論文2000-37SP-5-6

다 해상도 프레임 구조에 기반한 고속 움직임 추정 기법

(A Fast Motion Estimation Algorithm Based on Multi-Resolution Frame Structure)

宋 秉 哲 *, 羅 鍾 範 *

(Byung Cheol Song and Jong Beom Ra)

요 약

고속 움직임 추정을 위한 다 해상도 블록 정합 기법을 제안한다. 최저 해상도 계층에서 전역 탐색을 통해 최소 정합 오차를 갖는 움직임 벡터를 선택하고, 공간적으로 인접한 블록들의 움직임 벡터들 중에서 최소 정합 오차를 갖는 움직임 벡터를 찾는다. 이 때, 주변 움직임 벡터들의 보다 정확한 탐색을 위해 저 해상도 계층에서도 움직임 벡터의 양자화 없이 탐색을 할 수 있는 효과적인 방법을 제안한다. 이렇게 얻어진 2개의 움직임 벡터들은 중간 해상도 계층에서의 탐색을 위한 초기 탐색 중심점들로 사용된다. 중간 계층에서, 각 초기점을 중심으로 훨씬 좁아진 영역에서의 지역 탐색을 수행한다. 최저 해상도 계층에서 주변 움직임 벡터 탐색을 위해 사용했던 방법을 이용하면, 각 지역 탐색을 정수 화소 단위로 수행할 수 있다. 지역 탐색 영역 내에서 최소 정합 오차를 갖는 움직임 벡터를 찾고, 이 벡터를 중심으로 마지막 계층에서의 마지막 탐색을 수행한다. 그러나, 중간 해상도 계층에서 이미 정수 화소 단위의 정확한 움직임 벡터 추정을 수행했기 때문에, 마지막 최고 해상도 계층에서의 지역 탐색은 전체 성능에 미미한 영향을 주게 된다. 따라서 최고 해상도 계층에서의 탐색을 생략하더라도 성능 저하 없이 탐색 속도를 향상시킬 수 있다. 모의 실험을 통해 최고 계층에서의 지역 탐색을 생략하더라도 제안한 블록 정합 기법이 전역 탐색 기법에 비해 보편적인 MPEG2 부호화 환경 하에서 최대 0.2dB의 PSNR 저하만을 보이며, 200배 이상의 계산 속도를 가짐을 보인다. 또한, 제안한 기법은 규칙적인 데이터 흐름을 가지므로 하드웨어 구현에도 적합하다.

Abstract

We present a multi-resolution block matching algorithm (BMA) for fast motion estimation. At the coarsest level, a motion vector (MV) having minimum matching error is chosen via a full search, and a MV with minimum matching error is concurrently found among the MVs of the spatially adjacent blocks. Here, to examine the spatial MVs accurately, we propose an efficient method for searching full resolution MVs without MV quantization even at the coarsest level. The chosen two MVs are used as the initial search centers at the middle level. At the middle level, the local search is performed within much smaller search area around each search center. If the method used at the coarsest level is adopted here, the local searches can be done at integer-pel accuracy. A MV having minimum matching error is selected within the local search areas, and then the final level search is performed around this initial search center. Since the local searches are performed at integer-pel accuracy at the middle level, the local search at the finest level does not take an effect on the overall performance. So we can skip the final level search without performance degradation, thereby the search speed increases. Simulation results show that in comparison with full search BMA, the proposed BMA without the final level search achieves a speed-up factor over 200 with minor PSNR degradation of 0.2dB at most, under a normal MPEG2 coding environment. Furthermore, our scheme is also suitable for hardware implementation due to regular data-flow.

* 正會員, 韓國科學技術院 電子電算學科

(Dept. of Electrical Engineering and Computer Science, KAIST)

接受日字:1999年12月28日, 수정완료일:2000年6月20日

Ⅰ. 서 론

MPEG2나 H.261같은 동영상의 부호화에서 연속된 프레임들간의 시간적 중복성 (temporal redundancy)을 제거하기 위해 움직임 보상 예측 부호화 기법이 널리 이용되고 있다[1 2]. 블록 정합 기법 (block matching algorithm; BMA) 은 대표적인 움직임 보상 기법으로서, 이전 프레임의 탐색 영역 내에서 현재 프레임의 기준 블록과 가장 유사한 블록을 찾는 방식이다. 특히 전역 탐색 기법 (full search BMA; FSBMA) 은 탐색 영역 내의 모든 블록들을 탐색하기 때문에 최적의 방법이라고 할 수 있다. 그러나 FSBMA는 계산량이 방대하여 넓은 탐색 영역을 요구하는 MPEG2나 HDTV같은 영상 시스템에는 적합하지 않다. FSBMA의 방대한 계산량을 줄이기 위해 많은 고속 BMA들이 개발되었다[3-11].

고속 BMA들은 특성에 따라 다음과 같이 네 그룹으로 분류할 수 있다: 단조로운 오류 표면 모델에 기반한 LOGS (logarithmic search) 나 TSS (three step search) 같은 BMA 그룹[3 4], 화소 추출 (pixel sub-sampling) 을 이용한 BMA 그룹[5], 다 해상도 BMA 그룹[6-8], 그리고 시간적/공간적 상관성 (spatial-temporal correlation) 을 이용한 BMA 그룹[9-10]. 첫 번째 BMA그룹은 그 기본 모델이 실제 영상에 잘 맞지 않기 때문에 국소적 최소치 (local minimum) 에 빠지기 쉽다. 두 번째 그룹의 BMA들은 계산량 감소를 위해 과도하게 화소 추출을 할 경우 역시 국소적 최소치 문제가 생기므로 계산량 감소에 한계가 있다. 세 번째 그룹의 BMA들은 비교적 좋은 성능과 빠른 계산 속도를 갖는다. 그러나 대개의 경우 시간적/공간적인 상관성 같은 부가 정보를 이용하지 않고, 이전 계층의 움직임 정보만을 이용하여 현재 계층의 움직임을 추정하기 때문에 항상 믿을 만한 성능을 보이는 것은 아니다. 마지막 그룹의 BMA들은 예측 탐색에 기반한 기법이다. 이런 기법은 시간적/공간적 상관성을 이용함으로써 계산량을 줄일 뿐만 아니라 상관성이 강한 움직임 벡터 (motion vector; MV) 장을 갖는 영상들에 대해서는 상당히 믿을 만한 성능을 보인다. 그러나 이런 기법들은 기본적으로 움직임 벡터 장이 강한 상관성을 갖는다고 가정하기 때문에 가정이 맞지 않는 영상에서는 성능 저하가 심각해진다. 또한 불규칙적인 데이터 흐름 (data-flow) 으로 하드웨어 구현에 적합하지 않다.

위에서 언급한 기존 고속 BMA들의 단점을 극복하기 위해, 본 논문은 새로운 다 해상도 블록 정합 기법을 제안한다. 먼저 최저 해상도 계층에서 두 종류의 움직임 탐색을 수행한다. 최저 해상도로 줄어든 전체 탐색 영역 내에서 전역 탐색을 수행함과 동시에, 움직임 벡터장의 공간적 상관성을 이용하기 위해 공간적으로 인접한 블록들의 MV들을 후보로 한 탐색을 수행한다. 이때, 주변 MV들의 보다 정확한 탐색을 위해 저 해상도 계층에서도 MV의 양자화 (quantization) 없이 탐색을 할 수 있는 효과적인 정합 오차 (matching error) 예측 기법을 제안한다. 즉, 기준 블록으로부터 공간적으로 적당히 이동된 블록의 정합 오차를 기준 블록의 정합 오차대신 이용하는 방식으로, 실제 값에 가까운 정합 오차를 예측할 수 있다. 이를 위한 현재 프레임 구조로서 블록 평균 피라미드 프레임 구조를 도입한다. 해상도에 맞게 양자화된 MV의 정합 오차보다 제안한 방법에 의한 정합 오차가 실제 정합 오차에 훨씬 근접함을 실험을 통해 보인다. 최저 계층의 후보 MV들 중에서 최소 정합 오차를 갖는 2개의 MV들을 선택하여 다음 계층에서의 탐색을 위한 초기 탐색 중심점들로 사용한다. 중간 계층에서는, 각 초기점을 중심으로 훨씬 줄어든 영역에서의 지역 탐색을 수행한다. 각 지역 탐색은 제안한 정합 오차 예측 기법이 적용되어 정수 화소 단위로 수행된다. 따라서, 기존의 다 해상도 기법들처럼 많은 계산량을 요구하는 최고 계층 탐색을 반드시 수행할 필요가 없다. 모의 실험을 통해 최고 해상도 계층에서의 지역 탐색을 생략하더라도 제안한 블록 정합 기법이 FSBMA에 비해 보편적인 MPEG2 부호화 환경하에서 최대 0.2dB의 미미한 PSNR 저하만을 보이며, 200배 이상의 탐색 속도를 가짐을 보인다. 또한, 제안한 기법은 구조가 단순하고, 규칙적인 데이터 흐름을 가지기 때문에 하드웨어 구현에 적합하다.

본 논문은 다음과 같이 구성된다: 2절에서 새로운 다 해상도 BMA를 제안하며, 3절에서는 모의 실험 결과를 보인다. 마지막으로 4절에서 결론을 맺는다.

Ⅱ. 고속 다 해상도 블록 정합 기법

최근 몇 년 동안 움직임 추정을 비롯한 여러 응용

분야에 다 해상도 기법을 적용하는 연구가 활발히 진행되어 왔다. 일반적으로 다 해상도 BMA는 각 프레임을 여러 해상도의 계층들로 나누어 (그림 1), 저 해상도 계층에서는 해상도에 맞게 양자화된 MV들을 탐색하고, 고 해상도 계층에서는 이전 계층에서 얻어진 몇 개의 MV 후보들을 중심으로 주어진 영역 내에서 지역 탐색을 수행한다. 3 계층 구조의 경우, 계층 0에서는 4화소 간격으로 MV 들을 탐색하는 셈이며, 계층 1에서는 2화소 간격의MV 탐색만이 가능하다. 최고 계층인 계층 2에서만 1화소 간격의 MV들을 탐색할 수 있다. 따라서 정수 단위의 MV 추정을 위해서는 최고 계층인 계층 2에서의 탐색이 반드시 이루어져야 한다. 최근에 개발된 시간적/공간적 상관성을 함께 이용한 고속 다 해상도 BMA (이하 MRST)의 경우도[8], 비록 다른 기존의 고속 다 해상도 BMA들보다 좋은 성능을 보이기는 하지만, 저 해상도 계층에서의 움직임 추정 시 후보 MV들로 쓰이는 주변 MV들이 현재 계층에 맞게 양자화 된 후 대응하는 정합 오차를 계산하기 때문에 역시 추정의 정확성이 떨어지는 문제점이 있다.

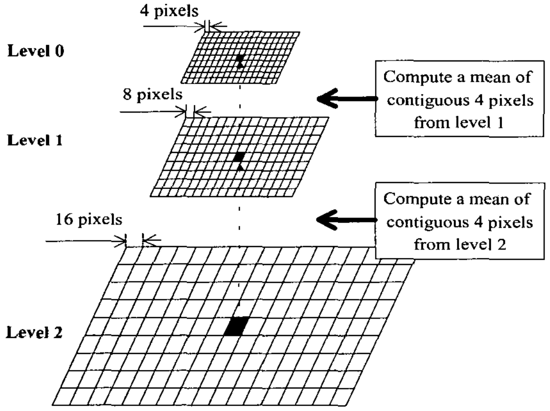


그림 1. 계층 수가 3인 경우의 이전 프레임의 다 해상도 구조. 각 사각형은 한 블록을 의미한다
Fig. 1. Multi-resolution structure of the reference frame with three levels. Each rectangle denotes a block.

만약 저 해상도 계층에서도 정수 단위의 MV를 탐색할 수 있다면, 지역 탐색에서 찾아진 MV의 정확도가 증가하여 전체 성능이 향상되고, 많은 계산량을 요구하는 최고 계층에서의 정수 단위 MV 탐색을 수행할 필요도 사라질 것이다. 이를 위해, 본 논문은 저 해상도 계층에서 정수 단위의 MV를 탐색할 수 있는 정합 오차 예측 기법을 제안하고, 이를 이용한 새로운 고속 다 해상도 BMA (multi-resolution BMA: MRBMA)를 선보인다. 한편, 제안한 움직임 추정 기법을 적용하기 위해 기존의 다 해상도 프레임 구조와 다른 현재 프레임 구조가 요구되며, 다음 절에서 이를 위한 블록 평균 피라미드 프레임 구조를 정의한다.

1. MRBMA의 프레임 구조

본 논문은 계층 수가 3인 MRBMA만을 다룬다. 계층 0을 최저 해상도 계층으로, 계층 2을 최고 해상도 계층으로 정의한다. MRBMA를 위한 이전 프레임 데이터 구조로는 그림 1처럼 기존의 다 해상도 구조가 이용된다. 저 해상도 계층의 각 화소는 인접한 상위 계층의 2×2 블록의 평균을 구함으로써 얻어진다.

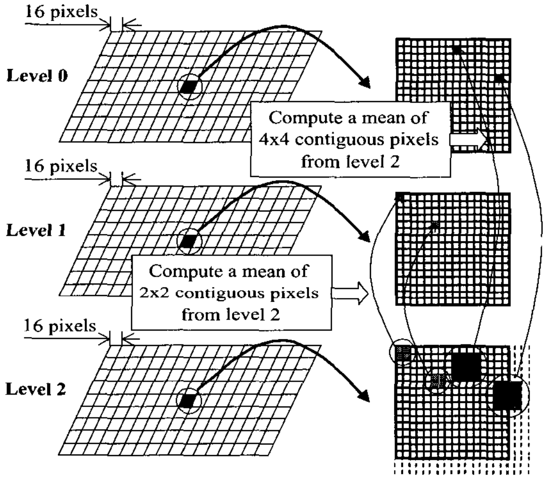


그림 2. 현재 프레임을 위한 블록 평균 피라미드 프레임 구조. 각 사각형은 한 블록을 의미한다
Fig. 2. Block mean pyramid structure of the current frame. Each rectangle denotes a block.

한편 현재 프레임의 구조로서 그림 2처럼 블록 평균 피라미드 구조를 도입한다[11]. 기존의 다 해상도 프레임 구조와 달리, 블록 평균 피라미드 구조의 프레임 및 블록 크기가 모든 계층에서 동일하다는 것이 특징이다.

2. SAD 예측 기법

본 논문에서는 정합 오차 척도로서 SAD (sum-of-absolute differences)를 사용한다. 계층 l에서, 현재 블록(우측 최상 단의 좌표가)의 SAD는 일반적으로 다음과 같이 구해진다.

$$SAD^{l}_{normal}(p,q)=\sum^{\frac{N}{2^{2-l}}-1}\sum^{\frac{N}{2^{2-l}}-1}\left|I^{l}_{k}(m\cdot 2^{2-l},n\cdot 2^{2-l})-I^{l}_{k-1}\left(m+\frac{p}{2^{2-l}},n+\frac{q}{2^{2-l}}\right)\right|,\ l=0,1,2. \quad (1)$$

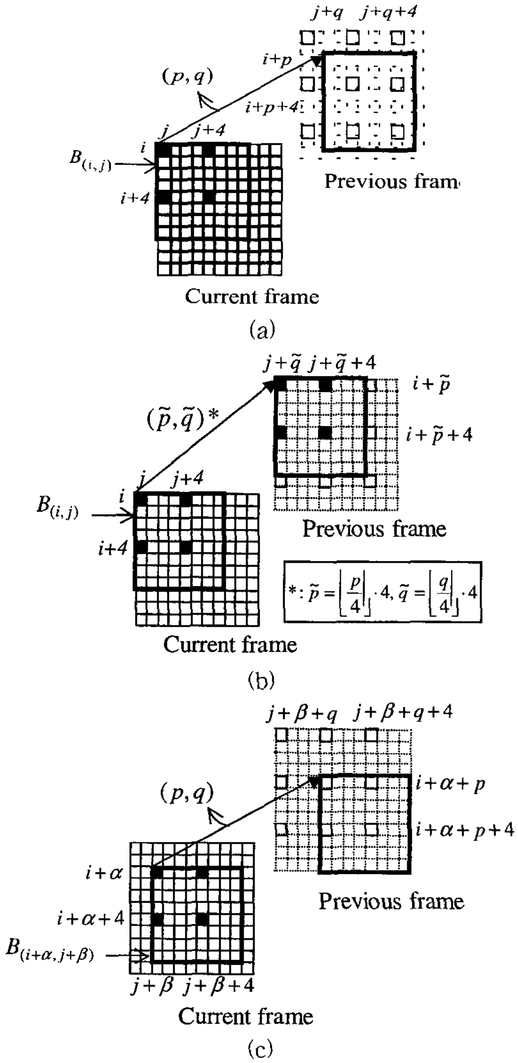


그림 3. 계층 0이고 $N=8$일 때, p와 q가 4의 배수가 아닌 MV 탐색. 검은 색 화소들이 SAD 계산 시 대응된다. 점선으로 표시된 화소들은 실제 메모리에 존재하지 않는 가상의 화소들이다. 편의상 모든 좌표는 계층 2의 관점에서 표기된다. (a) 이상적인 블록 정합, (b) 식 (1)이 적용되는 블록 정합 (MV양자화), (c) 식 (2)이 적용되는 블록 정합 (제안한 방식)

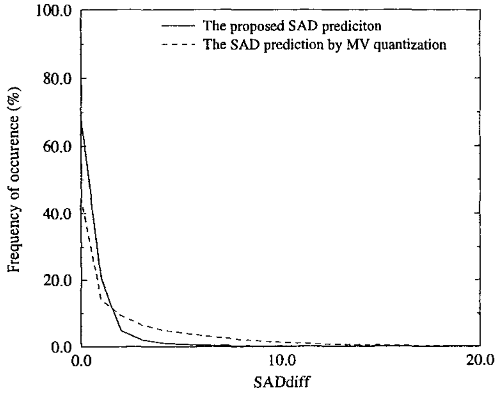
Fig. 3. The search of MV whose components are not 4-multiples at level 0 in the case of $N=8$. Black pixels are matched in computing SAD. Dotted pixels are imaginary ones which do not exist in the actual frame memory. For convenience, all the coordinates are written from the perspective of level 2. (a) Ideal block matching, (b) block matching applying Eq. (1) (MV quantization), (c) block matching applying Eq. (2) (the proposed method).

여기서 (p, q)는 탐색중인 MV이며, 블록의 크기는 $N\times N$이다. 편의상 블록의 좌표 는 식에 표기하지 않았다. $I_k^l(\cdot)$과 $I_{k-1}^l(\cdot)$는 각각 계층 l에서의 현재 프레임과 이전 프레임을 의미하며, 1절의 프레임 구조를 갖는다. 식 (1)을 통해 SAD를 계산할 경우 2^{2-l}의 배수인 MV들만 탐색이 가능하다 (그림 3 (b)). 다시 말해, $l=0$이고 $-3\le p,q\le 3$이면, 항상 $SAE_{normal}^0(p,q) = SAD_{normal}^0(0,0)$이다. 따라서, 일반 영상들의 경우 MV양자화에 따른 SAD오차가 매우 클 수 있다. 이러한 문제점을 해결하기 위해 2^{2-l}의 배수가 아닌 MV들에 대해서도 양자화 없이 SAD를 구할 수 있는 다음과 같은 SAD예측 기법을 제안한다.

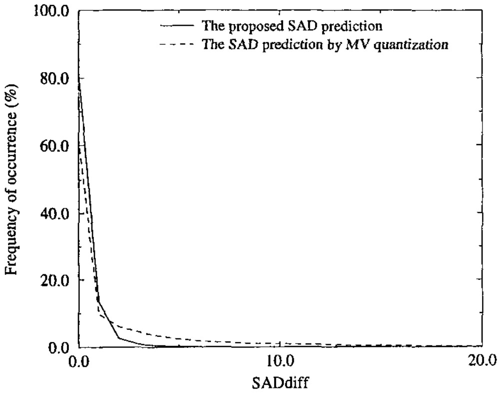
$$SAD_{pred}^l(p,q)=\sum_{m=0}^{\frac{N}{2^{2-l}}-1}\sum_{n=0}^{\frac{N}{2^{2-l}}-1}\left|I_k^l(m\cdot 2^{2-l}+\alpha, n\cdot 2^{2-l}+\beta)-I_{k-1}^l(m+\left\lceil\frac{p}{2^{2-l}}\right\rceil, +n+\left\lceil\frac{q}{2^{2-l}}\right\rceil)\right|, l=0,1,2 \quad (2)$$

$\alpha=\left\lceil\frac{p}{2^{2-l}}\right\rceil\cdot 2^{2-l}-p$, $\beta=\left\lceil\frac{q}{2^{2-l}}\right\rceil\cdot 2^{2-l}-q$이며, $\lceil A \rceil$는 A보다 크거나 같은 최소 정수이다. 그림 3 (a)는 p와 q가 2^{2-l}의 배수가 아닌 경우, 계층 1에서 정확한 SAD를 구하는 것이 이전 프레임 구조 때문에 불가능함을 보인다. 그래서, 본 논문은 식 (2)를 적용하여 공간적으로 약간 이동된 $B_{i+\alpha,j+\beta}$의 SAD를 $B_{(i,j)}$의 SAD대신 사용할 것을 제안한다 (그림 3 (c)). 보통 α와 β는 N보다 훨씬 작아 $B_{(i,j)}$의 대부분이 $B_{(i+\alpha,j+\beta)}$과 겹친다. 따라서, 식 (2)를 통해 얻어진 $B_{(i+\alpha,j+\beta)}$의 SAD는 의 실제 SAD에 상당히 근접하다. 그림 3 (b)와 (c)는 각각 계층 0에서 4의 배수가 아닌 MV들에 대해 기존의 방식으로 SAD를 구하는 경우와 제안한 예측 기법으로 SAD를 구하는 경우를 보여준다. 다음 실험 결과는 그림 3 (b)처럼 MV를 양자화하여 구한 SAD보다 그림 3 (c)처럼 MV의 해상도를 그대로 유지하며 제안한 방식대로 구한 SAD가 실제 SAD에 훨씬 가깝다는 것을 증명해준다. 먼저 III장의 실험에서 사용할 여러 실험 영상들에 대해 FSBMA를 통해 각 블록의 최적 MV를 찾는다. 계층 0에서, 4의 배수가 아닌 MV들에 대해서만 다음 두 종류의 SAD를 계산한다: 제안한 예측 기법에 의한 SAD (식 (2))와 MV양자화에 의한 SAD (식 (1)). 이 두 SAD들은 각각 계층 0에서의 실제 SAD와 비교된다. 이 실제 SAD값은 현재 프레임

구조를 이전 프레임 구조와 동일하게 함으로써 얻을 수 있다. 그림 4 (a)는 그 실험 결과를 보여준다. 그림에서 SAD_{diff}는 실제 SAD와 절대차 (absolute difference)를 의미하며, 세로축은 각 SAD_{diff}의 발생 빈도를 의미한다. 실제 영상에서 SAD_{diff}가 20이상인 경우는 거의 없으므로 명확한 비교를 위해 0부터 20까지의 SAD_{diff}에 대해서만 그래프들을 도시하였다. 그림 4 (b)는 계층 1에서의 실험 결과를 보여준다. 계층 1에서는 2의 배수가 아닌 MV들에 대해서만 실험한 결과이다. 두 저 해상도 계층 모두에서 제안한 방법에 의한 SAD가 MV양자화에 의한 SAD보다 훨씬 정확함을 알 수 있다.



(a)



(b)

그림 4. 제안한 SAD 예측기법의 정확도 테스트 (a) 계층 0 (b) 계층 1

Fig. 4. Accuracy test of the proposed SAD prediction scheme (a) at level 0 and (b) at level 1.

3. MRBMA의 프레임워크

MRBMA는 2절에서 제안한 SAD예측 기법을 적용한 새로운 다 해상도 BMA이다. 기존의 다 해상도 BMA와는 달리 저 해상도에서도 정수 단위의 MV추정을 상당히 정확하게 수행할 수 있기 때문에, MRBMA는 최적에 가까운 추정 성능 및 속도를 갖는다. 또한 많은 계산량을 요구하는 최고 해상도 계층에서의 탐색을 성능 저하 없이 생략할 수 있다. MRBMA의 전체적인 구조가 그림 5에 나타나 있다.

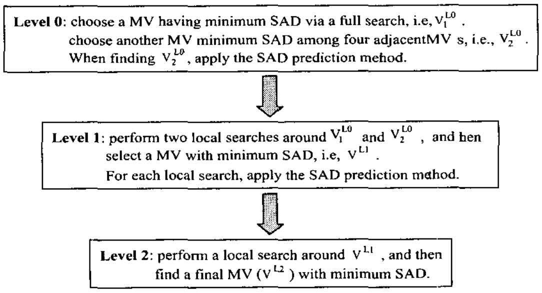


그림.5. MRBMA의 전체 동작 흐름도

Fig. 5. The overall working flow of MRBMA.

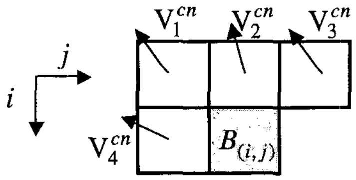


그림 6. 주변 블록들의 기하학적 위치

Fig. 6. Geometry of causally adjacent blocks.

계층 0에서 기본적으로 전역 탐색을 수행한다. 계층 2의 관점에서 볼 때, 계층 0의 전역 탐색은 전체 탐색 영역 $\Omega_S((2N_S+1)\times(2N_S+1))$ 내에서 일정한 간격을 가지고 추출된 MV들을 탐색하는 것과 동일하다. 탐색 영역 내에서 최소 SAD를 갖는 MV (V_1^{L0})를 선택하여, 중간 계층 탐색에서 초기 탐색점으로 삼는다. 여기서 위첨자 L0의 0는 계층 0를 의미한다. MRST처럼 좀더 나은 성능을 위해, $B_{(i,j)}$의 주변 블록들의 MV들 (V_m^{cn}, $1\le m\le 4$) 도 추가로 탐색한다 (그림 6 참조). 확률적으로 V_m^{cn}의 성분들이 4의 배수가 아닐 가능성이 많기 때문에, V_m^{cn}의 해상도를 유지하며 대응하는 SAD를 구하기 위해 SAD예측 기법을 적용한다. 네 개의 MV들 중 역시 최소 SAD를 갖는 MV (V_2^{L0})를 선택한다. 이렇게 얻어진 두 MV들, 즉 V_1^{L0}과 V_2^{L0}은 계층 1에서의 초기 탐색 중심점들로 이용된다. 한편, 계층

0의 모든 블록 정합 과정에서 이전 프레임의 탐색 영역으로부터 데이터를 규칙적으로 읽어오는 것은 하드웨어 구현 관점에서 매우 중요하다. 이를 위해 전역 탐색에서 양자화된 V_m^{cn}의 탐색은 생략되고, 대신 해당 위치에서 양자화되지 않은 V_m^{cn}에 대해 제안한 SAD 예측을 수행한다. 이렇게 함으로써, 이전 프레임의 데이터 흐름의 규칙성을 유지할 수 있다. 그림 7의 예는 N_S가 7이고 네 개의 주변 MV들이 주어졌을 때, 계층 0에서의 탐색 과정을 보여준다. 편의상 블록 크기가 8×8인 예를 보인다. 왼편에 네 개의 주변 MV들이 계층 2의 관점으로 표시되어 있다. 오른편에는 9개의 블록 정합이 순서대로 나타나 있으며, 실제 이전 프레임 메모리에 존재하는 화소들만을 나타내었다. 결과적으로, 이전 프레임으로부터의 데이터 흐름이 규칙적으로 이루어지고, 현재 블록 데이터에서 MV에 따라 화소 추출 위치가 조금씩 달라진다. 따라서 하드웨어 구현에도 적합하다.

계층 1에서는 두 초기 탐색점들을 중심으로 지역 탐색을 수행하며, 계층 2에서의 탐색을 위한 초기점 V^{L1}을 찾는다. 계층 1에서의 탐색 영역 Ω_{L1}은 $(2N_{L1}+1)\times(2N_{L1}+1)$으로 고정된다 (본 실험에서는 3×3이나 5×5). 이 탐색 영역은 전체 탐색 영역 Ω_S에 비해 현저하게 줄어든 것이다. 한편, 기존의 다 해상도 BMA들에서는 계층 2에서만 정수 화소 단위의 지역 탐색이 가능했던 반면, MRBMA에서는 중간 계층의 탐색도 제안한 SAD 예측 기법을 도입함으로써 정수 화소 단위로 수행되어짐에 주목해야 한다. 그림 8의 예는 계층 1에서 $N_{L1}=1$일 때의 탐색 과정을 보여준다. 모든 MV들의 해상도가 유지되며 SAD값들이 정해짐을 알 수 있다. 이전 프레임과 현재 프레임 모두 주어진 데이터 내에서 추출되는 화소 위치만이 약간 달라짐에 주목하라.

계층 2에서, V^{L1}중심의 지역 탐색 후 최종적인 MV인 V^{L2}을 구한다. 여기서 탐색 영역 Ω_{L2}는 $(2N_{L2}+1)\times(2N_{L2}+1)$이다 (우리 실험에서는 3×3).

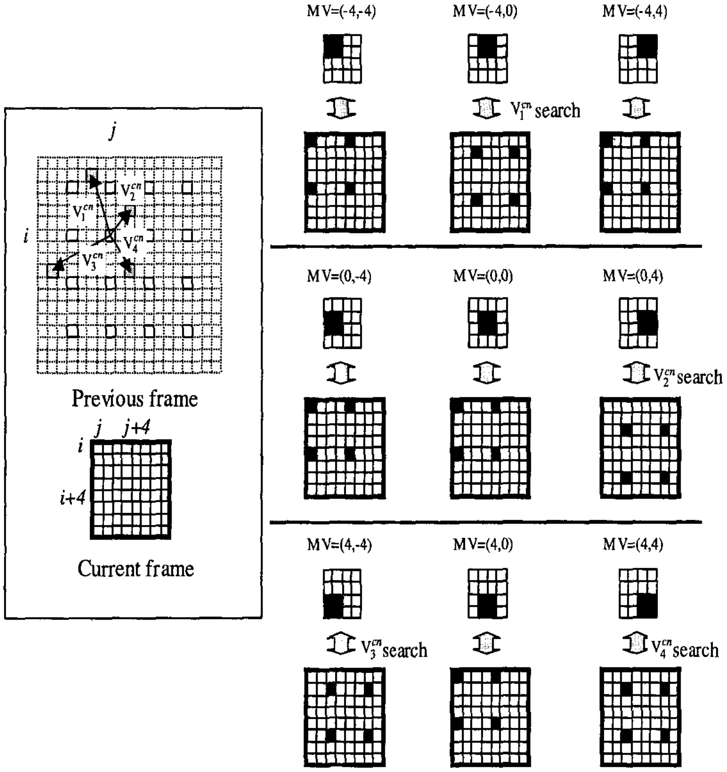


그림 7. N_S가 7이고 네 개의 주변 MV들이 주어졌을 때, 계층 0에서의 탐색 과정

Fig. 7. The search procedure at level 0 when N_S=7 and four adjacent MV's are given.

앞서 언급했듯이, 계층 1에서 이미 정수 단위의 탐색을 수행했기 때문에, 계층 2에서의 지역 탐색은 MRBMA의 전체 성능에 미미한 영향을 주게 된다. 따라서, 상당한 계산량을 필요로 하는 계층 2에서의 탐색을 생략함으로써 성능 저하 없이 추가적인 속도 개선을 할 수 있다. 3절에서의 모의 실험 결과가 이를 뒷바침해준다.

Ⅲ. 모의 실험 결과

다음 네 가지의 MPEG2 동영상들이 사용되었다: "cheer leaders (cheer)," "car (car)," "football (foot)," "flower garden (flower)." 각각 720×480의 해상도를 갖는 100프레임들로 구성된다. 위 네 가지 영상들은 빠른 움직임, 연속적인 움직임, 불규칙한 움직임 등 다양한 움직임 종류들을 포함하고 있기 때문에 제안한 기법의 성능 평가에 적합하다. $MRBMA_{nS}$ 는 n-단계 MRBMA를 의미한다. 즉, $MRBMA_{3S}$는 마지막 계층 탐색이 있는 MRBMA이고, $MRBMA_{2S}$는 마지막 계층 탐색이 생략된 MRBMA를 의미한다. 블록의 크기는 16×16이다.

1. MRBMA의 움직임 보상 성능 평가

MRBMA를 MPEG2 부호기에 적용하기에 앞서, 위 네 가지 영상들에 대해 움직임 보상 성능을 실험해 본다. FSBMA, TSS를 N_S에 따라 확장한 nSS (n=3, 4, 5, …), MRST와 성능을 비교한다. 표 1으로부터 MRBMA가 여러 탐색 영역들에 대해 모두 FSBMA와 거의 동일한 최고의 성능을 가짐을 알 수 있다. 또한 마지막 계층에서의 탐색을 생략하더라도 성능 저하가

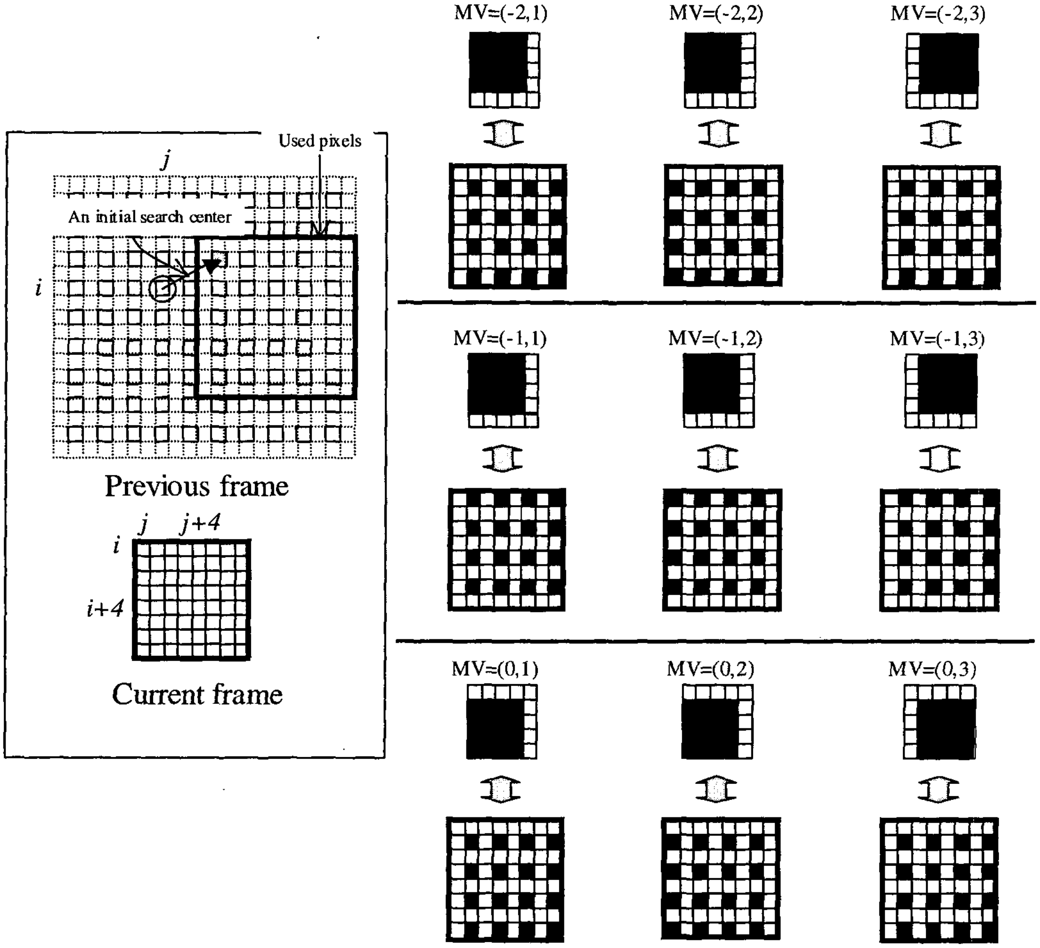


그림 8. N_S가 7이고 초기점이 (-1,2) 이라고 할 때, 계층 1에서의 탐색 과정

Fig. 8. The search procedure at level 1 when N_S=7 and an initial point is (-1,2).

거의 발생하지 않는 점에 주목할 만하다.

또한 탐색 영역 크기에 따른 MRBMA의 연산량을 표 I에 기술하였다. 블록 당 연산량를 기준으로 하여 FSBMA와의 상대적인 연산량을 속도향상도 (speed-up ratio: SUR)로 정의하였다. 연산량에는 다 해상도 프레임 구조를 생성하기 위해 요구되는 연산량도 포함되어 있다. 127×127의 탐색 영역에 대해 MRBMA2S는 200배 정도 FSBMA보다 적은 연산량을 갖는다. MRBMA의 3계층을 모두 사용할 때 연산량은 다음과 같다.

$$C_{3S} = \left\{ \left(\frac{2N_S + 1}{2^2} \right)^2 \frac{N^2}{16} + 2(2N_{L1} + 1)^2 \frac{N^2}{4} + (2N_{L2} + 1)^2 N^2 \right\} MN_{fr} \quad \text{operations/sec} \tag{3}$$

M과 Nfr은 각각 SAD연산 시 화소 당 연산량과 프레임율이다. MRBMA에서 다 해상도 프레임 구조를 생성하기 위해 요구되는 블록 당 연산량은 $8N^2$이다[11]. 표 II는 MRBMA에서 다 해상도 프레임 구조를 생성하기 위해 요구되는 오버헤드 연산량을 보여준다. $MRBMA_{3S}$의 오버헤드 연산량은 $8N^2/C_{3S}$로 정의된다. 탐색 영역이 커짐에 따라 오버헤드 연산량이 미미해짐을 알 수 있다.

표 1. 여러 기법들의 움직임 보상 성능 비교

Table 1. Motion compensation performance comparison for various algorithms.

(a) N_S =15

	FSBMA	4SS		MRST		MRBMA							
						3S (N_{L1}=2, N_{L2}=1)		3S (N_{L1}=2, N_{L2}=1)		2S (N_{L1}=2)		2S (N_{L1}=1)	
		PSNR	SUR	PSNR	SUR	PSNR	SUR	PSNR	SUR	PSNR	SUR	PSNR	SUR
car	25.5	24.4	31.3	25.2	25.6	25.4	25	25.4	38	25.3	41	25.3	90
foot	25.0	23.8	31.3	24.7	17.9	24.9	25	24.9	38	24.8	41	24.8	90
cheer	22.1	21.6	31.3	21.7	14.9	22.0	25	22.0	38	22.0	41	21.9	90
flower	27.7	24.2	31.3	27.7	19.6	27.7	25	27.7	38	27.6	41	27.6	90

(b) N_S =31

	FSBMA	5SS		MRST		MRBMA							
						3S (N_{L1}=2, N_{L2}=1)		3S (N_{L1}=2, N_{L2}=1)		2S (N_{L1}=2)		2S (N_{L1}=1)	
		PSNR	SUR	PSNR	SUR	PSNR	SUR	PSNR	SUR	PSNR	SUR	PSNR	SUR
car	29.5	25.8	100	29.2	71.4	29.4	81	29.4	109	29.3	114	29.3	176
foot	25.5	23.5	100	25.1	47.6	25.4	81	25.4	109	25.3	114	25.3	176
cheer	23.1	22.2	100	22.6	41.6	23.0	81	23.0	109	22.9	114	22.9	176
flower	27.8	21.0	100	27.7	55.5	27.8	81	27.8	109	27.7	114	27.6	176

(c) N_S =63

	FSBMA	6SS		MRST		MRBMA							
						3S (N_{L1}=2, N_{L2}=1)		3S (N_{L1}=2, N_{L2}=1)		2S (N_{L1}=2)		2S (N_{L1}=1)	
		PSNR	SUR	PSNR	SUR	PSNR	SUR	PSNR	SUR	PSNR	SUR	PSNR	SUR
car	30.2	26.9	333	29.9	167	30.1	167	30.2	192	30.1	195	30.1	229
foot	25.6	21.0	333	25.2	111	25.5	167	25.5	192	25.4	195	25.4	229
cheer	23.5	22.3	333	23.0	91	23.4	167	23.4	192	23.3	195	23.3	229
flower	27.8	20.1	333	27.7	100	27.8	167	27.8	192	27.7	195	27.7	229

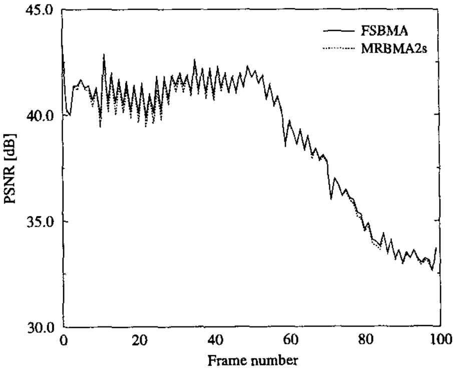
표 2. MRBMA에서 다 해상도 구조 생성을 위한 오버헤드 연산량 (%)

Table 2. Overhead computational cost (%) for multi-resolution frame structures of MRBMA.

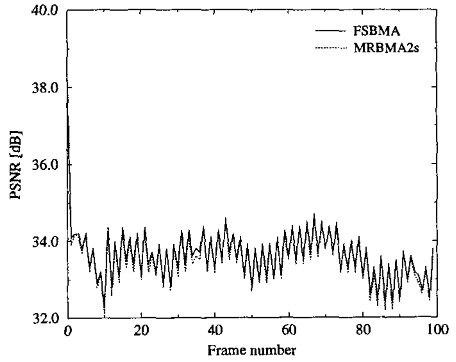
	MRBMA			
	3S (N_{L1}=2, N_{L2}=1)	3S (N_{L1}=1, N_{L2}=1)	2S (N_{L1}=2)	2S (N_{L1}=1)
N_S=15	10.6	15.4	16.4	32.2
N_S=31	5.4	9.2	9.5	13.2
N_S=63	3.2	3.5	3.5	4.0

2. MPEG2 동영상 부호화기에서의 MRBMA의 성능 평가

GOP 크기는 15프레임, P프레임 간 거리는 3 프레임, 비트 율은 4 Mbps로 고정한 후 실험하였다. Ω_S는 P프레임의 경우 127×127이며, B 프레임의 경우 63×63로 정했다. MRBMA의 성능을 FSBMA와 평균 PSNR ($PSNR_{avg}$)견지에서 비교하였으며, 추가로 프레임 별 PSNR 차이의 최대 편차 (maximum deviation: d_{max}) 및 최소 편차 (minimum deviation: d_{min})를 조사하였다. 표 III에 그 결과를 보였다. MRBMA가 FSBMA와 거의 동일한 $PSNR_{avg}$을 가짐을 알 수 있다. 최악의 경우인 "car"와 "football" 영상들에 대해 최대 0.2dB 정도 $PSNR_{avg}$가 FSBMA보다 떨어지지만, "flower garden" 영상에서는 $PSNR_{avg}$이 지역 탐색 영역 크기에 관계없이 FSBMA와 동일함에 주목하라. 한편, "car" 영상을 제외한 모든 영상들에 대해 FSBMA와의 PSNR편차가 0.2dB~0.1dB로 움직임에 거의 편차가 없음을 알 수 있다. "car"영상의 경우도 앞부분의 몇몇 프레임에 대해서만 0.5dB~0.6dB정도 차이가 날뿐, 움직임이 큰 뒷부분의 프레임들에선 FSBMA와 거의 차이가 존재하지 않는다 (그림 9(a)). 그림 9에서 두 개의 실험 영상들인 "car"와 "football"에 대해 MRBMA2S의 프레임 별 PSNR을 FSBMA와 비교하였다. FSBMA와 미세한 PSNR차이를 보이며 따라감을 확인할 수 있다.



(a)



(b)

그림 9. MRBMA2S과 FSBMA의 PSNR 비교 (a) "car," and (b) "football."

Fig. 9. PSNR comparisons of MRBMA2S with FSBMA (a) "car," and (b) "football."

IV. 결 론

본 논문은 저 해상도 계층에서도 정수 화소 단위의

표 3. MPEG2 부호화기에서의 MRBMA와 FSBMA와의 PSNR 비교. 여기서 PSNR은 복원된 영상에 대해 구해진다

Table 3. PSNR comparison of the MRBMA with the FSBMA in an MPEG2 encoder. Here PSNR is computed from the reconstructed frames.

	FSBMA	$MRBMA_{3S}$ (N_{L1}=2, N_{L2}=1)			$MRBMA_{3S}$ (N_{L1}=1, N_{L2}=1)			$MRBMA_{2S}$ (N_{L1}=2)			$MRBMA_{2S}$ (N_{L1}=1)		
		$PSNR_{avg}$	d_{max}	d_{min}	$PSNR_{avg}$	d_{max}	d_{min}	$PSNR_{avg}$	d_{max}	d_{min}	$PSNR_{avg}$	d_{max}	d_{min}
car	39.0	38.9	0.1	-0.4	38.9	0.1	-0.4	38.8	0.1	-0.6	38.8	0.1	-0.6
foot	33.7	33.6	0.1	-0.2	33.6	0.1	-0.2	33.6	0	-0.2	33.5	0	-0.2
cheer	31.3	31.2	0.1	-0.2	31.2	0	-0.2	31.2	0	-0.2	31.2	0	-0.2
flower	31.8	31.8	0.1	-0.1	31.8	0.1	-0.1	31.8	0.1	-0.1	31.8	0	-0.2

탐색을 할 수 있도록 SAD 예측 기법을 제안하고, 이를 적용한 새로운 고속 다 해상도 BMA (MRBMA) 를 제시하였다. 3계층 프레임 구조를 갖는 제안한 알고리즘은 중간 계층에서 이미 정수 화소 단위의 탐색을 수행할 수 있기 때문에 많은 계산량을 요하는 마지막 계층에서의 탐색을 생략함으로써 성능 저하 없이 속도를 개선할 수 있었다. 모의실험을 통해 MRBMA가 MPEG2의 일반적인 탐색 영역에 대해 탁월한 성능을 보임을 확인했다. 따라서, MRBMA는 실시간 동영상 부호화에 유용하게 적용될 수 있다. 또한 MRBMA는 전체 탐색 과정동안 규칙적인 데이터 흐름을 갖기 때문에 하드웨어 구현에도 적합하다.

참 고 문 헌

[1] MPEG2, "Information technology-generic coding of moving pictures and associated audio," Tech. Rep., ISO/IEC 13818-2, Committee Draft, Mar. 1994.

[2] CCITT Study group XV, "Draft revision of recommendation H.261-Video codec for audio visual services at px64 kbps," Temporary Document 5-E, July 1990.

[3] J. R. Jain and A. K. Jain, "Displacement measurement and its application in inter-frame image coding," *IEEE Trans. Commun.*, vol. COM-29, no. 12, pp. 730-741, Dec. 1981.

[4] T. Koga, K. Iinuma, A. Hirano, Y, Iijima, and T. Ishiguro, "Motion compensated inter-frame coding for video conferencing," *in Proc. Nat. Telecommunication Conf.*, pp. G5.3.1-5.3.5, New Orleans, USA, Nov. 29-Dec. 3, 1981.

[5] B. Liu and A. Zaccarin, "New fast algorithms for the estimation of block motion vectors," *IEEE Trans. Circ. Sys. Video Technol.*, vol. 3, no. 2, pp. 148-157, Apr. 1993.

[6] M. Bierling, "Displacement estimation by hierarchical block matching," in *Proc. SPIE Visual Communications and Image Processing,* SPIE-vol. 1001, pp. 942-951, Boston, USA, Nov. 1988.

[7] K. M. Nam, J. S. Kim, R. H. Park, and Y. S. Shim, "A fast hierarchical motion vector estimation algorithm using mean pyramid," *IEEE Trans. Circ. Sys. Video Technol.*, vol. 5, no. 4, pp. 344-351, Aug. 1995.

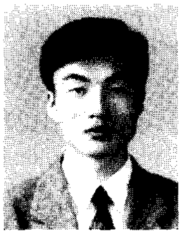
[8] J. Chalidabhongse and C.-C. J. Kuo, "Fast motion vector estimation using multiresolution-spatio-temporal correlations," *IEEE Trans. Circ. Sys. Video Technol.*, vol. 7, no. 3, pp. 477-488, June 1997.

[9] S. Zafar, Y. Q. Zhang, and J. S. Baras, "Predictive block matching motion estimation for TV coding part I: Inter-block prediction," *IEEE Trans. Broadcasting,* vol. 37, no. 3, pp. 97-101, Sep. 1991.

[10] B. C. Song and J. B. Ra, "Hierarchical block matching algorithm using partial distortion criterion," *in Proc. SPIE Visual Communications and Image Processing,* SPIE-vol. 3309, pp. 88-95, San Jose, USA, Jan. 1998.

[11] W. Li and E. Salari, "Successive elimination algorithm for motion estimation," *IEEE Trans. Image Processing,* vol. 4, no.1, pp. 105-107, Jan. 1995.

저 자 소 개



宋 秉 哲(正會員)

1972년 11월 8일생. 1994년 2월 한국과학기술원 전기 및 전자공학과 공학사. 1996년 2월 한국과학기술원 전기 및 전자공학과 공학석사. 1996년~현재 한국과학기술원 전기 및 전자공학과 박사과정 재학 중. 주관심분야는 동영상 부호화 및 영상 표현 및 탐색.

羅 鍾 範(正會員)

1953년 5월 23일생. 1975년 2월 서울대학교 전자공학과 공학사. 1977년 2월 한국과학기술원 전기 및 전자공학과 공학석사. 1983년 2월 한국과학기술원 전기 및 전자공학과 공학박사. 1983년 7월~1987년 6월 미국 Columbia 대학교 연구 조교수. 1987년 7월~현재 한국과학기술원 전기 및 전자공학과 교수. 주관심분야는 디지털 영상처리, 동영상 신호처리, 3차원 시각화, 의료영상 시스템 등