

단일 보간을 이용한 화면내 복수 샘플 예측

이진호, 김휘용, 임성창, 이하현, 최진수
한국전자통신연구원
jinossoul@etri.re.kr,

Single Interpolation for Multi-sample Prediction in Intra Coding of HEVC

Jinho Lee, Hui Yong Kim, Sung-Chang Lim, Hahyun Lee, and Jin Soo Choi
ETRI

요 약

본 논문에서는 화면내 예측의 계산 복잡도를 감소시키기 위한 방법을 제안한다. HM2.0 의 화면내 예측 방법인 각도에 의한 예측(Angular prediction)은 예측 대상블록 내부의 각 샘플이 예측하기 위해 사용하는 참조샘플로부터 정수단위의 위치에 있지 않을 경우 참조 샘플들을 보간 하여 예측을 수행한다. 이러한 방법은 예측 대상 블록의 크기가 커질수록 보간 하는 횟수가 증가하고 계산 복잡도가 증가하는 단점이 있다. 제안하는 방법은 한번의 보간으로 복수의 샘플들을 예측함으로써 보간 횟수를 줄이고 결과적으로 계산의 복잡도를 줄인다. PU_32x32 에 대해 제안하는 방법을 적용하였을 경우, 보간하는 횟수는 HM2.0 대비 1/2 배 또는 1/4 배 가까이 감소하고 성능 저하는 High-Efficiency (HE)에서 0.1% 또는 0.2% 정도로 미미하게 나타났다.

1. 서론

최근 고해상도의 영상에 대한 수요가 증가하면서 이를 충족하기 위해 ISO/IEC JTC1/SC29/WG11 MPEG(Moving Picture Experts Group)과 ITU-T Q.6/16 VCEG(Video Coding Experts Group)이 협력하여 JCT-VC(Joint-Collaborative Team on Video Coding)^[1]라는 협력팀을 조직하고 새로운 표준화를 진행하고 있다. JCT-VC 표준화 단체는 HEVC(High Efficiency Video Coding)^[2]라고 불리는 새로운 비디오 표준을 위해 부호화 효율을 향상시킬 수 있는 새로운 알고리즘들을 요구하고 있다.

최신 비디오 코덱인 H.264/AVC^[3]에서는 시간적, 공간적 중복성을 제거함으로써 압축을 수행하고 있다. 이러한 하이브리드 구조는 HEVC 에서도 유지되고 있으며, 공간적 중복성을 제거하기 위한 화면내 예측(Intra prediction)은 방향성 예측에 기반하고 있다. H.264/AVC 의 화면내 예측 모드는 DC 모드 및 8 가지의 방향성 모드, 즉, 총 9 가지를 포함한다. 반면에, H.264/AVC 대비 2 배의 압축 효율을 목표로 하는 HEVC 표준화에서는 최근 HM2.0^[4](HEVC Model)이라는 테스트 모델을 결정했는데, 화면내 예측 방향의 수를 증가하여 총 34 가지의 화면내 예측모드를 수행하고 있다. 또한 부호화 대상 블록의 크기도 기존의 16x16, 8x8, 4x4 뿐만 아니라 64x64, 32x32 등도 포함하고 있으며 부호화 유닛(Coding Unit)과 예측 유닛(Prediction Unit), 변환 유닛(Transform Unit) 등의 구조를 가진다.

이처럼, 화면내 부호화 효율을 향상시키기 위해 예측 모드

수를 늘리고 블록의 크기를 다양화함으로써 상대적으로 복잡도가 증가하는 문제를 갖게 되었다. 따라서, HEVC 는 부호화 효율을 저하시키지 않으면서 부/복호화기의 복잡도를 낮출 수 있는 기술들을 요구하고 있으며 이러한 노력을 표준화가 종료될 때까지 계속 될 것이다.

본 논문에서는 HM2.0 의 화면내 예측의 복잡도를 줄이기 위한 방법을 제안한다. 구체적으로, 각도에 의한 예측을 수행할 때에 하나의 샘플을 예측하기 위해 한 번의 보간을 하는 대신에 복수의 샘플들을 한 번의 보간으로 예측함으로써 보간의 횟수를 줄이고 계산 복잡도를 줄일 수 있다. 또한 이 방법은 부호화기 뿐만 아니라 복호화기의 복잡도도 감소시킬 수 있다.

본 논문의 구성은 다음과 같다. 2 절에서는 HM2.0 의 화면내 예측에 대하여 살펴본 후, 3 절에서는 제안하는 방법을 설명한다. 4 절에서는 제안하는 방법을 적용한 실험 결과를 살펴보고 마지막으로 5 절에서 결론을 맺는다.

2. HM2.0 의 화면내 예측

HM2.0 의 화면내 예측은 기본적으로 H.264/AVC 와 동일한 개념을 지닌다. 즉, 비방향성 예측 모드 (DC 모드)와 방향성 예측 모드를 가지고 있다. 차이점은 방향성 모드 수가 33 가지로 증가하였고, 블록의 크기가 보다 다양화 되었으며, 매크로블록(Macroblock)을 대체하는 CU 와 PU, TU 등의 구조가 생겨났다. 또한, 방향성 예측은 각도에 의한 예측을

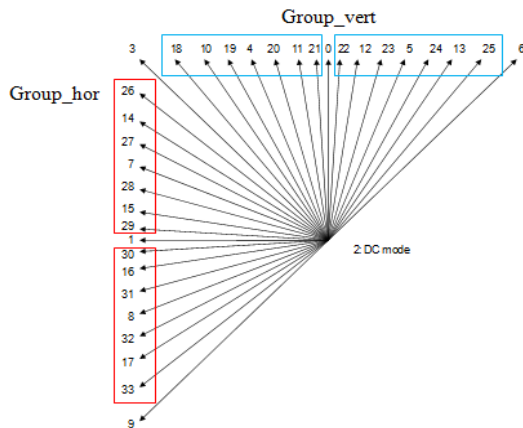


그림 1. HM2.0 의 화면내 예측 모드

수행하는데 샘플 단위의 예측을 수행하며, 예측 대상 샘플이 주변의 복원된 참조 샘플들로부터 정수위치에 존재하지 않을 경우에는 참조 샘플들을 보간하여 예측 값을 생성한다. 그림 1 은 HM2.0 의 화면내 예측 모드를 나타낸다. 모드 0 은 수직 모드(Vertical mode), 모드 1 은 수평 모드(Horizontal mode)를 의미하며, 나머지는 각도에 의해 정의되어져 있다. 이때, 45 도 각을 가지는 모드 3 과 5, 9 를 제외한 방향성 모드, 즉 그림 1 에서 'Group_vert' 와 'Group_hor' 에 해당하는 모드들은 앞에서 설명한 보간 예측을 수행한다.

보간하는 방법은 아래의 식 (1)과 같다.

$$\text{predSamples}[x, y] = ((32 - iFact) * \text{refMain}[\text{refMainIndex}] + iFact * \text{refMain}[\text{refMainIndex} + 1] + 16) >> 5 \quad (1)$$

여기서, $\text{predSamples}[x, y]$ 는 $[x, y]$ 위치의 예측된 샘플값을 나타내며, $iFact$ 는 참조 샘플들 사이의 가중치값, $\text{refMain}[\text{refMainIndex}]$ 은 현재 예측 모드에서 사용하는 참조 샘플을 나타낸다. 식 (1)의 계산은 하나의 샘플을 예측하기 위해 필요하기 때문에 기본적으로 블록 내부의 샘플 개수 만큼 보간을 수행하며 예측 블록의 크기가 커질수록 많은 연산량이 요구된다.

표 1 은 PU_32x32 에서 하나의 모드를 예측할 때 수행하는 보간 횟수 및 연산의 횟수를 나타낸다. 식 (1)에서와 같이 한번의 보간을 수행하는데 네 번의 덧셈, 뺄셈 ('+', '-') 연산과 두 번의 곱셈 (*) 연산, 한 번의 쉬프트 (>>) 연산이 실행된다. PU_32x32 의 경우 예측 대상 샘플의 수는 모두 1024(32x32)개가 되는데, 이들 중 모드에 따라 블록 내부의 한 개 또는 두 개의 라인이 참조 샘플로부터 정수 위치에 존재하여 보간 과정 없이 복사하여 예측한다. 따라서 992(32x31) 또는 960(32x30)개의 샘플들이 보간을 통하여 예측된다. 결론적으로, 표 1 에서와 같이 하나의 예측 모드를 수행하는데 필요한 연산량을 계산할 수 있다.

3. 제안하는 알고리즘

표 1. PU_32x32 에서 하나의 모드를 예측하기 위해 수행하는 보간 및 연산 횟수 (HM2.0)

| 연산자 | 보간에 사용되는 연산 횟수 (A) | 하나의 모드 예측시 필요한 보간 횟수 (B) | 하나의 모드 예측시 사용되는 연산 횟수 (C=A*B) |
|------|--------------------|--------------------------|-------------------------------|
| +, - | 4 | 992 or 960 | 3968 or 3840 |
| * | 2 | | 1984 or 1920 |
| >> | 1 | | 992 or 960 |

표 2. PU_32x32 에서 하나의 모드를 예측하기 위해 수행하는 보간 및 연산 횟수 (SIMP_2)

| 연산자 | 보간에 사용되는 연산 횟수 (A) | 하나의 모드 예측시 필요한 보간 횟수 (B) | 하나의 모드 예측시 사용되는 연산 횟수 (C=A*B) |
|------|--------------------|--------------------------|-------------------------------|
| +, - | 4 | 512 | 2048 |
| * | 2 | | 1024 |
| >> | 1 | | 512 |

영상을 부호화하는데 있어 움직임이 적고 평평한 영역은 큰 블록으로 부호화되는 경향이 있다. 큰 블록으로 부호화 된다는 것은 화소간의 상관도가 작은 블록에 비해 상대적으로 높기 때문이라고 할 수 있다. 이러한 관점에서 보았을 때, 블록의 크기가 큰 경우 HM2.0 의 각도에 의한 예측에서와 같이 하나의 샘플을 한 번의 보간에 의해 예측하는 것이 복잡도 대비 효율이 작을 수 있다.

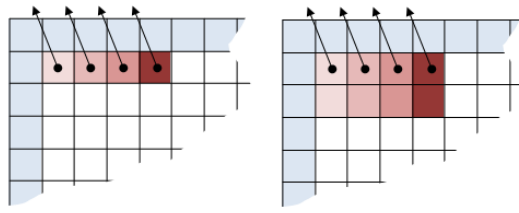
본 논문에서는 앞서 제기한 문제를 해결하기 위한 방법으로 단일 보간을 이용한 복수 샘플 예측(SIMP: Single Interpolation for Multi-sample Prediction)을 제안한다. SIMP 는 각도에 의한 화면내 예측에서 보간 횟수를 줄임으로써 계산 복잡도를 감소시키는데 목적이 있다. 이 방법은 부호화기 뿐만 아니라 복호화기의 복잡도를 감소시킬 수 있다. 구체적으로, 기존의 하나의 샘플을 예측하기 위한 한 번의 보간과 달리 복수개의 샘플들을 한 번의 보간을 통하여 예측함으로써 복잡도를 낮출 수 있다. 본 논문에서는 복수 샘플의 개수와 단일 보간을 위한 대표 위치의 변경을 고려한 4 가지의 SIMP 방법을 살펴본다.

3.1. 단일 보간을 통한 2 개의 샘플 예측 (SIMP_2)

단일 보간을 통하여 2 개의 샘플을 예측할 수 있다. 이때의 보간 방법은 식 (2), 식 (3)과 같다.

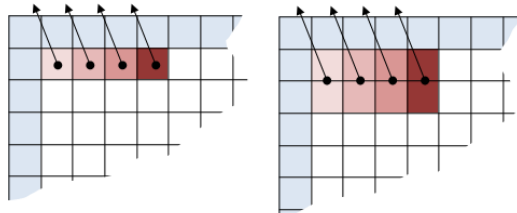
$$\begin{aligned} \text{predSamples}[x, y] &= \quad (2) \\ \text{predSamples}[x, y + 1] &= ((32 - iFact) * \text{refMain}[\text{refMainIndex}] + iFact * \text{refMain}[\text{refMainIndex} + 1] + 16) >> 5 \end{aligned}$$

$$\begin{aligned} \text{predSamples}[x, y] &= \quad (3) \\ \text{predSamples}[x + 1, y] &= ((32 - iFact) * \text{refMain}[\text{refMainIndex}] + iFact * \text{refMain}[\text{refMainIndex} + 1] + 16) >> 5 \end{aligned}$$



(a) HM2.0 (b) M1

그림 2. HM2.0 과 M1 의 각도에 의한 예측 방법.



(a) HM2.0 (b) M2

그림 3. HM2.0 과 M2 의 각도에 의한 예측 방법.

예측 모드가 그림 1 의 'Group_vert'에 속할 경우, 식 (2)와 같이 나타낼 수 있으며 'Group_hor'에 속할 경우, 식 (3)과 같이 나타낼 수 있다.

표 2 는 PU_32x32 에 SIMP_2 를 적용하였을 경우 수행하는 보간 횟수 및 연산의 횟수를 나타낸다. 한 번의 보간에 사용되는 연산의 횟수는 HM2.0 과 동일하고, 하나의 모드 예측시 필요한 보간의 횟수는 HM2.0 대비 1/2 배 정도인 512(32x16)번이 된다. 따라서 연산량도 감소하게 된다.

여기서 SIMP_2 를 수행할 때 한 번의 보간을 수행하는 위치를 고려하여 두 가지 방법을 살펴본다.

3.1.1. 방법 1 (M1)

방법 1 은 두 개의 예측 대상 샘플 중 참조 샘플로부터 가까운 거리에 있는 샘플의 위치에서 보간을 한 값을 예측값으로 하는 방법이다. 그림 2 는 예측 모드가 'Group_vert'에 속하였을 경우 HM2.0 의 각도에 의한 예측과 M1 예측의 예를 나타낸다. 즉, 두 샘플 중 상단 샘플의 위치에서 해당하는 참조 샘플들을 이용하여 보간을 수행하고, 이 때 보간된 값을 두 샘플의 예측값으로 사용한다. 'Group_hor'에 속하는 모드도 같은 방법으로 예측할 수 있다.

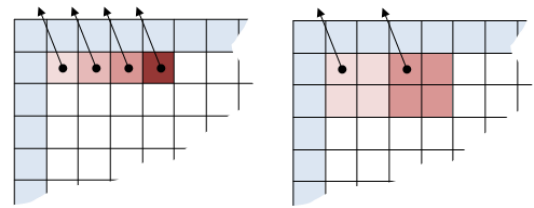
3.1.2. 방법 2 (M2)

단일 보간한 값이 두 샘플들의 예측값이 되기 때문에 두 개의 샘플을 한 번의 보간으로 보다 정확하게 예측하기 위하여 방법 2 에서는 두 개의 대상 샘플의 중간을 보간 위치로 한다. 예를 들어, 그림 3 에서와 같이 'Group_vert'에 속하는 예측 모드일 경우 두 샘플이 만나는 지점에서 대응하는 참조 샘플들의 보간을 통하여 예측을 수행한다.

3.2. 단일 보간을 통한 4 개의 샘플 예측 (SIMP_4)

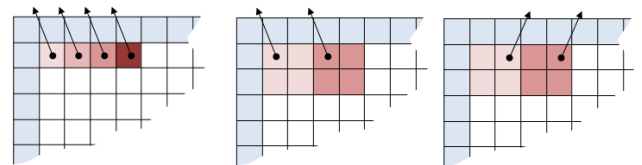
표 3. PU_32x32 에서 하나의 모드를 예측하기 위해 수행하는 보간 및 연산 횟수 (SIMP_4)

| 연산자 | 보간에 사용되는 연산 횟수 (A) | 하나의 모드 예측시 필요한 보간 횟수 (B) | 하나의 모드 예측시 사용되는 연산 횟수 (C=A*B) |
|------|--------------------|--------------------------|-------------------------------|
| +, - | 4 | 256 | 1024 |
| * | 2 | | 512 |
| >> | 1 | | 256 |



(a) HM2.0 (b) M3

그림 4. HM2.0 과 M3 의 각도에 의한 예측 방법.



(a) HM2.0 (b1) M4 (b2)

그림 5. HM2.0 과 M4 의 각도에 의한 예측 방법.

SIMP 의 다른 방법으로 4 개의 샘플을 단일 보간을 통하여 예측하는 방법을 살펴본다. 식 (4)는 SIMP_4 의 계산식을 나타내며, 표 3 은 PU_32x32 에 SIMP_4 를 적용하였을 경우 수행하는 보간 및 연산의 횟수를 나타낸다.

하나의 모드를 예측할 때 수행되는 보간 횟수는 HM2.0 대비 1/4 배 정도인 256(16x16)번이 된다. 여기서도 앞서와 같이 보간을 위한 위치에 따른 두 가지 방법을 살펴본다.

$$\begin{aligned}
 \text{predSamples}[x, y] &= \\
 \text{predSamples}[x, y + 1] &= \\
 \text{predSamples}[x + 1, y] &= \\
 \text{predSamples}[x + 1, y + 1] &= ((32 - iFact) * \text{refMain}[\text{refMainIndex}] + \\
 &\quad iFact * \text{refMain}[\text{refMainIndex} + 1] + 16) >> 5
 \end{aligned} \tag{4}$$

3.2.1. 방법 3 (M3)

방법 3 에서는 보간을 위한 위치를 하나의 샘플 위치로 고정한다. 구체적으로 그림 4 에서와 같이 'Group_vert'에 속하는 예측 모드인 경우 4 개의 샘플 중 좌상의 샘플을 보간을 위한 위치로 고정한다. 'Group_hor'에 속하는 모드도 같은 위치를 고정하여 보간을 수행한다.

3.2.2. 방법 4 (M4)

방법 4 에서는 보간을 위한 대표 샘플 위치를 모드에 따라 다르게 한다. 예를 들어, 그림 5 에서의 (b1)과 같이, 'Group_vert' 모드 중 좌상의 참조 샘플을 사용하는 모드의 경우 4 개의 예측 대상 샘플 중 좌상의 샘플 위치에서 보간을 수행한다. 또한 (b2)와 같이 우상의 참조 샘플을 사용하는 모드의 경우 4 개의 예측 대상 샘플 중 우상의 샘플 위치에서 보간을 수행하여 예측의 정확도를 높이고자 한다.

4. 실험 결과

제안하는 알고리즘을 HM2.0 에 적용하여 실험하였다. 실험 환경은 기본적으로 표준화에서 사용하는 JCTVC-D600^[5]의 공통 실험 조건으로 하였으며 High-Efficiency (HE)에 대하여 실험하였다. SIMP 는 PU_32x32 에만 적용하였으며, 실험에 사용된 시퀀스는 표준화 테스트 영상이며, 부호화 성능을 파악하기 위한 방법으로는 BD-rate(Bjontegaard delta bitrate)^[6]를 측정하였다.

표 4 는 본 논문에서 제안하는 각 방법을 구현하였을 때의 실험 결과를 나타낸다. 여기서 양의 값을 가지는 Luma(Y) BD-rate 는 부호화 손실을 의미한다. SIMP_2 를 실험한 M1 과 SIMP_4 를 실험한 M3 의 실험결과를 비교하였을 경우, 평균적인 부호화 손실은 각각 0.1%와 0.2%로 미미하게 나타났다. 하지만, 4 개의 샘플을 한 번의 보간으로 예측한 M3 의 경우 'Nebuta'와 'Vidyo3' 시퀀스에서 1%에 가까운 손실을 보였다.

다음으로, 단일 보간을 위한 대표 위치 변경에 의한 영향을 살펴보기 위한 M1 과 M2 의 부호화 손실을 비교하였을 경우 두 방법이 동일한 결과를 보였다. 이러한 양상은 M3 와 M4 를 비교할 때도 같음을 알 수 있다. 결론적으로, 블록의 크기가 상대적으로 큰 PU_32x32 에서는 샘플들 간의 상관도가 높기 때문에 복수 샘플들 내부에서 단일 보간을 위한 위치를 이동하는 것이 부호화 효율에 큰 영향을 주지 않음을 알 수 있었다.

5. 결론

본 논문은 화면내 예측의 복잡도를 줄이기 위한 단일 보간을 통한 복수 샘플 예측 방법을 제안한다. 기존의 각도에 의한 화면내 예측 수행할 때 요구되는 보간의 횟수를 줄임으로써 계산 복잡도를 감소시킨다. 제안하는 방법을 적용하였을 경우, 부호화 효율의 큰 손실 없이 기존 대비 1/2 배 또는 1/4 배로 연산량을 줄이는 효과를 보였다.

6. 감사의 글

본 연구는 방송통신위원회의 ETRI 연구개발지원사업의 연구결과로 수행되었음 (KCA-2011-11921-02001)

표 4. SIMP 실험 결과

| Resolution | Sequence | YBD-rate | | | |
|-------------------|-----------------|----------|-----|-----|-----|
| | | M1 | M2 | M3 | M4 |
| 2560 x 1600 | Traffic | 0.0 | 0.0 | 0.2 | 0.2 |
| | PeopleOnStreet | 0.0 | 0.0 | 0.1 | 0.1 |
| | Nebuta | 0.4 | 0.4 | 0.8 | 0.8 |
| | SteamLocomotive | 0.1 | 0.1 | 0.4 | 0.4 |
| 1920 x 1080 | Kimono | 0.1 | 0.1 | 0.5 | 0.5 |
| | ParkScene | 0.0 | 0.0 | 0.1 | 0.1 |
| | Cactus | 0.0 | 0.0 | 0.2 | 0.2 |
| | BasketballDrive | 0.0 | 0.0 | 0.3 | 0.3 |
| 832 x 480 | BQTerrace | 0.0 | 0.0 | 0.2 | 0.2 |
| | BasketballDrill | 0.1 | 0.1 | 0.4 | 0.3 |
| | BQMall | 0.0 | 0.0 | 0.0 | 0.0 |
| | PartyScene | 0.0 | 0.0 | 0.0 | 0.0 |
| 416 x 240 | RaceHorses | 0.0 | 0.0 | 0.0 | 0.0 |
| | BasketballPass | 0.0 | 0.0 | 0.1 | 0.1 |
| | BQSquare | 0.0 | 0.0 | 0.0 | 0.0 |
| | BlowingBubbles | 0.0 | 0.0 | 0.0 | 0.0 |
| 1280 x 720 | RaceHorses | 0.0 | 0.0 | 0.0 | 0.0 |
| | Vidyo1 | 0.0 | 0.0 | 0.3 | 0.3 |
| | Vidyo3 | 0.3 | 0.3 | 0.8 | 0.8 |
| | Vidyo4 | 0.0 | 0.0 | 0.3 | 0.2 |
| Average | | 0.1 | 0.1 | 0.2 | 0.2 |

7. 참고문헌

- [1] ISO/IEC JTC1/SC29/WG11, "Terms of Reference of the Joint Collaborative Team on Video Coding Standard Development," N1112, Jan. 2010.
- [2] JCT-VC, "Architectural Outline of Proposed High Efficiency Video Coding Design Elements", document JCTVC-A202 of JCT-VC, Apr. 2010.
- [3] ITU-T and ISO/IEC JTC 1, Advanced video coding for generic audiovisual services, ITU-T Recommendation H.264 and ISO/IEC 14496-10 (MPEG4-AVC), Fifth edition: May. 2009.
- [4] JCTVC-D503, "Summary of HEVC working draft 2 and HEVC test model (HM)", Jan 2011.
- [5] JCTVC-D600, "Common test conditions and software reference configurations", Jan 2011.
- [6] G. Bjontegaard, "Calculation of average PSNR differences between RD-curves", ITU-T SG16 Q.6 VCEG, Doc. VCEG-M33, 2001.