

VVC의 In-Loop Filter 기술

□ 박도현, 윤용욱, 김재곤 / 한국항공대학교

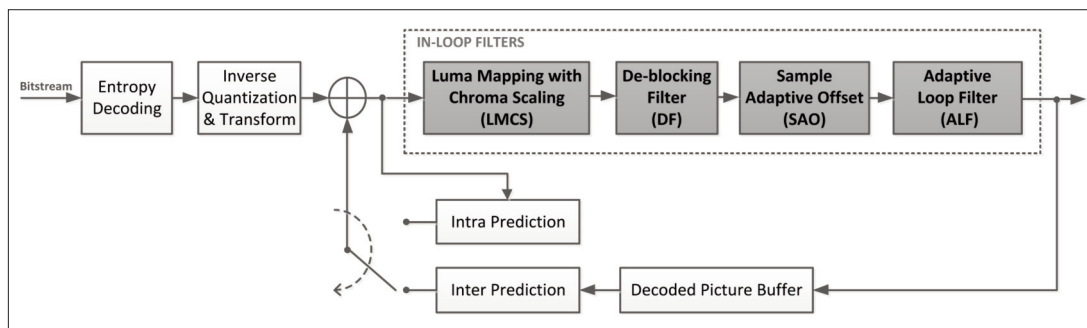
요약

JVET(Joint Video Experts Team)에서 새로운 비디오 압축 표준으로 진행 중인 VVC(Versatile Video Coding)에서는 HEVC(High Efficiency Video Coding)[1]의 기술을 근간으로 부호화 효율을 높일 수 있는 다양한 새로운 기술들을 채택하고 있다[2][3]. 인루프 필터(In-Loop Filter)는 복원영상의 화질을 향상시키기 위한 기술로 주관적 화질 개선뿐만 아니라 부호화 효율을 향상시키는 기술로 기존 HEVC의 확장 기술 및 새로운 인루프 필터 기술을 채택하고 있다. 본 고에서는 VVC의 CD에 채택되어 있는 인루프 필터 기술들을 소개한다 [4]. 인루프 필터 기술은 HEVC에 채택되어 있는 디블록킹 필터(Deblocking Filter: DF)와 SAO(Sample Adaptive Offset), 새로이 추가된 ALF(Adaptive Loop Filter)의 3가지의 필터와 LMCS(Luma Mapping with Chroma Scaling) 기술을 포함하고 있다. 이들 인루프 필터 기술은 주관적 화질 개선과 부호화 효율을 크게 개선하고 있으며, 2020년 7월 FDIS(Final Draft International Standard) 완료를 앞두고 인루프 필터링의 다양화로 인한 성능과 복잡도를 고려한 간소화 및 병렬처리 등의 고속화에 대한 표준화가 지속적으로 이루어질 전망이다.

1. 서론

인루프 필터링은 부/복호화가 끝난 복원영상의 화질을 향상시키기 위한 기술로 화질이 개선된 영상이 참조영상으로 이용됨으로써 부호화 효율 또한 향상된다. VVC에서는 HEVC에 채택된 디블록킹 필터, SAO와 함께 ALF가 추가로 채택되어 3가지 인루프 필터를 포함하고 있다. 또한, VVC의 참조 소프트웨어인 VTM(VVC Test Model)5.0부터 채택된 LMCS를 포함하고 있다.

〈그림 1〉은 VVC의 인루프 필터의 구성도이며 DF, SAO, ALF 순서로 적용된다. DF는 양자화로 인한 블록화 왜곡(blocking artifact)을 완화하기 위하여 블록 경계에 있는 화소에 적용되며, SAO는 DF가 수행된 후 모든 화소에 적용되어 링잉 현상(ringing artifacts)을 개선할 수 있다. DF와 SAO는 HEVC의 DF와 SAO와 거의 유사



〈그림 1〉 VVC 복호화기에서의 인루프 필터 구성도

하며 성능 및 복잡도가 개선된 형태로 VVC에 채택되었다. SAO 필터는 복원영상의 모든 화소를 정해진 기준에 따라 특성별로 분류하고, 그 분류에 따라 적응적으로 오프셋 값을 적용하여 원본영상과의 차이를 감소시킨다. 윤곽선(edge)에 대하여 오류를 보정하는 에지 오프셋(Edge Offset)과 특정 밝기 값 대역에 대하여 왜곡을 보상하는 밴드 오프셋(Band Offset)을 적응적으로 적용한다. ALF는 SAO가 수행된 복원영상에 적용되어 최종적으로 압축 손실을 보상해주는 기술이다. ALF는 Wiener Filter를 기반으로 원본영상과의 오차를 최소화함으로써 부호화 복잡도가 크지만 우수한 부호화 효율을 가진다. HEVC 표준화 당시에 논의되었던 ALF를 근간으로 복잡도 개선과 보다 정밀하게 필터링을 수행할 수 있도록 설계되었다.

VVC에 새로이 채택된 LMCS는 복원된 영상에 직접적으로 적용하는 일반적인 인루프 필터링과 달리, 영상 신호의 동적 범위(dynamic range)에 대한 코드워드(codeword)를 재분배함으로써 효율적인 예측 및 양자화가 가능하게 하여 부호화 성능 향상 및 화질 향상을 얻는다. LMCS 휘도신호 매핑과 색차신호 스케일링을 포함하며 휘도신호 매핑은 입

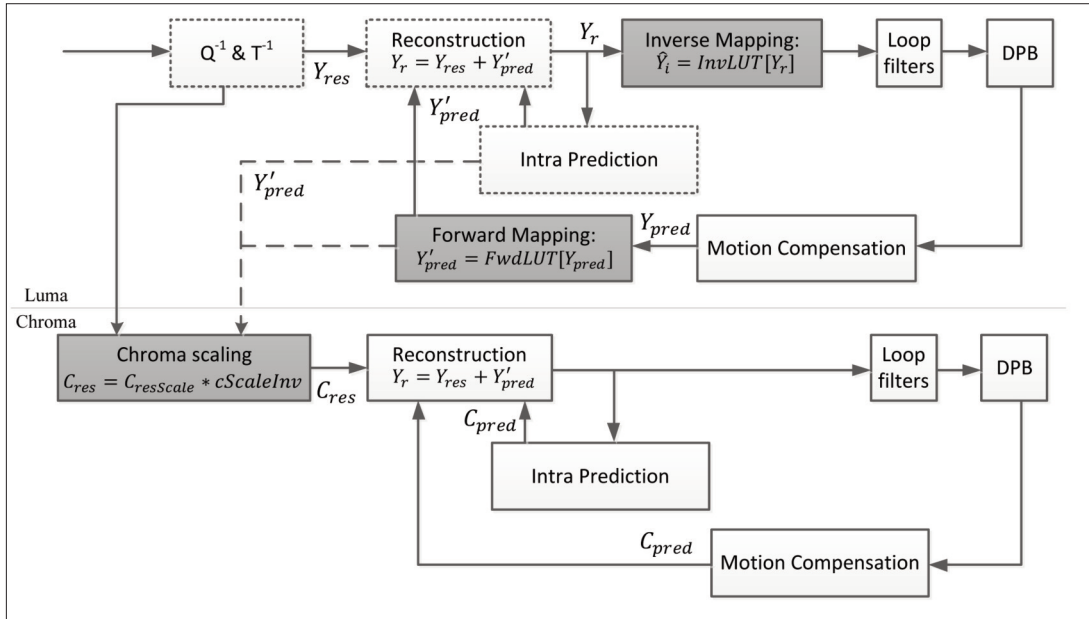
력 영상의 휘도신호에 대한 동적 범위를 매핑을 통하여 재구성하는 기술이고, 색차신호 스케일링은 매핑된 휘도성분과 색차신호 사이의 간극을 보상하기 위한 기술이다.

II. VVC의 인루프 필터링

본 장에서는 최근에 배포된 VTM6.0에 채택되어 있는 VVC의 인루프 필터링 기술인 DF, SAO, ALF, LMCS 기술의 상세내용을 살펴본다[5].

1. LMCS(Luma Mapping with Chroma Scaling)

LMCS는 보다 효율적인 화소 값의 동적 범위에서 부호화를 수행하게 하여 부호화 성능 및 주관적 화질을 향상하기 위한 기술로 휘도신호 매핑과 색차신호 스케일링을 포함한다. 〈그림 2〉는 LMCS에 따른 부호화 구조의 동적 영역을 나타낸 것이다. 〈그림 2〉에서 점선 블록으로 표시된 역양자화, 역변환, 색차신호의 화면내 예측 및 휘도신호의 복원 프로세스는 매핑된 동적 영역에서 수행되고, 실



〈그림 2〉 LMCS에 따른 복호화 구조의 동적 영역

선 블록으로 표시된 DF, ALF 및 SAO의 루프 필터, 움직임 보상, 색차신호의 화면내 예측 및 색차신호의 복원 프로세스는 기존 동적 영역에서 수행된다. 또한, 회색 블록과 같이 입력 영상의 동적 영역과 매핑된 동적 영역은 휘도신호 매핑과 색차신호 스케일링을 통하여 연결된다. LMCS는 SPS (Sequence Parameter Sets)에 포함된 선택스 정보 (sps_lmcs_enabled_flag)를 통해서 시퀀스 단위로 활성화시킬 수 있다.

1) 구간 선형모델(piece-wise linear model)을 이용한 휘도신호 매핑

휘도신호 매핑은 입력되는 휘도신호의 동적 범위에 대한 코드워드를 부호화 성능을 향상시킬 수 있는 코드워드로 재분배하는 인루프 매핑을 의미한다. 순방향 매핑(forward mapping)과 이에 대응하는 역방향 매핑(backward mapping)을 통해 휘

도신호 매핑을 수행할 수 있다. 순방향 매핑은 기존 동적 영역을 동일한 16개의 구간으로 분할한 뒤 각 구간에 대한 선형모델을 통해서 입력 영상의 코드워드 재분배를 수행하고 역방향 매핑은 매핑된 동적 영역에서 기존 동적 영역으로 역매핑을 수행한다. 예를 들어, 10비트 입력 영상에 대한 16개의 구간은 각각 64개의 코드워드가 기본적으로 할당된다. 실제로 전송되는 데이터는 16개의 각 구간에 대한 코드워드의 개수이며 이에 대한 선형모델로 매핑된 동적 영역 구간의 코드워드를 계산한다. 순방향 매핑을 위한 선형모델은 슬라이스(Slice) 레벨로 전송되며 역방향 매핑을 위한 선형모델은 순방향 매핑의 선형모델로부터 유도된다. 수식 (1)은 순방향 매핑의 선형모델을 계산하기 위한 식이고 이를 통해 입력 휘도신호의 코드워드를 전송된 동적 영역의 코드워드로 매핑한다.

$$FwdMap(Y_{pred}) = ((b2 - b1) / (a2 - a1)) * (Y_{pred} - a1) + b1 \quad (1)$$

Y_{pred} 는 입력 영상의 휘도성분에 대한 코드워드를 의미하고, $a1$, $a2$ 는 Y_{pred} 를 포함하는 특정 동적 영역 구간의 첫 코드워드와 마지막 코드워드이며, $b1$, $b2$ 는 해당하는 매핑된 동적 영역 구간의 첫 코드워드와 마지막 코드워드이다. 순방향 매핑과 역방향 매핑을 포함하는 휘도성분의 매핑 프로세스는 위 수식을 통해 직접 계산하거나 LUT(Look-Up Table)로 구현될 수 있다.

〈그림 2〉와 같이, 두 위치에서 휘도신호 매핑을 수행한다. 첫번째로 휘도신호에 대한 예측 블록이 DPB(Decoded Picture Buffer)에 존재하며 입력 신호의 동적 영역을 갖는 참조 신호로부터 생성된 후, 선형모델 매핑을 수행한다. 즉, 움직임 보상은 매핑된 동적 영역에서 수행된다. 하지만, 화면내 예측의 경우 매핑된 동적 영역에서 수행되기 때문에 순방향 매핑을 수행하지 않는다. 두번째로 매핑된 동적 영역에서 화면내 예측 또는 화면간 예측으로 복원된 휘도신호를 원본 신호의 동적 영역으로 다시 변환하기 위해서 역방향 매핑을 수행한다.

2) 색차 잔차신호 스케일링

색차신호 스케일링은 휘도신호와 대응하는 색차신호 사이의 상호 관계에 따라 색차신호를 보정한다. 색차신호 스케일링의 수행 여부는 타일 그룹 레벨로 휘도 매핑이 수행되고 현재 타일 그룹이 듀얼 트리(dual tree)일 경우 전송된다. 또한, 면적이 4 이하인 색차신호 블록에 대해서 색차신호 스케일링은 비활성화된다.

색차 잔차신호(residual signal) 스케일링 요소, C_{Scale_Inv} 는 대응되는 휘도 예측블록의 평균값에 기반하여 유도된다. 휘도 예측신호의 평균값의 역방

향 매핑을 통해서 얻은 동적 범위에서의 예측블록의 평균값을 얻고 휘도신호와 스케일링 요소 사이의 관계를 정의한 테이블에 따라 C_{Scale_Inv} 를 유도한다. 유도된 스케일링 팩터를 이용하여 다음과 같이 색차 잔차신호에 대한 스케일링을 수행한다. 수식 (2)는 스케일링을 위한 수식이며 수식 (3)은 역스케일링을 위한 수식이다. C_{Res} , $C_{ResScale}$ 는 색차 잔차신호와 스케일링된 잔차신호를 의미한다.

$$C_{ResScale} = C_{Res} * C_{Scale} = C_{Res} / C_{Scale_Inv} \quad (2)$$

$$C_{Res} = C_{ResScale} / C_{Scale} = C_{ResScale} * C_{Scale_Inv} \quad (3)$$

2. 디블로킹 필터(Deblocking Filter: DF)

손실 압축을 수행하는 각 블록의 양자화 파라미터(QP) 값의 차이에 따라 양자화로 인한 오차가 다르게 발생하게 되고, 이로 인하여 복호화된 영상에 블록 경계가 보이는 블록 열화(blocking artifact)가 발생한다. 이러한 블록 경계에서 발생하는 블로킹 아티팩트를 제거하기 위해 복원영상에 DF를 수행한다. 필터링은 수직 경계와 수평 경계에서 수행되며 수직 경계에서 먼저 필터링이 수행된 후 수직 경계 필터링의 결과에 수평 경계 필터링이 수행된다. VVC의 DF 기술은 전반적으로 HEVC의 DF 기술과 유사하지만 새로운 비디오 코덱에 적합하도록 세부적인 요소에서 발전이 이루어졌다.

1) 경계 및 필터 길이 결정

HEVC는 8x8 화소 단위의 필터링을 수행한다. 따라서, 블록 경계의 최소 단위는 4 화소이기 때문에 실제 블록의 경계임에도 불구하고 필터링이 수행되지 않는 경계가 발생할 수 있다. 이러한 문제를

해결하기 위하여 VVC는 블록 경계의 최소 단위인 4x4 화소 단위로 필터링을 수행하며 변환 블록 경계에서 필터링을 수행한다. 또한, 수행할 블록에 따라 최대 필터링 길이를 제한하여 발생 가능한 모든 블록 경계에 대해 병렬 처리가 가능한 필터링이 수행될 수 있도록 설계되었다. 필터링을 수행할 샘플의 위치는 <그림 3>과 같이 정의된다.

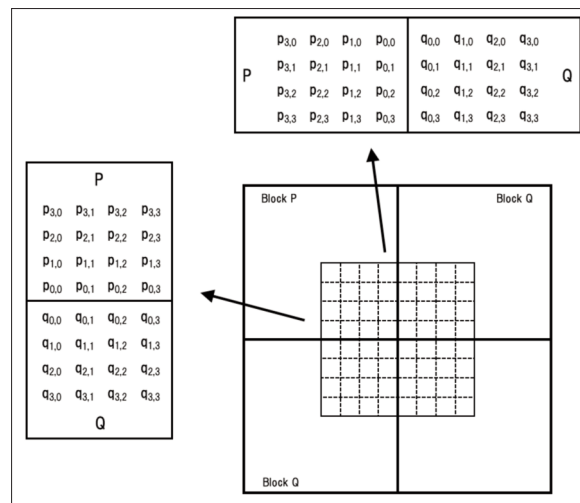
필터링을 수행할 경계가 정해졌다면, 블록 P와 Q의 적용될 필터의 최대 길이를 결정한다. 필터 길이는 각 블록의 크기와 영상 성분에 따라 다르게 결정된다. 필터링을 수행할 블록이 휘도성분일 때, 블록의 크기가 32 화소 이상이라면 최대 필터 길이는 7로 결정되고 그렇지 않으면 3으로 결정된다. 색차성분일 때, 블록 크기가 8 화소 이상인 경우 3으로 결정되고 8 화소 미만의 경우 1로 결정된다.

VVC는 SbTMVP(Subblock-based Temporal MV Prediction)와 Affine 예측과 같이 서브 블록 단위로 예측을 수행하는 블록 사이의 경계에 대해

서도 DF를 수행한다. 서브 블록 사이의 경계에 대한 필터링은 휘도성분인 경우에만 수행되며 최대 필터 길이는 필터링을 수행할 서브 블록의 경계와 변환 블록 경계 사이의 거리에 따라 최대 필터 길이가 결정된다. 두 블록의 경계가 같은 경우, 최대 필터 길이는 변환 블록과 같은 방법으로 결정되지만 필터의 최대 길이가 5로 제한된다. 경계 사이의 거리가 4 화소 떨어진 경우 최대 필터의 길이는 1로 결정되고 경계 사이의 거리가 8 화소 이상 떨어진 경우 최대 필터의 길이는 2로 결정된다.

2) 필터링 파라미터 유도

필터링이 수행될 수 있는 경계가 결정되면, 필터의 세기와 화소의 변화에 대한 클리핑(clipping) 값을 결정하는 파라미터 bS , β , tc 를 유도한다. 필터링 파라미터 bS 는 블록 경계의 강도를 의미하는 파라미터이다. 값이 클수록 블록 경계가 강하게 나타날 수 있으며 bS 가 0일 경우에는 DF를 수행하지 않



<그림 3> LMCS에 따른 복호화 구조의 동적 영역

는다. bS는 다음 표에 따라 결정된다.

필터링 파라미터 β , t_c 는 필터의 강도를 결정하거나 필터링을 수행할 때 적용되는 임계값을 의미한다. β' , t_c' 는 <표 1>과 같이 QP에 기반하여 정의된다. β' , t_c' 는 8 비트 영상 기준에 대한 값이고 입력 영상의 비트 심도에 따라 스케일링되어 사용된다. VVC는 51까지의 QP를 제공하는 HEVC보다 확장된 63까지의 QP를 제공하기 때문에 이에 따라 매핑 테이블도 확장되었다.

VVC에서는 복원 신호의 평균 휘도레벨(LL)에 따른 오프셋을 QP에 더해 DF를 조절할 수 있다. 평균 휘도레벨에 따른 QP 오프셋은 매핑 테이블을 통해 결정될 수 있으며, 영상특성(HDR, SDR)에 따라 다른 매핑 테이블이 적용된다. DF를 위한 QP와 평균 휘도레벨은 다음과 같이 계산된다.

$$QP_L = ((QP_Q + QP_P + 1) \gg 1) + qpOffset \quad (4)$$

$$LL = ((P_{0,0} + P_{0,0} + P_{0,0} + P_{0,0}) \gg 2) / (1 \ll \text{bitDepth}) \quad (5)$$

3) 필터 강도 결정 및 필터링

기존 HEVC의 DF는 휘도영상에 적용하는 2개의 필터와 색차영상에 적용하는 1개의 필터가 정의되어 있다. VVC의 DF는 영상 성분마다 필터가 하나씩 추가되어 휘도성분에 대한 3가지의 필터, 색차성분에 대한 2가지의 필터가 존재한다. 휘도성분에선형 필터가 추가되었으며 <그림 4>와 같이 필터의 강도가 선택된다.

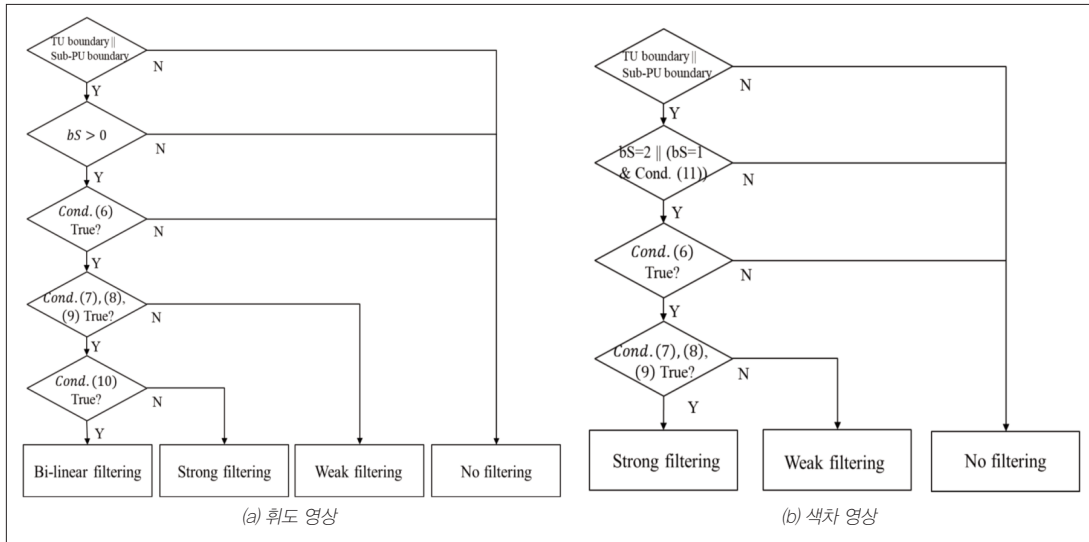
<그림 4>에 있는 조건들은 아래와 같다. 아래의 조건에서의 화소 값 및 위치는 블록의 최대 필터 길이에 따라서 추가되거나 확장될 수 있다.

<표 1> 블록 경계 강도(bS) 값 및 조건

Priority	Conditions	bS (Luma)	bS (chroma)
1	$(P == \text{intra}) \vee (Q == \text{intra})$	2	2
2	$(\text{cbf}_P == 1) \vee (\text{cbf}_Q == 1)$	1	1
3	$ MV_P - MV_Q \geq 1/2 \text{ sample}$	1	N/A
4	$MVP_P \neq MVP_Q$	1	N/A
5	Otherwise	0	0

<표 2> 필터링 파라미터(β , t_c) 매핑 테이블

QP	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
β'	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6
t_c	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
QP	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33
β'	7	8	9	10	11	12	13	14	15	16	17	18	20	22	24	26	28
t_c	0	3	4	4	4	4	5	5	5	5	7	7	8	9	10	10	11
QP	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
β'	30	32	34	36	38	40	42	44	46	48	50	52	54	56	58	60	62
t_c	13	14	15	17	19	21	24	25	29	33	36	41	45	51	57	64	71
QP	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65		
β'	64	66	68	70	72	74	76	78	80	82	84	86	88	-	-		
t_c	80	89	100	112	125	141	157	177	198	222	250	280	314	352	395		



〈그림 4〉 필터 강도 결정

$$|p_{2,0}-2p_{1,0}+p_{0,0}|+|p_{2,3}-2p_{1,3}+p_{0,3}|+|q_{2,0}-2q_{1,0}+q_{0,0}|+|q_{2,3}-2q_{1,3}+q_{0,3}| < \beta \quad (6)$$

$$|p_{2,i}-2p_{1,i}+p_{0,i}|+|q_{2,i}-2q_{1,i}+q_{0,i}| < \beta/8, \quad i=0,3 \quad (7)$$

$$|p_{3,i}-p_{0,i}|+|q_{0,i}-q_{3,i}| < \beta/8, \quad i=0,3 \quad (8)$$

$$|p_{0,i}-q_{0,i}| < 2.5t_c, \quad i=0,3 \quad (9)$$

$$\begin{aligned} \text{height(or width)} &\geq 32, \\ \text{when horizontal(or vertical) edge} \end{aligned} \quad (10)$$

$$\begin{aligned} \text{height(or width)} &\geq 8, \\ \text{when horizontal(or vertical) edge} \end{aligned} \quad (11)$$

필터 강도가 결정되면, 결정된 필터에 맞는 필터링을 수행한다. 휘도성분에 대한 강한 필터링과 약한 필터링은 HEVC와 같다[5]. 선형 필터링은 블록

의 최대 필터 길이 안에 해당하는 화소를 선형 보간된 값으로 대체하는 필터링 기술이다. 선형 보간된 화소 값은 다음과 같이 계산된다. 다음 수식은 P 블록에 대한 경우이며 Q 블록도 마찬가지로 수행된다.

$$p'_i = (f_i * \text{Middle}_{s,i} + (64 - f_i) * P_s + 32) \gg 6 \quad (12)$$

f 는 선형 필터의 계수이며 최대 필터 길이에 따라 필터의 길이와 계수가 결정된다. 최대 필터 계수가 7, 5 또는 3일 경우 각각 $f=\{59,50,41,32,23,14,5\}$, $f=\{58,45,32,19,6\}$, $f=\{53,32,11\}$ 과 같이 정의된다. $\text{Middle}_{s,i}$ 는 필터링이 수행되는 위치의 화소들의 값의 가중 평균으로 계산된다. P_s 는 필터링이 수행되는 화소에서 경계로부터 가장 먼 화소와 그 이전 화소의 평균이다. i 는 1에서 최대 필터 길이까지를 의미한다.

색차성분에 대한 약한 필터링은 HEVC의 방법과

같다[5]. 색차성분에 대한 강한 필터링은 다음 수식을 통해 수행된다. 다음 수식은 P 블록 필터링에 대한 경우이고 Q 블록도 마찬가지로 수행된다.

$$p_2' = (3*p_3 + 2*p_2 + p_1 + p_0 + q_0 + 4) \gg 3 \quad (13)$$

$$p_1' = (2*p_3 + p_2 + 2*p_1 + p_0 + q_0 + q_1 + 4) \gg 3 \quad (14)$$

$$p_0' = (p_3 + p_2 + p_1 + 2*p_0 + q_0 + q_1 + q_2 + 4) \gg 3 \quad (15)$$

3. SAO(Sample Adaptive Offset)

SAO는 원본영상과 복원영상 간의 왜곡을 화소 단위로 보상함으로써, 주관적 화질 향상뿐만 아니라 부호화 효율 또한 향상시킨다. 비디오 부호화 과정에서, 변환 및 양자화 과정은 원본 화소의 손실을 야기한다. 이때, 높은 QP로의 양자화는 입력영상의 에지 영역에서 링잉 현상 및 평탄한 영역에서 밴딩(banding) 현상이 발생할 수 있다. 이러한 왜곡들은 주관적 화질 저하를 야기할 수 있다. 이에, SAO는 에지의 방향성을 구분하고 주변 화소에 따라 현재 화소를 오프셋으로 보상함으로써 원본영상과의 왜곡을 최소화한다. 따라서, 주관적 화질뿐만 아니

라 객관적 화질(PSNR) 또한 향상시킬 수 있다. VVC의 SAO는 HEVC의 SAO와 동일하게 수행되며 DF가 완료된 블록에 대해 수행된다[6].

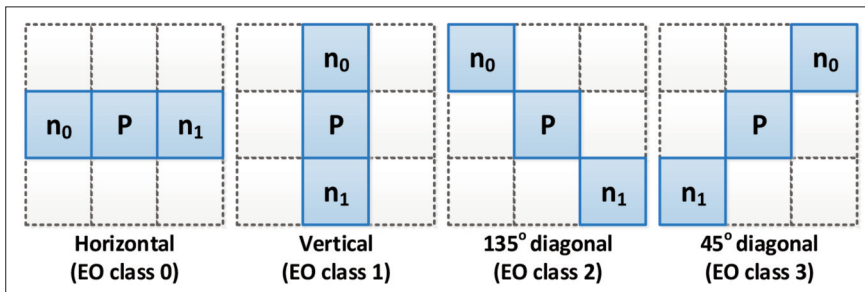
1) SAO Offset의 종류

SAO는 원본영상과 복원영상 간의 왜곡을 최소화하기 위한 오프셋을 계산하여 전송한다. 이때, 화소 단위로 결정된 오프셋이 아닌 CTU(Coding Tree Unit) 단위로 각 성분에 대한 4개의 오프셋이 결정된다. 다시 말해, SAO는 CTU 단위로 수행되며, 휘도 및 색차성분 각각에 대한 4개의 오프셋을 전송한다.

SAO는 CTU 내 화소의 특성에 따라 EO(Edge Offset)와 BO(Band Offset)로 세분화된다. EO는 CTU 내의 에지를 판단하여 에지에 대한 왜곡을 보정하고, BO는 특정 밝기 영역에 대한 왜곡을 보정하는데 사용된다. 따라서, 인코더는 SAO의 수행 여부를 결정하는데 있어, EO와 BO 그리고 SAO를 적용하지 않았을 때의 윌-왜곡 비용(R-D cost)을 비교하여 SAO 수행여부를 결정한다.

2) EO(Edge Offset)

EO는 CTU의 에지 방향에 따라, 화소의 왜곡을



〈그림5〉 Edge Offset의 패턴 클래스

〈표 3〉 Edge Offset의 카테고리 분류 방법

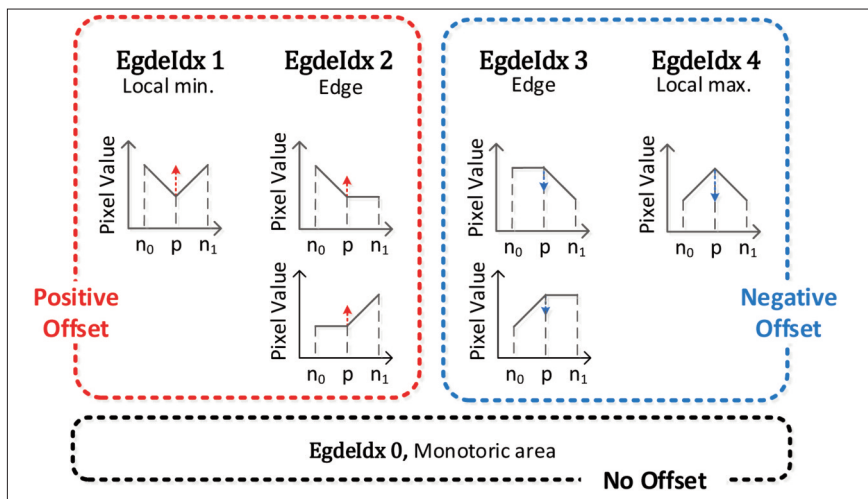
Idx	Category conditions	offset
1	$p < n_0 \ \&\& \ p < n_1$	O1
2	$(p < n_0 \ \&\& \ p == n_1) \ \ (p == n_0 \ \&\& \ p < n_1)$	O2
3	$(p > n_0 \ \&\& \ p == n_1) \ \ (p == n_0 \ \&\& \ p > n_1)$	O3
4	$p > n_0 \ \&\& \ p > n_1$	O4
0	otherwise	-

보정한다. EO의 에지 패턴은 〈그림 5〉와 같이 수평 에지, 수직 에지, 2개의 대각선 에지로 총 4개의 패턴으로 분류된다. 따라서, CTU가 EO를 수행하는 경우 선택된 클래스 정보와 해당 클래스에 대한 4개의 오프셋이 전송된다. 즉, 인코더에서는 CTU를 부호화함에 있어, 유효-왜곡 비용 비교를 통해 가장 최적의 클래스를 결정하며 결정된 클래스 정보와 해당 클래스의 4개의 오프셋을 전송한다. 전송되는 4개의 오프셋은 주변 화소의 패턴에 의해 분류된다. 〈표 3〉은 오프셋이 적용되는 카테고리 분류 방법을 보여주며, 〈그림 6〉은 오프셋

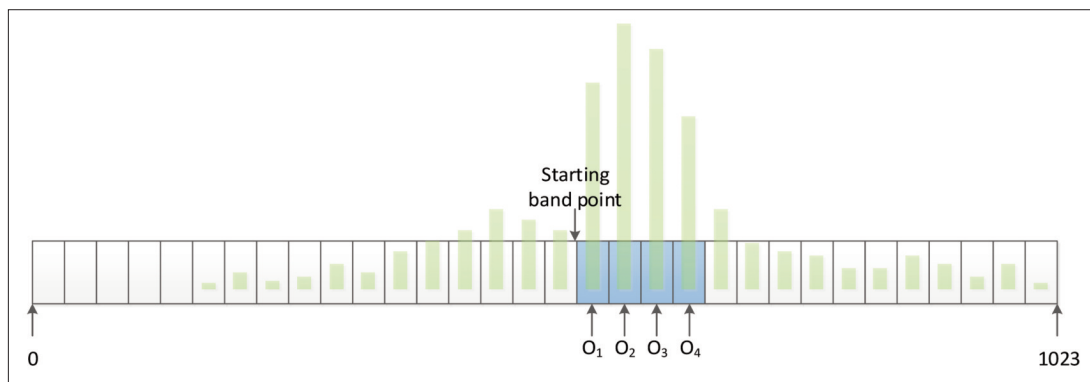
의 적용 예를 보여준다. 디코더는 먼저 클래스 정보를 파싱(parsing)하고, 〈표 3〉에 따라 CTU 내의 화소마다 카테고리를 결정하고 알맞은 부호를 적용한 오프셋을 더해줌으로써 영상 복원을 수행하게 된다.

3) BO(Band Offset)

BO는 부호화하는 CTU 내 화소들이 갖는 밝기 값을 특정 대역으로 분류하여 화소의 왜곡을 보정한다. 밝기 대역은 입력영상의 비트 심도를 32개의 구간으로 분할했을 때, 연속된 4개의 구간으로 구성된다. 따라서 전송되는 4개의 오프셋은 연속된 4개의 구간에 대한 오프셋으로 구성된다. 입력 영상이 10비트 심도를 갖을 경우, 0~1023의 값을 가질 수 있으며 이는 동일한 화소 범위를 갖는 32개의 구간으로 분할된다. 이 중 연속된 4개의 구간에 대해 오프셋이 부여된다. 이때, 선택된 대역의 시작구간을 알려주기 위한 정보가 전송된다. 즉, 인코더에서는



〈그림 6〉 Edge Offset의 적용 예



〈그림 7〉 Band Offset의 예

울-왜곡 비용 비교를 통해 가장 최적의 대역을 결정하며 결정된 대역의 시작점과 4개의 구간 각각에 대한 오프셋을 전송한다. 디코더는 시작점 정보를 파싱하고, CTU 내 화소마다 해당하는 구간에 대한 오프셋을 더해줌으로써 영상 복원을 수행하게 된다. 〈그림 7〉은 10 비트 영상에서 BO의 예를 보여준다.

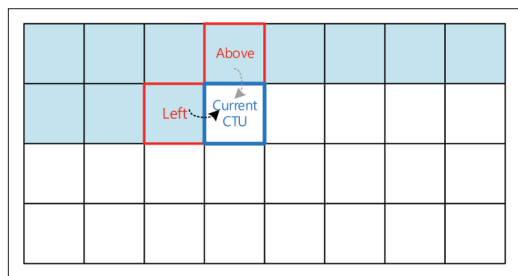
4) SAO Merge

SAO는 CTU단위로 SAO의 사용여부를 나타내는 정보, 클래스 정보, 4개의 오프셋 정보 등 다수의 정보를 전송해야 한다. 이러한 시그널링 오버헤드를 줄이고자 주변에 존재하는 SAO 정보를 활용하는 기법이 적용된다. 즉, 이미 부/복호화된 왼쪽

혹은 위쪽의 CTU로부터 SAO 정보를 받아 현재 CTU의 SAO를 수행한다. 〈그림 8〉은 SAO Merge의 예를 보여준다. 인코더는 왼쪽 혹은 위쪽 CTU의 SAO 정보를 이용하여 현재 SAO를 수행해보고 울-왜곡 비용 비교를 통해 최적일 경우 왼쪽 혹은 위쪽의 정보를 사용한다는 지시자를 전송한다. 이에 디코더는 전송된 지시자를 파싱하여 왼쪽 혹은 위쪽의 정보를 사용할 경우, 해당 위치의 SAO 정보를 이용하여 현재 CTU에 SAO를 수행하여 복원한다.

4. ALF(Adaptive Loop Filter)

ALF는 오랜 기간 비디오 부호화 효율 향상 및 주관적 화질 개선을 위해 연구되어져 왔고, HEVC 표준에서 또한 논의가 되었던 바가 있다[8]. ALF는 SAO와 마찬가지로 원본영상과 복원영상 간의 왜곡을 직접 계산하여 보상한다. SAO와의 차이점이 라면 블록 단위의 필터링을 수행하며 해당 블록의 에지 및 변화량의 정도를 구분하여 서로 다른 필터를 적용하여 왜곡을 보상한다. ALF는 VVC의 인트루프 필터링 기술 중 가장 마지막에 수행되며, SAO



〈그림 8〉 SAO merge의 예

가 수행된 샘플에 대해 ALF가 적용된다. HEVC 표준 당시에 논의되었던 ALF를 근간으로 복잡도 측면에서의 감소와 주변 화소와의 상관성을 높여 보다 정밀하게 필터링을 수행할 수 있도록 설계되었다[4][8].

ALF는 Wiener filter를 기반으로 동작하는 필터링 방법으로, 원본영상과 복원영상과의 오차를 최소화시키는 최적의 필터 계수를 유도하여 필터링을 수행하며 유도된 필터 계수를 인코딩하여 전송한다. 필터 계수는 식 (16)에 의해 유도된다.

$$c_i = \operatorname{argmin}_E \left[\left(x - \sum_{i=0}^{N-1} y_i c_i \right)^2 \right] \quad (16)$$

여기서, x 는 원본 영상의 화소값, c_i 는 i 번째 필터 계수, N 는 필터 계수의 수, y_i 는 복원영상의 화소값을 의미한다.

VVC의 ALF는 CTB 단위로 하나의 필터셋을 가지고 있으며, 필터셋은 최대 25가지의 클래스로 구성되어 있다. 4x4 블록 단위로 블록의 특성을 파악하여 클래스를 분류하고, 필터 계수를 계산 및 산출

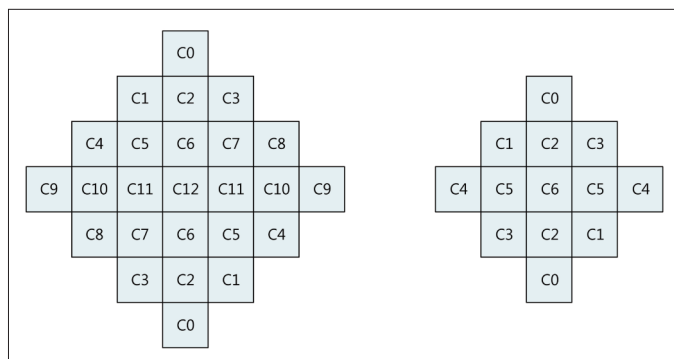
하여 ALF를 수행한다.

1) 필터 모양

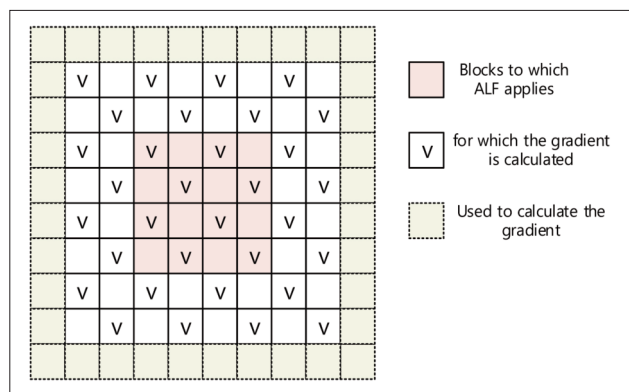
VVC의 ALF는 <그림 9>와 같이 휘도 성분에 대해서는 7x7 다이아몬드 모양의 필터를 사용하고, 색차 성분에 대해서는 5x5 다이아몬드 모양의 필터를 사용하여 필터링을 수행한다.

2) 블록 분류

VVC의 ALF는 4x4 블록 단위의 기울기(gradient)를 활용한 블록내 화소의 방향성(D, Direction)과 변화량(A, Activity)에 따라 25가지의 클래스로 분류된다. 방향성은 4x4 블록을 포함한 주변 화소들로부터 1D-Laplacian을 이용한 수평, 수직, 그리고 2 방향의 대각선 기울기를 계산하여 5가지의 방향성으로 분류된다. 클래스 분류의 복잡도를 고려하여 <그림 10>과 같이 주변 전체의 기울기가 아닌 일정 위치에서의 기울기가 계산된다. 변화량은 방향성 분류를 위해 계산된 기울기를 활용해 2D-Laplacian 연산을 수행하여 5가지의 변화량으로 분류된다. 따라서, 각 4x4 블록은 5가지의 방향성과 5가지의 변화량에 의



<그림 9> ALF 필터 모양 (7x7: 휘도 성분, 5x5: 색차 성분)



〈그림 10〉 ALF의 블록 분류를 위한 기울기 계산 샘플 위치

해 총 25가지의 클래스 중 하나의 클래스로 분류된다.

3) 기하학적 변환

4x4 블록의 특성에 따라 다양한 필터를 적용하기 위해, 블록 분류에서 계산된 기울기에 따라 선택된 필터에 대해서 회전, 수직 대칭, 대각선 대칭을 수행한다. 〈표 4〉는 기하학적 변환 조건을 보여준다. 여기서, g_{dl} , g_{d2} , g_v , g_h 는 각각 2개의 대각선과 수직, 수평 기울기를 의미한다.

〈표 4〉 기울기에 따른 변환 방법

Gradient values	Transformation
$g_{d2} < g_{d1} \ \&\& \ g_h < g_v$	No transformation
$g_{d2} < g_{d1} \ \&\& \ g_h < g_v$	Diagonal
$g_{d2} > g_{d1} \ \&\& \ g_h < g_v$	Vertical flip
$g_{d2} > g_{d1} \ \&\& \ g_h > g_v$	Rotation

4) 비선형 필터링

ALF는 원본 영상과 복원영상으로부터 오차를 최소화시키는 최적의 계수를 유도하여 필터링을 수행한다. 유도된 계수를 이용하여 4x4 블록의 각 화소

는 식 (17)에 의해 필터링된다.

$$O(x, y) = w_{(0,0)} \cdot I_{(x,y)} + \sum_{(i,j) \neq (0,0)} w_{(i,j)} \cdot I_{(x+i,y+j)}, \quad w_{(0,0)} \quad (17)$$

$$= 1 - \sum_{(i,j) \neq (0,0)} w_{(i,j)}$$

여기서, (x,y) 는 현재 화소의 위치, (i,j) 는 주변 화소의 위치, w 는 필터의 계수, I 는 화소값을 의미한다. 즉, 선형 방정식에 의해 현재 화소가 필터링된다. 이 때, 식 (17)은 식 (18)과 같이 표현할 수 있다.

$$O(x, y) = I_{(x,y)} + \sum_{(i,j) \neq (0,0)} w_{(i,j)} \cdot (I_{(x+i,y+j)} - I_{(x,y)}) \quad (18)$$

식 (18)과 같이 필터링된 화소 $O(x,y)$ 는 현재 화소, 현재 화소와 주변 화소의 차이의 선형 조합으로 표현할 수 있다. 식 (18)로부터 만약 현재 화소와 주변 화소의 차이가 클 때, 주변 화소의 영향을 제한

$$O(x,y) = I_{(x,y)} + \sum_{(i,j) \neq (0,0)} w_{(i,j)} \cdot \text{clip}(-cValue, cValue, I_{(x+i,y+j)} - I_{(x,y)}) \quad (19)$$

해주기 위하여 현재 화소와 주변 화소의 차이에 클리핑을 수행하여 필터링되는 화소에 비선형성을 부여할 수 있다. 따라서 필터링되는 화소 $O(x,y)$ 는 최종적으로 식 (19)와 같이 표현된다.

여기서, $\text{clip}()$ 은 클리핑 연산을 의미하며, $cValue$ 는 클리핑의 정도를 나타내주는 값이며, 현재 VVC에서는 입력 영상 비트심도에 따라 4개의 후보를 구성하여, 인코더에서는 가장 최적의 $cValue$ 를 선택하여 디코더로 전송한다.

5) 필터 파라미터 전송

ALF의 필터 파라미터는 APS(Adaptive Parameter Set)로 전송된다. 하나의 APS는 최대 25가지의 휘도성분 필터 계수와 클리핑 값, 최대 8가지의 색차성분 필터 계수를 포함하여 전송된다. ALF는 CTB 레벨에서 수행여부를 결정한다. 따라서 매 CTB에 대하여 ALF의 적용여부를 알려주는 플래그(flag)가 전송된다. 또한, 하나의 CTB는 16개의 고정된 필터와 APS로 구성된 필터 중 어떤 필터를 사용할 지 결정하여 결정된 인덱스가 전송된다. APS는 최대 7개로 구성될 수 있으며, 해당 인덱스는 슬라이스 헤더로 전송된다. 이때, 16개의 고정된 필터

는 인코더 및 디코더에 사전 정의가 되어 있다.

III. 결 론

본 고에서는 VVC의 VTM6.0에 포함된 인루프 필터 기술에 대해 살펴보았다. 더블로킹 필터링과 SAO만을 수행하던 HEVC 표준의 인루프 필터링을 확장하여 압축 왜곡을 보다 효과적으로 보상하는 ALF 필터가 추가되었다. 또한, 화소값의 동적 영역을 조정을 통하여 예측, 변환, 양자화 부분에 직접적으로 관여하여 부호화 성능 자체를 향상시킬 수 있는 LMCS 기술이 새로이 채택되었다. VVC의 인루프 필터링은 전반적으로 HDR(High Dynamic Range) 영상과 10비트 영상에도 효율적으로 적용될 수 있도록 설계되어, 다양한 영상 콘텐츠에 적합하게 발전되고 있다. 그러나, 이러한 다양한 인루프 필터링으로 VVC의 복잡도는 크게 증가하게 되었으며 이에 따라 부호화 성능을 최대한 유지하며 복잡도를 감소시킬 수 있는 연구가 필요하다[9]. 따라서, VVC의 인루프 필터 기술의 간소화 및 고속화에 대한 표준이 추가 진행될 전망이다.

참고 문헌

- [1] High Efficiency Video Coding, Version 1, Rec. ITU-T H.265, ISO/IEC 23008-2, Jan. 2013.
- [2] B. Bross, J. Chen, and S. Liu, "Versatile Video Coding (Draft 6)," Joint Video Exploration Team of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, JVET-O2001, Jul. 2019.
- [3] [Online]. Available at: https://jvet.hhi.fraunhofer.de/svn/svn_VVCSoftware_VTM/tags/
- [4] J. Chen, Y. Ye, and S. Kim, "Algorithm Description for Versatile Video Coding and Test Model 6 (VTM6)," Joint Video Exploration Team of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, JVET-O2002, Jul. 2019.
- [5] A. Norkin, G. Bjøntegaard, A. Fuldseth, M. Narroschke, M. Ikeda, K. Andersson, M. Zhou, and G. Auwera, "HEVC Deblocking Filter," IEEE Transactions on Circuits and Systems for Video Technology, Vol. 22, No. 12, Dec. 2012.
- [6] C. Fu et al., "Sample Adaptive Offset in the HEVC Standard," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, no.12, pp.1755-1764, Dec. 2012.
- [7] T. Wiegand et al., "WD3: Working Draft 3 of High-Efficiency Video Coding," JCT-VC document, JCTVC-E603, Mar. 2011.
- [8] M. Karczewicz et al., "Improvements on adaptive loop filter," Joint Video Exploration Team of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, JVET-B0060, Feb. 2016.
- [9] W. Chien, J. Boyce, Y. Chen, R. Chernyak, K. Choi, R. Hashimoto, Y. Huang, H. Jang, R. Liao, S. Liu, and D. Luo, "JVET AHG report: Tool reporting procedure (AHG13)," Joint Video Exploration Team of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, JVET-O0013, Jul. 2019.

필자 소개



박도현

- 2016년 2월 : 국립한밭대학교 멀티미디어공학과 학사
- 2018년 2월 : 한국항공대학교 항공전자정보공학과 석사
- 2018년 3월 ~ 현재 : 한국항공대학교 항공전자정보공학과 박사과정
- ORCID : <https://orcid.org/0000-0002-5873-0132>
- 주관심분야 : 비디오 부호화, 영상처리, 360 VR 비디오



윤용욱

- 2017년 2월 : 한국항공대학교 항공전자정보공학과 학사
- 2019년 2월 : 한국항공대학교 항공전자정보공학과 석사
- 2019년 3월 ~ 현재 : 한국항공대학교 항공전자정보공학과 박사과정
- ORCID : <https://orcid.org/0000-0002-5105-5437>
- 주관심분야 : 비디오 부호화, 영상처리, 360 VR 비디오

필자소개



김재곤

- 1990년 2월 : 경북대학교 전자공학과 학사
- 1992년 2월 : KAIST 전기 및 전자공학과 석사
- 2005년 2월 : KAIST 전기 및 전자공학과 박사
- 1992년 3월 ~ 2007년 2월 : 한국전자통신연구원(ETRI) 선임연구원/팀장
- 2001년 9월 ~ 2002년 11월 : Columbia 대학교 연구원
- 2007년 9월 ~ 현재 : 한국항공대학교 항공전자정보공학부 교수
- 2015년 12월 ~ 2016년 1월 : UC San Diego 방문교수
- ORCID : <https://orcid.org/0000-0003-3686-4786>
- 주관심분야 : 비디오 신호처리, 비디오 부호화 표준, Immersive Video