.

3. 계층적-B 부호화 구조를 위한 B-프레임의 QP 결정

JM의 초기 버전은 비참조 B-프레임을 포함하는 부호화 구조를 고려한 것으로, 계층적-B 부호화 구조를 위한 비트율 제어 기법은 추가로 제안되었다. 초기 버전에서 고려하던 비참조 B-프레임을 포함하는 구조^[7]와 달리, [8]에서는 B-프레임이 다른 시간계층의 프레임을 부호화하는데 참조 영상으로 사용되는 계층적-B 부호화 구조를 고려하여 B-프레임의 QP를 결정한다. 식 (8)에 의해 이전 QP의 값 QP_{prev} 이 먼저 계산된다.

$$QP_{prev} = \max(QP_1, QP_2) \quad (8)$$

이 때의 QP_1, QP_2 는 참조 프레임 리스트인 List 0 와 List 1 예측 방향을 구성하는 I- 혹은 P-프레임이다. 다음 단계로, 계층적 부호화 구조일 때 계층적 레벨 수 H_{levels} 이 계산된다. 예를 들면, 그림 1과 같이 I- 또는 P-프레임 사이에 3개의 B-프레임이 있다면 $H_{levels} = 2$ 가 된다. 부호화 구조와 재생 순서의 쌍이 $I_0 b_1 B_2 b_3 P_4$ (B는 참조 프레임으로 사용되고, b는 비참조 프레임을 의미함)로 주어질 때, b_1 과 b_3 는 레벨 0에 속하는 반면, B_2 는 레벨 1에 속한다. 그러나, 부호화 순서는 $I_0 P_4 B_2 b_1 b_3$ 이므로 B-프레임 부호화 순



	I_0	b_1	B_2	b_3	P_4
B_{order}	x	1	0	2	x
L(B_{order})	x	0	1	0	x
QP	10	12	11	12	10

그림 1. JM에서의 계층적 부호화 구조를 위한 QP 결정
Fig. 1. QP decision for hierarchical coding structure in JM

서 B_{order} 는 0, 1, 그리고 2가 B_2, b_1, b_3 에 해당된다. $L(n)$ 을 B-프레임의 부호화 순서 n 일 때의 계층적 레벨이라고 할 때, 식 (9)과 같이 계층적 부호화 구조의 B-프레임에 대한 QP 값이 계산된다. 즉, 계층적 부호화 구조를 고려하여, 시간계층이 높은 레벨에 포함되는 프레임에 대해 높은 QP를 할당하여 부호화 효율을 최대화하였다.

$$\tilde{Q}B_{B_{order}} = QP_{prev} + H_{levels} - L(B_{order}) \quad (9)$$

4. 비트율 제어 결과

[8]에서 제안한 계층적 부호화 구조를 위한 비트율 제어 기법의 성능을 확인한다. JM에서는 비트율 제어를 위한 총 4개의 모드를 제공하고 있는데, 비참조 B-프레임 구조를 위해 제안된 초기 모델인 0번 모드, 프레임 타입이 모두 같은 경우인 I only, IPPPP, 혹은 IBBBB와 같은 구조를 위한 1번 모드, 계층적 부호화 구조를 위한 2, 3번 모드를 제공하고 있다. 앞서 설명한 비트율 제어 방식은 2번 모드에 해당된다. 2, 3번 모드를 통한 고정 비트율 제어를 통한 성능을 확인해본다. JM 16.2버전을 이용하였고, 비트율 제어를 위한 실험 조건은 표 1과 같다. 표 2의 결과와 같이 목표 비트율 대비 실제 발생 평균 비트율은 2번 모드의 경우, 5.52%

표 1. 비트율 제어 실험 조건
Table 1. Test conditions for rate control

부호화 파라미터	설정값
Resolution (framerate)	1920x1080(Kimono, ParkScene(24), BasketballDrive(50), BQTerrace(60)) 1920x544(KBS2, MBC2(60))
GOP size	8
Intra period	24 for 24 fps 48 for 50 fps 64 for 60 fps
Length	10 초
Reference frames	4
Search mode	EPZS
Search range	64
Symbol mode	CABAC
Bit depth	8

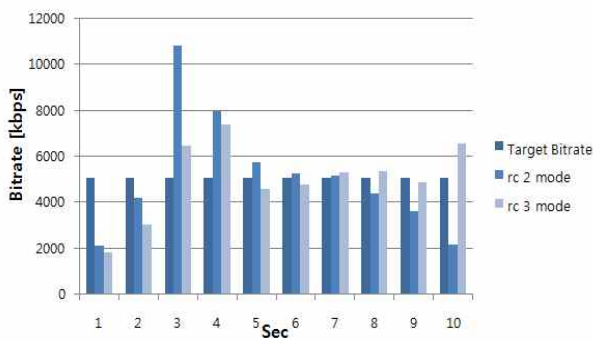
표 2. JM을 이용한 비트율 제어 결과

Table 2. Experimental results of rate control using JM

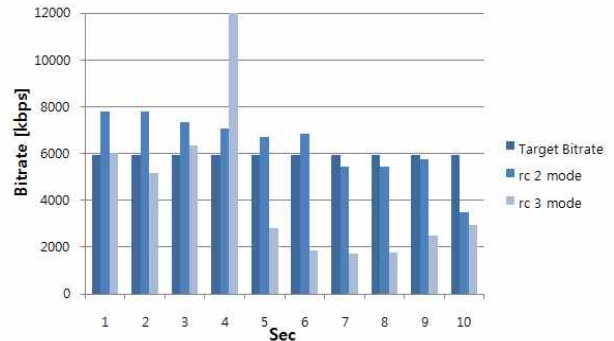
Sequence (Framerate)	Target Bitrate[kbps]	2 번 모드			3 번 모드		
		Actual Bitrate[kbps]	PSNR [Y-dB]	Error [%]	Actual Bitrate[kbps]	PSNR [Y-dB]	Error [%]
Kimono (24)	1800	1795.58	36.98	0.25	2225.63	34.65	23.65
ParkScene (24)	5050	5137.51	37.30	1.73	5003.45	36.18	0.92
BasketballDrive (50)	5940	6358.08	36.03	7.04	5909.31	33.80	0.52
BQTerrace (60)	6000	6724.62	34.25	12.08	5952.47	33.62	0.79
KBS2 (60)	11000	11827.10	36.90	7.52	11506.21	36.53	4.60
MBC2 (60)	6000	6269.12	41.33	4.49	6524.05	38.89	8.73
Average				5.52			6.54

의 오차율을 보이고, 3번 모드의 경우, 6.54%의 오차율이 발생함을 확인하였다. 추가적으로, 고정 비트율의 성능을 확인하기 위해, 그림 2와 같이, 초당 발생 비트량을 도식화

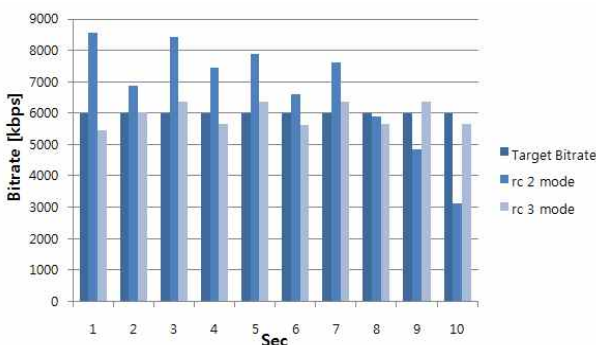
하였다. 설정한 초당 비트율 대비 실제 발생 비트량은 비트율 변동이 큰 것을 확인할 수 있다. 이는 JM에서 계층적 부호화 구조에 따른 GOP 단위의 효율적인 비트율 제어가



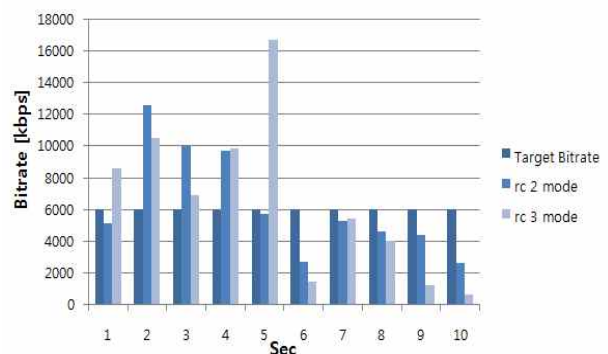
(a) ParkScene, 24 fps



(b) BasketballDrive, 50 fps



(c) BQTerrace, 60 fps



(d) MBC2, 60 fps

그림 2. JM을 이용한 초당 발생 비트량

Fig. 2. Bits per sec using JM

이루어지지 않기 때문이다.

III. 제안하는 비트율 제어 기법

H.264/AVC에서 사용되고 있는 비트율 제어 알고리즘은 앞서 기술한 바와 같이, 2차 R-D 모델을 이용하여 P-프레임의 QP 값을 계산하고, I와 B-프레임의 QP 값들은 P-프레임들의 부호화 결과와 현재 프레임의 계층을 기준으로 결정함에 따라 I와 B-프레임의 비트율 제어를 정확하게 수행하지 못한다. 현재 HEVC의 공통 실험조건과 같이 GOP의 크기를 8로 설정한 경우에는 B-프레임의 수가 증가함에 따라 정확한 비트율 제어를 어렵게 한다. 계층적-B 부호화 구조에서는 각 프레임 타입, 시간계층에 따라 다른 R-D 특성을 가지기 때문에 각각의 특성을 반영하는 알고리즘의 적용이 요구된다.

따라서, 본 논문에서는 계층적-B 부호화 구조에서의 모든 시간계층과 프레임 타입을 고려한 프레임에 대한 목표 비트량을 할당하고, 기존의 2차 R-D 모델을 시간계층과 프

레이프 타입에 따라 달리 적용하여 각 프레임의 QP를 결정하는 기법에 대해 기술한다.

제안하는 방식은 크게 1) 목표 비트량 계산, 2) 목표 비트량에 대한 QP 값 결정의 2 가지 주요 처리 단계로 구성된다. 그림 3과 같이 허용되는 채널 대역폭의 비트율을 입력하면 GOP, 프레임, 그리고 LCU 단위 순으로 목표 비트량 할당을 한다. 본 논문의 프레임 목표 비트 할당은 계층적-B 부호화 구조를 고려한 SVC(Scalable Video Coding) 비트율 제어방식^[9]을 이용한다.

1. 프레임 단위 비트율 제어

계층적-B 부호화 구조의 프레임 목표 비트할당은 다음 식 (10)에 의해 계산된다.

$$R_{\Sigma} = (1 - \alpha) \times R_1 + \alpha \times R_2 \quad (10)$$

이 때, α 는 [9]에서 고정 상수인 0.75으로 설정하고, R_1 은 식 (11)과 같다.

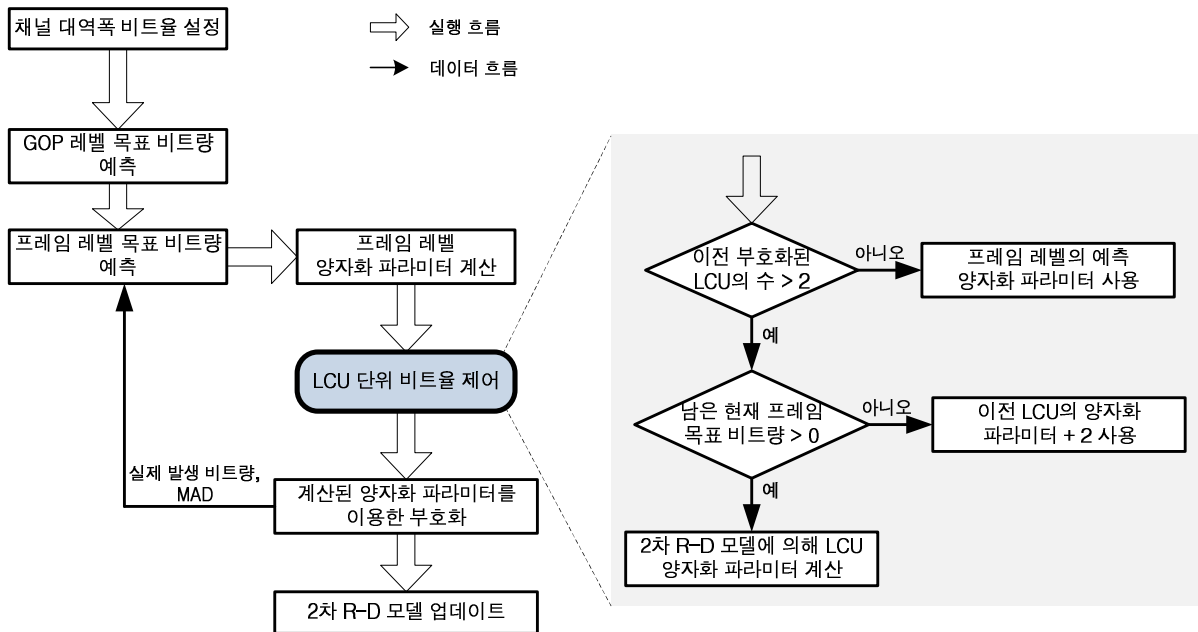


그림 3. 제안하는 비트율 제어의 전체적인 흐름도
Fig. 3. Flow chart of the proposed rate control algorithm

$$R_1 = \frac{W}{f} + \beta \times (B_t - B_c) \quad (11)$$

β 도 [9]의 상수 0.5로 설정하고, 이용 가능한 채널 대역폭 W , 프레임율 f , 목표 버퍼 점유도 B_t , 그리고 이전의 실제 버퍼 점유도 B_c 를 이용하여 R_1 을 구한다. 그리고, R_2 는 식 (12)와 같다.

$$R_2 = R_r \times \frac{w_k}{w_{sum}} \quad (12)$$

현재 프레임이 부호화되기 전의 GOP 내의 잔여 비트량 R_r , 현재 프레임의 시간계층 k 에서의 가중치 w_k , 그리고 GOP 내에 부호화되지 않은 남은 프레임들의 가중치 합 w_{sum} 을 이용하여 식 (12)를 통해 R_2 를 구한다. 가중치 합 w_{sum} 은 식 (13)과 같이 GOP 내에 부호화되지 않은 남은 프레임들의 합으로 계산된다.

$$w_{sum} = \sum_{k=0}^{L_{max}} N_{r,k} \times w_k \quad (13)$$

L_{max} 는 계층적 레벨의 최대값을 의미하고 $N_{r,k}$ 는 k 번째 계층적 레벨에 부호화되지 않고 남아있는 프레임의 수를 말한다. 이 때, 가중치 인자 w_k 는 다음 (14)와 같이 실험적으로 도출할 수 있다.

$$w_k = \begin{cases} 6.0 & k=0, I\text{frame} \\ 4.0 & k=0 \\ 3.5 & k=1 \\ 2.0 & k=2 \\ 1.0 & k=3 \end{cases} \quad (14)$$

이와 같이, 현재 프레임의 목표 비트량을 할당하면 계층적 부호화 구조에 따른 각 시간계층과 프레임 타입의 특성을 반영한 비트량을 할당 가능하다.

현재 프레임의 QP를 결정하기 위해서, H.264/AVC의 2차 R-D 모델을 확장하여 시간계층 l , 프레임 타입 t 를 반영한 식 (15)의 R-D 모델을 제시한다.

$$\frac{R_{sum}^{l,t}}{MAD_c^{l,t}} = \frac{X_1^{l,t}}{Q_{l,t}} + \frac{X_2^{l,t}}{Q_{l,t}^2} \quad (15)$$

식 (15)으로부터 식 (10)에서 도출한 프레임 목표할당 비트량 $R_{sum}^{l,t}$ 을 이용하여 해당 프레임의 QP를 결정한다. 이때의 $MAD_c^{l,t}$ 는 현재 프레임과 같은 시간계층, 프레임 타입을 가지는 이전의 실제 발생 MAD값들에 의해 선형 예측된 값이며, 2차 R-D 모델의 파라미터인 $X_1^{l,t}$, $X_2^{l,t}$ 는 같은 시간계층, 프레임 타입에서 발생한 이전의 실제 발생 비트량, 왜곡값, 그리고, 양자화 파라미터에 의해 갱신된 값이다. 이를 통해, 현재 프레임의 양자화 파라미터 $Q_{l,t}$ 값을 결정하여 현재 프레임에 적용한다.

2. LCU 단위 비트율 제어

HEVC에서는 부호화 단위로 MB를 확장하여 다양한 블록 크기를 갖는 CU를 정의하고 있으며, 현재 LCU를 64x64 크기로 설정하여 표준화가 진행 중이다. LCU는 기존 MB의 16 배의 크기에 해당하는 것으로 프레임과 MB 사이의 크기이며, 본 논문에서는 부호화기에서의 처리가 용이하도록 LCU를 BU로 고려한 비트율 제어 알고리즘을 제시한다.

프레임 내의 번째 LCU의 목표 비트량 b_l 은 프레임 레벨에서 할당된 남은 목표 비트량 T_r 을 기반으로 프레임 내의 LCU 간의 예측된 MAD 가중치를 이용하여 다음 식 (16)과 같이 할당한다.

$$b_l = T_r \times \frac{MAD_{pred,l}}{MAD_{pred,l} + N_{remainLCU} \times MAD_{mean}} \quad (16)$$

$$MAD_{mean} = \frac{\sum_{k=1}^{l-1} MAD_k}{l-1} \quad (17)$$

현재 부호화하고자 하는 LCU의 예측된 왜곡값 $MAD_{pred,l}$ 과 식 (17)과 같이 이전 $l-1$ 번째까지의 평균 왜곡값 MAD_{mean} , 그리고 프레임 내에 부호화되지 않고 남은

LCU의 수 $N_{remainLCU}$ 를 고려하여 전체 복잡도를 도출하고, 그 중에 현재 LCU의 복잡도의 가중치를 고려하여 비트량을 할당한다.

현재 부호화하는 LCU를 위한 QP 결정 방법은 앞서 기술한 프레임 단위의 기법과 동일하게 2차 R-D 모델을 이용하고 위의 식 (16)에서 도출한 LCU의 목표 비트량 b_l 을 적용하여 현재 LCU의 QP를 결정한다.

이와 같이 현재 LCU에 대한 예측된 QP는 위의 그림 3과 같이 다음 3 가지의 경우에 따라 조정된다. 첫 번째, 현재 프레임 내에서 두 번째 이하의 LCU의 경우에는 축적된 데이터가 없기 때문에 식 (18)과 같이 프레임 레벨에서 예측한 QP를 사용한다.

$$QP = \hat{QP}_{c,frm} \tag{18}$$

두 번째, 현재 프레임 내에 부호화되지 않고 남은 LCU에 대해 남은 비트 할당량이 $T_r \leq 0$ 일 때, 식 (19)와 같이 이전 LCU의 QP 값 $QP_{p,LCU}$ 보다 높은 QP값을 할당해 준다.

$$\hat{QP}_c = QP_{p,LCU} + 2 \tag{19}$$

$$QP = \max\{0, QP_{apf} - 3, \min\{51, QP_{apf} + 3, \hat{QP}_c\}\} \tag{20}$$

화질의 급격한 변화를 막기 위해 식 (20)과 같이 이전 프레임 내 LCU의 QP 평균값 QP_{apf} 을 기준으로 한정한다. 세 번째 경우는 그 외의 일반적인 경우로 2차 R-D 모델에 의해 계산된 QP 값을 이용한다.

IV. 실험 및 결과

본 장에서는 앞에서 제안한 비트율 제어 기법을 HM에 구현하여 계층적-B 부호화 구조에 대한 비트율 제어 결과를 확인한다. 또한, 비트율 제어 알고리즘을 프레임 단위와 LCU 단위로 설계 및 구현하여 그 결과를 비교한다.

계층적-B 부호화 구조에서의 비트율 제어 실험을 위해 III 장에서 기술한 제안 비트율 제어 알고리즘을 HM 2.0에

구현하였고, HEVC 표준화의 공통 실험조건에 의한 계층적-B 부호화 구조인 RA 모드를 사용하였고, 실험조건은 앞의 표 1과 같다.

표 3은 비트율 제어의 결과로 프레임 단위의 평균 비트율과 비트율의 정확도를 나타낸 것이고, 표 4는 LCU 단위의 결과를 보여준다. 본 논문에서는 기존 JM과 달리, 모든 시간계층에 대해 비트율 제어 모델을 이용하여 QP를 결정하기 때문에 목표 비트율에 근접하는 정확한 평균 비트율을 얻을 수 있음을 확인하였다. 추가적으로, 프레임 단위의 정확도를 측정하기 위해, 식 (21)에 의해 프레임 단위의 평균 오차율 m_E 과 표준 편차 σ_E 를 사용하였다.

표 3. 프레임 단위 비트율 제어 실험결과
Table 3. Experimental results of frame-level rate control

시퀀스 (IntraPeriod)	Target Bitrate [kbps]	Actual Bitrate [kbps]	Bitrate Diff. (m_E/σ_E)	PSNR(Y) [dB]
Kimono(24)	920	918.65	0.32/0.29	35.19
ParkScene(24)	2525	2521.48	0.47/0.58	34.26
BasketballDrive(48)	2970	2972.78	0.19/0.19	35.00
BQTerrace(64)	3000	2990.86	0.89/2.42	32.44
KBS2(64)	5500	5501.09	0.41/1.76	35.46
MBC2(64)	3000	3002.06	0.52/2.14	40.66
평균			0.47/1.23	35.50

표 4. LCU 단위 비트율 제어 실험결과
Table 4. Experimental results of LCU-level rate control

시퀀스 (IntraPeriod)	Target Bitrate [kbps]	Actual Bitrate [kbps]	Bitrate Diff. (m_E/σ_E)	PSNR(Y) [dB]
Kimono(24)	920	921.33	0.41/0.52	35.17
ParkScene(24)	2525	2527.72	0.40/0.48	34.16
BasketballDrive(48)	2970	2975.11	0.21/0.29	34.76
BQTerrace(64)	3000	3046.61	0.81/2.21	32.58
KBS2(64)	5500	5500.65	0.23/0.31	35.43
MBC2(64)	3000	2999.41	0.40/0.71	40.56
평균			0.41/0.75	35.44

$$E = \sum_{i=1}^{N_{data}} \frac{|T_i - S_i|}{T_i} \quad (21)$$

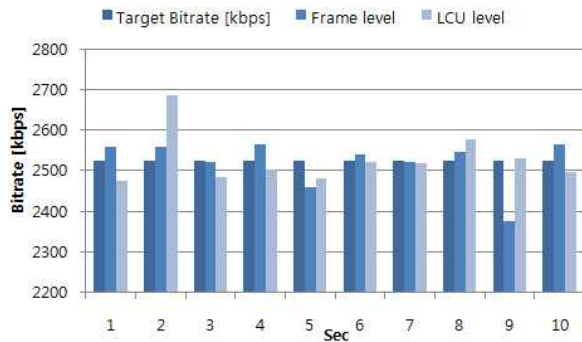
T_i 는 프레임 단위의 목표 비트 할당량, S_i 는 비트율 제어에 의해 발생하는 실제 발생 비트량을 나타낸다. 표 3과 표 4를 비교해보면 LCU 단위로 비트율 제어를 수행한 결과의 평균 오차율과 표준 편차가 프레임 단위로 수행한 결과보다 개선된 것을 확인할 수 있다.

그림 4는 LCU 단위로 비트율 제어를 수행했을 때, 초당 발생하는 비트량(kbits/sec)을 나타낸 것으로 발생하는 초당 부호화 비트량이 목표 비트율에 근접하게 발생함을 확인할 수 있다.

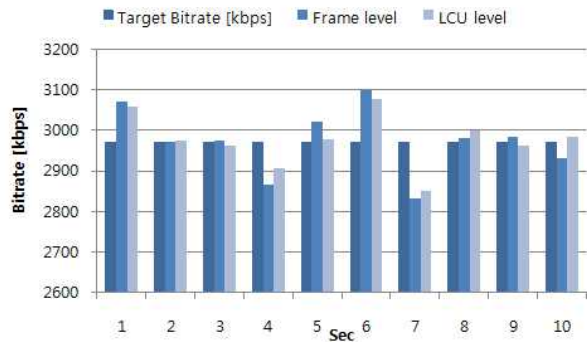
본 논문에서 제안한 방법을 이용한 비트율 제어 기법은 기존 JM의 방법과 다르게, P-프레임뿐만 아니라 I, B-프레

임에 대해서도 비트율 제어 모델을 적용하여 QP를 결정한다. 계층적-B 부호화 구조를 고려하여 각 시간계층과 프레임 타입에 따라 각기 다르게 발생하는 R-D 특성을 2차 R-D 모델에 반영하여 비트율 제어 정확도를 향상시켰다. 또한, 프레임 단위의 QP 할당방식에 추가적으로 LCU 단위까지 목표 비트량을 할당하고 QP를 결정함으로써 프레임 단위의 발생 비트량이 목표 비트량에 대해 평균적으로 5% 정도의 정확도가 향상되었지만 0.12dB의 PSNR 손실이 있음을 확인하였다. 이는 목표 비트량에 도달하기 위해 LCU단위로 QP를 조절함으로써 나타나는 상충관계를 확인할 수 있다.

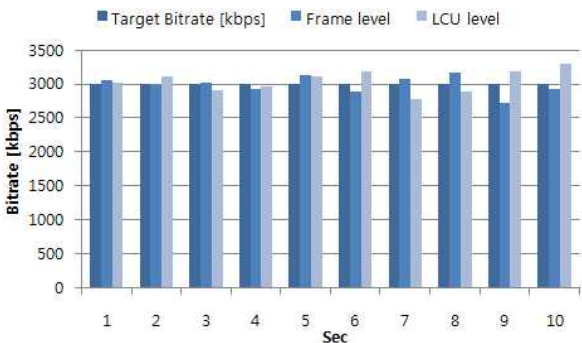
추가적으로, 프레임 단위의 비트율 제어 정확도가 향상되어 표 5의 결과와 같은 결과를 확인하였다. 표 5에 나타난 비율은 프레임 목표 비트량이 0으로 설정되는 경우에



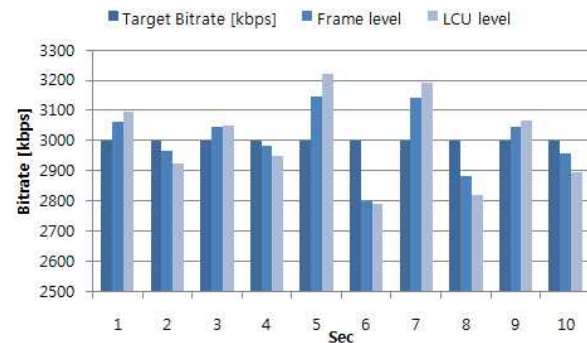
(a) ParkScene, 24 fps



(b) BasketballDrive, 50 fps



(c) BQTerrace, 60 fps



(d) MBC2, 60 fps

그림 4. HM을 이용한 초당 발생 비트량

Fig. 4. Bits per sec using HM

대한 것으로 이는 GOP에 할당된 비트량을 GOP 내의 프레임들의 잘못된 QP 예측으로 모두 소진한 경우를 나타낸다. 시퀀스에 따라 차이가 있지만 평균적으로 LCU 단위의 비트율 제어의 경우, 약 2% 정도 향상이 있음을 확인할 수 있다. 즉, 프레임 단위의 비트율 제어의 경우, 한 프레임에 대해 하나의 QP 값을 할당하므로, 목표로 한 비트량에 근접하기 어려운 반면, LCU 단위의 비트율 제어의 경우, 현재 프레임에 할당된 비트량을 LCU 단위로 조절함으로써 향상된 정확도를 얻을 수 있음을 알 수 있다.

또한, 프레임과 LCU의 2 가지 기본단위에 대한 비트율 정확도와 화질간의 상충관계를 분석하였다. LCU 단위의 비트율 제어는 프레임 단위의 비트율 제어 대비 평균적으로 5% 정도의 정확도 향상을 보였고 0.12dB의 PSNR 손실을 보임을 확인하였다.

이와 같은 비트율 제어의 정확도와 화질열화 간의 상충관계가 최소가 되는 기본단위의 설정에 대한 추가적인 연구가 필요하다. 본 논문의 비트율 제어 기법은 추후 표준화 단계에서 계층적-B 부호화 구조를 위한 비트율 제어를 위한 시험 및 검증에 활용될 수 있을 것으로 기대한다.

표 5. 프레임 목표 비트량이 0인 비율
Table 5. Percentage of the exhausted frame target bits

시퀀스	Frame level[%]	LCU level[%]
Kimono	2.08	4.58
ParkScene	4.58	1.67
basketballDrive	0	0
BQTerrace	18.50	11.83
KBS2	4.33	0
MBC2	4.17	3.83
평균	5.61	3.65

V. 결론

본 논문에서는 계층적-B 부호화 구조를 갖는 HEVC의 RA 모드에서의 비트율 제어 알고리즘을 구현하고 그 성능을 분석하였다. 계층적-B 부호화 구조의 시간계층 및 프레임 타입을 반영하여 기존의 H.264/AVC에 제안된 2차 R-D 모델을 확장한 양자화 파라미터 결정 기법을 제시하였다. 제시된 2차 R-D 모델을 바탕으로 프레임 단위와 LCU 단위의 비트율 제어 기법을 HM에 구현하여 모의실험을 통해 HEVC RA 모드에서 비교적 정확한 비트율 제어가 가능함을 확인하였다.

참고 문헌

- [1] Joint Call for Proposals on Video Compression Technology, ITU-T Q6/16 VCEG 39th meeting, VCEG-AM91, Kyoto, Japan, 17-22 Jan., 2010.
- [2] T. K. Tan, Gary, J. Sullivan and Jens-Rainer Ohm, The summary of HEVC working draft 1 and HEVC test model (HM), JCTVC-C405, JCT-VC 3rd meeting, Guangzhou, 7-15 Oct., 2010.
- [3] HEVC reference software (svn): https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags
- [4] JCT-VC, Common test conditions and software reference configurations, JCTVC-E700, JCT-VC 5th meeting, Geneva, Mar., 2011.
- [5] ISO/IEC 14496-10: Information technology - Coding of audio-visual objects - Part 10: Advanced video coding, 2008.
- [6] H.264/AVC reference software: <http://iphome.hhi.de/suehring/tml/>
- [7] Z. Li, W. Gao, F. Pan and K. Pang, Adaptive Basic Unit Layer Rate Control for JVT, Joint Video Team of ISO/IEC MPEG and ITU-T VCEG, JVT-G012, 7th meeting, Pattaya II, Thailand, 7-14 Mar., 2003.
- [8] A. Leontaris and A. M. Tourapis, Rate Control reorganization in the Joint Model (JM) reference software, Joint Video Team of ISO/IEC MPEG and ITU-T VCEG, JVT-W042, 23rd meeting, San Jose, California, USA, 21-27 Apr., 2007.
- [9] Y. Liu, Z. G. Li and Y. C. Soh, Rate Control of H.264/AVC Scalable Extension, IEEE Trans. Circuits Syst.Video Technol., vol. 17, no. 5, pp. 517-529, May 2007.
- [10] 서찬원, 한종기, 계층적 부호화 구조의 H.264/AVC에서 효율적인 비트율 제어를 위한 향상된 비트율-양자화 모델링, 한국방송공학회 추계학술대회, 2010.
- [11] JCT-VC, WD3: Working Draft 3 of High-efficiency video coding, JCTVC-E603, JCT-VC 5th meeting, Geneva, Mar., 2011.

— 저 자 소 개 —



박 동 일

- 2010년 2월 : 한국항공대학교 항공전자 및 정보통신공학부 학사
- 2010년 3월 ~ 현재 : 한국항공대학교 항공전자 및 정보통신공학부 석사과정
- 주관심분야 : 영상부호화, UHDTV, High Efficiency Video Coding



김 재 곤

- 1990년 2월 : 경북대학교 전자공학과 학사
- 1992년 2월 : KAIST전기 및 전자공학과 석사
- 2005년 2월 : KAIST전기 및 전자공학과 박사
- 1992년 3월 ~ 2007년 2월 : 한국전자통신연구원(ETRI) 선임연구원/팀장
- 2001년 9월 ~ 2002년 11월 : 뉴욕 콜롬비아대학교 연구원
- 2007년 9월 ~ 현재 : 한국항공대학교 항공전자 및 정보통신공학부 조교수
- 주관심분야 : 비디오 신호처리/부호화, 디지털방송 미디어, 미디어 컨버전스, 멀티미디어 통신



임 성 창

- 2006년 2월 : 세종대학교 컴퓨터공학과 학사
- 2008년 2월 : 세종대학교 컴퓨터공학과 석사
- 2008년 3월 ~ : 한국전자통신연구원 실감미디어연구팀 연구원
- 주관심분야 : 영상 압축, 영상 처리



김 종 호

- 2005년 2월 : 한국해양대학교 제어컴퓨터공학과 학사
- 2007년 8월 : 과학기술연합대학원대학교 (UST) 컴퓨터소프트웨어 및 공학 전공 석사
- 2008년 9월 ~ : 한국전자통신연구원 실감미디어연구팀 연구원
- 주관심분야 : 영상 부호화, 영상 처리



김 휘 용

- 1994년 8월 : 한국과학기술원 전기및전자공학과 학사
- 1998년 2월 : 한국과학기술원 전기및전자공학과 석사
- 2004년 2월 : 한국과학기술원 전기및전자공학과 박사
- 2003년 8월 ~ 2005년 10월 : ㈜에드팩테크놀러지 멀티미디어 팀장
- 2005년 11월 ~ : 한국전자통신연구원 실감미디어연구팀 선임연구원
- 2006년 9월 ~ 2010년 8월 : 과학기술연합대학원대학교(UST) 이동통신및디지털방송공학과 겸임교수
- 주관심분야 : 컴퓨터비전, 영상압축, 멀티미디어시스템