

# 웨이브릿 변환 영역에서의 움직임 정보를 이용한 비디오 압축

정회원 김 동 욱\*, 김 진 태\*\*

## Video compression using motion information in Wavelet transform domain

Dong-wook Kim\*, Jin-tae Kim\*\* *Regular Members*

요 약

본 논문에서는 움직임에 따른 인간시각 시스템의 인지 특성을 이용하여 동영상상을 효과적으로 압축할 수 있는 기법을 제안한다. 비디오 신호에 대한 인간 시각 인지 특성을 분석하여 큰 이득을 주지 못하는 성분, 즉 인간의 눈에 민감하지 않은 부분을 감축 내지는 제거함으로써 인지 능력에서 볼 때 주관적 화질의 저하 없이 비디오 신호를 낮은 비트율로 압축할 수 있다. 제안된 기법은 움직임이 발생하는 정도와 크기에 매우 의존적으로, 움직임이 많은 비디오 신호에 더욱 적합하나 움직임이 작은 경우에도 적합한 움직임 보상을 통해 높은 이득을 얻기 때문에 모든 영상에 대해서 적용 가능한 방법이다. 몇 가지 영상을 대상으로 한 컴퓨터 시뮬레이션 결과, 프레임당 21-30:1 정도의 압축하에서 높은 재생화질을 얻을 수 있다.

### ABSTRACT

A technique for an efficient video coding based on characteristics of human visual response in relation to motion is described in this paper. An input frame is segmented into low frequency bands and high frequency bands by wavelet transformation. The non-sensitivity parts of the segmented bands are removed according to spatial and directional frequency sensitivity, which is related to motion property in a frame. Experimental results of the proposed method show good performance in PSNR with keeping on without degradation of subjective quality with 21-30:1 coding rate.

### I. 서 론

영상 압축의 목적 중의 하나는 가능한 한 가장 적은 비트량으로 영상을 표현하는 것이며, 이를 통해 전송 속도를 개선시키고 또한 저장 요구량을 최소화시킬 수 있다. 또다른 측면에서는 영상 압축은 주어진 통신 수단이나 저장 용량에 대해 가능한 한 최적의 충실도를 얻는 것이다. 이 두 가지 면을 충족시키기 위해 정지 영상 및 동영상에 대해 여러 가지 압축 기법들이 제시되었으며<sup>[1,2,3]</sup>, 표준화의 일

환으로서 대표적으로 JPEG, MPEG 1,2<sup>[4]</sup> 등이 제시되었다. 일반적으로 압축영상을 바탕으로 얻어지는 복원 영상의 화질은 PSNR을 중요한 파라미터로 적용한다. PSNR은 계산에 의해 얻어지는 객관적인 수치이지만, 이에 반하여 복원 영상의 주관적인 화질과 일치하는 것은 아니다. 영상 통신 시스템의 설계에서 가장 중요한 것 중의 하나는 인간의 눈을 통해 볼 수 있는 정보를 전송하고 디스플레이하는 것이다. 따라서, 영상 정보를 전송하는데 있어서 인간의 시각 특성을 이해하고, 이를 적절하게 이용함으로써 효과적인 정보 전송을 달성할 수 있다. 인간

\* 전주대학교 전자매체공학부(dwkim@jeonju.ac.kr), \*\* 한서대학교 컴퓨터과학과 (jtkim@gaya.hanseo.ac.kr)

논문번호 : 98018-0120, 접수일자 : 1999년 1월 20일

\*본 연구는 전주대학교 교내연구비지원으로 수행되었습니다.

시각특성을 이용한 영상압축은 여러 분야에서 적용되었다. 영상 부호화와 관련하여 인간 시각 반응에 대한 문제는 Clarke<sup>[5]</sup>에 의해 고려되었는데, 홍채에서 망막을 통하는 광학적 전송의 물리적인 특성은 주파수 영역에서 저역(low pass) 필터 특성을 갖는다고 보였다. 이러한 특성은 S. H. Tan<sup>[6]</sup>에 의해 양자화 과정에서 유용하게 적용되었다.

동영상 신호의 각 프레임은 시간축상에서 샘플링된 일정한 해상도를 가진 영상 데이터로서, 영상의 밝기와 명확도에 따라서 인간의 눈이 느끼는 감도가 다르다. 비디오와 같이 프레임율(frame rate)가 높아 어떤 밝기에서도 인간의 시각이 영상의 시간축상 불연속을 느끼지 못할 경우에는 영상내 영역의 상세한(detail) 정도에 따라서 인간 시각의 감도가 다르다. 인간의 눈은 시간축상으로 저역통과 필터(low-pass filter)의 특성을 갖고, 그 필터링 효과는 낮은 공간 주파수(평탄한 영역)보다는 높은 공간 주파수(섬세한 영역)에 대해서 더 크게 나타난다<sup>[2]</sup>. 본 논문에서는 움직임의 속도가 클수록 시각이 갖는 감도는 떨어진다는 시각특성을 이용하여 효율적인 동영상 압축을 달성하고자 한다. 먼저, 입력되는 각 프레임에 대해 정해진 대역으로 분할하고, 분할된 각 대역에 대해 고주파 대역들에 대해서는 인간시각의 시간축상의 저역통과 필터 특성과 공간 주파수 영역상에서의 움직임 속도에 따른 인간시각의 대역통과 필터를 적용할 수 있다. 즉, 입력 영상을 움직이는 영역과 배경영역, 배경영역은 평탄한 부분과 거친부분 등으로 구분하여 처리하고자 한다. 이러한 개념을 바탕으로 영상의 각 영역이 일정한 화질을 유지하게 되고 전체적으로는 주관적 고화질을 유지하게 한다.

본 논문에서는 이러한 목적을 달성하기 위해 대역분할 기법을 이용하며, 대역분할 방법으로서 웨이브릿(wavelet) 이론<sup>[7]</sup>을 적용한다. 대역분할 부호화 기법은 첫째, 신호의 대역분할 및 복원의 개념이 간단하며, 둘째 각 대역신호를 부호화할 때 발생하는 오차가 해당 주파수 대역에 국한되어 복원 후 전체 영상에 미치는 영향이 제한된다는 점, 셋째 분할된 각 대역에 할당되는 비트수를 조절함으로써 양자화 잡음 스펙트럼의 형태를 변화시켜 주관적으로 우수한 신호를 재생시킬 수 있다는 점, 마지막으로 병렬 처리가 가능하다는 점 등의 여러 가지 장점을 가지고 있다. 웨이브릿(wavelet) 이론은 다양한 스케일(scale)과 해상도(resolution)에서 신호를 고찰, 분석하고자 하는 생각으로 수학, 공학분야에서 발전되어

왔다<sup>[7]</sup>. 이러한 시도는 1980년대 중반 Morlet, Grossmann, Meyer 등에 의해 개괄적인 수학 체계가 이루어졌으며, 이후 Doubechies와 Mallat에 의해 이산 신호처리로 확대, 발전되었다. 일반적으로 정재적(stationary)인 신호의 분석은 푸리에(Fourier) 변환을 사용하고 있지만, 이 방법은 전체 공간영역에서 적분을 행하므로 국부적(local)인 특성을 만족하지 않는다. 대상 신호가 비정재적(non-stationary)인 경우에는 신호가 시변-스펙트럼 특성을 갖는데, 기존에 사용되었던 Short Time Fourier 변환(STFT)은 시간 해상도가 주파수에 관계없이 항상 일정하므로 신호 해석에 불리하다. 반면에 웨이브릿 변환은 주파수에 따라 시간 해상도가 변하므로 신호 해석에 훨씬 유용하다. 따라서, 본 논문에서는 TV 신호와 비디오 신호 같은 영상신호(sequence images)를 웨이브릿 변환함으로써 고화질을 유지한 고압축을 이루고자 한다. 웨이브릿 변환을 이용하여 영상 신호의 대역을 압축하고자할 때, 대역 세분에 의한 공간영역상의 압축으로는 충분한 압축을 이룰 수 없으므로 시간축상의 중복성을 이용하는 방법도 도입되고 있다. 웨이브릿 변환에 움직임 보상을 함으로써 시간축상의 중복성을 이용한 데이터 압축을 이룰 수 있다<sup>[8]</sup>.

본 논문의 구성은 다음과 같다. 먼저 제 2장에서 웨이브릿 변환의 원리 및 움직임과 인간 시각의 인지 및 반응능력에 관해 살펴보고, 제 3장에서 시각특성을 이용한 동영상 압축기법을 제시하며, 제 4장에서 제시된 기법에 대한 실험결과 및 검토를 살펴 보며, 마지막으로 제 5장에서 결론을 맺는다.

## II. 대역분할 및 시각특성

### 1. 다해상도 표현

다해상도(multi-resolution) 표현은 영상 정보 분석에 유용한 계층적인 처리 기법을 제공한다. 저해상도에서 세부 신호(detail signal)는 영상의 윤곽을 나타내는 보다 큰 구조를 갖고 있다<sup>[7]</sup>. 따라서 낮은 해상도에서 영상의 세부 신호를 분석하고 점진적으로 해상도를 증가시켜 영상 신호를 해석하는 것이 신호 처리에 보다 효율적이다. 이러한 다해상도 기법은 저단계(low-level) 신호 처리 분야에서 폭넓게 연구되고 있으며, 후술될 계층적 움직임 추정 기법도 이와 유사한 신호 분석 기법이다.

먼저 웨이브릿 변환의 개괄적 내용을 수식으로 살펴보면,

$$V_{2^j} \subset V_{2^{j+1}} \quad \forall j \in \mathbb{Z} \quad (1)$$

$\mathbb{Z}$  : 정수 집합

와 같다. 해상도  $2^{j+1}$ 의 근사 신호  $V_{2^{j+1}}$ 가 보다 작은 해상도  $2^j$ 에서 계산된 모든 정보  $V_{2^j}$ 를 갖게 됨을 나타낸다. 또한 해상도  $2^j$ 상에서  $f(x)$ 의 근사치를 계산할 때, 원래 신호  $f(x)$ 의 몇몇 정보는 소실되며, 해상도를 점차 증가시키기에 따라 근사 신호는 원래 신호에 수렴하게 되고, 반대로 해상도를 감소시키기에 따라 근사 신호는 점차적으로 정보를 잃어 가고, 결국 0으로 수렴하게 됨을 나타낸다.

$$\lim_{j \rightarrow +\infty} V_{2^j} = \bigcup_{j=-\infty}^{+\infty} V_{2^j} \text{ is dense in } L^2(\mathbb{R}) \quad (2)$$

$$\lim_{j \rightarrow -\infty} V_{2^j} = \bigcap_{j=-\infty}^{+\infty} V_{2^j} = \{0\} \quad (3)$$

근사 연산(approximation operation)은 모든 해상도에서 유사하게 행하여지는데, 근사 함수 공간은 해상도 정도에 따라 각각의 근사 함수를

$$f(x) \in V_{2^j} \Leftrightarrow f(2x) \in V_{2^{j+1}} \quad \forall j \in \mathbb{Z} \quad (4)$$

와 같이 스케일링(scaling)함으로써 얻어진다. 신호의 이동(translation)은

$$f(x) \in V_{2^j} \Rightarrow f(x - 2^{-j}k) \in V_{2^j} \quad \forall k \in \mathbb{Z} \quad (5)$$

와 같으며, 해상도 변화에는 영향을 주지 않는다.  $\mathbb{R}$ 을 실수,  $L^2(\mathbb{R})$ 를 유한 에너지를 갖는 벡터 공간이라 하면 다해상도 근사는 위의 조건을 만족하는 폐 부분공간(closed subspace)  $(V_j)_{j \in \mathbb{Z}}$  계열(sequence)이다<sup>[11]</sup>.

벡터 공간  $V_{2^j}$ 상에 원 신호  $f(x)$ 의 직교 투영 성분은  $f(x)$ 를 정규 직교 기저  $\phi_{2^j}(x - 2^{-j}n)$ 으로 분해하되로서 계산할 수 있다. 즉, 벡터 공간  $V_{2^j}$ 로의 직교 투영 연산자를  $A_{2^j} f(x)$ 라 하면,

$$A_{2^j} f(x) = 2^{-j} \sum_n \langle f(u), \phi_{2^j}(u - 2^{-j}n) \rangle \cdot \phi_{2^j}(x - 2^{-j}n) \quad (6)$$

$$\forall f(x) \in L^2(\mathbb{R})$$

의 관계가 성립하고, 여기에서 해상도  $2^j$ 에서의 신호  $f(x)$ 의 이산적 근사는

$$A_{2^j}^d f = (\langle f(u), \phi_{2^j}(u - 2^{-j}n) \rangle)_{n \in \mathbb{Z}} \quad (7)$$

로 표현되는 내적(inner product)집합으로 나타나게 된다. 실제로는

$$A_{2^j}^d f = \sum_{k=-\infty}^{+\infty} \tilde{h}(2n-k) A_{2^{j+1}}^d f \quad (8)$$

단,  $\tilde{h}(n) = h(-n)$

와 같은 피라미드 변환으로 나타나게 된다.  $j < 0$ 일 때, 모든 이산적 근사  $A_{2^j}^d f$ 는  $A_{2^1}^d f$ 로부터 반복적 변환을 통해 계산된다. 반면에  $D_{2^j} f = (\langle f(u), \psi_{2^j}(u - 2^{-j}n) \rangle)_{n \in \mathbb{Z}}$ 를 해상도  $2^j$ 에서의 이산 세부(discrete detail)라 한다. 이것은  $A_{2^{j+1}}^d f$ 와  $A_{2^j}^d f$  사이의 정보차를 나타내며,

$$D_{2^j} f = \sum_{k=-\infty}^{+\infty} \tilde{g}(2n-k) A_{2^{j+1}}^d f \quad (9)$$

단,  $\tilde{g}(n) = g(-n)$

와 같이 표현된다. 원래 신호는 필터링(filtering)과 부표본화(subsampling)에 의해 이산적 근사와 세부 신호로 분해(analysis)되며, 반대 과정으로 합성(synthesis)된다.

## 2. 쌍직교 웨이브릿 변환

Mallat에 의해 구성된 필터는 무한 차수의 경우에만 완전 재구성 조건을 만족하므로, Daubechies는 완전 재구성 특성을 만족하는 유한 차수의 필터를 설계했다. 그러나 대칭성(symmetric)은 불연속적인 기저 이외에는 만족시킬 수 없었다. 실제 영상처리에서 요구되는 필터는 대칭성을 만족하는 유한 차수의 필터이므로 최근에는 정규 직교(orthogonal) 조건을 완화하고,

$$\langle \psi_{1,j}^n, \psi_{1,j}^{n'} \rangle = \delta_{1,j} \cdot \delta_{n,n'} \quad (10)$$

단,  $\delta_{n,n} = 1, \quad n = n'$

0, otherwise

와 같은 쌍직교성(biothogonality) 조건을 만족하도록 하는 쌍직교 웨이브릿 변환<sup>[9]</sup>이 행해지고 있다. 즉, 신호를 분해할 때는  $\psi(x)$ 가 사용되고 합성할 경우에는  $\psi^*(x)$ 를 사용한다. 따라서 신호를 분해할 경우에는 식(8), (9)가 그대로 사용되지만, 재구성할 경우에는

$$A_{2j+1}^d f = 2[\sum_k h^*(n-2k)A_{2j}^d f + \sum_k g^*(n-2k)D_{2j}f] \quad (11)$$

가 되며, 위의 식에서 사용된 필터 특성은

$$\begin{aligned} g^*(n) &= (-1)^n h(-n+1) \\ g(n) &= (-1)^n h^*(-n+1) \\ \sum_n h(n)h^*(n+2k) &= \delta_{k,0} \\ \sum_n h(n)g(n+2k) &\neq 0 \end{aligned} \quad (12)$$

$\sum_n h^*(n)g(n+2k) \neq 0$ 의 조건을 만족해야 한다.

### 3. 인간 시각 특성

영상신호의 한 프레임은 시간축상에서 샘플링된 일정한 해상도를 가진 영상 데이터로서, 영상의 밝기와 명확도에 따라서 인간의 눈이 느끼는 감도가 다르다. 비디오와 같이 프레임율(frame rate)이 높아 어떤 밝기에서도 인간의 시각이 영상의 시간축상 불연속을 느끼지 못할 경우에는 영상내 영역의 상세한(detail) 정도에 따라서 인간 시각의 감도가 다르다<sup>[1]</sup>. 그림 1에서 볼 수 있는 바와 같이, 인간의 눈은 시간축상으로 저역통과 필터(low-pass filter)의 특성을 갖고, 그 필터링 효과는 낮은 공간 주파수(평탄한 영역)보다는 높은 공간 주파수(섬세한 영역)에 대해서 더 크게 나타난다. 즉, 높은 공간 주파수 영역은 낮은 프레임율로 전송되어도 인간의 눈은 거의 감지하지 못한다는 것을 의미하고, 이를 이용하여 효율적인 전송을 할 수 있다<sup>[12]</sup>. 영상신호에 대한 인간 시각 특성을 정량적으로 표현하기는 어려운 일이나, 인간 시각 특성의 영상신호 대역압축에의 적용은 인간의 눈이 감지할 수 없는 정보를 제거할 수 있게 한다.

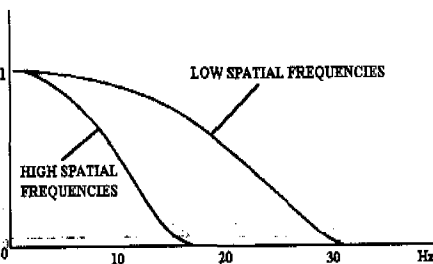


그림 1. 인간시각의 저역통과특성

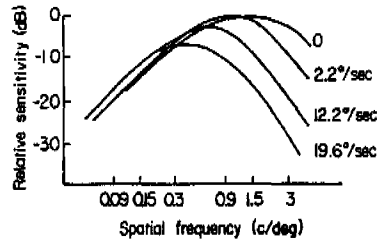


그림 2. 인간시각의 대역통과특성

그림 2는 영상신호에 대한 인간 시각 특성을 공간 주파수 영역상에서 표시한 것이다. 공간 주파수 영역상의 인간 시각 특성은 대역통과 필터(band-pass filter)의 특성을 갖고 있으며, 영상내 각 부분의 움직임 정도에 따라서 피크(peak)의 크기와 통과 대역을 달리한다. 움직임의 속도가 클수록 피크의 위치는 영(zero) 주파수쪽으로 이동하고 상대적 감도도 크게 떨어지는 것을 볼 수 있다<sup>[12]</sup>. 다시 말하면, 영상내 움직임이 큰 부분은 어느 정도 거친 처리를 하여도 인간 시각에는 별로 영향을 주지 않는다. 이러한 인간 시각 특성의 이용은 주관적 고품질을 유지한채 영상신호를 고압축할 수 있는 가능성을 제시한다.

### III. 시각특성을 이용한 동영상 압축

시각 특성을 영상신호의 압축에 이용한 기존의 연구들이 정지영상을 대상으로 각 주파수 영역에 대한 인간 시각의 상대적 감도(relative sensitivity)에 관심을 둔 반면에 본 논문에서는 일정한 프레임율을 갖는 비디오 신호에 대해서 인간 시각 특성을 적용하였다. 본 논문에서는 영상을 웨이브릿 이론을 바탕으로 16개의 동대역으로 분할한다<sup>[7, 8]</sup>. 4개의 저주파 대역들은 에너지가 높고 재생화질에 지대한 영향을 미치는 대역들이므로 움직임 보상을 하고 DCT하여 에너지 압축을 행한다. 나머지 12개 대역에 대해서는 저주파 대역의 움직임 벡터 크기에 근거하여 인간 시각 특성을 고려한 처리를 한다. 고주파 대역이 갖고 있는 성분들은 대체로 현재 프레임의 에지 성분들이다. 이들 에지 성분들은 해당 부위의 움직임의 크기에 따라서 중요도가 결정되고, 이렇게 함으로써 움직임이 작거나 정지한 영역에 대해서는 그 선명도를 살리면서 움직임이 큰 영역에 대해서는 시각적 지장을 주지 않는 범위내에서 데

이터를 크게 줄일 수 있다. 또한 제안한 방법은 움직임이 발생하는 정도와 크기에 매우 의존적이나, 움직임이 작은 영상의 경우에는 고주파 대역에서의 이득은 낮으나 4개의 저주파 대역들의 움직임 보상에 의한 이득이 크기 때문에 모든 영상에 대해서 적용 가능한 방법이다.

B- 11	B- 12	B- 13	B- 14
B- 21	B- 22	B- 23	B- 24
B- 31	B- 32	B- 33	B- 34
B- 41	B- 42	B- 43	B- 44

그림 3. 16-대역 분할

그림 3에서 B-11, 12, 21, 22 대역들은 재생영상의 화질에 매우 큰 영향을 미치는 대역이고, 나머지 12개의 대역들은 높은 공간 주파수 성분들로 구성된 대역 영상들로서 비교적 작은 영향을 미친다. 앞으로 이들 대역들을 각각 저주파수 대역과 고주파수 대역으로 구분하여 칭하겠다. 고주파수 대역들에 대해서 인간 시각의 시간축상의 저역통과 필터 특성과 공간 주파수 영역상에서의 움직임 속도에 따른 인간 시각의 대역통과 필터 특성을 적용할 수 있다. 즉, 영상은 움직이는 영역과 배경 영역, 배경 영역은 평탄한 영역과 세밀한 영역으로 구분되어 처리될 필요가 있다. 또한 시각적 중요도 측면에서 저주파수 대역은 고주파수 대역보다 충실히 처리해야 한다. 이렇게 함으로써 영상의 각 영역이 일정한 화질을 유지하게 되고, 전체적으로는 주관적 고품질을 유지할 수 있게 된다.

영상신호는 프레임 단위로 처리를 하는 프레임과 프레임간 처리를 하는 프레임으로 나눌 수 있다. 각각을 프레임내(intraframe) 처리와 프레임간(inter-frame) 처리로 부르기도 한다. 장면이 바뀔 때의 첫 번째 프레임과 그후 몇 프레임중 한 프레임은 프레임내 처리된다.

1. 프레임내 처리

프레임내 처리에 대한 프레임 복원 화질은 연이 어지는 다음 프레임들의 복원화질을 기본적으로 좌우한다. 프레임내 처리의 경우, B-11,12,21,22의 저주파수 대역이 다른 대역에 비해 월등히 높은 에너지를 갖고 있고, 복원화질을 크게 좌우하므로 신호

성분의 손실을 최소로 하는 압축기법을 채택하여야 한다. 따라서 이 대역들에 DCT를 적용하였다. B-11, 12, 21, 22를 제외한 나머지 고주파수 대역은 인간 시각의 저역통과 필터 특성을 적용하여 거칠게 처리되어도 주관적 화질에 큰 영향을 주지 않는다. B-11, 12, 21, 22 대역의 DCT 계수들은 좁은 양자화 계단폭을 갖는 균일 양자화기로 양자화된다. 나머지 대역들은 에너지 압축을 하지 않고 바로 균일 양자화되며, 양자화 계단폭의 조절에 의해 일정한 전송율을 유지하게 된다. 최종 양자화된 대역영상은 VLC(variable length coding)된다.

2. 프레임간 처리

프레임간 처리에서는 시간축상으로 나타나는 중복성을 이용하기 위해서 저주파수 대역에 움직임 보상을 한다. 즉, B-11,12,21,22 대역신호가 움직임 보상에 의한 예측 오차인 반면 나머지 대역신호들은 현재 프레임에 대한 고주파 성분들이고 시간축상의 변화에 대한 정보를 포함하고 있지 않다. 영상신호에 대한 인간 시각 특성의 적용 측면에서, 이들 고주파수 대역의 데이터 압축은 저주파수 대역에서 구한 움직임 벡터(움직임의 크기)에만 의존한다. 각 위치의 움직임 벡터의 크기에 따라서 움직임 영역과 배경 영역으로 구분된다. 각 영역별로 유효한 대역을 달리하여 유효하지 않은 대역의 해당위치 신호성분을 제거한다. 저주파수 대역신호는 DCT되어 좁은 양자화 계단폭을 갖는 균일 양자화기로 양자화되고, 고주파수 대역신호는 DCT변환을 적용하지 않고 바로 균일 양자화되며, 양자화 계단폭의 조절에 의한 일정한 전송율을 유지하게 된다. 최종 양자화된 대역영상은 VLC된다.

IV. 컴퓨터 시뮬레이션 및 검토

대상영상으로는 'Football'과 'Popple'을 사용하였다. 'Football'은 움직임이 상당히 크고 배경은 전체적으로 고르며 고주파 성분을 많이 포함한 영상이며, 'Popple'은 배경 부분이 매우 평탄하며 극히 저주파 성분으로 이루어져 있고 고정된 위치에서의 회전 운동만이 있는 영상이다. 대상영상 신호들은 4:2:2의 디지털 신호로 휘도신호(Y)와 색차신호(Cb,Cr)로 구성되어 있고, 비월주사된 형태이다. 그러나 본 실험에서는 홀수번째 필드를 세로축으로 보간하여 순차주사된 형태로 바꾸어 그것을 한 프레임으로 하였고, 휘도신호만을 대상으로 하였다.

영상의 크기는 720×480이고, 영상신호의 프레임율은 30frames/sec이고, 한 프레임의 intraframe처리 후, 9 프레임의 interframe처리가 뒤따른다. 표 1은 영상의 대역분할에 사용한 웨이브릿 계수이다.

표 1. 분해와 합성의 필터 계수

Kernel=0.75				Kernel=0.6			
분 해		합 성		분 해		합 성	
h*(n)		h(n)		h*(n)		h(n)	
h*(-2)		h(-3)	0.0	h*(-2)		h(-3)	-0.0107142
h*(-1)	-0.125	h(-2)	0.0	h*(-1)	-0.05	h(-2)	0.05357142
h*(0)	0.25	h(-1)	0.25	h*(0)	0.25	h(-1)	0.26071428
h*(1)	0.75	h(0)	0.5	h*(1)	0.6	h(0)	-0.6071428
h*(2)	0.25	h(1)	0.25	h*(2)	0.25	h(1)	0.26071428
h*(3)	-0.125	h(2)	0.0	h*(3)	-0.05	h(2)	0.05357142
		h(3)	0.0			h(3)	-0.0107142

한 프레임의 대역 신호는 저주파수 대역신호와 고주파수 대역신호로 나눌 수가 있다. 영상의 16-대역분할을 기본으로 하여 시뮬레이션 모델 SM1을 설정하였다. SM1에서 저주파수 대역의 움직임 보상은 B-11의 움직임 벡터를 B-12, 21, 22에 그대로 적용하는 방법을 채택했다. 이때, B-11 대역신호는 원영상에 비해 가로·세로 1/4의 해상도를 가지므로 움직임 추정 블록의 크기를 4×4, 움직임 추정 변위를 ±3으로 하여 BMA의 전역탐색 (full search) 방법을 사용하였다. 또한 블록 경계의 불연속으로 인한 부수적 고주파 성분의 발생을 최소화 하기 위해 DCT 블록의 크기도 4×4로 같게 했다. 고주파수 대역에서도 마찬가지로 4×4 블록단위로 각 대역 해당위치의 중요도가 B-11의 움직임 벡터에 의해 결정된다. 그러나 SM1의 저주파수 대역의 경우, 4×4의 작은 블록에 대해서 DCT를 적용하였기 때문에 블록의 크기가 작은 만큼 DCT의 압축효율이 떨어지고 모든 DCT 계수가 복원된 영상에 민감하게 영향을 미칠 수 있다. 이에 대한 대안으로 시뮬레이션 모델 SM2를 설정하였다. SM2에서는 저주파수 대역영상을 원영상의 가로·세로 1/2의 해상도를 유지하게 하였다. 따라서 움직임 추정 및 DCT 블록의 크기를 8×8로 하였다. 이 대역의 움직임 벡터 크기를 1/2로하여 SM1의 고주파수 대역 처리에 사용될 수 있다. 시뮬레이션 모델 SM3는 고주파수 대역의 좀더 높은 비율의 데이터 압축을 위한 것이다. 실제로 양자화된 후의 고주파수 대역의 많은 부분은 영(zero)의 값을 갖으나, 블록 단위의 정보전송

으로 인해 불필요한 비트가 할당되고 있다. B-23, 32, 33, 34, 43, 44의 대역들을 전송하지 않음으로써 고주파수 대역에서의 데이터를 압축할 수 있다. 시뮬레이션 모델들에 대해서 표 2에 나타내었다.

표 2. 시뮬레이션 모델

모 델	특 정
SM1	움직임추정 블록크기: 4×4 움직임 추정 변이: ±3 DCT 블록크기: B-11의 움직임 벡터
SM2	움직임 추정 블록크기: 8×8 움직임 추정 변이: ±7 DCT 블록 크기: 저주파대역의 움직임 벡터(크기의 1/2로)
SM3	전송되는 고주파대역: B-13,14,24,31,41,42 전송되지않는 고주파대역: B-23,32,33,34,43,44

영상처리 방법은 시뮬레이션 모델(SM1, SM2, SM3)의 혼성 형태를 취하였다. 다음 표현의 L과 H는 각각 저주파수 대역과 고주파수 대역을 나타낸다.

- Type-1: SM1,L+SM1,H
- Type-2: SM2,L+SM1,H
- Type-3: SM1,L+SM3,H
- Type-4: SM2,L+SM3,H

표 3 'Football' 영상의 화질과 비트율

Frame	Type-1	Type-2	Type-3	Type-4
1	34.81, 0.355	34.82, 0.415	34.8, 0.30,8	34.82, 0.368
2	36.42, 0.406	36.41, 0.490	36.36, 0.359	36.36, 0.443
3	36.33, 0.384	36.37, 0.452	36.28, 0.337	36.32, 0.405
4	36.61, 0.325	36.63, 0.377	36.55, 0.278	36.56, 0.330
5	36.45, 0.374	36.48, 0.428	36.38, 0.327	36.41, 0.381
6	36.20, 0.401	36.23, 0.467	36.14, 0.354	36.16, 0.420
7	36.38, 0.363	36.42, 0.427	36.31, 0.316	36.35, 0.380
8	36.44, 0.377	36.45, 0.433	36.35, 0.331	36.36, 0.387
9	36.48, 0.376	36.52, 0.437	36.42, 0.329	36.46, 0.390
10	36.07, 0.481	36.07, 0.568	36.01, 0.434	36.01, 0.521

각 시뮬레이션 모델의 성능을 비교하기 위해서 저주파수 대역과 고주파수 대역의 양자화 계단폭을 각각 2와 8로 고정시켜 놓고 각 대역을 전송하는데 필요한 비트를 구하였다.

표 4 'Popple' 영상의 화질과 비트율

Frame	Type-1	Type-2	Type-3	Type-4
1	33.77, 0.352	33.80, 0.359	33.66, 0.351	33.69, 0.356
2	39.22, 0.279	39.33, 0.273	37.36, 0.277	37.43, 0.271
3	39.23, 0.290	39.31, 0.285	37.36, 0.288	37.42, 0.283
4	39.20, 0.293	39.31, 0.286	37.33, 0.291	37.40, 0.284
5	39.21, 0.294	39.26, 0.287	37.28, 0.292	37.31, 0.285
6	39.21, 0.292	39.28, 0.284	37.26, 0.290	37.31, 0.282
7	39.16, 0.291	39.28, 0.286	37.26, 0.289	37.33, 0.284
8	39.17, 0.292	39.26, 0.282	37.24, 0.290	37.30, 0.280
9	39.08, 0.295	39.22, 0.286	37.24, 0.293	37.24, 0.284
10	39.03, 0.298	39.17, 0.291	37.13, 0.296	37.22, 0.289

표3, 4, 그림 4는 Type-1에서 Type-4까지의 방법으로 영상신호를 처리한 결과이다. 처리된 영상은 전반적으로 높은 PSNR과 주관적으로도 좋은 화질을 나타내었다. 저주파 대역과 고주파대역의 비트할당 결과는 대상 영상의 특징에 따라서 차이가 있었다. SM1, L+SM1, H의 방법을 사용하였을 때 움직임이 매우 많은 'football' 영상의 경우는 저주파 대역과 고주파 대역에 각각 3:1의 비트 할당이 되었고 전체적으로 21:1정도의 압축율을 보인다. 고주파 대역에 SM3, H를 사용하면 전체 압축율은 24:1 정도가 된다. 영상 전체가 매우 정적이고 평탄한 'Popple'의 경우 저주파 대역과 고주파 대역의 비트 할당은 거의 비슷하며 전체적인 압축율은 약 27:1이다. 고주파 대역에 SM3, H 방법을 적용한 결과는 거의 이득을 주지 못했으며, 움직임 블록이 매우 적은 영상이므로 일부 고주파 대역의 제거는 PSNR의 저하를 가져왔다. 이것은 영상이 포함하고 있는 공간주파수 성분이 매우 적기 때문이라 판단된다.

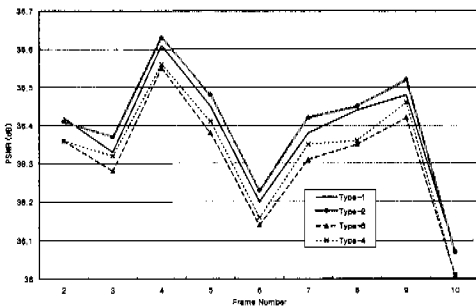


그림 4. 'Football' 영상의 화질

저주파 대역에 SM2,L의 방법을 적용한 결과는

SM1,L의 적용 결과와 비슷하다고 할 수 있다. DCT 블록 크기의 확대로 더 큰 압축 효과를 기대했으나 저주파 대역의 대역 세분으로 인한 압축 효율을 증가하지는 못했다. 또한 예측 오차를 DCT하기 때문에 DCT 블록 크기에 따른 영향은 매우 적다고 할 수 있다. 본 시뮬레이션에서는 균일 양자화기를 사용하여 일정 범위의 화질을 유지한 채 양자화 계단폭을 조정하는 것에 큰 제한을 받았는데, 불균일 양자화기를 사용하면 이런 문제점은 어느 정도 해결되고 압축율도 크게 높일 수 있을 것이라 생각한다.

V. 결론

본 논문에서는 인간시각특성을 고려한 대역 분할 부호화 기법을 제안하였다. 대역 분할 부호화는 대역 분할 특성으로 인하여 인간 시각특성을 적용하기가 매우 용이하다. 인간의 눈은 시간축 상으로 저역 통과 필터의 특성을, 공간 영역상으로는 대역 통과 특성을 갖는다는 점을 이용하여 분할된 각 대역 별로 움직임 정보를 이용하여 움직임 정도에 따라 적응적인 처리를 함으로써 주관적인 화질의 열화없이 고압축을 이룰 수 있다.

참고 문헌

- [1] A. N. Netravali and B. G. Haskell, *D-igital Pictures*, Plenum Press, 1988.9
- [2] A. N. Netravali and J. D. Robbins, "Motion compensated television coding Part I," *Bell Systems Technol. J.*, vol. 58, no. 3, pp.629-668, 1979.
- [3] A. Netravari, A. Lippman, "Digital T- elevision: a perspective," *Proc. of IEEE*, vol. 83, no. 6, pp. 834-842.
- [4] ISO-IEC/JTC1/SC29/WG11 *Recommendation H.262* Mar. 1994.
- [5] R. J. Clarke, *Transform coding of images*, Academic Press, 1985.
- [6] S. H. Tan, K. K. Pang, and K. N. Ngan, "Classified Perceptual coding with Adaptive Quantization", *IEEE C&S VT*, vol. 6, no.4, pp.375-388, 1996.8.
- [7] O. Rioul and M. Vetteri, "Wavelets and signal processing," *IEEE Signal Processing Mag.*,

