

스캔 포맷 변환을 지원하는 효율적인 MPEG-2 동영상

트랜스코딩

송병철, kangwook

디지털미디어 연구소, 삼성전자

전화: 031-200-3743

Efficient MPEG-2 Video Transcoding with Scan Format Conversion

Byung Cheol Song, Kang Wook Chun

Digital Media R&D Center, Samsung Electronics Co., Ltd
E-mail: bcsong@samsung.com

ABSTRACT

General-purpose MPEG-2 video transcoders must be able to achieve any conversion between 18 ATSC (Advanced television system committee) video formats for DTV (digital television), e.g., scan format, size format, and frame rate format conversion. Especially, scan format conversion is hard to implement because frame rate and size format conversion often happen together. This paper proposes a fast motion estimation (ME) algorithm for MPEG-2 video transcoding supporting scan format conversion. Firstly, we extract and compose a set of candidate motion vectors (MV's) from the input bit-stream to comply with the re-encoding format. Secondly, the best MV is chosen among several candidate MV's by using a weighted median selector. Simulation results show that the proposed ME algorithm reduces significantly transcoding complexity with a minor PSNR degradation.

I. 서론

최신 동영상 압축 기술과 함께, 멀티미디어 네트워크 기술의 급속한 발전은 원거리 화상회의, VOD (Video-On-Demand), 원거리 교육 등의 동영상 서비스를 일상화하는데 기여했다. 때때로 상기 동영상 응용 분야들은 이질적인 네트워크 환경 하에서 이미 특정 비트율로 압축되어 있는 비트스트림을 채널이 허락하는 가능한 비트율로의 변환을 필요로 한다 [1]. 이를 위해서는 현존하는 동영상 압축 표준들에는 효과적인 비트율 변환을 위한 스케일러블 동영상 부호화 기법 (Scalable video coding)들이 포함되어 있다 [2]. 그러나, 스케일러블 부호화는 상위 계층 (enhancement layer)의 수의 제한 때문에 최대 세 스케일러빌리티 (scalability) 계층만을 가질 수 있다 [3]. 많은 네트워크 동영상 응용 분야에서 보다 세밀한 스케일러빌리티가 요구되기 때문에, MPEG-4에서는 fine-granular scalable coding (FGS) 기술을 채택하였다 [4]. 그러나, 동영상 부호기는 FGS 비트율을 복호하기 위해서는 상위 계층의 복호를 위한 추가적인 복호 모듈을 갖추어야 하는 단점이 있다.

위 문제를 해결하기 위해, 여러 동영상 트랜스코딩 기법들이 제안되어 왔다. 동영상 트랜스코딩은 이미 압축되어 있는 동영상 비트율을 다른 비트율을 갖거나 [5], 다른 프레임 크기를 갖거나 [6], 다른 표준의 비트율로 변환시켜주는 [7, 8] 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 파라미터 (QP)로 재 양자화 (re-quantization)하는 방식이다 [5]. 열린 루프 트랜스코딩은 완전한 복호와 재 부호화 과정 없이 압축 영역에서 이루어지기 때문에, 간단하고 매우 빠르다. 그러나, 열린 루프 트랜스코딩은 원래 복원 영상들과 트랜스코더에서 사용되는 복원 영상들이 다르기 때문에 발생하는 drift 문제를 갖는다. 이로 인해 상당한 화질 저하가 생긴다. 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 [5]. 그래서 본 논문에서는 공간 영역 닫힌 루프 트랜스코더만을 고려한다. 공간 영역 트랜스코더는 입력 비트열이 완전히 복호한 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 [6].

일반적으로, FSA는 계산량이 많고 부호기 전체 연산량의 60% 이상을 차지할 만큼 부호기에서 차지하는 비중이 매우 크다 [9]. 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 [10]. 이런 문제를 해결하기 위해 Shen은

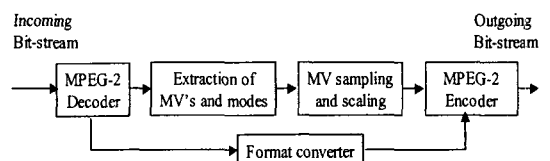


그림 1. 열린 루프 MPEG-2 동영상 트랜스코더.

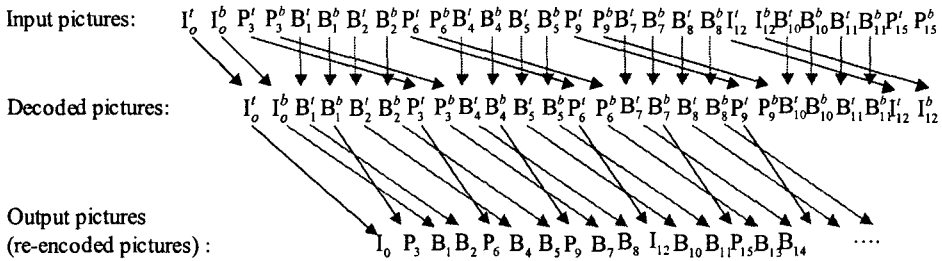


그림 2. 부호화 순서에 따른 입출력의 픽처 타입 (도시 순서). 위치자 't' 혹은 'b'는 탑 필드 및 바텀 필드를 의미한다.

적용적인 MV 재사용 방법을 제안하였다 [6]. Sugano는 각 매크로블록 (macroblock; MB)의 '공간적 활성화도'로서 그 QP값을 채택하여 새로운 움직임 벡터장을 구하는데 이용하였다 [8]. Hashemi는 최대 평균 상관도 (maximum average correlation; MAC) 방법을 제안하였다 [11]. MAC에서는 입력 비트열로부터 네 개의 주변 움직임 벡터들 중에서 최대 MAC를 갖는 하나를 택한다.

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

표 1. ATSC DTV 포맷. 't'는 비월 주사를 'p'는 순사 주사 방식을 의미한다.

Vertical Lines	Pixels Per Line	Aspect Ratio	Frame Rate and Scan Format		
			60I	30P	24P
1080	1920	16:9		30P	24P
720	1280	16:9		30P	24P
480	704	16:9/4:3	60I	60P	30P
480	640	4:3	60I	60P	30P

II. 스캔 포맷 변환이 있는 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 어려워진다. 예를 들어, 그림 2에서처럼 I-프레임이 B나 P 프레임으로 변환되는 경우가 발생한다. 또한, 입력과 출력의 움직임 보상 (motion compensation; MC) 타입도 일치하지 않게 된다. 입력에서는 프레임 MC와 필드 MC가 공존하는 반면, 재 부호화된 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 MPEG-2 복호기, 포맷 변환기, 그리고, MPEG-2 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 MPEG-2 TM5에 있는 간단한 IPC기법을 사용한다 [3]. 공간영역에서의 down-sampling을 위해서는 잘 알려진 cubic convolution scaler를 사용한다 [12].

제안한 기법의 설명을 위해 입력 동영상의 group of pictures (GOP) 크기 (M)는 12이며, sub-GOP (M; P프레임 간 거리)는 3, 그리고 출력 동영상의 스캔 포맷은 순사 방식이라고 가정한다. 또한, 출력의 M도 3이라고 하자 (그림 2 참조).

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 벡터들은 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.

A. 필드 MV들의 프레임 MV들로의 변환

입력의 각 MB의 MC 타입은 필드 MC이거나 프레임 MC가 된다. 그러나, 순사 방식 동영상에서는 프레임 MC만이 존재한다. 그래서, 필드 MV들이 프레임 MV로 변환되어야 한다. 입력 P-픽처에서, 필드 MC모드를 갖는 inter MB는 두 개의 MV를 갖는다. 바로 탑 필드 MV (MV^t)와 바텀 필드 MV (MV^b)이다. 또한, 각 필드 MB는 참조 필드가 탑 필드인지 바텀 필드인지를 나타내는 필드 선택 모드를 갖는다. 즉, MV^t 는 탑 필드의 참조 필드에서 찾아지거나 (T2T mode), 바텀 필드에서 찾아질 수도 있다 (T2B mode). 유사하게, MV^b 는 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 MB는 각각 T2T 모드와 B2B 모드를 갖는다고 가정한다. 마지막으로, 탑 필드 MB이 움직임 벡터 MV^t 를 가지면, 그 MV도 후보 벡터로

삼는다 (바텀 필드 MB의 경우, MV^b).

필드 MC 모드의 경우, MV^a 및 MV^b 이 각각 T2T 모드와 B2B 모드를 따르면, MV^a 및 MV^b 은 실제로 프레임 MV와 같다고 할 수 있다. 그러나, MV^a 이 T2B 모드이거나, MV^b 이 B2T 모드이면, MV^a 및 MV^b 의 수직 성분들은 다음과 같이 재조정되어야 한다: MV^a 가 T2B 모드를 따르면,

$$MV_y^a = MV_y^a + 1 \quad (1)$$

MV^b 가 B2T 모드를 따르면,

$$MV_y^b = MV_y^b - 1 \quad (2)$$

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.

B. 보조 후보 MV들의 추출

그림 2와 같은 입력 비트열의 첫번째 sub-GOP를 예로 들자 (그림 3 참조). 그림 3에서, 각 MB은 수직방향으로만 표기되어 있다. P_3' 에 원으로 표시된 MB이 T2T 모드를 가지고 있다고 가정하자. 목표는 원 표시된 MB의 MV가 통과하는 P_3' 와 I_0' 사이의 필드들의 MB들 (그림 3에서 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.

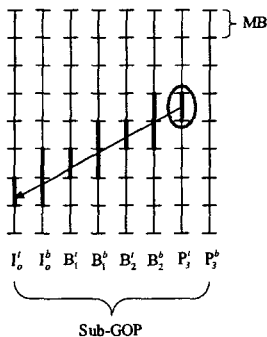


그림 3. 보조 후보 MV들의 생성.

2. 가중치 중간값에 기반한 MV 선택

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.B절에서 기술된다.

A. Down-scaling이 없는 경우의 동영상 트랜스코딩

적절한 가중치를 다음과 같이 구한다.

● MV^a 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:

$$MV_f = \frac{\sum_{i=1}^K w_i MV_i}{\sum_{i=1}^K w_i} \quad (3)$$

MV_f 과 K 는 각각 최종적으로 추정된 MV와 후보 MV들의 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, 본 논문은 다음과 같은 가중치 중간값 선택기를 채택한다:

$$MV_f = MV_{i_{min}} \quad (4)$$

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

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

B. Downscaling이 있는 경우의 동영상 트랜스코딩

Shen은 가중치 평균에 기반한 MV 선택 방법을 제안한 바 있다 [13]. Shen은 해당 MB과 겹치는 MB의 겹침 면적을 가중치 삼아 평균을 구하였다. 저해상도 프레임 크기를 갖는 동영상으로의 트랜스코딩을 위해서는, 새로운 MV가 III.2.A절에서 얻어진 또 다른 후보 MV들로부터 계산된다 (그림 4 참조). 가중치가 겹치는 면적에 따라 결정됨에 주목하라. 모든 후보 MV들은 가중치를 갖는 MV들로부터 얻어진다. 따라서, 각 inter MB의 최종 MV는 가중치들과 후보 MV들을 이용하여 식 (4)에 의해 얻어진다.

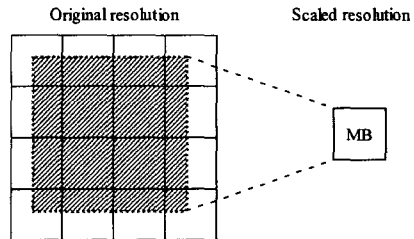


그림 4. 스케일된 MB 및 원 해상도 버전.

3. MV refinement

보다 나은 성능을 위해, 위에서 얻어진 최종 MV를 중심으로 반화소 단위로 ± 0.5 내지는 최대 ± 1.5 정도의 탐색 영역 내에서 refinement를 수행할 수 있다. 다음 절의 모의 실험 결과는 refinement 과정의 효과를 보인다.

표 2. PSNR (dB) 성능.

Mb ps		Tennis				Basketball			
		0	± 0.5	± 1.0	± 1.5	0	± 0.5	± 1.0	± 1.5
7	FS	31.1				32.1			
	WA	28.7	30.4	30.9	31.0	31.0	31.8	31.9	32.0
	WM	29.2	30.7	31.0	31.0	31.3	31.9	32.0	32.0
8	FS	31.7				32.8			
	WA	29.3	31.1	31.6	31.7	31.8	32.6	32.7	32.8
	WM	29.8	31.3	31.6	31.7	32.0	32.7	31.7	32.8
9	FS	32.3				33.5			
	WA	29.9	31.6	32.2	32.2	32.5	33.2	33.4	33.4
	WM	30.4	31.9	32.2	32.2	32.7	33.3	33.4	33.4

IV. 모의 실험 결과

1. 실험 조건

MPEG-2 TM5 부호기 및 복호기 [3]에 기반한 트랜스코더에 제안한 기법을 적용하였다. 실험을 위해 두 가지 입력 비트열을 사용한다. *Tennis*와 *Basketball* 두 영상의 처음 30프레임들을 부호화하여 입력 비트열들이 각각 만들어진다. 동영상들은 1920x1080크기를 가지며 비쵸 주사 방식이다. M 이 15, N 이 3, 30Hz의 프레임율, 비트율은 20Mbps이다. I_B, B_1, P 의 GOP 구조를 가지며, P -프레임의 움직임 탐색 범위는 ± 63 , B_1 -프레임의 순방향은 ± 31 역방향은 ± 63 , B_2 -프레임의 순방향은 ± 63 역방향은 ± 31 화소이다. 여기서, B_1 , B_2 -프레임은 각각 sub-GOP 내 첫 번째 B , 두 번째 B -프레임을 의미한다.

수사 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 480/1080 이며, 수평 방향이 720/1920 이다. 포맷 변환과 상관없는 모든 부호화 파라미터들은 입력과 동일하게 설정된다.

제안한 ME 기법은 PSNR 측면에서 FSA와 비교된다. 더불어서, III.2.A절에서 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.

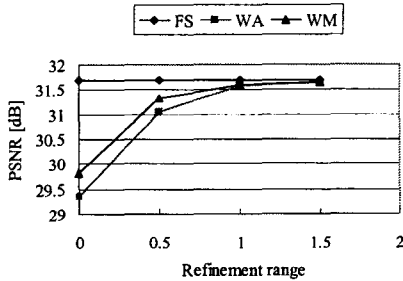


그림 5. *Tennis* 영상에서 비트율이 7Mbps일 때의 PSNR 성능.

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 5은 PSNR 그래프이다. 그림 5(a)에서, 어떤 refinement도 수행하지 않는 경우, 제안한 기법은 WA보다 0.48dB 우수하다. 그러나, ± 1.0 이상의 국부 탐색 영역에 대해서는 WA와 제안한 기법은 큰 PSNR 차이가 없다. 단지 ± 1.0 화소의 refinement만으로 제안한 기법은 FSA에 비해 불과 0.08dB만의 화질 저하를 갖음에 주목하라. ± 0.5 정도의 refinement만으로도 제안한 ME 방법은 FSA에 비해 0.38dB정도만 떨어진다. 반면, WA는 0.64dB의 비교적 심각한 화질 저하를 보인다. 결과적으로, 제안한 기법은 ± 0.5 보다 큰 refinement 탐색 영역을 가질 경우 FSA에 근접한 성능을 보임을 알 수 있다. 또한, 제안한 기법은 항상 WA보다 나은 성능을 보인다.

그림 5(b)의 *Basketball* 시퀀스에서는, 제안한 기법과 WA와의 PSNR 차이가 상당히 작다. 예를 들어, refinement가 없을 경우, 두 기법간의 PSNR차는 0.23dB에 불과하다. 그러나, 제안한 기법은 여전히 항상 WA보다 우수함에 주목하라.

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.

참조 문헌

- [1] J. Moura, R. Jasinschi, 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] 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.
- [14] S. J. Wee, J. G. Apostolopoulos, and N. Feaster, "Field-to-frame transcoding with spatial and temporal downsampling," *IEEE Conf. ICIP*, Oct. 1999.
- [15] N. Feaster and S. J. Wee, "An MPEG-2 to H.263 transcoder," *SPIE Intl. Symp. Voice, Video, and Data Commun.*, Sept. 1999.