

특집논문 (Special Paper)

방송공학회논문지 제22권 제3호, 2017년 5월 (JBE Vol. 22, No. 3, May 2017)

<https://doi.org/10.5909/JBE.2017.22.3.321>

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

DST 기반 보간 필터를 이용한 인터 코딩

김명준^{a)}, 이영렬^{a)†}

Inter Coding using DST-based Interpolation Filter

MyungJun Kim^{a)} and Yung-Lyul Lee^{a)†}

요약

HEVC 표준은 정수 화소로 표현된 신호에 DCT-II를 기반으로 하는 보간 필터를 사용하여 부화소 신호를 생성한다. 이러한 방법으로 생성된 신호는 움직임 보상 및 예측의 성능 향상을 가져온다. HEVC 표준은 부화소를 생성하기 위해서 길이가 다른 각각의 DCT 보간 필터를 사용하고 있다. 1/2-화소를 생성할 땐, 필터의 길이가 8인 DCT 기반 보간 필터를 사용하며, 1/4-화소와 3/4-화소의 경우에는 필터의 길이가 7인 DCT 기반 보간 필터를 사용한다. 본 논문에서는 DST-VII을 기반으로 하는 보간 필터를 제안하여, 움직임 보상 및 예측의 성능 향상을 가져온다. 본 논문에서 제안하는 방법은 HEVC 표준보다 BD-rate가 Random Access와 Low Delay B configurations에서 각각 0.45%와 0.5%의 성능 향상을 가져온다.

Abstract

High Efficiency Video Coding (HEVC) adopted the Discrete Cosine Transform-II (DCT-II) based interpolation filter to improve coding efficiency in motion compensation and estimation. In HEVC, the interpolation filters based on the DCT-II are composed of 8-point for half-pixel and 7-point for 1/4-pixel and 3/4-pixel. In this paper, a DST-VII based interpolation filter is used to improve motion compensation and estimation. The experimental results which applied the DST-VII interpolation filter are presented. They show the 0.45% of average bitrate reduction in Random Access configuration and 0.5% of average bitrate reduction in Low Delay B configuration, respectively.

Keyword : HEVC, Video Coding, DCT (Discrete Cosine Transform), DST (Discrete Sine Transform)

I. 서론

a) 세종대학교 컴퓨터공학부(Sejong University, Dept. of Computer Eng.)
† Corresponding Author : 이영렬(Yung-Lyul Lee)

E-mail: yllee@sejong.ac.kr
Tel: +82-2-3408-3753

ORCID: <http://orcid.org/0000-0003-2709-8282>

※ 이 논문은 일부분 2015년도 정부(미래창조과학부)의 지원으로 한국연구재단의 지원을 받아 수행된 연구임(NRF-2015R1A2A2A01006085).

Manuscript received March 6, 2017; Revised May 12, 2017; Accepted May 23, 2017.

JCT-VC(Joint Collaborative Team on Video Coding)^[1]는 ITU-T VCEG(Video Coding Expert Group)와 ISO/IEC MPEG(Moving Picture Expert Group)에 의해 공동으로 만들어진 그룹이다. JCT-VC 그룹은 2010년 1월에 결성되었으며 2010년 4월 독일 드레스덴의 1차 회의를 시작으로 차

Copyright © 2017 Korean Institute of Broadcast and Media Engineers. All rights reserved.

"This is an Open-Access article distributed under the terms of the Creative Commons BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited and not altered."

세대 표준화 비디오 코덱인 H.265/HEVC(High Efficiency Video Coding)^[2]의 개발을 진행하였다. HEVC 표준 기술에서의 움직임 예측 및 보상(MCP, Motion Compensation Prediction)이 영상 신호의 중복을 줄여 줌으로써 성능 향상에 상당히 많은 부분에 기여한다. 움직임 예측 및 보상은 예측 단위인 PU(Prediction Unit)와 가장 유사한 블록을 참조된 픽쳐들 내에서 찾는다. 이러한 유사한 블록을 찾는 과정에서 HEVC 표준 기술에서의 입력된 신호는 디지털화된 신호 형태를 나타내고 있기 때문에 정확한 위치의 움직임 벡터를 찾기 어렵다. 이러한 문제점을 해결하기 위해서 HEVC 표준 기술은 정수 화소로 표현된 신호에 DCT 기반 보간 필터^{[3][4][5]}를 사용하여 부화소 신호를 생성한다. 그리하여 더 정확한 위치의 움직임 벡터를 찾아 최소의 SAD (Sum of Absolute Difference)를 갖는 참조 픽쳐의 블록을 찾아낸다. HEVC 표준 기술은 1/2-화소와 1/4-화소 단위 보간 샘플 값을 구할 때 길이가 각각 8과 7인 보간 필터를 사용한다. 본 논문에서는 DST-VII을 활용한 기반 보간 필터를 제안한다. 제안된 DST 기반 보간 필터의 길이도 HEVC 표준 기술에서 사용되는 DCT 기반 보간 필터와 같이 1/2-화소와 1/4-화소 단위 보간 샘플 값을 구할 때 필터의 길이를 각각 8과 7인 보간 필터를 사용한다. 제안된 방법은 HEVC 시험 모델인 HM16.6 보다 BD-rate(Bjøntegaard Delta-rate)^[9]이 Random Access와 Low Delay B configurations에서 각각 0.45%와 0.5% 감소하였다. 본 논문의 구성은 다음과 같다. 2절에서는 HEVC내에서 사용되는 기존 연구에 관하여 설명하며, 3절에서는 본 논문에서 제안하는 방법을 설명한다. 4절에서는 실험 및 결과, 5절에서는 결론으로 이루어진다.

II. 기존 연구

1. DCT 기반 보간 필터(Discrete Cosine Transform based Interpolation Filter)

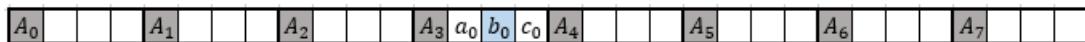


그림 1. 움직임 보상 및 참조 샘플 생성에서 부화소 위치
Fig.1. Fractional position in luma motion compensation

표준화 비디오 코덱인 H.265/HEVC에서는 DCT 기반 보간 필터^[3]를 적용한다. 1과 2의 식은 각각 DCT-II과 IDCT-II(Inverse DCT-II)을 나타낸다.

$$X(k) = \sqrt{\frac{2}{N}} \sum_{n=0}^{N-1} c_k x(n) \cos \frac{(n+1/2)\pi k}{N} \quad (1)$$

$$x(n) = \sqrt{\frac{2}{N}} \sum_{k=0}^{N-1} c_k X(k) \cos \frac{(n+1/2)\pi k}{N} \quad (2)$$

식 1에서 $X(k)$ 는 DCT-II 계수를 나타내며, $x(n)$ 은 IDCT-II 계수를 나타낸다. c_k 는 k 가 0일 때 $\frac{1}{\sqrt{2}}$ 값을 가지며, k 가 0 이외의 다른 값을 가질 경우 0을 갖는다. HEVC에서 사용되는 DCT 기반 보간 필터를 적용하기 위해선 식 1을 식 2에 대입한다. 식 1이 대입된 식 2는 식 3과 같이 표현할 수 있다.

$$= \frac{x(n)}{N} \sum_{m=0}^{N-1} x(m) \sum_{k=0}^{N-1} c_k^2 \cos \frac{(m+1/2)\pi k}{N} \cos \frac{(n+1/2)\pi k}{N} \quad (3)$$

예를 들어, 식 3에서 1/2-화소 필터 계수를 구하기 위해서 $N=8$, $n=3.5$ 가 사용된다. 그리고 1/4-화소 필터 계수를 구하기 위해서 $N=7$, $n=3.25$ 가 사용된다. 위와 같은 과정을 수행하면 DCT 기반 보간 필터에 사용되는 계수들 표 1로 나타낼 수 있다.

표 1. DCT-II 기반 보간 필터 계수

Table 1. DCT-II based Interpolation Filter Coefficients

| Index i | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---------------------|----|---|-----|----|----|-----|---|----|
| 1/2-pixel filter[i] | -1 | 4 | -11 | 40 | 40 | -11 | 4 | -1 |
| 1/4-pixel filter[i] | -1 | 4 | -10 | 58 | 17 | -5 | 1 | |

$$b_0 = (-1 \cdot A_0 + 4 \cdot A_1 - 11 \cdot A_2 + 40 \cdot A_3 + 40 \cdot A_4 - 11 \cdot A_5 + 4 \cdot A_6 - 1 \cdot A_7 + 32) \gg 6 \quad (4)$$

표 1은 HEVC에서 사용되는 DCT 기반 보간 필터의 계수들의 값이다. 그림 1에서는 A_0, A_1, \dots, A_7 은 정수 화소 위치를 나타내며, a_0, b_0, c_0 는 각각 1/4, 1/2, 3/4 화소 위치를 나타낸다. 표 1의 DCT 기반 보간 필터를 사용하여 그림 1에 나타낸 b_0 를 식 4로 나타낼 수 있다.

III. 제안하는 방법

1. DST 기반 보간 필터(Discrete Sine Transform based Interpolation Filter)

본 논문에서는 DST^[6] 기반 보간 필터를 제안한다. 5와 6의 식은 각각 DST-VII와 IDST-VII(Inverse DST-VII)을 나타낸다.

$$X(k) = \sqrt{\frac{2}{N+1/2}} \sum_{n=0}^{N-1} x(n) \sin \frac{(n+1)(k+1/2)\pi}{N+1/2} \quad (5)$$

$$x(n) = \sqrt{\frac{2}{N+1/2}} \sum_{k=0}^{N-1} X(k) \sin \frac{(n+1)(k+1/2)\pi}{N+1/2} \quad (6)$$

식 5에서 $X(k)$ 는 DST-VII 계수를 나타내며, $x(n)$ 은 IDST-VII 계수를 나타낸다. 본 논문에서 제안하는 DST 기반 보간 필터를 적용하기 위해선 식 5을 식 6에 대입한다. DCT 기반 보간 계수를 구하는 과정과 유사하다. 식 5가 대입된 식 6은 식 7과 같이 표현할 수 있다.

예를 들어, 식 7에서 1/2-화소 필터 계수를 구하기 위해서 $N=8, n=3.5$ 가 사용된다. 그리고 1/4-화소 필터 계수를 구하기 위해서 $N=7, n=3.25$ 가 사용된다. 위와 같은 과정을 수행하면 DST 기반 보간 필터에 사용되는 계수들 표 2로 나타낼 수 있다.

$$x(n) = \frac{2}{N+1/2} \sum_{m=0}^{N-1} x(m) \sum_{k=0}^{N-1} \sin \frac{(m+1)(k+1/2)\pi}{N+1/2} \cos \frac{(n+1)(k+1/2)\pi}{N+1/2} \quad (7)$$

$$b_0 = (-2 \cdot A_0 + 6 \cdot A_1 - 13 \cdot A_2 + 41 \cdot A_3 + 41 \cdot A_4 - 13 \cdot A_5 + 6 \cdot A_6 - 2 \cdot A_7 + 32) \gg 6 \quad (8)$$

표 2. DST-VII 기반 보간 필터 계수

Table 2. DST-VII based Interpolation Filter Coefficients

| Index i | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---------------------|----|---|-----|----|----|-----|---|----|
| 1/2-pixel filter[i] | -2 | 6 | -13 | 41 | 41 | -13 | 6 | -2 |
| 1/4-pixel filter[i] | -2 | 5 | -11 | 58 | 18 | -6 | 2 | |

표 2는 본 논문에서 제안하는 DST 기반 보간 필터의 계수들의 값이다. 표 2의 DST 기반 보간 필터를 사용하여 그림 1에 나타낸 b_0 를 식 8로 나타낼 수 있다.

2. Sinc 기반 보간 필터(Sinc based Interpolation Filter)

신호처리에서는 Sinc를 기반으로 하는 보간 필터가 가장 이상적이다. 또한 Sinc는 이상적인 저역 필터(low-pass filter)의 기능을 한다. Sinc기반 보간 필터를 식 5로 나타낼 수 있다.

$$x(t) = \sum_{k=-\infty}^{\infty} x(kT_s) \frac{\sin \frac{\pi}{T_s} (t - kT_s)}{\frac{\pi}{T_s} (t - kT_s)} \quad (9)$$

$x(t)$ 에서 t 는 서브샘플 및 정수샘플 위치, k 는 정수값을 나타내며 샘플링 주기 T_s 는 1로 하였다. Sinc 기반 보간 필터의 경우 필터의 길이를 $-\infty$ 에서 ∞ 까지 봐야지만 이상적인 저역 필터의 기능을 할 수 있다. 하지만 H.265/HEVC에 적용되는 DCT 기반 보간 필터의 경우 필터의 길이인 N 을 8 또는 7로 제한하고 있기 때문에, H.265/HEVC내에 Sinc 기반 보간 필터를 적용하기 어렵다. Sinc 기반 보간 필터^{[7][8]}와 DCT 기반 보간 필터를 주파수응답함수(frequency response function)를 통해서 볼 때, Sinc 기반 보간 필터가 DCT 기반 보간 필터보다 더 많은 물결 파동을 갖는다. 제

안하는 방법인 DST 기반 보간 필터와 DCT 기반 보간 필터를 주파수응답함수로 비교해보았을 때, DST 기반 보간 필터가 이상적인 필터인 Sinc 기반 보간 필터의 물결 파동과 더 유사하였다. 그러므로 DST 기반 보간 필터가 DCT 기반 보간 필터보다 더 강한 저역 필터의 특성을 갖는다.

IV. 실험 및 결과분석

1. 실험 조건

본 논문은 HEVC 참조 소프트웨어인 HM 16.6에서 구현되었으며, 제네바에서 발표된 HEVC공통 실험조건^[10]에 맞추어 실행하였다. Low Delay P, Low Delay B, 그리고

표 3. DST 기반 보간 필터를 적용한 실험 결과

Table 3. DST-IF experimental results for luma components

| Classes | Sequences | Low Delay B Pictures(%) | Low Delay P Pictures(%) | Random Access(%) |
|----------|-----------------|-------------------------|-------------------------|------------------|
| B | Kimono | 0.3 | 1.2 | 0.2 |
| | ParkScene | 0.8 | 2.1 | 0.3 |
| | Cactus | 0.8 | 2.3 | 0.2 |
| | BasketballDrive | 0.1 | 1.2 | 0.1 |
| | BQTerrace | 1.5 | 5.3 | 1.0 |
| C | BasketballDrill | -0.9 | 0.3 | -0.2 |
| | BQMall | -0.2 | 1.3 | -0.5 |
| | PartyScene | -1.7 | -0.2 | -2.5 |
| | RaceHorses | 0.6 | 1.7 | 0.4 |
| D | BasketballPass | 0.1 | 0.8 | -0.1 |
| | BQSquare | -4.1 | -0.4 | -5.2 |
| | PartyScene | -1.5 | 0 | -1.8 |
| | RaceHorses | 0.5 | 1.2 | 0.2 |
| E | FourPeople | 0.6 | 2.4 | |
| | Johnny | 0.5 | 4.4 | |
| | KristenAndSara | 0.6 | 2.4 | |
| Averages | Class B | 0.7 | 2.4 | 0.4 |
| | Class C | -0.5 | 0.8 | -0.7 |
| | Class D | -1.3 | 0.4 | -1.7 |
| | Class E | 0.5 | 3.1 | |
| | All | -0.1 | 1.6 | -0.6 |

Random Access Configurations에 맞추어 실험을 진행하였으며 사용한 QP는 22, 27, 32, 37이다. 시퀀스는 총 16개로 구성하여 B-E클래스에 속하는 모든 시퀀스들에 대하여 실험하였다. B-E 클래스의 각각의 해상도는 1080p, 832x480, 416x240, 720p에 해당된다. 또한, 각 시퀀스에 해당하는 프레임은 10초인 전체 분량으로 정하여 실험하였다. 본 논문에서의 실험 결과를 BD-rate^[9]로 표현하며 성능의 향상을 음의 값으로 표현하였다.

표 4. DST 기반 보간 필터를 적용한 실험 결과 (쌍방향 예측의 경우에만 적용)

Table 4. DST-IF experimental results for luma components (only applies to bi-prediction)

| Classes | Sequences | Low Delay B Pictures(%) | Random Access(%) |
|----------|-----------------|-------------------------|------------------|
| B | Kimono | 0.1 | 0.1 |
| | ParkScene | 0.2 | 0.1 |
| | Cactus | 0.2 | 0 |
| | BasketballDrive | 0 | 0 |
| | BQTerrace | 1.1 | 0.8 |
| C | BasketballDrill | -0.7 | -0.2 |
| | BQMall | -0.3 | -0.6 |
| | PartyScene | -1.5 | -2.4 |
| | RaceHorses | 0.2 | 0.2 |
| D | BasketballPass | -0.1 | -0.2 |
| | BQSquare | -3.7 | -5.0 |
| | PartyScene | -1.2 | -1.7 |
| | RaceHorses | 0 | -0.1 |
| E | FourPeople | 0.4 | |
| | Johnny | -0.4 | |
| | KristenAndSara | 0.2 | |
| Averages | Class B | 0.3 | 0.2 |
| | Class C | -0.5 | -0.8 |
| | Class D | -1.2 | -1.8 |
| | Class E | 0.1 | |
| | All | -0.3 | -0.7 |

2. 실험 결과

표 3에서 Random Access의 경우 0.6%의 BD-rate 성능 향상을 보였다. 또한, Low Delay B의 경우 0.1%의 BD-rate

성능 향상을 가져왔다. 특히 클래스 C에 속하는 BQSquare라는 시퀀스에서 Random Access와 Low Delay B에서 각각 4.1%와 5.2%의 BD-rate 감소를 가져왔다. 하지만 Low Delay P의 경우에는 B-E 클래스에서 PartyScene과 BQSquare 시퀀스들을 제외한 모든 시퀀스들의 BD-rate가 증가하였다. Low Delay P에서는 단방향 예측(Uni-prediction)^{[4][5]}으로 이루어져 있다.

단방향 예측과 다르게 쌍방향 예측(Bi-prediction)에서는 2개의 움직임 예측을 통하여 각각의 예측 값을 구할 수 있다. 2개의 예측 값을 평균을 취함으로써 저역 필터(Low-pass filter)의 효과가 가져와 고주파 성분들을 없앨 수 있다. 또한, DST-IF는 DCT-IF보다 더 강한 저역 필터의 기능을 하기 때문에 DST-IF를 쌍방향 예측인 경우 사용하면 DCT-IF를 사용하는 경우보다 좋아진다. 표 3의 실험 결과를 보았을 때, 단방향 예측의 경우 DST 기반 보간 필터의 성능 감소를 가져옴에 따라 단방향 예측의 경우 DCT 기반 보간 필터를 적용하고 쌍방향 예측^{[4][5]}의 경우에만 DST 기반 보간 필터를 적용하였다. 표 4는 쌍방향 예측의 경우에만 적용한 실험 결과이다.

표 4에서 Random Access의 경우 0.7%의 BD-rate 성능 향상을 보였다. 또한, Low Delay B의 경우 0.3%의 BD-rate 성능 향상을 가져왔다. 표 3의 결과와 유사하게 클래스 C에 속하는 BQSquare라는 시퀀스에서 Random Access와 Low Delay B에서 각각 5.0%와 3.7%의 BD-rate 감소를 가져왔다. Low Delay P의 경우 단방향 예측으로만 이루어져 있기 때문에 성능의 변화가 없었다.

V. 결 론

본 논문에서는 DST 기반 보간 필터를 적용하는 방법을

제안하였다. 하지만 단방향 예측의 경우 성능 감소를 가져옴에 따라 쌍방향의 경우에만 적용하는 방법을 함께 제시하였다. 단방향과 쌍방향 예측에 모두 적용한 DST 기반 보간 필터의 경우 Low Delay B와 Random Access에서 각각 0.1%와 0.6%의 성능 향상을 가져왔다. 쌍방향 예측에만 적용한 DST 기반 보간 필터의 경우 Low Delay B와 Random Access에서 각각 0.3%와 0.7%의 성능 향상을 가져왔다.

참 고 문 헌 (References)

- [1] B. Bros, W.-J. Han, J.-R. Ohm, G. J. Sullivan, Y.-K. Wang, T. Wiegand, "High Efficiency Video Coding (HEVC) text specification draft 10 (for FDIS & Consent)," document JCT-VC-L103, January 2013.
- [2] Gary J. Sullivan, Jens-Rainer Ohm, Woo-Jin Han, and Thomas Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard," IEEE Transactions on Circuits and Systems, No.12, pp.1649-1668, December 2012.
- [3] Kemal Ugur, Alexander Alshin, Elena Alshina, Frank Bossen, Woo-Jin Han, Jeong-Hoon Park, Jani Lainema, "Motion Compensated Prediction and Interpolation Filter Design in H.265/HEVC", IEEE in Signal Processing, Vol. 7, No.6, pp.946-956, December 2013.
- [4] Mathias Wien, *High Efficiency Video Coding - Coding Tools and Specification*, Springer, Berlin Heidelberg, 2015.
- [5] Vivienne Sze, Madhukar Budagavi, Gary J. Sullivan, *High Efficiency Video Coding (HEVC) - Algorithms and Architectures*, Springer, Switzerland, 2014.
- [6] Stitch (2013) Dicrete Sine Transform, <http://planetmath.org/sites/default/files/texpdf/39764.pdf> (accessed Sept. 1, 2016).
- [7] McClellan, James H., Ronald W. Schafer, and M. A. Yoder, *Signal Processing First*, Upper Saddle River, NJ: Pearson/Prentice Hall, pp.93-94, 2003.
- [8] Simon Haykin and Barry Van Veen, *Signals and Systems*, 2nd edition, Wiley, pp.234-235, 2003.
- [9] G. Bjøntegaard, "Calculation of Average PSNR Differences between RD-curves", ITU-T VCEG Meeting, Austin, TX, USA, Tech. Rep. SG16 Q.6 Doc., VCEG-M33, April 2001.
- [10] Frank Bossen, Common test conditions and software reference configurations, ITU-T/ISO/IEC Joint Collaborative Team on Video Coding (JCT-VC) document JCTVC-L1100, January 2013.

저자소개

김명준



- 2016년 8월 : 세종대학교 컴퓨터공학과 (학사)
- 2016년 9월 ~ 현재 : 세종대학교 컴퓨터공학과 (석사과정)
- ORCID : <http://orcid.org/0000-0002-4248-2285>
- 주관심분야 : Future Video Coding, HEVC(High Efficiency Video Coding), Image compression, Image processing

이영렬



- 1985년 2월 : 서강대학교 전자공학과 (학사)
- 1987년 2월 : 서강대학교 전자공학과 (석사)
- 1999년 2월 : 한국 과학기술원 전기·전자공학과 (박사)
- 1987년 1월 ~ 1994년 2월 : 삼성전자 R&D 센터 Digital Media Lab.
- 1999년 3월 ~ 2001년 8월 : 삼성전자 R&D 센터 Digital Media Lab. 수석연구원
- 2001년 9월 ~ 현재 : 세종대학교 전자정보대학 컴퓨터공학부 컴퓨터공학과 교수
- ORCID : <http://orcid.org/0000-0003-2709-8282>
- 주관심분야 : Future Video Coding, HEVC(High Efficiency Video Coding), Scalable Video Coding(SVC), Multi-view video coding(MVC), H.264/AVC, Image processing, Frame rate up-conversion, Transcoding, image transform