

논문-03-08-3-08

## 스캔 포맷 변환이 있는 효율적인 MPEG-2 동영상 트랜스코딩을 위한 고속 움직임 추정 기법

송병철\*, 천강욱\*

### Fast Motion Estimation Algorithm for Efficient MPEG-2 Video Transcoding with Scan Format Conversion

Byung Cheol Song\* and Kang Wook Chun\*

#### 요 약

ATSC (Advanced Television System Committee)에서는 18가지의 DTV (Digital Television)를 위한 동영상 포맷들을 제정하였다. ATSC 포맷들은 스캔 포맷, 크기 포맷, 프레임율 포맷 등의 적당한 조합들이다. 효과적인 MPEG-2 트랜스코딩을 위해서는 이런 포맷들 간의 변환도 지원할 수 있어야 한다. 여러 포맷 변환들 중에서 특히 스캔 포맷 변환은 프레임율과 크기 변화가 함께 일어나기 때문에 상대적으로 구현하기가 힘들다. 스캔포맷 변환으로 픽처 타입도 함께 변하기 때문에 트랜스코더의 움직임 추정부 (motion estimation: ME)에 상당한 연산량 부담을 주게 된다. 이런 문제를 해결하기 위해 본 논문은 스캔포맷 변환을 지원하는 MPEG-2 동영상 트랜스코딩을 위한 고속 움직임 추정 알고리즘을 제안한다. 먼저, 트랜스코더의 수신단은 입력 비트열로부터 재 부호화 (re-encoding)에 적합한 후보 움직임 벡터들을 추출한다. 그런 다음, 가중치 중간값 선택기(weighted median selector)를 이용하여, 여러 후보 움직임 벡터들 중에서 최적의 움직임 벡터를 선택한다. 모의 실험 결과는 제안한 움직임 추정 알고리즘이 전역 탐색 기법 (Full Search Algorithm: FSA)에 비해 현저하게 적은 연산량을 가지면서, FSA와 거의 동일한 PSNR 성능을 가짐을 증명한다.

#### Abstract

ATSC (Advanced Television System Committee) has specified 18 video formats for DTV (Digital Television), e.g., scan format, size format, and frame rate format conversion. Effective MPEG-2 video transcoders should support any conversion between the above-mentioned formats. Scan format conversion is hard to implement because it may often induce frame rate and size format conversion together. Especially, because of picture type conversion caused by scan format conversion, the computational burden of motion estimation (ME) in transcoding becomes serious. This paper proposes a fast ME algorithm for MPEG-2 video transcoding supporting scan format conversion. Firstly, we extract and compose a set of candidate motion vectors (MVs) from the input bit-stream to comply with the re-encoding format. Secondly, the best MV is chosen among several candidate MVs by using a weighted median selector. Simulation results show that the proposed ME algorithm provides outstanding PSNR performance close to full search ME, while reducing the transcoding complexity significantly.

Key Words: MPEG-2, transcoding, motion estimation, weighted median

## I. 서 론

\* 삼성전자 디지털미디어 연구소  
Digital Media R&D Center, Samsung Electronics Co., Ltd

기술이 급속히 발전함에 따라, 원거리 화상회의, VOD (Video-On-Demand), 원거리 교육 등의 동영상 서비스가 상당히 일상화되었다. 그러나 동영상 전송은 여전히 채널의 용량의 한계를 겪고 있다. 이로 인해, 이질적인 네트워크 환경 하에서 특정 비트율로 압축되어 있는 동영상 비트스트림을 채널이 허용하는 비트율로 변환하는 과정이 필요하게 되었다<sup>[1]</sup>. MPEG2나 MPEG4 같은 동영상 압축 표준들은 효과적인 비트율 변환을 위한 스케일러블 동영상 부호화 기법 (Scalable video coding) 들을 포함하고 있다<sup>[2]</sup>. 그러나, 일반적인 스케일러블 부호화는 상위 계층 (enhancement layer)의 수가 몇 개로 제한되어 있다<sup>[3]</sup>. 많은 네트워크 동영상 응용 분야들에서 보다 세밀한 스케일러빌리티가 요구되기 때문에, MPEG-4는 fine-granular scalable coding (FGS) 기술을 표준으로 채택하였다<sup>[4]</sup>. 그러나, 수신단의 동영상 복호기는 FGS 비트열을 복호하기 위해 상위 계층의 복호를 위한 추가적인 복호 모듈을 갖추어야 하는 단점이 있다.

위 문제를 해결하기 위한 대안으로 동영상 트랜스코딩 기법들이 제안되어 왔다. 동영상 트랜스코딩은 압축 비트열을 다른 비트율을 갖는 비트열로 변환하거나<sup>[5]</sup>, 다른 프레임 크기를 갖는 비트열로 변환하거나<sup>[6]</sup>, 다른 표준의 비트열로 변환시켜주는<sup>[7][8]</sup> 프로세스를 의미한다. 첫 번째나 두 번째 트랜스코딩 타입을 동질적인 트랜스코딩이라고 하며, 마지막 방식의 트랜스코딩을 이질적인 트랜스코딩이라고 부른다.

가장 간단한 트랜스코딩 구조 중 하나는 열린 루프 트랜스코딩 (open-loop transcoding)으로 DCT 계수들을 입력과 다른 양자화 계수 (QP)로 재 양자화 (re-quantization) 하는 방식이다<sup>[5]</sup>. 열린 루프 트랜스코딩은 완전한 복호 및 재 부호화 과정 없이 압축 영역에서 이루어지기 때문에, 간단하고 매우 빠르다. 그러나, 열린 루프 트랜스코딩은 원래 복원 영상들과 트랜스코더에서 사용되는 복원 영상들이 달라지기 때문에 drift 문제를 갖는다. Drift는 상당한 화질 저하를 유발한다. 복호기와 부호기가 cascade 형태로 연결된 트랜스코더인 소위 닫힌 루프 트랜스코더는 drift가 없는 트랜스코딩을 가능하게 한다 (그림 1 참조). 닫힌 루프 트랜스코더는 공간 영역 및 DCT 영역 모두에서 구현 가능하다. 일반적으로 DCT 영역 트랜스코더는 움직임 추정부에서 상당한 행렬 연산을 필요로 하기 때문에, 그 복잡도가 공간 영역 트랜스코더에 비해 큰 단점이 있다<sup>[5]</sup>. 그래서 본 논문에서는 공간 영역 닫힌 루프 트랜스코더만을 고려한다. 공

간 영역 트랜스코더는 입력 비트열이 완전히 복호된 후 복원된 영상을 가지고 재부호화를 수행한다. 그림 1에서 알 수 있듯이, 복호기로부터 얻어진 MV정보 및 모드 정보들을 이용하면 트랜스코더에서 움직임 벡터를 찾는 데 상당한 연산량을 줄일 수 있다<sup>[6]</sup>.

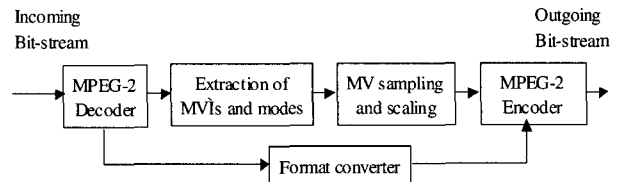


그림 1. 닫힌 루프 MPEG-2 동영상 트랜스코더.  
Fig. 1. A closed-loop MPEG-2 video transcoder.

일반적으로, FSA는 계산량이 많고 부호기 전체 연산량의 60%이상을 차지할 만큼 부호기에서 차지하는 비중이 매우 크다<sup>[9]</sup>. 그래서, 소프트웨어 동영상 트랜스코더들은 재부호기에서의 움직임 추정기로 FSA를 채택하지 않는다. 대신 복호기로부터 얻어진 움직임 벡터 정보들을 재사용하여 연산량을 줄이는 방법을 택한다. 그러나, 이런 단순한 MV 재사용은 그 부정확성 때문에 종종 심각한 화질 저하를 초래하기도 한다<sup>[10]</sup>. 이런 문제를 해결하기 위해 Shen은 적응적인 MV 재사용 방법을 제안하였다<sup>[6]</sup>. Sugano는 각 매크로블록(macroblock; MB)의 '공간적 활성화도'로서 그 QP값을 채택하여 새로운 움직임 벡터장을 구하는데 이용하였다<sup>[8]</sup>. Hashemi는 최대 평균 상관도(maximum average correlation; MAC) 방법을 제안하였다<sup>[11]</sup>. MAC에서는 입력 비트열로부터 네 개의 주변 움직임 벡터들 중에서 최대 MAC를 갖는 하나를 택한다.

한편, ATSC는 DTV를 위한 18가지 포맷을 규정하고 있다 (Table 1 참조). 임의의 복호기가 트랜스코딩된 비트열을 수신할 수 있도록 하기 위해서는 위 18가지 모든 포맷

표 1. ATSC DTV 포맷. 'I'는 비월 주사를 'P'는 순사 주사 방식을 의미한다.  
Table 1. ATSC DTV formats. 'I' stands for interlaced scan and 'P' stands for progressive scan

Vertical Lines1080	Pixels Per Line1920	Aspect Ratio	Frame Rate and Scan Format			
720	1280	16:9	60I		30P	24P
480	704	16:9		60P	30P	24P
480	640	16:9/4:3	60I	60P	30P	24P
		4:3	60I	60P	30P	24P

을 트랜스코더가 지원할 수 있어야 한다. 특히, 스캔 포맷 변환을 지원하는 동영상 트랜스코딩은 프레임율과 크기 포맷 변환이 동시에 발생하기 때문에 계산량이 많을 뿐만 아니라 구현도 어렵다.

본 논문은 비정수 (non-integer)의 프레임 크기 변환뿐 아니라 스캔 포맷 변환도 지원하는 동영상 트랜스코딩을 위한 고속 움직임 추정 기법을 제안한다. 먼저, 입력 비트열에서 움직임 벡터들을 추출하여, 재 부호화 포맷에 맞는 후보 움직임 벡터들을 구성한다. 다음으로, 가중치 중간값 선택기를 이용하여 여러 후보 움직임 벡터들로부터 최적의 움직임 벡터를 선택한다. 마지막으로 추가적인 추정 성능 향상을 위해 refinement 과정을 거친다.

본 논문은 다음과 같이 구성된다. II절에서는 스캔 포맷 변환을 갖는 MPEG-2 동영상 트랜스코딩을 소개한다. III절에서는 제안한 움직임 추정 방법을 기술하고, IV장에서 그 성능을 검증한다. 마지막 V장에서는 결론을 맺는다.

## II. 스캔 포맷 변환이 있을 때의 동영상 트랜스코딩

스캔 포맷 변환은 동영상 트랜스코딩에 있어 중요한 역할을 담당한다. 예를 들어, PC 상에서 HDTV 방송을 수신하기 위해서는, 입력 비월 방식의 동영상 비트열이 순사 방식 동영상 비트열로 변환되어야 한다. 이런 경우, 트랜스코딩에서 프레임 크기, 프레임율, 스캔 포맷 변환 등이 이루어져야 한다. 무엇보다도, 스캔 포맷이 변환되면, 입력과 출력의 픽처 타입이 서로 달라지기 때문에 ME 수행이 어려

워진다. 예를 들면, 입력 I-프레임이 재 부호화기에서는 B나 P 프레임으로 변환되어 MV재사용을 통한 움직임 추정이 불가능한 경우가 있다. 그림 2는 이런 문제점을 잘 설명하고 있다. 그림 2는 입력되는 각 필드가 재부호화될 때까지의 과정을 보인다. 비월 주사 방식의 입력 동영상의 group of pictures (GOP) 크기 (N)는 12이며, sub-GOP (M P프레임 간 거리)는 3, 그리고 출력 동영상의 스캔 포맷은 순사 방식이라고 가정한다. 또한, 출력의 M도 3이라고 하자. 그림 2에서  $I_0^t$ 와  $I_0^b$ 은 각각 0번째 프레임인 I-프레임의 탑 필드와 바텀 필드를 의미한다. 트랜스코더의 입력 영상들 (input pictures)은 부호화 순서를 따른다. 마찬가지로  $P_3^t$ 와  $P_3^b$ 은 각각 시간상으로 GOP 내 3번째 프레임인 P-프레임의 탑 필드와 바텀 필드를 의미한다. 트랜스코더 내 복호기를 거쳐 복호된 영상들 (decoded pictures)은 그림 2처럼 시간 순서에 맞게 재배열된다. 복호된 영상들은 다시 IPC를 거쳐 각 필드가 하나의 순사 방식 프레임으로 변환된다. 그런 다음 트랜스코더 내 부호기의 입력으로 입력된다. 그림 2에서 알 수 있듯이, 스캔 포맷이 변하게 되면, 픽처 타입이 바뀌는 문제가 발생한다. 예를 들어, 바텀 B 필드인  $B_1^b$ 은  $P_3$ 로 재부호화되며, 바텀 P 필드인  $P_3^b$ 은  $B_7$ 으로, 바텀 I 필드인  $I_0^b$ 은  $B_1$ 으로 재부호화된다. 이처럼, 픽처 타입이 바뀌므로 기존의 트랜스코딩 방식에서처럼 MV의 재사용에 어려움이 발생한다.  $I_0^b$ 의 경우 원래 어떤 MV도 없었으므로,  $B_1$ 으로 재부호화할 때 움직임 벡터의 재사용 자체가 불가능하다.

또한, 스캔 포맷이 변환되면, 입력과 출력의 움직임 보상 (motion compensation: MC) 타입도 일치하지 않게 된다.

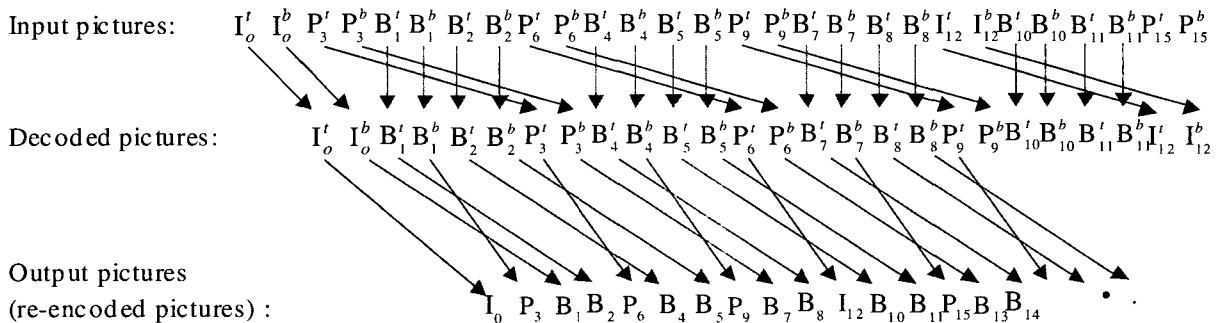


그림 2. 부호화 순서에 따른 입출력의 픽처 타입들. 도시 순서로 번호 매겨짐. 위첨자 't' 혹은 'b'는 탑 필드 및 바텀 필드를 의미한다. Fig. 2. Picture types of the input and output bit-streams in the encoding order, but numbered in the display order. Superscript 't' and 'b' denote top field and bottom field, respectively.

입력에서는 프레임 MC와 필드 MC가 공존하는 반면, 재부호화된 비트열에는 프레임 MC만이 존재한다. 상기 문제를 해결하기 위해서, 본 논문은 새로운 고속 움직임 추정 기법을 제안한다.

### III. 제안한 기법

그림 1에서처럼, 스캔 포맷 변환을 갖는 동영상 트랜스코더는 MPEG2 복호기, 포맷 변환기, 그리고, MPEG2 부호기 등 세 부분을 구성된다. 본 논문에서는 Interlaced-to-Progressive conversion (IPC)을 적용한 스캔 포맷 변환만을 고려한다. 그러나, 제안한 기법은 반대의 경우인 Progressive-to-Interlaced 변환에도 쉽게 적용 가능하다. 여기서는 MPEG2 TM5에 있는 간단한 IPC기법을 사용한다<sup>[3]</sup>. 공간영역에서의 down-sampling을 위해서는 잘 알려진 cubic convolution scaler를 사용한다<sup>[12][13]</sup>.

제안한 ME 기법은 크게 세 단계로 구성된다. 첫번째로, 재 부호화 포맷에 맞는 움직임 벡터 후보들을 입력 비트열에서 추출한다. 각 inter MB의 후보 벡터들은 복호기에서 추출 가능하다. 그러나, 출력 비트열의 어떤 프레임들은 픽처 타입 변환 때문에 어떤 움직임 벡터 후보들도 갖지 못한다. 그림 2의 예에서,  $I_0^b$ 는  $B_1$ 에 대응한다. 그래서,  $B_1$ 에 속한 모든 움직임 벡터들  $B_1$ 은  $I_0^b$ 에서 얻어지는 어떤 후보 벡터들도 갖지 못한다. 이런 프레임들에서는, 같은 sub-GOP 내의 주변 프레임들의 움직임 벡터들을 이용하여 보조 후보 움직임 벡터들을 구성한다. 결과적으로, 각 MB의 후보 움직임 벡터 집합은 두 그룹으로 나뉘는데, 하나는 현재 MB에 대응하는 원래 입력 MB들의 MV들이고, 다른 하나는 보조 후보 움직임 벡터들이다. 둘째로, 가중치 중간값 필터링을 이용하여 후보 움직임 벡터들 중에서 최적의 움직임 벡터를 선택한다. 마지막으로, 보다 향상된 화질 개선을 위해 refinement 과정을 거친다.

#### 1. 후보 움직임 벡터 집합의 구성

각 MB의 보조 후보 MV들은 sub-GOP 내의 주변 MB들의 MV들로부터 얻어진다. 보조 후보 MV들의 구성에 앞서, 필드 MV들은 프레임 MV들로 변환되어야 한다. 그런 다음, 보조 후보 MV들이 구해진다.

#### 1.1. 필드 MV들의 프레임 MV들로의 변환

입력의 각 MB의 MC 타입은 필드 MC이거나 프레임 MC가 된다. 그러나, 순차 방식 동영상에서는 프레임 MC만이 존재한다. 그래서, 필드 MV들이 프레임 MV로 변환되어야 한다. 입력 P-픽처에서, 필드 MC모드를 갖는 inter MB은 두 개의 MV를 갖는다. 바로 탑 필드 MV( $MV^t$ )와 바텀 필드 MV( $MV^b$ )이다. 또한, 각 필드 MB은 참조 필드가 탑 필드인지 바텀 필드인지를 나타내는 필드 선택 모드를 갖는다. 즉,  $MV^t$ 은 탑 필드의 참조 필드에서 찾아지거나 (T2T mode), 바텀 필드에서 찾아질 수도 있다 (T2B mode). 유사하게,  $MV^b$ 은 탑 필드의 참조 필드에서 찾아지거나 (B2T mode), 바텀 필드에서 찾아질 수도 있다 (B2B mode). 프레임 MC모드를 갖는 inter MB에 속한 탑 필드 MB 및 바텀 필드 MB은 각각 T2T 모드와 B2B 모드를 갖는다고 가정한다. 마지막으로, 탑 필드 MB이 움직임 벡터  $MV^t$ 을 가지면, 그 MV도 후보 벡터로 삼는다 (바텀 필드 MB의 경우,  $MV^b$ ). 그림 3은 필드 MC모드의 경우, 필드 MV들의 프레임 MV들로의 변환을 설명한다. 이 그림에서, 각 MB은 편의상 수직 방향만 나타나 있다. B필드 내 inter MB이 양방향 MC 모드를 가지면, 그 inter MB은 두 개의 후보MV (즉, 순방향 MV와 역방향 MV) 까지 가질 수 있다. 필드 MC 모드의 경우,  $MV^t$  및  $MV^b$ 이 각각 T2T 모드와 B2B 모드를 따르면,  $MV^t$  및  $MV^b$ 은 실제로 프레임 MV와 같다고 할 수 있다. 그러나,  $MV^t$ 이 T2B 모드이거나

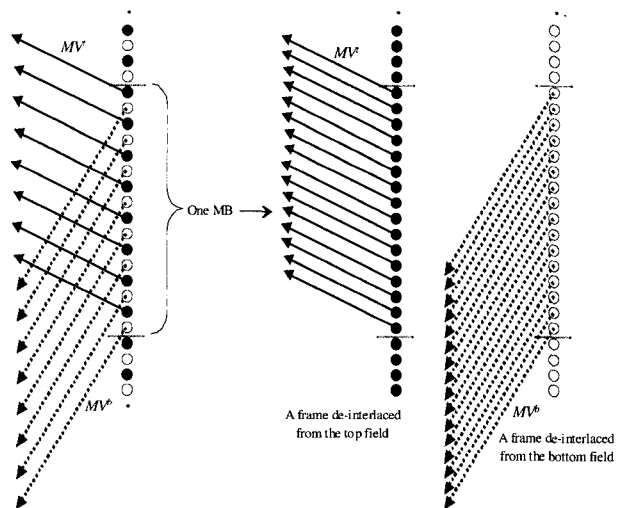


그림 3. 필드 MC 모드인 경우 필드 MV들의 프레임 MV들로의 변환.  
Fig. 3. Conversion of field MVs into frame MVs in case of a field MC mode

나.  $MV^b$ 이 B2T 모드이면,  $MV^t$  및  $MV^b$ 의 수직 성분들은 다음과 같이 재조정되어야 한다:

$MV^t$ 가 T2B 모드를 따르면,

$$MV_y^t = MV_y^t + 1 \tag{1}$$

$MV^b$ 가 B2T 모드를 따르면,

$$MV_y^b = MV_y^b - 1 \tag{2}$$

식 (1)과 식 (2)에서, 1은 한 화소를 의미하며,  $MV_y$ 은 각 MV의 수직 성분을 나타낸다. 한 화소만큼의 offset이 탑 필드와 바텀 필드간의 sampling 위치의 차에 따른 오류를 막아준다. 따라서, 앞서 기술한 일련의 과정을 거쳐 트랜스코딩할 모든 inter MB은 적어도 하나의 후보 프레임 MV를 갖는다. 이렇게 얻어진 후보 MV는 새로운 참조 프레임과 현재 프레임간의 거리를 고려하여 적당하게 스케일되어야 한다.

### 1.2 보조 후보 MV들의 추출

그림 2와 같은 입력 비트열의 첫번째 sub-GOP를 예로 들자 (그림 4 참조). 그림 4에서, 각 MB은 그림 3처럼 수

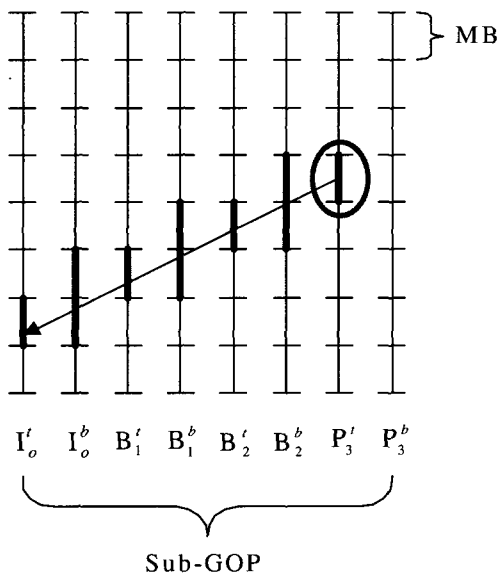


그림 4. 보조 후보 MV들의 생성.  
Fig. 4. The derivation of auxiliary candidate MV

직방향으로만 표기되어 있다.  $P_3^t$ 에 원으로 표시된 MB이 T2T 모드를 가지고 있다고 가정하자. 목표는 원 표시된 MB의 MV가 통과하는  $P_3^t$ 와  $I_0^t$ 사이의 필드들의 MB들 (그림 4에서 굵은 선으로 표시된 MB들)을 찾는 것이다. 이 때 찾는 기준은 다음과 같다: 해당 MB과 MV를 따라 MC된 MB이 겹치는 면적이 MB크기의 50%를 넘으면, 그 MB은 원 표기된 MB의 MV를 후보 벡터로 삼는다. Sub-GOP내에 존재하는 모든 MV들에 이 같은 과정을 반복한다. 결과적으로 sub-GOP 내의 모든 inter MB들은 자기 자신의 보조 후보 MV들을 얻게 된다. 그러나, 이런 과정에도 불구하고, 어떤 MB들은 아무 후보 MV들도 갖지 못할 수 있다. 이런 경우, (0,0)을 후보 MV로 삼거나, 해당 MB을 intra MB으로 처리할 수 있다. 본 논문에서는 전자의 방법을 택한다.

한편, 스캔 포맷 변환이 움직임 추정에 있어 해당 픽처간 거리도 함께 변화시키기 때문에, 후보 MV는 필드 간 변화된 거리에 따라 적당하게 스케일되어야 한다.

## 2. 가중치 중간값에 기반한 MV선택

IPC 후 downscaling에 앞서, 각 inter MB의 후보 MV들 중 최적 MV를 택한다. Down-scaling 후에는 또 다른 방식의 MV 선택 방법이 있으며, 이는 III.2.2절에서 기술된다.

### 2.1 Down-scaling이 없는 경우의 동영상 트랜스코딩

적절한 가중치를 다음과 같이 구한다.

- ▶  $MV^t$ 과  $MV^b$ 이 각각 T2T 모드와 B2B 모드를 따르면, 이들 MV들의 가중치는 모두 2로 선택된다. 그렇지 않으면, 가중치들은 1로 결정된다.
- ▶ 각 inter MB이 자체의 MV를 가지면 해당 MV의 가중치는 2이다.

$MV_i$ 와  $w_i$ 은 각각 i-번째 후보 MV이며, 그 가중치를 의미한다. 가중치 평균 기반 MV 선택기는 다음과 같이 표현된다:

$$MV_f = \frac{\sum_i^K w_i MV_i}{\sum_i^K w_i} \tag{3}$$

$MV_i$ 와 K는 각각 최종적으로 추정된 MV와 후보 MV들

의 수이다. 그러나, 이 방법은 MV 잡음에 강하지 못하다. 다시 말해 하나의 MV라도 잡음 때문에 전혀 다른 방향의 값을 가질 경우 평균값이 크게 달라질 수 있다. 그래서, 본문은 다음과 같은 가중치 중간값 선택기를 채택한다:

$$MV_f = MV_{i_{\min}}$$

$$i_{\min} = \arg \min_i \sum_{j=1}^K w_j \|MV_j - MV_i\|, i = 1, 2, \dots, K. \quad (4)$$

식 (4)에서  $\|\cdot\|$ 는  $L_1$ -놈 거리 측정자를 의미한다.

### 2.2 Downscaling이 있는 경우의 동영상 트랜스코딩

Shen은 가중치 평균에 기반한 MV 선택 방법을 제안한 바 있다<sup>[14]</sup>. Shen은 해당 MB과 겹치는 MB의 겹침 면적을 가중치 삼아 평균을 구하였다. 저해상도 프레임 크기를 갖는 동영상으로의 트랜스코딩을 위해서는, 새로운 MV가 III.2.1절에서 얻어진 또 다른 후보 MV들로부터 계산된다 (그림 5 참조). 가중치는 겹치는 면적에 따라 결정된다. 모든 후보 MV들은 가중치를 갖는 MV들로부터 얻어진다. 따라서, 각 inter MB의 최종 MV는 가중치들과 후보 MV들을 이용하여 식 (4)에 의해 얻어진다.

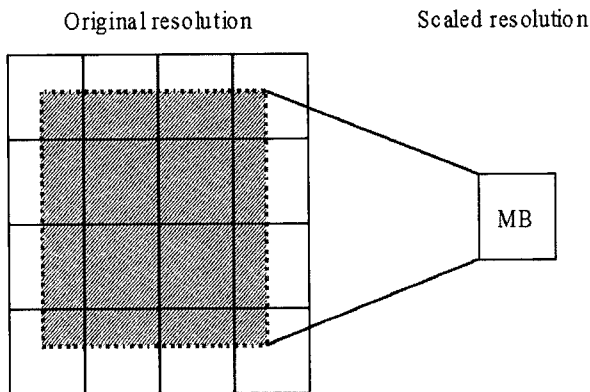


그림 5. 스케일된 MB 및 원 해상도 버전.  
Fig. 5. A scaled MB and its non-scaled version

### 3. MV refinement

보다 나은 성능을 위해, 위에서 얻어진 최종 MV를 중심으로 반화소 단위로 0.5내지는 최대 1.5 정도의 탐색 영역

내에서 refinement를 수행할 수 있다. III.1과 III.2절의 과정에 의한 연산량은 매우 작다. 제안한 고속 움직임 추정 방법의 연산량은 대부분 MV refinement과정 때문이다. 예를 들어, 0.5의 MV refinement 탐색 영역을 가질 경우, 반화소 기준으로 9번의 블록 정합과정이 필요하다. 1.0 이나 1.5의 경우는 각각 25번과 49번의 블록 정합 과정이 요구된다. 한편, FSA는 정수화소 단위 탐색과 반화소 단위 탐색로 구성된다. w의 정수화소 단위 탐색이라면, 반화소 탐색을 포함한 FSA의 MB당 블록 정합 수는  $(2w+1)^2+9$ 이다. 예를 들어 w가 32이고, 제안한 방법의 MV refinement의 탐색 영역이 1.0이라면, 제안한 방법의 연산량은 FSA의 0.6%에 불과하다. 다음 절의 모의 실험 결과는 refinement 과정의 효과를 보인다.

## IV. 모의 실험 결과

### 1. 실험 조건

MPEG-2 TM5 부호기 및 복호기<sup>[3]</sup>에 기반한 트랜스코더에 제안한 기법을 적용하였다. 실험을 위해 두 가지 입력 비트열을 사용한다. Tennis와 Basketball 두 영상의 처음 30 프레임들을 부호화하여 입력 비트열들이 각각 만들어진다. 동영상들은 1920x1080크기를 가지며 비율 주사 방식이다. M이 15, N이 3, 30Hz의 프레임율, 비트율은 20Mbps이다. IB1B2P의 GOP 구조를 가지며, P-프레임의 움직임 탐색 범위는  $\pm 63$ , B<sub>1</sub>-프레임의 순방향은 31 역방향은  $\pm 63$ , B<sub>2</sub>-프레임의 순방향은  $\pm 63$  역방향은  $\pm 31$ 화소이다. 여기서, B<sub>1</sub>, B<sub>2</sub>-프레임은 각각 sub-GOP 내 첫 번째 B, 두 번째 B-프레임을 의미한다.

순사 방식의 720x480 동영상으로의 트랜스코딩을 수행한다. 물론, 출력 동영상의 프레임율은 60Hz이다. 재 부호화를 위한 MV 탐색 범위는 자동적으로 입력 f-코드로부터 결정된다. 물론 영상 크기 변화에 맞게 스케일된 탐색 범위를 갖는다. 본 실험에서 스케일 비율은 수직 방향이 480/1080이며, 수평 방향이 720/1920이다. 포맷 변환과 상관없는 모든 부호화 파라미터들은 입력과 동일하게 설정된다. Inter/intra 모드 결정은 현재 MB 내의 intra 비중이 75% 이상일 경우 intra로 하고, 그 이하일 경우 inter로 하는 방식을 택한다. 기존의 방식들도 유사한 방법을 취하고 있다.

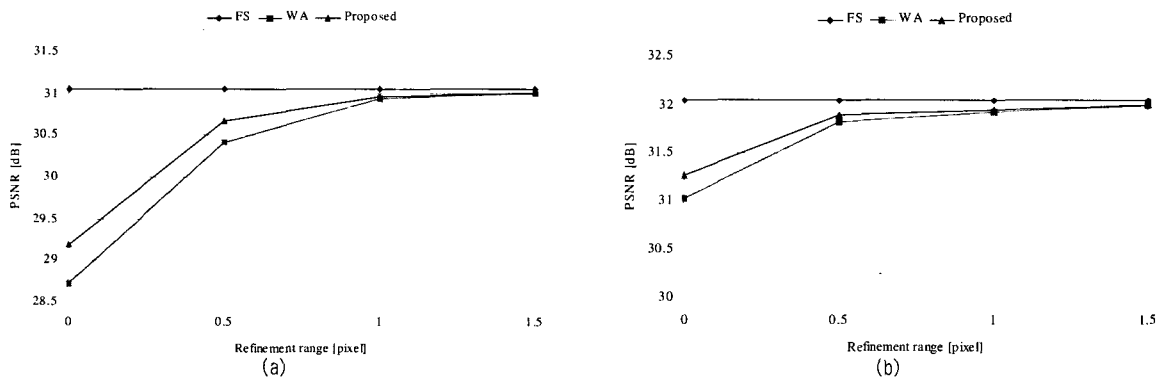


그림 6. 비트율이 7Mbps일 때의 PSNR 성능 (a) Tennis (b) Basketball.  
Fig. 6. PSNR performance for the target bit-rate of 7 Mbps. (a) Tennis (b) Basketball

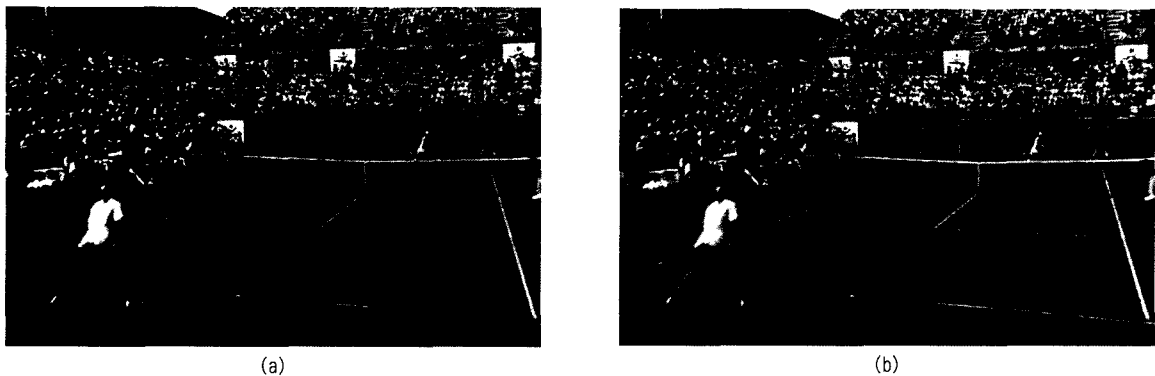


그림 7. Tennis의 9번째 프레임의 복원 영상 비교 (a) 제안한 방법이 적용된 트랜스코딩 후 복원된 영상 (b) FSA가 적용된 트랜스코딩 후 복원된 영상  
Fig. 7 Subjective quality comparison of the 9th reconstructed frame of Tennis; (a) Reconstructed frame using the proposed transcoding (b) Reconstructed frame using the conventional transcoding using FSA.

제안한 ME 기법은 PSNR 측면에서 FSA와 비교된다. 더불어서, III.2.1절에서 설명된 두 가지 MV 선택 방법이 비교된다. 가중치 평균 방식을 WA라고 부른다. WA 방식도 제안한 알고리즘에 적용이 되며, MV 선택 방법이 성능에 어떤 영향을 미치는지를 다음 절에서 보인다.

## 2. 성능 평가

IPC되고, down-scaling된 영상들 (재 부호화기의 입력 영상들)을 원영상이라고 간주하고 PSNR을 계산하였다. 7Mbps, 8Mbps, 및 9Mbps로 트랜스코딩된 결과를 표 2에 보였다. 그림 6은 PSNR 그래프이다. 그림 6(a)에서, 어떤 refinement도 수행하지 않는 경우, 제안한 기법은 WA보다 0.48dB 우수하다. 그러나, 1.0 이상의국부 탐색 영역에 대해

서는 WA와 제안한 기법은 큰 PSNR 차이가 없다. 단지 1.0화소의 refinement만으로 제안한 기법은 FSA에 비해 불과 0.08dB만의 화질 저하를 갖는다. 0.5 정도의 refinement만으로도 제안한 ME 방법은 FSA에 비해 0.38dB정도만 떨어진다. 반면, WA는 0.64dB의 비교적 심각한 화질 저하를 보인다. 결과적으로, 제안한 기법은 0.5나 이보다 큰 refinement 탐색 영역을 가질 경우 FSA에 근접한 성능을 보임을 알 수 있다. 또한, 제안한 기법은 항상 WA보다 나은 성능을 보인다.

그림 6(b)의 Basketball시퀀스에서는, 제안한 기법과 WA와의 PSNR 차이가 상당히 작다. 예를 들어, refinement가 없을 경우, 두 기법간의 PSNR차는 0.23dB에 불과하다. 그러나, 제안한 기법은 여전히 WA보다 우수한 성능을 보인다.

그림 7의 (a)와 (b)는 각각 제안한 방법과 FSA를 적용

표 2. PSNR (dB) 성능.  
Table 2: PSNR (dB) performance

Re-encoding target bitrate	Algorithms	Tennis				Basketball			
		Refinement search range				Refinement search range			
		0	0.5	1.0	1.5	0	0.5	1.0	1.5
7 Mbps	FS	31.06				32.05			
	WA	28.71	30.42	30.94	31.01	31.03	31.82	31.93	32.00
	Proposed	29.19	30.68	30.96	31.02	31.26	31.90	31.95	32.00
8Mbps	FS	31.69				32.81			
	WA	29.35	31.05	31.57	31.65	31.82	32.57	32.70	32.76
	Proposed	29.82	31.32	32.59	31.56	32.03	32.65	31.71	32.76
9Mbps	FS	32.28				33.47			
	WA	29.94	31.63	32.18	32.24	32.50	33.24	33.37	33.43
	Proposed	30.40	31.90	32.19	32.24	32.73	33.33	33.39	33.44

한 트랜스코딩으로 복원된 영상들이다. 모두 8Mbps로 트랜스코딩되었으며, 제안한 방법의 MV refinement 범위는 1.0으로 설정된 경우이다. PSNR 성능과 마찬가지로 주관적 화질 역시 차이가 거의 없음을 알 수 있다.

제안한 기법의 연산량은 상당히 줄어든다. 식 (4)에 따른 MV 선택기의 연산량 부담은 실제로 무시할 만하기 때문에, 재 부호화기의 ME 전체 연산량은 결국 반화소 단위의 MV refinement 연산량이 된다.

따라서, 제안한 ME 기법을 이용한 MPEG-2 동영상 트랜스코더는 매우 적은 연산량을 유지하면서, 전역 탐색 기법을 채용한 MPEG-2 동영상 트랜스코더와 거의 동일한 PSNR 성능을 보인다.

### V. 결론

본 논문은 공간적 해상도 변환은 물론 스캔 포맷 변환도 함께 고려한 MPEG-2 동영상 트랜스코딩을 위한 고속 ME 기법을 제안하였다. 먼저, 재 부호화 포맷에 맞는 후보 MV들을 입력 비트열에서 복호된 MV들을 이용하여 구성한다. 둘째로 가중치 중간값 선택기에 기반한 최적 MV를 선택한다. 모의 실험 결과로부터 제안한 기법은 트랜스코더의 연산량을 크게 줄이면서 매우 우수한 PSNR 성능을 보임을 알 수 있다. 동일한 탐색 전략이 Progressive to Interlaced 스캔 포맷 변환을 갖는 트랜스코딩에도 쉽게 적용 가능하다.

### 참 조 문 헌

- [1] J. Moura, R. Jasinski, H. Shiojiri, and C. Lin, "Scalable video coding over heterogeneous networks," Proc. SPIE, vol. 2602, pp. 294-306, 1996.
- [2] M. Ghanbari, "Two-layer coding of video signals for VBR networks," IEEE J. Select. Areas Commun., vol. 7, pp. 771-781, June 1989.
- [3] MPEG2, "Information technology-generic coding of moving pictures and associated audio," Tech. Rep., ISO/IEC 13818-2, Committee Draft, Mar. 1994.
- [4] W. Li, "Overview of fine granularity scalability in MPEG-4 video standard," IEEE Trans. Circ. Syst. Video Technol., vol. 11, no. 3, pp. 301-317, March 2001.
- [5] H. Sun, W. Kwok, and J. W. Zdepski, "Architectures for MPEG compressed bitstream scaling," IEEE Trans. Circ. Syst. Video Technol., vol. 6, no. 2, pp. 191-199, April 1996.
- [6] B. Shen, I. K. Sethi, and B. Vasudev, "Adaptive motion-vector resampling for compressed video down sampling," IEEE Trans. Circ. Syst. Video Technol., vol. 9, no. 6, pp. 929-936, Sept. 1999.
- [7] T. Shanableh and M. Ghanbari, "Heterogeneous video transcoding to lower spati-temporal resolutions and different encoding formats," IEEE Trans. Multimedia, vol. 2, no. 2, pp. 101-110, June 2000.
- [8] M. Sugano, Y. Nakajima, H. Yanagihara, and A. Yoneyama, "An efficient transcoding from MPEG-2 to MPEG-1," IEEE Conf. ICIP, pp. 417-420, Thessaloniki, Greece, Oct. 2001.
- [9] V. Bhaskaran and K. Konstantinides, Image and Video Compression Standards, Algorithms and Architectures, Boston, MA: Kluwer Academic, 1995.
- [10] N. Bjork and C. Christopoulos, "Transcoder architecture for video coding," IEEE Trans. Consumer Electronics, vol. 44, no. 1, pp. 88-98, Feb. 1998.
- [11] M. R. Hashemi, L. Winger, and S. Panchanathan, "Compressed



- domain motion vector downscaling of MPEG video," IEEE Conf. ICIP, pp. 276-279, Kobe, Japan, Oct. 1999.
- [12] J. K. Han and H. M. Kim, "Modified cubic convolution scaler with minimum loss of information," Optical Engineering, vol. 40, pp. 540 - 546, April 2001.
- [13] R. G. Keys, "Cubic convolution interpolation for digital image processing," IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 29, no. 6, pp. 1153-1160, Dec. 1981.
- [14] G. Shen, B. Zeng, Y.-Q. Zhang, and Ming L. Liou, "Transcoder with arbitrarily resizing capability", IEEE Conf. ISCAS, vol. 5, pp. 25-28, Sydney, Australia, May 2001.
- [15] S. J. Wee, J. G. Apostolopoulos, and N. Feaster, "Field-to-frame transcoding with spatial and temporal downsampling," IEEE Conf. ICIP, Oct. 1999.
- [16] N. Feaster and S. J. Wee, "An MPEG-2 to H.263 transcoder," SPIE Intl. Symp. Voice, Video, and Data Commun., Sept. 1999

---

## 저 자 소 개

---



### 송 병 철

- 1994년 2월 : 한국과학기술원 전기 및 전자공학과 학사 졸업
- 1996년 2월 : 한국과학기술원 전기 및 전자공학과 석사 졸업
- 2001년 2월 : 한국과학기술원 전기 및 전자공학과 박사 졸업
- 2001년 3월 ~ 현재 : 삼성전자 디지털미디어 연구소 책임연구원
- 주관심분야 : 동영상 부호화, 영상 표현 및 탐색, 동영상 CODEC 설계 및 구현



### 천 강 옥

- 1988년 2월 : 부산대학교 전자공학과 학사 졸업
- 1990년 2월 : 한국과학기술원 전기 및 전자공학과 석사 졸업
- 1995년 2월 : 한국과학기술원 전기 및 전자공학과 박사 졸업
- 1995년 3월 ~ 현재 : 삼성전자 디지털미디어 연구소 수석연구원
- 주관심분야 : 영상 신호 처리, 동영상 CODEC 설계 및 구현, 멀티미디어 시스템 설계 등