

## 학습된 선형 변환 기저를 이용한 VVC 잔차신호 변환

김남욱, \*이영렬

세종대학교, \*세종대학교

nukim@sju.ac.kr, \*yllee@sejong.ac.kr

### 요 약

본 논문에서는 비디오 코딩 잔차신호를 보다 효율적으로 변환하기 위하여 오프라인으로 잔차신호를 학습하여 RD(Rate Distortion) Cost 를 기반으로 분류된 몇 가지 변환 기저들을 생성하고, 비디오 복호화 과정 중 잔차신호를 역변환을 수행할 때 주변의 복호화가 완료된 신호들을 이용하여 최적의 변환 기저를 선택하여 해당 변환 기저로 역변환을 수행하여 효율적으로 잔차신호를 압축하는 방법에 대해 제안한다. 변환 기저 생성에는 분류된 잔차신호들에 대하여 2 차원 혹은 1 차원 KLT 를 계산함으로써 얻어내어진다. 제안하는 방법은 VTM(VVC Test Model) version 10 에서 실험하였으며 약 0.5% 정도의 성능향상을 보인다.

### 1. 서론

VVC(Versatile Video Coding)[1]는 가장 최근 표준화가 완료된 비디오 코딩 표준이다. VVC 에서는 하이브리드 블록 기반 코딩으로 영상을 압축하며, 다양한 직사각형 크기의 블록들로 영상을 분할하여 비트레이트(bit-rate)를 줄이면서 영상 화질은 올리는 성능을 보이고 있다. VVC 에서는 잔차신호(residual signal)를 효율적으로 압축하기 위하여 DCT-II(Discrete Cosine Transform type-II), DST-VII(Discrete Sine Transform type-VII), DCT-VIII 1 차원 선형 변환을 2 차원 잔차신호에 가로 방향으로 한번, 그리고 세로방향으로 한번 적용하여 에너지 압축(energy compaction)을 수행한다 [2]. 변환된 잔차신호는 변환계수(transformed coefficients) 라고하며 변환계수는 잔차신호보다 더 적은 차원에 더 많은 신호 에너지가 집중되기 때문에 엔트로피 코딩을 통한 압축 효율이 증가될 수 있다. VVC 의 예측성능의 증가로 잔차신호의 불규칙함이 조금 감소했지만 여전히 잔차신호의 패턴은 VVC 의 DCT/DST 변환으로 비상관화(decorrelation)하기에는 어려운 패턴들이 많이 존재한다. 본 논문에서는 보다 더 많은 잔차신호 패턴 분류를 통해 비상관화 성능을 높이기 위해 KLT(Karhunen Loeve Transform)을 활용하여 미리 실제 잔차신호들을 가지고 분류 및 학습을 하여 1 차원 선형의 KLT 기저들과 2 차원 비선형의 KLT 기저들을 생성하고 이러한 기저들을 가지고 VVC 에 적용하는 방법에 대하여 설명한다. 본 논문의 2 장에서는 제안하는 방법에 대하여 설명하고, 3 장에서는 실험결과 그리고 마지막 4 장에서는 결론으로 논문을 마친다.

### 2. 제안하는 방법

실제 영상들을 VVC 로 코딩하여 얻는 많은 수의 잔차신호 블록들을 이용하여 KLT 를 미리 학습할 수 있다. 본 논문에서는 학습을 보다 더 효율적으로 하기 위하여 화면 내 예측 방향과 잔차신호의 자기 상관관계(auto correlation)을 고려하여 학습할 잔차신호 블록들을 분류하였다.

---

#### Algorithm 1 KLT Training

---

**Input:** residual dataset  $R$

$QP = 22$

**for all**  $r_k \in R$  **do**

$orgcost = \min(\text{dctCost}(r_k, QP), \text{dstCost}(r_k, QP))$

$kltcost = \text{kltCost}(r_k, QP)$

**if**  $orgcost \leq kltcost$  **then**

remove  $r_k$  from  $R$

**end if**

**end for**

**for** ( $i = 0$  ;  $i < \text{numGroups}$ ;  $i++$ ) **do**

$R' = r$  in  $R$  with  $r$  in group  $i$

$K_i = \text{deriveKLT}(R')$

**end for**

**return**  $K$

---

세부적인 학습 진행과정은 위의 Algorithm 1 과 같다. 우선 잔차신호 집합을  $R$  이라고 하면  $R$  의 모든 잔차신호들에 대해서 DCT 로 변환 양자화 엔트로피 코딩 역양자화 역변환 과정을 통해 RD(Rate-Distortion) 비용을 계산한다. 그리고 DST 에 대해서도 동일한 과정을 수행하여  $orgcost$  를 얻는다. 그리고 잔차신호에 대하여 KLT 를 계산하여 마찬가지로 RD 비용을 계산하여  $kltcost$  를 얻는다. 그리고 만약  $orgcost$  보다

$kltcost$  가 크다면 KLT 의 효율이 떨어질 것으로 미리 예상하여 해당 잔차신호는 제거한다. 그리고 미리 정한 기준으로 분류한 잔차신호 그룹들 각각에 대하여 KLT 변환 기저를 계산하여 저장한다. 그룹에 대한 기준은 잔차신호의 화면 내 예측 모드와 잔차신호의 자기 상관 값으로 분류하였다.

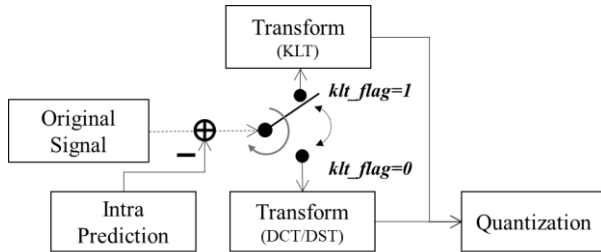


그림 1. 제안하는 방법의 부호화 다이어그램

위의 그림은 제안하는 방법을 VVC 에 적합한 인코딩 방법을 나타낸다. 우선 원본 신호에서 예측 신호를 뺀 잔차신호를  $klt\_flag$  를 통해 KLT 변환을 수행할지 아니면 기존 DCT-II 혹은 DST-VII 변환을 사용할지를 결정한다.

KLT 변환 기저 계산 방법에는 비선형 계산방식과 선형 계산방식을 사용할 수 있다. 우선 비선형 계산 방식은 2 차원 (N, M) 크기의 잔차신호를  $N*M$  크기의 1 차원으로 취급함으로써 계산할 수 있다. 하지만 비선형 기저는 학습이 잘 되지 않는 문제가 있고, 기저의 크기가  $N*M*N*M$  으로 학습하여 저장하기가 어려운 문제가 있어서 본 논문에서는 선형 기저만 사용하였다.

### 3. 실험 결과

제안하는 방법은 VTM10 에서 구현 및 테스트하였다. 테스트에 사용한 영상은 4K 영상으로는 Tango2, Campfire 이고 FHD 영상으로는 RitualDance, Cactus 를 사용하였다.

표 1. 실험 결과

영상	BD Rate	EΔT	DΔT
Tango2	-0.4%	128%	106%
Campfire	-0.3%	132%	99%
RitualDance	-0.8%	122%	101%
Cactus	-1.1%	137%	96%

BD-Rate 는 두 개의 코덱간의 압축성능 차이를 비교하는 metric 으로 [3]를 참조하여 계산하였고, EΔT 및 DΔT 는 다음 과 같다.

$$E\Delta T = \frac{\text{Proposed 인코딩시간}}{\text{HM 인코딩시간}} \% \quad (6)$$

$$D\Delta T = \frac{\text{Proposed 디코딩시간}}{\text{HM 디코딩시간}} \% \quad (7)$$

위의 실험 결과를 참조하면 테스트한 모든 영상에서 성능 향상을 보였음을 알 수 있다. 그리고 디코더 및 인코더에서 KLT 를 계산하지 않고 미리 계산된 값들을 사용하기 때문에 인코딩 시간 및 디코딩 시간이 크게 증가하지 않았음을 알 수 있다.

### 4. 실험 결과

제안하는 방법은 VTM10.0 에서 의미 있는 성능 향상을 보였으며 인코딩 및 디코딩 복잡도가 거의 증가하지 않았음을 보였다. 비선형 KLT 기저를 사용하기 위해서는 화면 내 픽처에서 유사블록 탐색과정이 필요할 것으로 보이고 현재 제안된 관련 기술의 복잡도 대비 성능 향상이 미비하기 때문에 인공지능을 활용한 방법을 통해 개선을 할 필요가 보인다고 생각한다.

### 5. 참조

[1] JVET-R2001, Versatile Video Coding(Draft 9); [http://phenix-it-sudparis.eu/jvet/doc\\_end\\_user/current\\_document.php?id=10155](http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=10155)

[2] X. Zhao, J. Chen, M. Karczewicz, L. Zhang, X. Li and W. Chien, "Enhanced Multiple Transform for Video Coding," 2016 Data Compression Conference (DCC), Snowbird, UT, 2016, pp. 73-82, doi: 10.1109/DCC.2016.9.

[3] G. Bjøntegaard, Calculation of Average PSNR Differences Between RD Curves, ITU-T Q.6/SG16, Document VCEG-M33, 2001.

### 6. 감사의 글

이 논문의 일부는 2020 년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (NRF-2018R1D1A1B07045156)