

# MPEG-2 TM5 부호기의 움직임 예측 처리 과정

김준기, 이호석

호서대학교 컴퓨터 공학과

## The analysis of motion estimation algorithm of MPEG-2 TM5 encoder

Jun-Ki Kim, Ho Suk Lee

Department of Computer Engineering, Hoseo University

### 요약

본 논문은 MPEG-2 TM 5 비디오 부호기의 움직임 예측(motion estimation) 처리과정을 소개한다. 비디오에는 공간의 중복성보다 시간의 중복성이 훨씬 많다. 따라서 시간의 중복성을 찾아내는 것이 압축의 효율을 높이는 중요한 척도가 된다. MPEG-2 부호기는 움직임 예측 알고리즘을 사용하여 시간의 중복성을 줄여 압축 효율을 높인다. 움직임 예측은 참조 블록의 위치로부터 원래 블록의 위치를 추정하여 최적의 움직임 벡터를 찾는 과정이다. MPEG-2에서의 움직임 예측은 full search 알고리즘을 사용하여 마지막으로 half pel로 산출한다. 본 논문에서는 MPEG-2 움직임 예측 과정을 frame estimation, field estimation, picture 타입에 따른 estimation, 움직임 예측을 위한 블록 매칭 알고리즘, full search 방법 및 움직임 벡터에 대하여 소개한다.

### 1. 서론

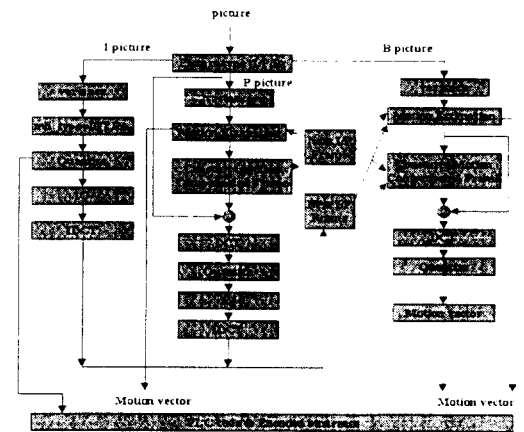
MPEG-2에서는 시간의 중복성(redundancy)을 줄이기 위하여 움직임 예측(motion estimation) 기법을 사용한다. 움직임 예측 기법은 frame 사이의 시간적인 중복성을 줄이기 때문에 비디오 압축에서 넓게 사용되어진다. 시간 축으로 연속된 picture 들은 주로 화면의 중앙부분에서 이미지의 움직임이 많기 때문에 움직임 예측에서는 이러한 성질을 이용하여 frame의 중복성을 제거한다. 즉 picture 사이에 가장 비슷한 블록을 찾는 것을 움직임 예측이라 하며, 얼마 만큼 이미지가 움직였는가 하는 변위를 나타내는 것을 움직임 벡터(motion vector)라고 한다. MPEG-2에서는 화면의 움직임을 찾아 전체 영상을 부호화 하는 것이 아니라 움직인 영상의 차이 만큼을 부호화 해서 데이터량을 큰 폭으로 줄이는 것이다.

움직임 예측을 위해 가장 널리 사용되는 알고리즘은 블록 매칭(block matching) 알고리즘이다. 즉 주어진 블록내의 모든 화소에 대하여 변위를 구하고 그 중 가장 작은 변위를 나타내는 탐색점의 값을 움직임 벡터로 추정하는 것을 full search 블록 매칭 알고리즘이라 한다. 이 방법은 탐색 영역 내에서 최적인 움직임 벡터를 추정할 수 있으나, 매칭 과정에서 많은 계산량을 필요로 하는 단점이 있다. 따라서 계산량을 줄이면서 거의 정확한 움직임 벡터를 찾을 수 있는 개선된 알고리즘에 대한 연구가 활발히 진행되고 있다. 이러한 알고리즘에는 고속 블록 매칭 알고리즘, 2차원 로그 탐색 알고리즘, 3단계 탐색 알고리즘, 교차 탐색 알고리즘, 개선된 3단 탐색 알고리즘 등이 연구되고 있다[1][4].

본 논문에서는 MPEG-2 TM5 부호기에서 사용하는 블록 매칭 알고리즘, full search 방법, 움직임 벡터 처리 과정을 각각의 picture 구조와 picture 타입으로 분리하여 소개한다.

### 2. 움직임 예측

MPEG-2 부호기의 움직임 예측 처리과정은 <그림 1> 과 같다[1].



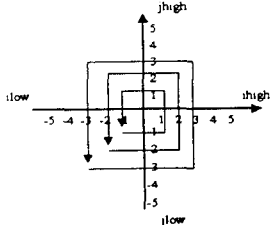
<그림 1> MPEG-2 부호기의 움직임 예측 처리과정

MPEG-2에서는 세가지 picture 형식을 지원하는데 움직임 예측을 사용하지 않는 I-picture(intra-coded picture)와 순방향 예측만을 지원하는 P-picture(predictive-coded picture) 그리고 순방향과 역방향 예측을 모두 사용하는 B-picture (Bidirectionally predictive-coded picture)가 있다. 또한 MPEG-2에서는 frame

picture 뿐만 아니라 field picture 까지 지원함으로써 더욱 압축의 효율성을 강조시켰다. 움직임 예측 과정은 각 picture 에서 매크로 블록 단위(16x16)단위로 처리한다[2].

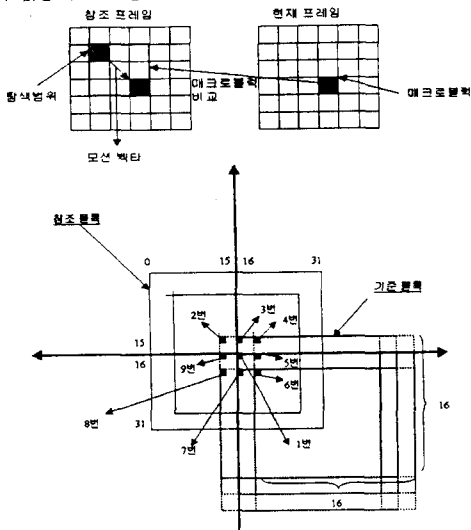
### 3. Full search 블록 매칭 알고리즘

Full search 블록 매칭 알고리즘은 입력 영상을 매크로 블록 (16x16) 단위로 나눈 다음, 기준 블록을 이전 frame 의 정해진 탐색 범위 내의 여러 개의 블록과 비교하여 가장 닮은 블록 (best matched block)을 찾아내 기준 블록에 대하여 매칭된 블록의 위치를 움직임 벡터로 정하는 기법이다. 부호기의 MPEG-2 파라미터 파일에는 각 picture 의 형식에 따라서 탐색 범위를 설정하여 준다. 즉 탐색 범위 안에서 어떤 블록이 현재 frame 의 정해진 블록으로 이동하였는가를 찾는 것이다. 블록 매칭을 위한 full search 화소의 위치 변화는 <그림 2>와 같다[1].



<그림 2> 탐색 범위 위치

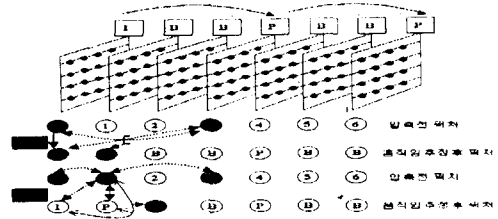
<그림 2>에서와 같이 탐색점의 위치는 픽처의 방향을 고려하여 각각의 좌표값 (-1,-1),(0,-1),(1,-1),(1,0),(1,1),(0,1),(-1,1),(-1,0)의 위치로 변화한다. MPEG-2 TMS 부호기에서 기준블록과 가장 닮은 매칭 블록을 찾아내는 방법으로는 MAD(mean absolute difference) 절대값의 합과 MSD(mean square difference) 차의 제곱의 합으로서 처음의 원래의 영상과 참조 블록사이의 full pel 탐색으로 MAD의 값을 구한다. 더욱 정확한 움직임을 구하기 위하여 half pel 의 방법으로 MSD 를 구하여 움직임 벡터의 값을 구하게 된다. 이 모든 과정은 <그림 3>과 같다.



<그림 3> 움직임 예측 과정

### 4. Frame 움직임 예측

Frame 에서 I,B,P-picture 의 움직임을 예측하는 과정은 <그림 4>와 같다[1][3].



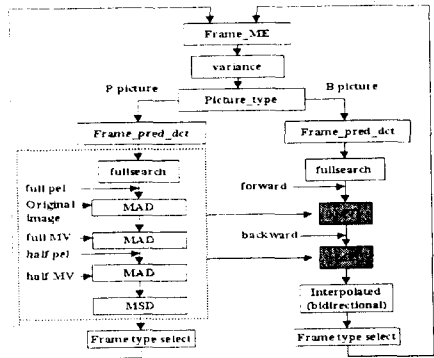
<그림 4> frame 움직임 예측

#### 4.1 I-picture

Frame, field 움직임 예측에서 I-picture 는 시간적 압축 기술을 사용하지 않고 공간적 압축 기술만을 사용하므로 움직임 예측 과정을 거치지 않는다. I-picture 는 다른 picture 의 참조 picture 로써의 역할을 한다. <그림 1>과 같이 I-picture 는 매크로 블록 단위의 편차(variance) 값을 저장하고 현재 블록의 매크로 블록 타입을 MB\_INTRA 라고 설정한다

#### 4.2 B,P-picture

P-picture 최근의 I-picture 나 P-picture 의 참조 picture 로 움직임 예측 기술을 사용한다. <그림 4>와 같이 첫번째 움직임 예측은 원래의 영상인 0번 picture 와 3번 picture 에 대한 full pel 블록 매칭을 적용한 다음 다시 half pel 로 I-picture 와 3번 picture 에 대한 최소 편차 값을 찾는다. 이후 P-picture 가 생성된다. B-picture 는 순방향 뿐만 아니라 역방향 예측을 한 picture 이다. 두번째 움직임 예측은 0번과 1번의 full pel 블록 매칭후 I-picture 와 1번 picture 에 대한 half pel 매칭 다음으로 1번과 3번의 full pel 블록 매칭과 1번과 최근의 P-picture 와의 half pel 매칭후 I,P,1 의 보간을 적용하여 B-picture 를 만들어 낸다. TMS 부호기에서의 P, B-frame picture 의 움직임 예측 블록도는 <그림 5>와 같다.



<그림 5> P,B-frame 움직임 예측 블록도

블록 매칭 움직임 예측은 최소 비용함수에 의해 이전에 처리된 frame 과 현재의 frame 사이에서 움직임 벡터를 얻는 것이다. 각각의 블록에서 최적의 매칭을 찾는 함수로는 MAD 와 MSD 를 사용하며 식은 다음과 같다[2][3].

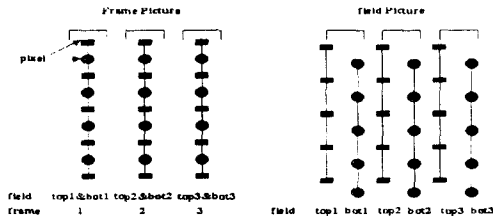
$$MSD(dx, dy) = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [F(i, j) - G(i+dx, j+dy)]^2$$

$$MAD(dx, dy) = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} |F(i, j) - G(i+dx, j+dy)|$$

dx, dy = motion vector of the 탐색 위치  
 F(i,j) = m x n 의 매크로 블록 위치, G(i, j) = m x n 의 참조 매크로 블록

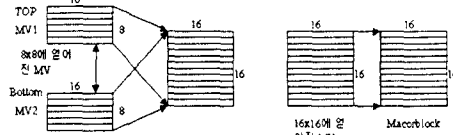
5. Field 움직임 예측

MPEG-2는 영상을 얻는데 있어서 프로그래시브와 인터레이스의 두가지 방법이 있다. 프로그래시브는 순차적으로 주사하는 것이고 인터레이스는 인접한 두 라인 간의 시간차가 있으므로 이를 두개의 필드를 top과 bottom으로 나누는 것이다. MPEG-2에서는 2개의 picture 구조의 선택을 제공한다[2].



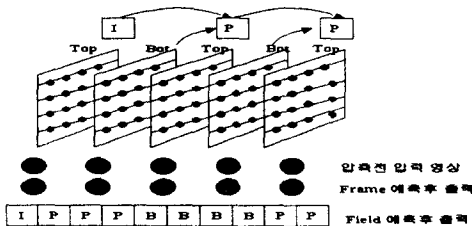
<그림 5> frame picture와 field picture 구조

<그림 5>와 같이 압축되기 많은 영상으로부터 입력을 받아 I.P.P.B.B.B.B.P.P.B.B의 형태로 출력되어진다. 즉 한 frame의 입력영상을 받아 상위 하위의 두 field로 처리되어진다. Field picture에서도 frame picture와 마찬가지로 I-picture는 움직임 예측 과정을 거치지 않는다. <그림 8>의 1번에서와 같이 I-picture는 0번 원래의 영상 TOP field로부터 매크로 블록 타입을 MB\_INTRA로 설정한 다음 바로 1 top picture가 만들어진다. P-picture는 움직임 예측 과정을 거치는데 <그림 8>의 2번에서와 같이 0번 원래 영상의 상위 field와 하위 field 간의 예측과정을 거친 후 P-picture의 영상을 만든다. 0번의 입력영상으로부터 I.P-picture를 만들고 다음 P-picture를 예측하기 위하여 3 field의 입력 영상을 읽어 들이고 같은 상위 field 들을 예측하므로써 3번의 P-picture를 만든다. 4번의 하위 P-picture를 만들기 위해서는 가장 최근에 예측하여 만든 영상 3번 field의 top과 P-picture의 top field로서 4번 P-picture가 생성되어진다.



<그림 6> frame과 field 움직임 예측

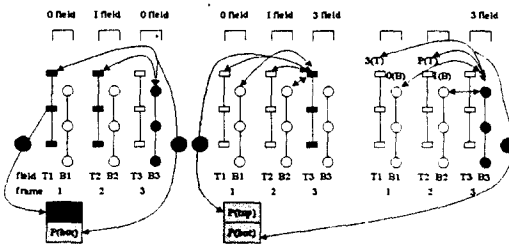
크로 블록을 <그림 6>과 같이 아래 위의 16x8 블록으로 나누어서 각각의 블록에 대하여 1개씩의 움직임 벡터를 예측한다. 두개의 Top과 Bottom field는 각각 분리되어 처리되어진다. TMS 부호기의 field 움직임 예측 과정은 <그림 7>과 같다.



<그림 7> 필드 예측 과정

5.1 I,P-picture 예측

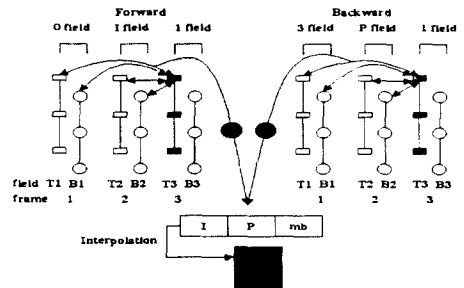
I-picture와 P-picture의 처리 구조는 <그림 8>과 같다.



<그림 8> I,P-picture 예측 처리

MPEG-2 field 예측의 처리결과와는 frame picture와 달리 <그림 7>과 같이 압축되기 많은 영상으로부터 입력을 받아 I.P.P.B.B.B.B.P.P.B.B의 형태로 출력되어진다. 즉 한 frame의 입력영상을 받아 상위 하위의 두 field로 처리되어진다. Field picture에서도 frame picture와 마찬가지로 I-picture는 움직임 예측 과정을 거치지 않는다. <그림 8>의 1번에서와 같이 I-picture는 0번 원래의 영상 TOP field로부터 매크로 블록 타입을 MB\_INTRA로 설정한 다음 바로 1 top picture가 만들어진다. P-picture는 움직임 예측 과정을 거치는데 <그림 8>의 2번에서와 같이 0번 원래 영상의 상위 field와 하위 field 간의 예측과정을 거친 후 P-picture의 영상을 만든다. 0번의 입력영상으로부터 I.P-picture를 만들고 다음 P-picture를 예측하기 위하여 3 field의 입력 영상을 읽어 들이고 같은 상위 field 들을 예측하므로써 3번의 P-picture를 만든다. 4번의 하위 P-picture를 만들기 위해서는 가장 최근에 예측하여 만든 영상 3번 field의 top과 P-picture의 top field로서 4번 P-picture가 생성되어진다.

5.2 B-picture 예측



<그림 9> field picture 예측 처리

B-picture의 처리과정은 <그림 9>와 같다. B-picture는 입력된 picture로부터 전방과 후방예측을 하고 이 두 picture의 보간을 통하여 생성되어진다.

6. 움직임 벡터

움직임 벡터는 한 매크로 블록당 최대 4개까지 나올 수 있다. 하지만 너무 비트량이 많으므로 바로 전 매크로 블록의 움직임 벡터와의 차이만을 부호화 한다. 즉, PMV(Prediction Motion vector)에 이전 매크로 블록의 움직임 벡터 값을 저장한다. 슬라이스의 시작에 PMV=0으로 설정되어진다. 최소의 비트율을 위하여 DPCM 과정을 거친다[4].

7. 결론

본 논문에서는 MPEG-2 TMS 부호기의 움직임 예측 전체 처리과정을 소개하였다. 움직임 예측의 단위는 매크로 블록단위이며 full search 블록 매칭 알고리즘을 사용하여 움직임 벡터를 계산한다. 움직임 예측은 frame picture, field picture에 따라서 다른 방법을 사용한다. 그리고 picture의 타입에 따라 다른 움직임 예측 방법을 적용한다. I-picture는 참조 picture의 역할을 함으로 움직임 예측을 하지 않으며 P-picture와 B-picture는 움직임 예측을 하여 picture의 움직임 벡터를 계산한다.

8. 참고 문헌

[1] MSSG, "MPEG-2 Source Code", 1994.  
 [2] Barry G.Haskell, Atul Puri, and Arun N.Netravali, Digital Video: An Introduction to MPEG-2, Chapman & Hall, 1997.  
 [3] Joan L.Mitchell, William B.Pennebaker, Chad Fogg, and Didier J.LeGall, MPEG Video Compression Standard, Chapman & Hall, 1997.  
 [4] ISO/IEC, "MPEG-2 Systems Working Draft", 1995.