

화면 간 예측에 사용되는 보간 필터에 관한 연구

\*임성원 \*\*문주희

세종대학교 정보통신연구소

\* [sw823@sju.ac.kr](mailto:sw823@sju.ac.kr) \*\* [jhmoon@sejong.ac.kr](mailto:jhmoon@sejong.ac.kr)

A Study about Interpolation Filter used for Inter Prediction

\*Lim, Sung-Won \*\*Moon, Joo-Hee

Information Telecommunication Research Institute, Sejong University

요약

동영상 부호화에서 중요하고 매우 높은 비중을 차지하는 화면 간 예측에서는 휘도와 색차 성분에서 다양한 tap의 보간 필터를 사용한다. 본 논문에서는, 현재 표준인 HEVC(High Efficiency Video Coding)에서 쓰이는 보간 필터인 DCT-IF(Discrete Cosine Transform-based Interpolation Filter)에 알아보고 tap 수에 따른 성능을 분석한다.

1. 서론

최근 인터넷에서는 동영상 서비스의 수요가 급격히 증가하였다. 인터넷 회선이 발전해 나가고 있지만 그에 따라 동영상의 해상도 역시 급격히 증가함에 따라 효율적인 동영상 부호화 방식이 필요하다. 그에 따라 2013년 1월, 표준화 단체인 ITU-T와 MPEG은 최신 동영상 부호화 표준인 HEVC[1]을 제정하였다.

HEVC는 이전 표준인 H.264/AVC와 비교하여 많은 기술들이 추가되었다. 부호화 블록의 크기는 최대 64x64로 늘어났으며 화면 내 예측의 모드의 수는 35가지로 늘어났고, 화면 간 예측에서는 머지 모드의 추가와 정교한 움직임 추정을 위해 보간 시 DCT-IF를 사용하였다. 변환과 루프필터에서는 RQT(Recursive Quadtree Transform)와 SAO(Sample Adaptive Offset)를 추가하여 부호화 효율을 H.264/AVC 대비 약 2배 가까이 증가시켰다. 부호화 효율에서 높은 비중을 차지하는 화면 간 예측에서, 정교한 움직임 추정을 위한 보간 기술은 매우 중요하다. 본 논문에서는 HEVC에서 사용되는 보간 기술인 DCT-IF 계수의 유도방식에 대해 설명하고 사용되는 탭수에 따른 성능변화를 살펴본다.

2. HEVC의 DCT-IF 계수 유도 방식

HEVC에서는 1/2위치에 존재하는 화소와 1/4, 3/4위치에 존재하는 화소를 위해 DCT-IF를 사용한다. 그림 1에서 A, B, C, D는 정수 위치에 존재하는 화소들이고, a~o는 분수 위치에 존재하는 화소들이다. a,b,c와 d,h,i 위치에 존재하는 화소들은 1D 포간 필터를 이용하여 생성되며 그 외 화소들은 2D 보간 필터가 이용된다. 2D 보간 필터는 세로로 1D 보간 필터를 사용하여 나온 값들에 다시 가로로 1D 보간 필터를 사용하는 방식이다. 그림 2는 1D 보간 필터를 설명하고 있다. 보간하고자 하는 화소의 왼쪽 4개의 정수 위치에 존재하는 화소와 오른쪽 4개의 정수 위치에 존재하는 화소들을 이용한다. 여기서  $x_0 \sim x_7$

은 정수 위치에 존재하는 화소를 의미하며,  $x_{q0}, x_h, x_{q1}$ 은 각각 1/4, 1/2, 3/4 위치에 존재하는 화소들을 의미한다.

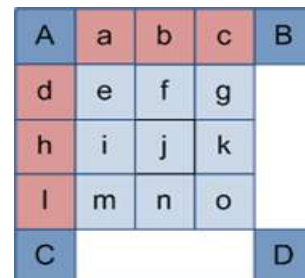


그림 1. 화소들의 위치

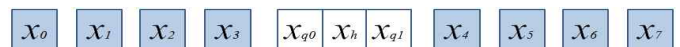


그림 2. 보간을 위한 정수 위치 화소와 분수 위치 화소들

분수 위치에 존재하는 화소를 보간하기 위해 위치의 화소들을 이용하는데, 예를 들어 1/2 화소 위치를 보간할 경우 수식 (1)과 같이 표현될 수 있다.

$$x_h = C_0x_0 + C_1x_1 + C_2x_2 + C_3x_3 + C_4x_4 + C_5x_5 + C_6x_6 + C_7x_7 \quad (1)$$

여기서  $C_0 \sim C_7$ 은 보간 필터 계수들이다. 해당하는 보간 필터 계수들은 다음에 설명되는 바와 같이 DCT(Discrete Cosine Transform)와 IDCT(Inverse DCT)식을 이용하여 유도한다. 유한개의 입력 데이터  $\{x(n), 0 \leq n \leq N-1\}$ 에 대한 1차원 DCT와 IDCT는 각각 수식 (2)와 수식 (3)과 같이 정의된다.

$$DCT : y(k) = \alpha(k) \sum_{n=0}^{N-1} x(n) \cos\left[\frac{\pi k(2n+1)}{2N}\right], 0 \leq k \leq N-1 \quad (2)$$

$$IDCT : x(n) = \sum_{k=0}^{N-1} \alpha(k)y(k) \cos\left[\frac{\pi k(2n+1)}{2N}\right], 0 \leq n \leq N-1 \quad (3)$$

$\alpha(0) = \frac{1}{N}$ ,  $\alpha(k) = \sqrt{\frac{2}{N}}$  이며, 보간 필터 계수를 유도하기 위해 식 (3)의 IDCT를 다시 한번 정리하게 되면 수식 (4)로 정리가 된다.

$$x(n) = \alpha(0)y(0)\cos\frac{\pi \times 0 \times (2n+1)}{2N} + \dots + \alpha(N-1)y(N-1)\cos\frac{\pi \times (N-1) \times (2n+1)}{2N} \quad (4)$$

수식 (4)에서  $y(0) \sim y(N-1)$ 에 대해 수식 (2)를 대입하고  $x(0) \sim x(N-1)$ 으로 정리하면 수식 (5)와 같이 정리된다.

$$x(n) = x(0)[\alpha^2(0)\cos\frac{\pi \times 0 \times (2 \times 0 + 1)}{2N} \cos\frac{\pi \times 0 \times (2n+1)}{2N} + \dots + \alpha^2(N-1)\cos\frac{\pi \times (N-1) \times (2 \times 0 + 1)}{2N} \cos\frac{\pi \times (N-1) \times (2n+1)}{2N}] + \dots + x(N-1)[\alpha^2(0)\cos\frac{\pi \times 0 \times (2 \times (N-1) + 1)}{2N} \cos\frac{\pi \times 0 \times (2n+1)}{2N} + \dots + \alpha^2(N-1)\cos\frac{\pi \times (N-1) \times (2 \times (N-1) + 1)}{2N} \cos\frac{\pi \times (N-1) \times (2n+1)}{2N}] = x(0)C_0(n) + x(1)C_1(n) + \dots + x(N-1)C_{N-1}(n) \quad (5)$$

최종적으로 유도된  $C_0(n) \sim C_{N-1}(n)$ 은 미리 계산되어 테이블로 저장된다.

### 3. 실험 결과

수식 (5)를 이용하여 DCT-IF 계수를 2-Tap에서 8-Tap까지 유도하여 각각 실험하였다. 실험 조건은 Lowdelay 메인 프로파일을 사용하였으며 실험 영상은 각 클래스마다 하나의 영상으로 실험하였고 각 100장씩 실험하였다. 참조 영상은 1장을 사용하였으며 그 외 실험 조건은 jct-vc에서 지정하는 공통 실험 조건[2]를 이용하였다. 표 1~ 표 3을 통하여 보간 필터의 tap수가 증가함에 따라 전체적인 BD-rate는 증가함을 확인할 수 있다.

표 1. 2-Tap vs 4-Tap

Class	Sequence name	BD-rate Y (%)	BD-rate U (%)	BD-rate V (%)
A	Traffic	-29.3%	-18.3%	-18.4%
B	Cactus	-15.3%	-9.8%	-9.7%
C	BasketballDrill	-13.8%	-6.0%	-5.9%
D	BasketBallPass	-6.6%	-4.7%	-5.2%
All		-15.1%	-9.0%	-9.3%

표 2. 2-Tap vs 6-Tap

Class	Sequence name	BD-rate Y (%)	BD-rate U (%)	BD-rate V (%)
A	Traffic	-26.1%	-17.7%	-17.9%
B	Cactus	-16.8%	-11.4%	-11.6%
C	BasketballDrill	-16.2%	-7.4%	-8.0%
D	BasketBallPass	-8.8%	-5.9%	-6.0%
All		-16.2%	-10.3%	-10.6%

표 3. 2-Tap vs 8-Tap

Class	Sequence name	BD-rate Y (%)	BD-rate U (%)	BD-rate V (%)
A	Traffic	-29.2%	-19.9%	-20.3%
B	Cactus	-17.6%	-12.1%	-12.1%
C	BasketballDrill	-18.5%	-9.0%	-9.7%
D	BasketBallPass	-9.9%	-6.5%	-7.1%
All		-18.1%	-11.6%	-12.0%

### 4. 결론 및 향후 연구 방향

표 1 ~ 표 3을 통하여, 보간 필터의 tap수를 증가시킬수록 전체적인 성능은 올라가는 것으로 확인되었다. 하지만 A 클래스의 Traffic영상의 경우 휘도 부분은 4-Tap 필터를 사용하는 것이 가장 좋은 효율이 나타남을 확인할 수 있으며, 색차 성분도 6-Tap 필터를 사용하는 것에 비해 효율이 좋음을 알 수 있다. 따라서 메모리 요구사항을 줄이고 더 나은 효율을 위하여 영상의 특성에 따라 적절한 tap수를 사용하여 연구가 필요하다.

### 5. 감사의 글

이 논문은 2015년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(No. 2014R1A1A2055351)

### 참고 문헌

[1] B. Bross, W.-J. Han, J.-R. Ohm, G. J. Sullivan, Y.-K. Wang, T. Wiegand, "High Efficiency Video Coding (HEVC) text specification draft 10" document JCTVC-L1003, ITU-T/ISO/IEC Joint Collaborative Team on Video Coding (JCT-VC), Jan. 2013

[2] F. Bossen, "Common HM test conditions and software reference configurations" document JCTVC-L1100, ITU-T/ISO/IEC Joint Collaborative Team on Video Coding (JCT-VC), Jan, 2013