

# 운동영역의 상관성을 선택적으로 이용한 고속 움직임 추정 기법

임 경 원, 송 병 철, 나 종 범  
한국과학기술원 전기 및 전자공학과

## Fast Hierarchical Block Matching Algorithm by Adaptively Using Spatial Correlation of Motion Field

K.W. Lim, B.C. Song, and J.B. Ra

Department of Electrical Engineering, Korea Advanced Institute of Science and Technology,  
373-1, Kusongdong, Yusonggu, Taejon, Korea.

### ABSTRACT

This paper describes a new hierarchical block matching algorithm especially appropriate for a large search area. The proposed algorithm consists of higher level search for an initial motion vector estimate by using a new matching criterion over the evenly subsampled search points, and lower level search for the final motion vector refinement. In the higher level matching criterion, mean absolute differences at the search points (or motion vector candidates) similar to motion vectors of causally neighboring blocks, are weighted properly so that these points can have a higher chance to being selected. The proposed algorithm outperforms existing hierarchical block matching algorithms, and its computational regularity makes hardware implementation simple.

### I. 서론

최근, 실시간 부호화가 필수적인 통신, 방송 매체와 함께 저장 매체를 지원하기 위하여 MPEG-2 동영상 압축 표준이 제정되었다 [1]. MPEG-2 동영상 압축 과정은 공간적인 중복성을 제거하기 위하여 DCT (discrete cosine transform)를 이용하는 변환 부호화 방식을 사용하고, 시간적인 중복성을 제거하기 위하여 움직임 추정 및 보상 기법을 사용한다. BMA (block matching algorithm)는 움직임 추정을 위하여 일반적으로 사용하는 방법으로서, BMA를 위한 직접적인 탐색 방법으로 전역 탐색 방법이 있다. 전역 탐색 방법은 탐색 영역내에서 최소의 정합 오차 (matching error)를 갖는 블럭을 찾기 위하여 모든 가능한 후보 위치의 정합오차를 계산한다. 그러나, 전역 탐색은 전체 부호화 과정의 50%를 상회하는 막대한 계산량을 필요로 하고, 구현 시에도 하드웨어 부담이 크며, 그러한 부담은 탐색 영역의 크기가 증가함에 따라 지수적으로 증가한다. 따라서, 디지털 TV, HDTV와 같은 응용을 위해서는 넓은 탐색 영역에 대해 실시간으로 움직임을 추정할 수 있는 고속 탐색 방법을 찾는 것이 중요하다.

기존의 고속 움직임 추정 기법중에서 간단한 2단계 계층적 탐색기법들이 이러한 목적을 위하여 적용될 수 있다 [2,3]. Wang *et al.*은 [2]에서 탐색영역 내의 모든 탐색점에서 추림 (subsampling)된 현재 블럭과 후보 블럭들 사이의 정합 오차를 비교하여 초기 움직임 벡터를 결정하고, 초기 움직임 벡터를 중심으로하는 국부 영역만을 탐색하여 최종 움직임 벡터를 결정한다. [3]에서는, 4개의 초기 움직임 벡터가 alternating pixel subsampling patterns 기법을 통하여 결정되고, 최종 움직임 벡터로서 그중 하나를 선택한다. [2]와 [3]에서는 공통적으로 초기 움직임 벡터를 결정하기 위하여 탐색 영역 내의 모든 탐색점을 탐색하며, 계산량을 제한하기 위하여 추림된 화소 (pixel)을 사용한다. 그러나, 최종 움직임 벡터에 가까운 초기 움직임 벡터를 결정하기 위해서 추림 간격을 충분히 작게 해야하므로, 넓은 탐색영역을 탐색하기에는 계산량이 여전히 많다. 또한, 계산량을 더욱 줄이기 위하여, subsampled motion-field estimation 혹은 sub-block motion-field estimation 방법이 결합된다면, uncausal 하고, 불규칙적인 data flow 때문에 하드웨어 구현에 적합하지 않게 된다.

이러한 문제를 해결하기 위해서, 본 논문에서는 새로운 2단계 계층적 탐색 방법을 제안한다. [2,3]에서와 같이 블럭내의 화소를 제한하는 대신에, 제안된 방법은 추림을 통하여 탐색점의 수를 제한하고, 움직임 벡터의 상관성을 적응적으로 이용하는 새로운 정합기준 (matching criterion)을 사용하여 초기 움직임 벡터를 결정하며, 초기 움직임 벡터를 포함한 국부 영역만을 탐색하여 최종 움직임 벡터를 결정한다. 제안된 방법은 기존의 방법 [2,3]보다 적은 계산량으로 우수한 움직임 추정 성능을 가질 수 있으며, 간단한 구조를 가지고 있기 때문에 하드웨어 구현이 용이하다는 장점을 갖는다.

### II. 제안된 방법

대부분의 동영상에서, 움직임 장 (motion field)은 연속, 불규칙, 또는 복잡한 움직임이 주를 이룬다. 만약, 하나의 물체가 연

속적으로 움직인다면, 그 물체 내부에 포함되어 있는 블록의 움직임 벡터들은 공간적으로 인접한 블록들의 움직임 벡터와 유사할 것이다. 따라서, 현재 블록의 움직임 벡터는 공간적으로 인접한 블록의 움직임 벡터들로부터 정확하게 예측될 수 있다 [4]. 그러나, 다중의 움직이는 물체 혹은 복잡한 움직임을 갖는 물체 상에 현재 블록이 있다면, 그러한 영역에 속한 블록의 움직임은 불규칙하거나 복잡할 것이고, 현재 블록의 움직임 벡터도 공간적으로 인접한 블록의 움직임 벡터와 유사하지 않을 수 있다. 이 경우, 상위 단계에서 규칙적으로 분포된 탐색점으로부터 초기 움직임 벡터를 구하는 계층적 탐색 방법은 보다 정확한 초기 움직임 벡터를 제공할 수 있다.

따라서, 연속, 불규칙, 또는 복잡한 움직임을 하는 모든 영역에서 정확하게 움직임 벡터를 예측할 수 있도록, 제안된 방법은 고정된 계층적 탐색 구조하에서 현재 블록의 움직임 벡터와 인접 블록의 움직임 벡터들 간의 상관성을 적용적으로 이용하였다.

## II. 1 단계 계층적 탐색 과정

제안된 방법은 2단계 계층적 탐색 기법이다 (그림 1 (a) 참조). 제안된 방법의 첫 단계에서는, 균일하게 수립된 탐색점들로부터 최소 정합 오차를 갖는 위치를 초기 움직임 벡터  $MV_{i,j}^0$ 로 정한다. 이때, 사용하는 정합 기준은 다음절에서 기술할 것이다. 두번째 단계에서는,  $MV_{i,j}^0 \pm L$ 의 국부 탐색점들에 대하여 MAD (mean absolute difference)를 계산하고 비교하여, 최소 MAD를 갖는 위치를 최종 움직임 벡터로 정한다.

## II. 2 움직임 벡터의 상관성을 고려한 변형된 정합 기준의 적용

계층적 탐색의 첫단계에서 움직임 벡터의 상관성을 이용하기 위하여, 현재 블록이 속해 있는 영역의 운동 형태를 인접한 블록의 움직임 벡터들간의 유사성으로부터 예측할 수 있다. 그림 2는 그러한 유사성에 따른 중요한 움직임 형태를 5가지로 분류한 것이다. 그림 2에서, class 1, class 2~4, 그리고 class 5는 현재 블록이 각각 움직이는 물체 내부, 움직이는 물체의 경계, 그리고 복잡한 움직임을 하는 영역에 속했을 때 발생할 가능성이 높다. 따라서, class 1에서는 현재 블록의 움직임 벡터  $MV_{i,j}$ 를 인접한 블록의 움직임 벡터들의 평균  $MV_{i,j}^0(0)$ 으로 예측할 수 있고, class 2~4에서는 쌍 (pair)을 이루는 인접 블록의 움직임 벡터의 평균  $MV_{i,j}^0(0)$ 과 남아있는 인접 블록의 움직임 벡터  $MV_{i,j}^0(1)$ 를 예측치로 사용할 수 있다. 그러나, class 5에서는 현재 블록이 불규칙한 움직임을 갖는 영역에 속하게 되므로  $MV_{i,j}$ 를 예측할 필요가 없다.

$MV_{i,j}^0(k)$  주위의 국부 영역내에 최소 정합 오차를 갖는  $MV_{i,j}$ 가 존재하는 확률 ( $P$ )는 각 class에 따라 다르며, 이러한 확률은

예측된 움직임 벡터의 신뢰성을 예상하기 위하여 이용할 수 있다.  $MV_{i,j}^0(k)$ 가 정의 되어진 각 4개의 class에 대하여  $P$ 를 조사하였다. 표 1은 블록 크기  $16 \times 16$  화소 ( $N=16$ ), 탐색 영역 크기  $65 \times 65$  화소 ( $S=32$ )를 사용한 실험 결과를 보여 준다. 표로부터, class 1의 예측 움직임 벡터는 최소 정합 오차를 갖는  $MV_{i,j}$ 를 찾는 데 매우 유용함을 알 수 있고, class 2~4의 예측 움직임 벡터, 특히 빗금친 인접 블록들의 평균 움직임 벡터  $MV_{i,j}^0(0)$ 도 움직이는 물체의 경계 부분에 대한 움직임 벡터를 찾는 데 유용하다는 사실을 확인할 수 있었다.

움직임 추정에 움직임 벡터의 연속성을 이용하기 위하여, 가장 효과적인 방법은 균일하게 수립된 탐색점들과 함께 예측된 움직임 벡터들을 최종 움직임 벡터를 위한 후보에 포함시키고, 각 후보들을 모두 탐색하는 방법이다 [5]. 그러나, 이것은 연산의 규칙성을 저해하고 하드웨어 구현을 복잡하게 만드는 문제를 가지고 있다. 이 문제를 피하기 위하여, class 1~4에 속하는 현재 블록 ( $i,j$ )에 대하여 예측된 움직임 벡터 주위의 탐색점들에 대하여 MAD를 확률  $P$ 에 비례하여 줄여주는 변형된 정합 기준을 적용한다. MAD를 비교하기 전에, 식 (1)의 조건을 만족하는 탐색점 ( $p,q$ )에 대하여 MAD에 가중치 (weighting factor)를 적용한다 (그림 1 (b) 참조).

$$\delta = \left| (p, q) - MV_{i,j}^0(k) \right| \leq D, \quad (1)$$

여기서,  $\|\cdot\|$ 는 Euclidian distance를 나타내고, class 1에 대하여  $k=0$ 이고, class 2~4에 대하여는  $k=0,1$ 이다. 적용한 가중치는 식 (2)와 같다.

$$f(\cdot) = \begin{cases} 1 - \eta_{c,k} \cdot \left[ \frac{1}{2} + \frac{1}{2} \cdot \cos\left(\frac{\pi \cdot \delta}{D}\right) \right] & \delta \leq D, \\ 1 & \text{else,} \end{cases} \quad (2)$$

여기서,  $c$ 와  $k$ 는 블록 ( $i,j$ )의 class 번호 및 예측 움직임 벡터에 대응되는 인덱스 (index)이다. 상수  $\eta_{c,k}$ 는 (0~1)의 범위를 가지며, 표 1에 주어진 확률  $P$ 에 비례하도록 선택된다. 예측된 움직임 벡터 주위의 수립된 탐색점의 MAD를 줄임에 의하여, 움직임 벡터의 상관성이 충분히 높은 경우에, 식 (2)는  $MV_{i,j}^0(k)$ 에 가까운 탐색점이 초기 움직임 벡터로서 선택될 기회를 증가시킨다. 이외의 경우 첫단계의 탐색은 일반적인 MAD비교에 의해 수행된다.

$(2S+1)^2$ 의 탐색점을 갖는 전역 탐색과 비교하여, 제안된 방법은 상위 단계에서  $(2S/D+1)^2$ 의 탐색점을 갖고, 하위단계에서  $(2L+1)^2$ 의 탐색점을 가지므로, 제안된 방법은 전역 탐색에 비해서  $(1/D)^2 + (L/S)^2$ 배의 속도 향상을 기대할 수 있다.

## III. 모의 실험 및 성능 평가

모의 실험을 통하여 제안된 방법의 성능을 다른 고속 BMA

방법들과 비교하였다. 사용된 영상은 MPEG-2 테스트 영상으로서 "Football" (fb), "Car" (car), "Susie" (sus), "Mobile and Calendar" (mob), 그리고 "Cheerleaders" (cheer)들이다. 모든 영상의 프레임률 (frame rate)은 30 프레임/초이고, 영상의 크기는  $720 \times 480$ 화소이다.  $\pm 32$ 의 넓은 탐색 영역이 사용되었고, 블록의 크기는  $16 \times 16$ 이다. 제안된 방법의 추림 간격  $D$ 는 8로 하였고,  $L$ 은  $D-1$  ( $=7$ )이다. 식 (2)의 상수  $\eta_{c,k}$ 는 실험적으로  $\eta_{1,0} = 0.6$ ,  $\eta_{2,0} = 0.3$ ,  $\eta_{3,0} = \eta_{4,0} = \eta_{3,1} = 0.2$ 로 결정하였다. class 2~4의  $MV_{i,j}^{(k)}$ 는 화질  $P$ 가 낮기 때문에 예측 움직임 벡터로 사용되지 않았다.

표 2에서는 제안된 방법의 평균 PSNR과 계산량에 전역 탐색 및 넓은 탐색 영역에 적합한 다른 고속 알고리즘 [2,3]과 비교되었다. 계산량을 맞추기 위해서 Wang *et al.*에 의해 제안된 방법 [2]의 추림 요소 (decimation factor)  $d$ 는 2로 고정되었고, Liu와 Zaccarin에 의해 제안된 방법 [3]중 motion field subsampling technique과 alternative pixel decimation patterns technique을 결합한 방법을 비교하였다. 표 2에서 보듯이, 제안된 방법은 연속적인 움직임 ("Car", "Susie"), 불규칙하고 복잡한 움직임 ("Cheerleaders"), 그리고 두 특성을 결합한 움직임 ("Football")을 가지는 영상들에 대하여 다른 BMA들 보다 우수한 움직임 추정 성능을 갖는다. "Mobile and Calendar" 영상의 경우, 제안된 방법의 성능은 다른 BMA들 보다 조금 (0.1dB) 낮지만 전역 탐색의 성능에 충분히 근접하고 있음을 알 수 있다. 또한, 제안된 방법의 계산량은 전역 탐색의 7.2%에 불과하며 비교된 방법들에서 가장 적은 계산량을 갖는다. 그림 3은 "Car", "Cheerleaders", 그리고 "Football" 영상의 처음 100 프레임에 대한 PSNR을 나타낸 것으로서, 제안된 방법이 다른 고속 BMA보다 항상 좋은 성능을 보이고 있음을 알 수 있다.

#### IV. 결론

본 논문은 움직임 벡터의 공간적인 상관성을 적절하게 이용하는 2단계 계층적 탐색기법을 제안하였다. MPEG-2와 같은 동영상 압축 표준에서 요구하는 넓은 탐색 영역에 대하여, 제안된 방법은 전역 탐색의 7.2% 계산량으로 다른 고속 알고리즘을 능가하는 성능을 보여 주었다. 또한, 알고리즘의 규칙성과 단순성 때문에, 제안된 방법은 하드웨어 구현에 적합한 특성을 갖는다.

#### 참고 문헌

[1] ISO/IEC 13818-2 | ITU-T H.262 (MPEG-2 Video), "Information technology- Generic coding of moving pictures and associated audio: Video," 1995.  
 [2] B.M. Wang, J.C. Yen, and S. Chang, "Zero waiting cycle hierarchical block matching algorithm and its array architectures," *IEEE Trans. Circ. and Syst. for Video*

*Technol.*, vol. 4, no. 1, pp. 18-28, Feb. 1994.

[3] B. Liu and A. Zaccarin, "New fast algorithms for the estimation of block motion vectors," *IEEE Trans. Circ. and Syst. for Video Technol.*, vol. 3, no. 2, pp. 148-157, Apr. 1993.  
 [4] M.C. Chen and A.N. Willson, Jr., "A high accuracy predictive logarithmic motion estimation algorithm for video coding," in *Proc. IEEE ISCAS'95*, Seattle, Washington, USA, pp. 617-620, 30 Apr. - 3 May 1995.  
 [5] B.C. Song, K.W. Lim, and J.B. Ra, "A fast motion estimation algorithm using spatial correlation of motion field and hierarchical search," in *Proc. SPIE Conf. Digital Comp. Technol. and Syst. for Video Commu.*, Berlin, FRG, vol. 2952, pp. 308-315, 7-9 Oct. 1996.

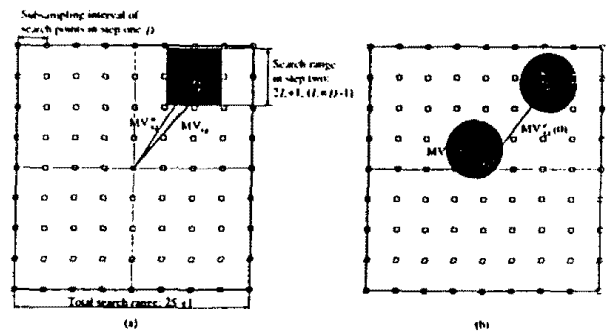


그림 1: (a) 제안된 2단계 계층적 탐색 기법.  $MV_{i,j}^{(k)}$ 와  $MV_{i,j}$ 는 각각 초기 움직임 벡터, 최종 움직임 벡터이다. (b) 첫 단계에서 MAD에 가중치가 더해지는 탐색점들 (검은 점들). 가중치는  $MV_{i,j}^{(k)}$ 로 부터의 거리와 표 1의 확률에 따라 다름.

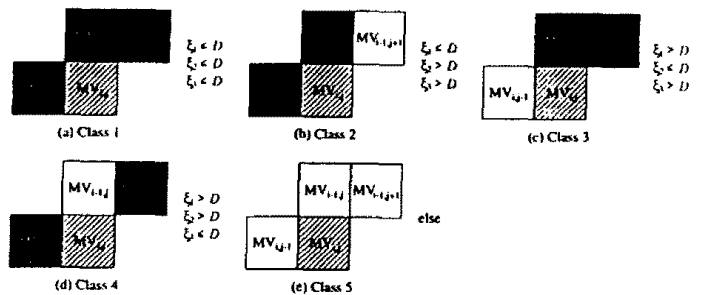


그림 2: 유사한 인접 움직임 벡터의 그룹 (group).  $\xi_1 = \|MV_{i,j} - MV_{i-1,j}\|$ ,  $\xi_2 = \|MV_{i,j} - MV_{i,j-1}\|$ , 그리고  $\xi_3 = \|MV_{i,j} - MV_{i-1,j-1}\|$ 이며,  $D$ 는 탐색점의 추림 간격이다.

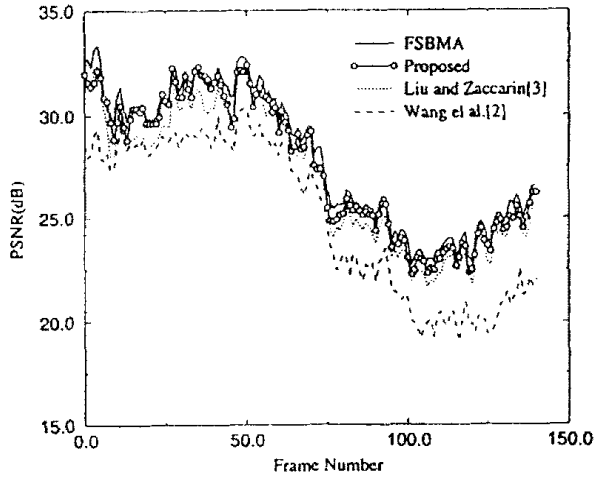
	Class 1	Class 2		Class 3		Class 4	
	$MV_{i,j}^{(0)}$	$MV_{i,j}^{(0)}$	$MV_{i,j}^{(1)}$	$MV_{i,j}^{(0)}$	$MV_{i,j}^{(1)}$	$MV_{i,j}^{(0)}$	$MV_{i,j}^{(1)}$
Probability (%)	95.2	70.5	8.8	48.6	23.8	64.1	11.6

표 1: 예측 움직임 벡터를 중심으로 하는 국부 영역에서 최소 오차를 갖는 움직임 벡터를 발견할 확률. 수치는 5가지 MPEG-2 테스트 영상으로 부터 얻어졌음.

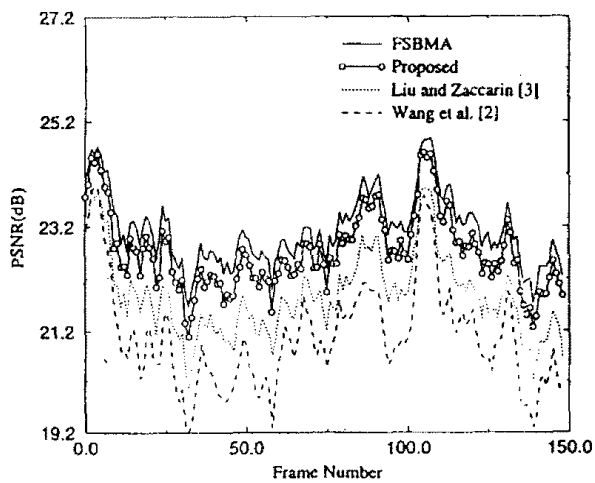
Algorithm	Average PSNR (dB)					Complexity (%)
	fb	car	sas	mob	cheer	
Full search method	25.5	29.5	36.5	24.8	23.1	100.0
Wang et al. [4]	22.8	27.0	34.6	24.6	21.0	25.2
Liu and Zaccarin [5]	24.4	28.5	35.9	24.6	21.8	12.5
Proposed method	24.9	29.1	36.3	24.5	22.7	7.2

표 2: MPEG-2 테스트 영상의 초기 100 프레임에 대한 다양한 탐색 방법의 성능 비교

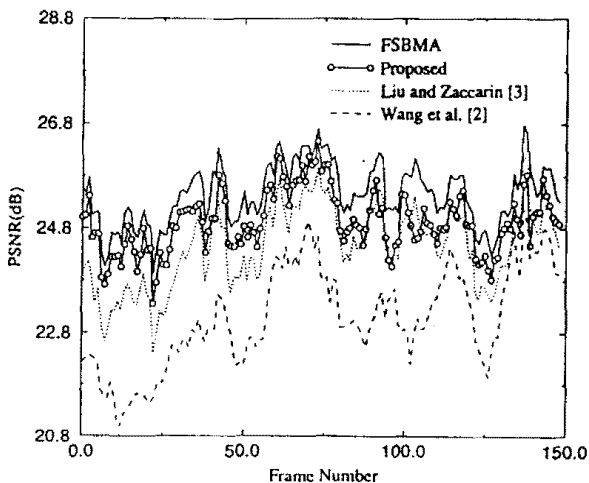
그림 3: MPEG-2 테스트 영상의 초기 100 프레임에 대한 PSNR 비교



(a) Car



(b) Cheerleaders



(c) Football