

ECM 인루프 필터링 기술

□ 김민섭, 심동규 / 광운대학교

요약

JVET(Joint Video Experts Team)은 VVC(Versatile Video Coding) 이후 세대의 비디오 압축 표준을 목표로, 신호처리 기반 압축 성능을 한층 더 향상시키기 위한 기술 탐색을 지속하고 있으며, 이를 위해 ECM(Enhanced Compression Model) 참조 소프트웨어를 중심으로 다양한 기술을 검증하고 있다. ECM은 휘도 및 색차 성분 전반에서 VVC 대비 우수한 부호화 성능을 보이며, VVC에 채택된 기존 기술의 개선뿐만 아니라 새로운 도구들을 포함하는 확장 구조로 구성되어 있다. ECM에서는 복호화 영상의 화질 개선 및 참조 영상 품질 향상을 수행하는 인루프 필터링 기술로서, VVC에 채택된 DF(Deblocking Filter), SAO(Sample Adaptive Offset), ALF(Adaptive Loop Filter), CCALF(Cross-Component ALF)를 기반으로 하되, 일부 알고리즘 구조와 적용 방식이 개선·확장되었다. 또한, 색차 성분 간 상관성을 보다 적극적으로 활용하기 위한 CCSAO(Cross-Component SAO)와 에지 보존 특성을 갖는 BIF(Bilateral Filter) 기술이 새롭게 도입되었다. 본 고에서는 2025년 10월에 개최된 제152차 JVET 회의 이후 배포된 ECM 19.0을 기준으로, ECM에 포함된 인루프 필터링 기술의 구성과 주요 특징을 정리한다. 특히, VVC 대비 구조적 변화와 성능 개선이 두드러진 ALF 및 CCALF와 함께, 신규 기술인 CCSAO 및 BIF의 동작 원리와 설계 방향을 중심으로 기술 동향을 분석하고자 한다.

I. 서론

JVET(Joint Video Experts Team)은 ISO/IEC MPEG과 ITU-T VCEG가 공동으로 운영하는 표준화 기구로, 차세대 비디오 압축 표준 기술의 개발과 검증을 담당하고 있다. JVET은 2020년 7월 VVC(Versatile Video Coding,

H.266)[1] 표준화를 완료한 이후, VVC보다 우수한 압축 성능을 가지는 차세대 비디오 부호화 기술에 대한 논의를 본격화하였다. 이러한 기술 탐색 과정에서 제안되는 다양한 도구들의 성능과 구조를 검증하기 위해 채택된 기술들을 통합한 ECM(Enhanced Compression Model) 참조 소프트웨어가 활용되고 있다. 2025년 10월에 개최

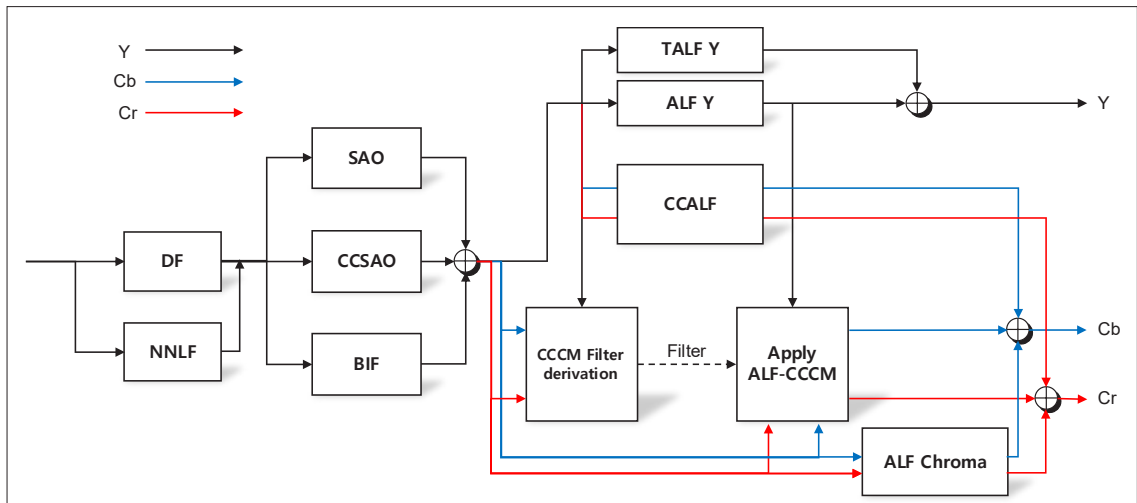
※ 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터사업의 연구결과 (IITP-2026-RS-2022-00258639) 및 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원-학·석사연계 ICT 핵심인재양성 지원을 받아 수행된 연구임. (IITP-2026-RS-2022-00156215)

된 제152차 JVET 회의 이후 배포된 ECM 19.0[2]은 VVC 참조 소프트웨어인 VTM(VVC Test Model) 11.0[3] 대비 AI(All Intra) 환경 및 RA(Random Access) 환경에서 높은 부호화 효율을 보이지만, 블록 기반 하이브리드 구조의 특성상 영상 분할, 예측, 양자화 과정에서 발생하는 블록킹(Blocking), 링잉(Ringing), 텍스처 및 에지 스무딩(Smoothing)과 같은 압축 왜곡 문제는 여전히 주요한 품질 저하 요인으로 남아있다. 이러한 압축 왜곡을 효과적으로 저감하고, 복호화 영상의 품질을 향상시켜 이후 예측 성능을 확보하기 위해 ECM에서는 예측 및 역변환 이후 복원된 영상이 참조 영상으로 사용되기 전에 부·복호화 루프 내(In loop)에서 수행되는 후처리 기법인 인루프 필터링(In-loop filtering) 기술이 적용된다. ECM의 인루프 필터링 구조는 VVC에 채택된 DF(Deblocking Filter), SAO(Sample Adaptive Offset), ALF(Adaptive Loop Filter), CCALF(Cross-Component ALF)를 기본 골격으로 하되, 필터 구조와 적용 방식이 일부 확장·개선된 형태로 구성되어 있다. 아울러, 색차 성분 간 상관성을 보다 적극적으로 활용하는 CCSAO(Cross-Component SAO), ALF 기반의 성분 간 결합을 확장한 ALF-CCCM(ALF-

Convolutional Cross-Component intra prediction Model), 참조 프레임을 이용해 휘도 성분을 추가로 보정하는 TALF(Temporal ALF), 에지 보존 특성을 고려한 BIF(Bilateral Filter), 그리고 신경망 기반 인루프 필터링 기술인 NNLF(Neural Network in-loop filter)[8] 등 신규 기술들이 ECM에 추가되었다. 본 고에서는 2025년 10월 제152차 JVET 회의 이후 배포된 ECM 19.0의 인루프 필터링 기술 동향을 새롭게 도입된 기술들 중심으로 살펴본다.

II. In loop filtering tools in ECM 19.0

본 장에서는 ECM 19.0에 포함된 주요 인루프 필터링 기술들의 구조와 적용 특성을 체계적으로 정리한다. 구체적으로는 기술의 개선·확장이 이루어진 DF, SAO, ALF, CCALF와 새롭게 추가된 BIF, CCSAO, ALF-CCCM, TALF 그리고 NNLF를 살펴본다.

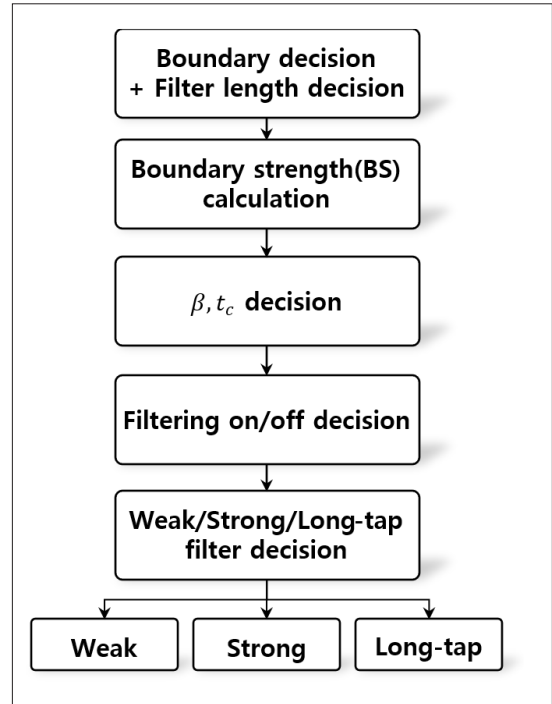


<그림 1> ECM의 인루프 필터링 블록도

1. Deblocking filter

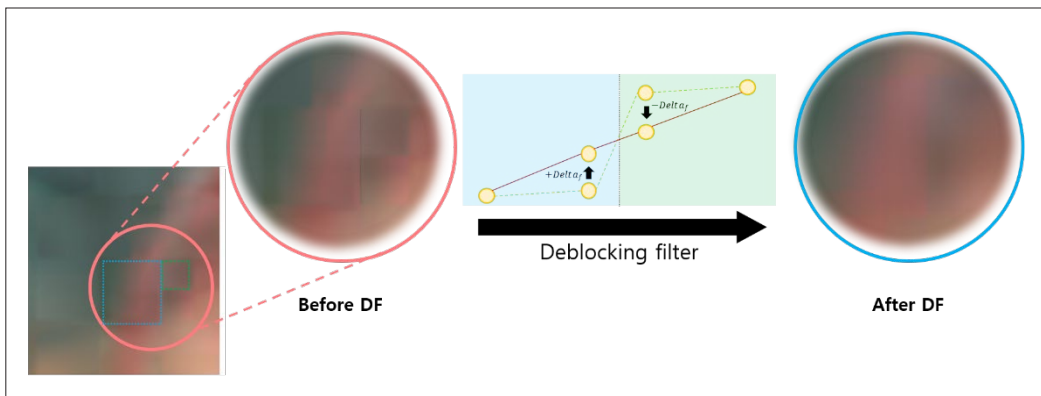
Deblocking filter는 VVC 이전 표준 코덱인 HEVC부터 채택된 기술로서, ECM에서도 수행되는 인루프 필터링 기술이다. 예측신호와 잔차신호의 합연산을 통해 생성된 복원 신호를 입력받아 수행되며, 블록 기반 하이브리드 압축 방식을 수행하는 과정에서 인접한 블록 간의 QP(Quantization Parameter) 차이에 의해 일어나는 블로킹 왜곡의 완화를 목적으로 한다. 부호화기로부터 별도의 시그널링을 받지 않고 수행되며, 인루프 필터링 기술들 중 가장 먼저 수행된다.

<그림 3>은 ECM의 deblocking Filter의 블록도를 나탄다. Deblocking Filter는 필터링 대상이 되는 블록 경계를 따라 경계 강도를 계산하고 필터링 수행 여부를 결정하여, 최종적으로 필터링 강도를 결정하여 필터링을 수행한다. 각 단계는 다음과 같은 과정으로 수행된다. 우선, 인접한 두 블록 사이의 경계가 변환 블록 경계인지 또는 서브 블록 경계인지 확인 후, 경계를 기준으로 필터링에 사용되는 샘플의 수를 {1, 2, 3, 5, 7} 중 하나로 결정한다. 이때, 블록 경계를 기준으로 양쪽 블록에서 사용되는 샘플의 수는 다를 수 있다. 다음으로 인접한 블록 간의 경계에 대해서 블로킹 왜곡 발생의 가능성을 나타내는 경계 강도(Boundary Strength)를 계산한다. 경계 강도는 인접한 두 블록의 예측 모드(Intra/Inter), 움직임 정보의 유



<그림 3> Deblocking filter의 블록도

사성, 참조 픽처의 POC(Picture Order Count) 등을 종합적으로 고려하여 {0, 1, 2} 중 하나로 분류된다. 휘도 성분인 경우 BS가 1 또는 2로 분류되면 Deblocking Filter가 적용될 수 있으며, 색차 성분에서는 BS가 2인 경우에만 필터링이 수행된다[4]. 특히, 두 블록 중 적어도 하나가 GPM(Geometric Partitioning Mode)을 이용하여 예측되



<그림 2> Deblocking filter의 개념도

었고, 서브 영역 간에 화면내 예측과 화면간 예측이 혼합되어 사용된 경우에는 강한 블록 경계 왜곡이 발생할 가능성이 높다고 판단하여 BS가 2로 설정된다. 이후, 해당 경계에서의 불연속성이 실제 영상의 에지인지 또는 블로킹 열화인지를 판단하기 위해서 BS와 각 블록의 QP에 따라 결정되는 필터 임계값 관련 파라미터 β 와 필터링 보정량 제한 관련 파라미터인 t_c 를 산출하고 필터링 여부를 결정한다. 마지막으로, 산출된 β 및 t_c 를 이용하여 해당 경계에 대해 적용할 필터의 유형을 약한 필터(weak filter), 강한 필터(strong filter), 또는 롱 탭(long-tap) 필터 중에서 선택한다.

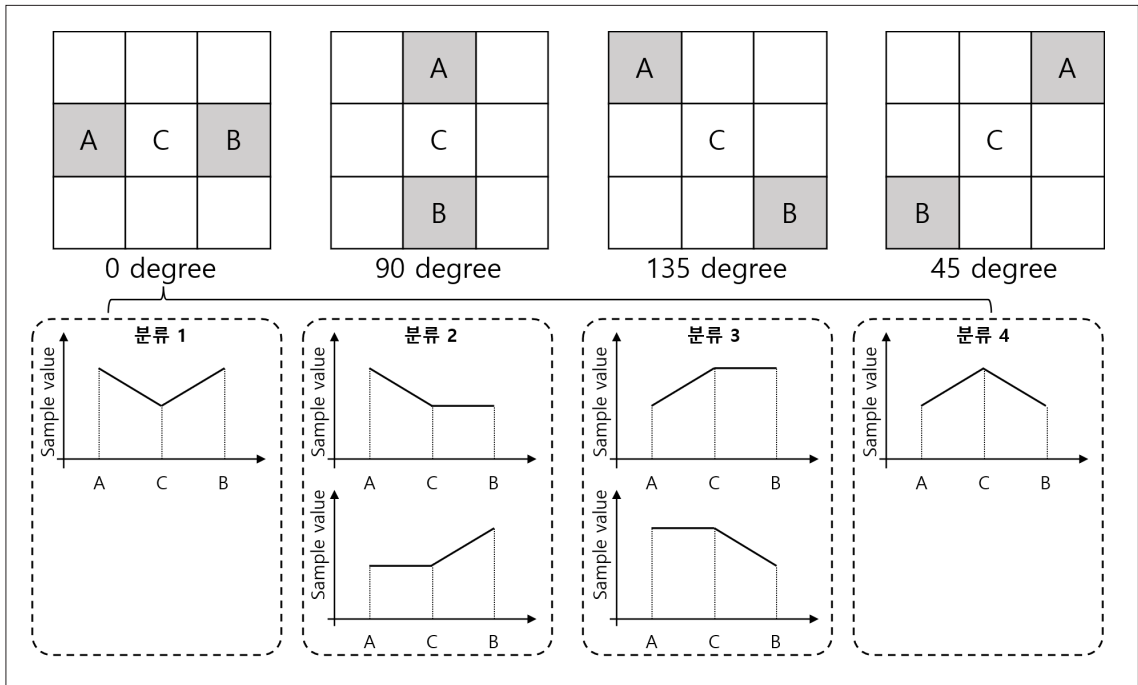
2. Sample adaptive offset

SAO(Sample Adaptive Offset)는 HEVC에서 도입되어 VVC 및 ECM에서도 적용되는 인루프 필터링 기술로서, Deblocking Filter 이후 단계에서 수행된다. SAO는

Deblocking Filter가 블록 경계에서 발생하는 블로킹 왜곡을 완화하는 것과 달리, 복원된 샘플 값이 특정 조건에서 반복적으로 가지는 평균 오차를 보정하는 것을 목적으로 한다. SAO는 복원된 샘플을 사전에 정의된 조건에 따라 분류한 후, 분류에 따라서 시그널링된 상수 오프셋을 결정한다. SAO는 부호화기에서 CTU 단위로 적용 여부 및 파라미터가 결정되어 시그널링되며, 인접한 상단 또는 좌측 CTU의 SAO 파라미터를 재사용하는 merge 방식 또한 지원한다. SAO는 크게 EO(Edge Offset)와 BO(Band Offset)의 두 가지 모드를 지원하며, 각 모드는 서로 다른 기준에 따라 샘플을 분류하고 오프셋(ΔI_{SAO})을 결정한다. 결정된 오프셋은 3장과 4장에서 설명할 새로 추가된 인루프 필터 기술인 CCSAO와 BIF로부터 유도된 오프셋들을 합산하여 5장에서 설명할 ALF의 입력으로 사용된다.

1) Edge offset

EO(Edge Offset) 모드는 CTU 단위로 시그널링되는 방



<그림 4> SAO의 Edge offset의 분류

향 정보에 따라 수평(0°), 수직(90°), 또는 두 대각 방향(45°, 135°) 중 하나의 방향을 선택하여 수행된다. 선택된 방향에 대해 현재 샘플을 기준으로 해당 방향상에 위치한 두 개의 인접 샘플을 정의하고, 이들 인접 샘플과 현재 샘플 간의 밝기값 크기 관계를 이용하여 현재 샘플을 분류한다. 복호화기는 부호화기로부터 전송된 분류별 오프셋의 절대값 4개와 각 절대값에 대응되는 부호 4개를 이용하여, 현재 샘플이 속한 분류에 해당하는 오프셋을 결정한다. 이를 통해 에지 주변에서 특정 방향 조건하에 반복적으로 나타나는 평균 복원 오차를 효과적으로 보정하기 위한 오프셋을 유도할 수 있다.

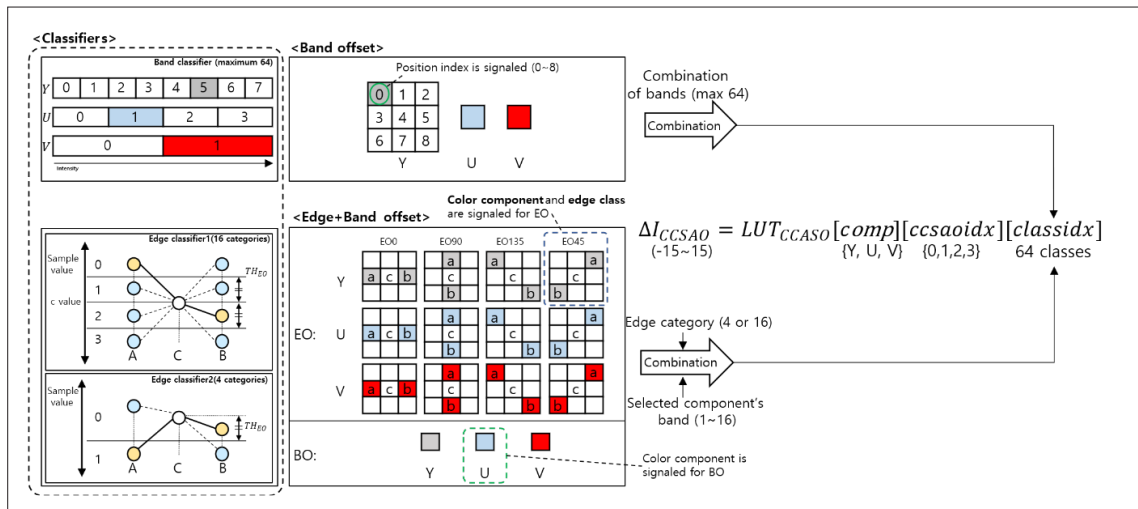
2) Band offset

BO(Band Offset) 모드는 CTU 단위로 시그널링되는 밝기 대역 정보를 기반으로 수행된다. BO에서는 샘플의 밝기값 범위를 비트 심도에 따라 균일한 개수의 밴드로 분할한 후, 이 중 연속된 네 개의 밴드에 속하는 샘플에 대해서 오프셋을 적용한다. 복호화기는 부호화기로부터 전송된 시작 밴드 위치 정보와 선택된 네 개 밴드에 대응되는 오프셋의 절대값 4개 및 각 절대값에 대응되는 부호 4개를 이용하여, 현재 샘플의 밝기값이 속한 밴드에 해당하는

오프셋을 결정한다.

3. Cross-component SAO

CCSAO(Cross-Component Sample Adaptive Offset)는 ECM에 새롭게 채택된 인루프 필터링 기술로, 기존 SAO를 확장하여 색차 성분 간 상관관계를 보다 효과적으로 활용하는 것을 목적으로 한다. CCSAO[12]는 Deblocking Filter된 신호를 입력으로 수행되며, SAO와 마찬가지로 복원된 샘플의 통계적 편향을 보정하기 위한 오프셋을 결정한다. 그러나 CCSAO는 단일 성분 기반 분류에 한정되었던 SAO와 달리, 휘도(Y) 및 색차(Cb, Cr) 성분을 교차적으로 활용하여 샘플을 분류함으로써 보다 정교한 왜곡 보정을 가능하게 한다. CCSAO 또한 분류 방식에 따라 BO(Band Offset)와 EBO(Edge+Band Offset)의 두 가지 모드를 지원하며, 각 모드별 기준에 따라 샘플을 분류하고 오프셋을 결정한다. 전체적인 CCSAO의 오프셋 결정 구조는 <그림 5>와 같다. -15~15의 값의 범위를 가지는 오프셋 후보들은 프레임단위로 부호화기로부터 시그널링되며, 각 성분별로 최대 4개의 EO, BO 공용 오프셋 세트를 가질 수 있다. 각 성분별 CTU에서는 부호



<그림 5> CCSAO의 오프셋 결정 구조

화기로부터 해당 선분의 오프셋 세트 중 어떤 세트를 사용할지 시그널링을 통해 결정된다.

1) Band offset

BO(Band Offset) 모드는 샘플 값의 분포 특성을 기반으로 CCSAO를 수행하는 방식으로, 각 성분의 밝기값을 일정 개수의 밴드로 분할한 후, 밴드 조합에 따라 샘플을 분류하고 이에 대응되는 오프셋을 적용한다. CCSAO의 BO 모드에서는 단일 성분에 대해서만 밴드 분류를 수행하는 기존 SAO BO와 달리, 휘도(Y) 및 색차(Cb, Cr) 성분의 밴드 정보를 동시에 고려함으로써 교차 성분 기반의 보다 정교한 분류를 수행한다. 복호화기는 각 성분별로 몇 개의 밴드로 분할할지 프레임 단위로 부호화기로부터 시그널링받으며, 성분별 밴드 인덱스의 조합에 따라 오프셋 값이 결정된다. 이때 휘도 및 색차 성분에 대해 사용 가능한 밴드 개수는 각각 최대 16, 4, 4로 제한되며, 세 성분의 밴드 개수 조합에 의해 정의되는 전체 클래스 수가 최대 64를 초과하지 않도록 부호화기에서 조합이 결정되어 시그널링된다. 추가적으로, 휘도 성분의 밴드 인덱스를 계산할 때 사용할 샘플의 공간적 위치를 프레임단위로 부호화기로부터 전송받을 수 있다. 공간적 위치 후보는 현재 샘플을 중심으로 한 3×3 이웃 영역으로, 총 9가지 위치 중 한 위치가 시그널링을 통해 결정된다. 반면, 색차 성분(Cb, Cr)의 경우에는 현재 위치의 샘플 값을 이용하여 각각의 밴드 인덱스를 계산한다. 이와 같이 휘도 성분에 대해서만 공간적으로 단일 이웃 샘플을 활용하는 구조를 통해, CCSAO BO는 휘도 성분의 공간적 구조 정보를 간접적으로 반영할 수 있다.

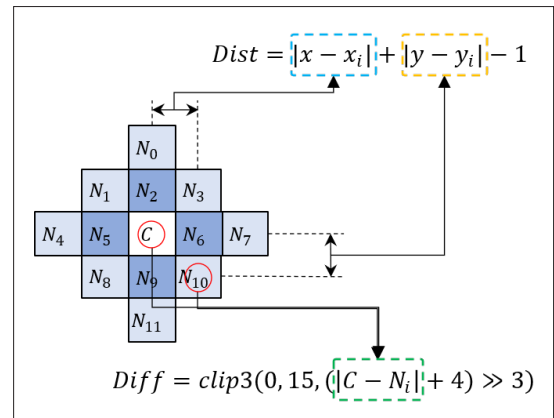
2) Edge + Band offset

CCSAO에서의 EBO(Edge Band offset) 모드는 SAO에서의 EO 구조를 기반으로 하되, 에지 분류 과정에서 현재 성분의 보정을 위한 오프셋을 결정할 때 현재 성분이 아닌 성분의 에지 정보와 밴드 정보를 함께 활용하도록 확장된 기술이다. EBO 모드는 에지 분류에 사용할 성분과 밴드 분류에 사용할 성분을 각각 독립적으로 선택할 수 있

으며, 이는 부호화기로부터의 시그널링으로 결정된다. 이에 따라, 현재 성분의 오프셋을 유도하는 과정에서 다른 성분의 에지 특성 또는 값 분포 특성을 기준으로 분류가 수행될 수 있다. 밴드 정보를 활용하는 과정에서는 현재 성분이 아닌 시그널링을 통해 선택된 성분의 밴드 인덱스를 기준으로 샘플이 분류가 될 수 있다. 에지 정보를 활용하는 과정에서는 기존 고정된 에지 분류 구조를 사용하는 SAO와 달리, CCSAO는 에지 분류 구조를 부호화기로부터의 시그널링을 통해 결정하며, 이를 통해 에지 분류 해상도를 유연하게 조절할 수 있다[2][5]. 최종적으로 EBO 모드에서는 에지 정보와 밴드 정보를 결합하여 샘플을 분류하며, 전체 분류 클래스 수는 시그널링 오버헤드 및 구현 복잡도를 고려하여 최대 64로 제한된다.

4. Bilateral filter

BIF(Bilateral Filter)는 변환 및 양자화 과정에서 발생하는 에지 주변의 링잉(ringing) 열화를 저감하기 위해 적용되는 인루프 필터링 기술로, ECM에 새롭게 채택되었다. BIF는 SAO 및 CCSAO와 동일하게 Deblocking Filter 이후의 복원 샘플을 입력으로 수행되며, 에지 보존 특성을 갖는 비선형 필터링을 통해 왜곡을 보정한다. 특히, BIF는 공간적 거리와 샘플 간 밝기 차이를 동시에 고려하는



<그림 6> BIF의 공간적 거리 및 밝기 차이

$$\Delta I_{BIF} = \left\{ \left\{ \sum_{i=0}^{11} (c_i \cdot f_i) \right\} + (1 \ll (5 - strength_{pps})) \right\} \gg (6 - strength_{pps}) \quad (1)$$

bilateral filtering 구조를 기반으로 하여, 에지 구조를 유지하면서도 링잉 성분을 효과적으로 억제하는 것을 목적으로 한다. BIF는 휘도 및 색차 성분에 대해 각각 독립적으로 적용되며, CTU 단위로 사용 여부가 결정된다. BIF의 적용 가능 여부는 예측 모드, 양자화 파라미터(QP), 변환 블록(TB)의 크기 및 변환 계수의 존재 여부(cbf) 등의 조건에 따라 제한된다. 예를 들어, 화면간 예측이 수행된 블록의 경우 QP가 일정 임계값 이상이고 TB 크기가 제한 조건을 만족할 때에만 BIF가 적용되며, 화면내 예측 블록에 대해서도 별도의 조건이 적용된다. BIF에서 결정되는 오프셋은 <그림 6>과 같이 현재 샘플과 그 주변에 위치한 복수의 인접 샘플 간의 관계를 기반으로 식 (1)과 같이 계산된다.

이때, 주변 샘플과의 공간적 거리(Dist)는 맨해튼 거리(Manhattan distance)로 정의되며, 공간적 거리와 Scale 파라미터를 통해 각 인접 샘플에 대한 가중치(c_i)가 식 (2)에 따라 계산된다. 이때 Scale은 현재 블록의 예측모드와 변환 블록의 크기에 따라 <표 1>과 같이 결정된다. 임계값 T_0 및 T_1 은 블록 분할 구조가 듀얼 트리(Dual tree)인 경우, 휘도 및 색차 성분 모두에 대해 $T_0=4$, $T_1=128$ 이 적용된다. 반면, 싱글 트리(Single tree) 구조에서는 휘도 성분에 대해 $T_0=4$, $T_1=128$ 이 적용되며, 색차 성분에 대해서는 $T_0=16$, $T_1=256$ 이 적용된다.

$$c_i = (1/Dist) \cdot Scale \quad (2)$$

<표 1> 현재 블록의 예측모드와 변환 블록의 크기에 따른 Scale 파라미터

	$TB_{size} \leq T_0$	$T_0 \leq TB_{size} \leq T_1$	$T_1 \leq TB_{size}$
Intra	3	2	1
Inter	2	2	1

또한, 현재 샘플과 주변 샘플 간의 밝기 차이(Diff)는 미

리 정의된 LUT(Look-Up Table)를 통해 비선형적으로 매핑되며, 이를 통해 각 인접 샘플에 의한 밝기 기반 오프셋(f_i)은 식 (3)에 따라 계산된다. 추가적으로, 양자화 왜곡 특성을 반영하기 위해 양자화 파라미터를 고려하여 LUT가 정의되어 있으며, 이때 사용되는 QP_{BIF} 는 현재 TU의 QP와 프레임 단위로 결정되는 오프셋 QP와의 합산을 클리핑하여 17-42 범위의 값을 가진다.

$$f_i = \text{sign}(C - N_i) \cdot \text{LUT}_{Diff}[\text{CompID}][QP_{BIF} - 17][Diff] \quad (3)$$

식 (1)에 포함된 $strength_{pps}$ 는 프레임 단위로 부호 화기로부터 시그널링되어 BIF의 전체적인 필터링 강도를 제어하는 파라미터로, 0-2 범위의 값을 가질 수 있다.

최종적으로, Deblocking Filter가 수행된 복원 샘플 I 에 대해 SAO, CCSAO, BIF를 각각 적용하여 세 종류의 오프셋 ΔI_{SAO} , ΔI_{CCSA} , ΔI_{BIF} 를 계산한 후, 이들을 모두 합산하여 현재 비트 심도에 따른 클리핑을 수행한 뒤, ALF의 입력으로 사용된다.

5. Adaptive loop filter

ALF(Adaptive Loop Filter)는 Deblocking Filter, SAO, CCSAO 및 BIF 이후 단계에서 수행되는 인루프 필터링 기술로, 영상의 지역적 특성에 따라 서로 다른 필터 계수를 적용한다는 점이 특징이다. LLS(Linear Least Square) 근사에 기반한 Wiener 필터링 방식을 사용한다. 해당 구조는 원본 영상과 복원 영상 간의 통계적 관계를 모델링하여, 부·복호화 과정에서 발생한 왜곡을 최소화하는 방향으로 필터 계수를 산출한다. 기존 VVC의 ALF의 경우 DF 및 SAO 수행 이후의 복원 샘플을 입력으로 수행되었으며, ECM에서는 ALF 구조가 전반적으로 확장되어,

필터에 입력되는 신호와 필터의 다양화를 통해 보다 정교한 왜곡 보정을 수행하는 기술이 추가되었다. 현재 ECM의 ALF에서 사용되는 입력 신호는 SAO, CCSAO, BIF 이후의 출력 신호, DF 이전 복원 신호, 잔차 신호 등이 사용되고 있으며, 지역적 특성을 분류하기 위한 분류기를 부호화기로부터 시그널링 받아 결정하거나 상기 입력 신호들에 대해서 다양한 크기와 모양을 가진 필터를 적용한다. 기존 VVC가 지역적 특성을 분류하기 위해 방향성과 활동성을 계산하여 4×4 크기의 블록별로 25개의 클래스로 분류하던 부분은 ECM에서 2×2 크기의 블록별로 분류를 수행하도록 변경되었다. 또한, 필터 크기가 휘도 성분과 색차 성분 각각 최대 13×13 , 9×9 까지 확장되었으며, 고정 필터(fixed filter)와 부호화 필터(signalled filter)를 병행 사용하는 구조를 채택함으로써, 성능이 크게 향상되었다. 이와 같은 구조적 확장을 통해 ECM의 ALF는 VVC 대비 향상된 왜곡 저감 성능을 보이지만, 동시에 필터 연산량 증가와 메모리 접근 증가를 동반하였다.

6. Cross-component ALF

CCALF(Cross-Component Adaptive Loop Filter)는 ALF가 수행된 색차 성분에 대해서, 추가적인 보정을 위한 신호를 휘도 성분과 색차 성분 간의 상관관계를 활용하여 생성하는 인루프 필터링 기술이다. VVC에서의 CCALF는 색차 성분의 복원 샘플과 이에 대응되는 휘도 샘플 주변의 복수 인접 샘플들을 입력으로 사용하며, 선형 필터링을 통해 ALF 수행 이후의 색차 성분 샘플을 추가로 보정한다. 이때 사용되는 휘도 성분은 SAO로부터 출력된 휘도 성분이다. ECM에서의 CCALF는 VVC에 채택된 CCALF와 달리, SAO로부터 출력된 휘도 성분뿐만 아니라 휘도 성분의 잔차 정보까지 함께 활용하여 색차 성분을 보정하는 구조로 확장되었다. 추가적으로, DF 이전의 색차 복원 샘플을 입력으로 사용하여, 한 색차 성분을 이용해 다른 색차 성분을 보정할 수 있도록 색차 간의 교차 성분 탭도 추가되었다. 이는 VVC의 CCALF에서는 고려되지 않았던 색차

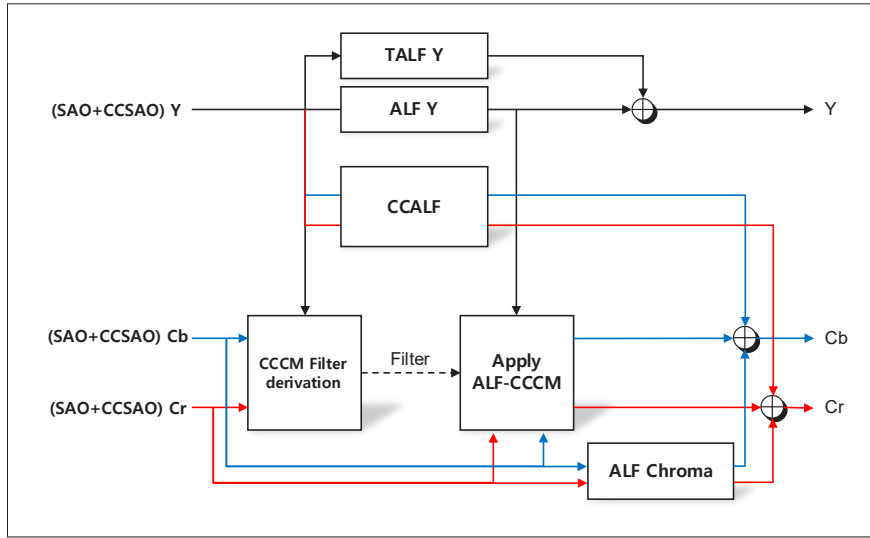
성분 간 상호 상관관계를 활용하는 확장된 구조이다. 마지막으로, 색차 성분의 SAO 출력 샘플에 3×3 비대칭 십자 형태의 4개 탭을 통해 필터링한 이후 CCALF의 입력으로 사용되도록 변경되었다.

ECM의 CCALF에서는 슬라이스 단위로 통계 정보를 수집하고 이를 기반으로, 최대 4개였던 VVC가 최대 16개의 CCALF 필터를 APS를 통해 신호화할 수 있도록 변경되었다. 또한, CCALF 계수의 소수부 비트 수는 성분별로 7비트에서 10비트 범위 내에서 적응적으로 선택될 수 있으며, 해당 정보는 APS 내에서 색차 성분별로 2비트 구문 요소를 통해 전달된다. 이와 함께, VVC에서 적용되던 계수 값의 2의 거듭제곱(power-of-2) 제약 조건이 제거되어, 필터 계수 표현의 자유도가 증가하였다.

7. ALF-CCCM

ALF-CCCM(Cross-Component Convolutional Model)은 인터 슬라이스의 경우, 복호기 측에서 유도된(convolutional) 교차 성분 필터를 이용하여, ALF가 적용된 색차 성분 출력에 대해 추가적인 보정 신호를 생성하는 인루프 필터링 기술이다[6]. ALF-CCCM에서는 휘도 성분과 색차 성분 간의 상관관계를 활용하여 교차 성분 필터 계수를 유도하며, 해당 필터링은 SAO 및 CCSAO 이후의 휘도 출력 샘플에 대해 수행된다. 이 과정은 <그림 7>과 같다.

ALF-CCCM에서는 각 블록에 대해 SAO 및 CCSAO 이후의 휘도 및 색차 출력 신호를 이용하여 교차 성분 필터 계수를 유도한다. 유도된 필터는 ALF가 수행된 휘도 출력 신호와 SAO 및 CCSAO 이후의 색차 신호를 활용하여 ALF-CCCM의 출력을 생성하는 과정에서 활용되며, 출력된 보정 신호는 가중치 0.5가 적용되어 ALF 및 CCALF와 합산된다. 색차 출력에 가산되어 색차 복원 샘플의 품질을 향상시킨다. 필터의 모델은 7-tap CCCM 모델과 동일한 구조를 사용한다. 필터 계수 유도에는 CCCM solver가 사용되며, 휘도 샘플을 색차 그리드로 매핑하기 위해



<그림 7> ALF-CCCM의 블록도

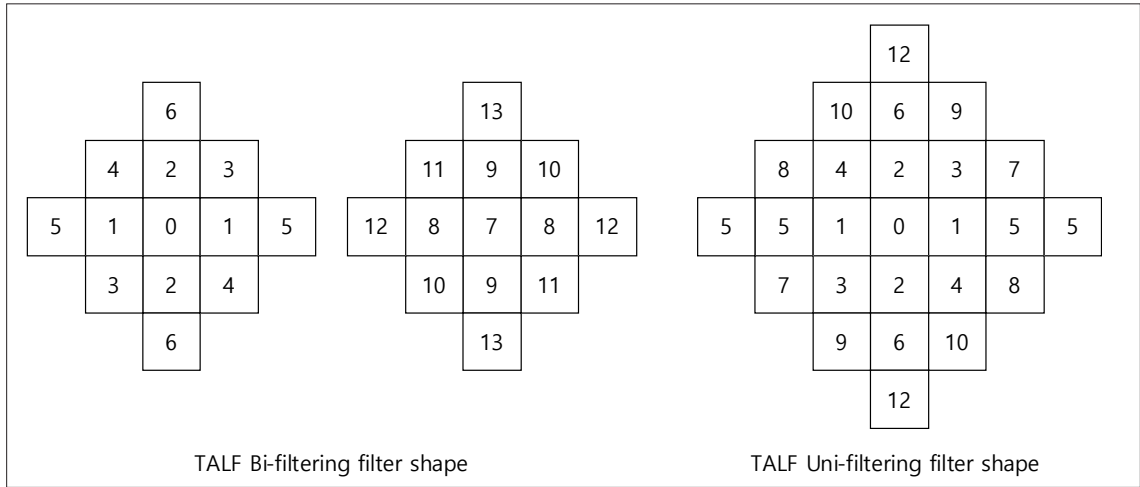
CCCM 및 CCLM에서 사용되는 것과 동일한 6-tap 다운샘플링 필터가 적용된다. CTU는 인접 CTU로부터 블록 크기와 모델 유형을 상속받을 수 있다. 좌측 또는 상단 CTU에 ALF-CCCM이 적용된 경우, 현재 CTU가 해당 정보를 상속할지 여부가 시그널링된다. 좌측과 상단 CTU가 모두 후보인 경우에는 추가적인 CABAC 플래그를 통해 상속 대상을 결정한다. 또한, 참조 프레임으로부터 모든 CTU의 블록 크기 및 모델 유형을 상속받을 수 있다. 참조 프레임은 L0 및 L1 리스트에 포함된 프레임 중, 최소 하나 이상의 CTU에서 ALF-CCCM이 적용된 프레임만을 대상으로 하며, 현재 프레임과의 POC 거리가 가장 작은 프레임이 선택된다. 프레임 단위 상속이 활성화되면, 해당 프레임에서는 CTU 단위의 시그널링은 수행하지 않는다. 추가적으로, ALF-CCCM에서는 보정 신호에 적용되는 가중치를 적응적으로 선택할 수 있다. 가중치 후보 집합은 {1/4, 3/8, 5/8, 3/4}이며, 고정 가중치 0.5를 포함하여 각 프레임에 대해 모든 후보에 대한 필터링 결과를 평가한다. 각 가중치에 대해 왜곡과 비트 비용을 계산한 후, 최소 RD 비용을 갖는 가중치가 해당 프레임의 최적 가중치로 선택된다. 적응적 가중치 사용 여부를 나타내는 플래그가 시그널

링되며, 적응 모드가 선택된 경우 가중치 인덱스를 나타내는 2비트가 추가로 시그널링된다.

8. Temporal adaptive loop filter

TALF(Temporal Adaptive Loop Filter)는 참조 프레임을 활용하여, 공간 기반 인루프 필터링 이후에도 잔존하는 왜곡을 추가적으로 보정하기 위한 인루프 필터링 기술이다[7]. TALF는 현재 프레임의 휘도 성분에 대해 ALF가 적용된 이후의 출력 값을 기준으로, 참조 프레임과의 시간적 상관관계를 이용해 추가적인 보정 오프셋을 생성한다. TALF의 적용 여부는 CTB 단위로 시그널링된다. TALF는 참조 샘플을 선택하는 방식에 따라 총 여섯 가지 필터링 모드를 지원한다.

첫 번째 및 두 번째 모드는 각각 단일 참조 화상에 대해 정수화된 MV0 또는 MV1을 사용하여 참조 위치를 결정하는 단방향 필터링 모드로, 해당 CTB 내에서 유효한 MV가 존재하지 않는 위치에 대해서는 TALF 처리가 수행되지 않는다. 세 번째 모드는 양방향 필터링 모드로, MV0과 MV1을 각각 이용해 참조 화상 리스트 0과 1에서 참조 샘플



<그림 8> TALF의 필터 종류

플을 획득한 후, 두 참조 샘플을 평균하여 필터 입력으로 사용한다. 이 경우 두 MV가 모두 존재하지 않는 위치는 필터링 대상에서 제외된다.

네 번째 및 다섯 번째 모드는 움직임 벡터를 사용하지 않고, 현재 화상과 동일한 위치에 대응되는 참조 화상의 복원 샘플을 직접 사용하는 단방향 필터링 모드이다. 이때 네 번째 모드는 시간적으로 가장 가까운 참조 화상을, 다섯 번째 모드는 두 번째로 가까운 참조 화상을 각각 사용한다. 여섯 번째 모드는 이 두 참조 화상의 동일 위치 샘플을 평균하여 사용하는 양방향 필터링 모드로, MV 기반 양방향 필터링과 동일한 개념의 시간적 보정을 수행한다. TALF는 두 가지 필터 형태를 가지며, 각 필터 형태는 13개의 필터 계수로 구성된다. 필터 형태는 공간적 탭 배치만 상이하며, 계수의 개수 및 연산 구조는 동일하다.

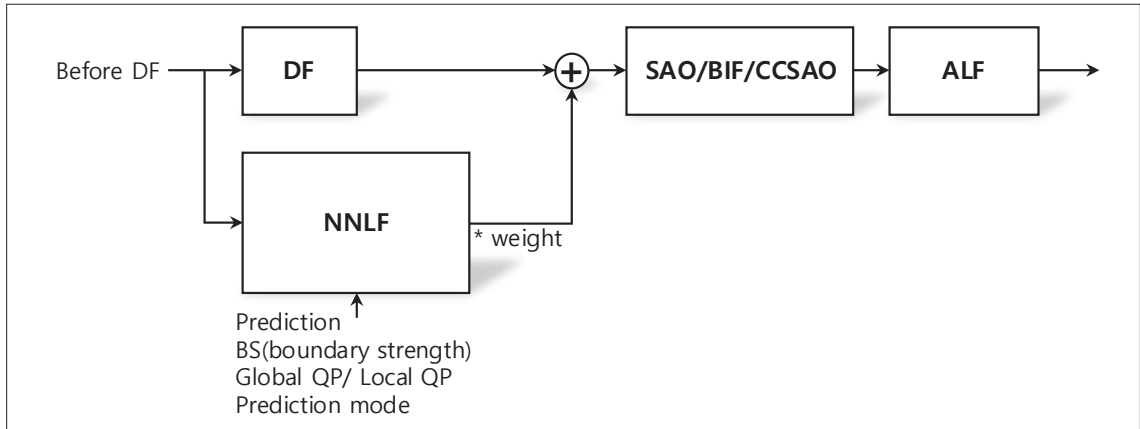
TALF 필터는 참조 화상으로부터 획득한 샘플과 현재 화상의 SAO 출력 휘도 샘플 간의 차이를 기반으로 동작하며, 해당 차이는 사전에 정의된 클리핑 범위 내로 제한된 후 필터 계수와 가중합을 통해 오프셋을 생성한다.

단방향 TALF의 경우, 참조 화상의 필터 입력 샘플과 현재 화상의 SAO 출력 휘도 샘플 간의 차이를 클리핑한 값을 공간적 필터 계수로 누적하여 오프셋을 계산한다. 반

면 양방향 TALF에서는 두 참조 화상으로부터 획득한 차이 값을 각각 클리핑한 후 평균을 취하여 필터 입력으로 사용함으로써, 시간적으로 양측에 위치한 참조 화상의 정보를 균형 있게 반영하도록 설계되었다. 최종적으로 계산된 TALF 오프셋은 ALF가 적용된 휘도 성분에 가산되어 복원 샘플을 형성한다.

9. NNLF

NNLF(Neural Network based Loop Filtering)는 NNVC(Neural Network Video Coding)[8]에서 논의되어 온 신경망 기반 복원 기술을 기존 블록 기반 하이브리드 코덱 구조에 단계적으로 통합하기 위한 시도의 일환으로, ECM의 인루프 필터 체계에 도입된 기술이다[9]. 기존의 규칙 기반 또는 선형 필터 기반 인루프 필터로는 충분히 제거하기 어려운 비선형 왜곡을 추가적으로 보정하기 위한 기술로써 DF 이후, SAO 전에 수행된다. ECM에서의 NNLF는 NNVC와 달리 전체 부호화 구조를 신경망 기반으로 대체하지 않고, 기존 표준 코덱 프레임워크를 유지한 상태에서 신경망 기반 필터를 선택적으로 결합할 수 있도록 설계되었다. 이를 통해 신경망 기반 기술의 성능



<그림 9> NNLF의 블록도

이점을 활용하면서도, 표준화 관점에서 요구되는 복호화 복잡도 제어를 동시에 만족시키는 방향으로 기술 기고가 되고 있다.

ECM에서 정의된 NNLF는 NNVC에서 제안된 NNLF 구조를 계승한 형태로, 인루프 필터를 위한 통합 필터 아키텍처(Unified Filter Architecture)의 일환으로 구성된다. 이 구조에서는 신경망 기반 필터가 독립적인 인루프 필터 블록으로 동작하며, 입력으로는 Deblocking Filter 이전의 복원 프레임, 예측 프레임, Deblocking Filter에서 계산된 boundary strength 정보, 전역 또는 국부 QP 정보, 그리고 예측 모드 및 부호화 모드 관련 정보가 사용된다. 이러한 다양한 입력 정보를 기반으로, 신경망은 현재 프레임의 복원 샘플에 대한 보정 출력을 생성한다.

III. 결론

본 고에서는 VVC 이후 차세대 비디오 압축 표준 기술을 목표로 JVET에서 탐색 중인 ECM 19.0을 기준으로, 인루프 필터링 기술의 구성과 확장 방향을 분석하였다. ECM의 인루프 필터링은 VVC에서 사용되던 기본 구성을 유지하면서도, 보다 정교한 왜곡 보정을 위해 구조적 확장

신규 기술 도입이 이루어진 것이 특징이다.

특히, SAO, CCSAO, BIF는 연산 복잡도를 고려하여 병렬적으로 수행되는 구조를 채택함으로써 다양한 형태의 양자화 왜곡을 효율적으로 보정하도록 설계되었다. 또한, ALF 및 CCALF는 VVC 대비 더 많은 입력 샘플과 확장된 필터 구성을 활용하여 원본 신호와의 차이를 효과적으로 줄이는 방향으로 발전하였다. 나아가, 비선형 보정 능력을 강화하기 위한 NNLF와 휘도-색차 성분 간 상관 관계를 보다 정밀하게 모델링하는 ALF-CCCM, 참조 프레임을 이용해 휘도 성분을 추가로 보정하는 TALF 기술이 추가됨으로써, ECM의 인루프 필터링은 기존 신호처리 기반 필터링을 넘어 보다 유연한 표현력을 갖는 구조로 확장되었다.

이와 같은 인루프 필터링 기술의 확장은 ECM 19.0에서 VVC 대비 우수한 부호화 효율을 달성하는 데 기여하고 있으나, 동시에 필터 구조의 복잡화와 이에 따른 부호화 연산량 증가를 동반한다. 향후 VVC 이후 차세대 비디오 압축 표준의 제정 과정에서는 이러한 인루프 필터링 기술들의 성능 기여도를 면밀히 분석하는 한편, 실용적인 복잡도 수준을 고려한 선택적 적용 및 구조적 단순화 방안에 대한 연구가 병행될 필요가 있을 것으로 판단된다.

참 고 문 헌

- [1] B. Bross, J. Chen, S. Liu, and Y.-K. Wang, "Versatile Video Coding (H.266) Specification," ITU-T Rec. H.266 | ISO/IEC 23090-3, Aug. 2020.
- [2] ECM 19.0, https://vcgit.hhi.fraunhofer.de/ecm/ECM/-/tree/ECM-19.0?ref_type=tags
- [3] VTM 11.0, https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware_VTM/-/tree/VTM-11.0?ref_type=tags
- [4] 심동규, 조현호, "HEVC 표준 기술의 이해," 홍릉 과학출판사, 2014.
- [5] M. Coban, R.-L. Liao, K. Naser, J. Ström, and L. Zhang, "Algorithm description of Enhanced Compression Model 19 (ECM 19)," 34th JVT Meeting, Teleconference, Doc. JVET-AN2025, Jan. 2024.
- [6] P. Astola, I. Jumakulyyev, D. Bugdayci Sansli, and J. Lainema, "AHG12: ALF-CCCM," 34th JVT Meeting, Teleconference, Doc. JVET-AJ0173, Jan. 2024.
- [7] L. Xu, Y. Yu, H. Yu, and D. Wang, "AHG12: Temporal ALF," 34th JVT Meeting, Teleconference, Doc. JVET-AJ0090, Jan. 2024.
- [8] F. Galpin, Y. Li, Y. Li, D. Rusanovskyy, T. Shao, J. Ström, and L. Wang, "Description of algorithms version 13 and software version 15 in neural network-based video coding (NNVC)," 34th JVT Meeting, Teleconference, Doc. JVET-AN2019, Jan. 2024.
- [9] T. Poirier, F. Galpin, G. Boisson, Y. Li, M. Karczewicz, J. Wang, L. Kerofsky, H. Wang, N. Hu, R. Yu, M. Coban, V. Seregin, D. Rusanovskyy, K. Panusopone, S. Hong, and L. Wang, "EE2-related: NNLF interface in ECM," 34th JVT Meeting, Teleconference, Doc. JVET-AL0228, Jan. 2024.

저 자 소 개



김민섭

- 2020년 2월 : 광운대학교 컴퓨터공학과 공학사
- 2022년 2월 : 광운대학교 컴퓨터공학과 공학석사
- 2023년 3월 ~ 현재 : 광운대학교 컴퓨터공학과 박사과정
- 주관심분야 : 영상신호처리, 영상압축, 컴퓨터비전



심동규

- 1993년 2월 : 서강대학교 전자공학과 공학사
- 1995년 2월 : 서강대학교 전자공학과 공학석사
- 1999년 2월 : 서강대학교 전자공학과 공학박사
- 1999년 3월 ~ 2000년 8월 : 현대전자 선임연구원
- 2000년 9월 ~ 2002년 3월 : 바로비전 선임연구원
- 2002년 4월 ~ 2005년 2월 : University of Washington Senior research engineer
- 2005년 3월 ~ 현재 : 광운대학교 컴퓨터공학과 교수
- 주관심분야 : 영상신호처리, 영상압축, 컴퓨터비전