

비디오 코딩을 위한 필터링

임수연, 이영렬

세종대학교

syylim@sju.ac.kr, ylle@sejong.ac.kr

Filtering for Video Coding

SuYeon Lim, Yung-Lyul Lee

Sejong University

요 약

본 논문에서는 VVC(Versatile Video Coding) 화면 내 예측에서 참조 샘플 생성의 정확도를 높이기 위해 블록의 크기와 방향성 모드에 따라 더 많은 정수 위치 참조 샘플을 이용하는 보간 필터를 추가적으로 사용하는 방법을 제안한다. VVC 표준에서 4-tap 보간 필터를 사용하는 기존의 방식에 추가로 8-tap 보간 필터를 함께 사용하여 VVC 참조 소프트웨어인 VTM(VVC Test Model) 14.2[1] 대비 평균 -0.16% 의 luma BD-rate 개선을 보였다.

1. 서론

ITU-T VCEG(Video Coding Experts Group)과 ISO/IEC MPEG(Moving Picture Experts Group)은 차세대 비디오 표준 개발을 위해 2015 년 10 월에 JVET(Joint Video Exploration Team)을 결성하였고, VVC는 2020년 10월에 표준화 완료되었다.

VVC 화면 내 예측에서는 65 개의 일반 방향성 모드, 20 개의 wide angle 모드와 DC 모드, Planar 모드를 포함하여 총 87 개의 화면 내 예측 모드를 사용한다. 화면 내 예측의 성능 향상을 위해 MIP(Matrix-Based Intra Prediction), MRL(Multiple Reference Line), ISP(Intra Subpartition), PDPC(Position Dependent Prediction Combination) 등의 다양한 기술들이 개발되었다.

본 논문에서는 화면 내 예측 과정에서 참조 샘플을 생성할 때 4-tap DCT(Discrete Cosine Transform) 보간 필터와 4-tap Gaussian 보간 필터를 사용하던 기존의 방법에 추가로 블록 크기와 방향성 모드에 따라 선택적으로 8-tap DCT 보간 필터와 8-tap Gaussian 보간 필터를 적용하는 방법을 제안한다.

본 논문의 구성은 다음과 같다. 2 절에서는 VVC 의 화면 내

예측에서 참조 샘플을 생성하는 방법, 3 절에서는 본 논문에서 제안하는 방법을 설명한다. 4 절에서는 제안하는 방법의 성능을 실험을 통해서 확인한다. 마지막으로 5 절에서는 본 논문에 대한 결론을 맺는다.

2. VVC 참조 샘플 생성 방법

VVC 에서 사용하는 87 개의 모드 중 Vertical 모드, Horizontal 모드는 예측 샘플을 생성할 때 필터를 사용하지 않는다. Planar 모드와 Integer slope 을 갖는 45 도의 배각에 해당하는 $-14, -12, -10, -6, 2, 34, 66, 72, 76, 78, 80$ 모드들은 [1, 2, 1] 참조 샘플 필터를 사용하며, 보간 필터는 사용하지 않는다. 이 외의 모드들은 Fractional slope 을 갖게 되며, 정수 위치 참조 샘플에 보간 필터를 적용하여 현재 예측 샘플의 참조 샘플을 생성한다. 이때 4-tap DCT 보간 필터 또는 4-tap Gaussian 보간 필터가 사용된다[2].

VVC 의 4-tap DCT 보간 필터 계수는 DCT-II 와 IDCT-II(Inverse DCT-II)를 통해 유도된다. 한 예로서 표 1 은 $1/32$

Pixel 4-tap DCT 보간 필터 계수들 중 특정 16/32-pixel 위치에서의 계수를 나타낸다. 표 2 는 1/32 pixel 4-tap Gaussian 보간 필터 계수들 중 특정 16/32-pixel 위치에서의 계수를 나타낸다. 표 1 과 표 2 에서 index i 는 정수 화소의 위치를 의미한다.

표 1. VVC 의 4-tap DCT 보간 필터 계수

Index i	0	1	2	3
16/32-pixel filter[i]	-4	36	36	-4

표 2. VVC 의 4-tap Gaussian 보간 필터 계수[3]

Index i	0	1	2	3
16/32-pixel filter[i]	8	24	24	8

블록의 크기와 방향성 모드에 따라 두 개의 필터 중 현재 블록에서 사용할 필터가 결정된다.

$$nTbS = (\text{Log2}(W) + \text{Log2}(H)) \gg 1 \quad (1)$$

$$\text{minDistVerHor} = \text{Min}(\text{Abs}(\text{predModelIntra} - 50), \text{Abs}(\text{predModelIntra} - 18)) \quad (2)$$

W 는 블록의 너비, H 는 블록의 높이를 의미하며 블록의 크기에 따라 nTbS 값이 결정된다. predModelIntra 는 현재 블록의 예측 모드를 의미하고, 현재 블록의 예측 모드와 Vertical 모드 혹은 Horizontal 모드와의 차이 값으로 minDistVerHor 값이 결정된다.

표 3. nTbS 에 따라 필터 선택에 사용되는 threshold[2]

	intraHorVerDistThres[nTbS]
nTbS = 2	24
nTbS = 3	14
nTbS = 4	2
nTbS = 5	0
nTbS = 6	0
nTbS = 7	0

식 (1), (2)를 통해 계산된 블록의 nTbS 와 minDistVerHor 를 사용하여 필터를 선택한다. minDistVerHor 가 표 3 의 intraHorVerDistThres[nTbS]보다 큰 경우에는 Gaussian 보간

필터가 사용되고, 그렇지 않은 경우에는 DCT 보간 필터가 사용된다.

3. 제안하는 방법

기존의 VVC 에서는 모든 블록에 4-tap 보간 필터만 사용한다. 본 논문에서는 블록의 크기와 방향성 모드에 따라 추가적으로 8-tap DCT 보간 필터와 8-tap Gaussian 보간 필터를 사용하는 방법을 제안한다.

nTbS 가 2 인 블록에는 기존의 4-tap DCT 보간 필터를 대체하여 8-tap DCT 보간 필터를 사용하였고, nTbS 가 5 이상인 블록에는 기존의 4-tap Gaussian 보간 필터를 대체하여 8-tap Gaussian 보간 필터를 사용하였다.

표 4 는 1/32 pixel 8-tap DCT 보간 필터 계수들 중 특정 16/32-pixel 위치에서의 계수를 나타낸다. 표 5 는 1/32 pixel 8-tap Gaussian 보간 필터 계수들 중 특정 16/32-pixel 위치에서의 계수를 나타낸다.

표 4. 8-tap DCT 보간 필터 계수

Index i	0	1	2	3	4	5	6	7
16/32-pixel filter[i]	-3	11	-24	80	80	-24	11	-3

표 5. 8-tap Gaussian 보간 필터 계수

Index i	0	1	2	3	4	5	6	7
16/32-pixel filter[i]	2	14	42	70	70	42	14	2

그림 1 의 그래프는 4-tap DCT 보간 필터, 4-tap Gaussian 보간 필터, 8-tap DCT 보간 필터, 8-tap Gaussian 보간 필터의 16/32-pixel 위치에서의 Magnitude response 를 나타낸다. X 축은 0 과 1 사이의 값으로 정규화된 Frequency 를 나타내며, y 축은 Magnitude response 를 나타낸다.

파란색 그래프인 4-tap DCT 보간 필터와 노란색 그래프인 8-tap DCT 보간 필터를 비교하면, 8-tap DCT 보간 필터가 4-tap DCT 보간 필터보다 High frequency response 를 가진다는 것을 알 수 있다. 또한, 빨간색 그래프인 4-tap Gaussian 보간

필터와 보라색 그래프인 8-tap Gaussian 보간 필터를 비교하면, 8-tap Gaussian 보간 필터가 4-tap Gaussian 보간 필터보다 Low frequency response 를 가진다는 것을 알 수 있다.

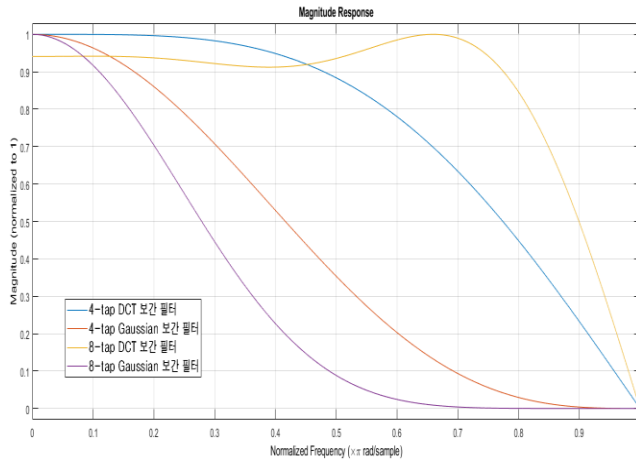


그림 1. 보간 필터의 Magnitude response

4. 실험 결과 및 분석

실험은 VVC 참조 소프트웨어인 VTM 14.2 버전에서 구현하였다. AI(All Intra) configurations 에서 진행되었으며, A1, A2, B, C, D 클래스들을 사용하였다. 시퀀스 마다 152 프레임 사용하였으며, AI configurations에서는 8 프레임마다 1 프레임을 인코딩하기 때문에 시퀀스 마다 19 프레임이 인코딩되었다. 시퀀스 마다 22, 27, 32, 37 QP(Quantization Parameter)를 사용하여 각각 실험하였다.

표 6 의 실험 결과를 통해, 본 논문에서 제안하는 방법을 사용할 때 평균적으로 -0.16% 의 luma BD-rate 성능 향상을 얻었음을 알 수 있다.

표 6. 제안하는 방법의 성능

All Intra Main 10					
Class	Y	U	V	EncT	DecT
A1	-0.16%	-0.35%	-0.23%	101%	103%
A2	0.00%	-0.05%	-0.09%	101%	102%
B	-0.02%	-0.06%	-0.05%	100%	105%
C	-0.31%	-0.25%	-0.25%	100%	103%
D	-0.29%	-0.47%	-0.47%	100%	103%
Average	-0.16%	-0.23%	-0.22%	100%	103%

5. 결론

본 논문에서는 기존 VVC 화면 내 예측에서 참조 샘플 생성에 사용하는 4-tap DCT 보간 필터와 4-tap Gaussian 보간 필터에 블록 크기와 방향성 모드에 따라 8-tap DCT 보간 필터와 8-tap Gaussian 보간 필터를 추가적으로 사용함으로써 성능을 향상시켰다.

감사의 글

이 논문의 일부는 2022 년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (NRF2018R1D1A1B07045156)

참고문헌

[1] Versatile Video Coding (VVC) reference software 14.2, "https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware_VTM/tags/VTM-14.2"

[2] J. Chen, Y. Ye, S. H. Kim "Algorithm description for Versatile Video Coding and Test Model 14 (VTM 14)", Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29 July 2021

[3] A. Filippov, V. Ruffitskiy "Non-CE3: Cleanup of interpolation filtering for intra prediction", document JVET - P0599, Oct. 2019