

Shifting Matrix를 이용한 DCT 기반 부화소 단위 움직임 예측 알고리즘

신재영*, 류철°

DCT-Based Subpixel-Accuracy Motion Estimation Utilizing Shifting Matrix

Jae-young Shin*, Chul Ryu°

요약

최근 동영상 압축 표준은 세밀한 움직임 정보를 확보하기 위해 1/4 화소 단위의 움직임 예측 알고리즘을 사용하고 있다. 일반적인 동영상 부호화기에서 사용하는 공간 영역에서의 움직임 예측은 부화소 단위 움직임 예측을 위한 보간 기술의 적용으로 인한 계산량 증가 문제가 발생한다. 본 논문에서는 주파수 영역에서 shifting matrix를 이용한 부화소 정밀도의 움직임 예측 알고리즘을 제안한다. 주파수 영역에서 shifting matrix 알고리즘을 사용함으로써 낮은 계산량으로 부화소 움직임 예측을 수행할 수 있었으며, 실험 결과 공간 영역에서의 움직임 예측 알고리즘에 비해 낮은 비트량과 높은 PSNR(peak signal-to-noise ratio)을 제공함을 확인하였다.

Key Words : motion estimation, shifting matrix, H.264

ABSTRACT

The latest video compression standard (such as H.264/AVC and HEVC) utilizes quarter-pel accuracy motion estimation in order to retain detailed motion information. Many sub-pixel motion estimation algorithms used in the spatial domain usually encounters increment of computational complexity due to embedded interpolation algorithm. In this paper, an approach to measure sub-pixel accuracy motion estimation in frequency domain using shifting matrix is proposed. Complexity can be reduced utilizing shifting matrix algorithm in frequency domain and simulation results demonstrate not only higher PSNR but lower bit rates than spatial domain algorithms.

I. 서론

최근 이동통신 단말기의 발전으로 인해 영상과 음향 서비스를 무선 환경에서도 즐길 수 있게 됨에 따라 사용자들은 무선 환경에서도 고품질의 동영상 서비스를 즐길 수 있길 기대한다. 고품질의 동영상을 한정된 대역폭의 통신환경에서 송수신하기엔 용량이 매우 방대하므로 사용자의 요구조건을 만족시키기 위해서는

동영상 압축이 필수적으로 사용되고 있다. ITU-T와 ISO에서 공동으로 제안한 동영상 압축 표준인 H.264/AVC는 현재 가장 널리 쓰이고 있는 코덱으로, 이전의 동영상 압축 표준에 비해 절반 이하의 비트율로 더 좋은 화질을 얻을 수 있도록 개발되었다. 데스크톱이나 노트북 뿐 만 아니라 디지털카메라, 휴대폰, CCTV등에도 H.264 기술이 널리 활용되고 있다.

동영상 압축은 기본적으로 공간적인 중복성과 시간

* First Author : Dongguk University Department of Information Communication Engineering, jyshin702@gmail.com, 학생회원

° Corresponding Author : Dongguk University Department of Information Communication Engineering, cryu@dgu.ac.kr, 종신회원

논문번호 : KICS2014-11-468, Received November 21, 2014; Revised January 29, 2015; Accepted January 29, 2015

적인 중복성을 제거하는 기술이 활용되며 그 중 영상 내 연속된 프레임 사이의 시간적인 중복성을 제거하는 방법으로 움직임 예측 기술을 사용한다. 움직임 예측은 동영상 부호화기에서 전체 계산량의 60-80%를 차지하는 부분이므로 뛰어난 압축 효율과 낮은 계산량을 갖는 움직임 예측 알고리즘을 개발하기 위해 끊임없이 연구되고 있다¹¹. 움직임 예측에서 기본이 되는 블록 정합 알고리즘은 현재 영상의 각 블록들의 탐색 영역에서 최소 정합 오차를 갖는 블록을 이전 프레임에서 찾아내어 움직임 정보를 예측하는 방법이다. 동영상 압축 시스템은 움직임 정보와 예측 블록의 오차만을 엔트로피 부호화하여 비트스트림으로 전송하게 되는데, 움직임 정보가 정확할수록 예측 블록의 오차가 줄어들어 시간적 중복성을 효과적으로 제거할 수 있다.

움직임 정보는 프레임간의 밝기 변화, 비선형성 움직임, 물체간의 경계 등으로 인하여 정수화소 단위로만 발생되지 않으므로 정수 화소 단위의 움직임 예측만으로는 정확한 움직임 정보를 기대할 수 없다. 부화소 단위 움직임 예측은 정수 화소 단위 움직임 예측에 비해 부화소 단위 움직임 정보까지 예측할 수 있으므로 보다 정확한 움직임 예측이 가능하다. 이러한 이유로 MPEG-1, 2와 H.263은 1/2 화소 단위의 움직임 예측을 사용하였고, H.264에서는 1/4 화소 단위까지 확장하여 기존 동영상 부호화 표준보다 움직임 예측 성능을 높였다.

일반적인 동영상 부호화기의 움직임 예측 알고리즘은 공간 영역에서 수행된다. 그러나 공간 영역에서의 부화소 단위 움직임 예측의 경우 보간 기술의 적용으로 인해 동영상 부호화기의 전체 계산량이 증가된다. 본 논문에서는 이러한 문제를 해결하고자 주파수 영역에서 shifting matrix를 사용하여 계산량을 줄일 수 있는 방법을 제안한다. 본 논문의 구성은 다음과 같다. 제2장에서는 주파수 영역에서의 움직임 예측 알고리즘에 대해 설명하고, 제3장에서는 1/4 화소 단위의 움직임 예측을 위한 shifting matrix 알고리즘을 소개한다. 제4장은 실험을 통하여 공간 영역과 주파수 영역에서의 움직임 예측 성능을 비교하였으며 제5장에서는 본 논문이 제안한 방식의 활용 방안에 대해 설명한다.

II. 주파수 영역에서의 움직임 예측

동영상의 부호화와 복호화는 모두 블록 단위로 이루어지므로 움직임 예측 또한 블록 기반의 알고리즘을 사용하고 있다. 블록 기반의 알고리즘은 크게 공간

영역에서의 알고리즘과 주파수 영역에서의 알고리즘으로 나뉠 수 있다. 공간 영역에서의 알고리즘은 블록 정합 알고리즘과 그래디언트 기반 알고리즘 등으로 구성되며 주파수 영역에서의 알고리즘은 DCT(discrete cosine transform) 혹은 웨이블릿 기반 움직임 예측 알고리즘 등으로 구성된다²⁻⁶. 주파수 영역에서의 알고리즘 중 웨이블릿 기반 알고리즘은 탐색과 정합시간이 적게 소모되지만 모든 부대역에서 움직임을 예측하므로 높은 계산량을 가지며 모든 부대역에 대한 움직임 정보를 전송해야 하므로 전송해야 하는 비트량이 증가한다는 단점을 가진다. 반면 DCT 기반 알고리즘은 기존 동영상 압축 표준이 DCT를 기반으로 한다는 점에서 호환성이 좋고 DCT 계수를 사용하여 움직임 예측을 수행하기 때문에 IDCT(inverse DCT) 과정을 생략할 수 있다는 장점을 가진다.

영상 압축에서 널리 사용되고 있는 방법인 DCT 변환은 공간 영역의 정보를 주파수 영역으로 변환시키는 기술로 에너지가 저주파대에 집중된다는 특성이 있어 영상 압축에 효율적이다. 일반적으로 동영상 부호화기는 공간 영역에서의 움직임 예측을 사용하는데, 참조 프레임으로 사용되는 복원 프레임의 정보를 공간 영역으로 저장해야하므로 IDCT 과정이 필요하다. 그러나 주파수 영역에서의 움직임 예측은 공간 영역의 정보들을 DCT 변환하여 움직임 예측을 수행하므로 IDCT 변환 과정이 생략될 수 있다⁷. 이는 공간 영역에서의 움직임 예측 방법에 비해 간소화된 부호화기의 구조로 인해 계산량이 낮아진다는 장점을 가지며 공간 영역과 주파수 영역에서의 움직임 예측을 사용하는 동영상 부호화기 구조는 그림 1과 같이 나타낼 수 있다.

Plompen이 제안한 shifting matrix 알고리즘은 예측 블록 f_{pred} 를 포함한 주변 4개 블록의 수평, 수직 변환된 형태의 합으로 나타낼 수 있음을 보여준다⁸. 그림 2는 현재 블록과 가장 비슷한 블록 f_{pred} 를 예측하기 위해 4개의 블록 f_0, f_1, f_2, f_3 이 사용되는 과정을 보여준다.

예측 블록 f_{pred} 는 수평, 수직 shifting matrix인 h_i 와 v_i 를 이용하여 다음과 같이 나타낼 수 있다.

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

h_i 와 v_i 는 다음과 같이 정의되며,

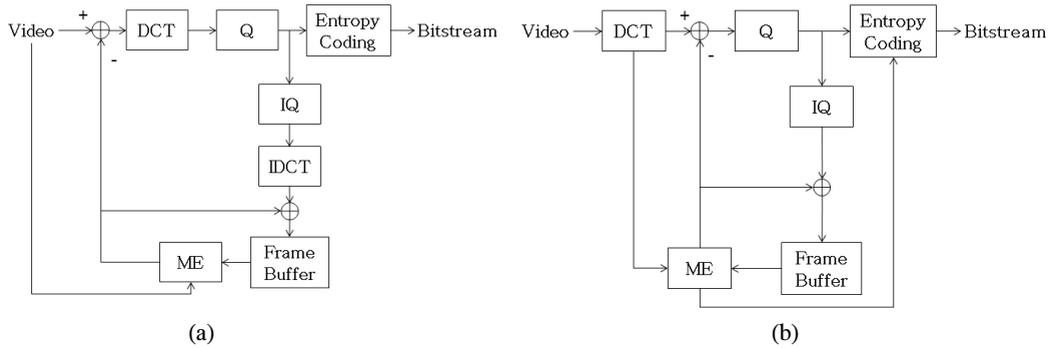


그림 1. (a) 공간 영역에서의 동영상 부호화기 구조 (b) 주파수 영역에서의 동영상 부호화기 구조
 Fig. 1. (a) ME/MC in spatial domain video coder architecture (b) ME/MC in transform domain video coder architecture

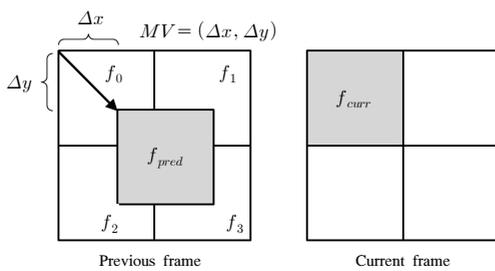


그림 2. Shifting matrix를 이용한 움직임 예측
 Fig. 2. Motion estimation using shifting matrix algorithm

$$\begin{aligned} h_0 &= D_{4-\Delta x}^t, & h_1 &= D_{\Delta x}, \\ h_2 &= D_{4-\Delta x}^t, & h_3 &= D_{\Delta x} \end{aligned} \quad (2)$$

$$\begin{aligned} v_0 &= D_{4-\Delta y}, & v_1 &= D_{4-\Delta y}, \\ v_2 &= D_{\Delta y}^t, & v_3 &= D_{\Delta y}^t \end{aligned} \quad (3)$$

변위행렬 D_n 은

$$D_n = \begin{bmatrix} 0 & I_n \\ 0 & 0 \end{bmatrix} \quad (4)$$

이고, I_n 은 $n \times n$ 의 단위행렬이다. 예를 들어, $\Delta x = 2, \Delta y = 0$ 일 경우 그림 3은 f_{pred} 가 예측되는 과정을 보여주며 식 (5)와 (6)은 사용된 h_i 를 나타낸다.

$$h_0 = h_2 = D_2^t = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (5)$$

$$h_1 = h_3 = D_2 = \begin{bmatrix} 0 & I_2 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (6)$$

DCT 변환의 직교성과 분리성을 이용하면 식 (1)은 다음과 같이 전개될 수 있다.

$$\begin{aligned} F_{pred} &= DCT(f_{pred}) \\ &= \sum_{i=0}^3 DCT(v_i) DCT(f_i) DCT(h_i) \\ &= \sum_{i=0}^3 (Tv_i T^t)(Tf_i T^t)(Th_i T^t) \\ &= \sum_{i=0}^3 V_i F_i H \end{aligned} \quad (7)$$

예측 블록의 DCT 계수 F_{pred} 를 구하기 위해서는 8개의 행렬 곱셈과 3개의 행렬 덧셈이 요구된다. 하지만 저주파대에 에너지가 집중되는 DCT 변환 특성을 이용하면 많은 연산량을 감소시킬 수 있다. 따라서 DCT 변환 후 양자화 된 블록 F_i 의 특성을 두어 계산상의 복잡도를 줄이기 위해 식 (7)을 다음과 같이 유도할 수 있다.

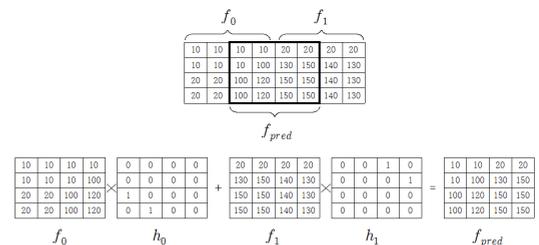


그림 3. $\Delta x = 2, \Delta y = 0$ 일 때 f_{pred} 의 예측 수행 과정
 Fig. 3. Estimation process of f_{pred} when $\Delta x = 2, \Delta y = 0$

$$\begin{aligned}
 F_{pred} &= \sum_{i=0}^3 V_i F_i H_i \\
 &= \sum_{i=0}^3 \sum_{m=0}^3 \sum_{n=0}^3 F_i(m,n) V_i(m) H_i(n) \quad (8) \\
 &= \sum_{m=0}^3 \sum_{n=0}^3 \sum_{i=0}^3 F_i(m,n) V_i(m) H_i(n)
 \end{aligned}$$

식 (8)에서 $F_i(m,n)$ 이 0값을 가질 경우 계산에서 제외될 수 있기 때문에 F_i 의 희소성을 이용하여 예측 블록을 효율적으로 계산할 수 있다. 따라서 일반적인 동영상 압축 알고리즘에서와 마찬가지로 DCT 변환 후 저주파 대역에 에너지가 집중되어 고주파 대역에 있는 계수들이 대부분 0값을 갖는다는 특성을 이용하여 F_i 의 계수들을 식 (9)와 같이 지그재그 스캔 순서로 정리할 필요가 있다.

$$\begin{aligned}
 F_{pred} &= \sum_{m=0}^3 \sum_{n=0}^3 \sum_{i=0}^3 F_i(m,n) H_i(m) V_i(n) \\
 &= F_0(0,0) V_0(0) H_0(0) + F_1(0,0) V_1(0) H_1(0) \\
 &\quad + F_2(0,0) V_2(0) H_2(0) + F_3(0,0) V_3(0) H_3(0) \\
 &\quad + F_0(0,1) V_0(0) H_0(1) + F_1(0,1) V_1(0) H_1(1) \\
 &\quad + F_2(0,1) V_2(0) H_2(1) + F_3(0,1) V_3(0) H_3(1) \\
 &\quad + F_0(1,0) V_0(1) H_0(0) + F_1(1,0) V_1(1) H_1(0) \\
 &\quad + F_2(1,0) V_2(1) H_2(0) + F_3(1,0) V_3(1) H_3(0) \\
 &\quad + \dots \\
 &\quad + F_0(3,3) V_0(3) H_0(3) + F_1(3,3) V_1(3) H_1(3) \\
 &\quad + F_2(3,3) V_2(3) H_2(3) + F_3(3,3) V_3(3) H_3(3) \quad (9)
 \end{aligned}$$

식 (9)에서 유도된 재귀방정식은 다른 주파수 대역을 갖는 영상에서 F_i 의 개수를 조절하여 효율적인 연산을 할 수 있으며, 블록의 크기와 F_i 에서 0이 아닌 성분의 개수에 따라 연산량을 효과적으로 줄일 수 있다는 장점을 가진다. 식 (9)에서 0이 아닌 계수만이 V_i, H_i 와의 외적을 수행하게 되며, 최종적으로 다음과 같이 나타낼 수 있다

$$F_{pred} = F_{pred}^{k-1} + F_i(m,n) V_i(m) H_i(n) \quad (10)$$

여기서 F_{pred}^0 는 4×4 의 영행렬이며, k 는 블록 내의 1부터 0이 아닌 DCT 계수들의 전체 개수까지 증가한다.

III. 부화소 움직임 예측을 위한 Shifting Matrix

공간 영역에서의 부화소 움직임 예측은 보간 기술을 사용하여 영상을 2배 또는 그 이상으로 확대시키는 과정을 거친 후 움직임 예측이 수행되며 bilinear 보간법을 사용한 반화소 단위 움직임 예측의 경우 보간 성능이 비교적 낮기 때문에 예측 성능이 낮다⁹⁾. 반면 b-spline 또는 cubic 보간법의 경우 식 (11), (12)와 같은 3차원 함수로 부화소를 예측해내는 방법으로 bilinear 보간법에 비해 보간된 값이 정확하여 예측 성능이 비교적 높지만 계산량이 증가한다는 단점이 있다. H.264 표준에서는 1/4 화소 단위 움직임 예측을 통해 영상의 화질을 향상시킬 수 있지만 보간 기술의 적용으로 인해 움직임 예측 과정의 계산량이 증대되어 동영상 부호화 시간에 영향을 주게 된다.

$$f(x) = \begin{cases} \frac{1}{2}|x|^3 - |x|^2 + \frac{2}{3}, & 0 \leq |x| < 1 \\ -\frac{1}{6}|x|^3 + |x|^2 - 2|x| + \frac{4}{3}, & 1 \leq |x| < 2 \\ 0, & 2 \leq |x| \end{cases} \quad (11)$$

$$f(x) = \begin{cases} (a+2)|x|^3 - (a+3)|x|^2 + 1, & 0 \leq |x| < 1 \\ a|x|^3 + 5a|x|^2 - 8a|x| + 4a, & 1 \leq |x| < 2 \\ 0, & 2 \leq |x| \end{cases} \quad (12)$$

제3장에서는 기존의 반화소 단위 shifting matrix 알고리즘을 1/4 화소 단위로 확장하고 bilinear 보간법 보다 보간 성능이 뛰어난 b-spline, cubic 보간법을 이용한 주파수 영역에서의 움직임 예측 기법을 제안한다. B-spline 또는 cubic과 같은 보간법을 shifting matrix에 적용하면 기존 shifting matrix는 다음과 같이 전개된다.

$$h_1 = h_3 = f(d)D_{\lfloor \Delta x \rfloor} + f(1-d)D_{\lceil \Delta x \rceil} + f(1+d)D_{\lfloor \Delta x \rfloor - 1} + f(2-d)D_{\lceil \Delta x \rceil + 1} \quad (13)$$

$$h_0 = h_2 = f(d)D_{4 - \lfloor \Delta x \rfloor} + f(1-d)D_{4 - \lceil \Delta x \rceil} + f(1+d)D_{4 - \lfloor \Delta x \rfloor - 1} + f(2-d)D_{4 - \lceil \Delta x \rceil - 1} \quad (14)$$

$$v_0 = v_1 = h_0^t, \quad v_2 = v_3 = h_1^t \quad (15)$$

1/4 화소 단위 움직임 예측이므로 $d = \left\{ \frac{1}{4}, \frac{1}{2}, \frac{3}{4} \right\}$ 가 되며, $n < 0$ 일 경우 D_n 과 D_n^t

표 1. 공간 영역과 주파수 영역 부화소 움직임 예측의 연산량 비교
Table 1. Comparison of complexity in spatial domain and frequency domain

	Spatial domain	Transform domain
Addition	$16p^2(32N^2-1) + (45Np)^2 + 2N^2(N-1)$	$16p^2(2N^2-1) + 16N^3p^2$
Multiplication	$(60Np)^2 + 2N^3$	$(16Np)^2N$

는 4×4 의 영행렬을 사용하고 $n > 4$ 일 경우 $D_n = D_{8 \bmod n}^t$ 그리고 $D_n^t = D_{8 \bmod n}$ 가 된다. 식 (15)에서 사용되는 움직임 정보는 Δx 대신 Δy 가 적용되어 사용되며, $\lfloor x \rfloor$ 는 버림함수를, $\lceil x \rceil$ 는 올림함수를 의미한다.

기존의 bilinear 보간법의 경우 부화소 예측을 위해 양쪽 2곳의 주파수 정보가 활용되는 반면, b-spline과 cubic 보간법을 사용할 경우 4곳의 주파수 정보가 필요하므로 $|d| < 1$ 또는 $|d| > 3$ 일 경우 참조 블록의 경계 밖에 위치하는 주파수 정보를 활용하게 된다. 주파수 영역으로 변환된 블록에서 특정 위치의 주파수 정보를 가져오기 위해선 특정 위치가 포함되어 있는 블록 전체의 정보가 필요하므로 불필요한 메모리가 요구된다. 이러한 문제를 해소하기 위하여 본 논문에서는 intra prediction 모드 중 vertical 또는 horizontal 모드를 적용하여 필요한 위치를 예측하는 방법을 사용하였다^[10].

예를 들어, f_{curr} 가 y축 방향으로 $-3 < d < -4$ 의 움직임이 발생할 경우 탐색 영역의 위쪽 경계 밖의 정보가 필요하게 되며, 이 경우 그림 4와 같이 탐색 영역의 정보를 위쪽 방향으로 복사함으로써 메모리 소모 문제를 해소할 수 있게 된다. 식 (16)과 (17)는 1/4 화소 단위 움직임 예측에서 그림 4와 같이 수직 방향으로 예측되는 경우 수직 shifting matrix의 예를 보여준다.

$$v_0 = v_1 = \begin{bmatrix} f(1+d) + f(d) & f(1-d) & f(2-d) & 0 \\ f(1+d) & f(d) & f(1-d) & f(2-d) \\ 0 & f(1+d) & f(1-d) & f(2-d) \\ 0 & 0 & f(1+d) & f(1-d) \end{bmatrix} \quad (16)$$

$$v_2 = v_3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ f(2-d) & 0 & 0 & 0 \\ f(1-d) & f(2-d) & 0 & 0 \end{bmatrix} \quad (17)$$

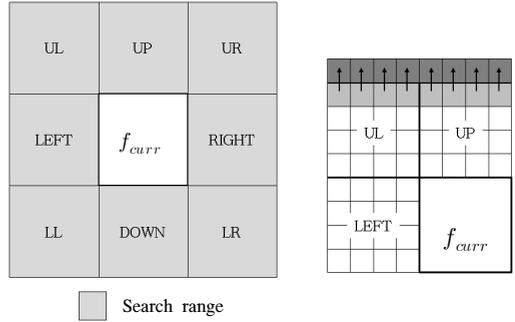


그림 4. Intra prediction을 이용한 예측 방법
Fig. 4. Estimation using intra prediction technique

본 논문에서 제안한 방식은 공간 영역에서의 움직임 예측에 비해 IDCT 과정이 필요하지 않으며, 재귀 방정식을 유도하여 연산량이 감소된다는 장점을 가진다. 표 1은 $N \times N$ 크기의 블록으로 공간 영역과 주파수 영역에서의 움직임 예측에 사용되는 계산량을 나타낸다. p 는 탐색구간, N 은 블록 크기를 나타내며, 실험을 위해 $N=4$, $p=8$ 로 고정하였다.

IV. 실험

본 논문에서 제안한 알고리즘의 성능을 평가하기 위해 H.264/AVC 비디오 부호화기를 사용하였다. 움직임 예측 방법으로는 1/4 화소 단위의 전역 탐색 기법을 사용하여 성능을 비교하였으며 움직임 정보를 결정하기 위한 기준으로는 SAD(sum of absolute difference)를 사용하였다.

표 2와 3은 CIF 해상도의 foreman과 paris 영상의 1, 2번째 프레임에 사용하여 양자화 파라미터(QP)를 변화시켜 공간 영역과 주파수 영역의 움직임 예측 결과를 보여준다. Foreman 영상의 경우 주파수 영역에서의 움직임 예측 방식이 공간 영역에서의 움직임 예측 방식보다 높은 PSNR을 제공하였으나 $QP \geq 21$ 의 경우 프레임별 총 비트수가 공간 영역에서의 움직임 예측보다 높아짐을 확인하였다.

Paris 영상의 경우 b-spline 보간법을 사용한 주파수 영역에서의 움직임 예측은 QP의 변화와 관계없이 모두 높은 PSNR을 제공하였으나, cubic 보간법을 사용한 주파수 영역에서의 움직임 예측은 $QP \geq 27$ 일 경우에 공간 영역에서의 움직임 예측보다 PSNR이 낮게 나타난다. 비트수는 b-spline 보간법의 경우 $QP=3$ 과 $QP \geq 21$ 에서, cubic 보간법은 $QP=15$ 와 $QP \geq 21$ 에서 공간 영역보다 높게 나타남을 확인할 수 있다. 특정

표 2. Foreman 시퀀스를 이용한 QP 조정에 따른 공간 영역과 주파수 영역에서의 1/4 화소 단위 움직임 예측 성능 비교
Table 2. Performance of motion estimation in frequency domain and spatial domain using foreman sequence

QP			3	6	9	12	15	18	21	24	27	30
Spatial domain	H.264	PSNR	61.46	55.90	52.99	50.46	48.11	45.66	43.65	41.58	39.90	38.27
		bits	343,183	301,121	258,644	214,380	159,479	115,206	75,996	46,775	30,216	18,514
	Cubic	PSNR	61.42	55.93	52.99	50.47	48.13	45.69	43.71	41.63	40.01	38.38
		bits	344,156	301,168	258,359	213,622	158,020	115,177	75,161	46,461	29,088	18,194
	B-spline	PSNR	61.61	55.92	52.99	50.46	48.09	45.62	43.63	41.54	39.86	38.22
		bits	344,202	302,705	260,707	216,729	160,335	116,678	76,186	47,558	30,391	18,637
Transform domain	B-spline	PSNR	61.75	56.09	53.18	50.81	48.57	46.30	44.27	42.31	40.52	38.78
		bits	347,719	298,433	251,082	198,875	152,680	109,822	76,398	49,704	32,353	20,864
	Cubic	PSNR	61.77	56.11	53.17	50.79	48.54	46.21	44.15	42.12	40.32	38.49
		bits	341,217	293,956	249,165	198,644	151,792	108,108	75,526	48,770	32,306	21,456

표 3. Paris 시퀀스를 이용한 QP 조정에 따른 공간 영역과 주파수 영역에서의 1/4 화소 단위 움직임 예측 성능 비교
Table 3. Performance of PSNR of motion estimation in frequency domain and spatial domain using paris sequence

QP			3	6	9	12	15	18	21	24	27	30
Spatial domain	H.264	PSNR	61.45	55.91	52.93	50.38	48.12	45.64	43.43	40.94	38.76	36.49
		bits	328,746	290,597	245,925	201,140	146,462	111,367	80,318	58,519	42,638	29,951
	Cubic	PSNR	61.39	55.94	52.93	50.38	48.11	45.65	43.50	41.00	38.86	36.60
		bits	329,250	291,531	245,985	199,714	145,831	110,075	79,254	58,002	41,900	29,751
	B-spline	PSNR	61.47	55.92	52.93	50.41	48.11	45.61	43.42	40.92	38.73	36.43
		bits	328,947	290,924	248,443	201,991	147,983	112,111	80,635	59,198	43,108	30,374
Transform domain	B-spline	PSNR	61.68	56.07	53.13	50.69	48.39	46.01	43.77	41.41	39.18	36.83
		bits	330,891	283,832	235,873	186,428	143,574	106,161	80,003	59,079	44,079	32,226
	Cubic	PSNR	61.60	56.04	53.14	50.65	48.35	45.88	43.61	41.16	38.95	36.56
		bits	327,233	282,450	235,541	190,540	146,558	109,340	82,058	60,428	45,196	33,260

값을 제외한 QP에서 이러한 결과가 나타나는 것은 주파수 영역의 움직임 예측이 공간 영역의 움직임 예측보다 정확한 움직임 정보를 찾아내어 부호화시 사용된 비트양이 적어짐에도 영상의 화질은 공간 영역과 유사하게 유지됨을 의미한다. 하지만 높은 QP에서 많은 비트율이 발생하는 이유는 QP가 높아질수록 양자화 오류가 커지기 때문으로 판단된다.

그림 5와 6은 QP=18에서 공간 영역과 주파수 영역의 움직임 예측 성능을 나타내고 있다. 실험 결과 공간 영역에서의 움직임 예측과 비교하여 주파수 영역에서의 움직임 예측의 PSNR이 더 높게 나타남을 확인할 수 있다.

전체적인 성능을 보면, 공간 영역에서의 움직임 예측에서 PSNR은 cubic(CSD), FIR 필터를 이용한 H.264, b-spline(BSD) 순으로 높게 나타나는 것을 확인할 수 있는 반면, 주파수 영역에서는

b-spline(BTD), cubic(CTD) 순으로 공간 영역과 다른 결과를 보이고 있다. 영상의 보간을 공간 영역에서 수

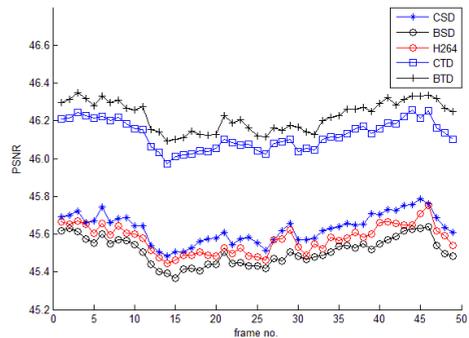


그림 5. Foreman CIF 비디오 시퀀스를 사용한 공간 영역과 주파수 영역에서의 움직임 예측 기법의 성능 비교 (QP=18)
Fig. 5. Comparison graph of ME in spatial and frequency domain performance using foreman sequence (QP=18)

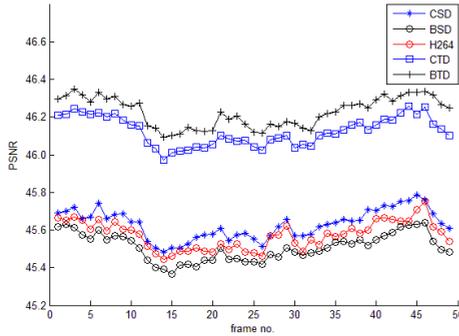


그림 6. Paris CIF 비디오 시퀀스를 사용한 공간 영역과 주파수 영역에서의 움직임 예측 기법의 성능 비교 (QP=18)
 Fig. 6. Comparison graph of ME in spatial and frequency domain performance using paris sequence (QP=18)

행하는 경우 화소를 생성하기 위해 그 주변 화소를 이용하게 되는데, 본 논문에서 제안하는 알고리즘의 경우 DCT 계수를 이용하게 된다. 블록 단위의 2D-DCT 변환의 경우 특정 주파수 값을 얻기 위해 블록 내의 대부분의 정보들이 사용되므로 공간 영역에서 특정 보간법을 사용한 움직임 예측 성능이 높은 PSNR을 제공하더라도 주파수 영역에서 그 성능이 동일하게 유지되지 않음을 알 수 있다.

V. 결 론

본 논문에서는 3차 함수 보간 방법을 적용한 주파수 영역에서의 1/4 화소 단위 움직임 예측 알고리즘을 제안하였다. 다양한 특성을 갖는 비디오 시퀀스를 이용한 모의실험을 통하여 제안한 알고리즘이 대부분의 QP 값에서 공간 영역의 1/4 화소 단위의 움직임 예측 알고리즘 보다 높은 PSNR을 제공함을 확인하였다. 특정 QP 값에서 공간 영역에서의 움직임 예측보다 PSNR이 감소하는 부분이 있었지만 그 차이가 무시할 수 있을 정도로 작다. 또한, IDCT의 생략과 재귀 방정식의 사용으로 공간 영역보다 낮은 계산량을 유지할 수 있었으며, DCT를 기반으로 하는 동영상 부호화기의 구성이 간단해지므로 실시간 동영상 서비스를 제공하기에 적합하다. 최근 확정된 H.265의 경우 블록 크기가 4×4 부터 32×32 까지 가변 블록을 사용하므로 shifting matrix를 수정하여 적용 가능할 것이며, 재귀방정식 활용으로 인해 32×32와 같이 큰 블록에서 더 낮은 계산량으로 연산이 가능하므로 효율적인 압축이 가능할 것이라 예상된다^[11].

References

- [1] S. Agha and V. M Dwyer, "Algorithm and VLSI architectures for MPEG-4 motion estimation," *Electronic System and Control Division Research*, 2003.
- [2] J. Oh, "Improved sub-block matching algorithm," *J. KICS*, vol. 33, no. 7, pp. 628-633, Jul. 2010.
- [3] Y. Keller and A. Averbuch, "Fast gradient methods based on global motion estimation for video compression," *IEEE, Circuits and Syst. for Video Technol.*, vol. 13, no. 4, pp. 300-309, Apr. 2003.
- [4] Q. H. Nguyen, V. A. Nguyen, C. V. Trinh, K. Q. Dinh, Y. Park, and B. Jeon, "Block-based transform-domain measurement coding for compressive sensing of images," *J. KICS*, vol. 39A, no. 12, Dec. 2014.
- [5] C. M. Yang and K. C. hung, "Embedded image compression scheme using rate-distortion optimized block coding of wavelet coefficients," *J. KICS*, vol. 39, no. 11, Nov. 2014.
- [6] J. Song and B. Yeo, "A fast algorithm for DCT-domain inverse motion compensation based on shared information a macroblock," *IEEE, Circuits and Syst. for Video Technol.*, vol. 10, no. 5, pp. 767-775, Aug. 2000.
- [7] R. Kleihorst and F. Caberera, "Implement of DCT-domain motion estimation and compensation," *IEEE Workshop, Signal Process. Syst. (SIPS)*, pp. 53-62, Oct. 1998.
- [8] R. H. J. M. Plompen, B. F. Schuurink, and J. Biemond, "A new motion-compensated transform coding scheme," *IEEE Acoustics, Speech, Signal Process*, vol. 1, pp. 371-374, 1985.
- [9] 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.
- [10] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC

video coding standard,” *IEEE Circuits and Syst. for Video Technol.*, vol. 13, no. 7, pp. 560-576, Jul. 2003.

- [11] D. Yoon and Y. Ho, “Fast mode decision method for HEVC in depth video,” *J. KICS*, vol. 37, no. 1A, pp. 51-56, Jan. 2012.

신 재 영 (Jae-young Shin)



2013년 : 동국대학교 정보통신
공학과 공학사
2015년 : 동국대학교 정보통신
공학과 공학석사
<관심분야> 영상처리, 동영상
압축

류 철 (Chul Ryu)



1989년 : Florida Institute of
Technology 컴퓨터공학과
공학사
1991년 : Ploytechnic Institute of
NYU 전기공학과 공학석사
1997년 : Ploytechnic Institute of
NYU 전기공학과 공학박사
1998년~1999년 : LG 정보통신연구소 선임연구원
1999년~현재 : 동국대학교 정보통신공학과 교수
<관심분야> Visual Communication, High
Resolution Reconstruction, Joint Source Coding