

論文96-33B-8-6

웨이브릿 변환된 다해상도 영상을 이용한 계층적 움직임 추정

(Multi-Resolution Hierarchical Motion Estimation in the Wavelet Transform Domain)

金珍台*, 張峻彌**, 金東郁***, 崔宗秀***

(Jin Tae Kim, Joon Pil Chang, Dong Wook Kim, and Jong Soo Choi)

요 약

본 논문에서는 웨이브릿 변환된 다해상도 영상을 이용한 새로운 계층적 움직임 추정 방법을 제안한다. 전역 탐색 움직임 추정 방법과 비교할 때, 기존의 계층적 움직임 추정 방법은 계산량을 크게 감소시킨다는 이점이 있지만 국부 최소의 문제로 인하여 추정 효율이 낮아진다. 제안된 방법에서는 웨이브릿 변환을 이용하여 영상을 다해상도의 계층 영상으로 분해하고, 영상 국부의 저주파 대역 성분과 고주파 대역 성분을 고려하여 그 부분의 움직임 추정에 참여하는 실제 계층 수를 결정함으로써 국부 최소의 문제를 해결하고자 한다. 고주파 대역 성분과 저주파 성분의 에너지의 비를 계산하여, 고주파 대역 성분이 많아 국부 최소에 빠질 수 있는 블럭에 대해서는 해상도가 높은 계층에서 보다 정확한 움직임 추정을 한다. 그 밖의 블럭에 대해서는 계층을 모두 사용하는 계층적 움직임 추정을 한다. 최종으로 구해진 움직임 벡터는 1차 웨이브릿 변환된 계층의 저주파 대역에 대한 것이다. 따라서, 전송되는 움직임 벡터량을 감소시키며, 수신단에서는 역 웨이브릿 변환하여 영상을 복구하게 됨으로써 움직임 추정으로 인한 블럭화 현상을 완화시킨다.

Abstract

In this paper, a new hierarchical motion estimation scheme using the wavelet transformed multi-resolution image layers is proposed. Compared with the full search motion estimation method, the existing hierarchical methods remarkably reduce the amount of the computation but their efficiencies are depreciated by the local minima problem. In order to solve the local minima problem, the multi-resolution image layers are composed using the wavelet transform and the number of layers participated in the motion estimation for a block is determined by considering of its low band energy and higher band energy on the first wavelet transformed layer. The ratio between higher band energy and low band energy of each block is evaluated and in the case of the blocks which include relatively large higher band energy, the motion estimation is carried out in the high resolution layer. Otherwise, all layers are used. The final motion vectors are obtained in the first wavelet transformed layer. So less bits for motion vectors are transmitted, and the decomposition of received image using inverse wavelet transform decreases the blocking effect.

* 正會員, 韓瑞大學校 電算情報學科

(Dept. Computer and Information Science, Hansoo Univ.)

** 正會員, LG電子 멀티미디어 研究所

(Multimedia Lab., LG Electronics INC)

*** 正會員, 中央大學校 電子工學科

(Dept. Electronics Eng., Chung-Ang Univ.)

※ 본 논문은 정보통신부의 대학기초 연구지원사업의 연구비지원으로 이루어졌음.

接受日字: 1995年9月28日, 수정완료일: 1996年7月31日

I. 서론

영상 부호화(image coding)의 목적은 원영상의 화질을 가능한한 유지하면서 영상 데이터간에 내재하는 중복성(redundancy)을 제거함으로써 적은 양의 정보로 원영상을 충실히 표현하고 재생하는 데 있다. 시간에 따라 변하는 동영상은 공간 상관성과 시간 상관성을 갖고 있어 효율적으로 많은 양의 데이터를 압축할 수 있다. 동영상의 경우 시간 상관계수가 공간 상관계수보다 훨씬 크기 때문에 시간 상관성을 이용하는 프레임간(interframe) 부호화가 효율적이다. 프레임간 부호화 방법은 움직임 추정(motion estimation) 및 움직임 보상(motion compensation) 부분과 예측 오차(prediction error)를 부호화하는 부분으로 구성된다. 대부분 시스템의 움직임 추정부는 블럭 정합 알고리즘(block matching algorithm: BMA)을 사용하는데, 이는 블럭 단위로 이전 프레임의 탐색 영역(search region)에서 평가 함수가 최적인 위치를 찾아 그 위치를 해당 블럭의 움직임 벡터(motion vector)로 결정하는 방식이다.¹¹⁾

BMA의 가장 일반적인 방법인 전역 탐색(full search)은 추정 블럭당 소요되는 계산량이 많으며, 대상 영상의 움직임에 대해 적절한 탐색 영역과 블럭 크기를 결정하기 어렵다. 전역 탐색이 갖는 계산량의 과다를 해결하기 위해 다양한 방법의 고속 알고리즘이 제안되었으나, 이들 방법은 탐색 영역에서 탐색 횟수를 줄여 가며 움직임 추정을 하므로 전역 탐색에 비해 예측 오차가 크고 추정 벡터의 신뢰성이 크게 떨어지는 문제점을 내포하고 있다.¹²⁾ 고속 알고리즘이 갖는 계산량의 이점을 유지하면서 탐색 영역과 블럭 크기 결정의 어려움을 극복하기 위해 계층적 움직임 추정(hierarchical motion estimation) 방법이 제안되었다.¹³⁻⁵⁾ 그러나 일반적인 계층적 움직임 추정은 해상도가 낮은 상위 계층에서의 부정확한 움직임 추정이 움직임 벡터에 결정적 오류를 제공하게 되는 국부 최소(local minima)의 문제점이 있다.

한편, 웨이브릿(wavelet) 변환은 시간과 주파수에 대한 국부성(locality)을 갖고 신호를 표현할 수 있어서 비정재성(non-stationary)을 갖는 영상 신호를 해석함에 유리하고, 이를 이용하여 표현된 영상은 인간의 시각 특성과 비슷하여 최근 영상처리 분야에서 각광을 받고 있다. 웨이브릿 변환을 이용한 다해상도 영상 처

리를 영상 부호화에 적용한 논문도 많이 발표되고 있다.¹⁶⁻⁸⁾

기존의 계층적 움직임 추정에서는 저역 통과 필터를 통과한 신호 성분만으로 계층이 구성된다. 복잡한 내용을 갖는 블럭의 경우, 고주파 성분이 많으므로 국부 최소에 빠질 위험이 높다. 따라서 본 논문에서는 계층적 움직임 추정 방법이 갖는 국부 최소 문제를 해결하기 위해 움직임 추정을 할 때 영상의 고주파 성분 에너지를 고려한다. 입력 영상을 먼저 쌍직교 웨이브릿(biorthogonal wavelet) 변환을 사용하여 다해상도의 계층을 구성한 후, 1차 웨이브릿 변환된 계층에서 블럭이 갖는 고주파 성분 에너지와 저주파 성분 에너지의 비를 평가한다. 상대적으로 고주파 성분 에너지가 많은 블럭은 그 해상도에서 움직임 추정을 하여 보다 정확한 움직임 추정 벡터 값을 갖도록 하고, 그 밖의 블럭에 대해서는 해상도를 한 단계 낮추어 다시 고주파 성분 에너지와 저주파 성분 에너지를 비교한다. 상대적으로 고주파 성분 에너지가 많은 블럭은 해당 계층부터 계층적 움직임 추정을 실시하며, 그렇지 않은 블럭은 해상도가 한 단계 더 낮은 계층부터 계층적 움직임 추정을 한다. 즉, 영상 국부의 저주파와 대역 성분과 고주파와 대역 성분을 고려하여 움직임 추정에 참여하는 계층 수를 결정한다. 이 때, 해상도가 낮은 계층으로 한 단계 옮길 때마다 블럭의 크기는 가로, 세로 각각 반으로 줄어든다. 최종 움직임 벡터는 1차 웨이브릿 변환된 계층의 저주파 대역에서 구해지고, 수신단에서는 역 웨이브릿 변환을 하여 영상을 재구성한다.

본 논문에서는 I장에 서론, II장에 웨이브릿 변환의 이론, III장에 기존의 계층적 움직임 추정 방법, IV장에 제안하는 다해상도 영상 계층을 이용한 계층적 움직임 추정 방법을 설명한다. 그리고 V장에서 컴퓨터 시뮬레이션과 그 결과를 검토하며, VI장에서 결론을 맺는다.

II. 웨이브릿 변환

웨이브릿 이론은 다양한 크기(scale)와 해상도(resolution)에서 신호를 고찰, 분석하고자 하는 생각으로 수학 및 공학 분야에서 발전되어 왔다.¹⁹⁾ 이러한 시도는 1980년대 중반 Morlet, Grossmann, Meyer 등에 의해 개괄적인 수학 체계가 이루어졌으며, 이후 Daubechies와 Mallat에 의해 이산 신호 처리(discrete signal processing)로 확대되고 발전되었다.

일반적으로 정재적인 신호의 분석은 푸리에(Fourier) 변환을 사용하고 있지만, 푸리에 변환은 전체 공간 영역에서 적분을 행하므로 국부적(local) 특성을 만족하지 않는다. 대상 신호가 비정재적인 경우에는 시변-스펙트럼 특성을 갖는데, 기존에 사용되었던 단시간(short time) 푸리에 변환은 시간 해상도가 주파수에 관계없이 항상 일정하므로 신호 해석에 불리하다. 반면에 웨이브릿 변환은 주파수에 따라 시간 해상도가 변하므로 신호 해석에 훨씬 유용하다.

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. 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,

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

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.

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

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

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

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

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

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

와 같으며, 해상도 변화에는 영향을 주지 않는다. R 을 실수, $L^2(R)$ 를 유한 에너지를 갖는 벡터 공간이라 하면 다해상도 근사는 위의 조건을 만족하는 폐 부분공간(closed subspace) $(V_{2^j})_{j \in Z}$ 계열(sequence)이다.^[10]

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 분해하므로서 계산할 수 있다. 즉, 벡터 공간 V_{2^j} 로의 직교 투영 연산자를 $A_{2^j} f(x)$ 라 하면,

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

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

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

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

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

와 같은 피라미드 변환으로 나타나게 된다.

$j < 0$ 일 때, 모든 이산적 근사 $A_{2^j}^d f$ 는 $A_1^d f$ 로부터 반복적 변환을 통해 계산된다. 반면에 $D_{2^j} f = \langle f(u), \psi_{2^j}(u-2^{-j}n) \rangle_{n \in Z}$ 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,

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

와 같이 표현된다. 원래 신호는 필터링(filtering)과 부표본화(subsampling)에 의해 이산적 근사와 세부 신호로 분해(analysis)되며, 반대 과정으로 합성(synthesis)된다. 그림 1은 다해상도 상에서의 분해와 합성을 나타낸다.

2. 쌍직교 웨이브릿 변환

Mallat에 의해 구성된 필터는 무한 차수의 경우에만 완전 재구성 조건을 만족하므로, Daubechies는 완전 재구성 특성을 만족하는 유한 차수의 필터를 설계했다.

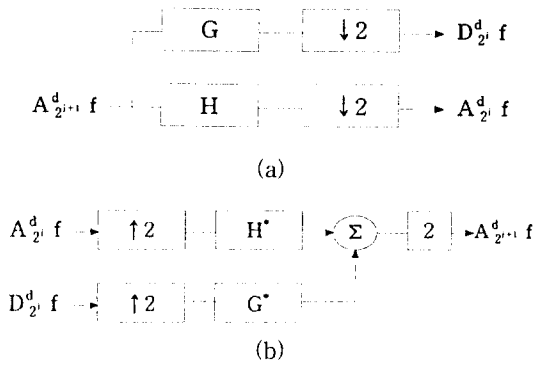


그림 1. 다해상도 웨이브릿 모델 : (a) 분해, (b) 합성
Fig. 1. Multiresolution wavelet model : (a) Analysis, (b) Synthesis.

그러나 대칭성(symmetric)은 불연속적인 Harr 기저 이외에는 만족시킬 수 없었다. 실제 영상처리에서 요구되는 필터는 대칭성을 만족하는 유한 차수의 필터이므로 최근에는 정규 직교(orthogonal) 조건을 완화하고

$$\langle \Psi_{i,j}^*, \Psi_{i',j'} \rangle = \delta_{i,i'} \cdot \delta_{j,j'} \quad \text{단, } \delta_{n,n'} = 1, n = n' \quad (10)$$

0, otherwise

와 같은 쌍직교성(biorthogonality) 조건을 만족하도록 하는 쌍직교 웨이브릿 변환^[11]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

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

가 되며, 위의 식에서 사용된 필터 특성은

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

의 조건을 만족해야 한다.

III. 기존의 계층적 움직임 추정 방법

계층적 움직임 추정(hierarchical motion estimation: HME) 방법은 다해상도 표현 기법의 기본 개념

과 마찬가지로 각 단계마다 탐색 영역 혹은 블록 크기를 다르게 적용하여 대상 영상의 다양한 움직임을 효과적으로 찾아내고자 하는 시도로서 도입되었다. 즉, 해상도가 가장 낮은 최상위 계층에서 거친(coarse) 움직임을 찾아내고, 그 움직임을 근거로 하여 해상도를 단계로 증가시키며 보다 세밀한(fine) 움직임을 추정하는 기법이다. 그림 2에 일반적인 계층적 움직임 추정 기법을 나타내었다.^[12] 그림 2는 영상을 4개의 계층으로 구성하여 움직임 추정을 하는 경우이다. 기존의 계층적 움직임 추정 방법에는 Biering이 제안한 기법^[3], 오차 함수의 계층적 근사를 이용한 기법^[4], 평균 피라미드를 이용한 계층적 고속 움직임 벡터 추정 방법^[5] 등이 있다.

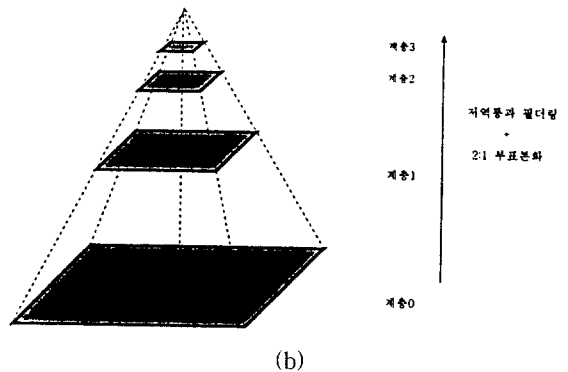
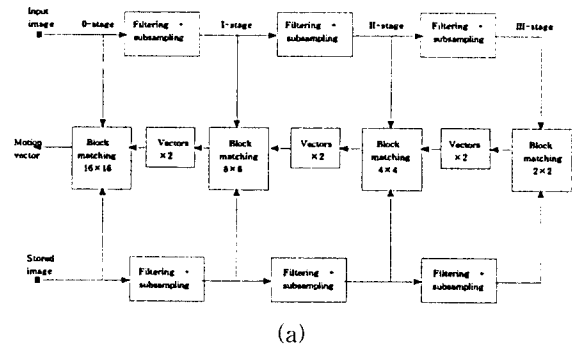


그림 2. 계층적 움직임 추정 : (a)블럭도 (b)계층구조
Fig. 2. Hierarchical motion estimation : (a)Block diagram, (b)Hierarchical layers.

IV. 웨이브릿 변환된 다해상도 영상을 이용한 계층적 움직임 추정

일반적으로, 계층을 구성하기 위해 여러 번의 저역 통과 필터를 통과한 영상은 뭉롱화(blurring)되고, 움직임이 큰 블록일 경우 국부 최소에 빠질 확률이 비교

적 크다. 특히, 에지 성분이 강한 영상의 경우에는 계층 구성 후 영상 전체가 몽롱화 되기 때문에 잘못된 움직임 추정을 하기 쉽다. 즉, 해상도가 낮은 상위 계층에서의 부정확한 움직임 추정은 연속적으로 하위 계층에 영향을 주어 최종적으로 부정확한 움직임 벡터를 전송하게 된다. 이러한 국부 최소의 문제를 해결하기 위해 본 논문에서는 각 블럭의 영상 신호 특성에 따라 적응적으로 움직임 추정에 참여하는 계층 수를 결정하여 보다 정확한 추정 벡터 값을 갖도록 하였다.

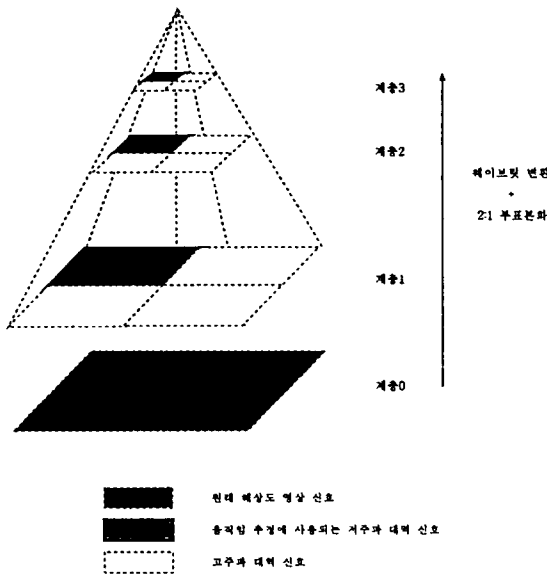


그림 3. 웨이브릿 변환된 다해상도 계층구조
Fig. 3. Hierarchical structure with the wavelet transformed multi-resolution layers.

제안된 알고리즘에서는 쌍직교 필터를 이용하여 영상을 웨이브릿 변환하여 다해상도 계층을 구성한다. 다해상도 영상 계층 구조는 그림 3과 같다. 먼저, 계층1의 저주파 신호에 대해서 블럭 단위로 메모리 내에 저장된 이전 프레임의 계층 신호와 현재 프레임의 계층 신호간의 차를 구하고, 각 계층간 저주파 대역 에너지 값들을 비교함으로써 계층적 움직임 추정시 고주파 성분의 손실로 움직임이 잘못 찾아질 가능성이 있는 블럭을 판단한다. 이 때, 차신호를 구하는 이유는 움직임이 없는 블럭을 배제하기 위함이다. 식(13)은 계층간의 에너지 관계를 나타낸다.

$$E_{k,L} = E_{k+1,L} + E_{k+1,H} \tag{13}$$

여기서 E_k 는 계층 k 의 에너지이고, L 은 저주파 대역

을, H 는 고주파 대역을 나타낸다. 식(13)에서 알 수 있듯이 $E_{k+1,H}$ 의 값이 커지면 $E_{k+1,L}$ 의 값이 작아지고, 이러한 신호에 의한 움직임 추정은 부정확한 국부 최소의 문제를 발생시킨다.

국부 최소로 인해 움직임 추정이 잘못될 가능성이 있는 블럭을 판단하기 위해서 이전 프레임과 현재 프레임 계층 신호의 저주파 대역 에너지와 고주파 대역 에너지를 비교한다. 어떤 계층의 저주파 대역 에너지와 고주파 대역 에너지의 비교는 그 계층과 그 보다 해상도가 한 단계 낮은 계층의 저주파 대역 에너지를 비교함으로써도 같은 결과를 얻을 수 있다. 계층1에서의 블럭 크기는 8×8 이고 계층이 한 단계 높아질수록 블럭의 크기는 가로, 세로 각각 반으로 줄어든다. 계층1과 계층2의 저주파 대역 에너지의 비, $(4 \times E_{2,L})/E_{1,L}$ 을 구한다. 그 값이 임계값보다 작으면 계층1의 저주파 대역의 블럭은 고주파 성분을 많이 잃은 블럭으로 움직임 추정에 적합하지 않다고 판단되므로 해상도가 높은 계층1에서 움직임 추정을 한다. 그렇지 않은 블럭은 계층을 한 단계 높여서 계층2와 계층3의 저주파 대역 에너지의 비, $(4 \times E_{3,L})/E_{2,L}$ 을 구한다. 그 값이 임계값보다 작으면 계층2와 계층1을 이용하여 계층적으로 움직임 추정을 하고, 그렇지 않은 블럭은 계층을 한 단계 올라간다. 이러한 판단은 계층 k 와 계층 $(k-1)$ 의 저주파 대역 에너지 비교까지 계속된다. 이 알고리즘의 흐름도를 그림 4에 보였다.

각 경우 계층1에서 구해지는 최종 움직임 벡터 $MV_1(x, y)$ 는

$$MV_1(x, y) = 2^{n-1} \cdot MV_n(x, y) + \sum_{m=2}^n 2^{n-m} \cdot \Delta MV_{n-m+1} \quad n \geq 2 \tag{14}$$

와 같이 나타낼 수 있다. 각 블럭의 움직임 추정에 참여하는 계층수 n 은 저주파 대역 에너지의 비에 따라 결정된다. 즉, 영상의 국부적 특성을 고려하여 그 해당 블럭의 움직임 추정시 참여하는 계층 수를 변화시켜 보다 효과적인 움직임 추정을 하도록 한다. 또한 영상의 복구시 움직임 보상된 계층1의 영상을 역 웨이브릿 변환함으로써 낮은 비트율 부호화(low bitrate coding)에서 인간 시각에 민감하게 반응하는 블럭화 현상(blocking effect)을 줄일 수 있다. 전송되는 최종 움직임 벡터는 계층1의 저주파 대역에 대한 것이므로 움직임 벡터의 전송량도 감소된다. 계층1의 고주파 대

역에 대해서는 저주파 대역의 움직임 벡터를 이용하여 움직임 보상을한다.

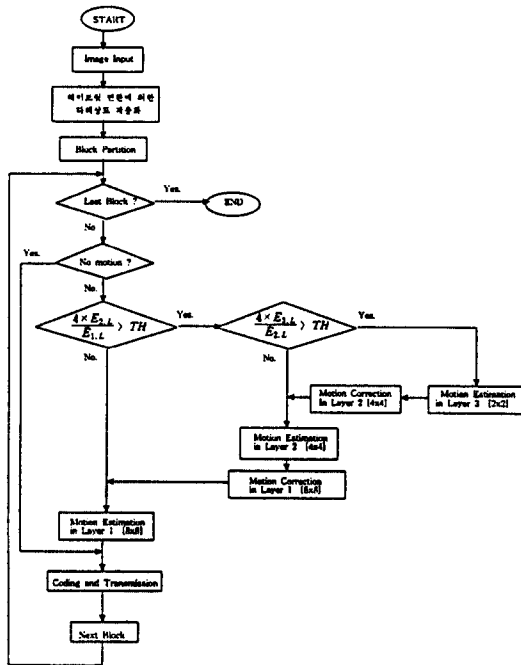


그림 4. 제안한 방법의 흐름도
Fig. 4. Flow chart of the proposed algorithm.

V. 컴퓨터 시뮬레이션 결과 및 검토

본 실험에서는 720×480 크기의 MPEG 표준 영상인 'Football' 영상을 사용하였다. 'Football' 영상은 고주파 성분이 많은 배경과 움직임이 큰 부분을 포함하고 있어 본 알고리즘의 실험을 위하여 선정하였다. 그림 5는 실험에 사용한 원영상이다. 웨이브릿 변환을 이용한 다해상도 영상 계층 구성에는 표 1의 쌍직교 분해 및 합성 필터를 사용하였다. 분해 및 합성 필터의 탭 수는 각각 5와 7이고, LPM(Laplacian Pyramid Model)의 커널(kernel)이 각각 0.75^[13]와 0.6^[14]인 두 경우를 고려했다. 본 실험에서 두 경우의 쌍직교 필터를 사용한 것은 다른 모 웨이브릿 선정에 따른 계층 구성시의 영향을 비교하기 위함이다. 그림 4의 저주파 대역 에너지를 비교하는 임계값 TH는 실험적으로 예측 오차가 큰 블록의 값을 고려하여 설정하였으며, 0.7(70%), 0.8(80%), 0.9(90%)를 사용하였다. 또한 탐색 영역을 벗어나는 움직임의 영향을 고려하기 위해서 계층을 3개인 경우와 4개인 경우로 각각 구성했다. 본 방

법과 다른 방법과의 성능 비교를 위해서 전역 탐색법, 3단계 탐색법, 기존의 계층적 움직임 추정법을 사용하였고, 각 방법 모두 원영상의 해상도에서 블록 크기는 16×16, 탐색 영역의 크기는 ±15가 되도록 하였다. 처리된 영상의 객관적 화질 평가 척도로는 식 (15)의 PSNR(peak signal-to-noise ratio)을 사용하였다.

$$PSNR = 10 \log_2 \frac{255^2}{\frac{1}{M \times N} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (f(x, y) - f'(x, y))^2} \quad (dB) \quad (15)$$

여기서, $f(x, y)$ 는 원영상이고 $f'(x, y)$ 는 복구된 영상 신호 값이다. 영상의 해상도는 $M \times N$ 이다.



그림 5. 'Football' 원 영상
Fig. 5. Original 'Football' image.

표 1. 분해와 합성의 필터 계수 : (a) Kernel = 0.75, (b) Kernel = 0.6.

Table 1. Filter coefficients for analysis and synthesis : (a) Kernel = 0.75, (b) Kernel = 0.6.

(a)				(b)			
분해		합성		분해		합성	
$h^*(n)$	$h(n)$	$h^*(n)$	$h(n)$	$h^*(n)$	$h(n)$	$h^*(n)$	$h(n)$
		$h(-3)$	0.0			$h(-3)$	-0.010714286
$h^*(-2)$	-0.125	$h(-2)$	0.0	$h^*(-2)$	-0.05	$h(-2)$	0.053571429
$h^*(-1)$	0.25	$h(-1)$	0.25	$h^*(-1)$	0.25	$h(-1)$	0.260714286
$h^*(0)$	0.75	$h(0)$	0.5	$h^*(0)$	0.6	$h(0)$	-0.607142857
$h^*(1)$	0.25	$h(1)$	0.25	$h^*(1)$	-0.25	$h(1)$	0.260714286
$h^*(2)$	-0.125	$h(2)$	0.0	$h^*(2)$	-0.05	$h(2)$	0.053571429
		$h(3)$	0.0			$h(3)$	-0.010714286

표 2는 필터, 계층 구성, 임계값을 각각 달리하여 제안한 방법에 적용했을 때의 각 계층에서 그 움직임이 추정되는 블록 수를 나타낸다. 임계값을 0.7(70%)로 했을 경우, 해상도가 낮은 계층에서 추정되는 블록이 많

으므로 계산량 측면에서 이득을 얻을 수 있지만 전역 탐색에 비해 화질이 떨어진다. 반면 임계값이 0.9 (90%)일 경우에는 해상도가 높은 계층에서 추정되는 블럭수가 전역 탐색과 비슷하게 되므로 계층적 움직임 추정 목적에 부합되지 않는다. 따라서 임계값의 결정을 0.8(80%)로 하는 것이 계층적 움직임 추정을 가장 크게 얻으면서 비교적 우수한 화질을 얻을 수 있었다.

표 2. 각 계층에서의 움직임 추정 블럭 수
Table 2. Block number of motion estimation in each layer.

		kernel=0.75			kernel=0.6		
		TH=0.7 (70%)	Th=0.8 (80%)	Th=0.9 (90%)	TH=0.7 (70%)	TH=0.8 (80%)	TH=0.9 (90%)
3계층 구성	계층2	1060	547	149	651	332	27
	계층1	290	803	1201	699	1018	1323
4계층 구성	계층3	644	362	173	505	244	70
	계층2	507	251	79	213	133	7
	계층1	199	737	1098	632	973	1273

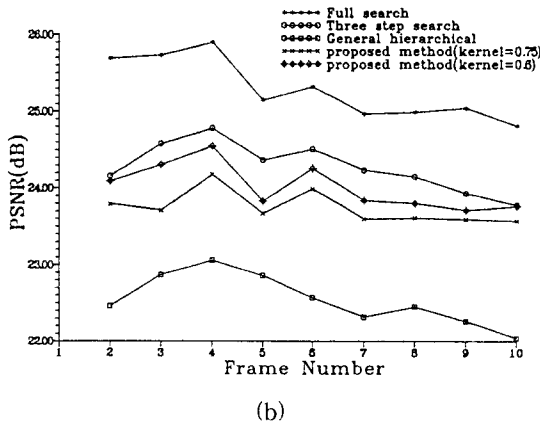
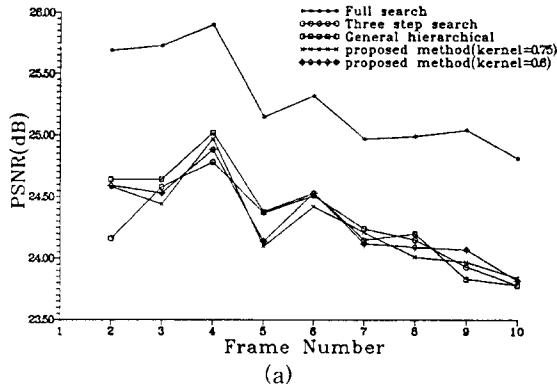
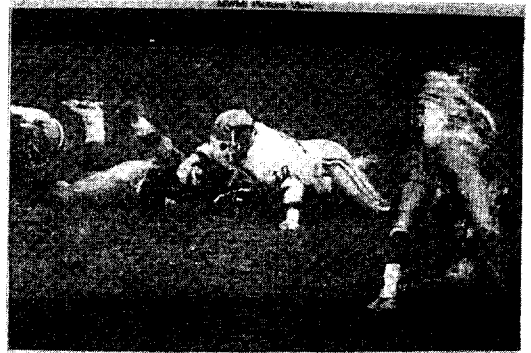


그림 6. 각 기법의 PSNR: (a) 3계층, (b) 4계층
Fig. 6. PSNR of each method: (a) 3 layer, (b) 4 layer.

그림 6은 임계값을 0.8로 고정하고 계층 수를 3개와 4개로 달리 구성한 경우의 실험 결과이다. 일반적으로 계층적 움직임 추정의 경우는 계층 수를 적게 할수록 좀더 정확한 움직임을 찾을 수 있으나 계산량이 늘어난다. 그림 6도 역시 같은 결과를 보이고 있다. 기존의 계층적 움직임 추정법과 제안한 방법을 비교해 보면, 기존의 방법은 계층수가 3인 경우에 비해 4인 경우 현저한 화질 저하를 보이나 본 방법의 경우 계층수가 3인 경우와 4인 경우 큰 차이를 보이지 않으며 이는 블럭 단위로 국부 영상 특성을 고려하여 실제로 움직임 추정에 참여하는 계층 수를 적응적으로 결정한 것의 효과이다. 계산량 측면까지 고려한다면 본 방법의 경우 계층 수를 4로 한 경우에 이득이 더 크다.



(a)



(b)

그림 7. 결과 영상: (a) Kernel=0.75, (b) Kernel=0.6
Fig. 7. Result images: (a) Kernel=0.75, (b) Kernel=0.6.

두 경우의 쌍직교 필터를 사용하여 다해상도 영상 계층을 구성한 실험 결과를 살펴보면, 3계층을 구성했을 때는 LPM 커널이 0.75인 경우가 움직임 추정에 소

오되는 계산량은 적지만 객관적 화질은 두 경우 모두 비슷하다. 4계층을 구성했을 때는 LPM 커널이 0.6의 경우에 해상도가 높은 계층에서 추정되는 블럭수가 더 많아 화질이 좋지만 계산량이 증가되었다. 따라서 계층 구성시 사용하는 필터의 선택이 화질에 미치는 영향은 거의 없다고 할 수 있다. 그림 7은 3개의 계층을 사용하고 두 종류의 필터를 적용한 경우의 결과 영상이다.

표 3. 하나의 움직임 벡터를 추출하는데 필요한 계산량

Table 3. The amount of the computation for each motion estimation.

(N=16, P=15)

	더하기 연산	절대값 연산
전역 탐색법	$(N^2 - 1) \times (2P + 1)^2$ 245055	$N^2 \times (2P + 1)^2$ 246016
5단계 탐색법	$(N^2 - 1) \times (9 \times 5 - 4)$ 10455	$N^2 \times (9 \times 5 - 4)$ 10496
기존의 계층적 탐색법 (4계층)	$[(N/8)^2 - 1 + (N/4)^2 - 1 + (N/2)^2 - 1 + N^2 - 1] \times 3^2$ 3024	$[(N/8)^2 + (N/4)^2 + (N/2)^2 + (N)^2] \times 3^2$ 3060
제안한 방법 (4계층)	모든 블럭이 계층3부터 움직임 추정되는 경우(최선의 경우)	
	$[(N/8)^2 - 1 + (N/4)^2 - 1 + (N/2)^2 - 1] \times (2 + 1)^2$ 729	$[(N/8)^2 + (N/4)^2 + (N/2)^2] \times (2 + 1)^2$ 756
	모든 블럭이 계층1에서 움직임 추정되는 경우(최악의 경우)	
$[(N/2)^2 - 1] \times (7 + 1)^2$ 4032	$(N/2)^2 \times (7 + 1)^2$ 4096	

계산량의 관점에서 살펴본 결과를 표 3에 보였다. 표 3은 전역 탐색법, 5단계 탐색법, 기존의 계층적 움직임 추정법, 제안한 방법을 사용해서 움직임 벡터를 추출할 때 필요한 더하기 연산량과 절대값 연산량을 블럭 당의 값으로 나타낸 것이다. 이 때, 블럭의 크기는 $N \times N$ 이고, 탐색 영역의 범위는 전역 탐색법의 경우 $\pm P$, 계층적 움직임 추정법의 경우 각 계층에서 ± 1 이다. 표 3에 $N=16, P=15$ 일 때의 계산량 예를 보였다. 제안한 방법에서, 모든 블럭이 해상도가 가장 낮은 계층3부터 계층적 움직임 추정된다고 하면 계산량은 최소가 되며, 729번의 더하기 연산과 756번의 절대값 연산이 필요하다. 또한 모든 블럭이 계층1에서 움직임 추정이 이루어진다고 하면 계산량은 최대가 되고, 4032번의 더하기

연산과 4096번의 절대값 연산이 필요하다. 즉, 제안한 방법의 경우 계산량은 각 블럭이 어느 계층부터 움직임 추정이 되느냐에 따라 최소값과 최대값 사이에 놓이게 된다. 기존의 계층적 움직임 추정법과 제안한 방법의 경우, 계층 구성시 탭수가 K 인 필터를 사용하여 계층을 구성하면, 해상도가 한단계 낮은 상위 계층을 구성할 때마다 화소당 $2 \times K$ 번의 곱하기 연산과 $2 \times (K-1)$ 번의 더하기 연산이 추가된다. 예를 들어, 계층을 4개로 구성하기 위해 5탭의 분해 필터를 사용하면 블럭당 3360번의 곱하기 연산과 2688번의 더하기 연산이 필요하다.

제안한 방법에서 사용되는 메모리는 그림2와 그림3에서 볼 수 있듯이 기존의 계층적 움직임 추정보다 적다. 그림 3은 4개의 계층으로 이루어졌으나 계층0는 움직임 추정에 사용되지 않는 가상의 계층이고 실제로는 계층1, 계층2, 계층3만이 사용된다. 임의의 계층에서의 저주파 대역과 고주파 대역의 에너지 비는 그 계층의 저주파 대역 에너지와 그 보다 해상도가 한 단계 낮은 바로 상위 계층의 저주파 대역 에너지의 비로 구할 수 있기 때문에 계층2와 계층3의 고주파 대역만큼 메모리가 절약된다. 또한 원영상을 저장하는 영상 버퍼를 1차 웨이브릿 변환후 계층 1의 버퍼로 다시 사용할 수 있다. 따라서 기존의 계층적 움직임 추정법에 비해 $[5/4 \times \text{원영상 해상도}]$ 만큼의 메모리를 줄일 수 있다. 이전 프레임에 대한 메모리까지 고려하면 이것의 두배만큼 움직임 추정부에서 사용되는 메모리가 줄어든다. 필터링에 필요한 1차원 버퍼의 크기는 기존의 계층적 움직임 추정법의 경우와 제안한 방법의 경우 모두 같다.

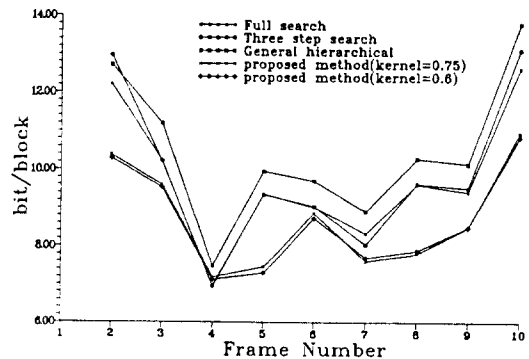


그림 8. 각 기법의 블럭당 평균 비트율
Fig. 8. Average bits per block in each method.

그림 8은 각 방법의 블럭당 평균 비트율을 나타낸다.

블럭당 평균 비트율은 움직임 추정후의 예측 오차를 DCT하고 선형 양자화하여 구한 비트 수와 움직임 벡터에 대한 비트 수를 더하여 평균한 것이다. 본 방법은 국부 영상 특성을 고려한 계층적 움직임 추정으로 인한 예측 이득과 최종 움직임을 1차 웨이브릿 변환된 계층의 저주파 대역에서 구함으로써 얻은 움직임 벡터의 절약으로 비트율을 낮출 수 있다. 또한 수신단에서는 역 웨이브릿 변환으로 영상을 복구하게 됨으로써 블럭 단위 움직임 추정으로 인한 블럭화 현상을 완화시켰다.

VI. 결 론

본 논문에서는 웨이브릿 변환을 이용하여 영상을 다해상도 계층으로 분해하고, 영상 국부의 저주파 대역 성분과 고주파 대역 성분을 고려하여 움직임 추정에 참여하는 계층 수를 적응적으로 결정함으로써 국부 최소의 문제를 해결하고자 하였다. 실제 영상을 적용하여 실험한 결과에 따르면 화질, 계산량, 비트율의 측면을 함께 고려할 때 기존의 계층적 움직임 추정법과 3단계 추정법에 비해 비교적 우수한 것으로 나타났고, 기존의 계층적 움직임 추정법에 비해 적은 메모리를 사용하여 본 방법을 구현할 수 있다. 또한 최종 움직임을 1차 웨이브릿 변환된 계층에서 구하고, 수신단에서 영상 복구시 역웨이브릿 변환을 하게됨으로써 인간 시각에 민감한 블럭화 현상을 완화시켰다.

참 고 문 헌

[1] H. G. Musmann, P. Pirsch, and H. J. Grallert, "Advanced in picture coding," *Proc. IEEE*, vol. 73, no. 4, pp. 523-548, Apr. 1985.

[2] T. Koga, K. Iinuma, A. Hirano, Y. Ijima, and T. Ishiguro, "Motion-compensated interframe coding for video conferencing," in *Proc. Nat. Telecommun. Conf.*, pp. G5.3.1-5.3.5, 1981.

[3] M. Biering, "Displacement estimation by hierarchical blockmatching," in *Proc. Visual Commun. and Image Process.*, vol. SPIE 1001, pp. 942-951, Nov. 1988.

[4] 천 강욱, 나 종범, "오차함수의 계층적 근사를 이

용한 고속 블럭 알고리즘에 의한 움직임 벡터의 추정", 제4회 신호처리 합동학술대회 논문집, pp. 5-10, 1991년 9월

- [5] 남 권문, 김 준식, 박 래홍, 심 영석, "평균 피라미드를 이용한 계층적 고속 이동벡터 추정", 대한전자공학회 논문지, 제30권 B편 제6호, pp. 35-48, 1993년 6월
- [6] 김 인겸, 정 준용, 류 충일, 이 광기, 박 규태, "스플라인 웨이브릿 변환을 이용한 영상의 다해상도 부호화에 관한 연구", 한국통신학회 논문지, 제19권 제12호, pp. 2313-2327, 1994년 12월
- [7] 김 웅태, 김 형명, "Biorthogonal 웨이브릿 변환 영역에서 HVS를 이용한 적응 영상 부호화 기법", 한국통신학회 논문지, 제18권 제 10호, pp. 1469-1482, 1993년 10월
- [8] 장 준필, 김 진태, 장 태규, 최 중수, "Wavelet 변환을 이용한 계층적 움직임 추정", 제5회 신호처리 합동학술대회 논문집, pp. 406-410, 1992년 9월
- [9] O. Rioul and M. Vetteri, "Wavelets and signal processing," *IEEE Signal Processing Mag.*, vol. 8, no. 4, pp. 14-38, Oct. 1991.
- [10] S. G. Mallat, "Multiresolution approximations and wavelet orthogonal bases of $L_2(\mathbb{R})$," *Tran. Amer. Math. Soc.*, vol. 315, no. 1, pp. 969-87, Sep. 1989.
- [11] J. C. Feauveau, P. Mathieu, M. Barlaud, and M. Antonini, "Recursive biorthogonal wavelet transform for image coding," in *Proc. ICASSP*, pp. 2649-2652, 1991.
- [12] H. Hölzlwimmer, A. V. Brandt, and W. Tengler, "A 64 Kbit/s motion compensated transform coder using vector quantization with scene adaptive codebook," in *Proc. Intl. Conf. on Commun.*, pp. 151-156, 1987.
- [13] C. H. Keon, "Subband image coding with biorthogonal wavelets," *IEICE Trans. Fundamentals.*, vol. E75-A, no. 7, pp. 871-881, Jul. 1992.
- [14] M. Antonini, M. Barlaud, P. Marthieu, and I. Daubechies, "Image coding using vector quantization in the wavelet transform domain," in *Proc. ICASSP*, pp. 2297-2300, 1992.

저 자 소 개



金 珍 台(正會員)
 第 26卷 第 9號 參照
 1995년 3월 ~ 현재 한서대학교 전
 산정보학과 전임강사. 주관심 분야
 는 초저속 동영상 부호화, 물체 지
 향 부호화, 움직임 추정 및 보상 등
 임



張 峻 弼(正會員)
 1966년 7월 18일생. 1991년 2월
 중앙대학교 전자공학과 졸업(공학
 사). 1993년 2월 중앙대학교 대학
 원 전자공학과 졸업(공학석사).
 1993년 2월 ~ 현재 LG전자 멀티
 미디어 연구소 근무. 주관심 분야
 는 영상 신호처리, 영상 압축 및 복원, 위성방송 시스템
 등임



金 東 郁(正會員)
 1987년 2월 성균관대학교 전자공
 학과 졸업(공학사). 1992년 2월 중
 앙대학교 대학원 전자공학과 졸업
 (공학석사). 1996년 8월 중앙대학
 교 대학원 전자공학과 졸업(공학박
 사). 주관심 분야는 영상처리 및 부
 호화, 컴퓨터 시각 등임

崔 宗 秀(正會員) 第 33卷 第 4號 參照
 1981년 9월 ~ 현재 중앙대학교 전자공학과 교수. 주관
 심 분야는 컴퓨터 시각, 영상 부호화, 적외선 신호 처리
 등임