

## 낮은 공간 해상도로의 디지털 비디오 변환 부호화기에 대한 움직임 벡터 추정 기법

김 봉 곤, 박 강 서, 박 상 회  
연세대학교 전기 전자 공학과

### Motion vector re-estimation technique for transcoding into lower spatial resolution

Bong-Gon Kim, Kang-Seo Park, Sang-Hui Park  
Dept. of Electrical & Electronic Eng., Yonsei Univ.

**Abstract** - 본 논문에서는 낮은 공간 해상도로의 디지털 비디오 변환 부호화기에 대한 움직임 벡터 추정 기법을 제안한다. 블록 정합 에러가 테일러 급수 전개에 의한 참조 블록의 복잡도에 비례한다는 사실에 근거하여 새로운 움직임 벡터 평가식을 유도하여 변환 부호화기에 대한 움직임 벡터를 이 평가식에 의해 추정한다. 제안한 방법에 의해 추정된 움직임 벡터들은 기존 방법에 의해 추정된 벡터보다 최적 벡터에 더 인접하였으며, 부호화된 영상은 더 좋은 화질을 보인다. 특히 빠른 움직임을 포함하는 영상에서 제안된 방법의 우수성은 더 두드러진다.

### 1. 서 론

멀티 미디어 기술이 계속 발전해감에 따라 멀티 미디어에 대한 통신 환경은 더욱 다양해지고 복잡해진다. 주문형 비디오(VoD)나 이동형 멀티 미디어 통신기기에서의 응용과 같이 이미 부호화 되어있는 비디오를 사용한 응용분야에서 이미 부호화된 비디오는 전송 채널 및 다양한 특성을 가지는 수신자측의 장치에 적합하도록 맞추어져야 한다. 이것은 이미 부호화된 비디오 시퀀스를 주어진 환경에 적합하도록 다양한 포맷, 비트율, 공간적 혹은 시간적 해상도를 가지는 새로운 시퀀스로 변환시키는 비디오 변환 부호화기에 의해 이루어질 수 있다[1][2].

효율적인 변환 부호화기 알고리즘은 단순히 계산량을 줄이는 것 뿐만 아니라 이미 부호화된 비디오의 화질을 최대한 유지시켜야 한다. 이를 위해서 대부분의 변환 부호화기는 부호화시 수행 처리 시간의 60%이상을 차지하는 움직임 벡터 추정에 의해 소비되는 시간을 절약하기 위해서 전역 탐색에 의한 움직임 벡터 추정기법 대신 입력 비디오 스트림의 움직임 벡터를 이용하여 움직임 벡터를 재추정한다. 입력 비디오 스트림을 이용한 새로운 움직임 벡터의 추정은 공간적 혹은 시간적 해상도의 감소와 관계되어 있는 변환 부호화기의 효율을 결정짓는 중요한 요소이다. 이러한 벡터 추정에 대한 여러 가지 해결책들이 제안되어 왔으나 기존에 제안된 대부분의 움직임 벡터 재추정 기법은 수학적 유도가 아닌 경험적 방법을 통해 개발되었다[1][2][3].

본 논문에서는 블록 정합 에러가 테일러 급수 전개(Taylor series expansion)에 의해 참조 블록의 복잡도에 비례한다는 사실로부터 변환 부호화기에 대한 움직임 벡터 평가식을 수학적으로 유도하고 제안된 기법의 성능과 기존 기법의 성능을 비교하여 본다.

### 2. 낮은 공간 해상도로의 변환 부호화

전송 채널 용량이 이미 부호화된 비디오의 비트율에 비교하여 극히 작거나 수신자측 장치의 해상도가 제한되어 있는 경우 낮은 공간 해상도로의 변환 부호화기가 필요하다. 이같은 경우, 변화 부호화되는 비디오의 움직임

벡터가 이미 부호화된 비디오의 움직임 벡터와 다른 크기를 가지기 때문에 이미 부호화된 비디오의 움직임 벡터를 직접적으로 재사용 하고 정제 하는 것은 불가능하다[2]. 낮은 공간 해상도로 축소하는 변환부호화에서는 일반적으로 2의 배수에 의한 축소가 이루어지게 되고 변환 부호화되는 비디오의 움직임 벡터는 그림 1과 같이 네 개의 입력 움직임 벡터로부터 구해진다. 이러한 변환 부호화를 위해서 중간 벡터 선택법(MVS-median vector selection), AWW(aligned to worst weighting), 평균법(averaging)과 같은 경험적 방법을 통한 여러 가지 방법들이 개발되어 있다.

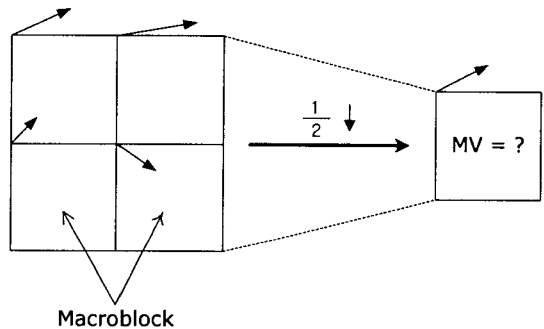


그림 1 공간 해상도 축소시 움직임 벡터 재구성

### 3. 제안된 알고리즘

블록 정합 과정에서 정합 에러는 참조 블록의 그래디언트 값 및 최적 벡터와 추정 벡터의 거리에 비례한다는 것이 테일러 급수 전개를 이용하여 증명되었다[4]. 블록  $p = (x, y)$ 에서의 화소값을  $f_i(p)$ 라 하고  $p$  위치에서의 최적 움직임 벡터를  $mv = (mvx, mvy)$ 라 하면 주어진 후보 움직임 벡터  $cmv = (cmvx, cmvy)$ 의 정합 에러는 식 (1)과 같이 참조 블록의 그래디언트 값으로 표현된다.

$$d_i(p, cmv) = |f_i(p) - f_{i-1}(p + cmv)| \\ \approx \left| \frac{\partial f_i(p)}{\partial x} (cmvx - mvx) \right| + \left| \frac{\partial f_i(p)}{\partial y} (cmvy - mvy) \right| \quad (1)$$

식 (1)은  $p$  위치의 블록 정합 에러가 영상의 복잡도를 나타내는 블록의 그래디언트 값에 비례한다는 사실을 보여준다.

블록  $p$ 의 수평방향의 그래디언트 값은 다음과 같이 정의된다.

$$|G_x| = \left| \frac{\partial f_i(p)}{\partial x} \right| = \sum_x \sum_y |f_i(x, y) - f_i(x+1, y)| \quad (2)$$

여기서  $x$ 와  $y$ 는 각각 블록  $p$ 의 수평 및 수직방향의 화소 위치이다. 수직방향의 그래디언트 값은 이와 유사한 방법을 통해 구할 수 있다.

2의 배수로 축소되는 낮은 공간 해상도로의 변환 부호화에서 새로운 움직임 벡터는 높은 해상도의 입력 시퀀스에서 사용된 네 개의 움직임 벡터로부터 재추정된다. 낮은 해상도를 가지는 시퀀스의 한 매크로 블록을 네 개의 하위 블록으로 구성하면 매크로 블록의 에러는 다음과 같이 네 개의 하위 블록에 대한 에러의 합으로 표현된다.

$$d_i(p) = d_i(p_1) + d_i(p_2) + d_i(p_3) + d_i(p_4) \quad (3)$$

여기서  $p_1 \sim p_4$ 는 매크로 블록  $p$ 의  $8 \times 8$  하위 블록이다. 각각의 하위 블록의 최적 벡터는 입력 시퀀스의 움직임 벡터를 반으로 나눈 값이라고 가정하고 이 벡터를 입력 벡터라고 부른다. 하위 블록의 정합 왜곡 에러는 식 (1)에서 최적 벡터( $mv = (mv_x, mv_y)$ ) 대신 입력 벡터를 사용하여 얻어진다. 일반적으로 하나의 매크로 블록에는 단지 하나의 움직임 벡터가 할당되므로 동일한 매크로 블록내의 각 하위블록은 동일한  $cmv$ 를 갖게 된다. 그러면 주어진  $cmv$ 에 대한 블록  $p$ 의 정합 왜곡 에러는 다음과 같이 표현된다.

$$d_i(p, cmv) = \sum_{i=1}^4 \left\{ \left| \frac{\partial f_i(p_i)}{\partial x} (cmv_x - mv_{x_i}) \right| + \left| \frac{\partial f_i(p_i)}{\partial y} (cmv_y - mv_{y_i}) \right| \right\} \quad (4)$$

여기서  $(mv_{x_i}, mv_{y_i})$ 는 각 하위 블록의 입력 벡터이다. 변환 부호화되는 낮은 해상도를 갖는 시퀀스의 매크로 블록에서 식 (5)와 같이  $d_i(p, cmv)$ 를 최소화 시키는  $cmv$ 가 재추정되는 움직임 벡터로 선택이 된다.

$$\text{estimated } mv \text{ of } p = \arg \min [d_i(p, cmv)] \quad (5)$$

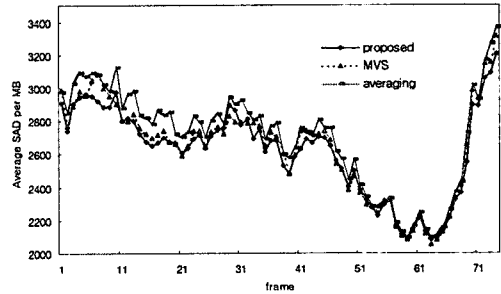
식 (4)를 통해서 네 개의 입력 벡터중 하나의 벡터에 대해서 왜곡이 최소가 됨을 알 수 있다. 그리고  $cmv_x$ 과  $cmv_y$ 는 각각 독립적으로 선택이 된다. 이 식은 또한 매크로 블록에서 최적 벡터는 입력 벡터와의 거리뿐만 아니라 하위 블록의 복잡도와도 관계가 있음을 보여준다. 만약 각 하위 블록이 동일한 그래디언트를 갖는다면 입력 벡터중에서 가장 가까운 벡터가 새로운 벡터로 재추정될 것이다.

#### 4. 실험 및 결과

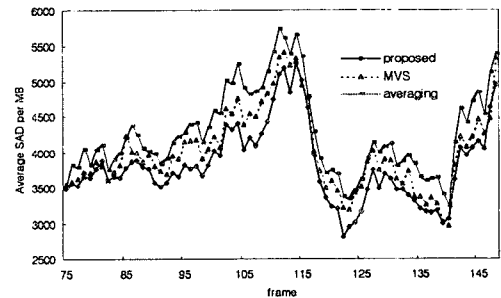
성능을 측정하기 위해 먼저 CIF의 Football 시퀀스를 사용하여 부호화 하였다. 150 프레임의 시퀀스를 사용하였고 양자화 파라미터(QP)는 5로 고정을 한 상태에서 TMN 5를 이용하여 H.263 시퀀스로 부호화 하였다 [5][6]. 그 다음에 양자화 파라미터를 10으로 고정하여 부호화된 시퀀스를 평균법, MVS 및 제안된 방법으로 부호화된 비디오를 QCIF로 변환 부호화하였다. 기존 방법중에서 MVS가 가장 단순하고 가장 좋은 화질을 보였

기 때문에 제안된 방법과 MVS를 주로 비교하였다.

그림 2(a)와 (b)는 주어진 시퀀스에 대하여  $\pm 0.5$  화소의 정제를 한 후의 매크로 블록당 절대 차이의 합인 평균(SAD-average sum of absolute difference)을 보여준다. 낮은 공간 해상도로의 변환 부호화에서 높은 해상도를 가진 입력 시퀀스의 움직임 벡터를 2로 나누는 과정에서 반올림 에러가 흔히 일어나기 때문에  $\pm 0.5$  화소의 정제 과정이 필수적이다.



(a) 1~74번째 프레임



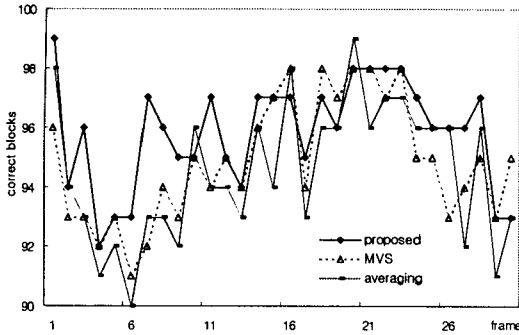
(b) 75~149번째 프레임

그림 2 각 프레임에 따른 매크로 블록당 평균 SAD

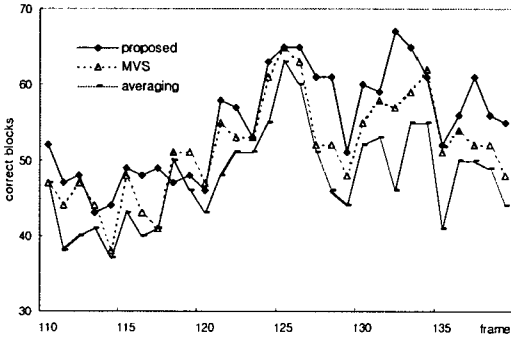
그림 2(a)는 상대적으로 움직임이 느린 전반부 75 프레임의 SAD를 나타내고 그림 2(b)는 움직임이 빠른 후반부 75 프레임을 나타낸다. 제안된 방법이 전반부 75 프레임에서는 약간 더 좋은 성능을 보이고 후반부 프레임에서는 확연하게 우수한 성능을 보인다. 제안된 방법의 평균 SAD가 MVS방법보다 199만큼 작고 평균법보다는 357만큼 작다. 그래디언트 계산에 의한 부가적 계산량은 2 포인트에 대한 SAD의 계산량밖에 되지 않는다. 정제 과정을 포함하지 않더라도 제안된 방법으로 구한 SAD는  $\pm 0.5$  화소의 정제 과정을 거친 MVS의 SAD보다 단지 12.7밖에 높지 않았고 제안된 방법에 정제 과정을 포함시켰을 경우에는 훨씬 작은 SAD값을 얻을 수 있었다.  $\pm 0.5$  화소의 움직임 벡터 정제 과정은 9 포인트에 대한 SAD의 계산량 정도만을 필요로 한다.

각 알고리즘의 정확도를 평가하기 위해서 각 알고리즘에 의해 추정된 움직임 벡터와 전역 탐색에 의해 추정된 최적의 움직임 벡터를 비교하였다. 그림 3은 최적의 움직임 벡터와 일치하는 각 알고리즘에 의해 추정된 움직임 벡터 수를 나타낸다. 제안된 알고리즘에 의해 추정된 움직임 벡터가 기존의 다른 알고리즘에 의해 추정된 움직임 벡터보다 더 많은 수가 최적의 움직임 벡터와 일치

함을 알 수 있다. 정확도면에서 또한 제안된 방법이 움직임이 빠른 후반부 프레임에서 더 우수한 성능을 갖는다.



(a) 1~30번째 프레임



(b) 110~139번째 프레임

그림 3 최적 벡터와 일치하는 재추정된 움직임 벡터 수

제안된 알고리즘 및 기존 알고리즘을 고정 비트율에서도 실험을 하였다. 입력 시퀀스로서 CIF의 Football 시퀀스를 QP=5로 고정시켜 부호화 하였다. 이 입력 시퀀스를 200Kbit/s의 비트율에서 QCIF 시퀀스로 변환 부호화 하였고 이를 다양한 정제 범위범위에 따라 정제를 하였다. 그림. 4는 각각의 정제 범위에 따라 각 알고리즘에 의해 변환 부호화된 시퀀스의 평균 PSNR을 보여준다. 변환 부호화시 입력 시퀀스의 움직임 벡터를 이용하지 않고 입력 시퀀스를 복호화한 뒤 낮은 해상도의 시퀀스 상에서 전역 탐색에 의해 재부호화를 한 비디오의 평균 PSNR은 25.99dB이다. 그림 4를 보면 제안된 방법을 통해  $\pm 0.5$  화소의 정제를 한 비디오의 PSNR이 다른 알고리즘을 통해  $\pm 1.5$  화소의 정제를 한 비디오의 PSNR과 거의 같음을 알 수 있다. 이 같이 동일한 범위내에서 정제를 하는 경우 제안된 방법에 의해 변환 부호화 된 비디오가 더 좋은 화질을 보여준다.

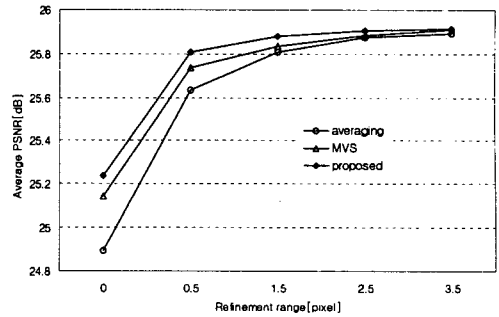


그림 4 각 정제 범위에 따른 평균 PSNR

## 5. 결 론

본 논문에서 낮은 공간 해상도로의 디지털 비디오 변환기에 대한 새로운 움직임 벡터 재추정 기법을 제안한다. 움직임 벡터 평가식이 테일러 급수 전개에 의해 유도되었고 변환 부호화시 이 평가식에 의해 움직임 벡터를 재추정 하였다. 제안된 방법에 의해 재추정된 움직임 벡터가 기존 방법에 의한 움직임 벡터보다 최적 벡터에 더 인접하였고 더 좋은 화질을 보여주었다. 특히 빠른 움직임이 있는 프레임에서 제안된 방법의 우수성은 더 두드러진다.

## [참 고 문 헌]

- [1] T. Shanableh and M. Ghanbari, "Heterogeneous video transcoding to lower spatio-temporal resolutions and different encoding formats," IEEE Transactions on Multimedia, vol. 2, no. 2, pp. 101-110, 2000
- [2] N. Bjork and C. Christopoulos, "Transcoder architectures for video coding," IEEE Transactions on Consumer Electronics, vol. 44, no. 1, pp. 88-98, 1998
- [3] B. Shen, I. K. Sethi, and B. Vasudev, "Adaptive motion-vector resampling for compressed video downscaling," IEEE Transactions on Circuits and Systems for Video Technology, vol. 9, no. 6, pp. 929-936, 1999
- [4] J. N. Kim, S. C. Byun, Y. H. Kim, and B. H. Ahn, "Fast full search motion estimation algorithm using early detection of impossible candidate vectors," IEEE Transactions on Signal Processing, vol. 50, no. 9, pp. 2355-2365, 2002
- [5] ITU-T Rec. H.263, "Videocodex for low bit rate communication," 1996
- [6] Telnor Codec, ITU-T/GS-15, "Video codec test model, TMN5," Telnor research, 1995