

특집논문 (Special Paper)

방송공학회논문지 제25권 제5호, 2020년 9월 (JBE Vol. 25, No. 5, September 2020)

<https://doi.org/10.5909/JBE.2020.25.5.698>

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

허프 변환을 이용한 VVC의 기하학 분할 모드 및 블록 분할 고속 결정 방법

이민훈^{a)}, 박준택^{b)}, 방건^{c)}, 임웅^{c)}, 심동규^{b)}, 오승준^{a)†}

Fast Decision Method of Geometric Partitioning Mode and Block Partitioning Mode using Hough Transform in VVC

Minhun Lee^{a)}, Juntaek Park^{b)}, Gun Bang^{c)}, Woong Lim^{c)}, Donggyu Sim^{b)}, and Seoung-Jun Oh^{a)†}

요약

현재 차세대 부호화 표준으로 진행 중인 VVC (Versatile Video Coding)는 재귀적 블록 분할 구조 및 GPM (Geometric Partitioning Mode)과 같은 다양한 예측 방법들의 채택으로 HEVC (High Efficiency Video Coding) 대비 RA (Random Access) 환경에서 약 34%와 LDB (Low-Delay B) 환경에서 약 30%의 부호화 성능 향상을 보이지만 부호화 복잡도는 약 10배, 7배 증가를 보인다. 본 논문에서는 VVC의 부호화 복잡도 개선을 위하여 블록 내 방향성을 이용한 GPM 모드 고속 결정 및 블록 분할 고속 결정 방법을 제안한다. 제안하는 방법은 현재 블록에 허프 변환을 적용하여 블록 내의 방향성을 파악하고, 이를 통해 울-왜곡 비용 탐색 과정에서 생략할 GPM 모드와 특정 블록 분할 방법을 결정하는 방법이다. 실험 결과로써 제안하는 방법은 VTM8.0 대비 RA 환경에서 2.48%의 부호화 성능 감소와 31.01%의 부호화 시간 감소의 효과를 얻고 LDB 환경에서 2.69%의 부호화 성능 감소와 29.84%의 부호화 시간 감소의 효과를 얻었다.

Abstract

VVC (Versatile Video Coding), which has been developing as a next generation video coding standard. Compared to HEVC (High Efficiency Video Coding), VVC is improved by about 34% in RA (Random Access) configuration and about 30% in LDB (Low-Delay B) configuration by adopting various techniques such as recursive block partitioning structure and GPM (Geometric Partitioning Mode). But the encoding complexity is increased by about 10x and 7x, respectively. In this paper, we propose a fast decision method of GPM mode and block partitioning using directionality of block to reduce encoding complexity of VVC. The proposed method is to apply the Hough transform to the current block to identify the directionality of the block, thereby determining the GPM mode and the specific block partitioning method to be skipped in the rate-distortion cost search process. As a result, compared to VTM8.0, the proposed method reduces about 31.01% and 29.84% encoding complexity for RA and LDB configuration with 2.48% and 2.69% BD-rate loss, respectively.

Keyword : VVC, Hough transform, Fast algorithm, Block partitioning, Geometric Partitioning Mode

I. 서 론

최근 실시간 스트리밍 서비스의 확대와 함께 4K UHD (Ultra High Definition), 8K UHD의 고화질 비디오와 기존의 2차원 비디오를 넘어선 360도 비디오를 비롯한 3차원 비디오 콘텐츠에 대한 사용자의 요구가 꾸준히 증가하고 있다. 이에 따라 기존 비디오 압축 표준인 HEVC (High Efficiency Video Coding)^[1]보다 높은 성능을 가지는 비디오 압축 기술을 필요로 하게 되었다. 이를 위해 ISO/IEC MPEG과 ITU-T VCEG은 JVET (Joint Video Experts Team)을 구성하여 360 비디오와 HDR (High Dynamic Range) 비디오 등을 고려하여 HEVC보다 두 배 이상의 압축 효율을 목표로 하는 VVC (Versatile Video Coding)^[2] 표준화를 진행하고 있다. 그리고 JVET은 매 회의의 결과로 VVC 표준화를 위한 공통 실험을 위하여 참조 소프트웨어인 VTM (VVC Test Model)^[3]을 배포하고 있다. 현재 VTM8.0 버전이 배포되었고 HEVC대비 RA (Random Access) 환경에서 약 34%와 LDB (Low-Delay B) 환경에서 약 30%의 부호화 성능 향상을 보이며 이에 따른 부호화 복잡도는 약 10배, 7배 증가를 보이고 있다^[4]. VVC는 HEVC보다 다양한 기술이 채택되었고, 특히 더욱 유연해진 블록 분할 구조와 GPM (Geometric Partitioning Mode)과 같은 다양한 예측 모드의 증가로 인해 높은 성능 향상을 보인다.

HEVC에서 하나의 슬라이스 (Slice)는 다수개의 CTU (Coding Tree Unit)으로 분할되고, 하나의 CTU는 쿼드 트리 (Quad-Tree, QT)로 재귀적으로 분할되며 분할의 말단

a) 광운대학교 전자공학과(Dept. of Electronic Engineering, Kwangwoon University)

b) 광운대학교 컴퓨터공학과(Dept. of Computer Engineering, Kwangwoon University)

c) 한국전자통신연구원(Electronics and Telecommunications Research Institute)

‡ Corresponding Author : 오승준(Seoung-Jun Oh)

E-mail: sjoh@kw.ac.kr

Tel: +82-2-940-5102

ORCID: <https://orcid.org/0000-0002-5036-376>

* 이 논문은 2020년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원 (No. 2018-0-00207, [전문연구실] 이머시브 미디어 전문연구실)의 지원 및 2020년도 광운대학교 연구년에 의하여 연구되었음.

※ This work was supported by the Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No. 2018-0-00207, Immersive Media Research Laboratory) and conducted during the sabbatical year of Kwangwoon University in 2020.

· Manuscript received July 13; Revised August 31, 2020; Accepted September 9, 2020.

노드를 CU (Coding Unit)이라 하였다. 이후 CU는 예측 단위인 PU (Prediction Unit)와 변환 단위인 TU (Transform Unit)로 분할되어 예측과 변환 단위로 수행되었다. VVC에서는 HEVC와 다르게 정사각형에 형태와 함께 직사각형의 CU 형태도 사용하는 보다 유연한 분할을 사용하고, 이를 위해 각 CTU를 QT에 더불어 BT (Binary-Tree)와 TT (Ternary-Tree)로 구성된 MTT (Multi-Type Tree)로 재귀적으로 분할되는 구조를 사용한다. 또한 CU, PU, TU의 개념을 단일화하여 예측 및 변환 부호화를 수행한다. 그림 1은 VVC의 재귀적 블록 분할 구조를 나타낸다.

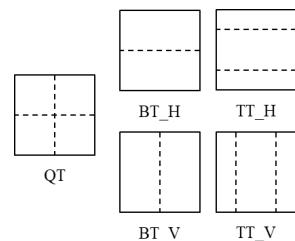


그림 1. VVC의 재귀적 블록 분할 구조

Fig. 1. The recursive structure of block partitioning in VVC

이렇듯 VVC에서는 MTT 구조의 도입으로 직사각형의 CU를 이용해 효과적인 예측을 수행할 수 있었지만, 실제 영상 내부의 물체 경계 부분은 다양한 각도로 존재하기 때문에 이를 보다 정밀하게 예측하기 위해서는 더 깊은 블록 분할이 수행되어야 하는 한계점이 존재하여 더욱 유연한 블록 분할의 필요성이 제기되었다. 이를 위해 화면간 예측 수행 시 사각형 형태의 예측자만을 만들어 수행하는 HEVC 와 달리, 사각형 형태에 더불어 사다리꼴 및 삼각형 형태의 예측자를 이용해 예측을 수행하는 GPM이 2020년 1월 회의에서 채택이 되어 VVC에 포함되었다^[5]. 그림 2 (a)는

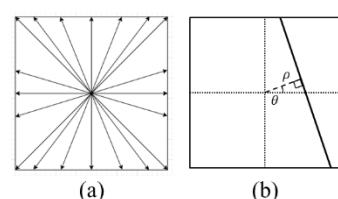


그림 2. GPM 모드의 각도 분포와 예시 (a) GPM 모드의 각도 분포; (b) GPM 모드의 예시

Fig. 2. Angles distribution and example of GPM mode (a) Angles distribution of each GPM modes; (b) Example of GPM mode

GPM 모드를 결정하는 20개의 각도이다.

GPM은 그림 2 (a)의 20개의 각도와 블록의 크기에 따른 4개의 거리를 이용하여 구성된 64개의 모드 중 그림 2 (b)의 예시와 같이 한 가지를 선택하여 CU를 두 가지 영역으로 분할하고 분할된 두 개의 영역에 대하여 서로 다른 움직임 정보로부터 두 개의 예측자를 만든 후에 분할된 영역의 모양에 따른 마스크 (Mask)를 이용하여 두 개의 예측자를 가중 합하여 하나의 예측자를 생성함으로써 하나의 CU를 서로 다른 사다리꼴 모양으로 각각 화면간 예측을 수행하는 것과 같은 효과를 얻을 수 있는 예측 기술이다. VVC는 이와 같은 다양한 CU의 재귀적 분할과 GPM을 포함한 여러 가지 화면간 예측 기술에 의해 부호화 성능이 크게 증가하였다. 하지만, 재귀적 분할로 인해 생성되는 다양한 CU에 대하여 여러 가지 예측 모드 중 최적의 모드를 결정하기 위한 복잡한 율-왜곡 비용 탐색 과정, 특히 GPM의 경우 64개의 모드에 대하여 모든 움직임 벡터 후보군의 조합에 대한 율-왜곡 비용 탐색 과정을 거친으로써 부호화 복잡도가 상당히 증가하게 되었다. 이러한 문제를 해결하기 위해 다양한 비디오 압축 기술의 부호화 복잡도 감소 기술들이 꾸준하게 연구되고 있다^{[6][7][8][9][10]}. Shen^[6]은 주변 블록의 깊이 정보를 이용하여 현재 블록의 분할 깊이의 화면내 예측 과정을 생략하는 방법을 제안하였다. 그러나 주변 블록의 정보를 이용하기 때문에 주변 블록에 대한 의존성이 존재할 수 있다. Li^[9]은 VVC의 재귀적 분할을 빠르게 결정하기 위하여 딥러닝 기반의 블록 분할 고속 결정 방법을 제안하였다. 높은 정확성으로 블록 구조를 결정할 수 있지만, 학습에 의존적일 수 있다.

본 논문에서는 현재 블록 내 방향성을 파악하여 GPM 모드 결정과 블록 분할을 빠르게 하는 고속 결정 방법을 제안한다. 현재 블록 내의 방향성을 파악하기 위해 보편적으로 직선 검출에 사용되는 알고리즘인 허프 변환 (Hough transform)^[11]을 이용한다.

본 논문의 구성은 다음과 같다. 2장에서는 허프 변환과 제안하는 고속 결정 방법에 대해 설명한다. 그리고 3장에서는 제안하는 방법의 성능을 평가하고 4장에서 결론을 맺는다.

II. 고속 결정 방법

1. 허프 변환

허프 변환이란 일반적으로 2차원 영상에서 빠르게 직선을 검출할 수 있는 알고리즘이다. 그림 3은 허프 변환의 예시이다. 그림 3 (a)를 보면, (x, y) 좌표계에서 임의의 한 점 (a, b) 를 지나는 직선의 방정식은 식 (1)과 같이 ρ 와 θ 로 표현할 수 있고, 이때 θ 의 범위는 (x, y) 좌표계의 모든 기울기의 직선을 표현하기 위해 $\theta \in [0, \pi]$ 을 사용한다. 식 (1)로부터 얻은 ρ 와 θ 를 (ρ, θ) 좌표계에 나타내면 그림 3 (b)와 같이 하나의 곡선을 얻을 수 있다. 즉, 이 곡선은 (x, y) 좌표계에서 점 (a, b) 를 지나는 모든 직선을 의미한다. 이러한 관계를 이용하여 (x, y) 좌표계의 점들을 (ρ, θ) 좌표계로 변환 후, 일정 빈도수 이상 누적된 교점을 $(\rho_{\text{accumulate}}, \theta_{\text{accumulate}})$ 라고 할 때 이를 식 (2)와 같이 (x, y) 좌표계로 다시 변환하여 해당 영역의 방향성을 나타내는 직선을 얻을 수 있다.

$$x \cos \theta + y \sin \theta = \rho \quad (1)$$

$$y(x) = -x \cot \theta_{\text{accumulate}} + \frac{\rho_{\text{accumulate}}}{\sin \theta_{\text{accumulate}}} \quad (2)$$

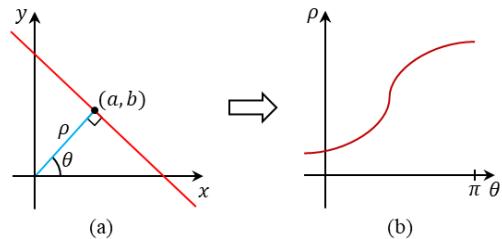


그림 3. 허프 변환의 예시 (a) (x, y) 좌표계; (b) (ρ, θ) 좌표계
Fig. 3. The example of hough transform (a) Coordinate system of (x, y) ; (b) Coordinate system of (ρ, θ)

본 논문에서는 상기 언급한 허프 변환을 각 CU에 적용하여 검출된 직선으로부터 블록 내 방향성을 파악한다. 그리고 블록 내 중요한 특성과 성질을 보존하면서 직선 검출에 필요하지 않은 데이터양을 줄이기 위한 방법으로 블록의 에지 맵 (Edge map)을 추출한 후에 허프 변환을 적용한다. 이때 사용되는 에지 맵은 보편적으로 사용되는 알고리즘인 잡음에 강하고 정확한 에지를 추출하는 캐니 에지 (Canny edge) 방법^[12]을 이용하여 추출한 결과이다. 이때 필요한 두

임계값은 효과적인 에지 맵 추출을 위하여 Fang^[13]이 제안한 방법을 사용한다. 이는 두 임계값을 영상의 효과적인 이진화를 위한 방법인 Otsu 이진화^[14]방법을 현재 CU에 적용하여 이를 통해 구한 임계값을 통해 식 (3), (4)와 같이 정하여 사용하는 방법이다.

$$\text{High threshold} = \text{Otsu's threshold value} \quad (3)$$

$$\text{Low threshold} = \text{Otsu's threshold value} \gg 1 \quad (4)$$

그리고 허프 변환을 적용할 블록의 크기는 CU의 크기로써 서로 다른 크기를 가지므로 누적 빈도수 임계 값을 고정된 값으로 사용할 경우, 큰 값이 사용되면 상대적으로 크기가 작은 CU에 대하여 방향성이 있더라도 직선이 검출되지 않을 수 있고 작은 값이 사용되면 상대적으로 크기가 큰 CU에 대하여 잡음에 의한 직선이 검출될 수 있다. 따라서 사용될 누적 빈도수 임계 값을 Guil^[15]이 제안한 방법에 따라 크기의 CU에 대하여 식 (5)와 같이 적응적으로 정하여 사용한다.

$$((\log_2 M + \log_2 N) \gg 1) + 1 \quad (5)$$

그림 4는 상기 언급한 방법을 통하여 캐니 에지 기법과 허프 변환을 수행하는 예시로써 휴도 성분의 32×32 크기인 CU에 대해 적용하는 과정을 나타낸다.

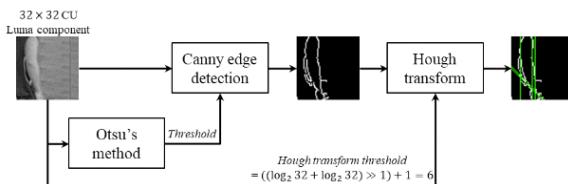


그림 4. CU에 대하여 캐니 에지 기법과 허프 변환을 적용하는 예시
Fig. 4. The example of applying the canny edge detection and hough transform to CU

2. GPM 모드 결정 방법

VVC의 GPM은 현재 블록을 미리 정의된 20개의 각도와 블록 크기에 따른 4개의 거리에 따른 모드에 의해 두 개의 블록으로 분할하여 각각 독립적으로 서로 다른 움직임 정

보를 이용하여 예측자를 만들고, 가중합을 이용하여 예측을 수행하는 방법이다. 그림 5는 VTM8.0을 이용하여 QP (Quantization Parameter)=22에서 LDB 환경으로 부호화된 BQmall 시퀀스의 7번째 프레임이고, 파란색 블록은 GPM이 선택된 CU이다. GPM은 특히 그림 5와 같이 영상 내 물체의 경계에서 많은 선택이 되고, 선택되는 모드는 경계와 유사한 직선 형태를 보이는 경향을 보인다. 그러나 현재 VTM8.0의 부호화 과정에서 이를 고려하지 않고 그림 6과 같이 64개의 모드에 대하여 각 모드마다 모든 움직임 벤더 후보군 조합에 대한 율-왜곡 비용 탐색 과정을 거쳐 최종적으로 GPM 모드를 결정한다.

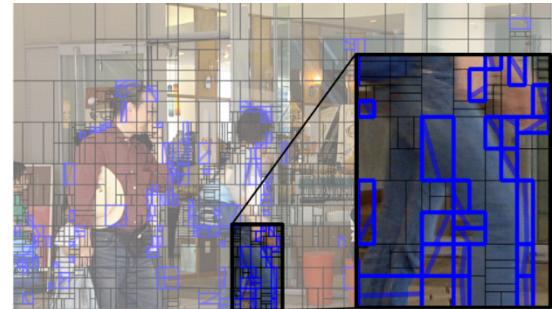


그림 5. VTM8.0에서 QP=22, LDB 환경으로 부호화된 BQmall 시퀀스의 7번째 프레임

Fig. 5. The 7th frame of the BQmall sequence encoded as QP=22, LDB configuration in VTM 8.0

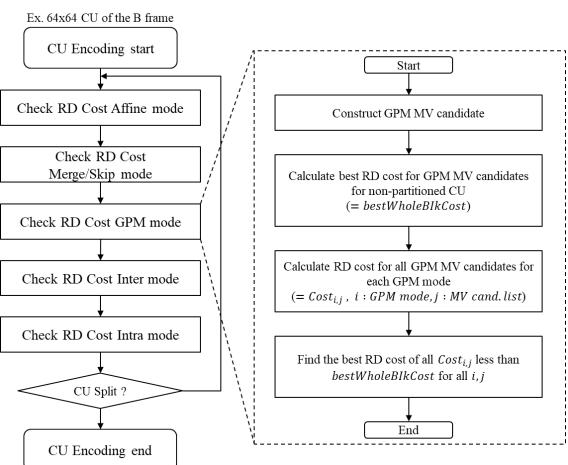


그림 6. VTM8.0에서 GPM의 율-왜곡 비용 탐색 과정

Fig. 6. The process of rate-distortion cost search process for GPM in VTM8.0

본 논문에서 제안하는 방법은 현재 블록에 허프 변환을 적용하여 검출된 직선의 정보로부터 현재 블록 내 방향성을 얻고, 이를 통해 율-왜곡 비용 탐색 과정에서 생략할 GPM 모드를 결정하는 방법이다. 하나의 CU에 존재하는 방향성을 수평, 수직, 두 개의 주 대각, 그리고 무방향성으로 분류할 수 있다^[16]. 이를 토대로 현재 CU에 방향성을 얻기 위한 허프 변환은 수평, 수직 두 개의 주 대각에 대하여 검출하고, 직선이 검출되지 않는 경우는 현재 CU가 무방향성을 갖는다고 분류하였다. 그리고 검출된 직선의 각도를 통해 각 GPM 모드가 갖는 20개의 각도를 그림 7과 같이 분류하였다. 이를 바탕으로 현재 블록에 허프 변환을 수행한 뒤, 검출된 직선의 각도에 해당하는 GPM 모드에 대해서만 율-왜곡 비용 탐색 과정에 사용함으로써 부호화 시간을 줄일 수 있다. 예를 들어, 현재 블록 내에 수평 방향의 직선이 존재하지 않는 경우, 현재 블록의 수평성이 약하다고 볼 수 있으므로 GPM 모드 중 수평 방향의 모드를 생략할 수 있고 직선이 검출되지 않는 경우, GPM 모드의 율-왜곡 비용 탐색 과정을 생략할 수 있다. 그리고 4가지 방향에 대한 직선이 모두 검출된 경우, 현재 CU가 네 가지 방향성 모두를 갖는다고 할 수 있고 모든 GPM 모드에 대하여 율-왜곡 비용 탐색 과정을 수행한다. 표 1은 제안하는 방법의 검출되는 직선에 따라 율-왜곡 비용 탐색 과정을 수행할 GPM 모드와 수행할 GPM 모드의 개수를 나타낸 것이다.

표 1. 제안하는 방법의 각 경우에 따라 선택되는 GPM 모드와 개수

Table 1. The number and mode of selected GPM mode for each case of proposed method

Case	Selected GPM mode	The number of selected GPM modes
No line	Skip GPM	0
Only vertical	0,1,2,3,4,5,32,33,34,35,36,37,38,39,40,61,62,63	18
Only horizontal	14,15,16,17,18,19,20,21,22,23,47,48,49,50,51,52,53,54	18
Only 45°	20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,52,53,54,55,56,57,58,59,60,61,62,63	28
Only 135°	2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,38,39,40,41,42,43,44,45,46,47,48,49	28
No 45° & 135°	0,1,2,3,4,5,14,15,16,17,18,19,20,21,22,23,32,33,34,35,36,37,38,39,40,47,48,49,50,51,52,53,54,61,62,63	36
No vertical & 45°	2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54	39
No vertical & 135°	14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,47,48,49,50,51,52,53,54,55,56,57,58,59,60,61,62,63	39
No horizontal & 45°	0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54,61,62,63	39
No horizontal & 135°	0,1,2,3,4,5,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,52,53,54,55,56,57,58,59,60,61,62,63	39
No diagonal (45°)	0,1,2,3,4,5,14,15,16,17,18,19,20,21,22,23,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,57,58,59,60,61,62,63	50
No diagonal (135°)	0,1,2,3,4,5,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,47,48,49,50,51,52,53,54,55,56,57,58,59,60,61,62,63	50
No horizontal & vertical	2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,57,58,59,60,61,62,63	56
No vertical	2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,57,58,59,60,61,62,63	60
No horizontal	0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,57,58,59,60,61,62,63	60
All line	All modes	64

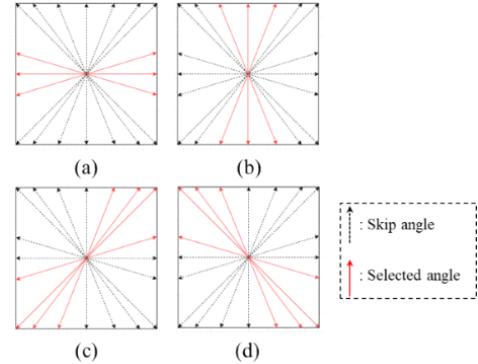


그림 7. 각 방향에 따른 GPM 모드의 각도의 양자화 (a) 수평; (b) 수직; (c) 45°; (d) 135°

Fig. 7. Quantization of the angle of GPM mode for each direction (a) horizontal; (b) vertical; (c) 45°; (d) 135°

3. 블록 분할 결정 방법

VVC의 블록 분할 결정 방법으로는 QT, BT_H (Binary Tree_Horizontal), BT_V (Binary Tree_Vertical), TT_H (Ternary Tree_Horizontal), TT_V (Ternary Tree_Vertical)가 있고, 상기 다섯 가지 방법들을 통해 CU를 분할한다. 현재 VTM8.0에서는 분할 가능한 각 CU에 대해 최소한의 율-왜곡 비용을 갖는 CU를 선택하기 위해 율-왜곡 비용 탐색 과정을 수행하므로 이러한 과정에서 상당한 복잡도를 야기한다. 본 논문에서는 이러한 복잡도를 줄이기 위해 현

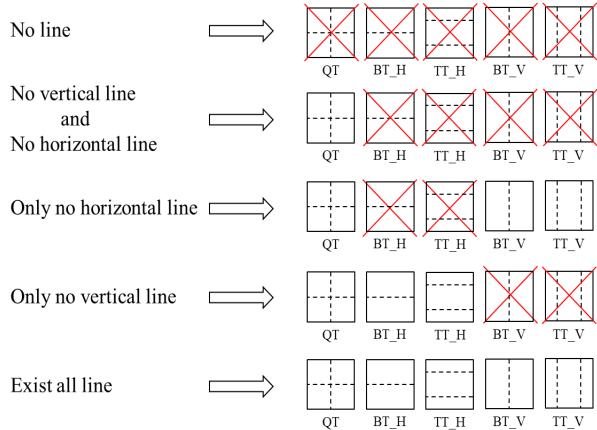


그림 8. 각 조건에 따른 현재 블록의 율-왜곡 비용 탐색 과정에서 생략할 블록 분할 방법

Fig. 8. Block split method to be skipped during rate-distortion cost search process of the current block under each condition

재 블록 내 방향성에 따라 블록 분할 방법을 결정한다. 제안하는 블록 분할 결정 방법은 현재 블록의 방향성을 상기 언급한 다섯 가지로 분류하고, 그에 따라 현재 블록에 허프 변환을 수행한 결과로 수평, 수직과 두 개의 주 대각 방향의 직선을 얻고, 검출된 직선의 정보를 통해 현재 블록이 갖는 방향성을 파악함으로써 생략할 블록 분할 방법을 결정하는 방법이다. 하나의 예시로써, 허프 변환의 결과로 현재 블록에 수직 방향의 직선이 존재하지 않는 경우, 현재 블록의 율-왜곡 비용 탐색 과정에서 BT_V와 TT_V의 생략을 결정할 수 있다. 그림 8과 그림 9는 각 조건에 따른 현재 블록의

율-왜곡 비용 탐색 과정에서 생략할 블록 분할 방법과 본 논문에서 제안하는 방법의 순서도를 나타낸다.

III. 실험 환경 및 결과

본 논문에서 제안하는 방법의 성능 평가를 위하여 VTM 8.0에서 JVET-CTC (Common Test Condition)^[17]의 RA 환경과 LDB 환경에서 실험을 진행하였다. 성능의 비교는 부호화 복잡도와 BD-rate^[18]을 이용하였고 성능 평가에 사용한 테스트 시퀀스는 클래스 B, C, D 영상을 이용하였고, 모든 실험은 각 영상의 전체 프레임에 대하여 실험을 진행하였다. 그리고 감소한 부호화 시간의 정도 (T_{Red})는 식 (6)과 같이 측정하였다. 이때 T_{Anc} 은 기존 방법의 부호화 복잡도, T_{prop} 는 제안하는 방법의 부호화 복잡도를 의미한다.

$$T_{Red} = \frac{T_{Anc} - T_{prop}}{T_{Anc}} \times 100(\%) \quad (6)$$

표 2는 제안하는 방법의 VTM8.0대비 성능을 보여준다. 제안하는 방법은 RA 환경에서 평균적으로 31.01%의 부호화 복잡도 감소를 보이고 2.48%의 부호화 성능 감소를 보인다. LDB 환경에서는 평균적으로 29.84%의 부호화 복잡도 감소와 2.69%의 부호화 성능 감소를 보인다.

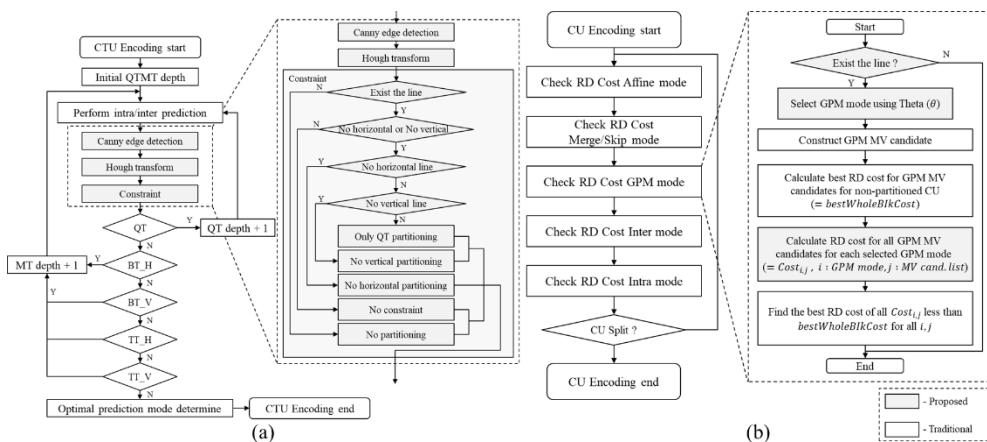


그림 9. 제안하는 방법의 순서도 (a) 블록 분할 결정 방법; (b) GPM 모드 결정 방법

Fig. 9. The flow-chart of the proposed method (a) The method of decision of block partitioning; (b) The method of decision of GPM mode

표 2. 제안하는 방법의 VTM8.0 대비 성능

Table 2. Performance of the proposed method over VTM8.0

Class	Sequence	Random Access		Low-Delay B	
		BD-rate Y	T_{Red}	BD-rate Y	T_{Red}
Class B (1920 × 1080)	RitualDance	2.82%	30.23%	-	-
	Cactus	2.30%	31.68%	2.33%	32.59%
	BasketballDrive	1.91%	29.98%	-	-
	BQTerrace	2.42%	32.54%	-	-
Class C (832 × 480)	RaceHorsesC	3.19%	35.73%	4.05%	34.81%
	BQMall	3.08%	36.15%	3.32%	29.05%
	PartyScene	2.52%	41.02%	1.86%	29.28%
	BasketballDrill	3.60%	36.46%	4.84%	36.39%
Class D (416 × 240)	RaceHorses	3.96%	32.88%	3.60%	28.24%
	BQSquare	1.12%	21.60%	2.08%	21.11%
	BlowingBubbles	1.28%	25.08%	1.61%	20.42%
	BasketballPass	1.56%	18.75%	1.63%	28.42%
Average	Class B	2.36%	31.11%	2.33%	32.59%
	Class C	3.10%	37.34%	3.52%	32.38%
	Class D	1.91%	24.58%	2.23%	24.55%
	Overall	2.48%	31.01%	2.69%	29.84%

표 3. 제안하는 방법으로 인해 변화한 GPM 부호화 복잡도와 고속 결정 과정에 필요한 추가 계산 복잡도의 비율

Table 3. The ratio of the GPM encoding complexity by proposed method and the additional complexity required for fast decision process

Sequence	Random Access		Low-Delay B	
	Reduced GPM_encT_rate	Additional complexity	Reduced GPM_encT_rate	Additional complexity
RaceHorses	29.11%	39.04%	32.99%	41.02%
BQSquare	18.20%	36.23%	15.47%	33.98%
BlowingBubbles	20.24%	43.61%	25.91%	40.99%
BasketballPass	21.58%	37.01%	27.56%	33.31%
Average	22.28%	38.97%	25.48%	37.33%

표 3은 Class D 영상에서 기존 방법의 GPM 부호화 복잡도 대비 제안하는 방법의 GPM 모드 부호화 복잡도의 비율 (Reduced GPM_encT_rate)과 제안하는 방법을 적용하여 GPM 모드를 결정하기 위한 부호화 시간 중 고속 결정을 위한 추가 계산 복잡도의 비율 (Additional complexity)을 보여준다. RA 환경에서는 평균적으로 22.28%의 GPM 모드 부호화 복잡도 감소가 있었으며, 이 중 평균적으로 38.97%는 제안하는 방법으로 인한 추가 계산 복잡도이며, LDB 환경에서는 평균적으로 25.48%의 GPM 모드 부호화 복잡도 감소가 있었고, 이 중 평균적으로 37.33%는 제안하는 방법으로 인한 추가 계산 복잡도이다.

그림 10은 LDB 환경에서 BQSquare 시퀀스와 RA 환경

에서 RaceHorses 시퀀스의 VTM8.0과 제안하는 방법의 RD 곡선 (Rate-Distortion Curve)를 나타낸 것이다. 그림 10을 보면, 두 그래프가 매우 근접한 것을 확인할 수 있다. 이는 제안하는 방법의 부호화 복잡도 감소로 인한 부호화 성능 감소가 무시할 수 있는 수준이라고 볼 수 있고, 낮은 비트율과 높은 비트율에서 모두 효과적이라는 것을 보여준다. 또한 RD 곡선은 그림 10의 두 시퀀스를 제외한 다른 시퀀스에서도 비슷한 경향을 보인다.

그림 11과 그림 12는 제안하는 방법의 주관적 화질 비교를 위한 그림이다. 비교를 위한 그림은 각 환경에서 부호화 성능 감소의 정도가 가장 큰 LDB 환경으로 부호화된 BasketballDrill과 RA 환경으로 부호화된 RaceHorses 영상

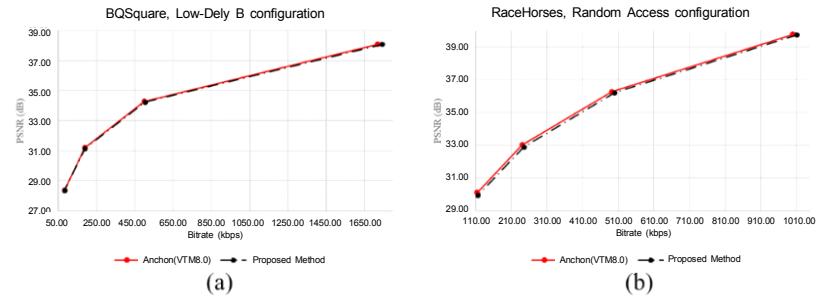


그림 10. LDB 환경에서 BQSquare와 RA 환경에서 RaceHorses 시퀀스의 VTM8.0과 제안하는 방법의 RD 곡선

Fig. 10. RD curves of VTM8.0 and the proposed method, BQSquare for LDB configuration and RaceHorses for RA configuration

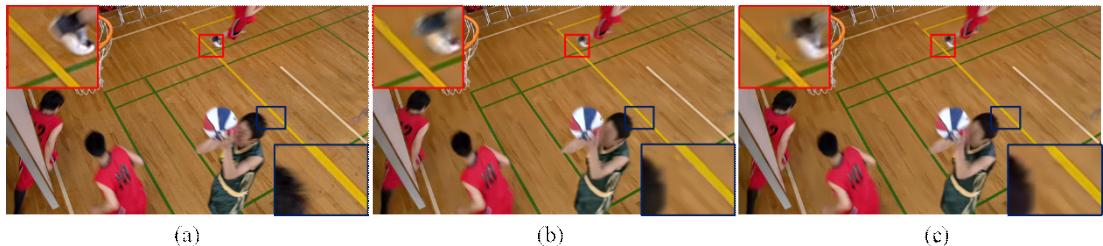


그림 11. LDB 환경, QP=37으로 부호화된 BasketballDrill 시퀀스의 93번째 프레임에서 주관적 화질의 비교 (a) 원본; (b) 기존 방법; (c) 제안하는 방법

Fig. 11. Visual quality comparison for BasketballDrill with LDB configuration, where the 93th frame is shown (QP=37), (a) Original; (b) Existing method

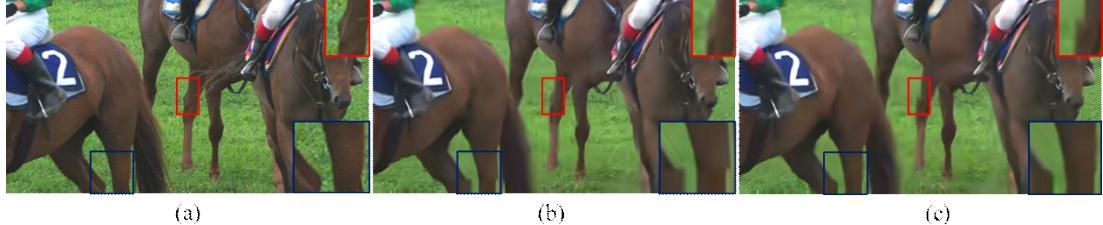


그림 12. RA 환경, QP=37으로 부호화된 RaceHorses 시퀀스의 100번째 프레임에서 주관적 화질의 비교 (a) 원본; (b) 기존 방법; (c) 제안하는 방법

Fig. 12. Visual quality comparison for RaceHorses with RA configuration, where the 100th frame is shown (QP=37), (a) Original; (b) Existing method;

(c) Proposed method

의 한 프레임이 사용되었고, 각각 원본과 기존 방법, 제안하는 방법에 대한 그림이다. 또한 효과적인 주관적 화질 비교를 위해 비교하는 부분을 확대하여 표시하였다. 적색으로 표시된 영역은 기존 방법이 제안하는 방법보다 더 좋은 화질을 가진 영역이고, 남색으로 표시된 영역은 제안하는 방법이 기존 방법보다 더 좋은 화질을 가진 영역이다. 제안하는 방법으로 인해 부호화 성능이 감소하지만, 부호화 성능 감소가 큰 영상에 대하여 주관적 화질 비교를 진행하였을 때 기존 방법보다 더 좋은 화질을 갖는 영역이 존재하는

것을 확인하였다. 또한 낮은 QP값을 사용하여 부호화된 영상에서는 주관적 화질의 차이가 미미하게 존재하였고 높은 QP값을 사용하여 부호화된 영상에서 그림 11과 그림 12와 같은 주관적 화질 차이가 보이는 것을 확인하였다.

IV. 결 론

본 논문에서는 VVC의 부호화 복잡도 개선을 위하여 블

록 내 방향성을 이용한 GPM 모드 고속 결정 및 블록 분할 고속 결정 방법을 제안하였다. 제안하는 방법은 현재 블록에 허프 변환을 적용하여 블록 내의 방향성을 파악하고, 이를 통해 율-왜곡 비용 탐색 과정에서 생략할 GPM 모드의 선택과 특정 블록 분할 방법을 생략함으로써 부호화 복잡도를 감소시키는 방법이다. 제안하는 방법의 성능 평가를 위하여 JVET-CTC의 RA 환경과 LDB 환경에서 VTM8.0을 사용하여 구현하고 제안하는 방법과 기존 방법을 비교하였다. 실험 결과로써 제안하는 방법은 RA 환경과 LDB 환경에서 각각 평균적으로 31.01%와 29.84%의 부호화 복잡도 감소와 2.48%와 2.69%의 부호화 성능 감소를 보였다. 또한, VTM8.0과 제안하는 방법에 대한 RD 곡선 확인과 주관적 화질을 비교함으로써 제안하는 방법의 부호화 복잡도 감소로 인한 부호화 성능 감소가 무시할 수 있는 수준임을 확인하였다. 본 논문에서 제안하는 방법을 바탕으로 부호화 복잡도 감소율을 보다 증가시키면서 부호화 성능 감소를 최소화할 수 있는 추가적인 복잡도 개선 연구가 필요하다.

참 고 문 헌 (References)

- [1] G. J. Sullivan, J. R. Ohm, W. J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," IEEE Transactions on circuits and systems for video technology, Vol.22, No.12, pp.1649-1668, Dec. 2012.
- [2] B. Bross, J. Chen, S. Liu, and Y. K. Wang, JVET-O2001, "Versatile Video Coding (Draft 8)," Jan. 2020.
- [3] VTM, https://vegit.hhi.fraunhofer.de/jvet/VVCSoftware_VTM
- [4] F. Bossen, X. Li, and K. Suehring, JVET-R0003, "AHG report: Test model software development (AHG3)," Apr. 2020.
- [5] H. Gao, S. Esenlik, E. Alshina, A. M. Kotra, B. Wang, R. L. Liao, J. Chen, Y. Ye, J. Luo, K. Reuze, C. C. Chen, H. Huang, W. J. Chien, V. Seregin, Z. Deng, L. Zhang, H. Liu, K. Zhang, Y. Wang, J. Li, C. S. Lim, Y. L. Hsiao, C. C. Chen, C. W. Hsu, Y. W. Huang, S. M. Lei, L. F. Chen, X. Li, C. Li, S. Liu, L. P. Van, G. V. Auwera, A. K. Ramasubramonian, H. Huang, W. J. Chien, M. Karczewicz, M. Blaser, J. Sauer, H. Chen, and H. Yang, JVET-Q0806, "Integrated Text for GEO," Jan. 2020.
- [6] L. Shen, Z. Zhang, and P. An, "Fast CU size decision and mode decision algorithm for HEVC intra coding," IEEE Transactions on consumer Electronics, Vol.59, No.1, pp.207-213, Feb. 2013.
- [7] S. K. Na, W. J. Lee, and K. W. Yoo, "Edge-based fast mode decision algorithm for intra prediction in HEVC," IEEE International Conference on Consumer Electronics, pp.11-14, Jan. 2014.
- [8] S. H. Park, and J. W. Kang, "Context-Based Ternary Tree Decision Method in Versatile Video Coding for Fast Intra Coding," IEEE Access, Vol.7, pp.172597-172605, Nov. 2019.
- [9] T. Li, M. Xu, and R. Tang, "DeepQTMT: A Deep Learning Approach for Fast QTMT-based CU Partition of Intra-mode VVC," arXiv preprint arXiv:2006.13125. 2020.
- [10] Y. U. Yoon, D. H. Park, and J. G. Kim, "Gradient-Based Methods of Fast Intra Mode Decision and Block Partitioning in VVC," Journal of Broadcast Engineering, Vol.25, No.3, pp.338-345, May. 2020.
- [11] P. V. C. Hough, "Method and means for recognizing complex patterns," US Patent 3,069,654, Patent and Trademark Office, Washington D.C., 1962.
- [12] J. F. Canny, "A Computational Approach to Edge Detection," IEEE Transactions Pattern Analysis and Machine Intelligence, Vol.8, No.6, pp.679-698, Nov. 1986.
- [13] M. Fang, G. Yue, and Q. Yu, "The Study on An Application of Otsu Method in Canny Operator," In Proceedings. The 2009 International Symposium on Information Processing, pp.109-112, Aug. 2009.
- [14] N. Otsu, "A threshold selection method from gray-level histograms," IEEE Transactions on systems, man, and cybernetics, Vol.9, No.1, pp.62-66, Jan. 1979.
- [15] N. Guil, J. Villalba, and E. L. Zapata, "A fast Hough transform for segment detection," IEEE Transaction on Image Processing, Vol.4, No.11, pp.1541-1548, Nov. 1995.
- [16] C. S. Won, D. K. Park, and S. J. Park, "Efficient Use of MPEG-7 Edge Histogram Descriptor," ETRI Journal, Vol.24, No.1, pp.23-30, Feb. 2002.
- [17] F. Bossen, J. Boyce, K. Suehring, X. Li, and V. Seregin, JVET-N1010, "JVET common test conditions and software reference configurations for SDR video," Mar. 2019.
- [18] G. Bjøntegaard, VCEG-M33, "Calculation of average PSNR differences between RD-curves," Apr. 2014.

저자 소개

이민훈



- 2019년 2월 : 광운대학교 수학과 학사
- 2019년 3월 ~ 현재 : 광운대학교 전자공학과 석사과정
- ORCID : <https://orcid.org/0000-0001-8165-5380>
- 주관심분야 : 영상신호처리, 영상압축, 컴퓨터비전

박준택



- 2017년 2월 : 광운대학교 컴퓨터공학과 학사
- 2019년 2월 : 광운대학교 컴퓨터공학과 석사
- 2019년 3월 ~ 현재 : 광운대학교 컴퓨터공학과 박사과정
- ORCID : <https://orcid.org/0000-0001-5647-7260>
- 주관심분야 : 영상신호처리, 영상압축, 컴퓨터비전

방건



- 2014년 2월 : 고려대학교 컴퓨터학과 박사
- 2008년 ~ 2013년 : MPEG FTV AhG 그룹 EE 코디네이터
- 2011년 9월 ~ 2012년 10월 : MIT RLE ATSP 그룹 방문 연구원
- 2002년 10월 ~ 2006년 2월 : ATSC T3/S2 ACAP 데이터방송 표준화 에디터
- 2000년 4월 ~ 현재 : 한국전자통신연구원 미디어부호화연구실 책임연구원
- ORCID : <https://orcid.org/0000-0003-4355-599X>
- 주관심분야 : 2D/3D 비디오 부호화, 영상처리, 컴퓨터비전, 인공지능

임웅



- 2008년 : 광운대학교 컴퓨터공학과 학사
- 2008년 3월 ~ 2010년 2월 : 광운대학교 컴퓨터공학과 석사
- 2010년 3월 ~ 2016년 : 광운대학교 컴퓨터공학과 박사
- 2012년 2월 ~ 2013년 : Simon Fraser Univ. 방문연구원
- 2016년 4월 ~ 현재 : 한국전자통신연구원 미디어부호화연구실 선임연구원
- ORCID : <https://orcid.org/0000-0002-1772-0683>
- 주관심분야 : 영상처리, 비디오 압축, 기계 학습, 컴퓨터 비전

저자소개



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



오승준

- 1980년 2월 : 서울대학교 전자공학과 학사
- 1982년 2월 : 서울대학교 전자공학과 석사
- 1988년 5월 : 미국 Syracuse University 전기/컴퓨터공학과 박사
- 1982년 3월 ~ 1992년 8월 : 한국전자통신연구원 멀티미디어연구실 실장
- 1986년 7월 ~ 1986년 8월 : NSF Supercomputer Center 초청 학생연구원
- 1987년 5월 ~ 1988년 5월 : Northeast Parallel Architecture Center 학생연구원
- 1992년 3월 ~ 1992년 8월 : 충남대학교 컴퓨터공학부 겸임교수
- 1992년 9월 ~ 현재 : 광운대학교 전자공학과 교수
- 2002년 3월 ~ 2017년 12월 : SC29-Korea 전문위원회 대표위원
- ORCID : <https://orcid.org/0000-0002-5036-3761>
- 주관심분야 : 비디오데이터처리, 영상압축, 컴퓨터비전