

H.264/AVC 스케일러블 비디오 코딩에서 빠른 부호화를 위한 단계적 모드 선택 기법

고경은¹ 강진미¹ 조미숙² 정기동¹

¹부산대학교 컴퓨터공학과 ²부산대학교 멀티미디어협동과정

oyoubi@melon.cs.pusan.ac.kr⁰¹, {wolf98¹, mscho², kdchung¹}@pusan.ac.kr

Fast Mode Decision Scheme for Inter Frame Coding in H.264 Scalable Video Coding

Gyeongun Goh⁰¹ Jinmi Kang¹ Misook Cho² Kidong Chung¹

¹Dept. of Computer Science and Engineering, Pusan National University

²Dept. of Multimedia, Pusan National University

1. 서론

H.264/AVC Scalable Video Coding(SVC)는 다양한 재생장치와 네트워크 환경에 적합한 멀티미디어 서비스를 위해 제안된 표준으로 공간 확장성(Spatial Scalability), 시간 확장성(Temporal Scalability), 화질 확장성(Quality Scalability)을 지원한다. 공간 확장성과 화질 확장성은 하나의 압축된 비트스트림으로 각각 다양한 해상도 또는 다양한 화질을 지원하기 위해 하나의 기본계층(Base Layer)과 다수의 확장계층(Enhancement Layer)으로 구성되며 계층 간의 중복성을 줄이기 위해 계층 간 예측(Inter Layer Prediction) 등 다양한 기법이 추가되었다. 시간 확장성은 다양한 프레임 임율을 지원하기 위해 계층적 B 구조(Hierarchical B Structure)를 사용한다. 이렇듯 SVC는 여러 가지 환경에 적합한 하나의 비트스트림을 보다 짧은 비트스트림으로 제공하면서 고품질의 영상으로 제공하기 위해 다양한 기술을 적용함으로써 부호화 과정의 복잡도가 증가하게 되는 문제점이 발생한다.

SVC 부호화 과정은 기존 H.264/AVC에서 확장되었다. 기본계층의 부호화는 최저의 화질과 해상도, 프레임율을 지원하며 H.264/AVC 부호화 방식과 동일하게 이루어진다. 향상계층의 경우 기존 H.264/AVC의 인터예측과 인트라예측뿐만 아니라 계층 간 중복성을 줄이기 위한 계층 간 예측이 추가되었다. 계층 간 예측은 하위 계층의 텍스처 정보를 사용하는 계층 간 인트라예측, 하위 계층의 잔여신호 정보를 사용하는 계층 간 잔여신호 예측, 하위 계층의 움직임 벡터와 매크로블록 크기를 사용하는 "Base layer mode"와 움직임 벡터에 1/4 단위 움직임 보상 과정을 거치는 "Quarter Per Refinement mode"가 있다.

부호화의 예측과정은 GOP단위로 이루어지며 프레임당 계층별로 기본계층부터 진행되고 계층의 종류에 따라 예측모드의 후보군이 달라진다. 예측모드가 가장 다양한 향상계층의 Non-Key picture의 예측과정은 그림1과 같다. 부호화하려는 해당 프레임의 매크로블록과 대응하는 하위계층의 매크로블록이 존재하면 Base layer mode 또는 계층 간 인트라예측을 수행한다. 또한 같은 계층의 이웃한 움직임 벡터를 그대로 이용하는 Direct 예측, 같은 프레임 내의 픽셀을 이용하는 인트라예측, 계층적 B 구조에 따른 B 픽처 예측이 수행된다. B 픽처 예측은 16x16, 16x8, 8x16, 8x8(8x4, 4x8, 4x4)의 블록 크기마다 "base layer mode", "quarter per refinement mode", 인터 B 픽처 예측을 각각 수행한다.

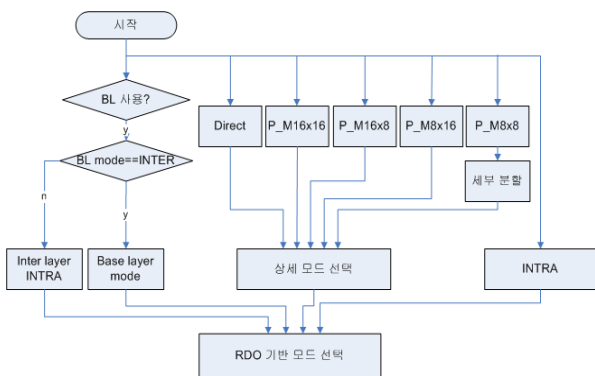


그림 1. Non-key picture의 예측모드 선택

2. 본론

본 논문은 JSVM11을 기반으로 향상계층의 RD 비용 분포와 예측모드를 분석하여 단계적으로 예측모드 후보군을 제한하고 RD 비용 기반 최적화 모드 선택 기법을 제안한다. 그림 3은 기본계층과 하나의 향상계층으로 이루어진 다양한 영상에서 여러가지 QP값(26,30,32,34,36,38)으로 실험한 결과로써, Non-Key picture인 향상계층이 기본계층의 정보를 사용할 수 있는 경우의 RD 비용별 매크로블록 예측모드의 분포를 나타낸 그래프이다. Base layer mode와 계층 간 인

이 논문은 2단계 두뇌한국21사업에 의하여 지원되었음.

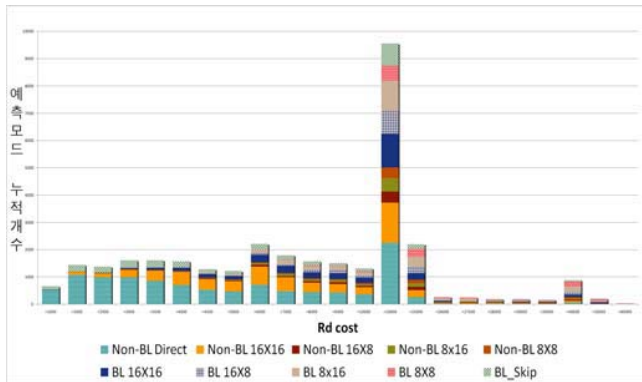


그림 3. RD 비용별 예측모드의 분포

트라예측을 BL_Skip으로 나타내고, 하위계층 정보를 사용하는 예측모드를 BL_‘블록 크기’, 하위계층 정보를 사용하지 않는 예측모드를 NON-BL_‘블록 크기’로 표현하였다. 이 그래프는 RD 비용에 따라 발생하는 예측모드의 후보군을 나눌 수 있음을 보여준다.

항상계층이 하위계층과 같은 블록 크기의 예측을 하는 경우는 약 68.1% 발생한다. 하위계층이 16x8, 8x16 블록 크기 예측을 수행한 경우에 현재 부호화하는 현재계층이 같은 블록 크기의 예측을 하거나 더 작은 블록인 8x8 블록 크기 예측을 하는 경우는 84.5%, 하위계층과 현재계층이 8x8 크기의 블록예측을 하는 경우는 약 78.5%의 확률을 보인다. 이는 하위계층의 모드에 따라 항상계층의 모드를 제한할 수 있음을 보여준다.

본 논문에서는 앞서 언급한 두 가지의 특성에 기반을 둔 항상계층의 Non-Key picture의 빠른 모드 선택 기법을 단계적으로 적용하는 기법을 제안한다. (경계값 TH1, TH2, TH3, TH4는 $TH1 < TH2 < TH3 < TH4$ 의 크기를 가진다.)

- 1단계: 전체 예측모드 중 선택되는 빈도가 두 번째인 BL_Skip 모드예측을 수행한다.
- 2단계: BL_Skip 모드의 RD 비용이 TH4 보다 작은 경우 Direct 예측을 후보군에 추가한다.
- 3단계: BL_Skip 모드의 RD 비용이 TH1보다 큰 경우 하위계층의 예측모드에 따라 후보군에 예측모드를 추가한다. 하위계층이 16x8(또는 8x16) 블록 크기 예측을 수행한 경우 부호화하는 현재계층은 16x8(또는 8x16) 블록 크기 예측모드와 8x8 블록 크기 예측모드를 후보군에 추가하고, 하위계층이 8x8 블록 크기 예측을 수행한 경우 현재계층은 8x8 블록 크기 예측모드를 후보군에 추가한 후 7단계로 진행한다. 그 외의 경우에는 계층 내 16x16 블록 크기 예측을 후보군에 추가하고 4 단계로 진행한다.
- 4단계: BL_Skip 모드의 RD 비용이 TH2 이상인 경우 16x16 블록 크기 계층 간 인트라예측모드를 후보군에 추가한다.
- 5단계: BL_Skip 모드의 RD 비용이 TH3 이상인 경우 16x8, 8x16 블록 크기 예측모드 중 RD 비용이 적은 예측모드를 후보군에 추가한다.
- 6단계: 5단계에서 선택된 예측모드가 16x16 블록 크기 예측모드 보다 RD 비용이 적은 경우 8x8 블록 크기 예측모드를 후보군에 추가한다.
- 7단계: 후보군 중 가장 적은 RD 비용을 가지는 예측모드를 매크로블록의 가장 적합한 예측모드로 선정한다.

3.결론

성능을 평가하기 위해 QCIF, CIF의 해상도로 구성된 2개의 공간 계층과 GOP 크기가 16인 시간 계층을 가진 4가지 영상에 대하여 제안하는 기법과 JSVM11을 비교하였다.

그림 4는 코딩효율과 화질의 감소 정도를 비교하기 위한 RD 비용 그래프이다. 비트율이 같을 때 PSNR의 차이가 0.1dB미만으로 원본과의 화질 차이가 거의 없음을 볼 수 있다. 제안하는 기법을 적용한 항상계층의 Non-Key picture를 부호화하는 시간은 최대 58.7%, 평균 38%의 감소량을 보였으며 비트율과 PSNR은 각각 0.08%증가, 0.01dB감소를 보이고 있으므로 화질의 차이는 거의 없으면서 부호화 시간이 크게 줄어 든 것을 볼 수 있다.

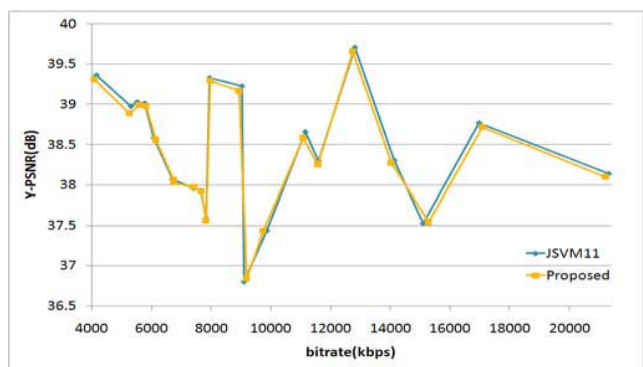


그림 4. RD 비용 그래프