

적응적 히스토그램 구성 기반의 암시적 화면내 예측 모드 유도 방법

이민훈, 심동규*
광운대학교 컴퓨터공학과

minhun@kw.ac.kr, *dgsim@kw.ac.kr

Decoder side Intra Mode Derivation Method based on Adaptive Histogram Construction

Lee Minhun, Sim Donggyu*

*Department of Computer Engineering, Kwangwoon University

요약

본 논문은 JVET (Joint Video Experts Team) 은 VVC (Versatile Video Coding)보다 높은 압축 성능을 가지는 신호처리 기반 차세대 비디오 압축 표준 기술 개발을 위해 ECM (Enhanced Compression Model) 참조 소프트웨어를 배포하여 기술 탐색을 수행 중이다. ECM 내 기술 중, 다양한 예측 모드에서 활용되는 DIMD (Decoder side Intra Mode Derivation)는 현재 블록 주변의 복원된 일부 영역을 템플릿으로 정의하고, 템플릿 내 각 서브 블록의 방향성으로 히스토그램을 구성하여 지배적인 방향성에 대응되는 화면내 예측 모드를 현재 블록의 예측 모드로 유도하는 기술이다. 본 논문에서는 DIMD 를 이용하여 효과적으로 화면내 예측 모드를 유도하기 위해 템플릿의 위치에 따라 적응적으로 히스토그램을 구성하는 방법을 제안한다. 제안하는 방법은 ECM10.0 대비 All Intra 환경에서 Class F 영상에 대해 추가적인 계산 복잡도 없이 휘도 및 색차 성분에서 각각 -0.06%, -0.08%, 0.09%의 부호화 효율을 얻는다.

I. 서론

JVET (Joint Video Experts Team)은 ISO/IEC MPEG 과 ITU-T VCEG 표준 기구에 의해 차세대 비디오 압축 표준 기술의 제정을 위해 설립된 단체로서, 2020년 7월 제정된 비디오 압축 표준 기술인 VVC (Versatile Video Coding)[1]는 JVET 에 의해 개발되었다. 이후 VVC 보다 우수한 압축 성능을 갖는 신호처리 기반의 차세대 비디오 압축 표준 기술을 준비하기 위한 논의를 시작하고, 기술 탐색 및 검증을 위해 ECM (Enhanced Compression Model) 참조 소프트웨어를 개발하고 채택된 기술을 통합하여 관리 중이다.

현재 ECM10.0 은 VVC 참조 소프트웨어인 VTM (VVC Test Model) 11.0 대비 AI (All Intra) 환경에서 휘도 성분에 대해 -12.49%의 부호화 효율을 얻는다[2]. ECM 에는 다양한 기술들이 채택되어 VVC 대비 우수한 압축 성능을 보이고 있으며, 대략 150 개 이상의 기술이 포함되어 있다. 그 중, 현재 블록 주변의 복원된 일부 영역을 템플릿으로 정의하고 이를 이용하여 복호화기에서 예측 모드를 유도하거나 예측 신호의 보정 등을 수행함으로써 압축 성능을 높이는 기술이 주를 이루고 있다. 템플릿을 활용하는 대표적인 기술인 DIMD (Decoder side Intra Mode Derivation)는 현재 블록의 화면내 예측 모드를 유도하여 예측을 수행하는 모드이며 다양한 예측 모드 및 변환 기술 등에서도 사용된다. DIMD 는 템플릿 내 각 서브 블록의 방향성으로 히스토그램을 구성하여 지배적인 방향성에 대응되는 화면내 예측 모드를 현재 블록의 예측 모드로 유도한다. 화면내 예측을 수행할 때, 방향성에 따라 사용되는 참조 라인의 범위가 상이하지만 현재 DIMD 는 템플릿의 위치에 관계없이 히스토그램을 구성하고 있다.

본 논문에서는 DIMD 를 이용하여 화면내 예측 모드를 유도할 때, 템플릿의 위치에 따라 적응적으로 히스토그램을 구성함으로써 효율적인 압축을 수행할 수 있는 방법을 제안한다.

II. ECM10.0 의 DIMD

그림 1 은 ECM10.0 의 DIMD 모드의 블록도이다. DIMD 가 수행되는 경우, 그림 2 와 같이 현재 블록 주변의 복원된 영역 중, 3 픽셀 넓이만큼 상단, 좌측 및 좌상단 영역을 템플릿으로 정의한다. 이때, 우상단 및 좌하단 픽셀이 복원되어 사용 가능한 경우, 추가로 4 픽셀 넓이만큼 템플릿으로 구성한다.

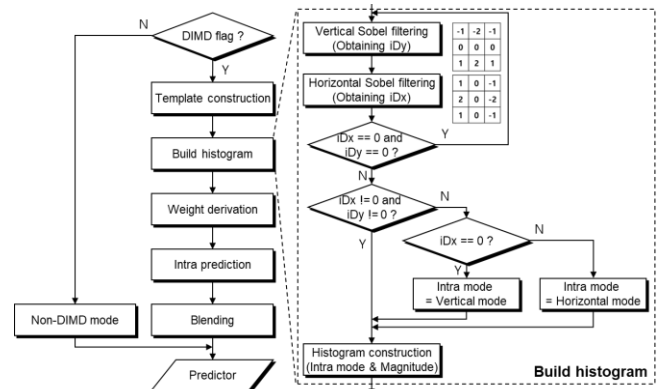


그림 1. DIMD 블록도

이후, 정의된 템플릿에서 화면내 예측 모드 유도를 위해 HoG (Histogram of Gradient) 알고리즘을 이용하여 히스토그램을 구성한다. 템플릿 내에서 각 3x3 블록 단위로 수직 및 수평 Sobel 필터링을 수행하여 각 블록의 그래디

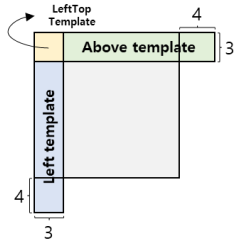


그림 2. 현재 블록의 템플릿

엔트를 얻고, 그래디언트의 크기와 방향에 맵핑되는 화면 내 예측 모드를 히스토그램으로 구성한다. 이때, 수직 방향의 그래디언트가 0 인 경우, 화면내 예측 모드는 수평 모드 (18 번)로 맵핑되며, 수평 방향의 그래디언트가 0 인 경우, 수직 모드 (50 번)로 맵핑된다. DIMD 모드로 예측이 수행될 때, Planar 모드를 포함하여 최대 5 개의 예측 모드로 생성된 예측 신호를 가중합하여 최종 예측 신호를 생성한다. 가중합에 사용되는 각 예측 신호의 가중치는 히스토그램 구성 후에 픽셀 단위로 계산되며, 각 예측 모드가 파생된 템플릿의 위치와 인접한 픽셀의 가중치가 크게 설정되도록 계산을 통해 구한다. 마지막으로, 각 예측 모드로 구한 예측 신호와 가중치를 사용하여 가중합을 통해 최종 예측 신호를 생성한다[3].

III. 제안하는 방법

DIMD 모드의 수행 과정에서 히스토그램을 구성할 때, 수직 및 수평 방향의 그래디언트가 0 인 경우 화면내 예측 모드가 수평 및 수직 모드로 맵핑된다. 하지만, 상단 템플릿 내 참조 라인은 현재 블록이 수직 모드로 예측될 때에만 사용되며, 좌측 템플릿 내 참조 라인은 수평 모드로 예측될 때에만 사용된다. 따라서 상단 템플릿을 이용하여 히스토그램을 구성할 때, 수평 모드의 정보를 히스토그램으로 구성하는 것은 화면내 예측 모드를 유도하는 데에 예측의 효율과는 무관한 예측 모드로 유도될 가능성이 있으며, 이는 좌측 템플릿의 경우도 마찬가지이다.

본 논문에서는 이러한 문제점을 보완하고자 적응적 히스토그램 구성 방법을 제안한다. DIMD 모드로 화면내 예측 모드가 유도될 때, 그림 3 과 같이 히스토그램을 구성하는 과정에서 수직 및 수평 방향의 그래디언트가 0 인 경우, 상단 템플릿이거나 좌측 템플릿인 경우에만 히스토그램을 구성함으로써 효율적으로 히스토그램을 구성할 수 있다.

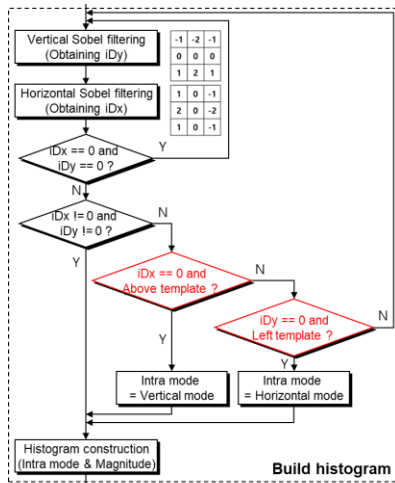


그림 3. 제안하는 적응적 히스토그램 구성 방법

IV. 실험 결과

본 논문에서 제안하는 방법의 성능 평가를 위해 제안하

는 방법을 ECM10.0[4]에 구현하였고, CTC (Common Test Conditions)[5]에 맞춰 AI 환경에서 부호화 효율을 확인하였다. 표 1 의 실험 결과와 같이, 제안하는 방법은 Class F 영상에 대해 휘도 및 색차 성분에서 추가적인 계산 복잡도 없이 각각 -0.06%, -0.08%, 0.09%의 부호화 효율을 얻었다.

표 1. 실험 결과 (BD rate, En/Decoding Time (%))

Sequence	Y	Cb	Cr	EncT/DecT
BasketballDrillText	0.00	-0.04	-0.04	100/100
ArenaOfValor	0.01	-0.04	0.02	100/100
SlideEditing	-0.12	-0.21	0.08	100/100
SlideShow	-0.14	-0.03	0.21	100/100
Average	-0.06	-0.08	0.09	100/100

BasketballDrillText 는 자연 영상에 추가로, 일부 영역에만 컴퓨터에서 생성된 그래픽스 및 텍스트가 혼합되어 있는 영상이고, ArenaOfValor 는 게임 화면 영상이며, SlideEditing, SlideShow 는 영상 전체가 컴퓨터에서 생성된 그래픽스 및 텍스트로만 구성된 영상이다. 따라서, SlideEditing, SlideShow 는 상대적으로 영상 내 평활성이 높기 때문에 제안하는 방법이 적용되는 영역이 더욱 많아, 높은 부호화 효율을 얻은 것으로 분석된다.

V. 결론

본 논문에서는 DIMD 기술의 히스토그램 구성 과정에서 템플릿의 위치에 따라 적응적으로 히스토그램을 구성하는 방법을 제안하였다. 실험 결과, Class F 영상에 대해 ECM10.0 대비 추가적인 계산 복잡도 없이, AI 환경에서 휘도 및 색차 성분에서 각각 -0.06%, -0.08%, 0.09%의 부호화 효율을 얻음을 확인하였다. 특히, 제안하는 방법은 영상 내 평활성이 높은 컴퓨터에서 생성된 그래픽스 및 텍스트로 구성된 영역이 많은 스크린 콘텐츠 영상에 대해 더욱 효과적으로 적용될 수 있을 것으로 사료된다.

ACKNOWLEDGMENT

본 연구는 과학기술정보통신부의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업 (NRF-2021R1A2C2092848) 및 과학기술정보통신부 및 정보통신기획평가원의 대학 ICT 연구센터육성지원사업 (IITP-2023-RS-2022-00156225)의 연구결과로 수행되었음.

참고 문헌

- [1] M. Lee et al., "Overview of Versatile Video Coding (H.266/VVC) and Its Coding Performance Analysis," IEIE Trans. Smart Process. Comput. 2023, 12, 2, 122-154.
- [2] V. Seregin et al., "JVET AHG report: ECM software development (AHG6)," 32nd JVET Meeting, Hannover, Doc. JVET-AF0006, Oct. 2023.
- [3] M. Coban et al., "Algorithm description of Enhanced Compression Model 10 (ECM 10)," 31st JVET Meeting, Geneva, Doc. JVET-AE2025, Jul. 2023.
- [4] ECM10.0, https://vcgit.hhi.fraunhofer.de/ecm/ECM/-/tree/ECM-10.0?ref_type=tags
- [5] M. Karczewicz and Y. Ye, "Common test conditions and evaluation procedures for enhanced compression tool testing," 31st JVET Meeting, Geneva, Doc. JVET-AE2017, Jul. 2023.