

# 웨이브렛 변환영역에서의 2단계 가변 블록 다해상도 움직임 추정

正會員 김성만\*, 이규원\*\*, 정학진\*, 박규태\*\*

## Two-Stage Variable Block-Size Multiresolution Motion Estimation in the Wavelet Transform Domain

SeongMan Kim\*, KyuWon Lee\*\*, HakJin Chong\*, KyuTae Park\*\* *Regular Members*

### 요 약

본 논문에서는 웨이브렛 분할 영역의 화면간 부호화기에서 사용될 수 있는 2단계 가변 블록 다해상도 움직임 추정 알고리즘(two-stage variable block-size multiresolution motion estimation algorithm)을 제안한다. 이 방법은 총 전송 비트율의 최소화를 통하여 움직임 벡터와 예측 오차간에 효율적인 비트 할당을 제공한다. 제안한 방법은 움직임 추정을 2단계로 분리하여 수행되며, 첫 단계의 움직임 추정부는 따로 독립적으로 적용될 수도 있다. 첫 단계 움직임 추정부에서는 평활한 움직임 필드(smooth motion field) 및 낮은 비트율의 변위차영상(displaced frame difference: DFD)을 제공하기 위한 새로운 방법을 소개한다. 두 번째 단계의 움직임 추정은 상세 영상 영역에서는 더욱 정확한 움직임 벡터를 갖도록 허용하고, 반면 균일 영상 영역에서는 움직임 벡터의 개수를 줄인다. 이러한 방법은 병합 연산에 기반을 둔 4진트리의 상향식 구축에 의하여 움직임 벡터와 DFD로 구성되는 총 전송 비트율을 최소화함으로써 달성된다. 여기에서 노드 병합에 대한 척도는 엔트로피가 사용된다. 본 논문에서 제안한 2단계 가변 블록 다해상도 움직임 추정 기법은 웨이브렛 분할 영역의 화면간 부호화기에서 기존의 다해상도 움직임 추정 기법보다 최대 0.276 bpp 만큼 낮은 비트율을 발생한다.

### ABSTRACT

In this paper, the two-stage variable block-size multiresolution motion estimation algorithm is proposed for an interframe coding scheme in the wavelet decomposition. An optimal bit allocation between motion vectors and the prediction error in sense of minimizing the total bit rate is obtained by the proposed algorithm. The proposed algorithm consists of two stages for motion estimation and only the first stage can be separated and run on its own.

\*한국통신 전송기술연구소

\*\*연세대학교 대학원 전자공학과

論文番號: 97072-0221

接受日字: 1997年 2月 21日

The first stage of the algorithm introduces a new method to give the lower bit rate of the displaced frame difference as well as a smooth motion field. In the second stage of the algorithm, the technique is introduced to have more accurate motion vectors in detailed areas, and to decrease the number of motion vectors in uniform areas. The algorithm aims at minimizing the total bit rate which is sum of the motion vectors and the displaced frame difference. The optimal bit allocation between motion vectors and the displaced frame difference is accomplished by reducing the number of motion vectors in uniform areas and it is based on a bottom-up construction of a quadtree. An entropy criterion aims at the control of merge operation. Simulation results show that the algorithm lends itself to the wavelet based image sequence coding and outperforms the conventional scheme by up to the maximum 0.28 bpp.

## I. 서 론

최근 널리 사용되고 있는 국제 표준인 H.261, MPEG-1 및 MPEG-2 등은 화면간 시간적 부호화 방법으로 움직임 보상(motion compensation) 기법을 사용하고, 화면내 공간적 부호화 방법으로 이산 역현 변환(discrete cosine transform: DCT)을 사용하는 동일한 범용 구조로 이루어져 있다. 움직임 보상 기법은 화면간 부호화 기에서 효율적인 이동량 예측으로 사용되며, 예측 오차인 DFD를 발생한다. DFD는 하드웨어 구현 용이성 때문에 작은 고정 블록(fixed block)으로 분할된 후 DCT를 거쳐 주파수 영역으로 변환된다. 이러한 블록 변환 부호화 방법은 특히 저 비트율 응용 분야에서 '블록화 현상(blocky effect)'을 초래하고, 이동량 검출 후 생기는 DFD에 대한 블록 변환에서는 낮은 공간적 상관 계수(correlation coefficient) 때문에 에너지의 압축 효율을 크게 저하시킨다. 변환 부호화 기법은 일반적으로 공간적 상관 계수 값이 0.9이상인 단일 계조도(gray level) 영상에 적용될 경우 좋은 압축 효율을 낸다[1][2]. Strobach[3]는 DFD에 대한 공간적인 상관 계수 값을 0.3에서 0.5 정도로 측정한 바 있다.

웨이브렛 변환 기법은 블록화 현상을 없애 주며, 영상 압축에 대한 성능 면에서 DCT보다 우수하다[4][5]. Zafar와 Zhang[6][7][8]은 실시간 비디오 압축을 위하여 웨이브렛 분할 영역에서 가변 블록(variable block-size)을 이용한 다해상도 움직임 추정(multiresolution motion estimation) 기법을 제안하였다. 웨이브렛 피라미드의 다른 계층상에 위치한 물체 움직임은 동일한 움직임 구조가 여러 스케일 및 주파수로 표현된 것이므로 상관관계가 매우 높다. 다해상도 움직임 추정 알고리즘[9]은 피라미드의 맨 위층에 위치한 가장 낮은 해상도의 부대역 영상에서 추정한 움직

임 필드(motion field)를 이용하여 피라미드의 나머지 계층에 위치한 부대역 영상의 움직임 벡터를 계산한다. 다해상도 움직임 추정 기법은 웨이브렛 표현의 고유 특성에 적합할 뿐만 아니라, 탐색(searching)과 정합(matching) 시간을 줄여 주고 평활한 움직임 필드(smooth motion field)를 제공한다.

하지만, 다해상도 움직임 추정 기법의 단점은 가장 낮은 해상도에 위치한 저주파 부대역 영상의 움직임 벡터를 자체 부대역 영상의 멀티스케일된 정보만으로 결정하고, 이렇게 구한 움직임 벡터는 다른 해상도의 타 부대역 영상의 움직임 벡터를 구하기 위한 초기 추정치(initial estimate)로 사용된다. 만약 멀티스케일된 정보로 결정된 움직임 벡터가 잘못되면 자체 부대역 영상 뿐만 아니라 다른 부대역 영상의 예측 오차에도 영향을 미쳐 전체 성능이 저하된다.

움직임 추정 면에서 살펴보면, H.261, MPEG-1, MPEG-2, 및 다해상도 움직임 추정 기법은 블록 정합 알고리즘[10]에 기반을 두고 있다. 비록 블록 정합 기법이 현재 동영상 부호화 응용 분야에서 가장 널리 사용되고 있지만, 그것은 불확실한 움직임 필드(unreliable motion field), 인조 블록(block artifact) 및 움직이는 물체 윤곽선 상의 불충분한 움직임 보상 예측(poor motion compensated prediction along moving edges)과 같은 중대한 단점을 지니고 있다[11]. 더욱 확실한 움직임 필드를 구하기 위하여 계층적 블록 정합 알고리즘[12]이 제안되었지만, 인조블록 문제를 해결하고 움직이는 물체 윤곽선상에서 더욱 정확한 예측을 제공하는 방법은 움직임 필드를 분할하는 것이다. 이러한 움직임 추정 방법으로 가변 크기 블록 정합(variable size block matching)[13], 움직임 필드 세분화(motion field segmentation)[14], 국부적 적용 멀티그리드 블록 정합(locally adaptive multigrid block

matching)[15] 등이 제안되었다.

본 논문은 앞에서 언급한 바와 같이 현재 널리 사용되고 있는 동영상 부호화 방법에서 발생하는 두 가지 측면의 문제점을 해결할 수 있는 움직임 추정 알고리즘을 제안한다. 첫 번째 측면에서는 공간적인 영역에서의 비효율적인 에너지 압축 및 블록화 현상을 해결하기 위하여 원영상에 대하여 전역 변환(global transform) 특성을 갖는 웨이브렛 분할을 시도하고, 웨이브렛 분할 영역에서 움직임 추정을 수행한다. 두 번째 측면은 불확실한 움직임 필드, 인조 블록 및 윤곽선상의 불충분한 움직임 보상 예측을 해결하기 위하여 한편으로는 웨이브렛 분할 영역의 다른 해상도 간에 다른 한편으로는 모든 부대역 영상 내에서도 가변 블록을 이용한 블록 정합을 실시한다. 이러한 이중 가변 블록을 사용한 다해상도 움직임 추정 알고리즘을 이후부터는 2단계 가변 블록 다해상도 움직임 추정으로 명명한다. 특히 제안한 방법의 첫 단계 움직임 추정부에서는 평활한 움직임 필드 및 낮은 비트율의 DFD를 제공하며, 둘째 단계에서는 상세 영상 영역에서 더욱 정확한 움직임 추정을 하고, 반면에 균일 영상에서 움직임 벡터의 개수를 줄이는 새로운 방법을 제안한다.

## II. 가변 블록 다해상도 움직임 추정(기존 알고리즘)

시간  $t$ 의 이차원 영상  $f(x, y)$ 는 해상도 깊이  $m$ 인 웨이브렛 분할에 의하여 여러 개의 부대역 영상으로 표현된다.

$$f \rightarrow \{P_m f, [D_m^k f]_{k=1, 2, 3, \dots}, [D_1^k f]_{k=1, 2, 3}\} \quad (1)$$

$m > 0$ 에 대하여,  $P_j f: j=0, 1, \dots, m$ 은 해상도  $\{1, \dots, 2^{-m}\}$ 에서 원영상  $f$ 에 대한 근사신호(approximation)의 집합을 나타낸다. 즉  $P_m f$ 는 가장 낮은 해상도  $2^{-m}$ 에서 원영상  $f$ 에 대한 근사신호이고,  $P_0 f$ 는 원영상  $f$ 를 표현한다.  $D_j^k f$ 는 해상도  $2^{-j}$ 에서의 상세 신호(detail signal)로 정의되며,  $P_{j-1} f$ 와  $P_j f$ 간에 차 정보(difference information)로부터 생성된다.

2차원 웨이브렛 변환은 2차원 영상  $f$ 를 가로와 세로 방향으로 분리하여 1차원 웨이브렛 필터에 의하여 계산되기 때문에 해상도  $2^{-(j-1)}$ 에서의  $P_{j-1} f$ 는 해상

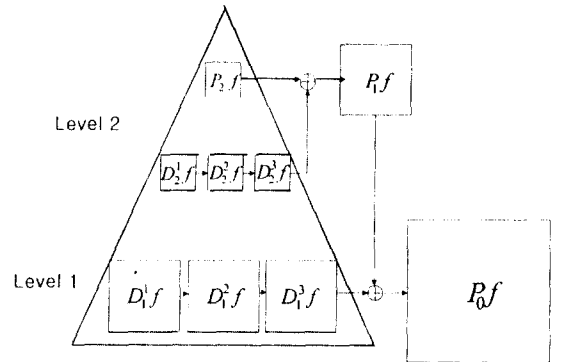


그림 1. 웨이브렛 분할과 복원에 대한 피라미드 구조  
Fig. 1 Pyramid structure of the wavelet decomposition and reconstruction

도  $2^{-j}$ 에서  $P_j f, D_j^1 f, D_j^2 f$  및  $D_j^3 f$ 로 분할된다. 여기에서 영상  $P_j f$ 는 가장 낮은 주파수 영역에 해당되는 부대역 영상이며,  $D_j^1 f$ 는 수직 방향의 고주파 성분(가로 방향의 경계선),  $D_j^2 f$ 는 수평 방향의 고주파 성분(세로 방향의 경계선) 및  $D_j^3 f$ 는 양방향의 고주파 성분(대각선 방향의 경계선)에 해당되는 부대역 영상이다. 그림 1은 해상도 깊이 2를 적용하여 원영상  $P_0 f$ 를 다해상도를 갖는 피라미드 형태의 계층적인 구조로 표현한 것이다.

여기에서 가장 높은 계층에 위치한 거친(coarse) 해상도의 부대역 영상은 물체의 문맥 정보를 제공하며, 낮은 계층으로 갈수록 해상도가 커지면서 물체의 상세 정보를 제공한다. 또한 다른 계층의 각 부대역 영상 내에서 동일한 장소에 위치한 블록은 다른 스케일(scale)과 다른 주파수 범위의 동일한 움직임 구조 특성을 묘사하기 때문에 상관관계가 매우 높다. 이러한 특성 때문에 탐색 및 정합에 소요되는 시간이 적고 평활한 움직임 필드를 구할 수 있는 다해상도 움직임 추정 방법이 효과적으로 사용된다.

먼저, 움직임 추정 방법을 설명하기 위해 기호 정의를 하면 다음과 같다. 해상도 깊이  $m$ 이 2에 대하여, 시간  $t$ 와 임의의 위치  $s=(s_x, s_y)$ 에서의  $P_2 f$ 의 화소값을  $I_{2, s}^0(s)$  또는  $I_{2, s}^0(s_x, s_y)$ 로 표시하면,  $D_2^k f$ 의 화소값은  $I_{j, s}^k(s)$  또는  $I_{j, s}^k(s_x, s_y)$ 로 표현된다. 여기에서  $I_{j, s}^k(s)$ 에서 아래첨자 ' $j$ ' 및 위첨자 ' $k$ '는 각각 해상도 레벨과 주파수 성분을 나타내고,  $I_{2, s}^0(s)$ 의 위첨자 '0' 및 아래

첨자 '2'는 저주파 대역과 해상도 깊이 2를 나타낸다. 따라서 시간  $t$  및 위치  $s$ 에서의 원상  $f$ , 즉  $P_0 f$ 는 다음과 같이 정의할 수 있다.

$$P_0 f \equiv I_{0,t}^0(s) \quad (2)$$

시간  $(t-1)$ 에서 시간  $t$ 동안 물체 움직임을 나타내는 움직임 벡터를  $v=(v_x, v_y)$ 로 정의하면,  $P_2 f$ ,  $D_j^k f$ 의 움직임 벡터는  $v_2^0$  및  $v_j^k$ 로 정의된다. 여기에서 위 첨자 및 아래 첨자는 주파수 방향 성분 및 해상도 레벨을 의미한다. 그리고 블록의 크기인  $X \times Y$ 을 측정 윈도우(measurement window)  $W$ 로 정의하고 탐색 영역(search area)을  $\Omega$ 로 정의하면,  $v \in \Omega$ 가 된다.

### 2.1 블록 정합 알고리즘

이 알고리즘은 최근 ITU-T의 H.261 권고와 ISO/IEC의 MPEG 표준화에서 채택된 DCT를 이용한 부호화 방법에 매우 적합하며, 또한 가변 블록 다해상도 움직임 추정 알고리즘 및 본 논문의 화면간 움직임 추정에서도 사용되고 있다. 블록 정합 알고리즘 원리[10]는 최소의 오차를 발생하면서 이전 프레임  $I_{0,t-1}^0(s-v)$ 로부터 현재 프레임  $I_{0,t}^0(s)$ 를 복원하는 움직임 벡터  $v$ 를 발견하는 것이다. 즉 다음과 같이 정합 기준(matching criterion)을 최소화시키는 이전 프레임의 탐색 영역 내의 움직임 벡터를 찾는다.

$$v = \underset{v \in \Omega}{\operatorname{argmin}} \sum_{s \in W} \|I_{0,t}^0(s) - I_{0,t-1}^0(s-v)\| \quad (3)$$

여기에서 정합 기준은 최대 상호 상관(maximum cross correlation: MCC), 최소 평균 자승 오차(minimum mean square error: MMSE) 및 최소 절대 오차(minimum absolute difference: MAD) 등이 있으나, 이후부터는 MAD를 사용한다.

식 (3)의 정합 기준에 대한 최소값을 발견할 수 있는 방법은 탐색 영역 내의 이산 후보 변위에 대하여 완전 탐색(exhaustive search)을 하는 것인데, 이러한 기법을 완전 탐색(full-search) 알고리즘이라 부른다.

### 2.2 가변 블록 다해상도 움직임 추정

가변 블록 다해상도 움직임 추정 방법[6][7][8]은 웨이브렛 피라미드 구조에서 가장 상위 계층에 위치한

$P_2 f$ 에서 제일 먼저 블록 정합 알고리즘을 이용하여 움직임 추정을 한다. 피라미드의 하위 계층의 움직임 벡터는 상위 계층에서 추정된 움직임 정보를 초기 추정치(initial estimate)로 사용하여  $D_j^k f$ 의 움직임 벡터를 상세하게 계산하는 것이다. 이 방법은 웨이브렛 표현의 본질적인 구조를 이용하기 위한 것으로, Zhang과 Zafar는 해상도 레벨  $m$ 인 웨이브렛 피라미드에서  $j$ 번째 레벨의 부대역 영상에 대하여  $p \cdot 2^{m-j} \times q \cdot 2^{m-j}$  크기의 가변 블록을 사용하여 블록 정합을 실시한다. 여기에서  $p$ 와  $q$ 는 가장 낮은 해상도의 부대역 영상에 해당되는 최소 블록의 크기를 의미한다.

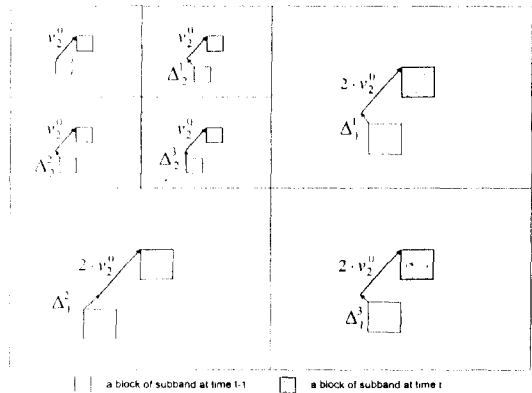


그림 2. 가변 블록 다해상도 움직임 추정  
Fig. 2 Variable block-size multiresolution motion estimation

Zafar가 제안한 움직임 추정 방법은 해상도 깊이 2에 대하여 그림 2처럼 설명된다. 가변 블록 다해상도 움직임 추정 방법에서는 제일 먼저 가장 높은 레벨에 위치한  $v_2^0$ 을  $2 \times 2$  크기의 블록과 탐색 영역  $\Omega$ 에 대하여 완전 탐색 알고리즘을 적용하여 계산한다.  $v_2^0$ 은  $D_j^k f$ 의 움직임 추정에서 초기 바이어스로 사용된다.  $j$ 번째 레벨의  $v_j^k$ 는  $v_2^0$ 을  $2^{2-j}$ 배 하여 초기 바이어스로 사용하고, 탐색 영역  $\Omega$  내에서 완전 탐색을 수행함으로써 더욱 정확하게 구해진다. 여기에서  $\Omega'$ 는 가장 낮은 해상도의 탐색 영역인  $\Omega$ 와 동일하지만, 블록의 크기 및 해상도가  $2^{2-j}$ 만큼 증가한 이유로 실제 탐색 영역은 감소한 것이다.  $D_j^k f$ 에서는  $p \cdot 2^{m-j} \times q \cdot 2^{m-j}$  크기의 가변 블록을 사용하고,  $v_2^0$ 은 다음과 같이 계산된다.

$$v_2^0 = \underset{v \in \Omega}{\operatorname{argmin}} \sum_{s \in W} |I_{2,t}^0(s) - I_{2,t-1}^0(s-v)| \quad (4)$$

따라서  $v_j^k$ 는 식 (4)를 사용하여 다음과 같이 구해진다.

$$v_j^k = v_2^0 \times 2^{2-j} + \Delta_j^k \quad (5)$$

여기에서  $\Delta_j^k$ 는 MAD 기준을 사용하여 계산된  $D_j^k f$ 의 오프셋(offset) 움직임 벡터이며,  $v_2^0$ 를 초기 바이어스로 사용하여 더욱 상세하게 다음과 같이 계산된다.

$$\Delta_j^k = \underset{v \in \Omega'}{\operatorname{argmin}} \sum_{s \in W} |I_{j,i}^k(s) - I_{j,i-1}^k(s - 2^{2-j} \times v_2^0 - v)| \quad (6)$$

### III. 2단계 가변 블록 다해상도 움직임 추정(제안 알고리즘)

가변 블록 다해상도 움직임 추정은 가장 낮은 해상도의 움직임 벡터가 잘못된 경우, 여기서 발생한 예측 오차는 모든 다른 부대역 영상까지 전파된다. 또한 다른 해상도의 부대역 영상간에는 가변 블록을 사용하여 블록 정합을 수행하지만, 각각의 부대역 영상 내에서는 고정 크기의 블록 정합을 하기 때문에 균일 영역에서 많은 움직임 정보가 발생되어 전송 비트율이 증가한다. 이러한 단점을 보완시킨 2단계 가변 블록 다해상도 움직임 추정 알고리즘을 제안한다.

#### 3.1 4진트리 표현

제안하는 움직임 추정 방법은 2단계로 구성되며, 첫 번째 단계에서는 웨이브렛 피라미드의 각 부대역 영상을 해상도에 비례하는 최소 블록으로 분할하여 각각의 최소 블록에 대하여 움직임을 추정한다. 두 번째 단계에서는 움직임 벡터의 개수를 줄이기 위하여 인접한 4개의 최소 블록에 대하여 병합 기준을 적용하여 균일한 영역에 해당하는 4개의 인접 블록을 단일 블록으로 합병한다. 따라서 각 부대역 영상은 최소 블록으로 사전에 분할되어 4진트리로 표현되어야 한다.

4진트리 발전에 대한 자세한 내력은 Klinger와 Dyer [19] 및 Samet[20] 논문에서 잘 설명되어 있고, 움직임 추정을 위한 4진트리 적용은 Strobach[3] 논문에서 참조될 수 있다.

시간적으로 연속하는 화면간에 움직임 추정을 위하여 웨이브렛 분할 영역의 각 부대역 영상은 해상도에 비례하는 최소 크기의 가변 블록으로 사전분할된

다. 4진트리는 이러한 가변 블록간에 존재하는 물체 움직임에 대한 상관관계를 쉽게 이용할 있는 구조로 구성된다.

원영상의 단위 블록은 그림 3과 같이 웨이브렛 분할 영역에서 다른 해상도와 다른 주파수 방향을 갖는 가변 블록으로 분할되고, 역과정으로 분할된 가변 블록은 원영상의 단위 블록으로 복원될 수도 있다. 여기에서 웨이브렛 해상도 깊이는 2이다. 원영상의 단위 블록에 해당하는 가변 블록은 비록 다른 해상도와 주파수 성분으로 표현되지만 동일한 물체를 표현하기 때문에 상관관계를 가진다.

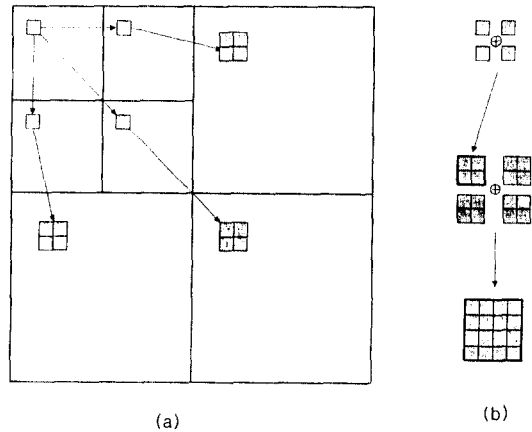


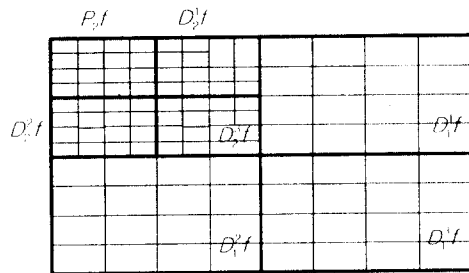
그림 3. 원영상의 단위 블록과 웨이브렛 분할 영역에서의 가변 블록간의 관계 (a) 웨이브렛 분할 영역에서의 가변 블록 (b) 복원된 원영상의 단위 블록

Fig. 3 The relationship between variable size blocks in the wavelet decomposition and unit block of the original image (a) Variable size blocks in the wavelet decomposition (b) The reconstructed unit block of the original image

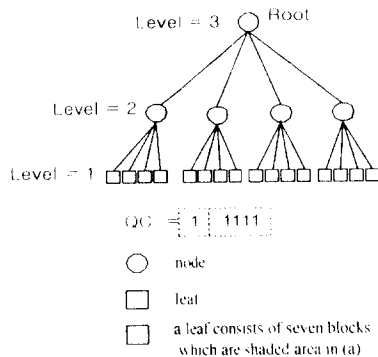
본 논문에서는 각 부대역 영상을  $p \cdot 2^{2-j} \times q \cdot 2^{2-j}$  크기의 가변 블록으로 정규 분할한다. 여기에서 '2'는 웨이브렛 분할의 해상도 레벨,  $j$ 는  $j$ 번째 레벨을 나타낸다. 즉 가장 낮은 해상도에서 최소 크기의 블록은  $p \times q$ 가 되고, 각 부대역 영상은 동일한 블록의 개수를 갖는다. 각 부대역 영상의 동일한 장소에 위치하는 가변 블록은 물체의 움직임에 대한 밀접한 상관관계를 가지며, 이러한 특성을 이용하기 위하여 4진트리 내에서 한 개의 노드로 표현된다. 마찬가지로

원영상에 해당되는 7개의 부대역 영상은 4진트리 내에서 루트 노드로 표시되기 때문에 원영상에 대하여 오직 한 개의 4진트리만이 존재한다. 이러한 방법으로 웨이브렛 분할 영상에 대한 4진트리를 구축하면, 리프 노드는 그림 4처럼 7개의 최소 가변 블록의 집합으로 구성된다. 따라서 전체 리프 노드를 재구성하면 웨이브렛 분할 영상이 복원된다.

만일 각 부대역 영상을 한 개의 4진트리를 표현하면 7개의 4진트리가 발생되기 때문에 전체 노드의 개수도 7배로 증가한다. 일반적으로 노드의 개수가 증가하면 할수록 메모리 용량도 비례적으로 증가하게 되고, 특히 가변 블록간에 존재하는 상관관계를 이용하기 위하여 7개의 4진트리를 독립적으로 탐색하여야 하기 때문에 알고리즘의 수행 시간이 그만큼 늘어나게 된다.



(a)



(b)

그림 4. 웨이브렛 분할 영역의 4진트리 표현 (a) 최소 블록까지 정규 분할된 각각의 부대역 영상 (b) 해당 4진트리

Fig. 4 The quadtree representation for the wavelet decomposition (a) Each subband is decomposed regularly up to the minimum blocks (b) The corresponding quadtree

한편 부호화기에서는 4진트리 부호를 최소의 비트율로써 표현하기 위하여 가장 낮은 레벨에 위치한 리프 노드는 0 부호를 할당하지 않는다. 예를 들어,  $l$  레벨의 4진트리에서  $l$ 에 해당되는 리프 노드는 4진 부호를 할당하지 않는다. 복호화기에서는 역과정으로  $l+1$  레벨에 위치한 노드의 4진 부호가 1인 경우, 4개의 자손 노드를 자동으로 생성하여야 한다.

### 3.2 움직임 추정

이 알고리즘은 단계별로 독립적인 방법으로 수행될 수도 있으며, 각 단계는 다른 목적의 움직임 추정을 한다. 첫 번째 단계의 움직임 추정은 그림 5와 같은 순서로 진행된다.

원영상은 웨이브렛 분할을 두 번 반복 수행함으로써 7개의 부대역 영상으로 분할되고, 이러한 부대역 영상은 피라미드 구조에서 바닥 계층에 3개, 꼭대기 계층에 4개가 그림 1처럼 위치한다. 이러한 웨이브렛 분할 영역은 한 개의 4진트리로 표현된다. 4진트리로 표현된 웨이브렛 분할 영상의 움직임 추정 알고리즘은 다음과 같은 순서로 진행된다. 즉 리프 노드로 순서대로 운행하여 7개의 가변 블록에 대하여 움직임 추정을 수행한다. 각 리프 노드는  $P_2f$ 에 해당되는 한 개의 블록과  $D_2^*f$ 에 해당되는 6개의 블록으로 구성된다. 제안한 움직임 추정 알고리즘은  $v_0^*$ 를 먼저 계산하고, 이것을 초기 바이어스로 이용하여 나머지  $v_1^*$ 를 더

$P_2f$				$D_2^*f$				$D_1^*f$			
1	8	29	36	2	9	30	37	5	12	33	40
22	15	50	43	23	16	51	44				
85	92	57	64	86	93	58	65	26	19	54	47
106	99	78	71	107	100	79	72				
3	10	31	38	4	11	32	39	89	96	61	68
24	17	52	45	25	18	53	46				
87	94	59	66	88	95	60	67	110	103	82	75
108	101	80	73	109	102	81	74				
6	13		34		41			7	14	35	42
27		20			48			28	21	56	49
90		97			62		69	91	98	63	70
111	104				76			112	105	84	77
				$D_2^*f$				$D_1^*f$			

그림 5. 웨이브렛 분할 영역에서 움직임 추정 순서

Fig. 5 The order of motion estimation in the wavelet decomposition domain

욱 상세하게 구한다. 현재의 리프 노드에 대한 움직임 추정이 완료되면 다음 리프 노드로 운행하여 동일한 방법으로 움직임 추정을 수행한다.

그림 5에서 각각의 숫자는 웨이브렛 분할 영상 내의 각각의 블록에 대한 움직임 추정 순서를 나타낸다. 이러한 순서로 첫 번째 단계의 움직임 추정에서는 평활한 움직임 필드와 낮은 비트율의 예측 오차를 제공하기 위하여 4가지의 새로운 기법이 제안되며, 각각의 방법에 대한 자세한 설명이 두 번째 단계의 알고리즘 이전까지 이어진다.

첫째, 가장 낮은 해상도에 위치한 4개의 부대역 영상의 정보를 이용하여  $P_2f$ 의 움직임 벡터를 추정한다.  $P_2f$ 는 어떤 물체에 대한 거친 정보를 제공하는 반면, 가장 낮은 해상도의 수평, 수직 및 대각선 방향의 주파수 성분을 나타내는  $D_2^1f$ ,  $D_2^2f$  및  $D_2^3f$ 는 해당 물체의 윤곽선 성분을 제공한다. 만약  $P_2f$ 의 움직임 벡터를 결정하기 위하여 가장 낮은 해상도의 모든 부대역 영상을 활용하여 움직임을 추정하면, 더욱 낮은 예측 오차가 발생한다. 즉  $P_2f$ 의 움직임 벡터인  $v_2^0$ 는 시간적으로 연속적인 웨이브렛 분할 영상에서  $P_2f$ ,  $D_2^1f$ ,  $D_2^2f$  및  $D_2^3f$  내의 동일 장소에 위치하는 4개의 블록간에 MAD 기준의 완전 탐색 블록 정합에 의하여 계산된다. 이러한 움직임 벡터  $v_2^0$ 는 가장 낮은 해상도에 위치한  $D_2^1f$ ,  $D_2^2f$ ,  $D_2^3f$  및 다른 해상도의  $D_1^1f$ ,  $D_1^2f$  및  $D_1^3f$ 의 움직임 벡터  $v_1^k$ 를 구하기 위한 초기 바이어스로서 사용된다.

초기 바이어스로 사용되는  $v_2^0$ 에 대한 움직임 추정이 그림 6에서 설명된다.  $v_2^0$ 는 가장 낮은 해상도의 4개의 부대역 영상의 정보를 활용하여 결정되며, 다음과 같이 표현된다.

$$v_2^0 = \underset{v \in \Omega}{\operatorname{argmin}} \sum_{s \in W} \sum_{k=0}^3 |I_{2,t}^k(s) - I_{2,t-1}^k(s-v)| \quad (7)$$

둘째,  $D_2^k f$  내의 블록에 대하여 유효성 검사(significance test)를 실시하여, 해당 블록에 대한 움직임 추정 수행 여부를 결정한다. 이것은 해당 블록의 에너지가 선정된 임계치보다 높은가를 점검한다. 만약 시험 결과가 긍정적이면 해당 블록에 대한 움직임 추정은 정상적으로 실시되며, 부정적이면 움직임 추정 대신 복호화기에서 해당 블록 내의 모든 화소값을 0으로 복원한다. 본 논문에서는 임계치를 양자화기의 영구간

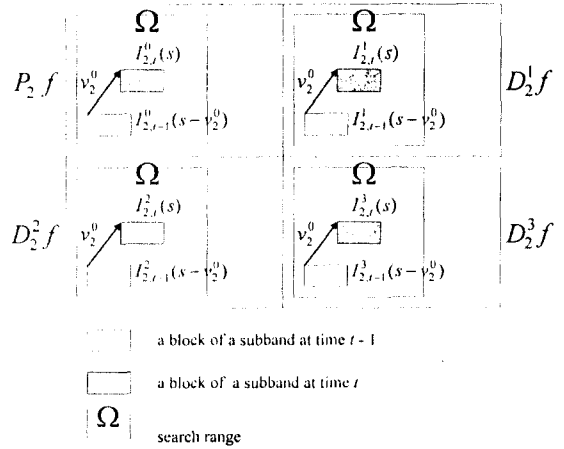


그림 6. 저주파수 부대역 영상에서 제안한 움직임 추정 방법  
Fig. 6 The proposed motion estimation in the subband with the lowest frequency

(deadzone)값으로 선택한다. 웨이브렛 분할 영역에서는 움직임이 적은 영상일수록  $P_2f$ 를 제외한  $D_2^k f$ 내의 블록은 미세한 에너지 분포만을 갖는다. 이러한 블록은 움직임 보상 과정에서 자체 에너지보다 더 큰 예측 오차가 발생할 수 있기 때문에 움직임 추정 대신 추가 심벌로 정의되어 복호화기로 전송되고, 복호화기에서는 추가 심벌을 수신하면 해당 블록 내의 모든 화소값을 0으로 복원한다.

셋째, 웨이브렛 분할 영역의 특성을 고려하여  $P_2f$ 를 제외한  $D_2^k f$ 내의 블록에 대하여 양자화를 고려한 MAD 기준의 블록 정합을 수행한다. 그림 7처럼 기존의 MAD 기준의 블록 정합 알고리즘의 결과에서 서로 다른 두 개의 움직임 벡터가 동일한 예측 오차를 발생한다고 가정하자. 예측 오차에 대하여 균일 양자화 구간 2를 적용하면, 'case 2'의 예측 오차는 'case 1'보다 낮은 에너지를 발생시킨다. 비록 양자화를 고려한 MAD 기준은 물체 움직임을 정확하게 결정하지는 않지만, 예측 오차를 감소시켜 DFD 비트율을 감소시킨다. 본 논문에서 제안하는 움직임 추정 알고리즘에서는  $P_2f$ 의 움직임 벡터만이 초기 바이어스로서 사용되기 때문에  $D_2^k f$ 의 움직임 추정에서 발생한 오차는 다른 부대역 영상으로 전파되지 않는다. 따라서  $D_2^k f$ 의 움직임 추정은 정확한 움직임 벡터 대신 더욱 낮은 예측 오차를 발생하는 움직임 벡터를 선택한다.

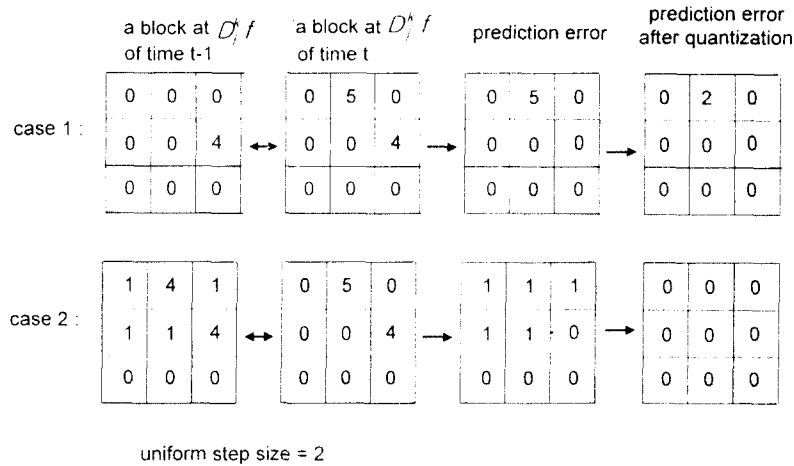


그림 7. 양자화를 고려한 MAD 기준의 블록 정합 예

Fig. 7 An example of a block matching with the MAD criterion reflecting a uniform quantization

넷째, 움직임 벡터에 대한 엔트로피 부호화를 고려하여 평활한 움직임 필드를 얻을 수 있는 움직임 추정을 수행한다.  $D_t^k f$  내의 어떤 블록이 동일한 예측 오차를 발생하는 여러 개의 움직임 벡터를 갖는 경우, 해당 후보 움직임 벡터와  $P_2 f$ 에서 대응되는 블록의  $v_2^k$  간에 노름(norm)을 계산하여 최소가 되는 움직임 벡터  $v_1^k$ 를 다음과 같이 선정한다.

$$v_1^k = \underset{v \in C}{\operatorname{argmin}} \|v_2^k - v\| \quad (8)$$

여기에서  $C$ 는 후보 움직임 벡터의 집합을 의미한다.  $D_t^k f$ 에서 양자화를 고려한 MAD 기준을 이용하여 블록 정합을 수행하면, 동일한 예측 오차를 갖는 여러 개의 후보 움직임 벡터가 동시에 발생될 수가 있다. 이러한 현상은 전 화면과 현재 화면의 블록 정합이 주로 윤곽선 정보만으로 비교되고,  $D_t^k f$ 는  $P_2 f$ 에 비하여 적은 에너지 분포를 갖는 특성 때문에 발생된다. 따라서 식 (8)에 의하여 본래의 움직임 구조 특성에 적합한 움직임 정보를 계산함으로써 더욱 평활한 움직임 필드를 구할 수 있다.

두 번째 단계의 움직임 추정에서는 각 부대역 영상 내에서도 가변 블록을 이용한 블록 정합을 수행함으로써 균일 영역에서 발생하는 움직임 벡터의 개수를 줄인다. 여기에서 가변 블록이란 이때까지 사용하던

해상도에 비례하는 가변 블록이 아니라, 부대역 영상 내의 균일 영역과 상세 영역에 따라 크기가 변하는 블록을 의미한다. 매우 상세한 영역에서는 더욱 정확한 움직임 벡터로써 낮은 예측 오차를 발생시킬 수 있도록 최소 크기의 블록을 사용하는 반면, 균일 영역에서는 큰 블록을 사용함으로써 움직임 벡터의 개수를 줄인다. 따라서 전체 웨이브렛 분할 영역에서 살펴보면 이러한 알고리즘은 이중 가변 블록으로 물체 움직임을 추정하는 것이다. 일반적으로 균일 영역은 움직이는 물체의 윤곽선이 없는 영역으로 정의될 수 있다. 이러한 영역에서는 큰 블록을 이용하여 블록 정합을 수행하더라도 예측 오차가 많이 증가하지 않는다. 반면 큰 블록을 사용하기 때문에 움직임 벡터의 개수는 많이 감소한다. 웨이브렛 분할 영상에서는 각 부대역 영상 내의 동일한 장소에 위치하는 모든 가변 블록이 동시에 윤곽선 성분이 아니면 균일 영역이 된다. 따라서 두 번째 단계의 움직임 추정은 4진트리에서 4개의 인접한 리프 노드에 속하는 모든 가변 블록을 점검하고, 균일 영역에 해당되는 인접한 4개의 리프 노드를 병합한다. 제안한 움직임 추정 알고리즘은 균일 영역을 판별하기 위하여 윤곽선 정보대신 전송 비트율을 사용하고, 균일 영역에 대한 병합 연산을 위하여 움직임 벡터 및 DFD 심벌에 대하여 확률 분포를 계산하여 비트율 예측을 한다. 즉 4개의



인접한 리프 노드에서 발생하는 예측 비트율과 부모 노드가 발생하는 예측 비트율을 비교하여, 부모 노드가 발생하는 예측 비트율이 작은 경우 4개의 리프 노드는 한 개의 부모 노드로 병합된다. 따라서 병합 연산에 적용되는 기준이 정확할수록 총 전송 비트율은 최적으로 감소된다.

두 번째 단계의 알고리즘은 병합 연산에 의한 DFD의 에너지 증가로 추가되는 전송 비트율이 움직임 벡터의 감소로 줄어든 전송 비트율을 초과하지 않으면 해당 영역을 균일 영역으로 판별하고, 실제 4개의 인접한 리프 노드를 한 개의 부모 노드로 병합한다. 총 비트율  $R$ 은 다음과 같이 정의된다. 여기에서  $R_{MV}$ 는 움직임 벡터의 전송 비트율이다.

$$R = R_{MV} + R_{DFD} \quad (9)$$

웨이브렛 분할 영상을 한 개의 4진트리만으로 구현하면, 4진트리 부호어는 움직임 보상 후 발생하는 움직임 벡터의 전송 비트율이나 예측 오차의 전송 비트율에 비하여 매우 적기 때문에 무시될 수 있다. 또한 식 (9)의 총 비트율은 다음과 같이 계산된다.

$$R = -1/N \cdot [\sum_A \log P_{MV}(A) + \sum_B \log P_{DFD}(B)] \quad (10)$$

여기에서  $P_{MV}$ 와  $P_{DFD}$ 는 움직임 벡터 및 DFD 심벌에 대한 확률을 나타내며,  $A, B, N$ 는 움직임 벡터, DFD 심벌 및 심벌의 개수를 나타낸다. 움직임 벡터와 DFD 심벌의 전송 비트율을 엔트로피에 의하여 정확하게 예측하기 위해서는 해당 심벌의 확률 분포 함수를 정확하게 계산하여야 한다.

Moscheni[21]는 DFD의 분포를 라플라시안 분포로 모델링한 멀티그리드 움직임 추정 알고리즘을 제안하였다. 그러나 Strobach[3]는 Alexis 시험 영상에서 DFD의 공간적인 상관관계를 통상 0.3에서 0.5 정도로 측정하였고, 선화(line drawing)의 특징을 발견하였다. 예측 오차는 주로 물체의 윤곽선과 움직이는 물체의 높은 대조(high contrast) 영역에서 유효 진폭(significant amplitude)을 가지며, 반면 균일하거나 움직임이 없는 영역에서는 거의 0에 접근한다. 이와 같이 DFD 신호는 완전히 비정상적 신호(nonstationary signal)에 해당되므로 DFD 심벌에 대한 모델링은 부적절하다

[3]. 또한 확률 분포 함수는 정의되는 심벌의 개수에 따라 조금씩 영향을 받는다. 실제 발생하는 심벌보다 많은 심벌을 정의하면, 각 심벌의 확률은 사용하지 않는 심벌의 개수에 영향을 받아 실제 발생 빈도보다 작은 값으로 나타난다.

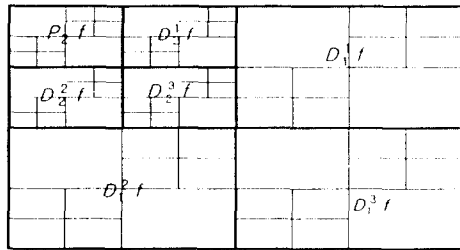
제안한 움직임 추정 알고리즘은 비정상적 통계 특성을 갖는 움직임 벡터 및 DFD 심벌에 대한 확률 분포 함수를 움직임 보상 후에 발생하는 해당 심벌에 대한 빈도수(frequency)를 계산하여 구하고, 여기서 고정 모델(fixed model)대신 적응 모델(adaptive model)을 사용한다. 따라서 병합 연산을 수행하기 위하여 4진트리 내에서 인접한 4개의 리프 노드가 발생하는 예측 비트율과 부모 노드가 발생하는 예측 비트율을 식 (11)을 이용하여 비교한다. 초기 상태에서는 가장 낮은 레벨에서만 리프 노드가 존재하고, 리프 노드는 첫 번째 단계의 움직임 보상 후에 발생하는 움직임 벡터 및 DFD 심벌로 표현된다.

$$R_{MV}(\text{parent}) + R_{DFD}(\text{parent}) < R_{MV}(4 \text{ leaves}) + R_{DFD}(4 \text{ leaves}) \quad (11)$$

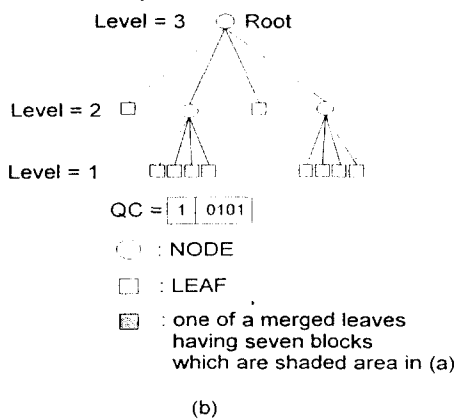
식 (11)이 성립되면, 4개의 인접한 리프 노드는 한 개의 부모 노드로 병합되어 4진트리로부터 제거된다. 만일 리프 노드의 레벨이  $l$ 이면 새로 병합된 부모 노드는  $l+1$  레벨에서 리프 노드로 변환되고, 부모 노드에 속하는 블록의 크기는 자손 노드의 4배가 된다. 따라서 움직임 벡터의 개수는 선행자보다 1/4 비율로 감소된다. 새로운 노드의 초기 바이어스로 사용되는  $v_1^*$ 는 4개의 선행자로부터 예측되고,  $v_2^*$ 는 마찬가지로  $v_1^*$ 를 초기 바이어스로 사용하여 더욱 상세하게 계산된다. 선행자의 움직임 벡터를 이용하여 새로운 병합 노드의  $v_2^*$ 를 예측함으로써 움직임 추정에 소요되는 계산량은 크게 감소된다.

또한 확률 분포 함수는 제거된 리프 노드에 해당되는 움직임 벡터와 DFD 심벌 및 새로 병합된 노드에서 발생하는 심벌을 이용하여 적응적으로 갱신된다. 이러한 절차는 가능한 최대 크기의 블록까지 혹은 더 이상의 해당 리프 노드가 없을 때까지 순환적으로 반복 수행되는데, 최종적으로 계산되는 4진트리에 대한 예가 그림 8에서 보여진다. 제안한 알고리즘은 이러한 절차에 의하여 매우 상세한 영역에서는 작은 블록

으로 더욱 정확한 움직임 추정을 함으로써 작은 예측 오차를 발생하고, 반면 균일한 영역에서는 큰 예측 오차의 증거 없이 병합 연산을 수행함으로써 움직임 벡터의 개수를 감소시킨다.



(a)



(b)

그림 8. 2단계 움직임 추정 후 결과 예 (a) 웨이브렛 분할 영역의 분할 구조 (b) 4진트리

Fig. 8 A resulting example of the two-stage motion estimation (a) The segmentation structure of the wavelet transform domain (b) The quadtree

#### IV. 화면간 부호화 시스템 모델

제안한 움직임 추정 알고리즘의 평가는 기존의 방법과의 성능 비교에 의하여 실시되고, 이러한 목적으로 웨이브렛 변환, 움직임 추정, 양자화 및 산술 부호화로 구성되는 화면간 부호화기를 구현한다.

웨이브렛 분할은 화면간 부호화기 내에서 원영상에 적용되거나 혹은 DFD에 적용될 수가 있다. Dufaux[16]는 DFD에 웨이브렛 분할을 적용하여 비디오 부호화

기를 구현한 바 있다. Zhang과 Zafar[6][7][8]는 원영상에 적용된 웨이브렛 분할이 DFD에 적용된 웨이브렛 분할보다 PSNR(peak signal to noise ratio)과 주관적 평가(subjective evaluation)에서 우수한 성능을 발휘함을 실험에서 보였다. 본 논문에서는 이러한 실험 결과를 고려하여, 그림 9처럼 화면간 부호화기를 구현한다. 여기에서  $Q$ 는 양자화기,  $Q^{-1}$ 는 역 양자화기, 그리고  $z^{-1}$ 은 프레임 저장을 의미한다.  $\hat{I}_{j,i}^k(s)$ 는  $I_{j,i}^k(s)$ 에 대한 복원 영상을 나타내고,  $v_{j,b}^k$ 는  $I_{j,i-1}^k(s-v)$ 로부터 예측되는  $I_{j,i}^k(s)$ 의 어떤 블록  $b$ 의 움직임 벡터를 의미한다. 이때 화소  $s$ , 움직임 벡터  $v_{j,b}^k$  및 시간  $t$ 에 의한 예측 오차 DFD는 복원된 이전 화면과 현재 화면간의 예측 오차를 나타내는 다음과 같이 정의될 수 있다.

$$DFD_t(s, v_{j,b}^k) = I_{j,i}^k(s) - \hat{I}_{j,i-1}^k(s - v_{j,b}^k) \quad (12)$$

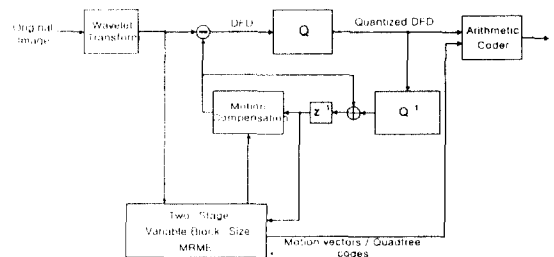


그림 9. 화면간 부호화기

Fig. 9 The interframe coder

#### 4.1 웨이브렛 분할 영역의 양자화

제안한 부호화기의 양자화 입력 심벌(symbol)은 DFD 심벌이 된다. 웨이브렛 분할 영역에 대한 부호화에서는 각각의 부대역 영상에서 다른 크기의 양자화 구간을 적용하여, 인간의 시각 민감도가 낮은 부대역 영상에서 더 큰 양자화 오차를 발생하여야 한다. Chen [17]은 웨이브렛 피라미드의 연속적인 두 레벨에서 최적의 양자화 구간은 균일 양자화기(uniform quantizer)에서 선행자(predecessor)의 1/2 크기임을 증명하였다. 균일 양자화기는 단순하게 구현할 수 있는 장점과 우수한 성능 때문에 실제 응용 분야에서 광범위하게 사용되고 있다. 웨이브렛 분할 영상에서 효율적인 양자화기는 분할 영상이 나타내는 특성을 반영함으로써

설계될 수 있다. 일반적으로 웨이브렛 분할 영상은 가장 낮은 해상도의 저주파 부대역 영상으로부터 고주파 부대역 영상으로 갈수록 계수값이 작아지는 특징을 나타낸다. 따라서 부호화 효율을 향상시키기 위해서는 대표값(representation value) 0에 해당되는 양자화 구간의 크기를 더 넓게 할당함으로써 0 근처의 입력신호를 대표값 0으로 할당할 수 있다. 즉 확률 분포 함수에서 대표값 0의 확률값이 상대적으로 높은 양자화기를 설계할 수 있다. 이러한 구간을 '영구간(deadzone)'[18]이라 한다. 본 논문의 화면간 부호화기에서 사용하는 균일 양자화기는 다음과 같이 정의된다.

$$\begin{cases} 0_j + (k-1)\Delta_j \leq I_k < 0_j + k\Delta_j & \text{if } k > 0 \\ -0_j < I_k < 0_j & \text{if } k = 0 \\ -0_j + k\Delta_j < I_k \leq -0_j + (k+1)\Delta_j & \text{if } k < 0 \end{cases} \quad (13)$$

여기에서  $\Delta_{j-1} = 2\Delta_j$ 이고,  $I_k$ 는 양자화기에서 모든 양자화 구간을 나타낸다. 또한  $\Delta_j$ 는 웨이브렛 피라미드의 해상도  $2^{-j}$ 에서 양자화 구간을 의미하며,  $0_j$ 는 해상도  $2^{-j}$ 에서 영구간에 해당되는 양자화 구간값(decision level)을 표시한다. 따라서 어떤 부대역 영상의 표본값  $x$ 에 대한 양자화기는 다음과 같이 설명될 수가 있다.

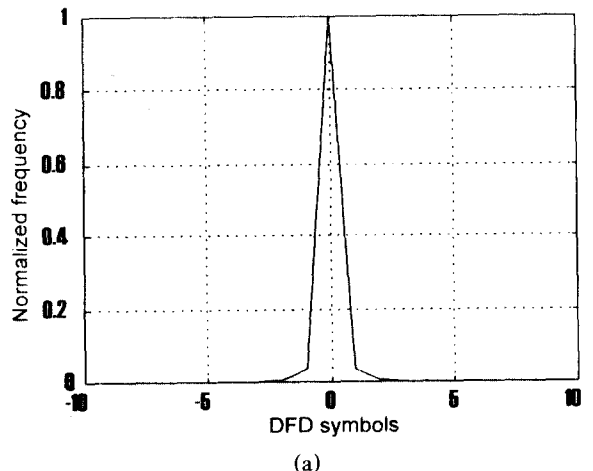
$$Q(x) = k, \quad \text{for } x \in I_k \quad (14)$$

식 (13) 및 식 (14)를 살펴보면, 이러한 균일 양자화기는 웨이브렛 분할 영역의 연속적인 두 레벨에서 상위 레벨의 양자화 구간이 하위 레벨보다 1/2 크기이고, 영구간의 크기를 조절할 수 있는 것이다.

### V. 모의 실험 및 고찰

제안한 알고리즘은 기존의 알고리즘과의 성능 비교를 통하여 평가된다. 성능 비교는 움직임 보상 후 발생하는 움직임 벡터 및 DFD 심벌의 비트율(bit-rate)을 이용하여 실시하고, 실제 비트율은 산술 부호화기(arithmetic coder)를 구현하여 계산한다. 시험 영상은 영상 부호화 분야의 표준 영상 중에서 Claire와 Football을 선택한다. Claire 시험 영상은 정상적인 상태에서 사람의 머리와 어깨가 보이는 비교적 느린 움직임

을 나타내고, Football 영상은 비교적 작은 물체, 빠른 움직임 및 카메라 팬 등을 포함한다. 시험 영상은 각각  $352 \times 288$ 과  $352 \times 240$  크기이고, 40장의 연속적인 화면이다. 두 종류의 시험 영상은 Antonini[22]의 9-7 탭을 갖는 스플라인(spline) 필터와 Barnard 알고리즘[23]을 이용하여 웨이브렛 분할 영상으로 변환된다. 해상도 깊이 2를 적용하여 원영상을 웨이브렛 분할을 하면, 웨이브렛 피라미드의 바닥 계층에 3개의 부대역 영상과 꼭대기 계층에 4개 부대역 영상을 합하여 전체 7개의 부대역 영상이 발생된다. 이러한 부대역 영상은 한개의 4진트리 루트 노드로 표현된다. 루트 노드에 속하는 각 부대역 영상은 최소 가변 블록으로 분할되고, 각각의 블록에 대하여 첫 번째 단계의 움직임 추정을 한다. 두 번째 단계의 움직임 추정부에서는 첫 번째 단계의 움직임 보상 후 생성된 움직임 벡터와 DFD의 확률 분포를 계산하여 산술 부호화기에서 발생될 최종 비트량을 예측한다. 비트량의 예측은 엔트로피 기준을 사용한다. 그림 10(a)와 그림 10(b)는 두 종류의 40장의 시험 영상에 대하여 제안한 움직임 추정 알고리즘을 수행하여 최종적으로 구한 DFD 심벌의 확률 분포 함수이다. 여기에서 0 심벌의 빈도수는 나머지 심벌의 그것보다 상대적으로 너무 크기 때문에 정규화하면 좌우 대칭 형태처럼 나타난다. 실제 움직임 추정 이후의 DFD는 매우 작게 나타나기 때문에 대부분 0 심벌에 해당되지만, 물체 윤곽선과 같은 복잡한 물체에 해당되는 영역에서는 약간의 예측 오차 성분이 존재하여 0 심벌이외의 다른 심벌이 발생된다.



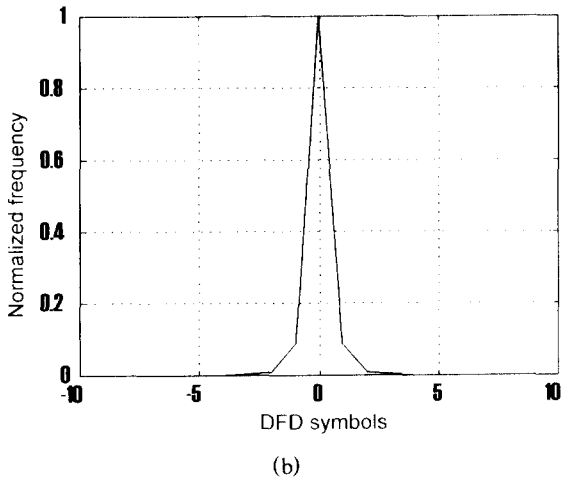
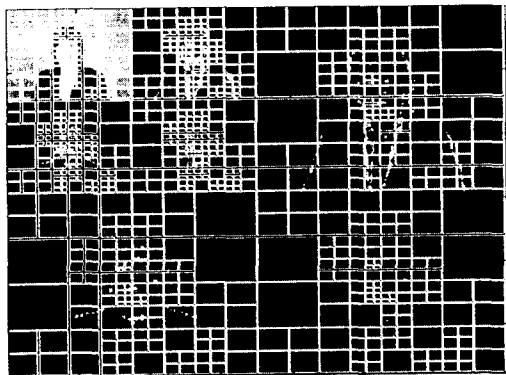


그림 10. DFD 심벌의 확률 분포 함수 (a) Claire (b) Football

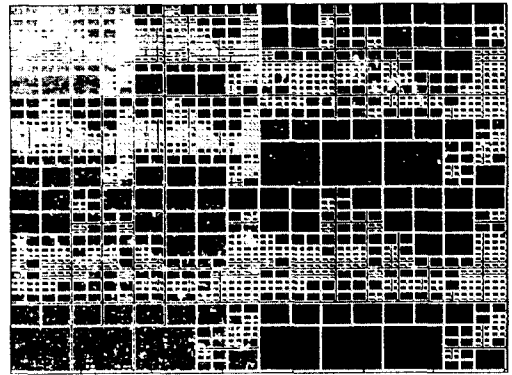
Fig. 10 Probability distribution function of the DFD symbols (a) Claire (b) Football

또한 함수의 형태는 실험 결과에서 쉽게 볼 수 있듯이 매우 밀집된 모양을 나타내는데, 이것은 제한한 알고리즘이 물체의 움직임을 정확하게 추정하는 것을 내포한다. 움직임이 복잡한 Football 시험 영상의 확률 분포 함수는 그림 10(b)처럼 Claire보다 0 심벌 근처에서 약간 높은 확률값을 갖는데, 이러한 결과는 Football 시험 영상의 예측 오차가 Claire보다 높은 것을 의미한다. 이러한 함수는 Moscheni[21]가 제안한 라플라시안 함수와는 다른 형태를 띠고 있다.

움직임 벡터 및 DFD 심벌의 분포 함수를 실제 계



(a)



(b)

그림 11. 2단계 움직임 추정 결과 (a) Claire (b) Football

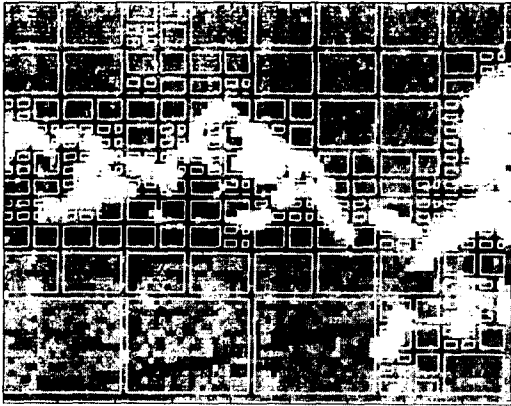
Fig. 11 The results of the two-stage motion estimation (a) Claire (b) Football

산하여 병합 연산에 사용한 결과, 그림 11(a) 및 그림 11(b)와 같은 움직임 추정 결과를 얻을 수 있었다.  $P_2f$  만을 확대하면 그림 12(a)와 그림 12(b)처럼 된다. 이와 같이 움직임 벡터 및 DFD 심벌의 정확한 분포 함수만이 정밀한 병합 연산을 수행하여 전체적으로 적은 비트율을 발생시킨다. 여기서 보면, Football 시험 영상의 결과에서는 움직임이 많고 작은 물체들이 많은 관계로 작은 크기의 블록들이 많이 발생하고, Claire 시험 영상에서는 얼굴의 일부를 제외한 나머지 영역에서 병합 연산에 의한 큰 블록이 많이 발생된다.

본 논문에서는 다음과 같은 4개의 움직임 추정 방법간에 동일한 PSNR 값에서 전송 비트율을 비교함으로써 제안한 알고리즘의 성능을 평가한다. 여기에



(a)



(b)

그림 12. 저주파수 부대역 영상만 확대한 결과 (a) Claire (b) Football  
Fig. 12 Zooming results for the low-pass subband (a) Claire (b) Football

서 제안한 알고리즘은 2단계 움직임 추정 알고리즘 중 첫 번째 단계만을 수행한 알고리즘과 첫 번째, 두 번째 단계를 모두 수행한 알고리즘이고, 기존의 알고리즘은 완전 탐색 블록 정합 알고리즘과 가변 블록 다해상도 움직임 추정 알고리즘이다.

• 방법 1(scheme 1): 가변 블록에 의한 완전 탐색 블록 정합 알고리즘

이러한 움직임 추정 기법은 웨이브렛 피라미드의 다른 해상도간에 존재하는 물체 움직임의 상관관계를 이용하지 않고, 현재 부대역 영상 내의 어떤 블록의 움직임을 찾기 위하여 이전 부대역 영상 내의 탐색 영역에 속하는 일련의 이산 후보 변위에 대하여 완전 탐색을 실시한다. 각 부대역 영상 내의 블록을  $b_k^s$ 라고 가정하면,  $b_k^s$ 의 움직임 벡터  $v_k^s$ 는 다음과 같이 계산된다.

$$v_k^s = \underset{v \in \Omega}{\operatorname{argmin}} \sum_{s \in W} \|I_{k,i}^s(s) - I_{k,i-1}^s(s-v)\| \quad (15)$$

여기에서  $j=1, 2$ 와  $k=0, 1, 2, 3$ 이고,  $k=0$ 는 저주파수 성분을 나타낸다. 완전 탐색 블록 정합은 식 (15)처럼  $P_2f$ 를 포함한 모든  $D_j^k f$ 에서 적용된다. 탐색 영역은 수평 및 수직 방향으로 동일한 크기를 갖고, 해

상도에 따라 비례하는  $\Omega_0 \cdot 2^{2-j}$ 로 주어진다. 여기에서  $\Omega_0$ 는  $P_2f$ 에서의 탐색 범위이고, 2은  $2^{-2}$ 해상도를 나타낸다. 또한  $j$ 는  $j$ 번째 계층을 표시한다. 이러한 방법은 탐색 영역의 증가로 계산량은 매우 증가하지만, 정합 기준에 대하여 최소의 예측 오차를 발생하기 때문에 비교 대상의 알고리즘으로 널리 사용된다. 블록 정합에서 사용되는 블록의 크기는 Zafar가 제안한 알고리즘과 동일하게 가변한다.

• 방법 2(scheme 2): 가변 블록 다해상도 움직임 추정 알고리즘

Zafar가 제안한 가변 블록 다해상도 움직임 추정 알고리즘이고, 그 중 예측 탐색 알고리즘은 최고의 성능을 나타내는 'A-III' 알고리즘[8]이다. 이러한 알고리즘에 대한 설명은 II-2 절에서 상세히 설명된다.

• 방법 3(scheme 3): 1단계 가변 블록 다해상도 움직임 추정 알고리즘

본 논문에서 제안한 2단계 움직임 추정에서 첫 번째 단계만 수행한다.

• 방법 4(scheme 4): 2단계 가변 블록 다해상도 움직임 추정 알고리즘

본 논문에서 제안한 두 번째 단계까지 움직임 추정을 수행한다.

여러 알고리즘간의 성능은 동일한 PSNR 조건에서 전송 비트율의 차이로 비교되고, 항목으로는 산술 부호화기에서 출력되는 움직임 벡터, DFD 및 총 전송 비트율을 사용한다. 여기에서 총 전송 비트율은 움직임 벡터와 DFD 비트율을 합친 것을 의미한다. 실제 연속적인 영상에 대하여 다른 알고리즘을 사용하는 다른 부호화기에서 동일한 PSNR 값을 만들기는 거의 불가능하다. 하지만, 본 논문은 웨이브렛 변환영역에서의 효율적인 움직임 추정 알고리즘의 개발에 목적을 두고, 개발된 알고리즘의 성능 평가를 위해 화면내 부호화(intraframe coding) 및 고정 비트율(constant bit rate) 부호화 등이 생략된 필수적인 기능만을 가진 영상 부호화기를 설계하였다. 이러한 성능 비교

를 위해 설계된 부호화기는 방법 1, 방법 2, 방법 3 및 방법 4에 대해 공통적으로 적용된다. 뿐만 아니라, 40개의 연속적인 시험 영상에 대한 복원 영상의 PSNR 값을 여러 알고리즘에서 동일하게 발생하기 위해서는 동일한 시험 영상과 동일한 부호화 매개변수를 사용하여야 한다. 본 논문에서는 복원 영상의 품질을 고려하여 부호화기에서 사용되는 매개변수를 다음과 같이 설정한다.

- 웨이브렛 피라미드의 해상도 레벨: 2
- $P_2^f$ 에서의 탐색 범위,  $\Omega_0: 8$
- $D_2^f$ 에서의 탐색 범위  $\Omega$ 
  - ▶ 방법 1:  $\Omega_0 \cdot 2^{2-j}$
  - ▶ 방법 2, 방법 3 및 방법 4:  $2^{2-j+1}$
- 양자화 구간: 5.0
- 영구간: 2.5

그림 13은 PSNR 곡선을 보여 준다. 여기서 Claire 영상은 평균 40.71 dB를 보여주고, Football 영상은 33.87 dB를 나타낸다.

동일한 PSNR 곡선을 갖는 조건에서 움직임 벡터, DFD 성분 및 총 전송 비트율을 bpp 곡선으로 나타낸다. 방법 1, 방법 2, 방법 3 및 방법 4에 해당되는 각각의 곡선은 성능 비교를 위하여 동시에 보여진다. 각각의 그림에서 'o' 표시는 방법 1을 나타내고, 'x'는 방법 2, '+'는 방법 3 및 '\*'는 방법 4에 대한 표시이다.

그림 14(a)와 그림 14(b)는 움직임 벡터의 bpp 곡선을 보여준다. 방법 1, 방법 2 및 방법 3은 부대역 영상 내의 균일 영역과 상세 영역간에 고정 블록에 의한

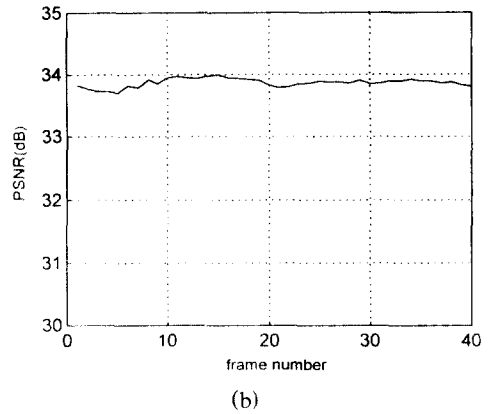
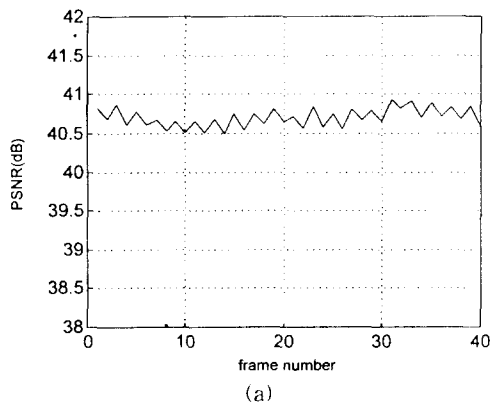


그림 13. PSNR 곡선 (a) Claire (b) Football  
Fig. 13 PSNR curves (a) Claire (b) Football

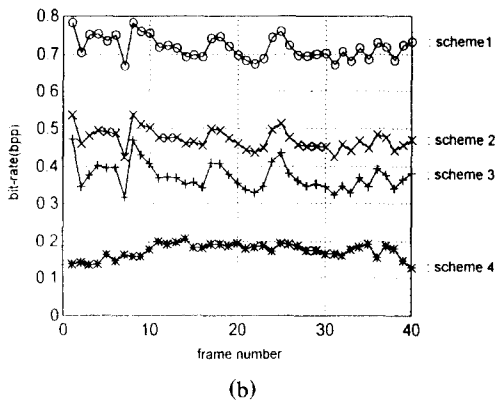
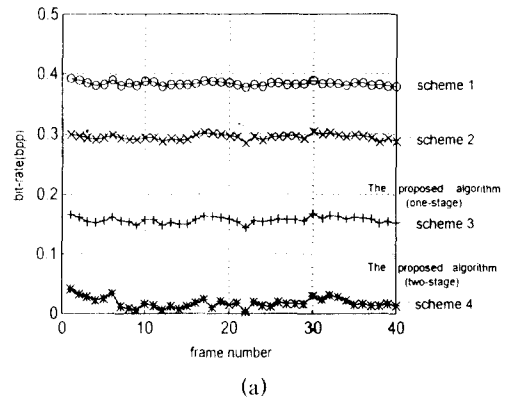


그림 14. 움직임 벡터의 bpp 곡선 (a) Claire (b) Football  
Fig. 14 The bit-rate curves of motion vectors (a) Claire (b) Football

블록 정합을 수행하기 때문에 움직임 벡터의 개수를 줄일 수가 없다. 또한 방법 1과 방법 2는  $D^*f$  영역에서 동일한 예측 오차를 갖는 여러 후보 움직임 벡터에 대하여 평활한 움직임 벡터를 결정하지 못한다. 방법 3은  $D^*f$  내의 낮은 에너지의 블록에 대하여 움직임 추정 대신 추가 심벌을 정의하여 전송하고, 같은 예측 오차를 갖는 여러 개의 후보 움직임 벡터 중에서 평활한 움직임 벡터를 선택한다. 방법 4는 방법 3에서 계산된 평활한 움직임 필드를 바탕으로 균일 영역에 해당하는 움직임 벡터의 개수를 더욱 줄임으로써 가장 낮은 비트율을 발생한다. 제안한 방법 4는 기존의 방법 2보다 Claire 시험 영상에서 평균 0.288 bpp, Football 시험 영상에서 평균 0.30 bpp 더 낮은 비트율을 발생한다. 그리고 제안한 방법 3은 방법 2보다

Claire 시험 영상에서 평균 0.14 bpp, Football 시험 영상에서 평균 0.1 bpp 더 낮은 비트율을 발생한다.

그림 15(a)와 그림 15(b)는 DFD의 bpp 곡선을 보여 준다. 일반적으로 완전 탐색 알고리즘은 이전 화면의 탐색 영역 내에서 완전 탐색을 통하여 최소 예측 오차를 갖는 움직임 벡터를 선택하기 때문에 가장 작은 예측 오차와 낮은 비트율의 DFD를 발생한다. 하지만 그림 15(a)에서는 제안한 방법 3이 기존의 방법 1보다 좋은 성능을 제공한다. 이러한 이유는 Claire 시험 영상을 웨이브렛 분할을 하면  $D^*f$  내의 대부분의 블록은 임계치보다 적은 에너지를 갖기 때문이다. 이러한 블록은 움직임 추정 대신 추가 심벌로 정의되고, 에너지가 0인 예측 오차를 발생한다. Football 시험 영상에서는 작은 물체가 많이 존재하고 움직임이 많은

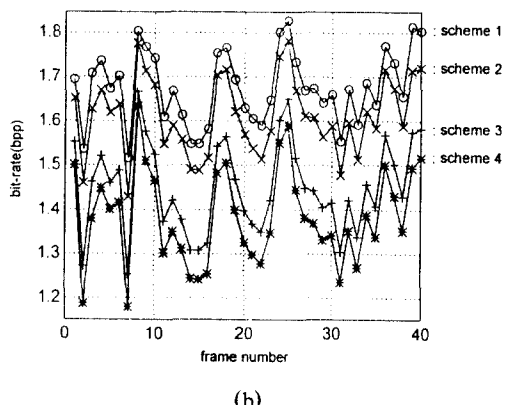
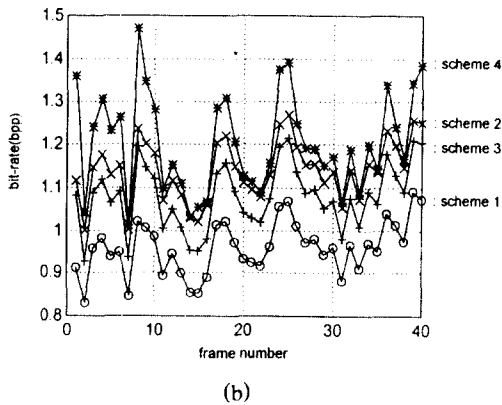
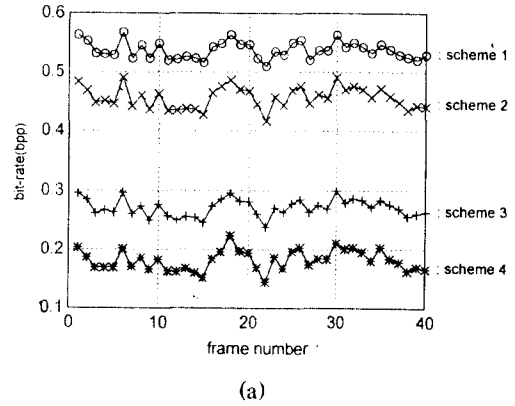
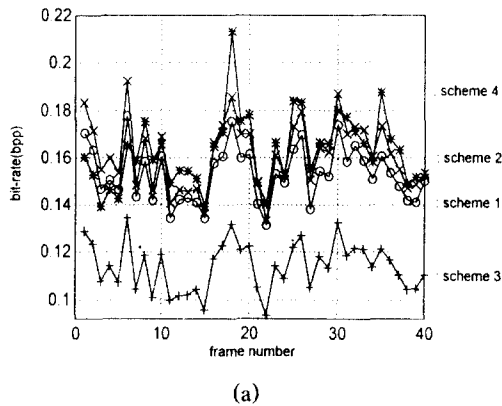


그림 15. DFD의 bpp 곡선 (a) Claire (b) Football  
Fig. 15 The bit-rate curves of resulting the DFD (a) Claire (b) Football

그림 16. 총 전송 비트율의 bpp 곡선 (a) Claire (b) Football  
Fig. 16 The total bit rate curves (a) Claire (b) Football

관계로  $D^*f$ 의 에너지 분포가 높아진다. 여기에서는 일반적인 경우처럼 방법 1이 가장 좋은 성능을 제공한다. 방법 3은 정확한 움직임 추정으로 기존의 알고리즘보다 낮은 비트율의 예측 오차를 발생한다. 방법 4는 움직임 벡터의 개수를 줄이기 위하여 균일 영역에서 큰 블록으로 움직임 보상을 수행하지만, 기존의 방법 2와 비슷한 비트율의 예측 오차를 발생한다. 제안한 방법 3은 기존의 방법 2보다 Claire 시험 영상에서 평균 0.05 bpp, Football 시험 영상에서 평균 0.06 bpp 더 낮은 비트율을 발생한다. 그리고 제안한 방법 4는 방법 2보다 Claire 시험 영상에서 평균 0.001 bpp, Football 시험 영상에서 평균 0.07 bpp 더 높은 비트율을 발생한다.

그림 16(a)와 그림 16(b)는 총 정보 비트율에 대한 bpp 곡선이다. 앞에서 언급한 바와 같이 총 전송 비트율은 움직임 벡터와 DFD의 비트율을 합한 것의 의미이다. 제안한 방법 4는 DFD 정보와 움직임 벡터에 대한 정확한 확률 분포에 의하여 최종 비트율을 예측한다. 이러한 비트율 예측에 의한 병합 연산으로 균일 영역에서는 거친 움직임을 추정하고, 반면 상세 영역에서는 정확한 움직임을 추정한다. 이러한 방법은 DFD와 움직임 벡터간에 최적의 비트 할당을 함으로써 총 전송 비트율은 최소가 된다. 실험 결과에서 볼 수 있듯이, 제안한 방법 3과 4는 기존의 방법보다 낮은 비트율을 제공한다. 방법 1은 낮은 비트율의 DFD를 발생하지만 높은 비트율의 움직임 정보로 인하여 총 전송 비트율은 증가한다. 방법 2는 비교적 높은 비트율의 DFD와 움직임 정보로 제안한 알고리즘보다 높은 전송 비트율을 발생한다. 제안한 방법 3은 기존의 방법 2보다 Claire 시험 영상에서 평균 0.19 bpp, Football 시험 영상에서 평균 0.16 bpp 더 낮은 비트율을 발생한다. 그리고 제안한 방법 4는 기존의 방법 2보다 Claire 시험 영상에서 평균 0.28 bpp, Football 시험 영상에서 평균 0.23 bpp 더 낮은 비트율을 발생한다.

## VI. 결 론

본 논문에서는 기존의 가변 블록 다해상도 움직임 추정 알고리즘의 성능을 향상시키고, 균일 영역에서 움직임 벡터의 개수를 줄이기 위하여 웨이브렛 영상

의 각 부대역 영상 내에서도 가변 블록에 의한 블록 정합을 실시한다. 제안한 2단계 가변 블록 다해상도 움직임 추정 알고리즘은 움직임 벡터 및 DFD로 구성되는 총 전송 비트율을 줄임으로써 양 성분간에 효율적인 비트 할당을 시도한다. 제안한 알고리즘은 2단계로 움직임을 추정하며, 첫 번째 단계에서는 화면간 평활한 움직임 필드와 정확한 움직임 예측을 제공하기 위하여 웨이브렛 변환 특성에 적합한 움직임 추정 방법을 개발한다.

두 번째 단계에서는 각 부대역 영상 내의 균일 영역에 할당되는 움직임 벡터의 개수를 줄인다. 이를 위하여 각 부대역 영상 내의 블록은 4진트리의 리프 노드로 할당되고, 4개의 인접한 리프 노드의 병합 연산에 의하여 균일 영역에 해당되는 움직임 벡터의 개수는 줄어든다. 병합 연산을 위하여 심벌에 대한 적응적 확률 분포를 계산하여 전송 비트율을 예측한다. 병합 연산이 정확할수록 총 전송 비트율은 최적으로 감소된다.

2단계 움직임 추정 알고리즘은 웨이브렛 변환 기반의 화면간 영상 부호화에서 기존의 알고리즘보다 낮은 전송 비트율을 제공한다. 하지만 2단계 알고리즘을 구현하기 위해서는 움직임 벡터와 DFD에 대한 확률 분포를 계산하여야 한다. 실제 두 성분에 대한 확률 계산을 할 수 없는 응용 분야에서도 첫 번째 단계의 움직임 추정 알고리즘이 기존의 알고리즘보다 낮은 전송 비트율을 제공한다.

본 논문에 대한 향후 연구는 엔트로피 기준 대신 물체 윤곽선 정보에 의한 병합 연산을 생각할 수가 있다. 즉 물체 윤곽선은 웨이브렛 분할 영상 자체가 제공하는 멀티스케일 정보에 의하여 검출될 수가 있고, 이러한 정보는 병합 연산의 균질성 기준으로 사용될 수도 있다. 다만 이 방법은 특수한 웨이브렛 필터에 대해서만 적용가능하므로 일반적인 웨이브렛 필터에서도 윤곽선을 쉽게 검출할 수 있는 간단한 방법이 개발되어야 한다.

## 참 고 문 헌

1. W. H. Chen and W. Pratt, "Scene adaptive coder," *IEEE Trans. Commun.*, vol. COM-32, pp. 225-232, Mar. 1984.



2. A. K. Jain, "Image data compression: A review," Proc., IEEE, vol. 69, pp. 349-389, Mar. 1981.
3. P. Strobach, "Tree-structured scene adaptive coder," IEEE Trans. Commun., vol. 38, pp. 477-486, Apr. 1990.
4. S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," IEEE Trans. Pattern Anal. Machine Intell., vol. 11, pp. 674-693, July 1989.
5. S. G. Mallat, "Multifrequency channel decompositions of images and wavelet models," IEEE Trans. Acoust., Speech, Signal Processing, vol. 37, pp. 2091-2110, Dec. 1989.
6. Y. Q. Zhang and S. Zafar, "Motion-compensated wavelet transform coding for color video compression," in SPIE Proc. Visual Commun. and Image Process., Boston, MA, Nov. 1991, vol. 1605, pp. 301-316.
7. Y. Q. Zhang and S. Zafar, "Motion-compensated wavelet transform coding for color video compression," IEEE Trans. Circuits Syst. for Video Tech., vol. 2, pp. 285-296, Sept. 1992.
8. S. Zafar and Y. Q. Zhang, "Multiscale video representation using multiresolution motion compensation and wavelet decomposition," IEEE J. Select. Areas Commun., vol. 11, pp. 24-35, Jan. 1993.
9. K. Uz, M. Vetteri and D. LeGall, "Interpolative multiresolution coding of advanced television and compatible subchannels," IEEE Trans. Circuits Syst. for Video Tech., vol. 1, pp. 86-99, Mar. 1991.
10. H. G. Musmann, P. Pirsh and H. J. Grallert, "Advances in picture coding," Proc. IEEE, vol 73, No. 4, pp. 523-548, Apr. 1985.
11. F. Dufaux and F. Moscheni, "Motion estimation techniques for digital TV: A review and a new contribution," Proc. IEEE, vol. 83, pp. 858-876, June 1995.
12. M. Bierling, "Displacement estimation by hierarchical block matching," in SPIE Visual Commun. and Image Process., Cambridge, MA, Nov. 1988, vol. 1001, pp. 942-951.
13. M. H. Chan, Y. B. Yu, and A. G. Constantinides, "Variable size block matching motion compensation with applications to video coding," IEE Proc., vol. 137, no. 4, pp. 205-212, Aug. 1990.
14. M. T. Orchard, "Predictive motion-field segmentation for image sequence coding," IEEE Trans. Circuits Syst. for Video Tech., vol. 3, pp. 54-70, Feb. 1993.
15. F. Dufaux and M. Kunt, "Multigrid block matching motion estimation with an adaptive local mesh refinement," in SPIE Proc. Visual Commun. and Image Process., Boston, MA, Nov. 1992, vol. 1818, pp. 97-109.
16. F. Dufaux, L. Moccagatta, B. Rouchouze, T. Ebrahimi and M. Kunt, "Motion-compensated generic coding of video based on a multiresolution data structure," Optical Engineering, vol. 32, pp. 1559-1570, July 1993.
17. J. Chen, S. Itoh and T. Hashimoto, "Scalar quantization noise analysis and optimal bit allocation for wavelet pyramid image coding," IEICE Trans. Fundamentals, vol. E76-A, pp. 1502-1514, Sept. 1993.
18. D. Taubman and A. Zakhor, "Multirate 3-D subband coding of video," IEEE Trans. Image Processing, vol. 3, pp. 572-588, Sept. 1994.
19. A. Klinger and C. R. Dyer, "Experiments on picture representation using regular decomposition," Comput. Graphics Image Processing, vol. 5, pp. 68-105, 1976.
20. H. Samet, "The quadtree and related hierarchical data structures," ACM Comput. Surveys, vol. 16, pp. 188-216, June 1984.
21. F. Moscheni, F. Dufaux, and H. Nicolas, "Entropy criterion for optimal bit allocation between motion and prediction error information," in SPIE Proc. Visual Commun. and Image Process., Cambridge, MA, Nov. 1993, vol. 2094, pp 235-242.
22. M. Antonini, M. Barlaud, P. Mathieu and I. Daubechies, "Image coding using wavelet transform," IEEE Trans. Image Processing, vol. 1, pp.

205-220, Apr. 1992.

23. H. J. Barnard, J. H. Weber and J. Biemond, "Efficient signal extension for subband/wavelet decomposition of arbitrary length signals," in SPIE Proc. Visual Commun., and Image Process., Cambridge, MA, Nov. 1993, vol. 2094, pp. 966-975.



김 성 만(SeongMan Kim) 정회원

1961년 10월 28일생

1984년 2월: 연세대학교 전자공학과(공학사)

1986년 8월: 연세대학교 대학원(공학석사)

1987년 7월~1990년 3월: 디지털 정보통신연구소 근무

1997년 2월: 연세대학교 대학원(공학박사)

1990년 5월~현재: 한국통신 전송기술연구소 근무

※주관심분야: 영상부호화, 영상통신



이 규 원(KyuWon Lee) 정회원

1963년 9월 25일생

1986년 2월: 연세대학교 전자공학과(공학사)

1988년 2월: 연세대학교 대학원 전자공학과(공학석사)

1988년 2월~1989년 8월: 금성산전(주) 연구소 연구원

1989년 9월~현재: 한국전자통신연구소 연구원(1993년 10월 휴직)

1993년 9월~현재: 연세대학교 대학원 전자공학과 박사과정

※주관심분야: 영상처리, Active Vision, Motion Analysis, 영상통신



정 학 진(HakJin Chong) 정회원

1959년 1월 5일생

1982년 2월: 경북대학교 전자공학과(공학사)

1984년 2월: 경북대학교 공과대학원 전자공학과(공학석사)

1984년 3월: 한국전기통신공사 연구원

1985년 5월~현재: 한국통신 전송기술연구소 팀장

※주관심분야: 영상통신, ATM Network

박 규 태(KyuTae Park)

정회원

현재: 연세대학교 전자공학과 교수

한국통신학회 논문지 제20권 4호 참조