

## 웨이블릿 영역에서 선택적 보간의 반화소를 이용한 효과적인 움직임 추정

이 대 호, 김 광 용, 정 태 연, 김 덕 규  
경북대학교 전자공학과  
전화 : 053-940-8625 / 핸드폰 : 018-530-0918

### Efficient Motion Estimation Using Half-pel Accuracy Motion Vector by Selective Interpolation in the Wavelet Domain

Tae-ho Lee, Kwang-Yong Kim, Tae-Yeon Jung, Duk-Gyoo Kim  
School of Electrioc and Electrical Engineering, Kyungpook Natinal University  
E-mail : zeronine@palgong.knu.ac.kr

#### 요약

본 논문은 웨이블릿(wavelet) 변환된 각 프레임의 모든 부대역의 블록들에 대해 계층적 움직임을 추정할 때 고해상도 계층에서는 기저대역에서 추정된 전역 움직임 벡터를 기초로 하여 국부 움직임을 추정한다. 이때 복원 영상에 미치는 영향이 가장 큰 기저대역에 대하여 반화소를 사용하면 더욱 최적의 움직임 벡터를 추정할 수 있으나 계산량이 증가하는 단점이 있다. 블록내에 인접한 화소들 간에는 상관관계가 높다는 사실을 이용하여 오차가 최소가 되는 방향을 예측하여 선별적인 보간을 행하여 반화소 움직임을 탐색하여 계산량을 줄였다. 그리고 더욱 향상된 화질을 얻기 위해서 에지 성분이 많은 고해상도 계층에서 저해상도 계층으로의 선택적 국부 움직임을 추정하였다. 모의 실험 결과 기존의 웨이블릿 변환을 이용한 움직임 추정 및 보상 방법보다 향상된 화질을 나타내었다.

#### I. 서론

고선명 TV, 주문형 비디오, 영상 회의 등의 시스템에서는 동영상에 제한된 용량의 채널을 통하여 전송하거나 저장 매체에 저장하기 위하여 동영상 압축(video compression)을 행한다. 동영상 압축 시스템에서는 동영상에 존재하는 중복성을 제거하여 주어진 비트율에서 복원 영상의 화질을 최적으로 하는데 높은 공간적

및 시간적 중복성을 제거함으로써 효과적인 압축을 행한다. 동영상 압축 표준인 H.263 및 MPEG-2에서는 더욱 정확한 움직임 추정 및 보상을 위하여 웨이블릿(wavelet) 변환과 분수 화소(fractional-pel) 단위의 움직임 추정 및 보상을 행한다.

본 논문은 기저대역에서 더욱 정확한 추정을 위해 반화소단위의 움직임 추정을 사용하면 더 향상된 화질을 얻을 수 있음을 알수 있었다. 그러나 이 방법은 계산량이 증가하므로 블록 내에 인접한 화소들 간의 상관관계가 높다는 사실을 이용하여 오차가 최소가 되는 방향을 예측하여 선별적인 보간을 행하여 정확한 움직임 벡터를 찾았으며 계산량을 상당히 줄였다. 또한 기존의 MRME에서는 저해상도 대역에서 고해상도 대역으로의 움직임 추정이므로 저해상도 대역의 블록의 크기가 적기 때문에 에일리싱(aliasing) 문제나 에러가 누적되는 단점이 있다. 따라서 이를 해결하기 위하여 에지 성분이 많은 고해상도 대역에서 저해상도 대역으로 선별적인 움직임을 추정하여 더 나은 화질을 얻을 수 있었다.

#### II. 웨이블릿 변환 영역에서의 움직임 추정 방법

움직임 보상 부호화에서는 이전 프레임을 이용하여 움직임 추정 및 보상을 행한 후, 움직임 보상된 영상과 원 영상과의 차영상과 추정된 움직임 벡터를 부호

화한다. 여기서 사용하는 움직임 추정 방법은 알고리즘이 간단하여 하드웨어 구현이 용이한 블록 정합 알고리즘(BMA)을 사용한다. 그러나 이 방법은 움직임 추정을 행하는 경우 대상 영상의 움직임에 대하여 적절한 탐색 영역과 블록의 크기를 결정하기 어렵고, 일정한 크기의 블록으로 나누어 움직임을 추정하기 때문에 구획화 현상(blocking effect)이 발생하는 단점이 있다. 이러한 단점을 제거하기 위하여 입력 영상을 웨이블릿 변환을 이용하여 다해상도 영상으로 분할한 후, 분할된 영상에 대하여 계층적으로 움직임 추정 및 보상을 행하는 방법들이 연구되고 있다. 웨이블릿 변환을 이용한 움직임 추정 및 보상 방법에는 Zhang<sup>[1]</sup> 등이 제안한 MRME(multi-resolution motion estimation) 방법이 대표적인 것이다. 그림 1은 MRME 방법을 나타낸 것으로 웨이블릿 변환된 각 프레임의 모든 부대역의 블록들에 대하여 블록 정합 알고리즘(BMA)을 적용하여 계층적으로 움직임 추정 및 보상을 행한다. 기저대역에서는 전역 움직임을 추정한 후, 고해상도 계층에서는 기저대역에서 추정한 움직임 벡터를 기준으로 작은 탐색 영역에서 미소 움직임 변위를 추정한다. 이러한 과정을 정제(refine)라 하며 이 정제에 의하여 블록 정합에 필요한 계산량을 감소시킬 수 있다. 이러한 방법은 BMA에 비하여 탐색과 정합시간이 적게 소모되며, 영상정보의 점진적인 전송이 가능하고 부드러운 움직임 벡터장을 만들 수 있으며 웨이블릿 역변환을 통하여 구획화 현상을 제거할 수 있다. 그러나 기저 대역에서의 움직임 벡터는 고해상도 계층에서도 사용되어지므로 부정확하게 추정되어 질 경우 움직임 추정 오차가 다음 고해상도 계층으로 과급되는 단점이 있다. 따라서 더욱 정확한 기저대역 움직임 추정이 필요함을 알 수 있다.

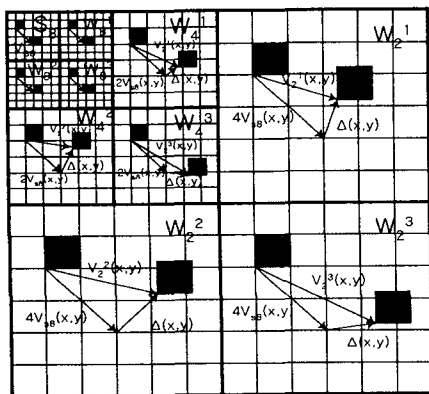


그림 1. 가변 블록 크기의 다해상도 움직임 추정

### III. 제안한 방법의 움직임 추정 방법

제안한 방법에서는 MRME 기법의 문제점을 개선하기 위하여 기저대역에서의 선별적인 보간을 통한 반화소 단위의 움직임 추정과 고해상도 계층에서 저해상도 계층으로의 선택적 국부 움직임 추정을 행하였다.

#### 3.1 기저대역에서의 선별적인 반화소 단위의 움직임 보간

웨이블릿 변환된 영상의 기저대역은 원영상의 대부분의 에너지를 차지하고 있으므로, 웨이블릿 역변환시의 복원 영상에 미치는 영향이 크다. 또한 기저 대역에서의 움직임 벡터는 고해상도 계층에서도 계속 사용되어지므로 부정확하게 추정되면 움직임 추정 오차가 다음 고해상도 계층으로 과급된다. 그러므로 기저대역에서의 정확한 움직임 추정은 매우 중요하다.

본 논문에서는 기저 대역에서의 정확한 움직임 추정을 위하여 반화소 단위(half-pel)의 움직임 추정을 사용하였다. Girod<sup>[2]</sup>는 TV 신호 및 비디오 폰 신호에 대하여 반화소로 향상된 성능을 얻을 수 있음을 확인하였고, MPEG-2 TM에서는 정확한 움직임 추정을 위하여 반화소 단위 움직임 추정을 사용하고 있다.

반화소 움직임 추정은 실제로 물체의 움직임이 반드시 카메라의 샘플링 격자 간격의 정수 배로만 일어나지 않는다는 사실에 근거한다. 즉 현재 프레임과 이전 프레임과의 관계를  $I_t(x, y) = I_{t-1}(x - d_x, y - d_y)$  와 같이 나타낼 때, 수평 및 수직 방향으로의 실제 움직임 거리인  $d_x$  및  $d_y$  는

$$d_x = (m_x + v_x)d, \quad d_y = (m_y + v_y)d \quad (1)$$

와 같이 반드시 샘플링 간격  $d$ 의 정수배로만 나타나지 않는다. 여기서,  $m_x$ 와  $m_y$  는 움직임 거리의 정수 성분이고,  $v_x$  및  $v_y$ 는 소수 부분을 나타낸 것으로 -0.5에서 +0.5사이의 값을 갖는다. 반화소 단위의 움직임 추정 기법에서는 전역 탐색 알고리즘으로 최적의 정화소 단위의 움직임 벡터를 찾은 후, 정화소 단위 움직임 벡터를 기준으로 하여 그림 2의 (a)에서처럼 이전 프레임의 영상을 보간(interpolation)하여 이웃 8개의 반화소 위치를 만든 뒤, 기준점과 이웃 8개의 반화소 위치에 대하여 정합 척도가 최적인 위치를 찾는다. 반화소 단위의 추정은 정확한 움직임 추정을 통하여 움직임 추정 오차의 분산을 줄일 수 있으나 전역 탐색 방법에 필요한 계산량 이외의 별도의 계산량이 필요해 계산량이 증가한다는 단점이 있다. 제안한 방법에서는 계산량을 줄이기 위해 블록 내에 인접한 화

소들 간에는 상관 관계가 매우 높다는 사실을 이용하여 오차가 최소가 되는 방향을 미리 예측하여 선별적인 보간을 통하여 반화소 단위 움직임을 탐색한다.

그림 2의 (b)에서처럼 수평이나 수직 방향을 기준 방향으로 잡으면 블록2, 블록7, 블록4, 그리고 블록5가 기준 움직임 블록이 된다. 먼저 이러한 기준 블록을 반화소 단위 움직임 보간하고 나서 수직방향의 블록 2와 블록 7을 비교하고, 수평방향의 블록 4와 블록 5를 비교한다. 각 방향당 최소의 오차(MSE)를 가진 두 개의 블록을 선정하면 블록내에 인접한 화소들 간에는 상관관계가 높으므로 오차가 최소가 되는 방향을 예측하여 그 사이에 있는 블록만을 보간한다. 예를 들면 각 방향당 블록 2가 블록 7보다 오차가 적고, 블록 5가 블록 4보다 오차가 적다면 최적의 반화소 단위 움직임 블록은 어떤 방향보다 위방향과 우측방향에 있을 확률이 크므로 블록 3만 보간하면 된다. 따라서 이러한 방법은 8개의 블록을 모두 보간할 필요없이 예측하여 블록2, 블록3, 블록 5만을 선별적으로 보간함으로써 많은 계산량을 줄일 수 있다.

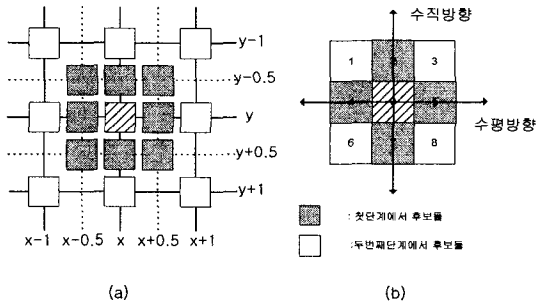


그림 2. (a) 기저 대역( $S_8$ )에서의 반화소 변위 (b) 수평 및 수직기준에서의 후보 반화소 단위 움직임 블록들

### 3.2 고해상도 계층에서 저해상도 계층으로의 선택적 국부 움직임 추정방법

계층적인 피라미드 구조의 다해상도 영상을 가지는 웨이블릿 변환 영역에서의 대표적인 움직임 추정 방법인 MRME는 분할된 영상에 대하여 기저 대역에서 추정된 움직임을 나머지 고해상도 부대역들의 초기 움직임으로 이용하여 정제(refine)을 통해 움직임을 추정한다. 그러나, 이 방법은 저해상도 계층의 작은 블록 크기에 의해 추정되므로 국부(local) 최소 오차 문제가 생길 수 있으며 고해상도 계층의 고주파 성분을 고려하지 않으므로 에일리싱 효과(aliasing effect)나 움직임이 많은 에지(edge)에 대한 정보를 잃을 수 있어 적은 계산량을 가지지만 정확한 움직임 추정이 되지 않

아서 화질을 저하시킨다. 그림 3에서처럼 Conklin<sup>[3]</sup> 등은 고해상도 계층에서 저해상도 계층으로의 MRME 움직임 추정 방법을 제안하였다. 이 방법은 정확한 움직임 추정을 할 수 있으나 고해상도에서의 많은 블록에 대하여 전역 탐색을 하므로 많은 계산량을 필요로 한다. 정확한 움직임을 추정하면서 이러한 계산량을 줄이기 위해 제한된 탐색 영역에서의 교번 부표본화(alternatively subsampling)움직임 추정을 제안하였다.

비록 기저대역에서의 움직임 추정은 고주파 성분의 부족과 에일리싱 효과로 인해 정확한 고해상도 계층의 움직임 추정은 어렵지만, 고해상도 계층에서의 움직임 추정의 중요한 단서가 된다. 따라서 Conklin의 방법에서는 독립적으로 고해상도 계층의 움직임을 추정했지만 탐색 영역을 줄이기 위해서 기저 대역에서의 움직임 벡터를 이용하여 제한시킬 탐색영역을 예측한다. 그림 4에서처럼 고해상도 계층에서의 원래 탐색 영역은  $N \times N$ 이지만, 연관성을 고려하여 기저 대역의 움직임 벡터가 우측 위쪽의 방향이라면 고해상도 계층의 탐색 영역도 우측 위쪽의 제한된 영역  $(N/2) \times (N/2)$ 으로 탐색하여 1/4의 계산량을 줄일 수 있다.

또한 제한된 영역에서의 움직임 벡터 탐색에서도 모든 탐색점에 대해 BMA를 하지 않고 교번적 부표본화를 함으로써 더욱 더 계산량을 줄였다. 이러한 과정은 아래와 같다.

단계 1 : 입력 영상을 이산적 웨이블릿 변환을 이용하여 N 계층의 다해상도 영상으로 분할한다.

단계 2 : N 계층의 부대역들을 일정한 크기의 블록으로 분할하고, 기저 대역( $S_8$ )의 블록들에 대하여 움직임 추정을 행한다.

단계 3 : 아래의 식을 이용하여, 고해상도 계층( $W_2^0$ )에서 적용적으로 움직임을 추정하고 계층을 감소시키면서 최저해상도 계층까지 스케일링과 정제를 행한다.

$$V_2^0(x, y) = V_8(x, y) \times p(x, y) + \Delta'(\hat{\alpha}x + \hat{\delta}y)$$

$$\bullet \text{ where } p(x, y) = \begin{cases} x = x \times 2 & \text{if } x = 2 \\ x = x \times 4 & \text{else} \\ y = y \times 2 & \text{if } y = 2 \\ y = y \times 4 & \text{else} \end{cases} \quad (2)$$

•  $\Delta'(\hat{\alpha}x + \hat{\delta}y)$ : alternating subsampling point.

$$V_j^0(x, y) = V_2^0(x, y) \div 2^{N-j} + \Delta(\hat{\alpha}x + \hat{\delta}y)$$

$$\text{for } j = 1, 2, \dots, N \quad (3)$$

단계 4 : 계층을 증가시키면서 최적의 움직임 벡터를 이용하여 움직임 보상된 예측 영상 ( $\hat{W}_{2^{N-j}}$ )과 차영상( $\bar{W}_{2^{N-j}}$ )을 형성한다.

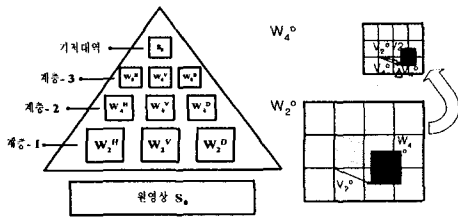


그림 3. 고해상도 대역에서 저해상도 대역으로 움직임 추정

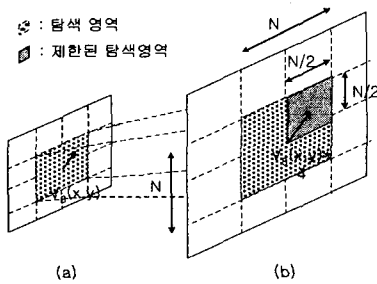


그림 4. 기저 대역(a)을 이용하여 고해상도 계층(b)의 선별적으로 제한된 탐색영역

#### IV. 실험 방법 및 결과

본 논문에서 제안한 방법의 성능을 평가하기 위해서 컴퓨터 모의 실험을 행하였다. 실험 영상으로는 720×480 크기의 football(Fb), table tennis(Tt), susie(Ss) 세영상을 각각 50프레임씩 사용하였다. 웨이블릿 변환은 좋은 성능을 나타낸다고 알려진 Daubechies의 8-tap필터를 사용하여 3계층 분할하였다. 기저대역에서의 블록의 크기는 2×2이고 Zhang의 MRME방법에서 정제시의 탐색 영역은 -2~+2화소로 하였고 제안방법에서 고해상도 계층의 탐색 영역은 가변적인 -8~+8로 하고 나머지 계층에서의 탐색 영역은 -2~+2로 하였다. 정합의 척도는 평균 절대치 오차 (mean absolute difference)을 사용하였고, 성능 평가 척도로 PSNR을 사용하였다. PSNR은 식(4)와 같다.

$$PSNR = 10 \log_{10} \frac{255^2}{MSE} \quad (4)$$

표 1에서처럼 제안한 방법을 사용하여 보간했을때의 예측율이 모든 방향으로 보간했을때의 95%이상으로 오차가 거의 없음에도 불구하고, 약 38%정도의 계산량이 감소하였음을 알 수 있었다. 표 2는 기존의 Zhang의 MRME방법과 제안한 방법의 PSNR과 계산량을 나타내었다. 계산량은 계산 비중이 가장 높은 고해상도 계층에서의 블록당 정합수로 하였다. 제안한 방법은 입력 영상에 따라 차이가 있으나, 기존의 방법에 비해

여 약간의 계산량 증가는 있었지만 정확한 움직임 추정으로 PSNR의 향상이 있음을 알 수 있었다.

표1. 기준 반화소 보간방법과 비교하여 제안한 선별적 반화소 보간방법의 예측율과 계산량

	예측율 (%)			계산량 (%)
	Fb	Tt	Ss	
기준 반화소방법	100	100	100	100
제안 반화소방법	94.7	95.1	95.0	62.5

표2. 기준 방법과 제안한 방법의 PSNR과 계산량

	PSNR (dB)			정합점/블록 (고해상도)
	Fb	Tt	Ss	
MRME 방법	23.69	25.56	33.55	25
제안 방법	24.83	26.34	34.27	27

#### V. 결론

본 논문은 웨이블릿 변환된 다해상도 계층의 기저대역에서 반화소단위 움직임 추정을 사용하면 더 나은 화질을 얻을 수 있음을 알 수 있었다. 그러나 이 방법은 계산량이 증가하므로 오차가 최소가 되는 방향을 예측하여 선별적인 보간을 행하여 정확한 움직임 벡터를 찾았으며 계산량을 상당히 줄였다. 또한 기존의 MRME에서는 저해상도 대역에서 고해상도 대역으로의 움직임 추정이므로 저해상도 계층 영역의 블록의 크기가 적기 때문에 에일리싱(aliasing)문제나 에러가 누적되는 단점이 있다. 따라서 이를 해결하기 위하여 에지 성분이 많은 고해상도 계층에서 저해상도 계층으로 선별적인 움직임을 추정하여 더 향상된 PSNR을 얻을 수 있었다. 모의 실험 결과 기존의 움직임 추정보다 약간의 계산량 증가로 더 좋은 화질을 얻을 수 있었다.

#### 참고문헌(또는 Reference)

- [1] Y. Zhang and S. Zafar, "Motion Compensated Wavelet Transform Coding for Color Video Compression," IEEE Trans. Cir&System for Video Tech, vol 2, no.3, pp. 285-296, Sept. 1992.
- [2] B. Girod, "Motion-compensating prediction with fractional-pel accuracy," IEEE Trans. on Comm., vol. 41, no. 4, pp. 604-612, Apr. 1993.
- [3] G.J. Conklin, S.S. Hemami, Multiresolution Motion Estimation, Proc. of ICASSP, pp. 2873-2876, 1997.