

## 새로운 율모델을 이용한 HEVC 율제어 기법

이범식, 김문철

한국과학기술원

lbs@kaist.ac.kr, mkim@ee.kaist.ac.kr

### A rate control scheme using a new rate model for the HEVC video codec

Bumshik Lee and Munchurl Kim

Korea Advanced Institute of Science and Technology (KAIST)

## 요 약

본 논문에서는 새로운 율모델을 기반으로한 프레임 단위 HEVC 율제어 기법을 제안한다. 기존의 비디오 압축 표준과는 달리 HEVC 는 계층 구조를 지닌 쿼드트리 기반 움직임 예측 및 변환 부호화를 수행한다. 본 논문에서는 쿼드트리 계층의 깊이에 따라 신호의 통계적 특성이 매우 달라지는 것은 이용하여 라플라시안 확률 모델을 각 쿼드트리 계층에 독립적으로 이용한 새로운 율모델을 이용한 율제어 기법을 제안한다. 제안 방법에서는 계층적 부호화 단위인 CU 를 계층 깊이에 따라 세 가지 카테고리로 분류하고 각 카테고리에 따라 변환 계수에 대한 라플라시안 확률 분포 함수를 율-양자화 모델을 만든다. 제안된 율모델은 특성이 매우 다른 각 CU 깊이에 따라 독립적인 라플라시안 확률 분포 함수를 사용하기 때문에 매우 정확하고 적응적인 비트율 예측이 가능하므로 보다 안정적이고 정확한 율제어가 가능하다. 실험결과는 제안된 율제어 기법이 단일 확률 분포 함수를 사용했을 경우보다 평균 0.16dB 의 PSNR 향상이 있었음을 보여주었으며 제안된 방법은 각 프레임에 대한 목표 비트에 보다 안정적으로 부호화하는 것을 보여주었다.

## 1. 서론

현재 HEVC[1] 는 위원회 초안(Committee Draft, CD)이 승인되어 2013 년 초에 최종 표준화 완료로 목표로 표준화 작업이 진행 중이다. 현재 HEVC 는 보다 유연한 계층 구조를 채택하여 작은 크기의 영상부터 초고해상도 영상까지 모두 효율적으로 부호화 할 수 있도록 설계 되어있다. 계층구조는 부호화 단위(Coding Unit, CU), 예측 단위 (Prediction Unit, PU) 그리고 변환 단위 (Transform Unit, TU)로 구성된다. CU 는 64×64 블록 크기부터 8×8 블록 크기까지 쿼드트리 구조로 부호화되며 H.264/AVC 의 매크로블록과 같은 역할을 수행한다. PU 는 CU 블록 내에서 예측을 수행하며 다양한 예측 모드로 수행된다. CU 블록 내에서 PU 를 이용해 잔차 신호가 얻어지면 TU 에서 변환 부호화 과정을 수행한다. TU 블록의 크기는 32×32 부터 4×4 블록 크기까지 계층적 쿼드 트리 구조를 지닌다. 기존의 비디오 표준에 비해 확장된 블록 크기를 갖는 예측 및 변환 구조, 그리고 계층 구조 때문에 HEVC 와 같은 구조에서 율제어 등에 적용 가능한 잔차 신호에 대한 정확한 율모델을 만드는 것이 어려워 졌다.

비디오 표준 코덱이 등장한 이후 율제어에 관한 많은 연구들이 수행되어 왔다. 기존 연구들은 MPEG-1, MPEG-2, H.263, H.264/AVC 등 특정 표준 코덱을 위한 기법들이 제안되어 왔으며, 잔차 신호의 주로 통계나 복잡도 분석을 통한 율모델을 이용하여 율제어 기법이 개발되었다. Chiang [2]은 이차 율-양자화 모델을 제안하였다. 제안된 모델은 잔차 신호가 라플라시안 확률 분포 모델을 따른다고 가정하였으며, 프레임의 복잡도를 MAD(Mean Absolute Difference)를 이용하여 측정하였다. 이 모델은 연산이 간단하며 비교적 높은 정확도 때문에 H.264/AVC 의 참조 소프트웨어에 적용되었다. 이 모델은 발표된 이후 프레임의 복잡도의 정확도를 향상시키거나, MAD 대신 SATD(Sum of Absolute Transformed Difference) [3] 등 다른

복잡도 예측 방법을 사용하는 방법이 제안되기도 하였다. 확률 분포 기반 방법으로는 Cauchy 확률 분포를 이용한 방법이 [4]에서 제안하였다. 이 논문들에서는 Cauchy 확률 분포가 H.264/AVC 의 잔차 신호를 표현하기에 다른 확률 분포보다 더 정확하다는 것을 실험 결과를 통해 보여주었다.

이러한 기존의 율모델들을 이용한 율제어 기법들은 기존의 비교적 작은 크기의 예측 및 변환 구조를 갖거나 단순한 블록 구조를 갖는 비디오 표준 코덱에 유용하며, 비교적 정확한 율제어 성능을 보여 주었다. 그러나 HEVC 와 같이 작은 크기부터 확장된 블록 크기를 갖는 예측 및 변환 구조에서는 정확하지 않은 비트율 예측으로 인해 율제어의 적용에 어려운 단점이 있다. 본 논문에서는 단일 확률로 인한 비트율의 부정확한 예측을 극복하기 위한 프레임레벨의 율제어 기법이 제안된다. 제안된 방법에서는 CU 의 깊이와 부호화 타입에 따라 CU 를 3 개의 카테고리로 구분하여, 독립적으로 라플라시안 확률 분포 모델을 만든 후, 엔트로피를 이용한 율-양자화 모델을 이용하여 율제어를 수행한다. 제안하는 방법은 다양하게 변화하는 잔차 신호를 보다 적응적으로 예측하여 표현하는 장점이 있다. 실험 결과는 제안하는 방법이 단일 확률 분포 모델을 이용하는 방법보다 월등히 우수한 율제어 성능을 보여 주었다.

본 논문은 다음과 같이 구성된다. 2 장에서는 새로운 율모델을 이용한 율제어 기법을 설명하며, 3 장에서 실험결과가 제시된다. 4 장에서 결론을 맺는다.

## 2. 새로운 율모델을 이용한 율제어 기법

### 제안하는 율모델:

HEVC 의 가장 큰 크기를 갖는 CU (LCU, the Largest CU)가 64×64 블록 크기를 갖고 가장 작은 크기를 갖는 CU 가 8×8

블록 크기를 가질 때 CU의 총 깊이는 4가 된다. 본 제안 모델에서는 CU 깊이가 0 또는 1을 갖는 CU를  $CU_l$ , CU 깊이가 2 또는 3을 갖는 CU를  $CU_m$ , 마지막으로 화면내 부호화되는 CU를  $CU_h$ 로 한다. 여기서  $l, m, h$ 는 각각 low-texture, medium-texture, high-texture를 나타낸다. 각 CU 카테고리에서 라플라시안 확률분포  $f(l) = (\lambda_c/2)\exp(-\lambda_c|l|)$ 를 이용하면 엔트로피를 이용한 율모델은 (1)과 같이 표현된다.

$$R(q) = \alpha_l N_l H_l(q, \lambda_l) + \alpha_m N_m H_m(q, \lambda_m) + \alpha_h N_h H_h(q, \lambda_h) \quad (1)$$

여기서  $H_l(q, \lambda_l)$ ,  $H_m(q, \lambda_m)$ ,  $H_h(q, \lambda_h)$ 는 각 CU 카테고리에서의 엔트로피를 나타낸다.  $N_l, N_m, N_h$ 는 각 CU 카테고리의 총 블록크기의 합,  $\alpha_l, \alpha_m, \alpha_h$ 는 모델 파라미터를 나타낸다. HEVC의 양자화 구조에서 양자화되는 변환 계수의 엔트로피는 (2)와 같다.

$$H_c(q, \lambda_c) = -P_{0,c} \cdot \log_2 P_{0,c} - 2 \sum_{i=1}^{\infty} P_{i,c} \cdot \log_2 P_{i,c} \quad (2)$$

$H_c(q, \lambda_c)$ 는 CU 카테고리  $c$ 에 대한 엔트로피,  $P_{0,c}$  and  $P_{i,c}$ 는 0으로 양자화되는 변환 계수의 확률과  $i$ 번째 양자화 간격으로 양자화되는 확률을 각각 나타낸다. (2)는 (3)과 같은 식으로 간략화 할 수 있다. [5]

$$R(q) = \alpha_l \hat{N}_l \left( e^{-(1-f)q\hat{\lambda}_l} \right) + \alpha_m \hat{N}_m \left( e^{-(1-f)q\hat{\lambda}_m} \right) + \alpha_h \hat{N}_h \left( e^{-(1-f)q\hat{\lambda}_h} \right) \quad (3)$$

여기서  $(\hat{N}_l, \hat{N}_m, \hat{N}_h)$ 와  $(\hat{\lambda}_l, \hat{\lambda}_m, \hat{\lambda}_h)$ 는 블록 크기와 라플라시안 모델파라미터의 예측치이고, 본 제안 방법에서는 이전 프레임의 가중 평균으로 구한다.

그림 2는 제안하는 율모델의 비트 예측 성능을 보여준다. *BlowingBubbles* 영상을 QP 33으로 부호화했을 때 실제 비트와 예측 비트, 그리고 단일 확률 분포를 이용하였을 때 예측 비트를 나타낸다.

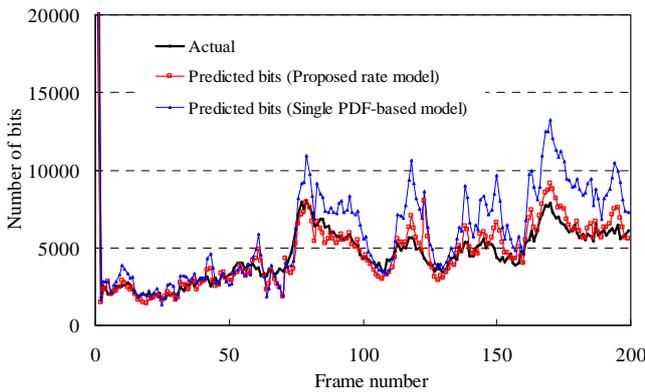


그림 1. 실제 비트와 제안하는 율모델에 의한 예측 비트 (*BlowingBubbles*\_416x240, QP=33)

그림 2에서 보는 것처럼 제안하는 율모델이 실제 비트와 매우 가까운 그래프를 보여준다. 이것은 비트량이 매우 크게 변화하는 구간에서 단일 모델에 의한 비트율 예측에 비해 월등히 높은 예측 성능을 보여주는 것을 알 수 있다.

### 율모델을 이용한 프레임 레벨 율제어 기법:

그림 1에서는 보는 것 처럼 제안하는 율모델은 비트 예측 성능이 매우 우수한 것을 알 수 있다. 먼저 GOP 레벨에서 비트할당은 (4)와 같다.

$$T_{GOP,i} = \begin{cases} \frac{R_T}{f} \times N_{GOP}, & \text{for } i=1 \\ \frac{R_T}{f} \times N_{GOP} + R_{r,i-1}, & \text{otherwise} \end{cases} \quad (4)$$

여기서  $T_{GOP,i}$ 는  $i$ 번째 GOP에 할당된 비트량을 나타낸다.  $f$ 는 프레임율(frame per second, fps),  $R_T$ 는 목표비트율,  $N_{GOP}$ 는 한 GOP 내에서 프레임 개수,  $R_{r,i-1}$ 는  $(i-1)$ 번째 GOP에서 GOP 목표 비트와 실제 비트의 차이를 각각 나타낸다. 다음으로 한 GOP 내의  $j$ 번째 프레임에서 비트할당은 다음과 같이 계산된다.

$$T_{f,j} = w\hat{T}_j + (1-w)\tilde{T}_j \quad (5)$$

여기서  $\hat{T}_j$ 와  $\tilde{T}_j$ 는 (6)으로부터 구할 수 있다.

$$\hat{T}_j = \frac{R_T}{f} + \gamma(T_{bl,j} - B_{C,j}), \quad \tilde{T}_j = \frac{R_{j,GOP}}{n_{rem}} \quad (6)$$

$T_{bl,j}$ 과  $B_{C,j}$ 는  $j$ 번째 프레임의 목표 버퍼레벨과 현재 버퍼레벨을 나타낸다.  $R_{j,GOP}$ ,  $n_{rem}$ 는 GOP 내에서 잔류 비트량 및 프레임수를 나타낸다. 또한  $\gamma$ 와  $w$ 는 모델파라미터로 상수이며 본 제안 방법에서는 0.25와 0.5가 각각 사용된다.

### 3. 실험 결과

제안하는 율제어 방법을 검증하기 위해 HEVC 참조 소프트웨어 6.0에서 실험하였다. GOP 구조는 IPPP, 화면내 프레임의 주기는 64가 사용되었다. 본 제안 방법은 단일 확률 모델을 이용한 방법과 비교 실험을 수행하였다. 표 1은 제안하는 새로운 율 모델을 이용한 HEVC 율제어 기법에 관한 실험 결과를 보여준다. 표에서 보는 것처럼 목표 비트율을 충족하면서 단일 확률 모델 기반 율제어 방법에 비해 평균 0.16dB의 성능향상을 보여준다. 이것은 그림 1에서 보는 것과 같이 제안 모델의 비트 예측 성능이 단일 확률 분포기반 방법에 비해 월등히 우수하기 때문이다. 또한 그림 2는 각 프레임에서 프레임 목표 비트 및 실제 부호화 비트를 보여준다.

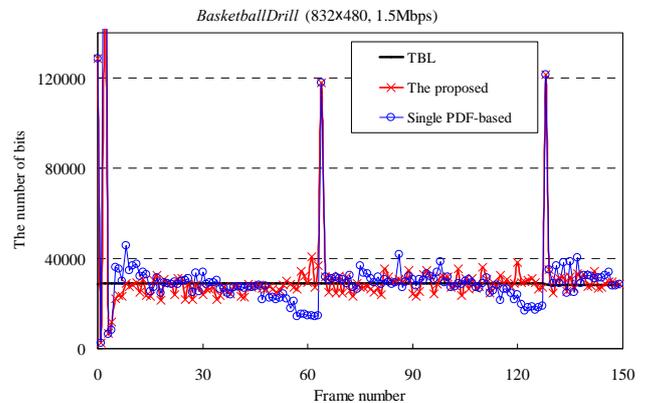


그림 2. 각 프레임에서 부호화 비트

표 1. 제안하는 율제어 기법의 실험 결과

Test sequences		fps	Target bitrate (kbps)	PSNR (dB)		Bit rate (kbps)	
				Proposed	Single-PDF	Proposed	Single-PDF
416×240	<i>BasketballPass</i>	50	750	37.47	37.40	748.2	738.8
			1500	41.29	41.19	1498.0	1488.1
	<i>BlowingBubbles</i>	50	750	33.65	33.50	760.2	749.2
			1500	36.33	36.43	1511.2	1499.2
832×480	<i>BQMall</i>	60	1500	32.81	31.79	1489.4	1490.2
			3000	35.96	35.81	2998.1	3010.1
	<i>BasketballDrill</i>	50	1500	35.28	35.11	1502.2	1499.2
			3000	37.87	37.91	3011.7	3100.2
1920×1080	<i>BQTerrace</i>	60	8000	38.33	38.21	7990.5	8010.2
			10000	39.21	39.22	9998.6	9988.6

그림 2 에서 보는 것처럼 제안하는 율제어 기법은 보다 프레임 목표 비트에 정확하고 안정적으로 비트를 할당하는 것을 알 수 있다. 반면 단일 확률분포 기반 방법은 비트할당이 제안방법보다 부정확하고 안정적이지 못할 것을 알 수 있다.

#### 4. 결론

본 논문에서는 새로운 율모델을 기반으로 한 율제어 기법을 제안하였다. 기존의 부호화 구조와 매우 상이한 HEVC 부호화 구조의 특성을 반영하도록 CU 의 깊이와 부호화 타입에 따라 CU 를 3 개의 카테고리로 구분하고 구분된 CU 에 따라 독립적인 라플라시안 확률 분포를 만들어 새로운 율-모델을 제안하였다. 본 제안 모델을 기반으로 율제어를 수행하면 단일 확률 모델 기반 율제어에 비해 목표 비트율을 충족하며 평균 0.16dB 의 PSNR 상승을 얻을 수 있었고 프레임당 목표 비트할당 역시 보다 안정적이고 정확하게 충족하는 것을 실험을 통해 확인하였다.

#### 참고 문헌

- [1] Bross, W.-J. Han, G. J. Sullivan, J.O.R. Ohm and T. Wiegand "High Efficiency Video Coding (HEVC) test specification draft 6", JCTVC-H1004, 8-th JCTVC meeting, San Jose, Feb. 2012.
- [2] T. Chiang and Y.-Q. Zhang, "A new rate control scheme using quadratic rate distortion model," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, no. 1, pp. 246-250, Feb. 1997.
- [3] D.-K. Kwon, M.-Y. Shen and C.-C. J. Kuo, "Rate control for H.264 video with enhanced rate and distortion models," *IEEE Trans. Circuits Syst. Video Technol.* vol. 17, no. 5, pp. 517- 529, May. 2007.
- [4] N. Kamaci, Y. Altunbasac and R. M. Mersereau, "Frame bit allocation for the H.264/AVC video coder via Cauchy-density-based rate and distortion models," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 8, pp. 994-1006, Aug. 2005.
- [5] B. Lee and M. Kim, "Modeling Rates and Distortions Based on a Mixture of Laplacian Distributions for Inter-Predicted Residues in Quadtree Coding of HEVC," *IEEE Signal Process. Letters*, vol. 18, no. 10, pp. 571 - 574, Oct. 2011.
- [6] HM reference software 6.0, [Online] Available: [https://hevc.hhi.fraunhofer.de/svn/svn\\_HEVCSoftware](https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware).