

일반논문 (Regular Paper)

방송공학회논문지 제30권 제5호, 2025년 9월 (JBE Vol.30, No.5, September 2025)

<https://doi.org/10.5909/JBE.2025.30.5.815>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

적응적 방향성 히스토그램 구성 기반의 화면내 예측 모드 결정 방법

이민훈^{a)}, 심동규^{a)*}

Intra Prediction Mode Decision Method based on Adaptive Directional Histogram Construction

Minhun Lee^{a)} and Donggyu Sim^{a)*}

요약

본 논문에서는 ECM (Enhanced Compression Model)의 방향성 히스토그램 (Histogram of Gradient, HoG) 기반의 화면내 예측 기법인 DIMD (Decoder-side Intra Mode Derivation) 및 MDIP (Most Dominant Intra Prediction)의 구조적 한계를 개선하기 위한 적응적 히스토그램 구성 방법을 제안한다. 기존 방법은 블록 주변의 참조 영역으로부터 방향성을 유도하되, 해당 방향에 따라 실제 예측에 사용될 참조 샘플의 존재 여부를 고려하지 않아 픽처 경계와 같은 조건에서 예측 성능 저하를 유발할 수 있다. 본 논문에서는 예측 수행 시 참조 샘플이 존재하지 않는 방향성에 대해서는 히스토그램 누적 과정에서 제외하거나, 방향성의 크기를 블록의 크기를 기반으로 축소시키는 방법을 통해 예측의 신뢰도를 개선한다. 제안하는 방법은 ECM-16.1에 구현되어 JVET 공통 실험 조건 하에서 평가되었으며, 부/복호화 런타임 측면에서 추가적인 부담 없이 평균적으로 휘도 성분 기준 0.03%의 부호화 성능 향상을 달성하였다.

Abstract

This paper proposes an adaptive directional histogram construction method to address the structural limitations of HoG (Histogram of Gradient)-based intra prediction techniques, DIMD (Decoder-side Intra Mode Derivation) and MDIP (Most Dominant Intra Prediction), used in ECM (Enhanced Compression Model). While conventional methods in ECM derive directional information from surrounding reference areas, the availability of actual reference samples in those directions is not considered, which can lead to degraded prediction performance, such as near picture boundaries. To address this limitation, the proposed method excludes directions lacking valid reference samples from the histogram accumulation process or attenuates their contribution based on block size. The proposed algorithm is implemented using ECM-16.1, and experimental results demonstrate that the algorithm achieves an average BD-rate reduction of 0.03% for the luma component under JVET common test conditions without additional burden in encoding/decoding runtime.

Keyword : Enhanced Compression Model, Intra Prediction, Joint Video Experts Team, Moving Picture Experts Group, Video Compression

1. 서론

디지털 영상 콘텐츠의 수요가 꾸준히 증가함에 따라, 고효율 비디오 압축 기술의 중요성은 더욱 부각되고 있다. 최근에는 2D 비디오뿐만 아니라, 3D 공간 데이터나 피쳐 맵 (Feature map) 등 다양한 형태의 미디어 데이터를 효율적으로 전송 및 저장하기 위한 표준 기술에서도 비디오 압축 표준 기술이 핵심적으로 활용되고 있다^{[1][2][3][4][5][6]}. 이를 반영하여 JVET (Joint Video Experts Team)은 2020년 7월 최신 비디오 압축 국제 표준 기술인 VVC (Versatile Video Coding)^[7]의 표준화 완료 이후, 차세대 비디오 압축 기술 (Beyond VVC)을 위한 탐색 작업에 착수하였다. 특히 JVET은 전통적인 신호처리 기반 부호화 기술을 중심으로 VVC 대비 향상된 압축 효율을 달성하고자 하는 탐색 실험 (Exploration Experiment 2: Enhanced Compression beyond VVC Capability)^[8]을 포함한 두 탐색 실험을 진행 중이고, 그 일환으로 2021년 5월부터 ECM (Enhanced Compression Model)^[9]이라는 참조 소프트웨어를 개발하여 운용 중이다. ECM은 EE2 실험의 핵심 도구로 활용되며, 각 회의에서 채택된 기술들을 통합하여 관리하고 있다. ECM의 개발 과정에서는 예측 및 변환뿐만 아니라, 인-루프 필터링, 엔트로피 부호화 등 폭넓은 분야의 기술들이 논의되고 있다^{[10][11][12][13][14][15]}. 특히 VVC 표준화 당시 구현 복잡도나 메모리 사용량 등의 제약으로 인해 채택되지 않았던 기술에 더불어 기존 VVC 기술의 개선안과 새롭게 제안된 부호화 기술들이 지속적으로 반영되고 있다. 현재 ECM-16.1 버전을 기준으로, VVC 대비 AI (All Intra) 구성에서 16.53%, RA (Random Access) 구성에서 27.06%의 BD (Bjontegaard Delta)-rate^[16] 절감 효과를 달성한 것으로 보고되고 있다^[17].

ECM에는 시그널링 오버헤드를 줄이고자 디코더 측에서 현재 블록을 복호화하는 과정에 사용될 일부 혹은 전체 정보를 예측하거나 유도하는 다양한 기술이 포함되어 있다. 대표적인 기술인 DIMD (Decoder-side Intra Mode Derivation)는 영상의 공간적 유사성을 기반으로 현재 블록 주변의 복원된 영역의 특성을 분석하여 현재 블록의 화면 내 예측 모드를 유도하고, 유도된 복수의 모드를 활용하여 예측을 수행하는 기술이다. 이는 ECM의 초기 버전에 포함되어 우수한 압축 효율을 보이고 있으며, 이후 예측뿐만 아니라 예측 신호의 보정, 변환, 등 다양한 분야에 접목되어 범용적으로 이용되고 있다^{[18][19]}. 비디오 코덱에서 영상의 부호화 및 복호화는 일반적으로 래스터 스캔 (Raster scan) 순서로 진행되며, 이로 인해 현재 블록의 복호화 시점에서 활용 가능한 주변의 복원된 영역은 그림 1의 예시와 같다.

그림 1의 (a)-(c)는 CTU (Coding Tree Unit) 내 임의의 현재 블록 위치에 대해 활용 가능한 주변 영역의 예시를 나타낸 것이다. 일반적으로 현재 블록의 상단 및 좌측 영역은 예측 및 복원이 완료되어 참조가 가능하며, 경우에 따라 우상단이나 좌하단 영역도 활용할 수 있다. 영상 신호는 지역적으로 공간적 연속성을 갖는 경향이 있기 때문에, 현재 블록과 인접한 복원 영역의 특성은 서로 유사할 가능성이 높다. 특히 복원 영역이 특정 방향성의 특성을 갖는다면 이는 원본 영상에서의 현재 블록이 가진 방향성과 유사할 확률이 높다. 이러한 특성을 바탕으로 DIMD 기술은 디코더에서 복원 영역의 방향성을 분석하여 방향성 히스토그램 (Histogram of Gradient, HoG)을 구성하고, 지배적인 복수의 방향성 화면내 예측 모드를 암시적으로 도출한 후, 이를 현재 블록의 예측에 효과적으로 활용한다. 화면내 예측 수행 시 방향성에 따라 사용되는 참조 라인의 위치와 범위가 상이하지만, 현재 DIMD와 같이 복원 영역의 지역적 방향성만을 고려하는 방식은 참조 영역의 실제 위치를 반영하지 않는 구조적 한계를 가진다.

이는 특히 그림 1의 (d) 및 (e)와 같이 사용 가능한 주변 영역이 상단 혹은 좌측 중 하나로 제한되는 픽처 (Picture) 경계에 위치한 블록에서 예측 정확도의 저하로 이어질 수 있다. 예를 들어, 그림 1의 (d)와 같이 픽처 좌측 경계에 위치한 블록에서, 상단 참조 영역의 방향성이 수평 방향으로 분류되어 수평 방향의 화면내 예측을 수행하는 경우, 좌측

a) 광운대학교 컴퓨터공학과(Department of Computer Engineering, Kwangwoon University)

✉ Corresponding Author : 심동규(Donggyu Sim)
E-mail: dgsim@kw.ac.kr
Tel: +82-2-940-5470

ORCID: <https://orcid.org/0000-0002-2794-9932>

※ 본 연구는 정부 (과학기술정보통신부)의 재원으로 정보통신기획평가원·석사연계ICT핵심인재양성의 지원 (IITP-2025-RS-2022 - 00156215) 및 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터육성 지원사업 (IITP-2025- RS-2022-00156225) 및 2025년도 광운대학교 교내학술연구비 지원 사업의 연구결과로 수행되었음

Manuscript September 11, 2025; Revised September 15, 2025;
Accepted September 15, 2025.

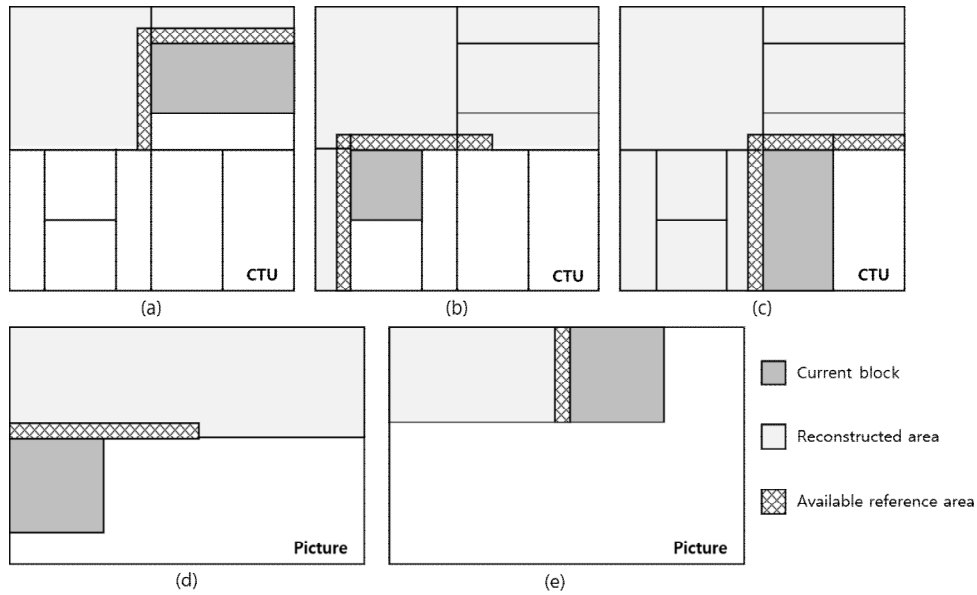


그림 1. 현재 블록의 활용 가능한 주변 영역 예시
Fig. 1. Example of the available reference area of the current block

의 참조 샘플이 필요하지만 해당 위치에는 사용 가능한 참조 샘플이 존재하지 않는다. 이 경우, 디코더는 특정 값을 패딩하여 좌측의 참조 샘플을 대체하게 되며, 이러한 샘플은 현재 블록과의 연관성이 낮을 수 있기 때문에 예측 정확도가 감소할 수 있다. 픽처 경계에서 발생하는 문제는 단일 픽처 내에서 상단과 좌측 경계에 한정되어 나타나지만, 영상 압축에서 필수적으로 고려되어야 할 핵심 요소 중 하나이다. 나아가, 비디오 압축 표준 기술이 실제 응용에 적용될 경우, 다양한 응용 목적 및 시스템 구조에 따라 슬라이스 (Slice), 타일 (Tile), 서브픽처 (Subpicture), VPDU (Virtual Pipeline Data Unit)와 같은 단위가 필수적으로 고려되기 때문에 픽처 경계뿐만 아니라, 각 단위의 경계를 고려한 부호화 기술의 설계가 필요하다.

본 논문에서는 이러한 문제를 해결하고자 참조 영역의 위치를 고려한 적응적 방향성 히스토그램 구성 방법을 제안한다. 제안하는 방법은 예측에 사용될 참조 샘플의 위치를 기반으로 방향성의 통계적 특성을 반영한 히스토그램을 구성하고, 이를 통해 유도된 하나 이상의 방향성 모드를 활용하여 화면내 예측을 수행한다. 결과적으로 현재 블록 주변의 방향성과 더불어 예측에 실질적으로 의미 있는 방향

성들이 반영될 확률이 증가하며, 이는 예측의 정확도와 압축 효율의 향상으로 이어질 수 있다. 본 논문의 2장에서는 DIMD를 포함한 ECM의 방향성 히스토그램 구성 기반 화면내 예측 유도 방법을 설명하고, 3장에서는 제안하는 적응적 히스토그램 구성 방법을 구체적으로 서술한다. 이후, 4장에서는 제안하는 방법의 성능을 제시하며 마지막으로 5장에서 결론을 맺는다.

II. ECM의 방향성 히스토그램 구성 기반 화면내 예측 모드 유도 방법

VVC는 HEVC 대비 부호화 효율 향상을 위해 CTU의 최대 크기를 64×64 에서 128×128 로 확대하고, HEVC에서 사용되던 QT (Quad Tree) 분할 방식에 추가로, 직사각형 블록 분할이 가능한 MTT (Multi-Type Tree) 블록 분할 구조를 도입하였다. 이러한 구조적 확장으로 인해 다양한 블록 형태와 크기에 대응하는 더 세분화된 방향 정보를 활용하는 화면내 예측의 필요성이 증가하였으며, 이에 따라 방향성 화면내 예측 모드의 수가 HEVC의 33개에서 VVC에서

는 65개로 확장되었다. 사용 가능한 예측 모드의 증가는 예측의 정확도 향상으로 이어질 수 있으나, 그에 따른 예측 모드 시그널링 오버헤드의 증가가 수반될 수 있다는 단점이 존재한다. 또한, 보다 초고해상도 영상의 수요가 지속적으로 증가함에 따라, 큰 블록 단위의 효율적인 예측이 더욱 중요해지고 있다. 기존 화면내 예측은 일반적으로 하나의 예측 모드를 하나의 블록에 적용하는 방식이었으나, 방향성이 상이한 영역을 포함하는 큰 블록에 대해 복수의 예측 모드를 효과적으로 적용한다면 블록 전체를 한 번에 처리할 수 있어 부호화 효율 측면에서 이점을 얻을 수 있다. 그러나 이러한 접근은 복수의 예측 모드 정보를 전송해야 하기 때문에 시그널링 오버헤드가 발생할 수 있다.

ECM에는 이러한 문제점을 해결하기 위한 두 가지 부호화 기술이 포함되어 있다. 첫째로, DIMD 기술은 현재 블록 주변 복원 영역의 통계적 특성을 분석하여, 추가적인 시그널링 없이 다수의 지배적인 방향성과 비방향성 예측 모드를 도출함으로써 예측 정확도를 향상시킨다. 둘째로, MDIP (Most Dominant Intra Prediction)는 유사한 통계적 분석을 기반으로 현재 블록에서 사용 가능성이 낮은 예측 모드를 사전에 제거하여 시그널링 오버헤드를 효과적으로 감소시킨다. 두 기술에 대한 상세한 내용은 이후 세부 절에서 다룬다.

1. DIMD (Decoder-side Intra Mode Derivation)

ECM의 DIMD 기술은 VVC의 표준화 과정에서도 논의된 기술이다^[20]. 당시에는 디코딩 복잡도 등으로 인해 채택되지 않았지만, ECM의 개발 과정에서 다양하게 개선된 형태로 제안되었고, 현재는 화면내 예측뿐만 아니라 화면 간 예측, 변환 등 다양한 분야에서 활용되고 있는 기술이다.

DIMD 기술의 개념도는 그림 2의 (a)와 같다. 예측에 사용될 비방향성 모드는 Planar 모드 혹은 BV (Block Vector) 기반 예측 모드 중 하나를 추가적인 시그널링 없이 암시적으로 결정된다. 이를 위해, 현재 블록 주변의 복원된 영역 (Template, 템플릿)을 이용하여 BV 기반 예측 모드 후보들과 Planar 모드에 대해 SATD (Sum of Absolute Transformed Differences) 비용을 계산하고, 그 중 예측 효율이 가장 높은 모드를 선택한다. 또한, 지배적인 방향성을 도출하기 위해 현재 블록 주변의 복원된 영역인 참조 영역에서 지역적인 특징을 분석한다. 이를 위해 주변 위치의 샘플에 대해 기울기 (Gradient) 정보를 추출하고, 이를 기반으로 HoG를 구성하여 지배적인 복수의 화면내 예측 모드를 유도한다. 최종 예측 신호는 도출된 비방향성 모드와 방향성 모드들의 가중합으로 구성되며, 이를 위해 각 예측 모드에

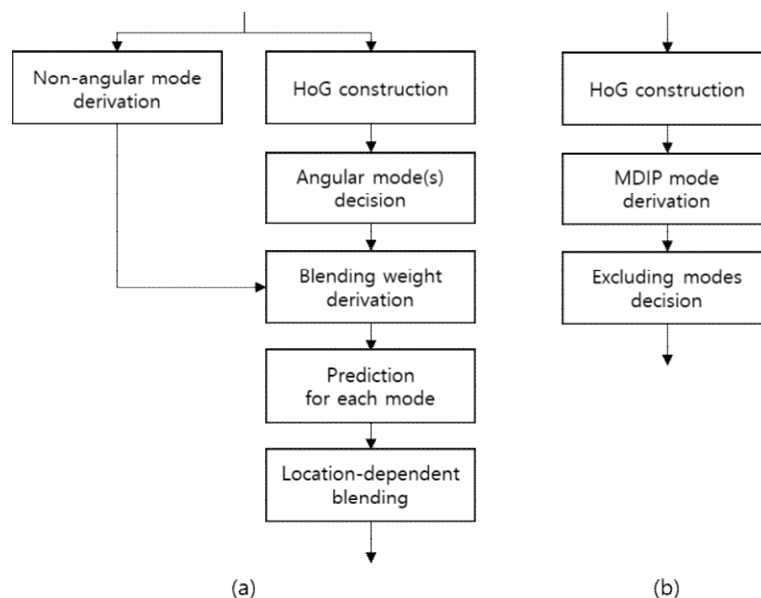


그림 2. (a) DIMD 기술의 개념도 (b) MDIP 기술의 개념도

Fig. 2. (a) Conceptual diagram of DIMD (b) Conceptual diagram of MDIP

대해 적절한 가중치를 결정한다. 비방향성 모드의 가중치는 고정된 값이 사용되며, 방향성 모드의 경우 해당 모드가 상단 복원 영역, 좌측 복원 영역, 혹은 양쪽 모두로부터 도출된 정도에 따라 샘플 단위의 적응적 가중치가 암시적으로 설정된다. 마지막으로, 유도된 가중치를 기반으로 각 예측 모드의 가중합을 수행하여 최종 예측 신호를 생성한다.

2. MDIP (Most Dominant Intra Prediction)

ECM에 포함된 DIMD 기술은 효과적인 예측을 위해 현재 블록 주변 복원 영역인 참조 영역의 지역적 특성을 분석하고, 이를 기반으로 지배적인 방향성을 도출하여 예측에 활용하는 반면, MDIP 기술은 유사한 분석 과정을 통해 비지배적인 방향성을 사전에 제거함으로써, 시그널링 오버헤드를 줄이는 데 중점을 둔다. MDIP 기술의 개념도는 그림 2의 (b)에 나타나 있으며, DIMD 기술과 마찬가지로 복원된 영역으로부터 기울기 정보를 추출하고 HoG를 구성하여 제거할 예측 모드를 결정한다. 구체적으로, MDIP는 HoG를 기반으로 최대 8개의 화면내 예측 모드를 도출하고, 현재 블록의 상단 및 좌측 1-라인 템플릿을 활용해 템플릿 비용을 계산한다. 이 중 최소 비용을 갖는 후보를 MDIP 모드로 지정한다. 이후 HoG 내 기울기의 크기가 작은 20개의 방향을 비지배적 후보로 판별하여 현재 블록의 화면내 예

측 모드 후보 리스트 (MPM (Most Probable Mode), SMPM (Secondary MPM), 및 비-MPM)에서 제외한다. 이 과정을 통해 후보 수는 단계적으로 축소된다. ECM에서 사용되는 전체 화면내 예측 모드는 67개이며, MPM과 SMPM 후보 22개를 제외하면 비-MPM 후보는 45개가 된다. 이후 HoG를 기반으로 제외된 20개의 비지배적 후보와 MDIP 모드 1개가 더해져 최종적으로 비-MPM 후보는 24개로 줄어든다. 결과적으로 MDIP는 선택 확률이 낮은 모드를 제거함으로써 비-MPM 후보를 45개에서 24개로 효과적으로 축소하고, 그 결과 시그널링 오버헤드를 줄임으로써 예측 효율을 향상시키는 데 기여한다. MDIP의 HoG 구성 방식은 DIMD 기술과 동일한 개념을 기반으로 수행하지만, 분석에 활용되는 복원 영역 선택 방식에서 두 기술 간 차이가 존재한다. 이에 대한 구체적인 내용은 이후 세부 절에서 상세히 다룬다.

3. DIMD 및 MDIP 기술의 HoG 구성 방법

DIMD 및 MDIP 기술에서 HoG 구성을 수행하기 위한 주변 참조 영역은 그림 3과 같다. 이때, a 와 b 는 각각 현재 블록의 너비와 높이인 W 및 H 로 결정되며, $c \sim h$ 는 적용되는 기술 (DIMD 또는 MDIP)과 현재 블록의 크기 등에 따라 적응적으로 결정되는 변수이다.

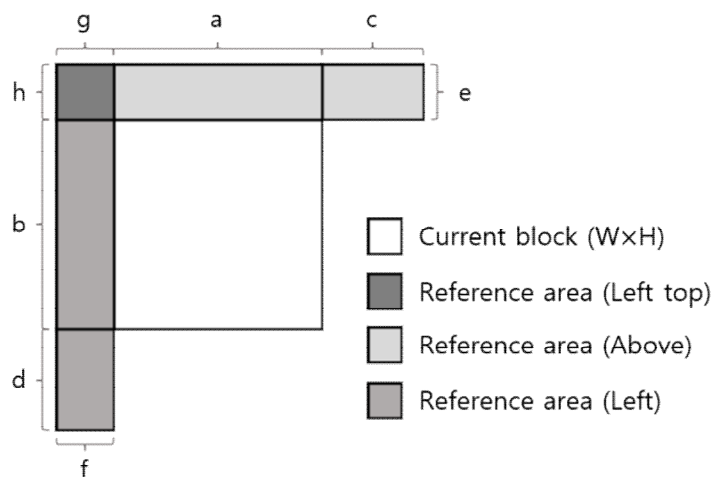


그림 3. HoG 구성을 위한 참조 영역

Fig. 3. The reference area for HoG construction

DIMD 및 MDIP 기술 모두 정의된 참조 영역 내에서 수평 및 수직 방향의 지역적인 기울기 정보 (G_{hor} 및 G_{ver})를 추출하기 위해, 수평 및 수직 소벨 (Sobel) 필터 또는 식 (1)과 (2)에 제시된 2×2 수평 및 수직 필터가 적용된다. 사용되는 필터의 종류는 현재 블록의 크기에 따라 결정되며, 블록 면적이 32보다 작은 경우, 연산 복잡도 등을 고려하여 2×2 필터를, 그 외의 경우에는 보다 정밀한 분석이 가능한 소벨 필터가 적용된다. 각 샘플 위치에서 계산된 기울기 정보는 식 (3)을 통해 기울기 각도 (θ)로 변환되고, 변환된 각도는 65개의 방향성 화면내 예측 모드 중 가장 유사한 하나의 모드로 맵핑된다. 이때, 각도의 크기 (Amp)는 식 (4)와 같이 계산되어 맵핑된 모드와 각도의 크기를 활용하여 HoG를 구성한다.

$$F_x = \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix} \quad (1)$$

$$F_y = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} \quad (2)$$

$$\theta = \arctan\left(\frac{G_{ver}}{G_{hor}}\right) \quad (3)$$

$$Amp = |G_{ver}| + |G_{hor}| \quad (4)$$

상기 과정은 그림 3에 도시된 상단, 좌측, 좌상단 참조 영역에 각각 수행되며, 이후 각 영역에서 구성된 HoG를 합산하여 지배적 혹은 비지배적인 예측 모드를 도출한다. DIMD 및 MDIP 기술 모두에서 필터 종류에 따라 좌상단 참조 영역의 크기를 결정하는 변수 g 및 h 의 값이 달라진다. 2×2 필터를 사용하는 경우 두 변수는 각각 2로 설정되며, 소벨 필터를 사용하는 경우에는 보다 넓은 범위의 기울기 계산을 위해 g 와 h 모두 4로 설정된다. 또한, 상단 및 좌측 참조 영역의 크기를 정의하는 변수 c ~ f 는 적용되는 기술과 블록의 크기에 따라 적응적으로 설정된다.

먼저 DIMD 기술에서는 c 와 d 는 항상 4로 고정되고, 현재 블록이 I-슬라이스에 속하고 면적이 128 미만일 경우, 기울기 분석에 사용되는 필터의 종류에 따라 변수 e 와 f 가 결정된다. 소벨 필터가 적용되는 경우에는 e 와 f 는 3으로

설정되며, 2×2 필터가 사용되는 경우에는 e 와 f 가 모두 2로 설정된다. 그 외의 경우에는 더욱 넓은 영역에서 방향성 분석을 수행하기 위해 e 와 f 값이 최대 12까지 확장될 수 있다. 넓은 참조 영역을 활용하면 현재 블록 주변의 방향성 특성을 보다 정밀하게 반영할 수 있다는 장점이 있지만, 그에 따른 계산 복잡도 증가가 불가피하다. 이를 완화하기 위해, DIMD 기술에는 누적된 기울기 크기의 합이 특정 임계값을 초과할 경우 HoG 구성을 조기 종료하는 메커니즘을 도입하였다. 해당 임계값은 고정된 값을 기본으로 사용하고, 현재 블록과 인접한 좌측 및 상단 블록의 크기를 고려하여 추가적인 임계값이 적용될 수 있다. 반면, MDIP 기술은 시그널링 오버헤드 감소를 목표로 비지배적인 방향성 예측 모드를 사전에 제거하는 것이 목적이며, 이에 따라 참조 영역 구성 방식도 DIMD와 상이하게 정의된다. MDIP 기술은 현재 블록과 인접하고 예측에 실제로 사용되는 샘플이 존재하는 영역을 중심으로 참조 영역이 설정되며, 구체적으로는 참조 영역의 범위를 결정하는 변수 c 와 d 가 각각 현재 블록의 너비 W 와 높이 H 로 설정된다. 또한, e 와 f 는 동일한 값을 사용하고, 현재 블록의 면적 및 기울기 계산에 사용되는 필터의 종류에 따라 다음과 같이 적응적으로 결정된다. 블록 면적이 128 이하인 경우, 소벨 필터가 적용되면 각각 3으로 설정되고, 2×2 필터가 사용되면 두 변수 모두 2로 설정된다. 반면, 블록 면적이 128을 초과하는 경우에는 보다 넓은 공간에서의 방향성 분석을 수행하고자 소벨 필터 사용 시 e 와 f 는 4로, 2×2 필터 적용 시 3으로 각각 확장된다.

앞서 언급한 바와 같이, DIMD 및 MDIP 기술에서의 HoG 구성 과정은 상단, 좌측, 좌상단의 세 참조 영역에 각각 수행되고, 이후 각 영역에서 유도된 HoG를 통합하여 지배적 또는 비지배적인 예측 모드를 도출한다. 그러나 이러한 방식은 방향성 정보의 통계적 특성만을 고려할 뿐, 해당 방향성 예측 모드를 적용할 경우 실제 사용될 참조 샘플의 존재 여부나 위치는 반영하지 않는 구조적 한계를 지닌다. 예를 들어, 그림 1의 (d)와 같이 픽처 좌측 경계에 위치한 블록에서 상단 참조 영역의 분석 결과 수평 방향의 지배적인 방향성이 도출되었을 경우, 수평 방향의 예측 수행에는 좌측 참조 샘플이 필요하다. 그러나 픽처 경계에 의해 좌측

참조 샘플이 존재하지 않으므로, 디코더는 이를 중간값 (예: 입력 영상의 비트 심도가 10비트인 경우, 512)으로 패딩하여 대체하게 된다. 그렇게 되면 해당 샘플들은 실제 블록 특성과의 상관성이 낮을 수 있기 때문에 예측 정확도 저하로 이어질 수 있다. 현재 ECM의 DIMD 및 MDIP 기술은 참조 샘플의 존재 여부나 위치 정보를 고려하지 않고 HoG를 구성하기 때문에, 특히 픽처 경계에 위치한 블록의 예측 성능에 부정적인 영향을 미칠 수 있다. 본 논문에서는 이러한 문제를 해결하기 위해 참조 영역의 위치 정보를 반영한 적응적 방향성 히스토그램 구성 방법을 제안한다. 제안하는 방식은 단순히 방향성의 통계적 강도만을 기반으로 HoG를 구성하는 기존 방식과 달리, 방향성 정보와 함께 해당 방향에 따른 참조 샘플의 가용성까지 고려하여, 실질적으로 예측 효율이 높은 방향성만을 반영함으로써 예측 성능을 개선할 수 있다.

III. 제안하는 적응적 방향성 히스토그램 구성 방법

본 장에서는 2장에서 언급한 기존 기술의 한계를 극복하

기 위해 제안하는 참조 샘플 가용성을 고려한 적응적 방향성 히스토그램 구성 방법에 대해 상세히 설명한다. 제안 방법은 기존 DIMD 및 MDIP 기술이 참조 영역의 방향성만을 기반으로 예측 모드를 유도하는 데 반해, 실제 예측 수행 시 사용 가능한 참조 샘플의 존재 여부와 위치를 함께 고려함으로써, 예측 신호의 정확도를 향상시키는 것을 목표로 한다.

제안하는 방법을 포함한 HoG 구성 방법의 블록도는 그림 4와 같다. 제안 방법은 기존 ECM-16.1의 HoG 구성 방법에 기반하되, 히스토그램 구성 시 각도의 크기를 참조 영역의 특성과 예측 샘플의 가용성에 따라 적응적으로 조정함으로써 보다 효과적인 방향성 분석이 가능하도록 한다. 구체적으로는 상단 및 좌측 참조 영역 각각에 동일한 매커니즘으로 다음 두 가지 제안하는 방법이 적용된다. 이는 각도 크기 조정 (그림 4의 Amplitude adjustment) 단계에 해당하며, 이에 대해 구체적인 설명은 다음과 같다.

그림 4의 각도 크기 조정 단계는 제안 방법의 핵심으로, 참조 영역의 유효성에 따라 각도의 크기 (*Amp*)를 조정하는 과정을 수행한다. 먼저, 현재 블록이 픽처 경계에 위치하는 등의 이유로 상단 혹은 좌측 참조 영역이 사용 불가능한 경우, 특정 방향의 예측에 필요한 참조 샘플이 존재하지 않

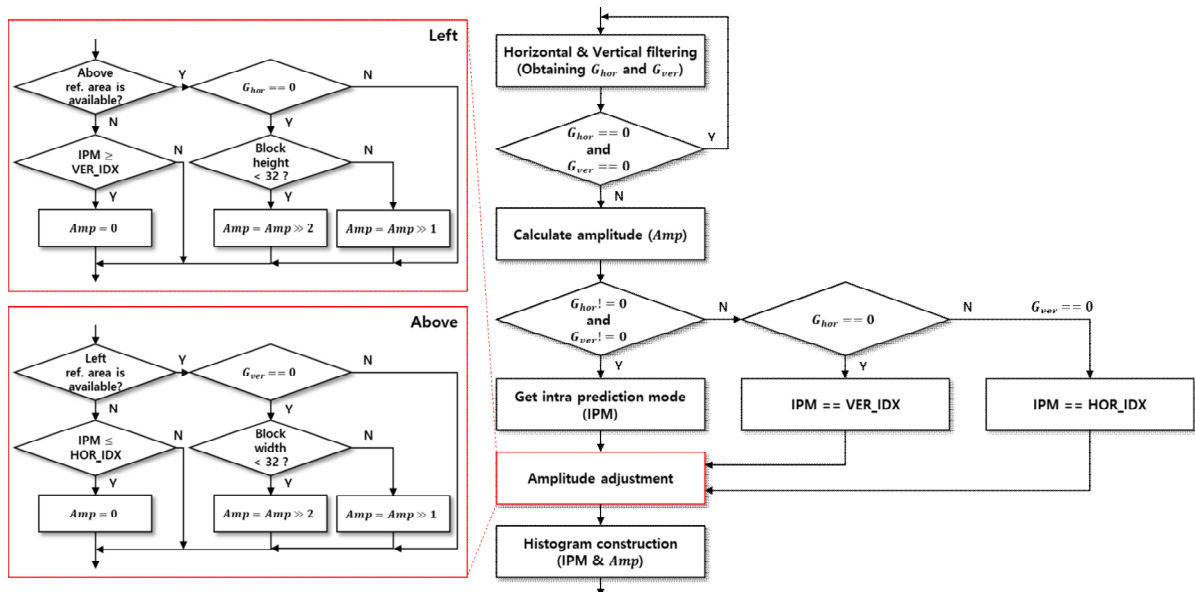


그림 4. 제안하는 방법의 블록도

Fig. 4. Block diagram of the proposed method

을 수 있다. 이러한 경우, 앞서 설명한 바와 같이 디코더는 일정한 고정값으로 패딩된 샘플을 이용해 예측을 수행하게 되며, 이러한 샘플들은 블록의 원본 신호와 낮은 상관성을 가질 가능성이 높다. 결과적으로는 해당 방향으로의 예측은 정확도가 낮아지고, 전반적인 부호화 효율의 저하로 이어질 수 있다. 이를 고려하여 제안하는 방법에서는 예측에 필요한 참조 샘플이 존재하지 않는 방향에 대해서는 해당 방향으로의 각도 크기를 0으로 설정 ($Amp = 0$)하여 HoG 누적 과정에서 해당 방향이 제외되도록 한다. 구체적으로, 상단 참조 영역에 대해 통계적 분석을 수행하는 경우 좌측 참조 영역이 비활성화된 상태라면, 수평 방향보다 작은 방향 인덱스를 갖는 모드들에 대해서 각도의 크기를 0으로 맵핑하여 해당 방향성을 히스토그램에서 배제한다. 반대로, 좌측 참조 영역을 분석하는 경우 상단 참조 영역이 비활성화되어 있다면, 수직 방향보다 큰 방향 인덱스에 해당하는 모드들의 각도 크기를 0으로 설정하여 제외 처리한다. 이와 같이 참조 영역의 가용성을 직접 반영하는 각도 크기 조정 과정을 통해, 예측 가능성이 낮은 방향성은 사전에 억제되고, HoG의 신뢰도를 높일 수 있다.

또한, 기존 ECM의 HoG 구성 방식에서는 수평 및 수직 기울기를 각각 계산한 뒤, 수평 기울기가 0인 경우 수직 방향 모드 (VER_IDX)로, 수직 기울기가 0인 경우 수평 방향 모드 (HOR_IDX)로 맵핑하는 단순 규칙이 적용된다. 그러나 앞서 설명한 바와 같이, 수평 방향 예측은 좌측 참조 샘플만을, 수직 방향 예측은 상단 참조 샘플만을 이용하므로 예측에 필요한 참조 샘플이 실제로 존재하지 않는 경우에도 해당 방향성이 유도될 수 있다는 구조적 문제가 존재한다. 특히 DIMD 기술은 유도된 복수의 방향성 모드로 현재 블록의 예측을 수행하기 때문에, 현재 블록의 특성과 무관한 유도된 예측 모드로 예측이 수행될 가능성이 있다. 이를 위해, 현재 ECM에는 HoG 정보를 기반으로 암시적으로 샘플 단위의 적응적 가중치를 결정하는 방법이 포함되어 있지만, 예측 정확도를 보다 실질적으로 개선하기 위해서는 HoG 구성 단계에서부터 방향성의 기여도를 현실적으로 반영할 수 있는 조정이 필요하다. 이를 해결하고자 참조 영역의 위치에 따라 예측 신뢰도가 낮은 수직 또는 수평 방향성에 대해서는 히스토그램 누적 시 각도의 크기를 축소시키는 방식을 제안한다. 구체적으로, 상단 참조 영역에서 수직

방향 예측에 필요한 참조 샘플이 존재하지 않는 경우, 해당 방향으로 맵핑되는 각도의 크기를 블록 크기에 따라 $1/2$ ($Amp \gg 1$) 또는 $1/4$ ($Amp \gg 2$)로 축소하여 히스토그램에 반영한다. 동일하게, 좌측 참조 영역에서도 수평 방향 예측이 비효율적인 경우 동일한 감쇠 방식을 적용한다. 이와 같은 방식은 단순히 방향성 빈도를 누적하는 기존 방식과 달리, 방향성과 참조 샘플의 실질적 유효성까지 고려하여 예측 가능성이 높은 방향에 상대적으로 더 높은 중요도를 부여하는 효과를 얻을 수 있다. 또한, 감소된 상태에서도 수직 및 수평 방향 모드가 여전히 지배적인 방향성으로 분류되는 경우, 이는 해당 방향성의 특성이 주변 영역에서 강하게 관찰되었음을 의미한다. 따라서 이러한 경우에는 오히려 현재 블록에 해당 방향성 모드를 적용하는 것이 공간적 연속성을 유지하며 예측 효율을 높이는 데 기여할 수 있다.

IV. 실험 결과

본 논문에서 제안하는 적응적 히스토그램 구성 방법의 성능을 평가하기 위하여 제안 방법을 ECM-16.1 버전^[9]에 구현하고, 기존 ECM 대비 성능 비교 실험을 수행하였다. 실험은 동일한 조건 하에서의 성능 비교를 위해, JVET에서 정의한 공통 실험 조건 (Common Test Conditions, CTC)^[21]에 따라 AI 부호화 환경에서 수행되었으며, 실험 영상으로는 Class A1, A2, B, C, D, E, F 및 TGM에 속한 영상들을 사용하였다. 또한, ECM-16.1은 VVC의 참조 소프트웨어인 VTM (VVC Test Model) 대비 10배 이상의 부호화 복잡도를 갖기 때문에, 이를 고려하여 블록 분할 깊이를 제한하는 인코딩 설정^[22]을 적용하였다. Class A1 및 A2의 영상은 부호화 복잡도 완화를 위해 1초 분량의 프레임만을 사용하여 비교 분석을 수행하였으며, 성능 평가를 위해 BD-rate^[16]를 정량적 지표로 사용하였고 부호화 및 복호화 시간 또한 함께 측정하여 비교 분석을 수행하였다.

표 1은 ECM-16.1 대비 제안하는 적응적 히스토그램 구성 방법의 실험 결과를 나타낸다. 전체 실험 영상에 대해 평균적으로 휘도 성분 기준 0.03%의 성능 향상을 보였으며, 이는 부/복호화 런타임 (Runtime) 측면에서 추가적인 부담 없이 달성된 결과이다. 추가적으로 색차 성분

표 1. 제안하는 방법의 ECM-16.1 대비 성능

Table 1. Performance of the proposed method over ECM-16.1

Class	Sequence	BD-rate (Y)	BD-rate (Cb)	BD-rate (Cr)	EncT	DecT
A1	Tango2	-0.01%	-0.10%	-0.09%	99.6%	99.9%
	FoodMarket4	0.02%	0.00%	0.06%	100.1%	99.5%
	Campfire	-0.03%	-0.02%	-0.06%	100.5%	100.1%
A2	CatRobot	-0.01%	-0.01%	-0.20%	99.5%	100.2%
	DaylightRoad2	-0.04%	0.02%	-0.12%	100.6%	99.8%
	ParkRunning3	-0.02%	-0.08%	-0.02%	99.6%	100.2%
B	MarketPlace	0.01%	-0.29%	-0.02%	99.3%	100.1%
	RitualDance	0.01%	-0.06%	-0.04%	100.3%	99.7%
	Cactus	-0.03%	-0.26%	0.11%	100.2%	99.6%
	BasketballDrive	-0.01%	0.10%	-0.01%	99.8%	99.4%
	BQTerrace	0.00%	0.02%	0.10%	100.1%	100.2%
C	BasketballDrill	-0.02%	-0.01%	0.15%	99.7%	100.3%
	BQMall	-0.02%	-0.03%	0.03%	100.5%	100.5%
	PartyScene	0.01%	-0.10%	-0.16%	100.2%	99.8%
	RaceHorsesC	-0.02%	-0.06%	-0.08%	99.8%	100.2%
D	BasketballPass	-0.06%	0.20%	-0.01%	100.3%	100.1%
	BQSquare	-0.04%	0.02%	-0.25%	100.4%	100.1%
	BlowingBubbles	-0.01%	-0.20%	0.21%	100.1%	100.2%
	RaceHorses	-0.06%	-0.02%	-0.04%	99.6%	100.5%
E	FourPeople	0.02%	-0.01%	-0.05%	99.7%	99.8%
	Johnny	0.01%	0.04%	0.03%	100.3%	100.2%
	KristenAndSara	-0.02%	-0.09%	0.01%	100.6%	100.3%
F	BasketballDrillText	-0.01%	0.20%	-0.01%	99.6%	100.1%
	ArenaOfValor	-0.01%	0.10%	0.01%	100.1%	100.2%
	SlideEditing	-0.02%	-0.03%	-0.08%	100.3%	100.6%
	SlideShow	-0.09%	-0.39%	-0.05%	99.9%	100.2%
TGM	FlyingGraphics	0.00%	-0.01%	-0.19%	100.2%	100.1%
	Desktop	-0.08%	-0.30%	-0.08%	99.7%	99.6%
	Console	-0.12%	-0.28%	-0.03%	100.2%	99.9%
	ChineseEditing	-0.16%	0.02%	-0.06%	100.1%	100.1%
Overall (A, B, C, D, E)		-0.02%	-0.04%	-0.02%	100.0%	100.0%
Overall (F, TGM)		-0.06%	-0.09%	-0.06%	100.0%	100.1%
Overall (All)		-0.03%	-0.06%	-0.03%	100.0%	100.1%

성능을 분석한 결과, ECM-16.1과 비교했을 때 유사한 수준을 유지하는 것으로 확인되었다. 이는 제안하는 방법이 참조 영역의 방향성 정보를 활용하는 과정에서 휘도 성분에서 효과가 두드러지기 때문에 판단된다. 다만, 제안하는

방법은 참조 영역의 특성과 예측 샘플의 가용성에 따라 각도를 적응적으로 조정하므로 계산 복잡도 측면에서는 일정 수준의 연산이 추가된다. 그럼에도 불구하고, 전체 실험에서 런타임 차이는 미미하여 실질적인 부호화 효율 향상을

표 2. CTC 조건에서 제안하는 방법의 ECM-16.1 대비 성능

Table 2. Performance of the proposed method under the CTC condition over ECM-16.1

Sequence	BD-rate (Y)	BD-rate (Cb)	BD-rate (Cr)	EncT	DecT
BasketballPass	-0.04%	0.20%	-0.01%	100.3%	100.2%
BQSquare	-0.02%	0.02%	-0.25%	100.2%	100.1%
BlowingBubbles	-0.02%	-0.20%	0.21%	100.3%	100.2%
RaceHorses	-0.04%	-0.02%	-0.04%	100.1%	100.3%
Overall (Class D)	-0.03%	-0.01%	-0.08%	100.2%	100.2%

달성했다고 볼 수 있다. 또한, 제안 방법의 성능이 인코딩 설정에 따라 달라질 가능성을 검증하기 위해 Class D 영상에 대해 CTC 조건에서도 별도의 실험을 수행하였으며, 그 결과는 표 2에 제시하였다.

기존에 분할 깊이를 제한하는 인코딩 설정에서는 휘도 성분에서 평균 0.04%의 성능 향상을 보였으며, 동일한 영상에 대해 CTC 조건을 적용한 경우에는 0.03%의 성능 향상을 확인하였다. 두 결과 간 성능 차이는 존재하나 그 폭은 매우 미미한 수준으로, 이는 블록 크기 분포 차이에 따른 자연스러운 결과로 해석될 수 있다. 즉, 보다 세밀한 분할로 작은 블록이 많이 생성되는 환경에서는 성능 향상 폭이 다소 줄어들 수 있으나, 전반적으로 제안 방법이 다양한 인코딩 설정에서도 안정적으로 동작함을 보여준다. 특히, Class F 및 TGM 실험 영상에서는 자연 영상 대비 상대적으로 높은 부호화 성능 향상을 확인하였고, 이는 해당 클래스가 포함한 스크린 콘텐츠에 기인한다. 스크린 콘텐츠 영상은 수직 및 수평 방향의 평활성이 강하게 나타나므로 제안하는 방법이 적용되는 빈도가 높고, 이로 인해 효율적인 방향성 예측이 가능해져 더 높은 부호화 효율을 얻은 것으로 분석된다. 예를 들어, PPT 슬라이드, 문서, 웹페이지 등과 같은 PC 화면으로 구성된 SlideShow, SlideEditing, Desktop, Console, ChineseEditing과 같은 영상은 높은 평활도를 가지며, 그 결과 제안하는 방법에 대한 성능이 다른 영상 대비 더욱 우수한 것으로 관찰되었다. 반면, 자연 영상과 스크린 콘텐츠가 혼합된 형태의 BasketballDrillText 영상이나 게임 화면 기반의 ArenaOfValor 영상의 경우, 보다 다양한 방향성이 존재하여 상대적으로 낮은 성능 개선 폭을 보였다. 이는 해당 콘텐츠의 공간적 방향성이 비교적 고르게 분포되어 있어 제안하는 기법의 적용 빈도가 낮았기 때문으로 해석된다. 제안하는 방법은 JVET의 CTC 하에서 주로 좌측 및 상단 픽처 경계 인접 블록을 중심으로 제한적으로 적용되었음에도 전체 실험 영상에 걸쳐 일관된 성능 향상을 보였다. 이는 부호화 기술 설계 시, 참조 샘플의 실제 가용성과 위치 정보를 반영하는 방향성 분석 방식의 타당성을 입증하며, 향후 차세대 비디오 압축 기술의 표준화 과정에서도 픽처, 슬라이스, 서브픽처, 타일, VPDU 등의 경계를 고려한 보다 정밀한 방향성 기반 예측 기술의 필요

성을 시사한다.

V. 결 론

본 논문에서는 ECM 기반의 화면내 예측 기술인 DIMD 및 MDIP의 방향성 유도 방식에서 발생하는 구조적 한계를 해결하고자, 참조 샘플의 가용성을 고려한 적응적 방향성 히스토그램 구성 방법을 제안하였다. 기존 기술은 복원된 참조 영역에서의 방향성 통계만을 기반으로 예측 모드를 유도함에 따라, 예측 수행 시 실제로 참조 샘플이 존재하지 않음에도 불구하고 해당 방향성이 선택되는 문제가 있었다. 이는 특히 픽처 경계와 같은 참조 샘플이 제한되는 위치에서 예측 정확도의 저하로 이어질 수 있다.

이를 해결하기 위해 본 논문에서는 참조 샘플의 존재 여부 및 위치 정보를 HoG 구성 단계에 반영하는 두 가지 방법을 제시하였다. 먼저, 예측 수행에 필요한 참조 샘플이 존재하지 않는 방향성에 대해서는 해당 각도의 크기를 히스토그램 누적 과정에서 제거함으로써 예측 신뢰도가 낮은 방향성이 선택되지 않도록 하였다. 또한, 수평 및 수직 방향 예측과 같이 참조 샘플의 존재 여부가 예측 성능에 직접적인 영향을 주는 경우에는, 블록 크기와 참조 샘플의 사용 가능 여부에 따라 각도의 크기를 축소시켜 히스토그램 구성 과정에서 현실적인 방향성 가중치를 적용하였다.

제안하는 방법은 ECM-16.1에 구현되어 JVET의 공통 실험 조건 (CTC) 하에서 다양한 영상을 대상으로 성능을 검증하였다. 실험 결과, 추가적인 부/복호화 시간의 증가 없이 평균적으로 휘도 성분 기준 0.03%의 BD-rate 감소를 달성하였으며, 특히 스크린 콘텐츠 기반 영상에서 상대적으로 높은 성능 향상을 확인하였다. 이는 제안 기법이 실제 예측에 의미 있는 방향성을 효과적으로 반영하였음을 보여준다. 제안하는 방법은 ECM의 DIMD 및 MDIP와 같은 방향성 히스토그램 구성 기반 화면내 예측 모드 유도 방법이 향후 표준에 채택된다면, 타일, 서브픽처, VPDU 등 다양한 단위의 경계에서 고려될 수 있음에 따라, 다양한 응용 시나리오에서 효율적인 예측을 수행하는 기술로 기여할 수 있을 것으로 기대된다.

참 고 문 헌 (References)

- [1] D. Graziosi, O. Nakagami, S. Kuma, A. Zaghetto, T. Suzuki, and A. Tabatabai, "An overview of ongoing point cloud compression standardization activities: video-based (V-PCC) and geometry-based (G-PCC)," *APSIPA Transactions on Signal and Information Processing*, Vol.9, p.e13, Apr. 2020.
doi: <https://doi.org/10.1017/ATSIP.2020.12>
- [2] J. Lee, "Introduction to Video-based Dynamic Mesh Coding Standardization Trends," *Broadcasting and Media Magazine*, Vol.28, No.3, pp.23-32, Jul. 2023.
- [3] M. Kim, J. Byeon, and D. Sim, "Valence based Lifting Wavelet Transform for Video-based Dynamic Mesh Compression," *Journal of Broadcast Engineering*, Vol.29, No.1, pp.42-56, Jan. 2024.
doi: <https://doi.org/10.5909/JBE.2024.29.1.42>
- [4] H. Choi and D. Sim, "Rate Control Method for Encoding Texture Map of Dynamic Mesh-based Video," *Journal of Broadcast Engineering*, Vol. No. pp.198-225, Mar. 2025.
doi: <https://doi.org/10.5909/JBE.2025.30.2.198>
- [5] M. Lee, S. Park, S.-J. Oh, Y. Kim, S. Y. Jeong, J. Lee, and D. Sim, "Transform-Based Feature Map Compression Method for Video Coding for Machines (VCM)," *Electronics*, Vol.12, No.19, pp.4042, 2023.
doi: <https://doi.org/10.3390/electronics12194042>
- [6] M. Lee, H. Choi, J. Kim, J. Do, H. Kwon, S. Y. Jeong, D. Sim, and S.-J. Oh, "Feature Map Compression for Video Coding for Machines Based on Receptive Block Based Principal Component Analysis," *IEEE Access*, Vol.11, pp.26308-26319, Mar. 2023.
doi: <https://doi.org/10.1109/ACCESS.2023.3254589>
- [7] M. Lee, H. Song, J. Park, B. Jeon, J. Kang, J.-G. Kim, Y.-L. Lee, J.-W. Kang, and D. Sim, "Overview of Versatile Video Coding (H.266/VVC) and Its Coding Performance Analysis," *IEIE Transactions on Smart Processing & Computing*, Vol.12, No.2, pp.122-154, Apr. 2023.
doi: <https://doi.org/10.5573/IEIESPC.2023.12.2.122>
- [8] V. Seregin, J. Chen, R. Chernyak, K. Naser, J. Strom, F. Wang, M. Winken, X. Xiu, and K. Zhang, "EE2: Summary report of exploration experiment on enhanced compression beyond VVC capability," 37th JVET Meeting, Geneva, Document JVET-AK0024, Jan. 2025.
- [9] Enhanced Compression Model 16.1 (ECM16.1). https://vcgit.hhi.fraunhofer.de/ecm/ECM/-/tree/ECM-16.1?ref_type=tags
- [10] S. Jeon, S. C. Lim, J. Kang, and J.-G. Kim, "Overview of ECM," *Broadcasting and Media Magazine*, Vol.29, No.1, pp.19-32, Jan. 2024.
- [11] J. Kang, J. Kim, H. Han, and H. Choi, "ECM Intra prediction," *Broadcasting and Media Magazine*, Vol.29, No.1, pp.33-43, Jan. 2024.
- [12] Y.-J. Ahn, "ECM Inter prediction I," *Broadcasting and Media Magazine*, Vol.29, No.1, pp.44-54, Jan. 2024.
- [13] K. Choi, "ECM Inter prediction II," *Broadcasting and Media Magazine*, Vol.29, No.1, pp.55-64, Jan. 2024.
- [14] K. Kim, D. Kim, J.-H. Son, and J. S. Kwak, "ECM Transform and Quantization," *Broadcasting and Media Magazine*, Vol.29, No.1, pp.65-75, Jan. 2024.
- [15] M. Lee and D. Sim, "In Loop Filtering Tools in ECM Reference Software for Beyond VVC," *Broadcasting and Media Magazine*, Vol.29, No.1, pp.76-88, Jan. 2024.
- [16] G. Bjøntegaard, "Calculation of average PSNR differences between RD-Curves," ITU-T SG16/Q6, Austin, TX, USA, Document VCEG-M33, Apr. 2001.
- [17] V. Seregin, J. Chen, R. Chernyak, F. L. Léannec, and K. Zhang, "JVET-AHG report: ECM software development (AHG6)," 38th JVET Meeting, Geneva, Document JVET-AL0006, Apr. 2025.
- [18] M. Coban, R.-L. Liao, K. Naser, J. Ström, and L. Zhang, "Algorithm description of Enhanced Compression Model 16 (ECM 16)," 37th JVET Meeting, Teleconference, Document JVET-AK2025, Jan. 2025.
- [19] M. Abdoli, R. G. Youvalari, K. Naser, K. Reuzé, and F. L. Léannec, "Video compression beyond VVC: Quantitative analysis of intra coding tools in Enhanced Compression Model (ECM)," *arXiv preprint*, 2024.
doi: <https://doi.org/10.48550/arXiv.2404.07872>
- [20] M. Abdoli, T. Guionnet, E. Mora, M. Raulet, S. Blasi, A. Seixas Dias, and G. Kulupana, "Non-CE3: Decoder-side Intra Mode Derivation (DIMD) with prediction fusion using Planar," 15th JVET Meeting, Gothenburg, Document JVET-O0449, Jul. 2019.
- [21] M. Karczewicz, and Y. Ye, "Common Test Conditions and evaluation procedures for enhanced compression tool testing," 31st JVET meeting, Geneva, Document JVET-AE2017, Jul. 2023.
- [22] T. Ikai, and K.-W. Liang, "[AHG7][AHG17] Further consideration of common test condition," 38th JVET Meeting, Teleconference, Document JVET-AL0245, Apr. 2025.

저 자 소 개



이 민 훈

- 2019년 2월 : 광운대학교 수학과, 전자공학과 (복수전공) 학사
- 2021년 2월 : 광운대학교 전자공학과 석사
- 2021년 3월 ~ 현재 : 광운대학교 컴퓨터공학과 박사과정
- ORCID : <https://orcid.org/0000-0001-8165-5380>
- 주관심분야 : 영상신호처리, 영상압축, 컴퓨터비전

저 자 소 개



심 동 규

- 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
- ORCID : <https://orcid.org/0000-0002-2794-9932>
- 주관심분야 : 영상신호처리, 영상압축, 컴퓨터비전