

카메라 패닝 보상에 기반한 계층적 블록 정합 알고리즘

곽 노 윤[†] · 황 병 원^{††}

요 약

본고에서는 움직임 추정 성능을 개선하고 과도한 연산량과 전송 부담을 경감시키기 위해 HBMA에 기반한 가변 움직임 추정 기법을 제안한다. 제안된 알고리즘은 크게 다음과 같이 네 단계로 구성된다. 우선, 연속된 두 프레임 간의 차영상 윤곽 정보에서 정의한 블록 활동도를 평균하여 현재 영상의 평균 블록 활동도를 산출한다. 두 번째로, 이렇게 산출한 평균 블록 활동도를 통해 카메라 패닝의 유무를 검출한 후, 웨이블릿 변환에 의해 구성된 피라미드 계층 구조상에서 카메라 패닝 벡터를 추정하여 보상한다. 다음으로, 카메라 패닝 보상 후에 정의한 블록 활동도를 토대로 각 블록을 움직임 블록, 준 움직임 블록, 비 움직임 블록 중 어느 하나로 분류한 검색 테이블을 작성한다. 마지막으로, 제안된 가변 HBMA는 검색 테이블을 참조하여 블록 크기를 가변시키고 초기 탐색 계층 및 탐색 영역을 적용적으로 선정함으로써 피라미드 계층 구조상에서 효율적인 고속 움직임 추정을 수행할 수 있다. 이상에서 설명한 각 단계에서 요구되는 비용함수는 차영상 윤곽정보를 통해 획득한 블록 활동도를 공통적으로 이용한다.

A Hierarchical Block Matching Algorithm Based on Camera Panning Compensation

No-Yoon Kwak[†] · Byong-Won Hwang^{††}

ABSTRACT

In this paper, a variable motion estimation scheme based on HBMA(Hierarchical Block Matching Algorithm) to improve the performance and to reduce heavy computational and transmission load, is presented. The proposed algorithm is composed of four steps. First, block activity for each block is defined using the edge information of differential image between two sequential images, and then average block activity of the present image is found by taking the mean of block activity. Secondly, camera pan compensation is carried out, according to the average activity of the image, in the hierarchical pyramid structure constructed by wavelet transform. Next, the LUT classifying each block into one among Moving, No Moving, Semi-Moving Block according to the block activity compensated camera pan is obtained. Finally, as varying the block size and adaptively selecting the initial search layer and the search range referring to LUT, the proposed variable HBMA can effectively carries out fast motion estimation in the hierarchical pyramid structure. The cost function needed above-mentioned each step is only the block activity defined by the edge information of the differential image in the sequential images.

[†] 정 회 원 : 한국항공대학교 대학원 항공전자공학과

^{††} 정 회 원 : 한국항공대학교 항공전자공학과 교수

논문접수 : 1999년 1월 11일, 심사완료 : 1999년 7월 14일

1. 서 론

오늘날, 프레임간 영상 부호화는 움직임 추정/보상 부호화(motion estimation/compensation coding)로 대표된다. 이는 현재의 영상과 시간적으로 전후에 인접한 영상간에서 움직임 정보를 추출해 내고, 이 움직임 정보를 이용하여 현재의 영상을 예측하는 방식으로, 화소 순환 방식(PRA: Pal Recursive Algorithm), 블록 정합 방식, 위상 상관 방식, 시공간 경사 방식, 객체 기반 움직임 추정 방식 등이 대표적이다. 다양한 움직임 추정/보상 부호화 방식들 중에서 블록 정합 알고리즘(BMA: Block Matching Algorithm)은 영상의 품질, 연산량, 전송 부담, 하드웨어적 복잡도 측면에서 적절한 상호 타협 관계(tradeoff)를 형성하는 것으로 인정되어 온 바, 기존의 각종 영상 표준 코덱(MPEG-1, MPEG-2, H.261, H.263, HDTV etc.)의 프레임간 부호화 방식으로 널리 채택되어 있다[1][2][3]. 최근에는 MPEG 4 표준화 작업[4]에 편승하여 객체 기반 움직임 추정 방식[5]이 프레임간 부호화 기법의 새로운 대안으로 부상하고 있는데, 블록 정합 알고리즘과 객체 기반 움직임 추정 방식은 각각 단위 블록과 객체에 포함된 모든 화소들이 동일한 움직임 벡터를 갖는 강체(rigid)로 가정한 움직임 추정 기법이라는 점에서 유사하며, 비록 후자가 임의의 모양을 갖는 객체에 기반하고 있지만, 일정 부분에 걸쳐 전자에 기초하고 있는 움직임 추정 방식이라고 말할 수 있다.

주지하다시피, 블록 정합 알고리즘은 다양한 방식으로 구현될 수 있으나, 그 중에서 전역 탐색(full search) 방식이 기설정된 탐색 영역 내에서 가장 정확한 움직임 벡터를 제공하는 것으로 알려지고 있다. 그러나, 전역 탐색 방식은 정합될 수 있는 경우의 수가 탐색 영역의 크기에 따라 기하급수적으로 증가함에 따라 정합 연산을 수행할 시에 많은 연산량을 요구하기 때문에 실시간적인 처리를 위해서는 탐색 영역의 확장 자유도가 심하게 제한받는 문제가 있다[1]. 따라서, 단계적으로 예상 범위를 줄어나감에 따라 정합 연산량을 줄임과 동시에 정확한 움직임 벡터를 탐색하려는 의도에서 출발한 고속 블록 정합 알고리즘들이 다수 제안되어 있는데, 대표적인 것으로는 삼단계 탐색(TSS: Three Step Search), 2-D 대수적 탐색(2-D logarithmic search), 공역측 탐색(conjugate direction search), 병렬 계층 1-D 탐색(parallel hierarchical 1-D search)[1][2] 등이 있다.

그러나, 이러한 고속 블록 정합 알고리즘들은 전역 탐색에 비해 추정 오차가 크고, 추정된 움직임 벡터의 신뢰성이 현저히 떨어지는 문제점을 내포하고 있음에 따라 추정된 움직임 벡터에 대한 신뢰성을 최대한 유지하면서 탐색 영역과 연산량에 대한 문제를 해결하기 위해 계층 탐색 방식(Hierarchical Search Algorithm)이 제안되었지만, 이 방식 또한 기존의 고속 알고리즘에서와 같이 서로 다른 해상도를 갖는 계층들 간에서 점진적으로 움직임을 추정해가는 방식이기 때문에 움직임 추정 벡터가 국부 최소(local minima)에 빠질 가능성을 배제시킬 수 없었다[1][6][7].

한편, 움직임 추정 방식과는 별도로 카메라의 움직임 측면에서 블록 정합 알고리즘을 고찰할 때, 기존의 블록 정합 알고리즘들은 카메라의 움직임을 고려하지 않기 때문에 카메라의 패닝(panning), 확대/축소(zooming), 회전(rotation)과 같은 카메라의 움직임이 발생할 경우, 카메라의 운동이 없었을 시에는 정지 영역으로 간주되던 배경 또한, 동영역으로 간주됨으로써 움직임 추정 성능을 저하시킬 뿐만 아니라 영벡터(zero vector)의 추정 빈도수를 감소시켜 전송 부담을 증가시키는 문제가 있었다[8].

본고에서는 이상에서 지적한 바와 같은 문제점들을 최소화하기 위해 움직임 추정 시간을 단축시키기 위한 고속 탐색 알고리즘을 제안함과 동시에 카메라 패닝을 보상함으로써 움직임 추정 효율을 개선시킨 계층적 가변 블록 정합 알고리즘을 제안하고 있다.

2. 기존의 계층적 블록 정합 알고리즘

일반적인 계층적 블록 정합 알고리즘(HBMA)은 점진적인(coarse-to-fine) 움직임 추정 방식으로써, 낮은 해상도를 갖는 영상에서는 전체적인 움직임의 오정합을 없애주고 높은 해상도로 감에 따라 초기에 추정된 값을 재조정함에 따라 고속으로 움직임을 탐색할 수 있는 블록 정합 알고리즘의 하나이다. 기존의 대표적인 계층적 블록 정합 알고리즘(HBMA)의 수행 절차를 (그림 1)을 참조하여 단계적으로 살펴보면 다음과 같다.

단계 1:

영상을 웨이블릿 변환을 이용하여 k 개 계층으로 분해한다. 그 결과, 최상위 계층(k 계층)에서는 4개의 부영상들이 존재하게 되고, $k-1$ 이하의 계층에서는 각

가 3개의 부영상들이 존재하게 되어 모두 $3k+1$ 개의 부영상을 얻게 된다.

단계 2:

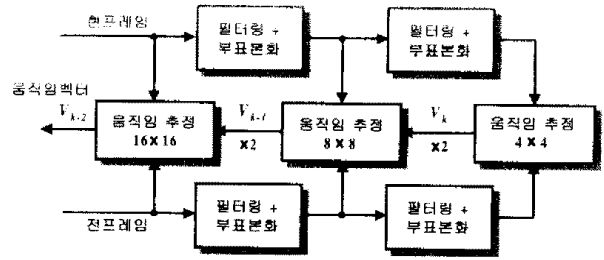
k 계층에 있는 4개의 부영상중 직주파 성분만을 가지고 있는 부영상을 $2^{-k}M_b \times 2^{-k}N_b$ 크기의 블록으로 분할하여 각 블록에 대한 움직임 벡터를 구한다. 이 움직임 벡터는 같은 계층의 다른 주파수 밴드 성분들 가지고 있는 부영상으로 전해져 움직임 탐색의 초기값으로 사용한다. 이는 각 계층 및 같은 계층에서의 부영상들은 움직임 벡터가 서로 다르나, 깊은 상관관계가 있고 움직임 정보의 중복성이 있다는 것을 이용하는 것이다.

단계 3:

k 계층에서 나온 주파수 밴드 성분을 가지고 있는 3개의 부영상들은 k 계층의 직주파 부영상에서처럼 $2^{-k}M_b \times 2^{-k}N_b$ 크기의 블록으로 분할한다. k 계층에서 나온 주파수 성분을 가지고 있는 3개의 부영상들은 영상의 수직, 수평, 대각선 고주파 성분으로 구성된다. 이상의 3개의 부영상 블록들의 움직임 벡터는 바로 직주파 부영상에서 구한 움직임 벡터의 근처에서 움직임 정보를 정제한다. k 계층에서 구한 움직임 벡터를 다음의 하위 계층인 $k-1$ 계층의 초기값으로 사용된다.

단계 4:

$k-1$ 계층의 부영상들은 k 계층의 부영상 보다 해상도가 2배 더 크기 때문에 그 영상들은 $2^{-(k-1)}M_b \times 2^{-(k-1)}N_b$ 크기의 블록으로 분할한다. $k-1$ 이하의 부영상에 대해서도 이와 같은 방식으로 부영상들을 여러 개의 블록으로 나눈다. 결과적으로, 모든 계층의 부영상의 블록의 갯수는 동일하게 된다. 나누어진 블록들에서의 움직임 벡터는 위 계층에서 전달된 움직임 벡터를 2배 한 후, 이 벡터 공간의 근처에서 구한다. 이 움직임 벡터는 다시 다음 하위 계층으로 전달되고, 그 계층에서는 그 이전 계층에서의 블록 크기의 2배로 영상을 여러 개의 블록으로 분할하고, 이전 계층에서 전달된 움직임 벡터를 2배 한 후, 그 근처에서 움직임 정보를 찾는다. 이러한 과정은 최하위 계층에 도착될 때까지 하위 계층으로 가면서 반복적으로 실행하게 된다.



(그림 1) 일반적인 HBMA의 블록도

상위 계층에서 추정된 움직임 벡터로부터는 전역적인 움직임에 대한 특성을 알 수 있으며, 하위 계층으로 갈수록 국부적인 움직임에 대한 상세한 움직임 벡터를 얻을 수 있다. 이와 같은 움직임 추정 기법상의 특성에 의해 큰 움직임에 대한 벡터 추정 성능이 우수하며, 계산 시간의 감소 측면에서 큰 장점을 갖는다. 하지만, 가장 상위 계층에서 추정한 움직임 벡터가 저해상도와 다단의 계층을 구성하기 위해 여러 번의 지역 통과 필터링에 의한 영상의 뭉롱화(blurring)와 탐색 영역을 벗어나는 큰 움직임에 기인하여 국부 최소(local minima)에 빠질 경우, 그 영향이 하위 계층에 전달되고 하위 계층에서는 그 영향을 상쇄시킬 수 없기 때문에 움직임 추정 성능에 심각한 문제를 초래할 확률이 높다. 또한, 모든 계층에서의 모든 부영상에서 구한 움직임 정보가 수신단에 전송됨으로써 전송 부담이 증가하고 다양한 부가 정보를 수반해야 하는 문제가 있다.

3. 제안된 계층적 가변 블록 정합 알고리즘

3.1 개요

제안된 움직임 추정 방식에서는 우선, 연속된 두 프레임 간의 차영상 윤곽정보에서 정의한 블록 활동도를 평균하여 현재 영상의 평균 블록 활동도를 산출한다. 다음으로, 이렇게 산출한 평균 블록 활동도를 통해 카메라 패닝의 발생 여부를 검출한 후, 웨이블릿 변환에 의해 구성된 피라미드 계층 구조상에서 계층적인 탐색 방법에 기초하여 카메라의 상하 좌우 병진 운동인 카메라 패닝 벡터를 추정하여 보상한다. 이후, 카메라 패닝 보상 후에 정의한 블록 활동도를 토대로 각 블록을 움직임 블록, 준 움직임 블록, 비 움직임 블록 중 어느 하나로 분류한 검색 테이블(LUT: Look Up Table)을 작성한 후, LUT에 따라 추정 오차가 크게 발생할 것으로 기대되는 블록에 대해서는 초기 탐색 계층을 걸

정할 시에 상대적으로 해상도가 높은 계층을 선택함과 동시에 넓은 탐색 영역을 선택한다. 이렇게 선택된 초기 탐색 계층 영상으로부터 출발하여 가장 해상도가 높은 원영상쪽으로 이동하면서 움직임 벡터의 추정 정밀도를 높여감으로써 국부 최소에 빠질 확률을 크게 줄일 수 있도록 적응성을 부여하고 이 블록은 복잡한 움직임 정보를 갖고 있을 확률이 매우 높기 때문에 추정 정밀도를 향상시키기 위해 쿼드트리 적용하여 추정 블록의 크기를 작게 함으로써 화면의 국부적인 움직임을 효과적으로 반영한다. 또한, 움직임 추정 시간에 대한 부담을 줄이기 위해 LUT상에 기설정된 임계값보다 작은 블록 활동도를 갖는 것으로 분류된 블록에 대해서는 움직임이 없는 것으로 간주하여 움직임 추정을 생략한 상태에서 영벡터(zero vector)를 할당함으로써 국부 최소와 연산량 그리고 전송 부담에 대한 고려를 동시에 하면서 보다 적절한 시간 안에 움직임 벡터값을 고속으로 추정할 수 있다.

3.2 웨이블릿 변환

웨이블릿 이론은 다중 해상도와 다중 주파수 측면에서 신호를 고찰·분석하고자 하는 생각으로 수학 및 공학 분야에서 발전되어 오늘에 이른 이론으로, 이러한 분석적 특성을 영상 부호화에 효과적으로 반영할 수 있는 직교 변환인 것으로 널리 평가받고 있다. 웨이블릿은 식 (1)의 $\psi(x)$ 로 정의된 모 웨이블릿(mother wavelet)을 변이(translation)시키고 확대/축소(dilation) 시킴으로써 얻어지는 함수의 집합을 말하는 것이다.

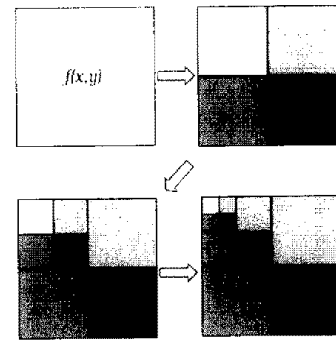
$$\psi_{a,b}(x) = a^{-1/2} \psi\left(\frac{x-b}{a}\right) \quad (1)$$

$$a \in R^+, b \in R$$

$\psi_{a,b}(x)$ 은 제곱적분가능(square integrable)한 공간 $L^2(R)$ 의 정규직교 기저를 형성한다.

통상, 영상에 2차원 웨이블릿 변환을 취하면, 웨이블릿 변환된 영상은 (그림 2)에 나타난 바와 같이, 다중 해상도와 다중 주파수 성분의 부영상들로 이루어진 다수의 계층들로 분해된다. 같은 계층에서 부영상들의 해상도는 동일하며 상위 계층으로 갈수록 해상도가 낮아진다. 이와 유사한 다중 해상도 표현 방법은 라플라시안 피라미드(Laplacian Pyramid)로 표현하는 것과 QMF 뱅크(Quadrature Mirror Filter Banks)를 이용한 대역 분할 부호화 방법이 있는데, 웨이블릿 변환 방식

은 이를 하나로 통합한 이론이라 할 수 있다[9][10]. 본고에서는 상대적으로 고속화를 기할 수 있는 하르 기저를 이용한 웨이블릿 변환을 수행하여 피라미드 계층 구조를 구성하였다.



(그림 2) 웨이블릿 변환에 의한 영상 분해

3.3 블록 활동도

연속되는 영상의 차영상은 식 (2)과 같이 표현할 수 있다.

$$D_{t,t-1}(i,j) = |I_t(i,j) - I_{t-1}(i,j)| \quad (2)$$

여기서, $I_{t-1}(i,j)$ 및 $I_t(i,j)$ 는 각각 이전 프레임의 화소값 및 현재 프레임의 화소값을 의미한다. 이렇게 얻어진 $D_{t,t-1}(i,j)$ 의 화소값은 이전 프레임과 현재 프레임간 차의 절대값으로 나타난다. 이 차영상을 임계치에 따라 이진화(binanzation)를 수행하여 이진화 영상을 획득한 후, 단위 블록에 포함된 '1'의 개수를 카운팅하여 이를 각 블록의 블록 활동도를 산출하기 위한 척도로 이용한다. 이때, 시간축에 대해 조도 및 이미지 센서의 미소 불균일에 의해 발생된 차분값이나 무시할 수 있을 정도의 미소한 객체(object)의 움직임 또는 손떨림 등에 의한 극소한 카메라의 움직임에 의해 발생된 차분값 등이 포함된 불필요한 차분 검출 영역을 제거하기 위해 윤곽 처리와 임계 처리를 수행한다. 이것은 컴퍼스 경사 마스크(compass gradient masks)를 이용하여 국부 8방향 윤곽 검출 처리를 차영상에 적용하여 두 영상 사이에 윤곽 정보가 있는 부분을 이진화하는 것으로, 본고에서는 Prewitt 연산자의 제2형(Prewitt operators type 2)[11]을 이용하여 각각의 주목 화소를 중심으로 주변의 8개 화소에 대한 윤곽의 방향성을 고려하여 최대 윤곽치를 검출한다. 이러한 과정을 통해 차영상에서 나타난 윤곽 부분을 이진화한

후, 각 블록에 대한 블록 활동도를 정량적으로 산출하기 위해 $M_b \times N_b$ 크기의 블록으로 분할하여 식 (3)과 같이 해당 블록에 포함된 1의 갯수를 카운팅한다. 이렇게 카운팅한 값을 각 블록에 대한 움직임의 복잡도를 대변하는 블록 활동도로 정의하여 윤곽 정보의 포함 정도를 나타내는 척도로 사용한다.

$$A(n) = \sum_{i=1}^{M_b} \sum_{j=1}^{N_b} B(i, j) \quad (3)$$

이후, 검색 테이블(LUT) 작성할 시에 사용되는 임계치(TH)를 산출하기 위해 식 (4)에서와 같이 모든 블록의 블록 활동도를 합산하여 평균함으로써 해당 프레임의 평균 블록 활동도 $E[A(n)]$ 을 산출한다. 여기서, M, N 은 각각 유효 수평 화소수 및 유효 수직 화소수이고, $M/M_b \times N/N_b$ 는 총 유효 블록수이다.

$$E[A(n)] = \frac{1}{M/M_b \times N/N_b} \sum_{n=1}^{M_b \times N_b} A(n) \quad (4)$$

비례 상수(μ)를 평균 블록 활동도 $E[A(n)]$ 에 승산하여 설정한 임계치(TH)와 각 블록의 블록 활동도 $A(n)$ 을 비교하여 조건식인 식 (5)에 따라 다음의 3가지 경우에 대한 LUT를 작성한 후, 각 경우에 대해서 움직임 추정을 달리해 줌으로써 화면 내의 전체 움직임과 국부적 움직임 모두를 고려해 줄 수 있다.

$$\begin{aligned} TH &= \mu \times E[A(n)] \\ A(n) > TH &: \text{움직임 블록} \\ A(n) = 0 &: \text{비 움직임 블록} \\ 0 < A(n) < TH &: \text{준 움직임 블록} \end{aligned} \quad (5)$$

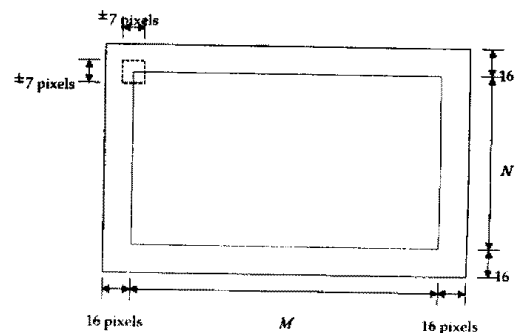
3.4 카메라 패닝 보상

연속된 실제 영상에서 카메라 패닝은 시현 영상에 따라 다르겠지만, 확률적으로 그 발생 빈도수가 낮기 때문에 카메라 패닝을 항상 고려하여 보상하는 것은 과도한 연산량을 수반하게 됨에 따라 전체적으로 움직임 추정 시간을 지연시키는 비효율적 요인이 된다. 따라서 적은 연산량을 필요로 하면서도 카메라 패닝의 발생 여부를 효과적으로 감지할 수 있는 판정 기준이 필요하다. 이를 위해, 제안된 움직임 추정 방식에서는 최하위 계층 영상에서 정의한 평균 블록 활동도 $E[A(n)]$ 가 실험적으로 정한 임계값 보다 클 경우에

일정한 카메라 패닝이 발생했 것으로 판단하고 카메라 패닝 보상을 위한 연산을 수행한다.

제안된 알고리즘은 웨이블릿 변환을 이용하여 계층의 구성을 k 단계로 분할하고, 이때 분해되는 과정에서 얻어지는 각 계층에서의 저주파 성분의 부영상들을 저장한다. 이 저주파 성분의 부영상들만을 이용하여 고속 카메라 패닝 보상 및 계층적 움직임 추정을 수행하기 위한 계층 공간을 구성한다.

제안된 방식에서는 웨이블릿 변환을 수행함에 따라 발생하는 각 계층의 저주파 부영상만을 이용하여 3단계의 계층으로 구성된 피라미드 계층 공간상에서 상위 계층에서 하위 계층 방향으로 진행하면서 점진적으로 카메라 패닝 벡터의 추정 정확도를 재조성하여 최종적으로 카메라 패닝 벡터를 결정한다. 명명의 편의상, k 계층 저주파 영상, $k-1$ 계층 저주파 영상 및 원영상을 각각 최상위 계층 영상, 중간 계층 영상 및 최하위 계층 영상으로 명하기로 한다. 카메라 패닝 벡터를 추정하기 위한 이후의 과정에서는 (그림 3)과 같이, 먼저, 탐색 공간에 대한 여유를 확보하기 위해 카메라 패닝 추정 블록을 원영상(즉, 최하위 계층 영상)에서 좌우 상하 ± 16 pixels를 들여서 정한 후, 이 블록을 가지고 피라미드 계층 공간의 최상위 계층 영상에서는 ± 1 pixel, 중간 계층 영상에서도 ± 1 pixel, 그리고 마지막 막으로, 최하위 계층 영상에서도 ± 1 pixel의 탐색 영역을 고려하여 카메라 패닝 벡터를 계층적으로 추정한다.

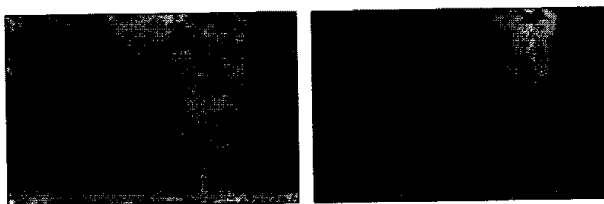


(그림 3) 카메라 패닝 보상의 탐색 영역

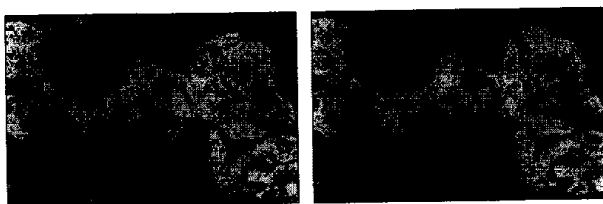
이에 따르면, 원영상을 기본 탐색 공간으로 가정할 때, 카메라 패닝 벡터의 탐색 영역을 ± 7 pixels로 설정한 것과 동일한 탐색 범위를 갖는데, 계층적인 추정 방식임에 따라 추정 속도를 고속화할 수 있으며 탐색 단위를 적절하게 표본화한 소수의 작은 크기의 블록들로 설정할 경우, 좀 더 빠른 고속화를 실현할 수 있다.

여기서, 탐색 영역을 $\pm 7\text{pixels}$ 로 고려한 것은 통상의 카메라 패닝은 한 프레임 사이에서 7개 화소 이상 이동하는 경우의 발생 빈도수가 매우 낮음을 실험적으로 확인한 바에 따른 것이며, 실사, $\pm 7\text{pixels}$ 이상의 카메라 패닝이 발생한다고 하더라도 이 경우에는 화면에 극도의 뭉뚱화(blurring) 현상이 발생됨에 따라 카메라 패닝을 보상하는 행위 자체가 무의미하기 때문이다.

제안된 알고리즘은 움직임 벡터 부호화 시, 카메라 패닝을 보상함에 따라 패닝이 보상된 영상의 배경(background)은 LUT상에 비 움직임 블록(Non-Moving Block)으로 분류되므로 영벡터(zero vector)를 많이 늘릴 수 있다. 따라서 움직임 탐색의 계산 시간을 감소시키고 전송 부담을 절감할 수 있는 결과를 기대할 수 있다. 예컨대, (그림 4)는 카메라 패닝 보상 전후의 블록 활동도의 변화를 시각적으로 표현하기 위해 움직임 블록(Moving Block)을 쿼드트리(quadtrees)하여 예시한 것이다. (그림 4)를 통해 살펴보면, 카메라 패닝 보상을 수행하기 전과 후에 있어서, 움직임 블록(Moving Block)으로 분류되는 블록수가 전체적으로 감소하고 있음을 알 수 있는데, 카메라 패닝 벡터가 상대적으로 큰 (a)의 BIKE 영상이 패닝 벡터가 상대적으로 작은 (b)의 FOOTBALL 영상에 비해 그 감소폭이 현저함을 쉽게 확인할 수 있다. 이에 따르면, 움직임 추정의 성능은 화질 측면에서 소폭 감소하거나 대등하게 유지하면서도 전체적인 탐색 시간은 큰 폭으로 감소시킬 수 있는 이점이 있다. 그리고, 카메라 패닝을 보상하게 되면,



① 카메라 패닝 보상 전 ② 카메라 패닝 보상 후
(a) BIKE(#52-#53)



① 카메라 패닝 보상 전 ② 카메라 패닝 보상 후
(b) FOOTBALL(#10 #11)

(그림 4) 카메라 패닝 보상 전후의 블록 활동도

초기 움직임 탐색점의 위치가 카메라 패닝 벡터만큼 보상되기 때문에 이렇게 보상된 탐색점을 기준으로 새로운 탐색 영역을 설정하게 되는 데, 이것은 좀 더 합리적인 탐색 공간에서 움직임 벡터를 추정할 수 있음을 의미하는 것이다. 통상, 카메라의 운동은 움직임 객체와 같은 방향으로 이루어지기 때문에 움직임 객체의 동적 범위가 클 경우에는 더욱 더 효과적이다. 이를 통해 결과적으로 움직임 벡터 추정의 최적화를 도모할 수 있으며, 움직임 벡터의 동적 가변 범위를 작게 할 수 있다.

3.5 초기 탐색 계층과 각 계층의 탐색 영역 선택 및 가변 블록 분할

일반적인 HBMA 기법에서는 각 계층의 모든 부영상에서 움직임 벡터를 구하기 때문에 전송 움직임 벡터 수가 너무 많고 다양하다는 결점이 있으며, 움직임 정보를 고주파 성분이 있는 부영상에 대해서도 추정함에 따라 고주파에 해당하는 잡음에 민감한 문제가 있음에 따라 이에 제안된 방식에서는 각 계층의 저주파 성분의 부영상만을 이용하여 3단계의 계층으로 구성된 파라미드 계층 공간상의 각각의 계층에서 추정한 움직임 벡터를 최하위 계층에서 최종적으로 재조정하여 구한 움직임 벡터만을 부호화하여 복호단에 제공한다. 움직임 추정을 위한 기본 단위는 최하위 계층 영상을 기준할 때, 현재 여러 표준 코덱들이 사용하는 $16\text{pixels} \times 16\text{pixels}$ 매크로 블록으로 정하였으며, 이에 따라 중간 계층 영상에서의 추정 블록의 크기는 $8\text{pixels} \times 8\text{pixels}$, 최상위 계층 영상에서는 $4\text{pixels} \times 4\text{pixels}$ 로, 계층이 증가함에 따라 추정 블록의 크기는 옥타브(octave)적으로 감소된다. 이러한 사실에 기초하여 제안된 방식에서 초기 탐색 계층의 선택과 선택한 계층에서의 탐색 영역을 설정하는 과정을 구체적으로 살펴보면 다음과 같다.

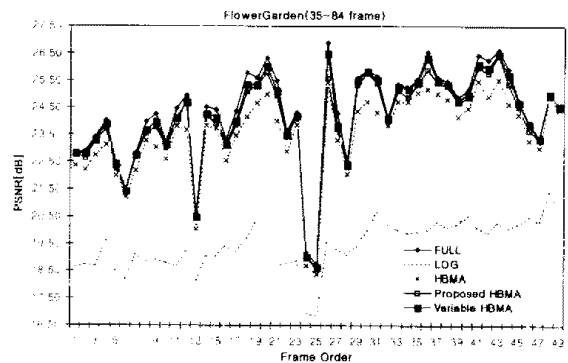
우선, LUT상에서 비 움직임 블록으로 분류된 블록은 움직임 벡터가 0일 것으로 가정하고, 움직임 탐색을 수행하지 않는 상태에서 영벡터를 할당함에 따라 불필요한 움직임 추정 과정을 생략함으로써 움직임 추정 속도의 고속화를 도모하는 한편, 준 움직임 블록으로 분류된 블록은 초기 탐색 계층을 최상위 계층으로 선택하고, 초기 탐색 계층으로 선택한 최상위 계층에서 $\pm 3\text{pixels}$, 중간 계층에서 $\pm 2\text{pixels}$, 최하위 계층에서 $\pm 1\text{pixel}$ 의 순으로 탐색 영역을 설정하여 움직임 벡터의 추정 정밀도를 재조정한다. 마지막으로, LUT상에

서 움직임 블록으로 분류된 블록은 국부 최소에 빠진 확률을 줄이기 위해 초기 탐색 계층을 중간 계층 영상으로 선택하고, 이 계층에서 ± 7 pixels, 다음으로 최하위 계층에서 ± 4 pixels의 탐색 영역을 설정하여 움직임 벡터를 검출한다. 이때, 이 블록은 복잡한 움직임 정보를 갖고 있을 확률이 매우 높기 때문에 쿼드트리 적용하여 추정 블록의 크기를 작게 함으로써 화면 내에 존재하는 국부적인 움직임을 적용적으로 반영할 수 있도록 한다. 즉, 최하위 계층 영상을 기준할 때, $16 \text{pixels} \times 16 \text{pixels}$ 블록을 $8 \text{pixels} \times 8 \text{pixels}$ 블록 크기로 쿼드트리하고, 중간 계층 영상에서는 $4 \text{pixels} \times 4 \text{pixels}$ 의 크기로 쿼드트리하여 움직임을 추정한다. 이때, 계층들 간의 움직임 추정은 일반적인 계층적 블록 정합 알고리즘과 마찬가지로, 상위 계층에서 추정한 움직임 벡터의 2배를 다음 계층의 움직임 추정 초기값으로 설정한다. 이러한 과정을 최하위 계층까지 반복하여 최종적으로 부호화할 움직임 벡터를 추정하게 된다. 이에 따르면, 원영상을 기본 탐색 공간으로 가정할 때, 움직임 블록은 한 프레임 안에서 최대 $\pm 18 \text{pixels}$, 준 움직임 블록은 최대 $\pm 17 \text{pixels}$ 의 탐색 영역을 가질 수 있음에 따라 $\pm 16 \text{pixels}$ 를 초과하는 움직임 벡터가 발생할 경우에는 그 값을 ± 16 또는 ± 16 으로 클램핑한다. 평가 함수로는 MSE보다 연산량이 상대적으로 적게 요구되는 MAE를 이용한다. 전술한 바와 같이, 각 경우에 대한 움직임 추정을 달리해 줌으로써 화면 내의 전체적인 움직임과 국부적인 움직임 모두를 고려해 줄 수 있기 때문에 국부 최소에 빠질 가능성을 현저하게 억제시킬 수 있다.

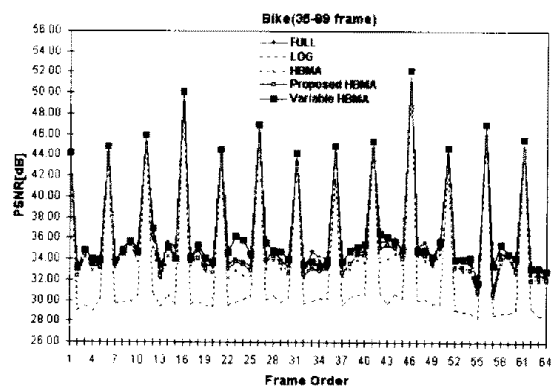
4. 시뮬레이션 결과 및 고찰

본 연구에서는 제안된 방식의 타당성과 보편성을 평가하기 위해, 서로 다른 동적 특성을 갖는 복수의 동영상에 대상으로 기존의 방식과 제안된 방식에 대한 컴퓨터 시뮬레이션을 수행하였다. 시험 영상 시퀀스로는 256 계조 (256 gray level)의 $352 \text{pixels} \times 240 \text{pixels}$ 크기를 갖는 카메라 패닝이 포함된 영상 시퀀스인 FLOWERGARDEN (#35~#84), BIKE(#35~#99), FOOTBALL(#2~#25)과 카메라 패닝이 없는 영상 시퀀스인 SALESMAN(#101~#130), TABLETENNIS1(#0~#30) 및 카메라 줌밍이 포함된 영상 시퀀스인 TABLETENNIS2(#90~#141) 등의

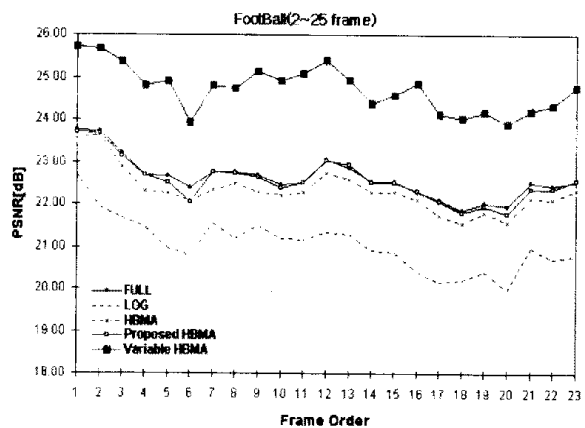
중 252 프레임의 휘도 영상을 이용하였으며, 시뮬레이션 결과에 대한 객관적인 성능 비교 결과를 제공하기 위해 판단 기준으로 PSNR(Peak Signal to Noise Ratio)를 채택하였다. 그림 5는 다양한 영상 시퀀스에 대한 각 방식의 PSNR 비교를 통한 성능 평가 결과를 보여주고 있다.



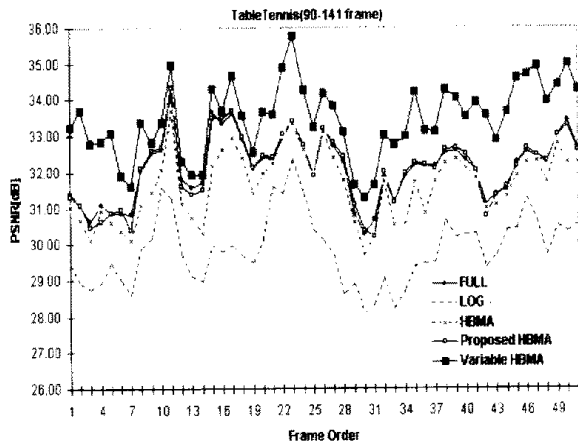
(a) FLOWERGARDEN(#35~#84)



(b) BIKE(#35~#99)



(c) FOOTBALL(#2~#25)



(d) TABLETENNIS2(#90~#141)

(그림 5) 영상 시퀀스들에 대한 각 방식의 PSNR 비교

평가 대상으로 삼은 각 방식들의 블록 크기(block size), 탐색 영역(search area), 정합 기준(matching criteria)에 대해 간략하게 살펴보면, 우선, 전역 탐색(Full Search; 'FULL') 및 2-D 대수적 탐색(2-D Logarithmic Search; 'LOG') 방식의 블록의 크기는 16pixels×16pixels, 탐색 영역은 ±16pixels이다. 다음으로, 기존의 계층적 블록 정합 알고리즘을 대변하기 위한 'HBMA'는 제안된 HBMA('Proposed HBMA')와의 공정한 성능 비교를 위해 다소 변형된 방식을 취하였다. 이를 좀 더 상세하면, 'HBMA'는 저주파 부영상만을 이용해 구성된 피라미드 계층 구조상에서 블록 활동도를 고려하지 않고, 제안된 HBMA('Proposed HBMA')에서 준 움직임 블록(Semi-Moving Block)으로 분류된 블록들에서처럼, 초기 탐색 계층을 최상위 계층으로 선택하고, 초기 탐색 계층으로 선택한 최상위 계층에서 ±3pixels, 중간 계층에서 ±2pixels, 최하위 계층에서 ±1pixel의 순으로 탐색 영역을 설정하여 높은 해상도로 감에 따라 초기에 추정된 움직임 벡터 값을 점진적으로 재조정한다. 이때, 고주파 부영상에서의 움직임 벡터의 재조정을 생략함으로써 제안된 HBMA('Proposed HBMA')와의 성능 비교 시에 공정성을 도모한다. 시뮬레이션 결과에 따르면, 기존의 HBMA를 일부 대변할 수 있는 이 변형된 HBMA('HBMA')를 비교 대상으로 삼음으로써 초기 탐색 계층의 선정 방식이 고정될 경우의 최상위 계층 또는 중간 계층 영상에서 국부 최소에 빠지는 정도를 간접적으로나마 확인할 수 있다.

(그림 5)의 그래프에서 'Proposed HBMA'라고 범

례(legend)가 붙은 곡선은 제안된 알고리즘에서 블록의 크기를 가변 하지 않고 16pixels×16pixels 크기로 블록을 고정시킨 HBMA에 대한 성능 평가이며, 'Variable HBMA'고 범례가 붙은 곡선은 제안된 알고리즘에 따른 가변 HBMA에 대한 성능 평가를 나타낸 것이다. 최하위 계층을 기준으로 할 때, 블록의 크기는 16pixels×16pixels, 탐색 영역은 준 움직임 블록에서 ±17pixels, 움직임 블록에서는 ±18pixels이며, 정합 기준은 MAE(Mean Absolute Error)이다.

<표 1> 각 방식의 평균 PSNR 비교[dB]

방식 시험영상	FULL	LOG	HBMA	Proposed HBMA	Variable HBMA
FLOWER-GARDEN	24.35	19.46	23.56	24.05	24.18
BIKE	33.68	28.88	32.86	33.02	34.16
FOOTBALL	22.62	21.04	22.32	22.56	24.73
SALESMAN	33.02	30.30	32.94	32.89	33.89
TABLE-TENNIS1	29.40	27.51	29.01	29.23	30.03
TABLE-TENNIS2	32.11	29.85	31.63	32.10	33.46
전체 평균	29.86	26.47	29.31	29.59	30.61

<표 1>에서 살펴볼 수 있는 바와 같이, PSNR 측면에서, 제안된 HBMA('Proposed HBMA')는 2-D 대수적 탐색('LOG'), 변형된 기존의 HBMA('HBMA')에 비해 각각 약 3.12dB 정도, 약 0.28dB정도만큼 평균적으로 우수하며, 전역 탐색('FULL')에 약 0.27dB정도만큼 떨어진 성능을 제공하고 있음을 컴퓨터 시뮬레이션 결과를 통해 확인할 수 있었다. 또한, 제안된 가변 HBMA('Variable HBMA')는 PSNR 측면에서 2-D 대수적 탐색('LOG'), 변형된 기존의 HBMA('HBMA')에 비해 각각 약 4.14dB정도, 약 1.30dB정도만큼 평균적으로 우수하고, 예측할 수 있는 바와 같이, 가변 블록 방식임에 따라 전역 탐색 비해서는 약 0.75dB정도 평균적인 성능 향상을 도모할 수 있었다.

제안된 HBMA('Proposed HBMA')와 제안된 가변 HBMA('Variable HBMA')는 움직임의 변위가 크고, 다양한 이동 객체(moving object)의 운동에 의해 움직임이 복잡하며, 이동 객체의 크기가 비교적 클수록 각각의 최적의 한계 성능인 16pixels×16pixels와 8pixels×8pixels 블록의 전역 탐색의 성능에 접근해 감을 확인할 수 있었다. 한편, 카메라 패닝을 보사한 후와 보

상 전에 있어서, PSNR은 거의 동일한 수준을 유지하면서 <표 2>에서 살펴볼 수 있듯이, 카메라 패닝을 보상하기 전에는 움직임 블록으로 분류된 다수의 블록들이 보상 후에는 준 움직임 블록과 비 움직임 블록으로 분류됨을 확인할 수 있는데, 이는 곧 움직임 탐색 시간을 현저하게 줄일 수 있거나 탐색 연산을 생략한 상태에서 영백터를 할당할 수 있는 블록이 증가했음을 의미하는 것이다.

<표 2> 카메라 패닝 보상 전과 보상 후의 블록의 분류 비교

블록 분류 영상	비 움직임 블록 (%)		준 움직임 블록 (%)		움직임 블록 (%)	
	보상후	보상전	보상후	보상전	보상후	보상전
FLOWER-GARDEN	8.3	6.2	38.2	29.5	53.6	64.3
BIKE	68.2	45.9	14.2	19.7	17.6	34.4
FOOTBALL	9.1	7.7	40.6	41.9	50.3	50.4

예컨대, 저주파 성분이 많은 배경을 포함하면서 일정 프레임 간격마다 카메라 패닝을 포함하는 BIKE (#35~#99) 영상 시퀀스의 경우, 카메라 패닝을 보상한 후에는 연산 시간을 약 0.5배 정도로 경감시킬 수 있었는데, 카메라 패닝의 발생 빈도수와 카메라 패닝 백터의 크기가 상대적으로 작은 FOOTBALL(#2~#25) 영상 시퀀스에 비해 평균 연산 시간의 상대적 감소폭이 현저함을 알 수 있다. 반면에 고주파 성분이 많은 배경을 포함하면서 연속적인 카메라 패닝을 포함하는 FLOWERGARDEN (#35~#84) 영상 시퀀스의 경우, 카메라 패닝을 보상한 후에는 연산 시간을 약 0.86배 정도로 소폭 경감시킬 수 있었으나, 그림 5의 (a)에서 살펴볼 수 있듯이, 여타의 영상 시퀀스에 비해 저급한 추정 성능을 제공함을 확인할 수 있다. 이러한 현상은 제안된 알고리즘을 적용함에 있어서, 고주파 성분을 많이 포함하면서 카메라 패닝을 포함하는 영상 시퀀스들로부터 공통적으로 나타나는 특징으로, 카메라 패닝 보상에 따라 추정 속도의 고속화를 도모할 수 있는 비 움직임 블록과 준 움직임 블록의 수가 증가함에 그 원인이 있다. 특히, 제안된 알고리즘은 저주파 부영상을 통해 피라미드 계층 공간을 구성하고 있기 때문에 임의의 블록에 대한 초기 탐색 계층이 최상위 계층이나 중간 계층으로 설정될 경우, 고주파 성분을 충분히 고려해 줄 수 없음을 기인하여 고주파 성분이 많은 영상 시퀀스에 포함된 블록들은 극부 최소에 빠지는 확률이

상대적으로 증가하여, 연산을 증가하고 있다.

하나의 추정 블록에 대한 전역 탐색('FULL')의 연산량을 ζ 라고 할 때, 2-D 대수적 탐색('LOG')의 연산량은 약 $1/30\zeta$ 이다. 이와 마찬가지로, 제안된 HBMA ('Proposed HBMA')의 경우, 준 움직임 블록 및 움직임 블록은 각각 약 $1/60\zeta$ 및 $1/8\zeta$ 의 연산량이 소요되고, 비 움직임 블록은 별도의 추가 연산이 필요 없음에 따라 이러한 사실에 기초하여 표 2와 같은 블록 분류 결과를 이용하고 더불어 각 평가 기준을 산출하기 위한 연산량을 포함시키면 각 시퀀스에 대한 평균 연산량을 산출할 수 있다. 시뮬레이션에 사용된 시험 영상 시퀀스를 기준할 때에 평균적으로 소요된 연산 시간을 비교해 보면, 제안된 HBMA('Proposed HBMA') 및 가변 HBMA ('Variable HBMA')는 2-D 대수적 탐색('LOG')과는 비슷한 연산 시간이 소요됐으며, 변형된 기존의 HBMA('HBMA')에 비해 각각 약 2배, 전역 탐색('FULL')에 비해서는 약 1/25배의 연산 시간이 소요됨을 컴퓨터 시뮬레이션을 통해 확인할 수 있었다. 여기서, 연산 시간을 비교한 이유는 제안된 방식은 일종의 적응형 알고리즘이기 때문에 정확한 연산량을 산출하기에 다소 어려움이 있음에 따라 간접적인 방법이지만 연산 시간은 연산량을 적절하게 반영할 수 있는 것임에 근거한다.

5. 결 론

본고에서는 움직임 추정 시간을 단축시키기 위한 고속 움직임 탐색 알고리즘을 제안함과 동시에 카메라 패닝을 보상함으로써 움직임 추정 효율을 개선시킨 카메라 패닝 보상 기반 계층적 가변 블록 정합 알고리즘을 제안하고 있다. 제안된 알고리즘은 피라미드 계층 구조를 구성하기 위해 웨이블릿 변환을 취해야 하고 차영상 윤곽정보를 이용하여 LUT를 작성하기 위한 추가적인 연산이 필요하지만, 기존의 알고리즘과 비교할 때, 추정 성능을 개선하면서도 전체적인 연산량은 현저하게 단축시킬 수 있었다. 또한, 카메라 패닝 발생 유무의 검출, 초기 탐색 계층과 각 계층에 대한 탐색 영역의 선택, 추정 블록의 크기를 가변하기 위한 쿼드트리 여부 등을 판단할 시에 이를 결정하기 위한 각각의 평가 함수로서, 차영상 윤곽정보를 통해 획득한 블록 활동도를 공통적으로 이용함에 따라 평가 함수를 정의하는데 소요되는 연산량과 하드웨어적인 복잡도를 간략화할 수 있는 장

잡이 있으며, 그 평가 함수의 우수성과 유용성을 컴퓨터 시뮬레이션을 통해 확인할 수 있었다.

그러나, 제안된 알고리즘은 적응적인 블록 정합 알고리즘의 일종이기에 움직임 추정을 위한 연산 시간이 입력 영상에 의존적이고, 특히, 제안된 가변 HBMA는 블록 분할을 위한 부가정보를 복호단에 전송해야 하는 부담이 있다. 그리고, 하드웨어적인 구현을 위해서는 기존의 블록 기반의 영상 코덱에 각 계층의 지주파 부영상으로 구성된 피라미드 계층 구조를 저장하기 위한 추가적인 저장 공간이 필요하고, 더불어 연속된 두 프레임 간의 차영상을 저장하기 위한 별도의 프레임 메모리가 필요하다.

향후, 고주파 성분이 많은 영상 시퀀스에 적절하게 대응할 수 있는 견고한 알고리즘의 개발과 차영상 윤곽 정보로부터 영상의 블록 활동도를 정의할 때에 적용되는 임계값들을 최적화하기 위한 추가적인 연구가 필요하며, 고속의 카메라 패닝 및 줌링(zooming) 보상 방식에 대한 연구가 차후로 진행되는 것이 바람직하다.

참 고 문 헌

[1] Borke Furht, Joshua Greenberg, Raymond Westwater, 'Motion Estimation Algorithms for Video Compression,' Kluwer Academic Publishers, 1997.

[2] Guido M. Schuster, Aggelos K. Katsaggelos, 'Rate-Distortion Based Video Compression,' Kluwer Academic Publishers, 1997.

[3] K. R. Rao, J. J. Hwang, 'Techniques and Standards for Image, Video, and Audio Coding,' Prentice Hall, 1996.

[4] ISO/IEC29/WG11, "Coding of Moving Pictures and Associated Audio Information," MPEG4 Verification Model Version 9.1, Feb. 1998.

[5] Luis Torres and Murat Kunt, 'Video Coding: The Second Generation Approach', Kluwer Academic Publishers, 1996.

[6] S. Panchanathan, E. Chan and X. Wang, "Fast Multiresolution Motion Estimation Scheme for a Wavelet Transform Video Coder," SPIE, Vol. 2308, pp.674-681, 1994.

[7] 김성만, 이규원, 정학진, 박규태, "웨이브렛 변환

영역에서의 2단계 가변 블록 다해상도 움직임 추정", 한국통신학회 논문지, Vol.22, No.7, pp.1487-1504, 1997.

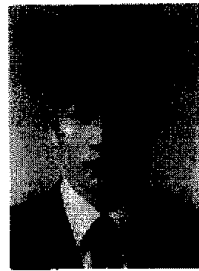
[8] M. Hoetter, "Differential Estimation of the Global Parameters Zoom and Pan," Signal Processing pp.249-265. 1989.

[9] Ingrid Daubechies, *Ten Lectures on Wavelets*, Society for Industrial and Applied Mathematics, Society for Industrial and Applied Mathematics, 1992.

[10] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image Coding Using Wavelet Transform" IEEE Trans on Image processing, Vol.1, No.2, pp205-220, Apr. 1992.

[11] Maher A. Sid-Ahmed, "Image Processing: Theory, Algorithm, and Architecture," McGraw-Hill, 1995.

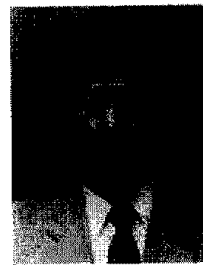
곽 노 윤



e-mail : nykwak@mail.hankong.ac.kr
 1994년 한국항공대학교 항공전자공학과 졸업(공학사)
 1996년 한국항공대학교 대학원 항공전자공학과(공학석사)
 1996년 3월~현재 한국항공대학교 대학원 항공전자공학과(박사과정)

관심분야 : 동영상 신호처리, 칼라 영상 부호화, 영상 디스플레이

황 병 원



e-mail : bhwang@mail.hankong.ac.kr
 1972년 한국항공대학 항공전자공학과 졸업(공학사)
 1981년 동경대학 대학원 전기공학과(공학석사)
 1984년 동경대학 대학원 전기공학과(공학박사)

1984년 5월~1985년 3월 국방과학연구소 선임연구원
 1993년 8월~1994년 8월 동경대학 전기공학과 교환교수
 1985년 8월~현재 한국항공대학교 항공전자공학과 교수
 관심분야 : 영상처리 및 영상인식, 동영상 신호처리, 교통류 제어