

MPEG의 양방향 이동벡터 추정을 위한 고속 블럭정합방식

正會員 金旻圭*, 金在均**

Image Compression using Hybrid Structure with Multi-Layer Perceptron and SOFM Coding

Min Kyu Kim*, Jae Kyoon Kim** Regular Members

要 約

양방향 이동보상예측은 매우 뛰어난 예측효율을 얻을 수 있지만, 이를 위한 양방향 이동 벡터추정은 엄청난 계산량이 요구된다. 본 논문에서는 MPEG-비디오 부호화 표준안의 양방향예측을 위한 새로운 블럭정합방식을 제안하는데, 이 방식은 적은 계산량으로 보다 정확한 순방향 및 역방향 이동벡터를 추정한다. 블럭정합방식(BMA)을 보다 가능성이 있는 후보블럭의 위치로부터 시작한다면, 추정된 이동벡터는 보다 작은 탐색영역 안에서 보다 빨리 실제 이동방향으로 수렴할 것이다. 이러한 개념에 근거하여 보다 신뢰도가 높은 초기 이동벡터를 추정하기 위하여, 제안된 방식에서는 시공간 방향에서 인접한 이동벡터는 높은 상관성을 갖는다는 성질과 MPEG의 양방향 이동벡터추정에 적합한 등속이동모델을 이용한다. MPEG-2의 TM5을 이용한 모의실험 결과, 제안된 방식은 매우 적은 계산량으로도 예측성능이 가장 우수한 전탐색 방식과 거의 같은 예측 및 부호화 성능을 얻었으며, 또한 보다 자연스러운 움직임 특성과 좋은 주관적인 화질을 얻을 수 있었다.

ABSTRACT

Even though the bi-directional motion-compensated prediction can obtain a remarkable prediction efficiency, the bi-directional motion estimation is a formidable computational challenge. This paper presents a novel block matching algorithm(BMA) for the bi-directional prediction of MPEG-video coding standard, which can estimate more accurate forward and backward motion vectors with low computational complexity. If BMA starts from more probable candidate location, the estimated motion vector converges faster to the true movement within a small search range. Our proposed algorithm is based on this concept. To estimate more reliable initial motion vector, the proposed algorithm exploits the property of high correlation among the motion vectors in

*대우전자 VCR연구소 뉴미디어개발부 DVCR연구팀
New Media Department, VCR Research Center DAE-
WOO ELECTRONICS CO., LTD

**한국과학기술원 전기 및 전자공학과
Dept. of Electrical Engineering
Korea Advance Institute of Science and Technology(KAIST)
論文番號 : 94258
接受日字 : 1994年 9月 27日

the spatio-temporal direction and transnational motion model suitable for bi-directional motion estimation. Compared with the full search algorithm, experimental results based on MPEG-2 TM-5 show that the proposed algorithm achieves almost the same prediction and coding performance with very low computational complexity, and also obtains slightly better subjective quality and motion rendition.

I. 서론

디지털 저장매체를 위한 ISO-IEC/MPEG(Moving Picture Experts Group)의 동영상 압축기법(MPEG-1⁽¹⁾ 및 MPEG-2^(2,3,4))을 실시간 구현함에 있어서 가장 큰 걸림들은 블럭단위의 이동벡터 추정에 요구되는 방대한 계산량이다. MPEG-비디오(이하 MPEG)에서는 임의접근(random access)과 영상데이터의 고압축율이라는 서로 상충되는 요건을 만족시키기 위하여 I-픽처(intra-coded picture), P-픽처(pure predictive-coded picture), 그리고 B-픽처(bi-directionally predictive-coded picture)로 구분되는 세 종류의 픽처를 주기적으로 사용한다.

그 중에서 B-픽처는 MPEG의 가장 큰 특징으로서, B-픽처 전후에 있는 두 기준픽처(I 또는 P)를 이용하여 다음의 세가지 모드 중의 하나로 이동보상 예측(motion-compensated prediction:MCP)된다: 이전 픽처로부터의 순방향 MCP, 이후 픽처로부터의 역방향 MCP, 그리고 전후의 두 기준픽처로부터의 양방향 MCP. Memory와 coding delay 등의 증가에도 불구하고, B-픽처의 세 MCP모드는 covered/uncovered area에서 예측 효율이 뛰어나거나 잡음의 영향을 줄일 수 있는 장점들이 있다. 따라서 전후의 두 기준픽처 사이에서 B-픽처의 사용 빈도에 비례하여 전체적인 부호화 성능이 향상된다. 그러나 B-픽처의 사용 빈도가 증가하면 블럭단위의 이동벡터추정(Motion estimation:ME) 기법인 블럭정합방식(block matching algorithm:BMA)을 사용함에 있어서 다음과 같은 문제점들이 야기된다. 첫째, P-픽처와 달리 B-픽처는 양방향 이동벡터를 추정해야 하기 때문에 계산량이 증가한다. 둘째, 전후의 기준픽처와 B-픽처 또는 전후의 두 기준픽처 사이의 픽처간격이 증가하기 때문에 BMA의 탐색영역이 증가한다. 따라서 BMA의 계산량은 픽처간의 제곱에 비례하여 증가한다. 셋째, ME의 대상이 되는 두 픽처 사이의 간격이 증가하면 할수록

시간방향의 상관성(temporal correlation)이 떨어지기 때문에 예측 효율과 추정의 정확도가 급격히 떨어진다.

B-픽처가 사용되는 경우, 이동벡터추정은 크게 B-픽처의 세 예측모드 모두에서 최적의 예측효율을 얻기 위한 방법^(6,6)과 B/P 픽처의 양방향 이동벡터를 적은 계산량으로 신속하게 추정하는 방법^(7,8)으로 나눌 수 있는데, 본 논문은 후자에 관한 것이다. BMA의 계산량을 줄이기 위한 기존의 고속 알고리즘들로는 3-단계 탐색⁽⁹⁾, 2-D 대수 탐색⁽¹⁰⁾ 등과 같이 탐색의 대상이 되는 후보블럭의 수를 제한하는 방식이 있다. 그러나 이들 알고리즘들은 정합점이 최소왜곡의 방향(direction of minimum distortion:DMD)으로부터 멀어질수록 예측오차가 증가한다는 가정에 근거하지만 실제 이러한 가정이 잘 맞지 않기 때문에, 이동벡터가 local optimum으로 수렴하는 오정합이 많이 발생한다. 따라서 DMD에 근거한 기존 방식들을 MPEG의 ME에 적용하기는 부적합하다.

본 논문에서는 MPEG과 같이 IBP 픽처가 주기적으로 사용되는 이동보상 예측 구조에서 실시간 구현이 가능할 정도의 매우 적은 계산량으로 보다 효과적인 양방향 이동벡터를 추정하기 위한 고속 블럭정합방식을 제안한다. 제안된 방식은 이동변위 제한식과 이동벡터의 시공간 상관성(spatio-temporal correlation)을 이용하여 각 블럭에 대한 BMA의 탐색영역, 즉 계산량을 적용적으로 줄인다. MPEG-2의 Test Model 5(TM5)(3)를 통한 모의실험 결과, 제안된 방식은 계산량, 예측효율, 이동벡터의 정확도, 그리고 부호화 성능 등에서 우수한 성능을 보였다.

II. MPEG을 위한 기존의 블럭정합방식

MPEG에서 임의접근의 단위에 해당하는 GOP(group of picture)의 크기와 인접한 두 기준픽처 사이의 B-픽처 수는 두개의 변수 (N,M)으로 나타내는데, 그림 1은 N=12, M=3일 때의 GOP 구

조를 나타낸 것이다. 그림에서 N는 GOP의 크기-즉 연속하는 두 I-픽처 사이의 픽처간격을, M은 sub-GOP의 크기- 즉 연속하는 두 기존픽처(I 또는 P)사이의 픽처간격을 의미한다. 따라서 두 기존픽처 사이에는 (M-1)개의 B-픽처가 존재한다. I-픽처는 이동보상을 하지 않는 원영상 데이터 만을 이용하여 부호화된다. 따라서 다른 픽처의 영향이 없기 때문에 부호화된 비트열로의 임의접근을 가능하게 하는 접근점(access point) 역할을 하며, 부호화 과정에서 발생하는 부호화 오류(coding error)가 후속 픽처로 전파되는 현상(error propagation)을 방지할 수 있다. P-픽처는 이전의 I-픽처 또는 P-픽처로 부터의 인과적(causal) 예측에 의한 이동보상(motion compensation:MC)후에 부호화된다. B-픽처는 이전 또는 이후의 I-픽처나 P-픽처로 부터의 비인과적(non-causal) 예측에 의한 이동보상 후에 부호화 된다.

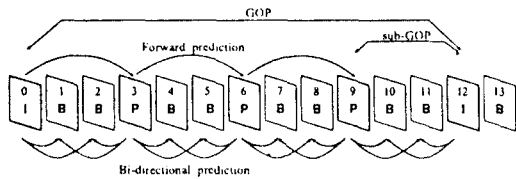


그림 1. MPEG-비디오의 GOP 구조(N=12, M=3)
Fig. 1. GOP structure of MPEG-video(N=12, M=3)

한편, 그림 2와 표1은 GOP내의 한 sub-GOP(I₀B₁B₂...B_{M-1}P_M)에서 B-픽처와 P-픽처에 대한 이동벡터 및 이동보상예측 관계식을 나타낸 것이다. 그림에서 v_f(m,n)(n>m)은 m번째 픽처로 부터 n번째 픽처로의 순방향 이동벡터(forward motion vector:FMV)를, v_b(n,m)은 역방향 이동벡터(backward MV:BMV)를 의미한다. B-픽처의 각 블럭은 표 1의 세 MCP 모드 중에서 예측오차의

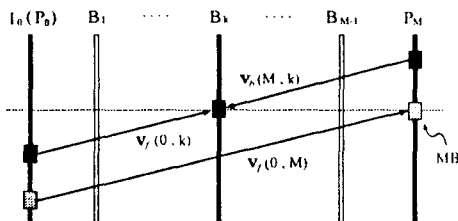


그림 2. sub-GOP의 B_k/P_k픽처에서의 FMV와 BMV
Fig. 2. FMV and BMV of B_k/P_k pictures in sub-GOP

표 1. B_k/P_k 픽처의 이동보상 관계식
Table 1. MCP equation of B_k/P_k pictures

Pictures	MCP mode	Equation
P _m	Forward	I(z, M)=I(z+v _f (0, M), 0)
B _k	Forward	I(z, k)=I(z+V _f (0, k), 0)
	Backword	I(z, k)=I(z+V _b (M, K), M)
	Bi-directional	I(z, k)=½ · I(z+V _b (M, K), M) +½ · I(z+V _b (M, K), M)

SSE(sum of square error)가 최소인 것을 그 블럭의 최종 MCP 모드로 결정한다.

MPEG에서는 macroblock(MB)이라 불리는 16×16크기의 블럭 단위로 이동벡터를 추정하는데, 블럭 단위의 이동벡터 추정 방식인 BMA는 예측 효율과 추정의 정확도, 계산상의 복잡도, 그리고 이동벡터 표현을 위한 데이터량 등의 서로 상충되는 요건들을 대체로 잘 만족시키며, 특히 MPEG과 같은 블럭 단위의 부호화기법에서 구현이 용이하기 때문에 널리 이용한다.

이제 MPEG의 양방향 MV를 추정하기 위한 기존의 대표적인 BMA 방식과 그 문제점을 살펴보면 다음과 같다.

1. 전탐색 방식(full search algorithm:FSA)^(1,2,3)

BMA에서 최대 계산량을 제한하기 위하여 미리 설정된 최대 이동변위 내에서 이동벡터를 추정하는데, 주어진 탐색 영역 내의 모든 블럭과 비교하여 이동벡터를 추정하는 전탐색 방식은 예측 이득의 측면에서 최고의 성능을 갖지만 지나치게 많은 계산량이 요구된다. 예를 들면, K×L 크기의 블럭에 대하여 픽처당 최대 이동변위를 ±p pels/frame이라고 할 경우, 기존픽처에서 탐색영역의 크기는 (K+2p)×(L+2p)가 된다. 즉 정합기준에 따라 비교해야 할 후보 블럭의 수가 C₀=(2p+1)²만큼 된다. 픽처 간격이 k일 때 탐색영역 내의 후보 벡터의 집합 R(k)는 다음과 같다.

$$R(k) = \{ (dx, dy) | -k \cdot p \leq dx, dy \leq k \cdot p \} \quad (1)$$

따라서 픽처간격이 k 일 때, 계산의 복잡도(computational complexity)에 해당하는 정합블럭의 수는 다음과 같다.

$$\begin{aligned}
 C_k &= (2 \cdot k \cdot p + 1)^2 \\
 &\approx k^2 \cdot 4p^2 \\
 &\approx k^2 \cdot C_0
 \end{aligned}
 \tag{2}$$

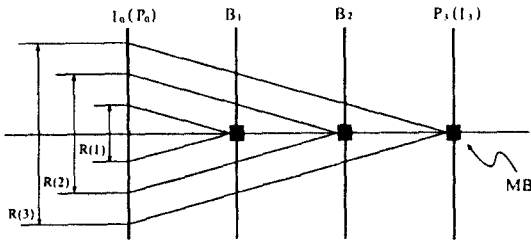


그림 3. 전탐색 방식(FSA)을 위한 탐색영역(M=3)
Fig. 3. Search range of full search algorithm(FSA)(M=3)

위 식에서 픽처 간격이 k 배 증가하면 계산량은 k^2 배로 증가함을 알 수 있다. 그림 3은 sub-GOP내의 각 픽처에 대한 순방향 이동벡터 추정(FME)할 때 FSA 방식의 탐색영역을 나타낸 것이다. 역방향 ME(BME)도 이와 유사하다. 따라서 P와 P, 또는 I와 P사이의 픽처간격이 M 일 때, sub-GOP내에서 동일한 위치의 각 블럭에 대한 전체 계산의 복잡도(=BMA의 정합블럭 수) C_{FSA} 는 다음과 같다.

$$C_{FSA} = C_0 \cdot \left\{ M^2 + \sum_{k=1}^{M-1} [k^2 + (M-k)^2] \right\} \tag{3}$$

식(3)에서 M^2 는 P_M 픽처의 FMV에 관한 항이고, k^2 과 $(M-k)^2$ 은 각각 B_k 픽처의 FME 및 BME에 관한 것이다. 따라서 p 또는 M 이 커질 경우, 계산량의 증가로 인하여 BMA의 실시간 구현이 용이하지 않게 된다.

2. Telescopic 탐색 방식(TSA)^(1,2,3,7)

영상 간에 운동체의 이동변위는 매우 높은 상관성을 갖는데, TSA 방식은 이러한 상관성을 이용하여

탐색영역을 줄이는 방식으로 MPEG에서 가장 널리 이용되는 고속 알고리즘이다. 그림 4는 M=3일 때의 FMV의 추정과정을 나타낸 것으로서 간략히 설명하면 다음과 같다.

- Step 1 : 기준픽처 P_0 와 현재 픽처 B_1 사이에서 B_1 픽처의 MB에 대한 이동벡터 $\mathbf{v}_r(0,1)$ 을 추정한다.
- Step 2 : $\mathbf{v}_r(0,1)$ 을 초기치로 하여 기준픽처 P_0 와 B_2 픽처 사이에서 B_2 픽처의 MB에 대한 이동벡터 $\mathbf{v}_r(0,2)$ 를 추정한다.
- Step 3 : $\mathbf{v}_r(0,2)$ 을 초기치로 하여 기준픽처 P_0 와 P_3 픽처 사이에서 P_3 픽처의 MB에 대한 이동벡터 $\mathbf{v}_r(0,3)$ 를 추정한다.

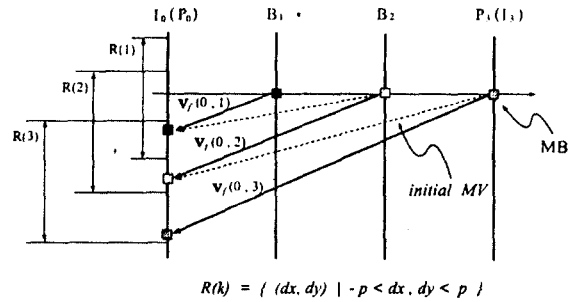


그림 4. Telescopic 탐색 방식(TSA)의 FMV 추정 과정(M=3)
Fig. 4. Forward ME of telescopic search algorithm (TSA)(M=3)

TSA 방식은 ME의 픽처간격에 관계없이 항상 일정한 탐색영역을 갖기 때문에 계산량은 C_0 와 같다. 또 TSA 방식의 각 단계에서 추정된 MV들이 바로 해당되는 픽처의 MV로 사용된다. BMV도 같은 방식으로 추정된다. 따라서 sub-GOP에 대한 전체 계산의 복잡도 C_{TSA} 는 다음과 같다.

$$C_{TSA} = C_0 \cdot \{ 2 \cdot (M - 1) + 1 \} \tag{4}$$

여기서 첫항은 B_1, B_2, \dots, B_{M-1} 에 대한 것이고, 두번째 항은 P_M 에 대한 것이다. 그러나 TSA 방식은 각 방향에서 순차적으로 ME를 해야하기 때문에 최악의 경우 M 배의 계산량이 요구된다. 또한 TSA 방식 역시 각 추정단계에서 픽처간격 $k=1$ 에 해당되는 탐색영역에서 전탐색 방식을 이용하기 때문에 전탐색 방식의 범주에 속하며 그 계산량도 방대하다. 특히 MPEG-2

의 경우 최대 이동 변위 p 가 ± 14 pels/frame 이상의 넓은 영역에서 추정하고 각 MB에 대한 필드 MV와 프레임 MV를 추정해야 하기 때문에 실시한 구현이 어렵다. 한편 운동체의 움직임이 크거나 에지(edge)에서는 운동체의 움직임계적 추정(motion trajectory tracking)이 용이하지 않기 때문에 이전 단계에서 추정된 부정확한 초기 MV 주위의 local minimum 방향으로 수렴하는 오정합이 발생한다.

이외에도 MPEG을 위한 ME 방식으로 3차원 이동계적 추정 방식(3-D motion trajectory tracking method)^[7], optical flow 추정 방식^[8] 등이 제안되었다. 그러나 이들 방식 역시 움직임이 빠른 영역에서 추적이 어렵거나 계산량이 증가하는 단점이 있다. 따라서 FSA 방식이나 TSA 방식에 비교될 만한 예측 효율을 얻을 수 있으면서, 한편으로 실시간 구현이 가능한 고속 알고리즘이 요구된다.

Ⅲ. 이동벡터의 시공간 상관성을 이용한 고속 블럭정합방식

영상 블럭의 이동 방향에 대한 정보를 미리 알 수 있으면, BMA에서 추정의 대상이 되는 탐색영역을 보다 효과적으로 줄일 수 있다. 블럭의 이동 방향에 대한 사전 정보(prior information)란 곧 추정될 블럭의 이동벡터에 대한 초기 벡터(initial vector)의 설정을 의미한다. 초기 이동벡터가 실제 움직임에 보다 접근하게 설정될수록 탐색영역, 즉 추정의 계산량을 더욱 효과적으로 줄일 수 있다. 일반적으로 이미 추정된 이동벡터들의 시공간 상관성을 이용한 초기 이동벡터 설정 방식이 널리 이용된다.^[11]

본 논문은 이러한 특성을 이용하여 MPEG과 같이 양방향 ME를 해야 하고, ME의 대상이 될 픽처 간의 간격이 일정하지 않은 IBP 픽처를 포함하는 GOP 구조에 적합한 고속 블럭정합 방식을 제안한다. 제안된 방식은 다음의 초기 이동벡터 추정, 탐색영역 결정, 그리고 추가적인 ME의 세 과정을 통하여 B-픽처 및 P-픽처에 대한 FMV 및 BMV를 추정한다.

1. 초기 이동벡터 추정

블럭의 크기가 충분히 작다면 동일한 움직임을 갖

는 운동체는 여러 블럭으로 나누어질 수 있기 때문에, 영상내에서 이웃한 블럭들의 MV들은 매우 높은 상관성(spatial correlation)을 갖는다. 따라서 현재 블럭의 MV는 그림 5에 나타낸 바와 같이 이미 추정된 주위 블럭들(causal blocks)의 MV들(v_1, v_2, v_3, v^*)로 부터 예측될 수 있다. 그러나 현재 블럭의 -45° 대각 방향에 있는 블럭의 MV(v^*)는 현재 블럭의 수평 또는 수직 방향에 있는 MV들(v_1, v_2)와 대부분 일치하는 중복된 결과를 갖는다. 따라서 제안된 방식에서는 공간영역에서의 v^* 를 제외한 세 MV들(v_1, v_2, v_3)을 현재 블럭에 대한 초기 MV를 설정하기 위한 후보벡터로 이용한다.^[11]

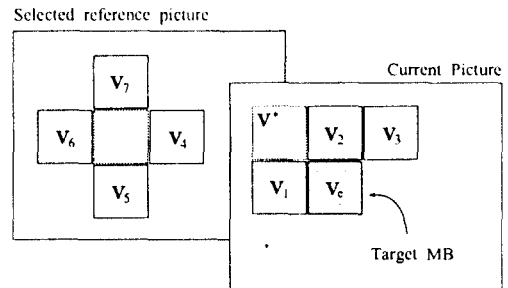


그림 5. 공간방향의 후보 MV와 시간방향의 추가적인 후보 MV
Fig. 5. Spatial candidate MVs and additional temporal candidate MVs

한편 영상 시퀀스에서 운동체의 움직임의 연속성으로 인하여, 시간 방향으로 이웃한 블럭들의 MV들 역시 매우 높은 상관성(temporal correlation)을 갖는다. 따라서 바로 이전 픽처에서 추정된 MV를 현재 블럭의 초기 MV로 사용할 수 있다. $M=1$ 인 경우, 즉 B-픽처를 사용하지 않는 IPPP...와 같은 GOP구조에서는 이동벡터가 추정될 픽처 간격이 항상 1이고 또 순방향 ME만을 하기 때문에, 이전 픽처의 MV를 현재 픽처의 ME를 위한 초기 벡터로 설정할 수 있다. 그러나 그림 2와 같이 $M>1$ 인 경우, 즉 하나 이상의 B-픽처가 사용될 경우 양방향 ME를 해야 하고, 또 현재 픽처와 기준픽처 사이의 간격이 일정하지 않기 때문에 바로 이전 픽처의 MV를 현재 픽처의 초기 벡터로 설정하는 것은 적합하지 않다.

따라서 제안된 방식에서는 영상 시퀀스의 이동 변위에 대한 속도 제한식(velocity constrained equation)으로 부터 MPEG에 적합한 시간 방향 초기 벡터를 설정한다. 등속도 및 가속도 운동을 포함

하는 영상에서 변위벡터는 다음과 같은 이동 모델로 표현할 수 있다.

$$d(p, n) = A \cdot (n - p)^2 + L \cdot (n - p) \quad (5)$$

여기서 $d(p, n)$ 는 p 번째와 n 번째 픽처 사이의 변위 벡터이며, A 와 L 은 연속하는 픽처간의 가속 및 등속 벡터이다. 두 픽처 간의 간격이 작을 경우에는 다음의 등속 모델로 제한할 수 있다.

$$d(p, n) = L \cdot (n - p) \quad (6)$$

식(6)을 그림 2와 같은 MPEG의 sub-GOP 구조에 적용하면 다음의 관계식을 얻을 수 있다.

$$v_f(0, k) = k \cdot v_f(0, 1) \quad (7)$$

$$v_f(0, M) = \frac{M}{k} \cdot v_f(0, k) \quad (8)$$

$$v_b(M, k) = -\frac{M - k}{k} \cdot v_f(0, k) \quad (9)$$

여기서 $v_f(0, k)$ 는 k 번째 픽처의 순방향 MV를, $v_b(M, k)$ 는 k 번째 픽처의 역방향 MV를 의미한다. 위의 관계식으로 부터 이미 추정된 다른 픽처의 MV들을 이용하여 현재 블럭에 대한 초기 이동벡터를 다음과 같이 설정한다.

(1) 순방향 초기 MV : v_f^o

순방향 MV는 인과적 예측이기 때문에 입력되는 영상의 순서(input order)에 따라 즉시 ME를 수행할 수 있다. 따라서 항상 이전의 픽처들의 MV들로부터 초기 MV, v_f^o 를 설정한다.

• $k=1$ 인 경우 : 현재의 sub-GOP내에 있는 첫번째 픽처에서의 이동벡터 추정이므로 바로 이전 sub-GOP에서 추정된 MV들을 이용하여 현재 블럭의 초기 MV 설정을 위한 후보 MV들을 구한다. 시간방향 후보 MV들은 식(10)과 같이 $v_f(0, 1)$ 에 가장 가까운 역방 MV($v_b^o(M, M-1)$)와 순방향 MV($v_f^o(0, M-1)$), 그리고 픽처간격이 같은 순방향 MV($v_f^o(0, 1)$)에 식

(6)의 등속 모델을 적용하여 구한다.

$$v_f^o(0, 1) \in \{ -v_b^o(M, M-1), \frac{1}{M-1} \cdot v_f^o(0, M-1), v_f^o(0, 1) \} \quad (10)$$

여기서 v_f^o 와 v_b^o 는 이전 sub-GOP의 MV들이다. 이전의 sub-GOP내에서 $v_f(0, 1)$ 에 가장 인접한 후보 벡터는 $v_f^o(0, M)$ 이지만, 현재의 sub-GOP내의 첫 픽처가 I-픽처일 경우 $v_f^o(0, M)$ 은 없기 때문에 이 MV는 초기 MV 설정을 위한 후보 MV에서 제외된다.

• $2 \leq k \leq M$ 인 경우 : 현재의 sub-GOP내에 이미 순방향 ME를 완료한 이전 픽처가 있으므로 그 FMV들을 이용하여 후보 MV들을 구한다. 시간방향 후보 MV들은 식(11)과 같이 추정하고자 하는 현 블럭의 MV($v_f(0, k)$)에 가장 가까운 MV($v_f(0, k-1)$)와 픽처 간격이 1일 때의 MV($v_f(0, 1)$)에 식(6)의 등속 모델을 적용하여 구한다.

$$v_f^o(0, k) \in \{ k \cdot v_f(0, 1), \frac{k}{k-1} \cdot v_f(0, k-1) \} \quad (11)$$

(2) 역방향 초기 MV : v_b^o

역방향 ME는 비인과적인 예측이기 때문에 B_k 픽처의 역방향 ME는 이후의 기준픽처 P_M 픽처(또는 I_M -픽처)가 입력될때 까지 기다린 후 입력의 역순으로 ME를 수행한다.

• $k=M-1$ 인 경우 : 현재의 sub-GOP 내에서 역방향 MV를 추정하는 첫번째 픽처이므로 이미 추정된 현재 또는 이전 픽처의 순방향 MV만을 이용할 수 있다. 시간방향 후보 MV들은 식(12)와 같이 추정하고자 하는 현 블럭의 순방향 MV($v_f(0, M-1)$)와 픽처 간격이 1일 때의 순방향 MV($v_f(0, 1)$)에 식(6)의 등속 모델을 적용하여 구한다.

$$v_b^o(M, M-1) \in \{ (0, 0), -\frac{1}{M-1} \cdot v_f(0, M-1), -v_f(0, 1) \} \quad (12)$$

• $1 \leq k \leq M-2$ 인 경우 (M)2) : 현 sub-GOP내에서 이미 추정된 순방향 및 역방향 MV들을 이용할 수 있다. 시간방향 후보 MV들은 식(13)와 같이 추정하고자 하는 현 블록의 순방향 $MV(\mathbf{v}_f(0, k))$ 와 가장 가까운 역방향 $MV(\mathbf{v}_b(M, k+1))$, 그리고 픽처 간격이 1일때의 역방향 $MV(\mathbf{v}_b(M, M-1))$ 에 식(6)의 등속 모델을 적용하여 구한다.

$$\mathbf{v}_b^o(M, k) \in \left\{ \begin{aligned} & (M-k) \cdot \mathbf{v}_b(M, M-1) \\ & \frac{M-k}{M-k-1} \cdot \mathbf{v}_b(M, k+1), \\ & - \frac{M-k}{k} \cdot \mathbf{v}_f(0, k) \end{aligned} \right\} \quad (13)$$

M=2인 경우에 위의 식(10)~식(12)에서 우변의 마지막 두 후보 MV는 동일한 것이기 때문에 시간방향 후보벡터는 한개 줄어든다. 그리고 식(13)에서는 M=3인 경우에 우변의 마지막 두 후보 MV가 동일한 값을 갖는다.

한편 블록에 moving edge가 포함될 때 같은 위치에 있는 이전 픽처의 MV들로는 적합한 초기 MV를 추정할 수 없다. 이와 같은 경우에는 그림 5에 나타난 바와 같이 시간방향에서 주위 블록의 MV들을 이용할 때 보다 정확한 MV를 추정할 수 있다. 제안된 방식에서는 MV들의 상관성을 고려하여 좌우상하에 있는 네 개의 MV, $(\mathbf{v}_4, \mathbf{v}_5, \mathbf{v}_6, \mathbf{v}_7)$ 을 이용한다.

이상의 관계식으로 부터 제안된 방식의 초기 MV 추정 과정을 정리하면 다음과 같다.

- Step 1 : 공간방향의 세 후보 MV(그림 5)와 시간방향의 두 개 또는 세 개의 후보 MV(식(10)~식(13))들을 ME를 하고자 하는 현재의 블록에 적용하여 이동보상 예측오차의 SAE(sum of absolute error)가 최소인 것을 잠정적인 초기 MV로 선택한다. 그리고 선택된 초기 MV로부터의 SAE를 SAE_0 라고 정의한다.
- Step 2 : Step 1에서 선택된 현재 블록의 초기 MV를 이용한 이동보상 예측오차 SAE_0 과 이동벡터 추정이 완료된 시공간 방향의 이웃한 블록들의 SAE로부터 Step 1에서 선택된 초기 MV의 적합성 여부를 아래와 같이 판단한다.

if (SAE0(T₁))

→ Step 1에서 선택된 MV를 최종 초기 MV로 결정

else

→ 그림 5의 시간방향의 네 후보 MV($\mathbf{v}_4, \mathbf{v}_5, \mathbf{v}_6, \mathbf{v}_7$)로 부터의 SAE와 Step 1에서 선택된 잠정적인 초기 MV로부터의 SAE_0 를 비교한 후, 최종 초기 MV 결정

여기서 T₁은 Step 1에서 선택된 초기 MV의 적합성 여부를 결정하는 기준으로서, 현재 블록의 SAE를 제외한 현재 블록의 초기 MV 설정에 이용된 시공간 후보 MV들이 소속된 블록들의 이동보상 예측오차의 SAE 중에서 최대값을 취한다. 즉 Step 1에서 선택된 초기 MV를 이용한 현재 블록의 SAE_0 가 시공간 방향의 이웃한 블록의 SAE의 최대치 T₁보다도 크다면 Step 1에서 선택된 초기 MV가 현재 블록의 실제 MV와 큰 차이를 갖을 가능성이 많다. 이 경우에는 오정합이 발생할 가능성이 크기 때문에 Step 2에서와 같은 추가적인 초기 MV 설정 과정이 필요하다. 이와 같은 판단의 근거는 이미 이동벡터 추정이 완료된 시공간 방향의 이웃한 블록들은 실제 움직임에 가장 근접하는 MV와 최적의 이동보상 예측오차를 갖는다고 가정할 수 있기 때문이다. 또 그림 5의 추가적인 시간방향 후보 MV들은 식(10)~식(13)의 시간방향 후보 MV들 중에서 최소의 SAE를 갖는 MV가 포함된 픽처의 MV들을 식(7)~식(9)의 등속 모델에 적용하여 구한다.

2. 탐색영역결정

추정된 초기 MV의 정확도는 탐색 영역의 범위를 결정한다. FSA 방식은 초기 MV로 원점, 즉 $\mathbf{v}^0 = (0, 0)$ 를 설정하기 때문에 탐색영역이 식(3)처럼 픽처 간격의 제곱에 비례하여 증가하고, TSA 방식은 초기 MV로 이전 픽처의 MV를 사용하기 때문에 식(4)처럼 항상 픽처 간격 1에 해당되는 탐색 영역을 갖는다. 본 논문에서 제안된 방식은 실제 움직임에 근접하는 초기 MV를 설정하기 때문에 탐색영역을 대폭 줄일 수 있다.

제안된 방식에서는 추정된 초기 MV의 정확도에

따라 픽처 간격 1일 때의 최대 이동 범위 p 의 $\frac{1}{3}$ 범위 내에서 탐색영역을 가변적으로 선택한다. 이미 설정된 초기 MV에 대한 정확도는 초기 MV의 결정과정과 마찬가지로 초기 MV를 이용한 현재 블럭의 SAE와 초기 MV의 후보로 이용된 시공간 방향의 이웃한 블럭들의 SAE들을 비교하여 판단한다. 먼저 설정된 초기 MV로 부터 이동보상된 현재 블럭의 예측오차의 SAE를 SAE_{init} 라 하고, 현재 블럭을 제외한 초기 MV 추정 과정에서 후보 MV로 이용된 시공간 방향의 이웃한 각 블럭들에 각각의 MV를 적용하여 계산한 이동보상 예측오차의 SAE중에서 최대값을 SAE_{max} , 최소값을 SAE_{min} 라고 정의한다. 그리고 현재 블럭의 SAE_{init} 를 이웃한 블럭의 SAE_{max} 및 SAE_{min} 과 비교하여 탐색영역을 결정한다. 이때 초기 MV를 중심으로 한 새로운 탐색영역 내의 이동변위, p' 는 다음과 같이 선택된다. ($p = \pm 14$ pels/frame 인 경우)

```

if (SAEinit < T0)                p'=1;
else if (SAEinit < 2×SAEmin OR
        SAEinit < 0.7×SAEmax)    p'=2;
else if (SAEinit < Emax)          p'=3;
else if (SAEinit < 1.25×SAEmax) p'=4;
else                                p'=5;
    
```

여기서 $T_0 = 3 \times (16 \times 16)$ 으로 선택하며, 일반적으로 이 값은 잡음의 영향으로 간주된다. 다른 문턱값들은 계산량과 SAE를 고려하여 여러가지 실험영상을 이용한 실험을 통하여 선택된 것이다. 특히 탐색영역을 결정하는 p' 는 ME될 두 픽처간의 간격과는 무관하게 결정된다. 본 논문에서는 $p = \pm 14$ pels/frame인 경우에 대하여 설명했지만, 일반적인 임의의 p 에 대해서도 유사하게 설정할 수 있다.

3. 추가적인 이동벡터 추정

마지막으로 제안된 방식에 의해서 결정된 초기 이동벡터 v^0 를 중심으로 한 작은 탐색영역 $R'(k)$ 내에서 BMA에 의한 최종 MV를 추정한다.

$$R'(k) = \{(dx, dy) | -p' \leq dx, dy \leq p'\},$$

$$1 \leq p' \leq \lfloor \frac{p}{3} \rfloor, \forall k \quad (14)$$

일반적으로 BMA에서 탐색영역의 범위 p 가 커질수록 예측효율은 높아진다. 그러나 초기 MV를 보다 정확하게 설정한다면 식(14)의 작은 탐색범위 p' 만으로도 만족할 만한 예측 효율을 얻을 수 있다. 본 논문에서는 실험적으로 이에 대한 타당성을 설명하였다. 그림 6과 그림 7은 $M=3$ 인 경우에 대하여 식(14)의 작은 탐색범위 p' 에 따른 예측이득(B-픽처)과 이동 벡터차(differential motion vector : DMV) 성분의 엔트로피(entropy)를 나타낸 것이다. 예측이득은 p' 에 크게 영향을 받지 않음을 볼 수 있다. 그러나 p' 가 커질수록 오정합의 경우가 증가하고, MV의 상관성이 떨어지기 때문에 DMV의 엔트로피 높아진다. 또한

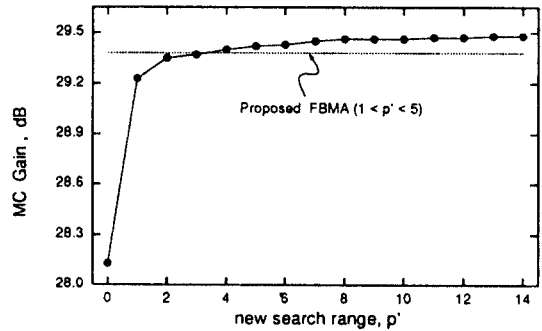


그림 6. $R'(k)$ 의 p' 의 크기에 따른 B-픽처의 평균 예측이득 ($M=3$, 실험영상: Flower Garden)
 Fig. 6. Average prediction gain of B-picture according p' ($M=3$, Test sequence: Flower Garden)

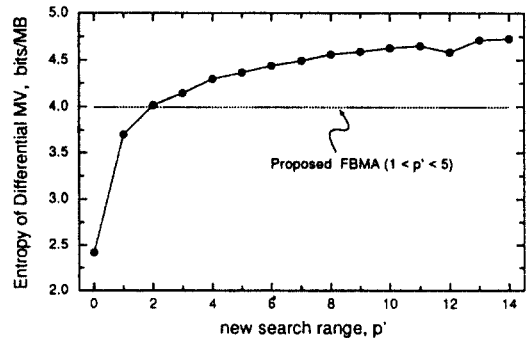


그림 7. p' 에 따른 MV 차성분의 평균 엔트로피 ($M=3$, 실험영상: Flower Garden)
 Fig. 7. Average entropy of MV difference according p' ($M=3$, Test sequence: Flower Garden)

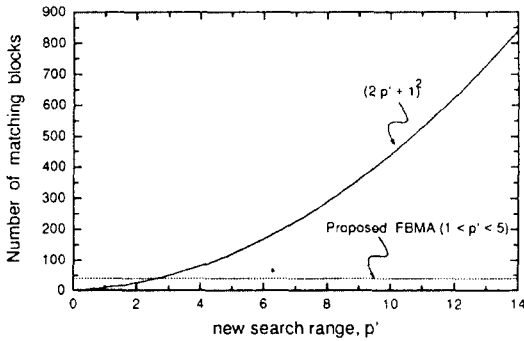


그림 8. p'에 따른 BMV의 정합블럭수
Fig. 8. Number of matching blocks for BMA according p'

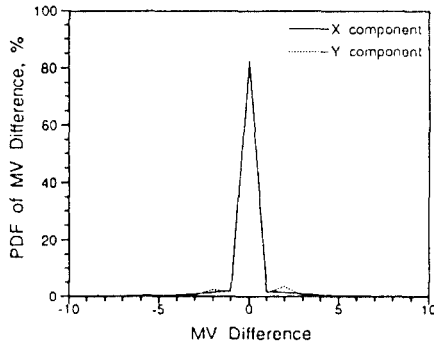


그림 9. FSA방식에 의한 FMV와 제안된 방식의 첫 단계에서 추정된 초기 FMV의 차 성분에 대한 확률분포함수 (M=3, 실험영상: Flower Garden)
Fig. 9. PDF of MV difference between MV of TSA and initial MV of proposed algorithm (M=3, Test sequence: Flower Garden)

그림 8에 나타난 것처럼 정합블럭의 수가 $(2p'+1)^2$ 에 비례해서 증가한다. 한편 그림 9는 FSA 방식에 의한 MV와 제안한 방식에 의한 초기 MV의 차에 대한 확률분포함수를 M=3인 경우에 대하여 나타낸 것으로, 약 94%가 $|DMV| \leq 5$ pel내에 분포한다. 따라서 ME의 계산량, 추정의 정확도, 그리고 예측효율 모두를 고려할 때, 식(14)와 같은 탐색범위 p'은 적합하다.

본 논문에서 제안한 이상의 세 과정을 통하여 결정된 식(14)의 탐색영역 내에서 B/P픽처의 이동벡터를 추정할 때, sub-GOP 내에서 동일한 위치의 각 블럭에 대한 전체 계산의 복잡도(정합블럭의 수), C_{NEW} 는 다음과 같다.

$$C_{NEW} = (2p' + 1)^2 \cdot (2M - 1)$$

$$= \left(2p \cdot \frac{p'}{p} + 1\right)^2 \cdot (2M - 1)$$

(15)

$$\approx \left(\frac{p'}{p}\right)^2 \cdot C_0 \cdot (2M - 1)$$

$$\approx \left(\frac{p'}{p}\right)^2 \cdot C_{TSA}$$

따라서 최대 이동변위가 $p = \pm 14$ pel/frame의 경우, 본 논문에서 제안한 고속 블럭 정합방식의 계산량은 TSA 방식에 비해서 약 $\frac{1}{7} \sim \frac{1}{196}$ 배 만큼 감소시킬 수 있으며, FSA 방식에 비해서는 더욱 큰 폭으로 감소시킬 수 있다.

IV. 모의실험 결과

MPEG-2의 TM5(3)를 이용한 모의실험을 통하여 본 논문에서 제안된 양방향 이동벡터 추정을 위한 고속 블럭정합방식의 성능을 비교 분석하였다.

1. 실험 조건

모의실험 조건은 표 2와 같다. half-pel accuracy의 MV는 두 단계로 추정하는데, 먼저 원영상을 이용한 open-loop ME에 의해서 integer-pel accuracy의 MV를 추정한 다음, local decoder에서 재생된 영상을 closed-loop ME에 의해 최종 half-pel accuracy의 MV를 추정한다. 또 B-픽처의 수에 따른 영향을 분석하기 위하여 표 2에 나타난 것과 같이 네 종류의 M에 대하여 모의실험을 수행하였다.

표 2. Test Model 5에 대한 모의실험 조건
Table 2. Simulation conditions of TM5

실험영상	Flower Garden 영상(60frames) ITU-R601(4:2:0)규격
TM 5의 픽처 구조	Frame-picture coding type
MV의 탐색범위	± 14.5 pel/frame
ME unit	16×16 크기의 frame-based MB
MV type	frame-based MV만 이용
Bit rate	2 Mbits/sec
GOP 구조	4 종류 이용 (14, 2), (15, 3), (16, 4), (20, 5)

성능의 비교를 위하여 다음의 네가지 ME방식을 TM 5에 적용하였다.

- (1) 제안된 고속 BMA(FBMA)
- (2) 전탐색 방식(FSA)
- (3) Telescopic 탐색 방식(TSA)
- (4) 2-D log 탐색+TSA(2-Dlog+TSA)

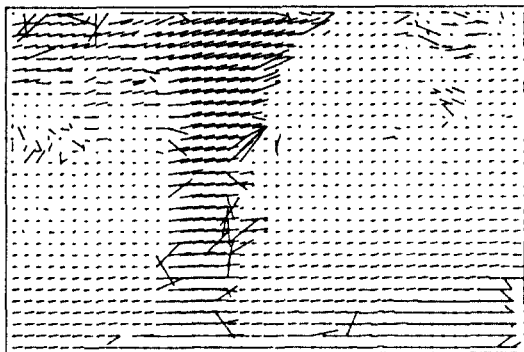
여기서 2-Dlog+TSA는 TSA 방식을 사용하지만 각 탐색단계에서 전탐색을 하지 않고 2-D Log 탐색 방식을 사용하는 방식으로, 2-D Log 탐색만을 사용하면 예측성능이 타방식에 비하여 현저히 떨어지기 때문에 보완한 것이다.

2. 실험 결과

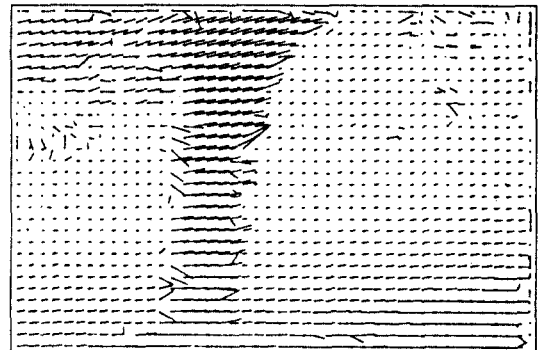
그림 10은 M=4일 때 픽처간격 k=3인 B₃ 픽처의

FMV 필드를 비교한 것으로 FSA과 TSA는 edge나 uncovered area등에서 오정합이 많이 발생하는 반면, 제안된 방식은 보다 정확한 MV를 추정함을 볼 수 있다. 2-log+TSA 방식은 최소 왜곡 방향(DMD)에 대한 가정의 한계로 오정합이 매우 많이 발생한다. 그림 11은 open-loop에서의 평균 예측이득을 비교한 것으로, 제안된 방식은 FSA 방식에 근접한 예측이득(-0.30~-0.53dB차)을 얻었다. 특히 ME의 픽처간격이 가장 큰 P-픽처의 경우, M이 커질 수록 제안된 방식이 TSA 방식 보다 예측이득이 오히려 더 높다. 이는 TSA방식의 특성상 edge나 uncovered area 에서 오정합이 많이 발생하기 때문이다.

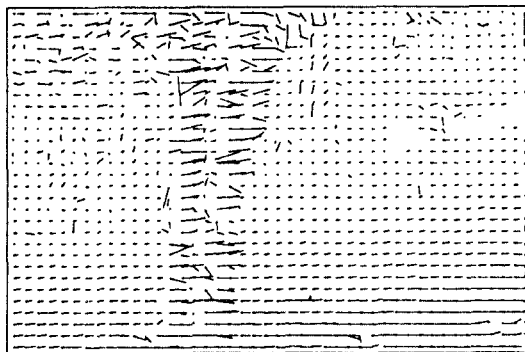
그림 12는 아래의 식(16)와 같이 MV를 DPCM한 후의 MV 차성분에 대한 평균 엔트로피를 비교한 것으로서, 제안된 방식은 타방식에 비해서 MV의 신뢰도와 균질도가 좋기 때문에 엔트로피가 가장 낮다.



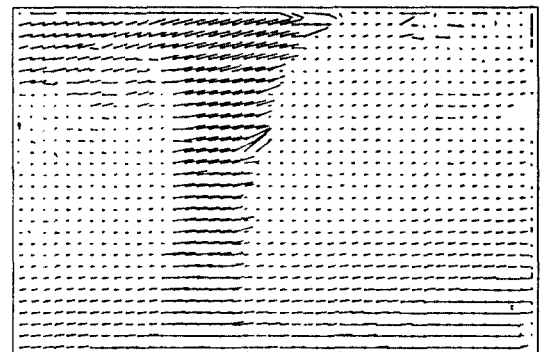
(a) FSA 방식



(b) TSA 방식



(c) 2-DLog+TSA 방식



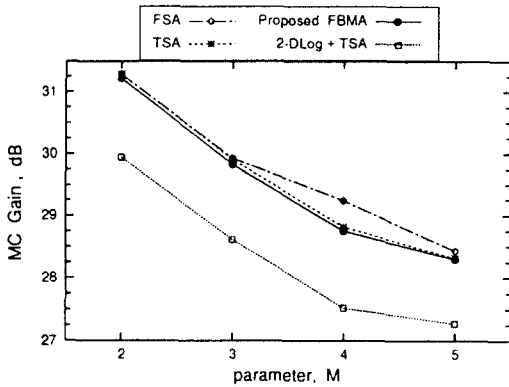
(d) 제안된 방식(FBMA)

그림 10. 픽처간격=3인 B₃ 픽처의 FMV 필드(M=4)

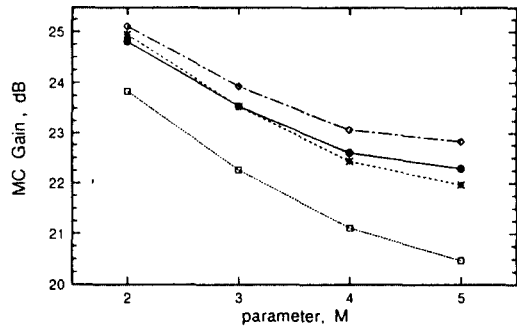
Fig. 10. FMV field of B₃ picture with picture distance=3 (m=4)

$$\begin{aligned}
 \text{DMV Entropy} &= (x \text{ 성분의 엔트로피}) \\
 &\quad + (y \text{ 성분의 엔트로피}) \\
 &= - [\sum p(\text{DMV}_x) \cdot \\
 &\quad \log p(\text{DMV}_x) + \sum p(\text{DMV}_y) \cdot \\
 &\quad \log p(\text{DMV}_y)] \quad [\text{bits/MB}]
 \end{aligned} \tag{16}$$

여기서 DMV_x 및 DMV_y 는 각각 MV의 수평(x), 수직(y) 성분의 차로 나타내며, $p(\cdot)$ 는 MV 차 성분의 PDF이다. 따라서 그림 13에 나타낸 것과 같이 TM 5를 통하여 발생된 MV의 프레임당 평균 비트량은 제안된 방식이 가장 적다.



(a) B-픽처



(b) P-픽처

그림 11. B/P 픽처의 평균 이동보상 예측이득(실험영상: Flower Garden)
Fig. 11. Average MCP gain of B/P pictures (Test sequence: Flower Garden)

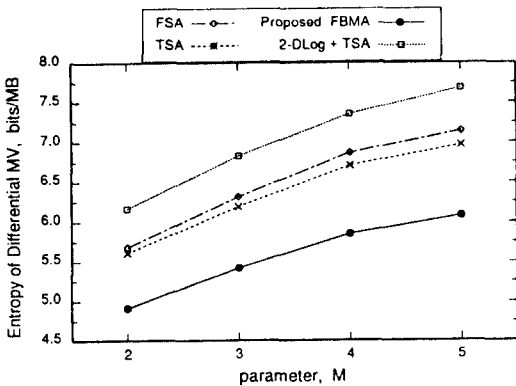


그림 12. DPCM된 MV 차성분의 평균 엔트로피 (실험영상: Flower Garden)
Fig. 12. Average entropy of MV difference (Test sequence: Flower Garden)

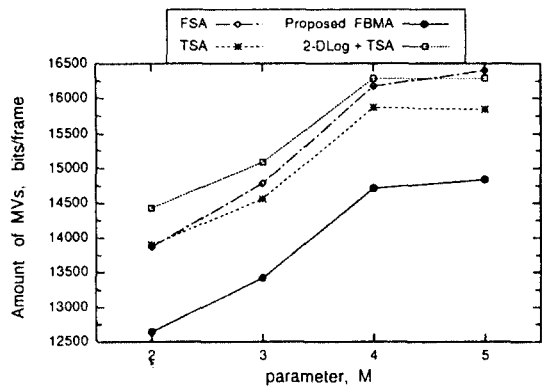


그림 13. TM5에 의한 픽처당 평균 MV 비트량 (실험영상: Flower Garden)
Fig. 13. Average amount of MV's bit per frame TM5 (Test sequence: Flower Garden)

그림 14는 TM5를 통하여 재생된 영상의 평균 PSNR을 비교한 것으로서, 제안된 방식은 FSA 방식보다 오히려 더 나은 성능을 얻을 수 있다. 이는 제안된 방식에 의해 추정된 MV는 보다 높은 신뢰도를 갖기 때문에 MV에서 절약된 비트량 만큼을 DCT 계수에 할당할 수 있고, 또 예측오차 성분의 통계적인 특성이 좋기 때문에 부호화 성능이 개선된다고 할 수 있다.

표 3은 각 ME 방식을 적용했을 때, sub-GOP의 각 블록에 대한 평균 계상량으로서 정합 블록의 수를 비교한 것이다. 제안된 방식은 TSA 방식에 비해서 약 $\frac{1}{20} \sim \frac{1}{23}$ 배, FSA 방식에 비해서 약 $\frac{1}{40} \sim \frac{1}{185}$ 배 만큼씩 계산량이 감소됨을 알 수 있었다.

한편 본 논문에서는 Flower Garden 영상 시퀀스

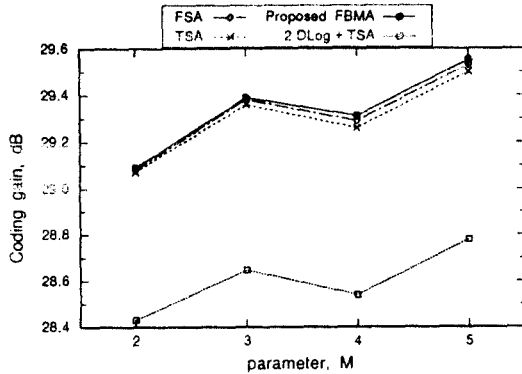


그림 14. TM5에 의해 재생된 영상의 평균 PSNR (실험영상: Flower Garden)
 Fig. 14. Average PSNR of reconstructed image TM5 (Test sequence: Flower Garden)

에 대한 실험 결과만을 보여주었지만, 다른 실험 영상에 대해서도 비슷한 결과를 나타내었다. 예를 들면, 운동체의 움직임은 비교적 적지만 다양하고 복잡한 움직임이 있는 'Mobile & Calendar' 영상 시퀀스나 운동체의 움직임이 매우 큰 'Football' 영상 시퀀스에 대한 모의실험 결과도 'Flower Garden' 영상 시퀀스에 대한 결과와 비슷한 양상을 보였다.

V. 결 론

본 논문에서는 MPEG과 같이 단방향 예측만을 하는 P-픽처와 양방향 예측이 가능한 B-픽처, 그리고 I-픽처가 주기적으로 사용되는 이동보상 예측 구조에서 실시간 구현이 가능할 정도의 매우 적은 계산량으로 보다 효과적인 양방향 이동벡터를 추정하기 위한 고속 블럭정합방식을 제안하였다. 제안된 ME 방식에서는 이동변위에 대한 등속 모델과 이동벡터의 높은 시공간 상관성을 이용하여 실제 움직임에 근접한 초기벡터를 추정하고, 이로부터 BMA의 탐색영역, 즉 계산량을 적용적으로 줄였다.

MPEG-2의 Test Model 5를 통한 모의 실험 결과, 제안된 방식은 기존의 ME 방식에 비해서 월등히 적은 계산량(TSA 방식의 약 $\frac{1}{20} \sim \frac{1}{23}$ 배, FSA 방식의 약 $\frac{1}{40} \sim \frac{1}{185}$ 배)으로 보다 높은 신뢰도와 정확도를 갖는 MV를 추정할 수 있었으며, B-픽처의 수가 많은 경우에는 타방식에 비해 더욱 우수한 성능을

표 3. sub-GOP의 각 MB에 대한 평균 정합블럭의 수
 Table 3. Average number of matching blocks of each MB in sub-GOP

ME method	M=2	M=3	M=4	M=5
FSA	5056	15979	37004	71485
TSA	2523	4205	5887	7569
Proposed FBMA	111	190	273	387
2DLog+TSA	57	95	133	71

보였다. 예측효율과 재생영상의 PSNR 측면에서도 가장 우수한 예측효율을 갖은 전탐색 방식(FSA)과 대등한 성능을 얻었다. 특히 제안된 방식은 실제 움직임에 보다 근접하는 MV를 추정하기 때문에 BMA의 단점인 오정합으로 인한 '블럭 현상'을 많이 줄일 수 있었고, 따라서 주관적인 화질도 다소 향상시킬 수 있었다.

참고문헌

1. MPEG Editorial Group, "Coding of moving pictures and associated audio for DSM at up to about 1.5 Mbit/s : ISO/IEC 111721-IS," ISO/IEC JTC1/SC29/WG11, April 1993.
2. MPEG Video Editorial Group, "Generic coding of moving pictures and associated audio: ISO/IEC 13818-2 CD," ISO/IEC JTC1/SC29/WG11, MPEG93/No.602, 1993.
3. MPEG Test Model Editing Committee, "Generic coding of moving pictures and associated audio : MPEG-2 Test Model 5(TM5)," ISO/IEC JTC1/SC29/WG11, MPEG93/No. 457, 1993.
4. 김재균, 김민규, "MPEG-2 표준 영상압축기법의 최근 동향," 전자공학회지, HDTV 특집, vol.19, No.12, Pp.1217-1231, Dec. 1992.

5. M.-K. Kim and J.-K. Kim, "Efficient motion estimation algorithm for bidirectional prediction scheme," IEE Electronics Lett. vol.30 No.8, pp.632-633, April 1994.
6. S.-W. Wu and A. Gersho, "Joint estimation of forward/backward motion vector for MPEG interpolative prediction," Proc. of Picture Coding Symposium, pp.14.3.1-14.3.2, Mar. 1993.
7. X. Lee and et al., "An efficient MPEG motion compensation scheme by motion trajectory tracking method," Proc. of SPIE: Visual Communication and Image Processing, vol.1818, pp.594-605, Nov. 1992.
8. N.Haddadi and C.-C. Jay Kuo, "Fast computation of motion vectors for MPEG," Proc. of SPIE: Visual Communication and Image Processing, vol.2094, pp.350-361, Nov. 1993.
9. T.Koga and et al., "Motion compensated interframe coding for video conferencing," Proc. of NTC, pp.G5.3.1-G5.3.5, Dec. 1981, New Orleans, LA, USA.
10. J.R.Jain and A.K.Jain, "Displacement measurement and its application in inter-frame image coding," IEEE Trans. on Comm., vol.COM-29, pp.1799-1806, Dec. 1981.
11. C.-H.Hsieh and et al., "Motion estimation algorithm using interblock correlation," Electrical Letters, vol.26, No.5, pp.276-277, Mar. 1990.



金 旻 圭 (Min Kyu Kim)

1961년 6월 10일생

1987년 2월 : 고려대학교 공과대학
전자공학과 졸업(공학
사)

1989년 2월 : 한국과학기술원 전기및
전자공학과 졸업(공학
석사)

1995년 2월 : 한국과학기술원 전기및전자공학교 졸업(공학박사)

1995년 2월~현재 : 대우전자(주) VCR연구소 뉴미디어 개발부
선임연구원

※ 주관심분야 : 영상부호화 및 디지털신호처리



金 在 均 (Jae Kyoon Kim)

1938년 9월 17일생

1962년 2월 : 한국항공대학교 응용전자
과 졸업(공학사)

1967년 2월 : 서울대학교 전자공학과
대학원졸업(공학석사)

1971년 8월 : 미국 남가주대학교 전자
공학과 졸업(공학박사)

1972년~1973년 : 미국 NASA ASFC 연구원

1973년~현 재 : 한국과학기술원 전기및전자공학과 교수

1984년~1985년 : 과학기술처 전기전자 연구조정관

1993년~1994년 : 한국통신학회 회장

1993년~현 재 : 한국과학기술원 정보통신연구센터 소장