

# 이종 동영상 부호화기의 트랜스코딩을 위한 움직임 벡터 재추정에 관한 연구

論文

51D-10-4

## A Study on Motion Vector Re-Estimation for Transcoding of Heterogeneous Video Coder

尹圭涉\* · 朴慷慨\*\* · 鄭泰允\*\*\* · 朴相嬉§

(Kyu-Seop Yoon · Kang-Seo Park · Tae-Yun Chung · Sang-Hui Park)

**Abstract** - This paper proposes new motion vector re-estimation techniques for transcoding of heterogeneous video coders. The first case is for transcoder of coding format. The coding algorithm with B-picture like MPEG standards is transcoded into the different algorithm without B-picture like H.261 standards. The second case is for transcoder of spatial resolution reduction. It generates a downsampled video bitstream at the video source. The third case is the integrated method that previous two cases are needed simultaneously. These three transcoding algorithms are compared and verified through reconstructed image quality in PSNR, motion vector errors between optimal motion vectors of cascaded transcoder and motion vectors estimated by proposed methods and the amount of operation.

**Key Words** : Video Coding, Trancoder, Format Conversion, Resolution Conversion

### 1. 서 론

최근 멀티미디어와 유·무선 정보 통신망의 급속한 발전으로 인해 일반 가정에서까지도 컴퓨터와 컨버터를 장착한 TV를 통해 고선명 TV를 감상할 수 있게 되었고, 휴대용 단말기에서도 음성정보 뿐 아니라 영상정보 서비스까지 제공되는 등 바야흐로 디지털 영상정보 혁명이 사회 전반에 확산되고 있다. 영상정보의 디지털화는 화질, 신뢰도, 보안성 그리고 정보조작의 편이성이 있어서 아날로그에 비해 월등히 우수한 성능을 갖고 있으나 그 정보량의 방대함 때문에 전송 및 저장 매체의 효율적인 사용이라는 측면에서 정보압축이 필수적이다[1]. 이러한 디지털 동영상정보의 압축 부호화 기술로는 MPEG1/2/4, H.261/263등의 국제규격의 표준안들이 대표적이며 응용 시스템의 목적, 사용범위 등에 따라 선택적으로 적용되고 있다[2].

그러나 이미 부호화된 동영상정보에 대해 전송 또는 저장 매체의 상태, 수신단의 특성, 화질에 대한 요구조건등이 변경될 경우, 목적 비트율, 영상의 크기 및 부호화 형식의 변환에 이루어져야 할 필요가 발생한다. 이와 같이 원래의 목적으로 이미 부호화된 영상을 다른 목적에 맞는 영상으로 변환하는 것을 일종의 트랜스코딩(transcoding)이라 한다.

이러한 목적의 트랜스코딩 중 가장 간단한 방법은 종속(cascaded) 트랜스코더로서, 미리 부호화된 영상정보를 완전히 복호화한 후 이를 다시 목적에 맞게 재부호화하는 과정

을 거치게 된다. 재부호화시 움직임 예측과정은 전체 부호화 과정의 60% 이상의 시간을 소모하게 되는데, 종속 트랜스코더는 기존의 부호화된 정보를 이용하지 않으므로 새로운 부호화 과정에서 많은 연산 시간을 필요로 한다[3]. 이러한 시간적 손실을 보상하기 위해서 이미 부호화된 영상의 움직임 벡터(MV, Motion Vector)를 재사용하는 기법에 대한 연구가 활발히 이루어지고 있다[4].

기존의 MV를 사용하는데 있어 영상의 크기와 부호화 형식이 달라지게 되면 MV의 수와 예측 방향이 달라지게 되므로 새로운 MV의 추출이 용이하지 않다. 그러나 영상의 크기가 달라지게 되더라도 이미 부호화된 영상에 존재하는 매크로블록의 MV와 목적 크기로 변환되는 영상에서 이 매크로블록에 상응하는 위치에서의 MV는 서로 큰 상관관계를 가지게 된다. 마찬가지로 영상의 목적 부호화 형식이 기존의 부호화 형식과 다른 이종 동영상(heterogeneous video) 트랜스코더의 경우에도 변환되는 영상에서의 MV는 이미 부호화된 영상의 MV와 큰 상관관계가 있다.

따라서 목적 비트율과 영상의 크기 및 부호화 형식의 변환을 위한 이종 동영상 트랜스코더의 효율적인 설계를 위해서는 먼저 이미 부호화된 영상정보의 MV로부터 목적 부호화 형식에 맞는 MV를 추정하고, 다시 추정된 MV를 이용하여 변화된 영상의 크기에 합당한 MV를 재추정한 후, 마지막으로 움직임 추정시 발생하는 오차를 최소화하는 정제(refinement) 과정을 수행해야 한다.

본 연구에서는 I, P, B picture의 부호화 모드를 갖는 이미 부호화된 MPEG 기반 동영상 비트열을, 목적 비트율과 영상의 크기가 약 1/4로 감소되고 또한 I, P picture의 부호화 모드로만 이루어진 MPEG 기반 비트열로 부호화 형식을 변환하는 이종 동영상 트랜스코더의 구현에 있어 종속 트랜스코더에 비해 화질 열화 및 수행 계산량과 시간 지연을 최소화 하는 새로운 움직임 벡터 재추정 기법을 제안한다.

\* 正會員 : LG 電子 情報通信 主任研究員 · 工碩

\*\* 正會員 : 延世大 電氣·電子工學科 博士課程

\*\*\* 正會員 : 江陵大 制御計測工學科 專任講師 · 工博

§ 正會員 : 延世大 電氣·電子工學科 教授 · 工博

接受日字 : 2002年 4月 19日

最終完了 : 2002年 8月 7日

## 2. 동영상 부호화기의 트랜스코딩

디지털 동영상 압축 부호화를 채용하는 매체의 사용이 급증하는 가운데, 다수의 전송 및 저장 매체간의 접속은 여러 가지 문제를 유발시킬 수 있다.

즉, 이미 부호화된 동영상 정보는  $R_1(\text{bits/sec})$ 의 비트율을 가지고 있는데 반해, 전송 및 저장 매체 용량의 한계로  $R_1$ 보다 작은  $R_2(\text{bits/sec})$ 의 비트율로 전송해야 할 경우가 생기고,  $M_1 \times M_2$ 의 영상의 크기로 이미 부호화된 영상이지만 그 크기를  $M_3 \times M_4$ 로 줄여야 하는 경우도 발생할 수 있다. 또한 MPEG1/2 형식으로 이미 부호화된 동영상 정보를 H.261/263 등의 기존 통신망에서 사용되는 형식으로 재부호화할 필요가 발생할 수도 있다. 이와 같이 이미 부호화된 동영상 정보를 새로운 목적에 맞는 동영상 정보로 변환하는 것이 동영상 부호화기의 트랜스코딩이다.

그림 1은 동영상 부호화기 트랜스코더의 일반적인 구조를 나타낸 것이다.

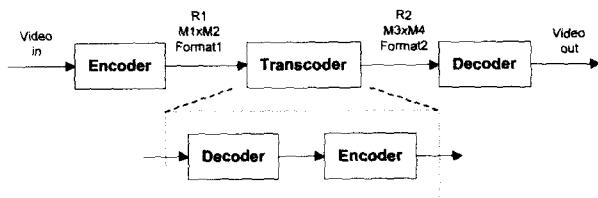


그림 1 동영상 부호화 트랜스코더의 일반적인 구조

Fig. 1 Structure of transcoder for video coding

### 2.1 종속 트랜스코더

종속 트랜스코더는 가장 간단한 구조의 트랜스코더로서 그림 2와 같이 복호화기의 출력단에 부호화기를 종속적으로 연결한 형태를 가진다. 즉, 이미 부호화된 정보를 완전히 복호화하여 영상을 재구성한 후 다시 부호화기를 거쳐 목적 비트율과 크기 그리고 목적 부호화 형식으로 재부호화한다.



그림 2 종속 트랜스코더

Fig. 2 Cascaded transcoder

그러나 이러한 종속 트랜스코더는 복호화 및 부호화의 전 과정을 다시 실행함으로서 최적의 화질을 얻을 수 있지만, 이미 부호화된 비트스트리밍(bitstream)에서 MV와 같은 정보를 트랜스코딩의 재부호화 과정에서 이용하지 않기 때문에 재부호화시 많은 연산 시간이 소모되며 또한 시스템 복잡도가 증가하는 단점이 있다.

### 2.2 움직임 벡터 재사용 트랜스코더

#### (1) 부호화 형식 변환을 위한 움직임 벡터 추정 방식

이종 동영상 트랜스코더는 이미 부호화된 영상의 크기 변

환 없이 다른 형식의 영상으로 재부호화하는데 사용된다. 즉, MPEG1으로 부호화된 영상을 H.261/263으로 변환하거나 MPEG2 MP@ML로 부호화된 영상을 MPEG2 SP@ML로 변환하는 것이 일종의 이종 동영상 트랜스코더이다. 이러한 여러 형식의 동영상 압축 부호화 방식은 움직임 예측 방법에 따라 전방향(forward) 예측으로만 이루어진 형식과 전방향, 후방향(backward) 예측, 혹은 양방향 예측으로 이루어진 형식으로 부호화 특징을 구분할 수 있다. 전자의 경우에는 I picture와 P picture만으로 영상을 부호화하고, 후자의 경우에는 I, P, B picture를 모두 사용하여 영상의 움직임을 예측/보상한 후 이를 부호화한다.

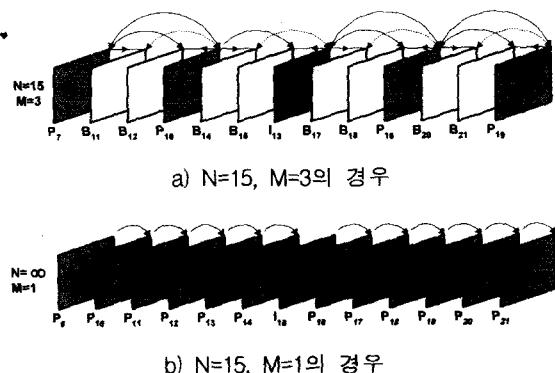


그림 3 입력 영상의 부호화 순서와 디스플레이 순서

Fig. 3 Coding and display order of input sequences

그림 3은 움직임 예측 방식에 따라 입력 영상에 대한 부호화 순서 및 복호화 후 디스플레이 되는 순서를 도시한 것이다. 그림 3 a)는 I picture와 P picture의 주기가 각각 15, 3 ( $N=15$ ,  $M=3$ )으로 부호화된 영상으로서 이미 부호화된 동영상의 MV는 각각의 영상마다 다르다. I picture는 MV를 수반하지 않고, P picture는 3번째 앞의 I 또는 P picture를 기준으로 전방향 MV를 추정하며 B picture는 1번째 혹은 2번째 앞, 뒤의 I 또는 P picture를 기준으로 양방향 MV를 추정한다. 반면에 그림 3 b)와 같이 부호화 모드가 I, P picture로 부호화 ( $N=15$ ,  $M=1$ )된 영상의 경우, P picture는 바로 앞의 I 또는 P picture를 기준으로 전방향 MV를 추정한다. 따라서 I, P, B picture로 이미 부호화된 영상을 I, P picture의 부호화 형식을 갖는 영상으로 트랜스코딩 해야 할 경우, 각기 다른 P, B picture의 MV를 사용하여 P picture의 MV 정보를 추정해야 한다[5][6].

이를 위해서 이종 동영상 트랜스코더는 각각의 영상들의 움직임을 같은 속도의 전체적인 움직임의 일부분으로 간주하여 앞, 뒤 영상들 사이의 움직임이 일정하다는 가정하에 새로운 MV를 추정한다. 추정된 MV의 수는 picture의 종류에 따라 다르며 추정된 MV를 중에서 부호화 오차에 대한 절대차의 합을 계산하여 가장 작은 오차를 가진 벡터를 선택한다.

#### (2) 부호화 크기 변환을 위한 움직임 벡터 추정

이미 부호화된 영상에 대해 영상의 크기 변환을 위한 트

랜스코딩이 필요할 경우가 발생한다. 특히 이미 부호화된 영상의 비트율과 목적 비트율의 차이가 매우 큰 경우나 수신단의 해상도에 한계가 있는 경우에는 비트율을 변환하기 전에 영상의 크기를 먼저 다운 샘플링(down sampling)하는 것이 일반적이다[4][7].

그림 4는 영상의 크기를 수평, 수직 각각 1/2로 축소하는 경우의 MV의 성질을 도시한 것이다. 이 경우, 이미 부호화된 영상의 4개의 매크로블록이 축소 변환된 영상의 1개의 매크로블록에 대응된다. 따라서 4개의 MV를 이용하여 새로운 MV 1개를 추정해야 한다.

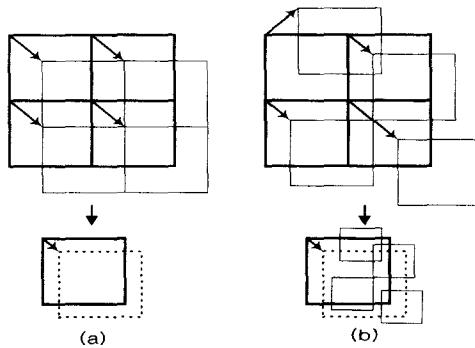


그림 4 1/2 크기 축소 움직임벡터

Fig. 4 Scale-down motion vector by two

그림 4 a)와 같이 4개의 모든 MV가 한 방향으로 잘 정렬되어 있는 경우에는 축소 영상의 MV를 추정하기가 비교적 용이하지만 그림 4 b)와 같이 4개의 MV가 각기 다른 경우, 4개의 MV와 매크로블록의 특성을 이용하여 새로운 벡터를 추정할 필요가 생기게 된다. 이러한 새로운 MV 추정법으로는 중간 위치 벡터 추정법, 평균 벡터 추정법 및 AWW(Align to Worst Weighting)법이 많이 사용되고 있다 [8].

#### • 중간 위치 벡터 추정법

이미 부호화된 영상의 4개의 공간적으로 인접한 매크로블록의 MV를  $V = \{v_1, v_2, v_3, v_4\}$ 이라 하고 각각의 1개의 MV에 대한 나머지 3개의 MV의 유클리디안 거리의 합을 식 (1)을 통해 얻는다.

$$d_i = \sum_{j=1, j \neq i}^4 \|v_i - v_j\| \quad (1)$$

이때 중간 위치 벡터는 4개의 MV 중에서 식 (2)와 같이 가장 작은 유클리디안 거리를 가지는 MV로 정의된다.

$$\text{med}(V) = v_k \in V \quad \min d_i = d_k \quad (2)$$

식 (2)로부터 얻어진 MV는 공간적인 거리상 나머지 벡터들 중간에 위치하게 되며 영상의 크기가 축소되면 같은 비율로 새로운 MV의 크기를 줄여야 한다.

#### • 평균 벡터 추정법

평균 벡터 추정법은 식 (3)과 같이 4개의 MV의 수평, 수직 성분의 산술 평균으로부터 얻어진다.

$$V = \frac{1}{4} \sum_{i=1}^4 v_i \quad (3)$$

4개의 MV가 한 방향으로 잘 정렬되어 있는 경우에는 매우 효과적이지만, 반대로 한 방향으로 잘 정렬되어 있지 않거나 4개의 MV중의 하나의 크기가 다른 3개의 MV의 크기보다 월등하게 크면 축소 영상에 대한 효과적인 움직임 추정을 할 수 없는 문제점이 발생한다.

#### • AWW 추정법

AWW법은 식 (4)와 같이 4개의 MV에 대한 적절한 가중치를 두어 축소 영상의 MV를 추정하는 방법이다.

$$\overline{mv}' = \frac{1}{2} \cdot \frac{\sum_{i=1}^4 \overline{mv}_i \cdot W_i}{\sum_{i=1}^4 W_i} \quad (4)$$

여기서  $\overline{mv}_i$ 와  $\overline{mv}'$ 은 각각 이미 부호화된 영상 및 축소된 영상에서의 MV를 나타내며  $W_i$ 는 각각의 MV에 대한 가중치를 나타낸다.

일반적으로 가중치 값은 설정할 때에는 식 (5)와 같이 축소된 영상의 매크로블록 위치에 해당하는 이미 부호화된 영상의 4개의 매크로블록의 공간 활성도(activity)를 고려한다.

$$W_i = \sum_{j=0}^3 \sum_{k=1}^{63} d(b_j(k)) \quad (5)$$

$$\begin{aligned} d(x) &= 0, & \text{if } x = 0 \\ &= 1, & \text{otherwise} \end{aligned}$$

식 (5)에서  $b_j(k)$ 는 이미 부호화된 영상의 비트열에서 얻은 원영상과 예측 영상과의 차영상에 대한 DCT 후의 양자화된 AC 계수로서, 예측 오차가 큰 매크로블록의 MV에 대해 가중치를 크게 하여 축소 영상에서의 추정 MV에 의한 오차를 작게 만들어 준다.

### 3. 제안된 움직임 벡터 추정법

이미 부호화된 영상정보를 수신단의 환경에 맞게 트랜스코딩할 경우, 목적 비트율의 변환만을 필요로 할 수 있으며 크기의 변환이 함께 수행되는 경우도 발생한다. 또한 수신측이 부호화 형식의 변환을 요구할 경우도 배제할 수 없다.

그림 5는 본 논문에서 제안한 영상 크기 및 부호화 형식 변환을 위한 이종 동영상 트랜스코더의 블록도를 나타낸 것이다. 일반적으로 동영상의 움직임은 인접한 시간과 공간에 대해 서로 큰 상관관계를 가지므로 트랜스코딩되는 영상의 MV는 이미 부호화된 영상의 MV로부터 추정할 수 있다.

#### 3.1 부호화 형식 변환을 위한 움직임 벡터 추정

I, P, B picture로 구성된 이미 부호화된 영상정보를 I, P

picture의 영상정보로 변환하는 이종 동영상 트랜스코더에서는 가능한 여러 가지 움직임 벡터를 추정한 후, 움직임 보상을 통해 영상을 재구성하여 최소 오차를 발생시키는 움직임 벡터를 선택하는 방법을 취하고 있다. 따라서 움직임 추정 방식에 따른 매크로블록의 재구성과 복호화된 매크로블록의 오차를 계산하는 데에는 많은 시간 지연과 연산이 필요하게 된다. 따라서 본 논문에서는 오차 계산의 시간적 소모와 B picture를 P picture로 재부호화하기 위한 시간 지연 및 연산량을 최소화하는 방법을 다음과 같이 3가지로 제안한다.

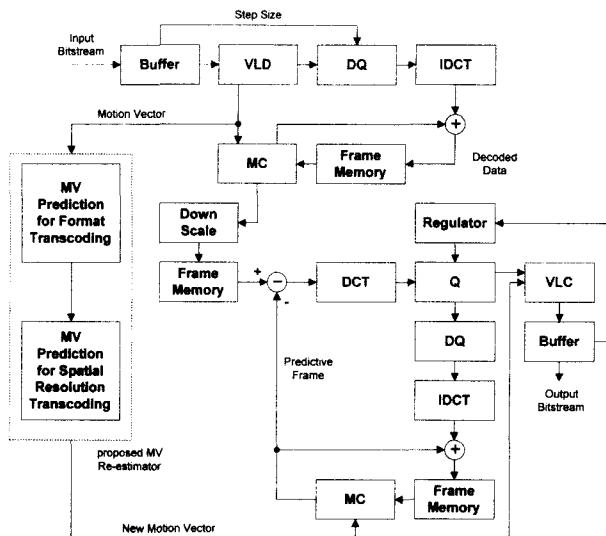


그림 5 제안된 transcoder 구조  
Fig. 5 Proposed trancoder block diagram

### (1) 방법 1 : P picture의 MV를 이용, 새로운 MV 추정

가장 간단한 MV 추정법으로 소그룹내의 P picture의 MV로 소그룹내의 다른 영상들, 즉 P, B picture의 MV를 추정한다. 그림 3 a)의 B<sub>11</sub>, B<sub>12</sub>, P<sub>10</sub>을 그림 3 b)의 P<sub>10</sub>, P<sub>11</sub>, P<sub>12</sub>의 영상으로 변환시키는 과정에서 새로운 MV는 소그룹내의 가장 먼저 부호화되어 있는 P picture의 MV를 이용하므로 상대적으로 나중에 부호화되어 있는 B picture의 MV 정보 없이도 MV의 추정이 가능하다. 따라서 B picture로 인한 시간 지연은 제거될 수 있다.

#### ① $V_{9 \rightarrow 10}^{out}$ , $V_{10 \rightarrow 11}^{out}$ 추정

소그룹내 첫 번째 부호화되어 있는 P picture의 MV의

$$1/3 : \frac{1}{3} V_{7 \rightarrow 10}^{fwd}$$

#### ② $V_{11 \rightarrow 12}^{out}$ 추정

소그룹내 첫 번째 부호화되어 있는 P picture의 MV와 추정된  $V_{8 \rightarrow 9}^{out}$ 과  $V_{9 \rightarrow 10}^{out}$ 의 합과의 차 :

$$V_{7 \rightarrow 10}^{fwd} - (V_{8 \rightarrow 9}^{out} + V_{9 \rightarrow 10}^{out})$$

위의 표현에서  $V^{out}$ 은 새롭게 추정해야 할 MV를 의미하며 화살표의 앞과 뒤는 각각 MV 추정시 사용되는 원영상

과 목적 영상의 번호이며 또한  $fwd$ 는 전방향 예측을 나타낸 것이다.

또한 I picture에는 MV가 없으므로 I picture가 포함된 소그룹의 B와 I picture는 바로 이전에 추정된 MV를 그대로 사용한다. 즉, 그림에서 I<sub>13</sub>의 MV는 존재하지 않으므로  $V_{12 \rightarrow 13}^{out}$ ,  $V_{13 \rightarrow 14}^{out}$ ,  $V_{14 \rightarrow 15}^{out}$ 는 모두 이전에 추정한  $V_{11 \rightarrow 12}^{out}$ 의 값을 사용한다.

이러한 추정법은 연속된 3개의 영상의 MV를 추정하기 위해서 1번의 추정 연산만이 필요하므로 시간적 지연은 제거될 수 있지만, B picture의 움직임의 변화가 일정하지 않을 경우 MV의 오차가 크게 되는 단점이 있다. 또한 P picture의 매크로블록 부호화 모드중에 Intra 모드가 많이 포함되어 있을 경우에는 시간적으로 인접한 매크로블록의 MV를 추정할 수 없는 경우도 발생할 수 있다.

### (2) 방법 2 : P, B picture의 MV를 이용, 부호화 순서로 트랜스코딩의 새로운 MV 추정

B picture를 포함하고 있는 영상은 복호화되어 화면으로 출력되는 순서와 부호화하는 순서가 다르다. 따라서 디스플레이를 위해서는 이미 부호화된 영상을 먼저 복호화해야 하고 트랜스코더에서도 복호화의 순서는 같다. 이 과정에서 구조적으로 먼저 복호화된 영상을 임시 저장할 수 있는 저장 공간 필요하고 시간적으로 지연이 발생한다. 따라서 만일 트랜스코더에 입력되는 비트열의 순서, 즉 부호화된 영상의 순서대로 MV를 추정할 수 있으면 트랜스코더의 구조는 간단해지고 시간 지연을 제거할 수 있다.

그림 3 a)에서 트랜스코더에 입력되는 소그룹내의 영상의 순서는 B<sub>11</sub>, B<sub>12</sub>, P<sub>10</sub>…이 아니라 P<sub>10</sub>, B<sub>11</sub>, B<sub>12</sub>…의 순서이다. 이들 영상에 대응되는 그림 3 b)의 영상은 P<sub>10</sub>, P<sub>11</sub>, P<sub>12</sub>이 되고 따라서 순서대로 P<sub>10</sub>, B<sub>11</sub>, B<sub>12</sub>의 MV를 각각 이용하여 새로운 MV인  $V_{8 \rightarrow 9}^{out}$ ,  $V_{9 \rightarrow 10}^{out}$ ,  $V_{10 \rightarrow 11}^{out}$ 을 추정할 수 있으면 구조적으로 복호화된 P picture를 저장할 공간이 따로 필요하지 않으며 시간 지연도 발생하지 않는다.

#### ① $V_{9 \rightarrow 10}^{out}$ 추정

- 소그룹내 첫 번째 부호화되어 있는 P picture의 MV의  $1/3 : \frac{1}{3} V_{7 \rightarrow 10}^{fwd}$

#### ② $V_{10 \rightarrow 11}^{out}$ 추정

- 소그룹내 첫 번째 B picture의 매크로블록이 전방향 예측인 경우 : 자신의 전방향 MV인  $V_{7 \rightarrow 11}^{fwd}$

- 소그룹내 첫 번째 B picture의 매크로블록이 후방향 예측인 경우 : 자신의 후방향 MV의  $-1/2$ 인  $-\frac{1}{2} V_{10 \rightarrow 11}^{bwd}$

- 소그룹내 첫 번째 B picture의 매크로블록이 양방향 예측인 경우 : 자신의 전방향과 후방향 MV의  $-1/2$ 의 산술 평균인  $\frac{1}{2} (V_{7 \rightarrow 11}^{fwd} - (\frac{1}{2} V_{10 \rightarrow 11}^{bwd}))$

#### ③ $V_{11 \rightarrow 12}^{out}$ 추정

- 소그룹내 두 번째 B picture의 매크로블록이 전방향 예측인 경우 : 자신의 전방향 MV의  $1/2$ 인  $\frac{1}{2} V_{7 \rightarrow 12}^{fwd}$

- 소그룹내 두 번째 B picture의 매크로블록이 후방향 예측인 경우 : 자신의 후방향 MV의 -1인  $-V_{10 \rightarrow 12}^{bwd}$
- 소그룹내 첫 번째 B picture의 매크로블록이 양방향 예측인 경우 : 자신의 후방향과 전방향 MV의 1/2의 산술 평균인  $\frac{1}{2}(-(\frac{1}{2})V_{7 \rightarrow 12}^{fwd} + V_{10 \rightarrow 12}^{bwd})$

I picture는 MV가 없으므로, 바로 전 영상의 MV를 그대로 다시 사용한다. 방법 2는 이미 부호화 과정에서 배치한 출력 영상의 순서를 트랜스코더에 입력되는 영상 순서로 재 배열하는 것이다. 따라서 사용된 MV는 정확하게 바로 다음 영상의 MV이므로 비록 오차가 발생할 가능성은 있지만 시간적으로 인접한 영상의 MV는 서로 높은 상관관계를 가지고 있으므로 정제과정을 거쳐 충분히 보상해 줄 수 있다. 또한 방법 1과는 달리 B picture의 MV를 사용함으로써 I picture가 포함된 소그룹에서도 보다 정확한 MV의 추정이 가능하다.

### (3) 방법 3 : P, B picture의 MV를 이용, 복호화 순서로 트랜스코딩의 새로운 MV 추정

방법 2에서 제시된 방법과 거의 동일하지만, 입력되는 영상의 복호화 순서대로 그림 3 a)에서의 B picture의 MV를 그림 3 b)에서의 P picture의 MV로 전환시키는 방법이다. 따라서 복호화된 P picture의 임시 저장 공간이 필요하고 또한 시간적 지연이 발생하게 된다. 그러나 방법 2와 비교해서 동일한 복호화 순서의 영상의 MV를 이용하기 때문에 최적의 MV에 더 근접한 MV의 추정이 가능하다.

#### ① $V_{9 \rightarrow 10}^{out}$ 추정

- 소그룹내 첫 번째 B picture의 매크로블록이 전방향 예측인 경우 : 자신의 전방향 MV인  $V_{7 \rightarrow 11}^{fwd}$
- 소그룹내 첫 번째 B picture의 매크로블록이 후방향 예측인 경우 : 자신의 후방향 MV의 -1/2인  $-\frac{1}{2}V_{10 \rightarrow 11}^{bwd}$
- 소그룹내 첫 번째 B picture의 매크로블록이 양방향 예측인 경우 : 자신의 전방향과 후방향 MV의 -1/2의 산술 평균인  $\frac{1}{2}(V_{7 \rightarrow 11}^{fwd} - (\frac{1}{2})V_{10 \rightarrow 11}^{bwd})$

#### ② $V_{10 \rightarrow 11}^{out}$ 추정

- 소그룹내 두 번째 B picture의 매크로블록이 전방향 예측인 경우 : 자신의 전방향 MV의 1/2인  $\frac{1}{2}V_{7 \rightarrow 12}^{fwd}$
- 소그룹내 두 번째 B picture의 매크로블록이 후방향 예측인 경우 : 자신의 후방향 MV의 -1인  $-V_{10 \rightarrow 12}^{bwd}$
- 소그룹내 첫 번째 B picture의 매크로블록이 양방향 예측인 경우 : 자신의 후방향과 전방향 MV의 -1/2의 산술 평균인  $\frac{1}{2}(-(\frac{1}{2})V_{7 \rightarrow 12}^{fwd} + V_{10 \rightarrow 12}^{bwd})$

#### ③ $V_{11 \rightarrow 12}^{out}$ 추정

- 소그룹내 첫 번째 부호화되어 있는 P picture의 MV의 1/3인  $\frac{1}{3}V_{7 \rightarrow 10}^{fwd}$

### 3.2 크기 변환을 위한 움직임 벡터 추정

이미 부호화된 영상의 크기가 축소되면 인접한 여러개의 매크로블록이 하나의 매크로블록에 대응되고, 따라서 여러개의 MV로부터 하나의 MV를 추출해야 한다. 기존의 평균 벡터 추정법, 중간 위치 벡터 추정법 및 AWW 추정법 등은 트랜스코딩을 위한 움직임 예측시의 계산량을 크게 줄일 수 있지만, 최적의 MV의 추정에는 다소 미흡하다. 따라서 본 논문에서는 계산량은 기존의 방법들과 비슷하고, MV들의 특성의 차이가 크더라도 차선의 MV를 추정함으로써 화질의 향상을 얻을 수 있는 최소 오차 추정법을 제안한다.

즉, 이미 부호화된 영상의 공간적으로 인접한 4개의 MV로 이루어진 공간이  $V_s = \{v_1, v_2, v_3, v_4\}$ 이라고 가정하고 4개의 MV 모두를 추정 가능한 MV의 후보로 선정한다. 각각의 MV를 이용하여 축소된 영상들과의 오차를 계산한 후 최소 오차가 발생하도록 하는 움직임 벡터를 축소영상의 추정 벡터로 최종 선택한다.

$$e_i = \sum_{m=0}^{15} \sum_{n=0}^{15} |mb_o(m, n) - mb_{mci}(m, n)| \quad (6)$$

여기서  $mb_{mci}$ ,  $mb_o$ 는 각각  $v_i$ 로 움직임 보상을 해준 매크로블록과 원영상의 매크로블록을 나타낸다.  $e_i$ 는  $v_i$  벡터로 예측된 매크로블록과 원영상 매크로블록과의 절대차의 합이다.

$$\text{Minimun Error}(V_s) = v_k \in V_s \quad \min e_i = e_k \quad (7)$$

기존의 수학적인 추정법과 달리 축소 영상의 추정 MV는 이미 부호화된 영상의 4개의 MV를 이용하여 움직임 예측 오차를 비교하여 선택되어지므로, 전역 움직임 추정법에서 얻은 최적의 벡터와 근접할 확률이 높아지게 된다.

## 4. 실험 결과 및 고찰

### 4.1 실험 환경

실험에 사용된 동영상은 CCIR601 4:2:0-525(704×480) 및 SIF(352×240) 규격의 표준 동영상 테스트 시퀀스인 Mobile & Calendar 영상 및 Football 영상 각 100 frame이다. 실험은 부호화 형식 변환의 트랜스코더를 위한 MV 추정, 목적 영상의 크기 변환의 트랜스코더를 위한 MV 추정과 부호화 형식 및 크기 변환의 트랜스코더를 위한 MV 추정으로 나누어 종속 트랜스코더와 본 논문에서 제안한 방식의 성능을 비교, 평가하였다.

제안 방식의 성능은 PSNR로 화질 평가를, 그리고 종속 트랜스코더에서 산출된 MV와 제안 방식에서 산출된 MV의 차이를 통해 MV의 정밀도 평가를 시행하였으며 마지막으로 MV 추정시 수행되는 연산량을 덧셈, 곱셈 그리고 비트연산의 개수로 비교, 평가하였다. 일반적으로 곱셈연산은 덧셈연산의 3배의 계산량과 같고 비트연산은 덧셈연산과 같은 계산량을 가지므로 총 수행 연산량은 식 (8)과 같다.

$$\text{Total\_Op} = (\text{Add\_Op}) + (3 \times \text{Multi\_Op}) + (\text{Shift\_Op}) \quad (8)$$

여기서  $\text{Add\_Op}$ ,  $\text{Multi\_Op}$ ,  $\text{Shift\_Op}$ 는 각각 덧셈, 곱셈, 비트연산의 개수이다.

## 4.2 결과 고찰

### (1) 부호화 형식 변환 트랜스코더의 MV 추정

표 1과 2는 각각 SIF 영상을  $N=15$ ,  $M=3$ , 목적 비트율 1.15Mbps로 MPEG1 SM3 방식에 의해 이미 부호화된 영상을 목적 비트율 변환없이  $N=15$ ,  $M=1$ 로 부호화 형식을 변환할 때, 종속 트랜스코더와 제안된 방식의 PSNR 및 MV 오차를 나타낸 것이다.

표 1 부호화 형식 변환 영상의 평균 PSNR (단위 dB)

Table 1 Average PSNR of format converted sequences

실험영상	MV 추출방법	방법 1	방법 2	방법 3	종속 트랜스코더
Mobile& Calendar	추정기법	21.44	22.14	22.14	22.68
	정제 범위	0.5	22.43	22.65	
	1.5	22.56	22.67	22.67	
	2.5	22.66	22.67	22.67	
Football	추정기법	25.43	26.16	26.22	26.99
	정제 범위	0.5	26.43	26.58	
	1.5	26.65	26.80	26.84	
	2.5	26.76	26.87	26.88	

표 2 부호화 형식 변환 영상의 움직임벡터 오차 비교

Table 2 MV error comparison of format converted  
sequences (단위 %)

실험영상	오 차		방법 1	방법 2	방법 3
Mobile & Calendar	오차 범위	±0.5이내	89.17	88.98	90.15
		±1.5이내	99.16	99.39	99.42
Football	오차 범위	±0.5이내	68.30	68.75	71.36
		±1.5이내	80.00	82.04	84.84

표 1의 결과에서 알 수 있듯이 제안 방법들은 종속 트랜스코더에 비해 약 0.5 ~ 1.25dB의 화질 감소를 보였으나 제안 방법들로부터 추정된 MV에 대해 ±2.5 화소의 움직임 정제과정을 수행할 경우, 방법 3은 종속 트랜스코더와 거의 대등한 화질을 보였다. 또한 표 2는 종속 트랜스코더에서

산출된 MV에 대해 제안 방법들로부터 추정된 MV의 정밀도를 각 오차 범위내에서 백분율로 표현한 것으로서 Mobile & Calendar 영상의 경우, 99% 이상의 MV가 ±1.5 화소의 오차 범위에 분포하는 결과를 보였다. 수행 연산량의 경우, 각 제안 방법들은 종속 트랜스코더에 비해 0.4 ~ 0.5%의 연산량을 필요로 하며 ±2.5 화소의 움직임 정제과정을 수행 하더라도 약 15%의 연산량으로 종속 트랜스코더와 대등한 성능을 보일 수 있었다.

### (2) 영상 크기 변환 트랜스코더의 MV 추정

표 3과 4는 각각 CCIR-601(704×480)의 영상을  $N=15$ ,  $M=3$ , 목적 비트율 5Mbps로 MPEG2 TM5 방식에 의해 이미 부호화된 영상을 부호화 형식 변환없이 목적 비트율 1.5Mbps, SIF 크기의 영상으로 크기를 축소할 때, 종속 트랜스코더와 제안된 최소 오차 벡터 추정법의 PSNR 및 MV 오차를 나타낸 것이다.

표 3 크기 축소 영상의 평균 PSNR (단위 dB)

Table 3 Average PSNR of resolution reduced sequences

실험영상	MV 추출방법	최소 오차	AWW	평균 벡터	종속 트랜스코더
Mobile& Calendar	추정기법	21.99	21.81	21.63	22.73
	정제 범위	0.5	22.67	22.54	
	1.5	22.68	22.63	22.65	
	2.5	22.68	22.65	22.66	
Football	추정기법	26.21	26.05	26.06	27.04
	정제 범위	0.5	26.97	26.76	
	1.5	26.99	26.92	26.93	
	2.5	26.99	26.96	26.97	

표 4 크기 축소 영상의 움직임벡터 오차 비교

Table 4 MV error comparison of resolution reduced  
sequences (단위 %)

실험영상	오 차		최소오차	AWW	평균벡터
Mobile & Calendar	오차 범위	±0.5이내	96.69	90.59	94.07
		±1.5이내	99.69	98.18	98.84
Football	오차 범위	±0.5이내	83.37	69.67	69.36
		±1.5이내	93.14	84.09	85.08

표의 결과에서 알 수 있듯이 크기 축소 트랜스코더의 MV 추정을 위한 제안된 최소 오차 벡터 추정법은 기존 방법들에 비해 평균적으로 0.2dB의 화질 향상을 보이고 있으

며, 추정된 MV 또한 종속 트랜스코더의 MV에 대해  $\pm 1.5$  화소의 오차 범위에 93% 이상 분포하는 등 기존 방법들보다 평균 7%이상의 정밀도 향상의 결과를 보였다.

### (3) 부호화 형식, 크기 변환 트랜스코더의 MV 추정

그림 6과 7은 각각 CCIR-601(704×480)의 Mobile&Calendar 영상을 N=15, M=3, 그리고 목적 비트율 5Mbps로 MPEG2 TM5 방식에 의해 이미 부호화된 영상을 N=15, M=1로 B picture 없이 형식 변환을 하고 또한 목적 비트율 1.5Mbps, SIF 크기의 영상으로 크기를 축소할 때, 종속 트랜스코더와 부호화 형식 변환에 대한 제안된 방법 3과 최소 오차 벡터 추정법으로 구해진 PSNR 및 MV 오차를 나타낸 것이고 그림 8은 각 방식의 연산량을 나타낸 것이다.

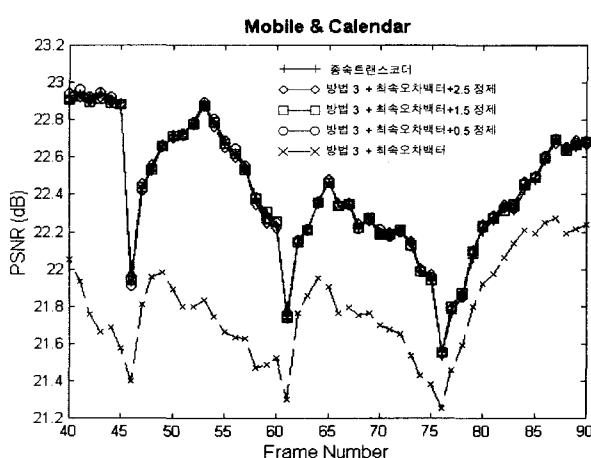


그림 6 부호화 형식 및 크기 변환의 PSNR 결과  
Fig. 6 PSNR of format/resolution reduced transcoder

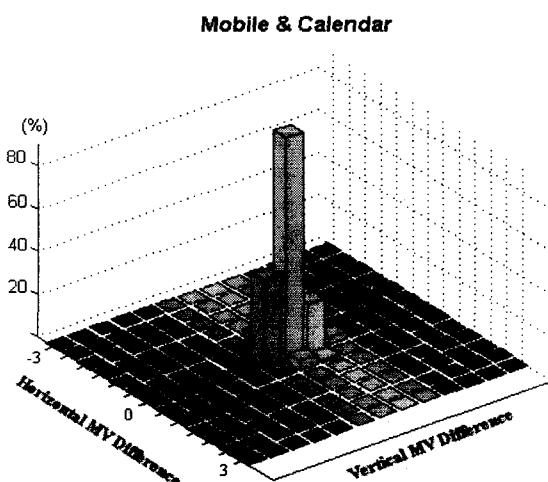


그림 7 부호화 형식 및 크기 변환의 MV 오차 분포  
Fig. 7 MV Error distribution of format/resolution reduced transcoder

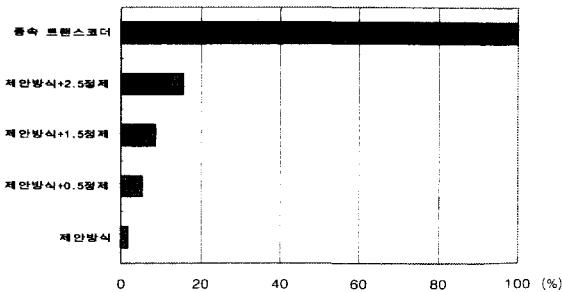


그림 8 부호화 형식 및 크기 변환의 연산량 비교  
Fig. 8 Operation count of format/resolution reduced transcoder

위의 결과에서 알 수 있듯이 이미 부호화된 영상의 비트열에서 추출한 MV를 이용하여 제안된 방법 3으로 부호화 형식을 변환하고 최소 오차 추정법으로 영상의 크기를 축소하는 트랜스코더를 구현할 경우, 제안된 방법에 의해 추정된 MV에 대해  $\pm 2.5$  화소의 움직임 정제과정을 수행하면 연산량은 종속 트랜스코더에 비해 약 15% 정도로 줄이면서 MV의 정밀도는 99%이상으로 접근하여 결과적으로 종속 트랜스코더의 화질에 거의 균접하는 좋은 화질 특성을 보였다.

## 5. 결 론

본 연구에서는 I, P, B picture의 부호화 모드를 갖는 이미 부호화된 MPEG 기반 동영상 비트열에 대해 I, P picture의 부호화 모드로만 부호화 형식을 변환하고 또한 영상의 크기가 수평, 수직 각 1/2로 축소되는 이종 동영상 부호화기의 트랜스코더를 위한 새로운 MV 재추정 기법을 제안하였다.

기존의 종속 트랜스코더는 화질 특성은 우수하지만 복호화와 부호화 전 과정을 수행하므로 새로운 MV 추정을 위해 많은 연산량 및 시간 지연이 필요하다. 제안된 방법은 이미 부호화된 영상의 비트열에 포함되어 있는 MV를 이용하여 목적한 트랜스코더에 적합한 새로운 MV를 추정함으로서 종속 트랜스코더에 비해 화질 열화 없이 연산량 및 시스템 복잡도를 대폭 감소시킬 수 있어서 부호화 형식 및 크기 축소 변환을 위한 이종 동영상 부호화기의 트랜스코더로서의 실용화 가능성을 확인하였다.

## 참 고 문 현

- [1] N. Farvardin, "Optimum Quantizer Performance for a Class of non-Gaussian Memoryless Source", IEEE Trans. Inform. Theory, Vol. 30, pp. 485-497, May 1984.
- [2] ISO/IEC13818-2, "Generic Coding of Moving Pictures and Associated Audio, Recommendation H.262 MPEG2", International Standard, 1995
- [3] V. Bhaskaran and K. Konstantinides, Image and Video Compression Standards, Algorithms and

- Architectures. Boston, MA: Kluwer Academic, 1995.
- [4] B. Shen, I. K. Sethi, and B. Vasudev, "Adaptive Motion-Vector Resampling for Compressed Video Downscaling," IEEE Trans. Circuits and Systems for Video Technology, Vol. 9, No. 6, Sep. 1999.
- [5] T. Shanableh and M. Ghanbari, "Heterogeneous Video Transcoding to Lower Spatio-Temporal Resolutions and Different Encoding Formats," IEEE Trans. Multimedia, Vol. 2, No. 2, pp. 101-110, June 2000.
- [6] S. J. Wee, J. G. Apostolopoulos, and N. Feamster, "Field-to-frame Transcoding with Spatial and Temporal Downsampling", Proceedings of ICIP 99, Vol. 4, pp. 271-275, 1999.
- [7] N. Bjok and C. Christopoulos, "Transcoder Architecture for Video Coding", IEEE Trans. Consumer Elec., Vol. 44, pp 88-98, Feb. 1998.
- [8] K. H. Tan and M. Ganbari, "Layered Image Coding using the DCT Pyramid", IEEE Trans. Image Processing, Vol. 4, pp 512-516, Apr. 1995.

## 저 자 소 개



### 윤 규 섭 (尹 圭 涵)

1972년 9월 17일 생. 1999년 연세대학교 전기공학과 졸업. 2001년 동 대학원 전기·컴퓨터공학과 졸업(석사). 1998년~현재 LG전자 정보통신 주임연구원  
Tel : 031-450-4043  
Fax : 031-450-2011  
E-mail : imque@lge.com



### 박 강 서 (朴 儉 繙)

1973년 2월 10일 생. 1996년 연세대학교 전기공학과 졸업. 1998년 동 대학원 전기공학과 졸업(석사). 1998년~현재 농대학원 전기전자공학과 박사과정  
Tel : 02-2123-2768  
Fax : 02-312-7735  
E-mail : psk@cyber.yonsei.ac.kr



### 정 태 윤 (鄭 泰 允)

1964년 11월 21일 생. 1987년 연세대학교 전기공학과 졸업. 1989년 동 대학원 전기공학과 졸업(석사). 2000년 동 대학원 전기·컴퓨터공학과 졸업(공박). 1989년~1996년 삼성종합기술원 선임연구원. 1996~2001년 삼성전자 중앙연구소 책임연구원. 2001년~현재 강릉대학교 정보전자공학부 제어계 측공학과 전임강사  
Tel : 033-640-2425, Fax : 033-640-2244  
E-mail : tychung@kangnung.ac.kr



### 박 상 희 (朴 相 喜)

1939년 8월 25일 생. 1962년 연세대학교 전기공학과 졸업. 1964년 동 대학원 전기공학과 졸업(석사). 1971년 동 대학원 전기공학과 졸업(공박). 1982년~1984년 Washington Univ. 방문교수. 1998년~1999년 대한전기학회 회장. 1970년~현재 연세대학교 전기전자공학과 교수. 2002년~현재 연세대학교 정보대학원 원장  
Tel : 02-2123-2768 Fax : 02-312-7735  
E-mail : psh@yonsei.ac.kr