

<https://doi.org/10.7236/JIIBC.2019.19.6.21>
JIIBC 2019-6-4

주파수 영역에서의 움직임 예측을 위한 8×8 크기의 DCT 스케일링 행렬 정의

Definition of 8×8 sized DCT Scaling Matrix for Motion Estimation in the Frequency Domain

김혜빈*, 류철**

Hye-Bin Kim*, Chul Ryu**

요약 동영상 압축 표준은 고해상도의 영상을 위한 처리 기술이 요구되면서 영상의 해상도 증가에 맞춰 부호화 크기를 증가시켰다. 정확한 움직임 예측과 증가된 부호화 크기는 높은 정확도와 압축률을 제공하지만 계산량 증가 문제가 발생한다. 본 논문에서는 복잡도를 줄이기 위해 주파수 영역에서 이동 행렬을 이용한 DCT 기반 움직임 예측을 사용한다. 하지만 일반적인 동영상 부호화기에 사용되는 DCT와 양자화 과정을 주파수 영역의 부호화기에 그대로 적용했을 때 스케일링 과정으로 인한 문제점이 발생함을 발견하였다. 따라서 본 논문에서는 DCT 단계에서 적용할 수 있는 스케일링 행렬을 추출하여 이를 해결하고, 증가된 부호화 크기를 이용해 움직임 예측의 성능을 높였다.

Abstract The video compression standard required a processing technique for a high resolution image and increased the coding size to increase the resolution of the image. Accurate motion estimation and increased coding size provide high accuracy and compression rate, but there is a problem of increased computational complexity. In this paper, we use DCT - based motion estimation in the frequency domain to reduce complexity. However, we found that the DCT and quantization process used in a general video encoder are applied to the frequency domain encoder, resulting in problems caused by the scaling process. Therefore, in this paper, we extract the scaling matrix that can be applied in the DCT step and resolve the, and improve the performance of motion estimation using increased coding size.

Key Words : video compression, frequency domain, motion estimation, integer DCT, scaling matrix

1. 서 론

최근 디지털 신호 처리와 저장매체 및 전송 기술의 급

격한 발전은 음성, 문자에 국한된 서비스에서 DMB (digital multimedia broadcasting), IPTV(internet protocol television), HDTV(high definition

*준회원, 동국대학교 정보통신공학과 석사과정

**정회원, 동국대학교 정보통신공학과 교수

접수일자: 2019년 9월 30일, 수정완료: 2019년 11월 2일

게재확정일자: 2019년 12월 6일

Received: 30 September, 2019 / Revised: 2 November, 2019 / Accepted: 6 December, 2019

*Corresponding Author: cryu@dongguk.edu

Major in Information Communication Engineering, Dongguk University, Korea

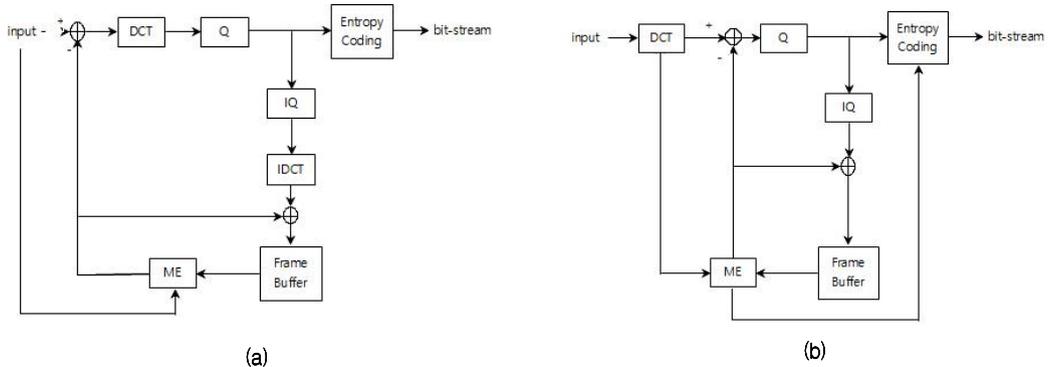


그림 1. (a) 동영상 압축 표준의 블록 기반 변환 부호화 구조 (b) 주파수 영역에서의 블록 기반 변환 부호화 구조
 Fig. 1. Block based transform coding architecture of video compression standard in spatial domain (b) Block based transform coding architecture in frequency domain

television) 등과 같이 다양한 매체를 이용한 멀티미디어 전송 서비스로 진화시켰다. 특히 HDTV가 일반화되면서 고화질, 고해상도의 영상을 위한 동영상 처리 기술들이 요구되고 있다^[1]. 그러나 고해상도 전송에 따른 데이터 송·수신의 증가로 한정된 대역폭의 통신환경에서 송·수신에 한계가 있어 고화질의 영상을 전송하기 위해 동영상 압축이 필수적으로 사용되고 있다^[2]. 동영상 압축 표준에는 MPEG-2, MPEG-4, H.263, H.264/AVC 등이 있다. 그 중 ITU-T의 VCEG와 ISO/IEC의 MPEG의 공동 작업으로 제정된 동영상 압축 표준인 H.264/AVC가 현재 가장 보편적인 포맷으로 사용되고 있다. H.264/AVC는 MPEG-4와 비교하여 2배 이상의 압축 효율을 얻을 수 있으며, 휴대폰과 같이 낮은 비트율을 요구하는 서비스부터 높은 비트 전송률이 필요한 HDTV 서비스까지 다양하게 이용될 수 있도록 규정하고 있다.

H.264/AVC를 포함한 동영상 압축 표준은 동영상 프레임의 연속된 시간 중복성을 제거하는 움직임 예측을 사용한다. 움직임 예측은 동영상 부호화기에서 전체 계산량의 60~80%를 차지하는 부분으로, 높은 압축률과 낮은 계산량을 갖는 움직임 예측 알고리즘을 개발하기 위해 계속해서 연구되고 있다^[3-4]. 움직임 예측의 기본이 되는 블록 정합 알고리즘은 이전 프레임에서 현재 프레임의 예측 대상 블록과 최소 정합 오차를 갖는 블록을 찾아내어 움직임 정보를 예측하는 방법이다. 움직임 예측으로 계산된 움직임 벡터와 예측 블록의 오차를 엔트로피 부호화하고, 비트 스트림으로 전송한다. 움직임 정보가 정확할수록 예측 블록의 오차가 감소하므로 시간적 중복성을 효과적으로 제거할 수 있다. 또한 영상의 움직임 정보가 정수 화소 단위로만 발생하지 않기 때문에 정수 화소 단위의 움직임 예측은 정확한 움직임 예측을 기대할 수

없다. 이러한 이유로 H.264/AVC는 부화소 단위의 움직임 예측을 적용하여 높은 움직임 예측 성능을 제공한다.

H.264/AVC 충실도 확장 규격에서는 해상도 증가에 맞춰 8×8 DCT가 추가됐다^[5]. 720×480인 SDTV와 비교하여 1920×1080의 해상도를 갖는 HDTV는 해상도의 증가로 이웃한 화소가 같거나 비슷한 화소 값을 갖기 쉽다. 따라서 8×8 DCT를 사용하면 비트스트림은 감소시킬 수 있다. 즉, 인접한 화소간의 값이 비슷할 때 부호화 블록의 크기를 증가시킬수록 압축률이 높아지기 때문에 나타나는 결과이다.

동영상 압축 표준에 사용되는 움직임 예측 알고리즘은 공간 영역에서 수행된다(그림 1-(a)). H.264/AVC의 확장 규격에서 추가된 8×8 DCT는 행렬 계산으로 인해 복잡도를 증가시킨다. 움직임 예측의 계산량 감소를 위해 주파수 영역에서 이동 행렬을 이용한 움직임 예측을 진행하는 방법이 제안되었다. 하지만 H.264/AVC에서 사용되는 DCT와 양자화 및 스케일링을 주파수 영역에서의 움직임 예측에 그대로 적용할 경우 문제점이 발생한다. 본 논문에서는 H.264/AVC의 8×8 DCT와 양자화 과정을 주파수 영역의 부호화기에 적용했을 때 어떤 문제점이 발생하는지를 정의하고, 이를 해결하기 위해 스케일링 행렬을 추출하여 주파수 영역에서 움직임 예측을 구현한다.

본 논문의 구성은 다음과 같다. 제2장에서는 주파수 영역에서의 움직임 예측 알고리즘을 설명하고, 제3장에서는 공간 영역의 DCT 및 양자화를 주파수 영역의 부호화기에 적용했을 때 문제점을 정의하고, 문제 해결을 위한 스케일링 행렬 정의 과정을 소개한다. 제4장은 실험을 통한 결과를 비교하였으며 끝으로 제5장에서 결론을 맺는다.

II. 주파수 영역에서의 움직임 예측

움직임 예측의 기본이 되는 블록 기반의 알고리즘은 크게 공간 영역에서의 움직임 예측과 주파수 영역에서의 움직임 예측으로 나눌 수 있다^[6-7]. 공간 영역에서의 움직임 예측은 이전 영상의 탐색 영역에서 현재 블록과 최소 오차를 갖는 블록을 찾는 방법으로, 탐색 영역의 블록을 계산하는 과정에서 많은 부호화 시간이 소비된다. 또한 움직임 정보의 정확도를 높이기 위해 보간법을 사용하게 되는데, 이는 부호화기의 계산량을 증가시키는 요인이 된다. 이를 해결하기 위해 주파수 영역에서의 움직임 예측 방법이 제안되었다. 다양한 주파수 영역에서의 움직임 예측 방법 중 본 논문에서는 DCT 기반 알고리즘을 사용한다^[8-10]. 기존 동영상 압축 표준이 DCT를 기반으로 있어 호환성이 좋고, DCT 영역에서 움직임 예측을 수행하므로 IDCT(inverse DCT) 과정을 생략하여 전체 부호화기의 계산량을 줄일 수 있다는 장점을 가지기 때문이다. 주파수 영역에서의 영상 부호화기 구조는 그림 1의 (b)에서 확인할 수 있다.

H.264/AVC는 기본 4×4 정수 DCT를 사용했으나, 해상도 증가로 인한 압축률을 낮추기 위해 충실도 확장 규격에서 8×8 정수 DCT가 도입되었다. 본 논문에서는 8×8 정수 DCT를 사용하며, 이에 맞게 이동 행렬과 움직임 예측 블록의 크기를 8×8 로 고정한다.

Plompen이 제안한 이동행렬 알고리즘을 이용하여 이전 블록과 이동 행렬의 곱으로 주파수 영역에서 움직임 예측을 수행할 수 있다^[10].

이동 행렬을 이용한 움직임 예측은 참조 프레임에서 예측한 f_{pred} 는 주변 4개 블록 f_0, f_1, f_2, f_3 의 수평, 수직 변환된 형태의 합으로 표현 가능하다. 예측 블록 f_{pred} 는 수평 이동행렬 H_i 와 수직 이동행렬 V_i 를 이용해 다음과 같이 나타낼 수 있다.

$$f_{pred} = \sum_{i=0}^3 V_i f_i H_i \quad (1)$$

$$V_0 = D_{8-\Delta x}, V_1 = D_{8-\Delta x}, V_2 = D_{\Delta x}^t, V_3 = D_{\Delta x}^t \quad (2)$$

$$H_0 = D_{8-\Delta y}^t, H_1 = D_{\Delta y}, H_2 = D_{8-\Delta y}^t, H_3 = D_{\Delta y} \quad (3)$$

식 (2)와 (3)에 사용된 변위 행렬 D_N 은

$$D_N = \begin{bmatrix} 0 & I_N \\ 0 & 0 \end{bmatrix} \quad (4)$$

이며, I_N 은 $N \times N$ 의 단위행렬이다.

DCT 변환의 직교성과 분리성을 이용하면 식 (1)을 다음과 같이 전개할 수 있다.

$$\begin{aligned} \hat{f}_i &= DCT(f_{pred}) = DCT\left(\sum_{i=0}^3 V_i f_i H_i\right) \\ &= \sum_{i=0}^3 DCT(V_i) DCT(f_i) DCT(H_i) \\ &= \sum_{i=0}^3 \hat{V}_i \hat{F}_i \hat{H}_i \end{aligned} \quad (5)$$

f_{pred} 의 DCT 계수를 구하기 위해서는 8개의 행렬 곱셈과 3개의 행렬 덧셈이 요구된다. 하지만 저주파 영역으로 에너지가 집중되는 DCT 변환 특성을 이용하면 연산량을 크게 줄일 수 있다. DCT 변환 후 양자화된 \hat{f}_i 의 특성에 기반을 두어 식 (5)를 아래 식 (6)과 같이 유도할 수 있다. DCT 변환한 계수 값은 저주파 대역에 에너지가 집중되고 고주파 대역의 계수들이 대부분 0 값을 갖는다는 특성을 이용하면, $\hat{f}_i(m, n)$ 이 0 값을 갖는 외적의 경우 계산 생략이 가능하기 때문에 f_{pred} 를 효과적으로 계산할 수 있다. $\hat{f}_i(m, n)$ 의 계수들은 식 (6)과 같이 지그재그 스캔 순서로 정리될 수 있다.

$$\begin{aligned} \hat{f}_{pred} &= \sum_{i=0}^3 \hat{V}_i \hat{F}_i \hat{H}_i = \sum_{m=0}^7 \sum_{n=0}^7 \sum_{i=0}^3 \hat{f}_i(m, n) \hat{V}_i(m) \hat{H}_i(n) \\ &= \hat{f}_0(0,0) \hat{V}_0(0) \hat{H}_0(0) + \hat{f}_1(0,0) \hat{V}_1(0) \hat{H}_1(0) \\ &\quad + \hat{f}_2(0,0) \hat{V}_2(0) \hat{H}_2(0) + \hat{f}_3(0,0) \hat{V}_3(0) \hat{H}_3(0) \\ &\quad + \hat{f}_0(0,1) \hat{V}_0(0) \hat{H}_0(1) + \hat{f}_1(0,1) \hat{V}_1(0) \hat{H}_1(1) \\ &\quad + \hat{f}_2(0,1) \hat{V}_2(0) \hat{H}_2(1) + \hat{f}_3(0,0) \hat{V}_3(0) \hat{H}_3(1) \\ &\quad + \dots \\ &\quad + \hat{f}_0(7,7) \hat{V}_0(3) \hat{H}_0(3) + \hat{f}_1(7,7) \hat{V}_1(3) \hat{H}_1(3) \\ &\quad + \hat{f}_2(7,7) \hat{V}_2(3) \hat{H}_2(3) + \hat{f}_3(7,7) \hat{V}_3(3) \hat{H}_3(3) \end{aligned} \quad (6)$$

표 1. 양자화 변수에 따른 스케일링 계수 (8×8)

Table 1. Scaling factor according to the quantization parameter (8×8)

QP mod 6	positions					other positions
	(0,0),(0,4),(4,0),(4,4)	(1,1),(1,3),(1,5),(1,7), (3,1),(3,3),(3,5),(3,7), (5,1),(5,3),(5,5),(5,7), (7,1),(7,3),(7,5),(7,7)	(2,2),(2,6),(6,2),(6,6)	(0,1),(0,3),(0,5),(0,7), (1,0),(1,4),(3,0),(3,4), (4,1),(4,3),(4,5),(4,7), (5,0),(5,4),(7,0),(7,4)	(0,2),(0,6),(2,0),(2,4), (4,2),(4,6),(6,0),(6,4)	
0	13107	11428	20972	12222	16777	15481
1	11916	10826	19174	11058	14980	14290
2	10082	8943	15978	9675	12710	11985
3	9362	8228	14913	8931	11984	11259
4	8192	7346	13159	7740	10486	9777
5	7282	6428	11570	6830	9118	8640

III. 스케일링 행렬 추출

1. 정수 변환 및 양자화

일반적인 동영상 압축 표준에서는 영상의 화소 간 상관관계를 없애고, 저주파와 고주파 성분으로 분리하여 저주파 대역에 에너지가 집중되는 특성을 갖는 DCT 변환 부호화 방법을 사용한다. 하지만 DCT 연산 과정에서 부동 소수점 계수 오류가 발생하여 DCT와 IDCT를 거친 결과가 소수점 계산으로 인한 오차로 원본 영상과의 차이가 발생한다. 이를 해결하기 위해 H.264/AVC에서는 정수 DCT를 사용한다. H.264/AVC는 기본으로 4×4 정수 DCT를 사용하지만, 해상도가 증가하면서 계산량도 증가하는 것을 해결하기 위해 본 논문에서는 8×8 정수 DCT를 고려한다. 일반적으로 정수 DCT는 2차원 변환 행렬을 이용하며, 8×8 정수 DCT 변환 식과 변환 행렬 T 는 아래와 같다.

$$f = T \cdot x \cdot T^T \quad (7)$$

$$T = \frac{1}{8} \times \begin{bmatrix} 8 & 8 & 8 & 8 & 8 & 8 & 8 & 8 \\ 12 & 10 & 6 & 3 & -3 & -6 & -10 & -12 \\ 8 & 4 & -4 & -8 & -8 & -4 & 4 & 8 \\ 10 & -3 & -12 & -6 & 6 & 12 & 3 & -10 \\ 8 & -8 & -8 & 8 & 8 & -8 & -8 & 8 \\ 6 & -12 & 3 & 10 & -10 & -3 & 12 & -6 \\ 4 & -8 & 8 & -4 & -4 & 8 & -8 & 4 \\ 3 & -6 & 10 & -12 & 12 & -10 & 6 & -3 \end{bmatrix} \quad (8)$$

식 (7)의 x 는 입력을 의미하며, T^T 는 변환 행렬 T 의 전치행렬을 의미한다.

그림 1의 (a) 동영상 부호화기의 구조에서 입력 영상은 DCT 변환 이후 양자화(Q)가 진행되고, 양자화 단계

에서 식 (9)와 같이 스케일링이 진행된다. 스케일링은 양자화 변수 QP 값에 따른 양자화 행렬 계수의 변경에 필요한 과정이다.

$$Z_{ij} = \text{round} \left(f_{ij} \cdot \frac{MF}{2^{q \text{ bits}}} \right), \quad q = 15 + \text{floor} \left(\frac{QP}{6} \right) \quad (9)$$

Z_{ij} 는 양자화 블록을 의미한다. MF 는 스케일링을 위한 행렬이고, QP 를 6으로 나눈 나머지 값에 의해 위치에 따라 계수가 달라진다. 위치에 따른 스케일링 행렬 MF 는 표 1에서 확인할 수 있다. q 는 QP 에 따른 비트 길이를 뜻한다.

이와 같은 정수 변환 및 양자화는 공간 영역에서 수행되는 과정이다. 이를 부호화기 구조가 다른 주파수 영역에서의 부호화기에 적용하면 문제가 발생한다. 공간 영역과 다르게 주파수 영역 부호화기는 역양자화(IQ)까지 적용된 참조 프레임과 DCT만 적용된 현재 프레임을 이용해 움직임 예측을 진행한다(그림 1-(b)). 이 때, 스케일링은 양자화(또는 역양자화)에서 진행되므로 참조 프레임과 현재 프레임 간의 스케일링 연관성이 사라지게 되어 정확한 움직임 예측을 기대할 수 없다. 따라서 움직임 예측에 사용되는 두 프레임의 스케일링을 일치시켜야 하며, 이를 위해 DCT 단계에서 스케일링이 수행될 수 있도록 구조를 변경하는 과정을 다음 절에서 제안한다.

2. DCT 스케일링 행렬 정의

그림 1의 (b) 구조를 수행하기 위해 이전 프레임과 현재 프레임에 동일한 스케일링을 적용해야 한다. 이를 위해 양자화 단계에서 진행되는 스케일링을 DCT 단계에서 진

행될 수 있도록 부호화 구조가 변경되어야 한다. DCT 단계에 적용될 스케일링 행렬을 정의하기 위해 양자화 단계에 적용될 스케일링 구조를 다음과 같이 표현할 수 있다^[11].

$$Z_{ij} = \text{round}(F_{ij} \cdot \frac{PF}{Qstep}) \quad (10)$$

$$\frac{MF}{2^{q \text{ bits}}} = \frac{PF}{Qstep} \quad (11)$$

식 (10)은 양자화 구조를 변형한 식으로, 식 (11)을 도출할 수 있다. 식 (11)의 좌변은 분자와 분모가 QP 에 따라 값이 변하기 때문에 양자화 과정에서 진행되어야 한다. 우변의 $Qstep$ 은 양자화 변수 값에 따라 달라지는 반면, PF 는 고정된 계수로 양자화 변수에 영향을 받지 않기 때문에 DCT 단계에 적용할 수 있다. 따라서 DCT에 적용될 스케일링 행렬 계수 PF 는 아래와 같이 정리할 수 있다. PF 의 계산을 위해서 $QP=0$ 으로 설정했다.

$$PF_{8 \times 8} = \frac{MF_{8 \times 8}}{2^{q \text{ bits}}} \cdot Qstep$$

$$= \begin{bmatrix} 0.2500 & 0.2331 & 0.3200 & 0.2331 & 0.2500 & 0.2331 & 0.3200 & 0.2331 \\ 0.2331 & 0.2180 & 0.2953 & 0.2180 & 0.2331 & 0.2180 & 0.2953 & 0.2180 \\ 0.3200 & 0.2953 & 0.4000 & 0.2953 & 0.3200 & 0.2953 & 0.4000 & 0.2953 \\ 0.2331 & 0.2180 & 0.2953 & 0.2180 & 0.2331 & 0.2180 & 0.2953 & 0.2180 \\ 0.2500 & 0.2331 & 0.3200 & 0.2331 & 0.2500 & 0.2331 & 0.3200 & 0.2331 \\ 0.2331 & 0.2180 & 0.2953 & 0.2180 & 0.2331 & 0.2180 & 0.2953 & 0.2180 \\ 0.3200 & 0.2953 & 0.4000 & 0.2953 & 0.3200 & 0.2953 & 0.4000 & 0.2953 \\ 0.2331 & 0.2180 & 0.2953 & 0.2180 & 0.2331 & 0.2180 & 0.2953 & 0.2180 \end{bmatrix} \quad (12)$$

이를 DCT와 양자화에 적용하여 식 (7)과 (9)는 다음과 같이 나타낼 수 있다.

$$F = (T \cdot x \cdot T^T) \cdot PF \quad (13)$$

$$Z_{ij} = \text{round}(F_{ij} \cdot \frac{1}{Qstep}) \quad (14)$$

IV. 실험 및 결과

주파수 영역에서의 움직임 예측을 위해 DCT와 양자화 단계에 변경된 구조를 적용하고, 이에 대한 결과를 확인하기 위해 모든 실험은 H.264/AVC 확장 규격인 8×8 정수 DCT를 기반으로 한다. 스케일링의 순서에 따른 결과를 비교하기 위해 공간 영역과 주파수 영역 부호화기의 DCT 또는 양자화 단계에 스케일링을 적용하였다. 구현을 위해 CIF 해상도를 갖는 flower 영상을 사용

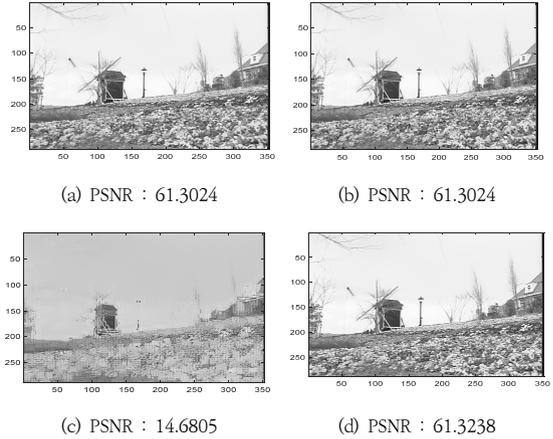


그림 2. 스케일링 순서에 따른 움직임 예측 화질 비교
 Fig. 2. Comparison of motion prediction quality due to scaling order

하였고, 공간 영역과 주파수 영역에서 이중 선형 보간법을 이용한 반 화소 단위 움직임 예측을 사용하였다^[12]. 스케일링 순서에 따른 결과를 확인하기 위해 양자화에서의 스케일링은 식 (7) ~ (9)를 이용했고, DCT 단계에서의 스케일링은 식 (12) ~ (14)를 사용했다.

그림 2는 스케일링에 따른 $QP=9$ 일 때의 해상도와 PSNR을 보여준다. 그림 2의 (a)와 (b)는 공간 영역, (c)와 (d)는 주파수 영역에서의 움직임 예측 결과로, (a)와 (c)는 양자화 단계에서 스케일링을 진행했고 (b)와 (d)는 DCT 단계에서 스케일링을 적용한 결과이다. 공간 영역에서의 움직임 예측은 각각 DCT나 양자화 과정에서 스케일링을 진행한 결과의 해상도 및 PSNR이 동일하다. 이는 도출된 스케일링 행렬 PF 를 DCT 단계에서 스케일링을 진행하는 것이 양자화 단계에서의 스케일링과 동일한 결과를 얻을 수 있음을 의미한다. 주파수 영역에서의 움직임 예측 결과에서 양자화 단계에서 스케일링을 진행한 그림 3의 (c)는 어느 정도의 윤곽은 표현할 수 있지만, 올바른 움직임 예측이 진행되지 않음을 확인할 수 있다. 스케일링 행렬 PF 를 사용하여 DCT에서 스케일링을 진행한 주파수 영역에서의 움직임 예측 결과인 (d)를 확인하면 공간 영역에서의 결과보다 더 높은 PSNR을 얻을 수 있다.

표 2는 CIF 해상도의 flower 영상의 1, 2번째 프레임을 사용하여 양자화 변수 값에 따른 블록 크기별 공간 영역 및 주파수 영역에서의 움직임 예측 결과를 PSNR과 비트 수로 나타내었다. flower 영상물의 움직임 예측의 결과는 공간 영역과 주파수 영역에서 PSNR은 비슷하지

표 2. flower 시퀀스 이용한 QP 조정에 따른 공간 영역과 주파수 영역에서의 블록 크기 별 움직임 예측 성능 비교
 Table 2. Motion estimation performance ratio by block size in spatial domain and frequency domain due to QP using flower sequence

QP		3	6	9	12	15	18	21	24	27	30
공간 영역 8×8	PSNR	91.20	70.58	61.30	55.95	53.20	51.08	48.85	46.30	43.77	41.01
	bits	144,982	128,239	112,457	96,221	81,311	68,103	58,170	49,840	42,459	36,317
주파수 영역 4×4	PSNR	61.23	56.38	53.61	51.40	49.14	46.53	43.99	41.20	38.62	35.88
	bits	373,629	308,694	257,869	218,251	187,183	158,817	133,970	111,339	89,427	67,133
주파수 영역 8×8	PSNR	93.42	70.43	61.32	56.08	53.40	51.19	48.93	46.38	43.82	41.06
	bits	143,762	126,491	110,014	93,595	78,330	65,654	56,444	48,508	41,894	35,388

만 최소 565에서 최대 2,991의 비트 수 감소를 확인할 수 있고, 4×4 블록 크기를 이용할 때보다 8×8 블록 크기 일 때 더 많은 비트 수 감소가 있었다. 영상의 정보에 따라 달라지지만 QP 값에 따른 결과를 확인했을 때, 주파수 영역에서 8×8 크기를 이용한 움직임 예측의 결과가 더 많은 비트 수의 감소를 확인할 수 있다.

V. 결론

H.264/AVC의 충실도 확장 규격에 기반하여 8×8 DCT를 적용한 주파수 영역에서 움직임 예측을 수행하는데 있어 스케일링의 순서가 움직임 예측에 영향을 끼친다는 것을 확인하였다. 움직임 예측에 사용될 두 프레임의 스케일링 연관성을 일치시키기 위해 DCT 단계에 적용될 스케일링 행렬을 정의하였다. 다양한 비디오 시퀀스를 이용한 결과물의 해상도와 PSNR을 고려했을 때, 주파수 영역에서 움직임 예측을 진행하기 위해서는 DCT 단계에서 스케일링이 진행되어야 하는 것을 확인하였다. 또한 주파수 영역에서 8×8 크기의 DCT와 움직임 예측이 대부분의 QP 값에서 PSNR이 높으면서 부호화 효율이 높음을 확인하였다.

References

- [1] B. Ye and B. Lju, "Rapid scene analysis on compressed video", IEEE Trans. Circuits and Systems for video Technology, Vol. 5, No. 6, pp. 533-540, 1995.
- [2] F. Dufaux and F. moscheni, "Motion Estimation Technique for Digital TV : A review and a new contribution," Proceedings of the IEEE, Vol. 83, pp. 858-876, 1995.
- [3] J. Oh, "Improved sub-block matching algorithm," J. KICS, Vol. 33, No. 7, pp. 628-633, Jul. 2010.
- [4] D. Song, "A Study on Fast Macroblock Partition Decision Method at H264," The Journal of the Institute of Internet, Broadcasting and Communication, Vol. 14, No. 6, pp. 99-105, 2014.
DOI : <http://dx.doi.org/10.7236/JIIBC.2014.14.2.43>
- [5] D. Marpe, T. Wiegand, and G. J. Sullivan, "The H.264/MPEG4 advanced video coding standard and its applications," IEEE Communications Magazine, Vol. 44, No. 8, pp. 134-144, Aug. 2006.
- [6] Y. Keller and A. Averbuch, "Fast gradient methods based on global motion estimation for video compression," IEEE Trans. Circuits and Systems for Video Technology, Vol. 13, No. 4, pp. 300-309, Apr. 2003.
- [7] S. Mallat, "A theory for multiresolution signal decomposition : The wavelet representation," IEEE Trans. Patter Analysis Mach. Interll, Vol. 11, pp. 674-693, Jul. 1989.
- [8] N. Merhav and V. Bhaskaran, "A fast algorithm for DCT-domain inverse motion compensation," IEEE International Conference on Acoustics, Speech and Signal Processing, Vol. 4, pp. 2307-2310, May. 1996.
- [9] J. Song and B. Yeo, "A fast algorithm for DCT-domain inverse motion compensation based on shared information in a macroblock," IEEE Trans. on Circuits and Systems for Video Technology, Vol. 10, No. 5, pp. 767-775, Aug. 2000.
- [10] R. H. J. M. Plompen, B. F. Schuurink and J. Biemond, "A new motion-compensated transform coding scheme", in Proc. IEEE Int. Conf. Acoust, Speech, Signal Process, Vol. 1, pp. 371-374, 1985.
- [11] Malyar, Henrique S., Hallapuro A., Karczewicz M., and Kerofsky L., "Low-complexity transform and quantization in H.264/AVC", IEEE Transactions on circuits and systems for video technology, Vol. 13, No. 7, pp. 595-603.

- [12] M. Kang, J. Heo, and C. Ryu, "Half-pixel accuracy motion estimation algorithm in the transform domain for H.264", J.KICS, Vol. 33, No. 11, pp. 917-924, Nov. 2011.

저 자 소 개

김 혜 빈(준회원)



- 2018년 2월 : 동국대학교 정보통신공학과 (공학사)
- 2019년 현재 : 동국대학교 정보통신공학과 석사 재학
- 주관심분야 : 영상 처리, 데이터 압축

• Email : halloqls@dongguk.edu

류 철(정회원)



- 1989년 : Florida Institute of Technology 컴퓨터공학과 (공학사)
- 1991년 : Polytechnic Institute of NYU 전기공학과 (공학석사)
- 1997년 : Polytechnic Institute of NYU 전기공학과 (공학박사)

- 1998년 ~ 1999년 : LG 정보통신연구소 선임연구원
- 1999년 ~ 현재 : 동국대학교 정보통신공학과 교수
- 주관심분야 : Visual Communication, High Resolution Reconstruction, Joint Source Coding
- Email : cryu@dongguk.edu