

論文2003-40SP-1-11

## 다해상도 움직임 예측을 이용한 동영상 물체탐지 알고리즘

(An algorithm for Video Object Detection using  
Multiresolution Motion Estimation)

趙 徹 勳 \* , 朴 章 漢 \* , 李 漢 佑 \* , 南宮 在 贊 \*

(Choul-Hun Cho, Chang-Han Park, Han-Woo Lee, and Jae-Chan Namkung)

### 요 약

본 논문에서는 웨이블릿 다해상도의 움직임 추정을 이용하여 물체탐지 알고리즘을 제안한다. 일반적인 물체추적 시스템에서 사용한 방법들은 부분적으로 움직임 추정의 특성을 지니고 있지만, 많은 계산량을 요구하는 단점이 있다. 웨이블릿 계수들의 부모 자식 관계를 이용한 다해상도 움직임 추정에서 해상도의 움직임 벡터가 상위 해상도의 움직임 추정에 대한 예측에 쓰이게 된다. 이 방법을 이용하여 상위 해상도에서 움직임 추정의 범위와 계산량을 줄일 수 있다. 제안된 방법은 3단계 SPIHT(Set Partitioning in Hierarchical Trees) 웨이블릿 변환했을 시 기존의 방법보다 약 40%의 계산량으로 움직임 추정을 할 수 있다. 실험을 통하여 제안한 방법은 동영상 비트율을 평균 11%의 MAD(Mean Absolute Difference)의 감소와 물체의 정확한 추적이 가능함을 보였다.

### Abstract

This paper proposes an object detection algorithm using the Multiresolution Motion Estimation(MRME) in wavelet domain. A existing motion estimation method has characteristics of motion estimation but it requires having computation. Motion estimation in higher resolution used the motion vector of the lower resolution with the MRME that has parent-child relationship on wavelet coefficients. This method reduces the search area of motion estimation in higher resolution and computational complexity. The computational complexity of the proposed method is about 40% of the existing method using 3-level Set Partitioning in Hierarchical Trees(SPIHT) wavelet transform. The experimental results with the proposed method showed about 11% decrease of Mean Absolute Difference(MAD) and gains able to precise tracking of object.

**Keyword** : object detection algorithm, wavelet, lower resolution motion vector, SPIHT, object tracking

### I. 서 론

동영상에서의 움직임 추적에 대한 연구는 오랫동안

\* 正會員, 光云大學校 컴퓨터工學科

(Department of Computer Engineering, Kwangwoon University)

接受日字:2002年5月28日, 수정완료일:2002年11月25日

컴퓨터 시각분야에서 연구되어 왔다. 인간의 시각 구조는 움직임에 매우 민감하여 움직임이 있는 곳에 초점을 맞추게 되며, 움직임이 있는 물체는 쉽게 감지할 수 있다. 움직임 추적을 위해 연속적으로 입력되는 동영상을 분석하여 실세계에 대한 정보를 추론하는 작업은 하드웨어 기술의 급속한 발전과 더불어 그 중요성이 강조되고 있다. 연속적으로 입력되는 영상에서 구해지는 이동 물체에 대한 특징을 이용하여 이동하는 물체

의 움직임 부분을 영역화 할 수 있다. 또한 물체의 인식과 움직임 분석 등을 수행할 수 있으며, 순차적인 영상에서 동일한 데이터 영역인 움직이지 않는 배경부분을 고려하여 동영상 전송 분야에서는 데이터 압축의 근거로 사용되고 있다. 특히 이동 물체를 탐지하고 탐지된 이동 물체의 동작 정보를 추출하는 작업은 군사 및 산업 분야에 적용될 수 있는 매우 흥미 있고 유용한 연구 분야로서 이를 위한 많은 방법들이 알려져 있다.

동영상에서의 움직임 추정 방법은 크게 세 가지로 분류할 수 있는데, 블록기반(block-based)추정 방법, 화소 기반(pixel-based)추정 방법, 광류 기반(optical flow-based)추정 방법 등이 있다. 블록기반 방법은 H.261, MPEG-1, MPEG-2등과 같은 영상 부호화 표준안에 적용되어 널리 사용되고 있지만, 회전(rotation)과 확대 및 축소(translation)를 보상하지 못하기 때문에 눈에 매우 거슬리는 블록화 잡음을 초래한다는 단점이 있다. 그리고 화소 기반 움직임 추정 방법은 객체의 움직임을 정확하게 표현할 수 있다는 장점이 있지만, 모든 화소의 움직임 벡터를 전송해야 한다는 어려움과 각각의 많은 화소에 대하여 실제 움직임을 찾아야 한다는 단점이 있다. 한편, 광류 기반 추정 방법은 시간에 따라 변화하는 영상의 강도패턴으로부터 2차원적 벡터 분포를 계산하여 움직임을 추정하는 방법으로 지금까지 제안되어 온 많은 동영상 해석 방법들의 근간이 되고 있다. 이러한 광류 기반 방법은 계산 속도가 오래 걸린다는 단점이 있다.

움직임 영상에 대한 압축 기법을 움직임 영상의 시간적, 공간적 중복성을 제거하기 위해 일반적으로 널리 사용되는 움직임 보상 부호(motion compensated coding)기법은 움직임을 추정(motion estimation)하여 보상하는 과정과 움직임이 보상된 차 신호를 부호화하는 과정을 거쳐 수행한다.

기존 MPEG등의 방식에서는 움직임 보상 부호화를 공간 영역에서 수행한 이후 변화 부호화를 함으로써 압축을 하게 되지만, 웨이블릿 변환의 주파수와 공간을 동시에 표현하는 특성을 이용하여 웨이블릿 변환영역에서 움직임 보상 부호화에 대한 많은 연구가 있었다 [7,10].

본 논문에서는 다해상도 움직임 추정의 개념을 사용하여 객체 검출시 부호화기의 계산량과 메모리를 줄이기 방법을 도입하였다. 움직임 추정시 만들어진 참조영

상과 부호화할 영상에서 웨이블릿의 계층적 상관성을 이용하여 한 단계 상위 해상도에서 구한 움직임 벡터를 기준으로 블록은 커지고 움직임 추정범위는 작게 하면서 움직임 벡터를 구한다. 이러한 방법으로 기존의 방법보다 보다 계산량을 줄이게 된다.

본 논문의 구성은 2장에서 웨이블릿과 SPIHT에 대하여 설명하며, 3장에서는 본 논문에서 제안한 시스템, 다해상도 움직임 추정방법과 물체의 후보 영역 생성방법에 대하여 설명한다. 4장에서는 실험 결과를 보이고, 5장에서 결론을 맺는다.

## II. 관련 연구

### 1. 웨이블릿 변환

웨이블릿 이론은 응용 수학에서 처음 소개된 후 인식 분야에서 연구되어온 다해상도 표현과 연관성이 있고, 필터 설계방법과 동일함이 밝혀진 후 현재 활발하게 연구되고 있는 변환 기법이다<sup>[4]</sup>.

현재 JPEG 2000의 표준 안으로 채택된 웨이블릿의 특징은 시간 축에서의 위치를 알 수 없는 푸리에 변환의 한계를 극복하고, 블록 단위로 처리하는 DCT (Discrete Cosine Transform)의 약점인 블록 효과가 없고 속도가 빠르며 압축율이 높다. 대표적인 웨이블릿 변환 기법으로 직교 웨이블릿(Orthogonal wavelet), 쌍직교 웨이블릿(Biorthogonal wavelet), 패킷 웨이블릿(Packet wavelet) 변환 등이 있다<sup>[11]</sup>. 웨이블릿 변환에서 하나의 기저함수를 웨이블릿이라 부르며, 웨이블릿은 하나의 대역 통과 필터이다. 푸리에 변환의 기저함수들과는 다르게 웨이블릿은 유한의 길이를 가지는 기저함수이므로 웨이블릿 변환은 모 웨이블릿(Mother Wavelet)라 불리는 원형의 웨이블릿의 수축과 팽창에 의해 얻어지는 웨이블릿들의 집합에 의해 구성된다. 그리고 웨이블릿 변환에서는 주파수라는 용어 대신에 스케일(scale)이라는 용어를 사용하며, 하나의 웨이블릿을 통과한 신호를 스케일의 상세 신호(detail signal)라 한다<sup>[6]</sup>.

웨이블릿 변환의 원리는 부대역 부호화(subband coding)를 하는 것이다. 이 부대역 부호화는 영상을 필터에 통과시켜 영상의 집합으로 만드는 것을 말한다. 각각의 부 영상은 공간적 주파수의 제한된 구간을 포함하고 있으며, 이들 각각의 영상 집합을 부대역이라고

부른다. 이렇게 변환된 부대역은 원 영상(Full band image)과 비교하였을 때, 대역폭이 줄어들게 되므로 이들을 다운샘플할 수 있다<sup>[2]</sup>. <그림 1>은 원 영상을 2개의 부대역으로 나누는 예이다.

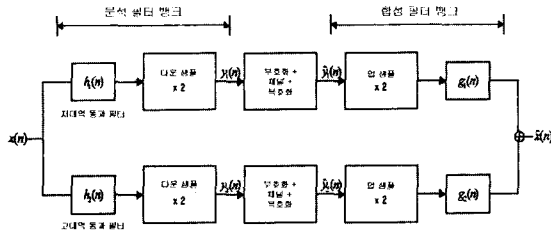


그림 1. 부대역 부호화(subband coding) 블록도  
Fig. 1. Subband coding block diagram.

2. SPIHT(Set Partitioning in Hierarchical Trees)

SPIHT은 낮은 비율 비트열(lower rate bit-stream)이 높은 비율 비트열(higher rate bit-stream)에 포함된다. 이 알고리즘은 트리 부호화(tree coding)를 기초로 하고 있다<sup>[1]</sup>. 웨이블릿 필터를 사용하여 변환된 계층적 구성의 계수들에 대해 임의의 계수  $x$ 가 임계값  $T$ 보다 작다면 중요도가 낮다고 가정하여 부호화를 하지 않고, 임계값  $T$ 보다 크다면 중요도가 크다고하여 부호화를 하게 되는 것이다.

이와 같은 알고리즘과 자기 유사성을 이용하여 다음과 같이 가정할 수 있다. 만약 계수  $x$ 가 임계값  $T$ 보다 작다면(중요도가 낮음:Insignificant),  $x$ 와 같은 방향성을 가지며, 같은 공간적인 위치를 가지는 하위레벨 계수들의 중요도 역시 낮다고 볼 수 있다. 따라서 부호화할 때 계수  $x$ 와 같은 공간적인 위치를 가지는 하위레벨 계수들을 대표하여 계수  $x$ 만을 부호화함으로써 원 영상의 화질이 많이 떨어지지 않고 압축율을 높일 수 있다. 이 때 계수  $x$ 를 부모(parent)라 하고,  $x$ 의 하위레벨 계수들을 자식(children 또는 descendent)이라고 한다.

자식들이 모두 임의 임계값  $T$ 보다 작은 부모를 ZTR(ZeroTree Root)이라고 하여 부호화하지만, 만약 자식들 중 하나의 계수라도 임계값  $T$ 보다 크다면 부모는 IZ(isolated zero)하여 자식들의 대표로써 부호화되지 않고 오직 IZ로써 부호화된다. Shapiro는 계수들을 부호화하기 위해 ZTR과 IZ, POS(Positive), NEG(Negative) 등 4개의 기호(symbol)를 사용하였다. 이와

같은 알고리즘을 통해 계층적 구조인 웨이블릿 계수들은 계수값의 절대값 크기에 따라 스캔(scan)이 이루어지므로 중요도에 따른 부호화를 할 수 있게 되는 것이다. 이것은 DCT에서 변환된  $8 \times 8$ 의 블록에 왼쪽 맨 위에 존재하는 DC계수를 많은 비트로 부호화하고, 나머지 AC계수에 대하여 지그재그 스캔으로 AC계수의 절대값이 큰것만 부호화 하는 것과 같은 맥락의 의미라고 할 수 있다. <그림 2>는 SPIHT 알고리즘을 이용한 데이터 구조를 보이고 있다.

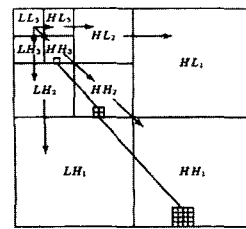


그림 2. SPIHT 알고리즘을 이용한 데이터 구조  
Fig. 2. Data Structure using SPIHT algorithm.

III. 제안한 시스템 설계

1. 제안한 전체 시스템

본 논문에서 제안한 물체추적 시스템으로써, <그림 3>은 전처리 과정으로 잡음 제거와 3단계 DWT(Discrete Wavelet Transform)로 처리하고, 계속해서 물체 인식 과정에서는 3단계 LL차 영상을 후보 영역을 생성하며, 물체 추적 과정에서는 움직임 평가를 통해 움직임을 추출한다. 따라서 물체 추적 영상을 얻을 수 있다.

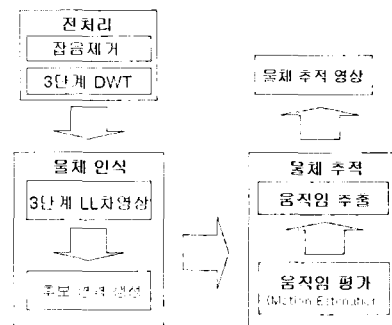


그림 3. 제안된 물체추적 시스템  
Fig. 3. Proposed object tracking system.

## 2. 다해상도의 움직임 추정

본 논문에서 제안하는 물체추적 시스템의 전체 알고리즘의 구성은 <그림 3>과 같으며, 다음과 같이 간략히 요약하면,

- a. 입력된 영상을 그레이 코드(gray code)로 변환,
- b. 메디안 필터(median filter)를 사용하여 잡음을 제거,
- c. 잡음이 제거된 영상을 3단계 웨이블릿 변환,
- d. 이전 프레임과 현재 프레임과의 3단계 저주파 대역 부분을 비교하여 움직임의 여부를 살핀다. 만약, 움직임이 없을 때는 다음 영상을 받아 들이고, 움직임이 검출 되었을 때는 그 움직임이 있는 부분에 후보 영역을 생성,
- e. 후보 영역 설정 후에 다 해상도 움직임 추정의 개념을 도입한다.

웨이블릿 변환의 다른 계층상에 위치한 물체의 움직임은 동일한 움직임 구조가 여러 스케일 및 주파수로 표현된 것이므로, 상관 관계가 매우 높은 성질을 갖고 있다. 웨이블릿 피라미드구조에서 각각의 계층의 움직임 경향은 다르지만, 매우 높은 연관성이 있다. 이 성질을 이용한 움직임 추정 방식을 다해상도 움직임 추정(multiresolution motion estimation)이라 한다<sup>[9]</sup>.

Uz등은 거친 영역에서 세밀한 영역으로(coarse-to-fine) 진행하는 접근방식을 제안하였다<sup>[5]</sup>. 이 방식은 최상위 단계의 부대역에 대하여 일반적인 블록 정합 기법을 사용하여 움직임 벡터를 구한 후, 다음 단계 부대역의 움직임 추정을 한 단계 위에 대응되는 부대역의 움직임 벡터로부터 예측(predict)을 하게 된다. 이 예측으로 인해 부대역의 움직임 추정은 작은 탐색 범위를 갖게 된다. 이 기법은 모든 웨이블릿 부대역에 대하여 고정된 블록 크기와 추정 범위를 가지는 항등 블록 움직임 추정 방법이다.

본 논문에서 제안한 움직임 추정 방법은 Uz가 제안한 것과 같이 항등 블록을 가지는 웨이블릿 구조의 계층적 상관성을 이용하여 물체의 움직임 벡터를 얻는다.

## 3. 물체의 후보 영역 생성

동영상 정보는 공간적인 중복성과 시간적인 중복성이 있으며, 공간적 중복성을 줄이는 부호화 기법을 프레임내 부호화(intra-frame coding)라 하고, 시간적 중복성을 제거하는 것을 프레임간 부호화(inter-frame

coding)라 한다. 동영상은 특히 시간적인 중복성을 효율적으로 줄여야 하는데, 이전 프레임의 영상 정보를 그대로 활용하면서 변화된 부분, 즉 움직임이 발생한 부분만을 찾아서 그 정보만을 보내 주면 뛰어난 부호화 효율을 얻을 수 있다. 이때 현재 영상에서 움직임이 발생한 영역과 가장 비슷한 부분을 이전 프레임 영상에서 찾는 과정을 움직임 추정이라고 하며, 이 추정된 정보를 현재 프레임에 예측하여 예측 오차만을 부호화하게 된다. 변위 차분 프레임(Displacement Frame Difference : DFD)은 움직임 추정 과정에서 얻어진 움직임 벡터에 의해 재구성된 참조 프레임과 현재 프레임간의 예측 오차를 의미하며, 두 프레임간의 시간적 중복성이 매우 높다면 이 변위 차분 프레임의 정보량은 매우 낮아지게 된다. 일반적으로 동영상 부호화 시스템에서는 이러한 움직임 보상 오차 프레임인 변위 차분 프레임과 움직임 벡터를 부호화하는 방식을 사용하므로 효율적인 부호화를 위해서는 정확한 움직임 추정이 필요하다<sup>[6]</sup>.

웨이블릿 변환 영역에서의 움직임 추정은 이동-변환 성질로 인하여 프레임간에 움직임이 일어났을 때, 그 계수값이 공간영역에서 이동된 위치에 따라 값이 달라지게 되므로 정확한 움직임 추정이 어렵다. 저 대역 이동법을 사용한다면 현재 프레임의 임의의 화소에서 이동된 화소의 변화된 웨이블릿 계수값을 참조할 수 있기 때문에 일반적인 웨이블릿 변환 영역에서의 움직임 추정보다 부호화면에서 더 효율적인 움직임 추정을 할 수 있다.

웨이블릿 변환 영역에서의 움직임 추정은 현재 프레임의 임의의 화소에서 이동된 화소의 변화된 웨이블릿 계수값을 참조한다. 웨이블릿 분해시 주어진 해상도에서 임의의 계수 값들은 <그림 4>와 같이 최대 주파수 대역을 제외하고, 한 단계의 세밀한(fine) 해상도의 대응되는 위치의 계수들과 공간영역에서 같은 화소이며, 그 계수간의 상관성(correlation)이 매우 높다. 거친(coarse) 해상도에 위치한 계수를 부모라 부르고, 그곳에 대응되는 위치의 다음 단계의 세밀한 해상도의 4개의 계수들을 자식이라고 부른다. 이 관계를 이용한 것이 웨이블릿 트리(wavelet tree)라고 불리는 자료 구조이다<sup>[9]</sup>.

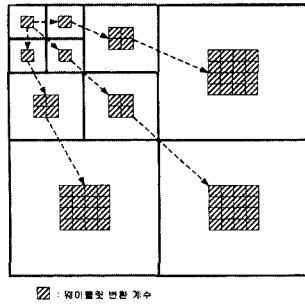


그림 4. 웨이블릿 계수의 부모-자식 관계  
Fig. 4. Parent-child relation of wavelet coefficient.

웨이블릿 블록에 대하여 움직임 추정을 하기 위해 MPEG이나 H.263등과 같은 동영상 부호화 기법에서 널리 쓰이는 전 영역 블록 정합 알고리즘(full-search block matching algorithms)을 사용한다. 이 기법은 프레임의 일정 크기의 블록으로 나눈 후 각 블록이 이전 프레임의 어떤 블록과 가장 유사한지 탐색하여 각 블록의 움직임을 결정하는 방식이다. 여기서 사용되는 정합 기준으로는 일반적으로 평균 절대 오차(MAD : Mean Absolute Difference)를 사용한다<sup>[4]</sup>. <그림 5>는 전영역 블록 정합 기법을 이용한 움직임 추정을 보인 그림이다.

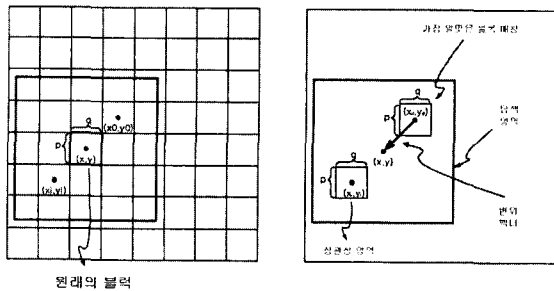


그림 5. 전영역 블록 정합 기법을 이용한 움직임 추정  
Fig. 5. Motion estimation using the full region block matching method.

평균 절대 오차 함수에 의한 움직임 벡터 (dx, dy)의 결정은 식 (1)과 같다.

$$(dx, dy) = \arg \min \sum_{p=-M/2}^{M/2} \sum_{q=-N/2}^{N/2} |I_{cur}(x+p, y+q) - I_{ref}(x+p+dx, y+q+dy)| \quad (1)$$

$I_{cur}(x+p, y+q)$  : 현재 프레임의 블록

$I_{ref}(x+p+dx, y+q+dy)$  : 이전 프레임의 블록

$M, N$  : 블록의 가로, 세로 크기

본 논문에서 이용한 움직임 추정은 기존의 방법으로, 3단계 웨이블릿 분해시 k번째 블록에 대한 블록 정합 움직임 추정 알고리즘을 적용한 식 (2)와 같다<sup>[3]</sup>.

$$MAD_k(dx, dy) = \sum_{i=1}^k \sum_{x_1=x_{i-1}}^{k+M-2^{i-1}k} \sum_{y_1=y_{i-1}}^{2^{i-1}k} \{a-b\} + \sum_{x_1=x_{i-1}}^{x_{i-1}+M-2^{i-1}k} \sum_{y_1=y_{i-1}}^{2^{i-1}k} \{c-d\} \quad (2)$$

(단,  $a = HH_{cur}^{(i)}(x_i, y_i)$ ,

$$b = HH_{ref}^{(i)}(dx \% 2^i, dy \% 2^i, x_i + \left\lfloor \frac{dx}{2^i} \right\rfloor, y_i + \left\lfloor \frac{dy}{2^i} \right\rfloor),$$

$$c = LL_{cur}^{(3)}(x_3, y_3),$$

$$d = LL_{ref}^{(3)}(dx \% 2^3, dy \% 2^3, x_3 + \left\lfloor \frac{dx}{2^3} \right\rfloor, y_3 + \left\lfloor \frac{dy}{2^3} \right\rfloor))$$

여기서,  $HH_{cur}^{(i)}$ 는 현재 프레임의 i번째 레벨의 고주파수 대역의 움직임 보상 블록이고, 마찬가지로  $LL_{cur}^{(3)}$ 은 3번째 단계의 저주파수 대역에 대한 움직임 보상 블록이다. 마찬가지로 여기서,  $HH_{ref}^{(i)}$ 는 이전 프레임에서 i번째의 웨이블릿 분해 단계의 고주파수 대역을 나타내며,  $LL_{ref}^{(3)}$ 값은 이전 프레임의 최저 주파수 대역을 나타낸다. i번째 웨이블릿 분해 단계의 저주파수 대역에서 k번째 웨이블릿 블록의 시작점  $(x_{i,k}, y_{i,k})$ 은 식 (3)과 같이 정의된다.

$$x_{i,k} = \frac{x_{0,k}}{2^i}, \quad y_{i,k} = \frac{y_{0,k}}{2^i} \quad (3)$$

(단,  $(x_{0,k}, y_{0,k})$ 은 공간영역에서 k번째 웨이블릿 블록의 시작점이다.)

k번째 웨이블릿 블록의 최적의 움직임은 식 (4)와 같다.

$$(dx_{\min}, dy_{\min})_k = \min_{(dx, dy) \in B} MAD_k(dx, dy) \quad (4)$$

움직임 벡터를 (dx, dy)라 놓을 때, 움직임 보상 블록에 대한 것은 식 (5), 식 (6)과 같다.

$$HH_{mc}^{(i)}(x_i, y_i) = HH_{ref}^{(i)}(dx \% 2^i, dy \% 2^i, x_i + \left\lfloor \frac{dx}{2^i} \right\rfloor, y_i + \left\lfloor \frac{dy}{2^i} \right\rfloor) \quad (5)$$

$$LL_{mc}^{(3)}(x_i, y_i) = LL_{ref}^{(3)}(dx \% 2^3, dy \% 2^3, x_3 + \left\lfloor \frac{dx}{2^3} \right\rfloor, y_3 + \left\lfloor \frac{dy}{2^3} \right\rfloor) \quad (6)$$

(단,  $i = 1, 2, 3$ ,  $x_{i,k} \leq x_i \leq x_{i,k} + \frac{M}{2^i} - 1$ ,  $y_{i,k} \leq y_i \leq y_{i,k} + \frac{N}{2^i} - 1$ )

$HH_{mc}^{(i)}$ 는  $i$ 번째 레벨의 고주파수 대역의 움직임 보상 블록이고, 마찬가지로  $LL_{mc}^{(3)}$ 은 3번째 단계의 저주파수 대역에 대한 움직임 보상 블록이다.

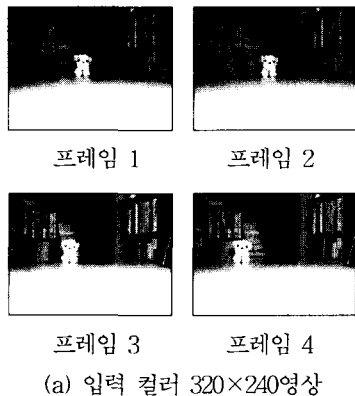
저주파 대역의 움직임 벡터는  $2 \times 2$ 블록에 대하여 움직임 벡터를 구하고, 같은 해상도의 고주파대역의 움직임 벡터를 구한다. 나머지 저주파 성분을 포함한 대역은 움직임 벡터를 구하지 않는다. 이렇게 하여 각 단계로 넘어가는 방법으로 계속 구하여 진다. 여기에서 한 단계 상위 움직임 추정시 블록의 길이는 2배로 커지게 되지만, 탐색 영역은 고주파 대역만을 하기 때문에 계산량을 줄이게 된다. 그러나 10개의 대역 중 4개의 대역만을 선택하기 때문에 세밀한 움직임 추정은 할 수 없지만, 현재 프레임의 예측 부호화가 더욱 용이하게 되므로, 변위 차분 프레임에 대한 부호화량은 줄어들게 되어 기존의 방법보다 부호화 효율이 더 좋아졌다고 할 수 있다.

IV. 실험 및 고찰

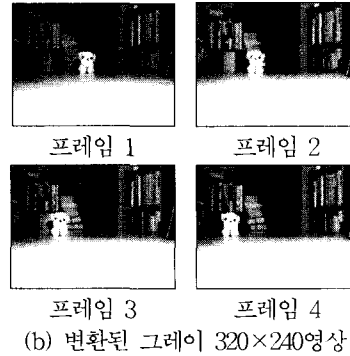
1. 실험

본 논문에서 제안하는 방법에 대한 검증으로 Pentium-III, 256Mb 환경에서  $320(240)$ 크기를 갖는 영상의 100프레임을 프로그램 Visual C++6.0을 이용하여 실험하였다.

<그림 6>은 실험에 사용된  $320 \times 240$ 크기를 갖는 영상을 보여주고 있다. (a)는 입력 컬러 영상이며, (b)는 그레이 영상으로 변환된 영상이다.



(a) 입력 컬러  $320 \times 240$ 영상



(b) 변환된 그레이  $320 \times 240$ 영상

그림 6. 실험에 사용된 영상  
Fig. 6. Used image for experimental.

<그림 7>은 <그림 6(b)>의 그레이 영상에 대한 3단계 웨이블릿 변환을 적용했을 때의 프레임이다.

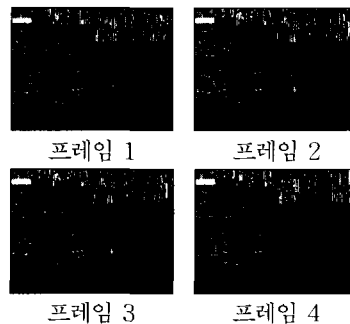
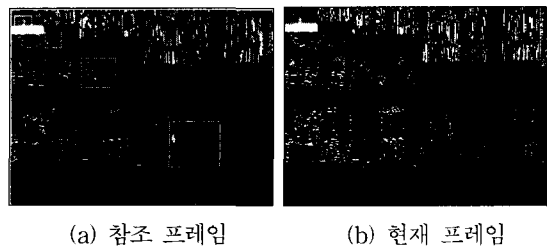


그림 7. 프레임에 대한 3단계 웨이블릿 변환  
Fig 7. 3-step wavelet transformation for frame.

움직임 추정할 때, 기존의 방법에는  $16 \times 16$ 크기의 블록에 대하여 탐색 범위를  $[-16, 16]$ 으로 하였지만, 제안한 방법은  $16 \times 16$  크기에 대응하여, 저주파수 대역부분의 블록을  $2 \times 2$  크기로 놓았으며, 참조 프레임은 현재 프레임의 바로 이전 영상으로 하였다. <그림 8>의 (a)는 참조 프레임이며, <그림 8>의 (b)는 현재 프레임이다.



(a) 참조 프레임 (b) 현재 프레임  
그림 8. 주파수 대역 부분의 블록  
Fig. 8. Block of low-subband region.

<그림 9> 이전 영상과 현재 영상에서의 후보 영역을 추출하는 화면이며, 현재 영역에서 움직임이 발생한 영역과 가장 비슷한 프레임을 영상에서 찾는 과정으로, 추정된 정보를 현재 프레임에서 예측하여 예측 오차만을 부호화하게 된다.

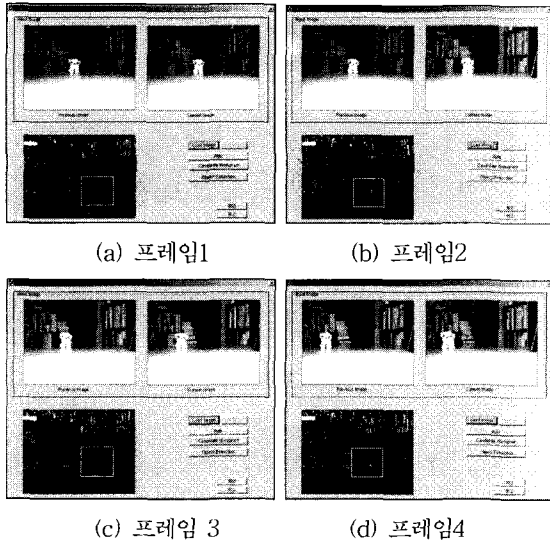


그림 9. 이전 영상과 현재 영상에서의 후보 영역 검출 화면 결과

Fig. 9. Candidate region extract display result for previous image and current image.

<표 1>은 기존의 움직임 추정 방법과 제안한 움직임 추정 방법의 프레임당 평균 MAD의 비교이다. 공간 영역 방법은 일반적인 공간영역에서 움직임 추정이고, 직접 방법은 웨이블릿 변환영역에서 두 프레임간에 공간영역 방식의 움직임을 추정을 하는 것이다. 여기에서 기존의 방법과 비교하기 위해 100프레임의 영상을 사용하였다. 또한 대표적인 영상인 Foreman과 Carphone의 동영상과도 비교해 보았다. 결과적으로 평균 MAD는

표 1. 여러 동영상에 대한 100프레임의 평균 MAD

Table 1. Average MAD of 100 frame for various motion picture.

순서	공간영역 방법	직접 방법	제안한 방법
Frame 1~ Frame 100	1.0283	1.0242	0.8756
Foreman	2.9869	3.4321	2.5769
Carphone	2.1324	2.4525	1.8765

직접 방법보다 제안된 방법이 0.1486낮은 값을 얻을 수 있으며, Foreman은 0.8552, Carphone은 0.576의 낮은 값을 얻을 수 있었다.

<표 2>는 실제로 물체의 움직임을 추적하는 데 소요되는 처리 내용으로, 그레이 영상 변화, 웨이블릿 변환, 후보 영역 설정, 움직임 벡터 계산이다. 본 시스템에서는 초당 3.7 프레임을 처리할 수 있다.

표 2. 프레임당 처리 시간  
Table 2. Frame per processing time

처리 내용	기존 방법	처리 시간	프레임당 처리 시간
그레이 영상 변화	0.33	0.33	0.33/100 = 0.0033
웨이블릿 변환	0.54	0.54	0.54/100 = 0.0054
후보 영역 설정	0.53	0.19	0.19/100 = 0.0019
움직임 벡터 계산	1.59	0.42	0.42/100 = 0.0042

<표 3>은 영상에서의 변이, 변이와 회전, 회전이 발생했을 때, 제안된 알고리즘과 공간적인 방법의 검출 성공과 검출 실패율을 비교 평가한 결과이다.

표 3. 실험 결과의 비교  
Table 3. Comparison of experiment result.

영상의 변화	전체 영상수	제안된 알고리즘		공간적인 방법	
		검출 성공	검출 실패	검출 성공	검출 실패
변이	50	48	2	43	7
변이와 회전	25	20	5	19	6
회전	25	18	7	18	7
백분율(%)		86.0	14.0	80.0	20.0

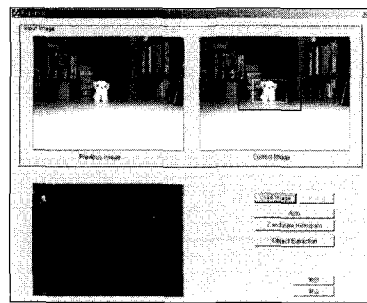


그림 10. 분리된 객체  
Fig. 10. Separated object.

<그림 10>은 움직임 벡터 계산이 끝난 뒤에 웨이블릿 역변환하여 저주파수 성분에서 정해진 부분을 계속 복원하면, 객체를 추출 할 수 있다. 여기에서 히스토그램을 사용하여 객체를 완전하게 추출할 수 있다.

## V. 결 론

본 논문에서는 웨이블릿 다해상도의 움직임 추정을 이용하여 영상의 객체 움직임에 따른 물체탐지 알고리즘을 제안하며, 물체탐지 시스템의 개발 가능성을 보였다. 일반적인 물체추적 시스템에서 사용한 방법들은 부분적으로는 최적의 특성을 지니고 있지만, 많은 계산량을 요구하는 단점이 있다.

따라서 제안한 방법은 참조 프레임과 현재 프레임에 대하여 웨이블릿 계수의 계층적 상관성을 이용한 다해상도 움직임 추정 개념을 적용하여 기존의 움직임 추정시 계산량과 메모리를 약 40%정도 줄일 수 있었다. 이 과정은 각 대역에 있는 움직임 벡터 중 저주파 성분이 포함된 고주파 대역의 일부분만을 사용하여 움직임 벡터의 부호화량의 감소와 변위 차분 프레임의 부호화량 감소의 이점이 있으나, 세밀한 움직임 추정의 단점이 있었다.

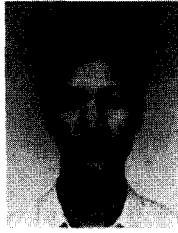
향후 연구 과제로 더 세밀한 움직임 추정을 위한 움직임 벡터의 선택 방법과 더 나은 부호화 효율을 위한 양자화기의 설계가 필요하다. 또한 한 영상에서의 실시간으로 복수의 객체를 탐지하는 연구가 계속되어야 할 것이다.

## 참 고 문 헌

- [1] A. Said and W. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees", IEEE Trans. Circuits and Systems for Video Technology, Vol. 6, No. 3, pp. 243~250, June 1996.
- [2] A. S. Lewis and G. Knowles, "Image compression using the 2-D wavelet transform", IEEE Trans. on Image Processing, Vol. 1, pp. 244~250, April 1992.
- [3] H.W.Park and H.S.Kim, "Motion Estimation using Low-Band Shift Method for Wavelet Based Moving Picture Coding", IEEE Trans. Image Processing, Vol. 9, No. 4, pp. 577~587, April 2000.
- [4] I. Daubechies, "The Wavelet Transform, Time-Frequency Localization and Signal Analysis", IEEE Trans. Information Theory, Vol. 36, pp. 961~1005, 1990.
- [5] K. Uz, M. Vetterli and D. Leg all, "Interpolative Multiresolution Coding of Advanced Television with Compatible Subchannels", IEEE Trans. On CSVT. Vol. 1, No. 1, pp. 86~99, March 1991.
- [6] S. G. Mallat, "A theory for multiresolution signal decomposition the wavelet representation", IEEE Trans. Patt. Anal. Mach. Intell., Vol. 11, pp. 674~693, July 1989.
- [7] S. Kim, T. Aboulnasr, and s. Panchanathan, "Adaptive Multiresolution Motion Estimation Techniques for Wavelet-Based Video Coding", Proc SPIE Visual Communications Image Processing, Vol. 3309, pp. 965~974, Jan. 1998.
- [8] S. Kim, S. Rhee, J. Jeon, and K. Park, "Interframe Coding Using Two-Stage Variable Block-Size Multiresolution Motion Estimation and Wavelet Decomposition", IEEE Trans. Circuits and Systems for Video Technology, Vol. 8, pp. 399~410, Aug. 1998.
- [9] V. Bhaskaran and K. Konstantinides, Image and Video Compression Standards, Kluwer Academic Publishers, 1995.
- [10] Y. Shi and H. Sun, 7/710 and Video Compression for Multimedia Engineering, CRC Press, 2000.
- [11] ISO/IEC CD15444-1, "JPEG2000 Image Coding System", ver 1.0, Dec. 1999.



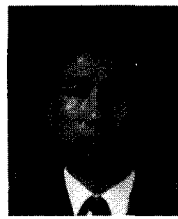
저 자 소 개



趙 徹 勳(正會員)

2000년 2월 : 광운대학교 컴퓨터공학과 졸업. 2002년 2월 : 광운대학교 대학원 컴퓨터공학과 석사. 2002년 3월~현재 (주)현대 시멘트 재직 중. <주관심분야 : 스테레오 영상, 신경망, MPEG, 영상처리, 영상인식, 영

상압축>



李 漢 佑(正會員)

1988년 2월 : 광운대학교 전산대학원 공학석사. 2002년 3월~현재 광운대학교 대학원 컴퓨터공학과 박사과정 수료 현재 인덕대학 겸임교수, 비전정보텍 대표 <주관심분야 : 신경망, 패턴인식, 이미지프로세싱>



朴 章 漢(正會員)

1997년 2월 : 방송통신대학교 전산계산학과 졸업. 1999년 8월 : 광운대학교 전산대학원 소프트웨어공학과 석사. 2001년 3월~현재 광운대학교 대학원 컴퓨터공학과 박사과정 재학 중. <주관심분야 : 스테레오 비전,

영상처리, 영상인식, 신경망, MPEG, 콤포넌트, UML>



南宮 在 贊(正會員)

1970년 2월 : 인하대학교 전기공학과 졸업. 1976년 8월 : 인하대학교 대학원 전자공학과 석사. 1982년 2월 : 인하대학교 대학원 전자공학과 박사. 1979년 3월~현재 광운대학교 컴퓨터공학과 교수. <주관심분야 :

신경망, 영상처리, 문자인식, 패턴인식, 영상통신, MPEG, 3D영상>