

논문 2006-43SP-3-6

H.264/AVC 콘텐츠의 공간해상도 트랜스코딩을 위한 고속 모드 결정 방법

(Fast Mode Decision for Spatial Transcoding of H.264/AVC Contents)

권 상 구*, 정 봉 수**, 전 병 우*

(Sang Gu Kwon, Bongsoo Jung, and Byeungwoo Jeon)

요 약

무선 네트워크 기술이 발전함에 따라, 모바일 환경을 통한 멀티미디어 콘텐츠 공급이 증가하고 있는 추세이지만, 공급자가 시시각각 변화하는 네트워크 상황과 다양한 사용자의 단말기 종류를 모두 고려하여 적당한 콘텐츠를 공급하는 것은 매우 어렵다. 이러한 문제를 해결하기 위해 트랜스코딩 기술이 제안되었지만 계산 복잡도가 크다. 본 논문에서는 다양한 트랜스코딩 기술 중 하나인 공간해상도 트랜스코딩 기술에서 2:1 화면 축소를 대상으로 H.264에서 매크로블록의 모드 결정 과정 중에 필연적으로 발생하는 계산 복잡도를 줄이기 위해 입력 비트스트림에서 얻은 매크로블록의 모드, 움직임 벡터, 그리고 차-영상 에너지 정보를 이용한 고속 모드 결정 방법을 제안한다. 제안한 고속 모드 결정 방법은 입력 정보들을 비교하여 재-부호화할 매크로블록의 가능한 모드 종류를 제한하여, 모드 결정 과정에서 소모되는 시간을 단축시키는 방법이다. 실험 결과는 영상에 따라 약간의 차이는 있으나 제안한 방법이 일반 부호화 과정에 비해 평균적으로 PSNR은 약 0.04dB 손실이 발생하고 비트율은 약 1.6% 증가하나, 부호화 속도에서는 약 2.66배 빨라짐을 보여준다.

Abstract

As wireless network technology has advanced, demands for multimedia contents through mobile environment have tendered to upward. Since network situation is changing every moment and types of user terminals are diverse, it is difficult for a content provider to consider network situation and type of user terminal to provide multimedia contents. As one solution, transcoding techniques have been proposed, but those have much complexity. In this paper, in order to reduce computational complexity, we propose a fast mode decision using input modes, motion vectors, and residual energies which are obtained from input bitstream for 2:1 down-scaling spatial transcoding application. The proposed method reduces processing time in mode decision by restricting possible mode types based on input information. Experimental results show that the proposed method achieves about 2.66 times improvement in encoding time compared to the normal encoding process while the PSNR is degraded by about 0.04dB, and bit-rate is increased by 1.6%.

Keywords : Mode Decision, 트랜스코딩, 공간해상도, H.264/AVC

I. 서 론

디지털 통신의 보편화에 따라 디지털 멀티미디어의 사용이 급증하고 있을 뿐 만 아니라, 무선 네트워크상에서도 양질의 멀티미디어 콘텐츠의 요구가 급증하고

있다. 하지만 네트워크 상황은 항상 일정한 것이 아니라 시시각각 변하게 되며, 사용자에 따른 단말기의 종류도 다양하다. 따라서 콘텐츠 공급자가 사용자의 단말기 및 네트워크 상황을 모두 고려하여 적당한 콘텐츠를 공급하는 것은 매우 어려운 일이며, 설사 공급이 가능하더라도 많은 비용이 요구된다. 이러한 문제를 해결하기 위해 서로 다른 네트워크 상황과 사용자 단말기 상황에 적합하게 영상 콘텐츠를 변환하는 트랜스코딩 기술이 필요하게 되었다^[1].

* 정회원, ** 학생회원, 성균관대학교 정보통신공학부 (Sungkyunkwan University, School of Information and Communication Engineering)
접수일자: 2005년12월9일, 수정완료일: 2006년4월17일

트랜스코딩 기술은 크게 비트율 변환, 공간해상도 변환, 프레임-율 변환, 그리고 압축 표준 변환 기술로 분류할 수 있다^{[1]~[3]}. 이 중 공간해상도 트랜스코딩 기법은 고화질의 영상을 화면의 크기가 제한된 단말에 제공하기 위해 영상의 크기를 줄이는 동시에 전송 시에 요구되는 비트율 또한 감소시켜 전송하는 기술이다. 가장 기본적인 트랜스코딩 시스템은 입력된 비트스트림을 읽어 복호하는 복호화기 부분과 복호된 영상을 원하는 형태로 재-부호화하는 부호화기 부분을 직렬로 (cascade) 연결하는 구조이다^[1]. 일반적인 복호화기와 부호화기를 단순히 연결하여 트랜스코딩할 경우, 시스템의 비트율 대비 화질 성능은 최적이지만, 높은 복잡도와 계산량으로 인해 실시간 구현이 어렵다. 이로 인해, 복호화와 재-부호화 과정에서 서로 중복되는 계산량을 감소시킬 필요성이 요구되었다.

MPEG-1/2/4 또는 H.263과 같은 기존의 압축 표준에서, 트랜스코더의 복잡도를 줄이기 위한 기술로 크게, 복호화기에서 복호된 움직임 벡터를 단순히 재사용하는 방법과 이때 기존 재사용으로 발생하는 화질의 열화를 줄이기 위해 움직임 벡터를 보정하는 기술들이 주로 연구 되어왔으나^{[4]~[7]}, 이 과정에 필요한, 매크로블록의 모드 결정 방법에 관한 연구는 미흡하였다. 하지만, H.264/AVC 부호화 표준에서는 부호화하는 가장 작은 블록 단위가 4x4이고, 인터 슬라이스에서 한 매크로블록이 가질 수 있는 인터 모드는 P16x16, P16x8, P8x16, P8x8이며, 부매크로블록 모드로 불리는 P8x8모드는 다시 P8x8, P8x4, P4x8, P4x4의 블록 등으로 세분화 될 수 있어 모드의 수로는 7가지이다. 그리고 인트라 모드는 Intra16x16, Intra4x4의 2가지를 가질 수 있다. 따라서 총 9가지나 되는 모드의 수는 재-부호화 시에 최적의 매크로블록 모드를 결정하기 위해 많은 계산량을 요구한다. 일반적인 부호화 과정의 모드 결정을 위해서는, 먼저 각각의 모드에 맞게 최적화된 움직임 벡터 값을 선택한 후, 각 모드에서 소비되는 비트의 발생량과 왜곡을 고려하여 최종적으로 비트-왜곡 측면에서 최적화된 모드를 결정한다. 이에 따라 H.264/AVC 표준 공간해상도 트랜스코더에서 매크로블록 모드 결정을 위한 계산량 감소는 실시간 동작을 위해 매우 중요한 요소이다.

이러한 복잡도 문제를 해결하고자 몇 가지 연구들이 이루어져왔다^{[8]~[10]}. Tan은 가능한 인터 모드를 강제적으로 부매크로블록 모드 없이 {P16x16, P16x8, P8x16, P8x8}의 4개로 줄여 매크로블록의 모드를 결정하는 방법을 제안하였다^[8]. 하지만 이 방법은 확일적으로 모드

의 수를 4개로 고정시키고 8x8블록 보다 작은 블록에 대한 움직임 보상이 없다는 단점이 있다. Zhang은 복호화기에서 얻은 모드 정보만을 이용하는 방법과 추가적으로 움직임 벡터를 이용하여 첫 번째 방법을 개선한 모드 결정 방법을 제안하였다^[9]. Zhang의 첫 번째 방법은 입력된 4개의 매크로블록 모드 정보를 종류별로 분류한 후 조건에 따라 재-부호화할 매크로블록의 모드를 결정한다. 그리고 Zhang의 두 번째 방법은 첫 번째 방법을 보완하기 위한 것으로, 첫 번째 방법에서 재-부호화할 매크로블록의 모드로 P8x8 모드가 선택되었을 경우에 한하여, 입력된 움직임 벡터의 분포를 고려하여 모드를 P16x16, P16x8, P8x16, P8x8, 또는 하위 모드로 할지 결정 한다. 하지만, Zhang의 두 방법 모두 재-부호화할 매크로블록으로 선택된 모드가 최적의 모드가 아닐 확률이 크다는 단점이 있다. Li는 단지 입력된 움직임 벡터 정보만을 고려하여 가능한 매크로블록의 모드를 결정하는 방법을 제안하였다^[10]. 이 방법은 입력된 4개 매크로블록의 움직임 벡터들의 분포 범위에 따라 P16x16, P16x8, P8x16 등으로 재-부호화할 매크로블록 모드를 결정 한다. 예를 들면, 입력된 4개 매크로블록 움직임 벡터들 간의 차이가 최대 첫 번째 경계치 값 이내에 있을 경우에는 P16x16을 재-부호화 매크로블록의 모드로 선택하고, 첫 번째 경계치 값보다는 크고 두 번째 경계치 값 이내에 있을 경우에는 매크로블록의 모드를 P16x8, P8x16으로 제한한다. 마지막으로 두 번째 경계치 보다 클 경우에는 P8x8이나 하위 모드로 제한하는 방식이다. 하지만 입력된 움직임 벡터들이 많이 떨어져 있을지라도 입력된 모드의 영향을 간과하는 단점이 있다.

위에서 살펴본 것과 같이, 기존의 방법들에서는 재-부호화할 매크로블록 모드를 결정하기 위하여 주로 입력된 모드 정보나 움직임 벡터를 사용하였다. 하지만 입력된 차-영상의 에너지 또한 매크로블록 모드 결정시 아주 중요한 요소이다. 따라서 입력된 매크로블록의 모드 정보와 움직임 벡터만을 고려한 것보다 차-영상의 에너지를 추가하여 구한 재-부호화 매크로블록의 모드가 최적에 더 근접할 것이다.

따라서 본 논문에서는 H.264/AVC 비트스트림의 2:1 공간해상도 변환 트랜스코딩 기술에서 계산량을 감소시키기 위해 입력된 4개의 매크로블록의 모드, 움직임 벡터, 그리고 차-영상 에너지 정보들을 통해 도출된 최적 매크로블록 모드의 확률적 분포를 기반으로, 매크로블록의 모드 결정시 최적 모드 하나를 선택하거나 발생

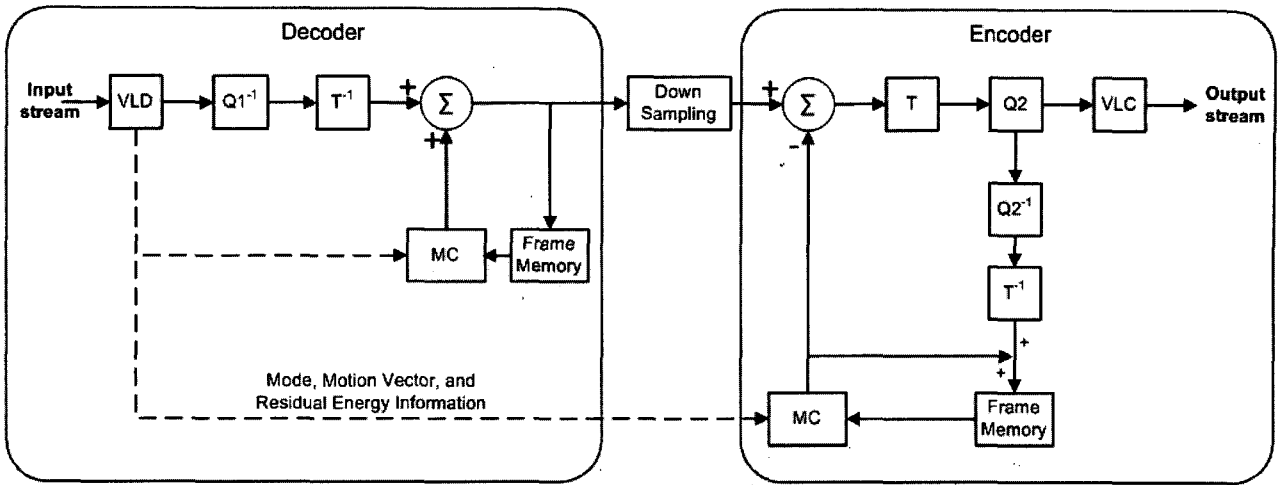


그림 1. 공간해상도 트랜스코딩 시스템
Fig. 1. Spatial resolution transcoding system.

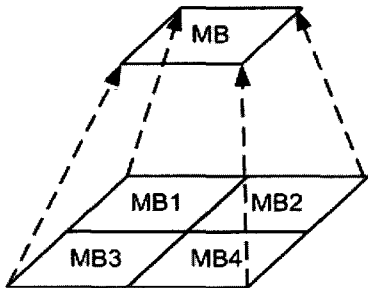


그림 2. 2:1 공간해상도 축소
Fig. 2. 2:1 Spatial down-scaling.

가능한 모드 형태를 제한하여 모드 결정에 필요한 계산량을 줄이는 고속 매크로블록 모드 결정법을 제안한다. 일반적으로 차-영상의 에너지는 화소 단위 영역이나 DCT 영역에서 power spectral density로 구하게 되지만, 이런 경우 많은 계산량이 필요하게 되고 또한 차-영상 에너지가 광범위하게 분포하여 분석하는데 어려움이 따르게 된다. 따라서 간단한 계산을 통하여 적절한 에너지범위를 알기 위해 매크로블록의 차-영상에서 Non-zero DCT 계수 값들의 수로 차-영상의 에너지를 나타내고자 한다. Non-zero DCT 계수가 존재한다는 것은 그 영역에서 에너지가 발생한다는 것을 의미하기 때문에 Non-zero DCT 계수로 에너지를 표현하였다.

본 논문은, 제안된 고속 모드 결정법에 대해 자세히 설명하기 전에 II장에서 공간해상도 트랜스코딩 기술과 H.264/AVC 표준에서의 부호화 모드 기술에 대해 간략하게 설명한 후, III장에서 입력된 정보에 대한 분석 및 출력 매크로블록의 최적 모드에 대한 확률적 상관성을 알아본다. IV장에서는 III장에서 도출해낸 확률적 상관성을 토대로 제안한 고속 모드 결정 방법을 살펴본 후,

V장에서는 제안된 방법에 대한 모의실험 결과를 설명하고, 마지막 장에서는 제안 방법에 대한 결론으로 마친다.

II. 공간해상도 트랜스코딩과 H.264/AVC

1. 공간해상도 트랜스코딩

공간해상도 트랜스코딩 기법은 영상의 크기가 제한된 단말환경에 적합하게 영상의 크기를 줄이는 동시에 전송 시 요구되는 비트율 또한 감소시켜 전송하는 기술이다.

그림 1은 공간해상도 트랜스코딩 시스템의 일반적인 구조를 도식화한 것이다. 그림 1에서 보느냐와 같이, 먼저 복호화기 부분에서 입력 비트스트림을 읽은 후, 복호화 과정을 통해 영상을 복원하여 메모리에 저장한다. 메모리에 저장된 복호 영상을 사용자의 단말기에 적합한 영상의 크기로 축소시켜 부호화기의 입력 영상으로 사용한다. 그리고 한 영상의 매크로블록의 정보, 즉 모드, 움직임 벡터, 차-영상 에너지 등은 부호화 과정에서 사용할 수 있도록 임시 저장한다. 부호화기에서는 크기가 축소된 복호 영상을 입력 영상으로 받아들여 재-부호화 과정에 들어간다. 재-부호화 과정의 움직임 보상 단계에서는 복호화기에서 얻은 매크로블록의 정보들을 이용하여 움직임 보상 과정의 복잡도를 감소시킨다.

공간 해상도 축소 비율에는 여러 가지가 있지만, 본 논문에서는 2:1 공간해상도 축소를 기반으로 하는 트랜스코딩 기술에 국한한다. 2:1 공간해상도 축소는 그림 2에서 보느냐와 같이 입력된 4개의 매크로블록이 재-부호화할 1개의 매크로블록과 대응된다. 영상의 크기만

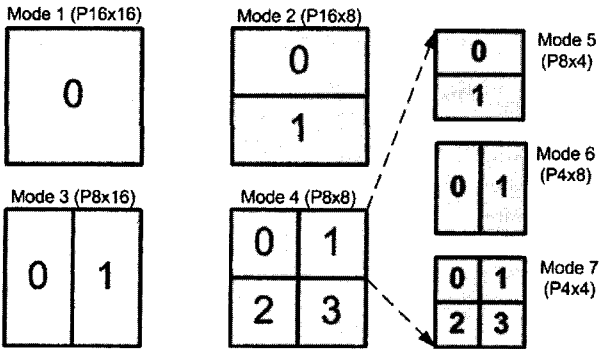


그림 3. 인터 매크로블록 모드: P16x16, P16x8, P8x16, P8x8, P8x4, P4x8, P4x4
 Fig. 3. Inter macroblock modes: P16x16, P16x8, P8x16, P8x8, P8x4, P4x8, P4x4.

다를 뿐, 입력된 영상과 재-부호화된 영상은 아주 유사한 특징을 가지고 있다. 예를 들어, 재-부호화할 영상의 에너지는 입력된 영상 에너지와 유사한 정보를 가지게 된다. 따라서 입력된 영상의 정보를 재-부호화 시에 이용한다면, 불필요한 계산량을 줄일 수 있다.

2. H.264/AVC 표준 기술에서의 부호화 모드

H.264/AVC는 기존의 H.263, MPEG-4처럼 영상을 작은 블록 단위로 처리하는 블록기반 부호화 방식이다. 하나의 매크로블록은 여러 개의 블록들로 구성되어 있으며, 최소 블록 단위는 4x4이다. 그리고 각각의 매크로블록은 인트라 또는 인터 모드로 처리될 수 있다. 인트라 모드는 동일 영상 내의 주변 화소간의 상관도를 적용한 공간 예측 기법을 이용한다. 인트라 모드는 Intra4x4와 Intra16x16이 있으며, Intra4x4에는 총 9개의 예측 방향, Intra16x16에는 총 4개의 예측 방향이 있다. 인터 모드는 시간축 상의 영상 상관도를 이용하여 블록 기반 움직임 보상 예측 기술을 이용하여 부호화 한다^[11]. 인터 모드는 그림 3과 같이 기본적으로 4가지의 분할 모드로 {P16x16, P16x8, P8x16, P8x8} 부호화할 수 있다^{[11],[12]}. P16x16은 SKIP을 포함하고 있으며, P8x8 모드가 선택 되면, 매크로블록 안의 4개의 8x8블록들은 다시 4개의 하위 블록 모드 {P8x8, P8x4, P4x8, P4x4}로 부호화할 수 있어, 하나의 매크로블록이 가질 수 있는 인터 모드의 총 수는 7가지이다^{[11],[12]}.

H.264/AVC는 기존 방법과는 다르게 여러 장의 이전 영상을 참조 영상으로 사용할 수 있다. 또한, 한 매크로블록 내에 있는 블록들이 각각 서로 다른 장에 있는 영상을 참조하거나, 이웃한 매크로블록에서 참조한 영상과는 다른 영상을 참조할 수 있다^[11]. 여러 장의 참조

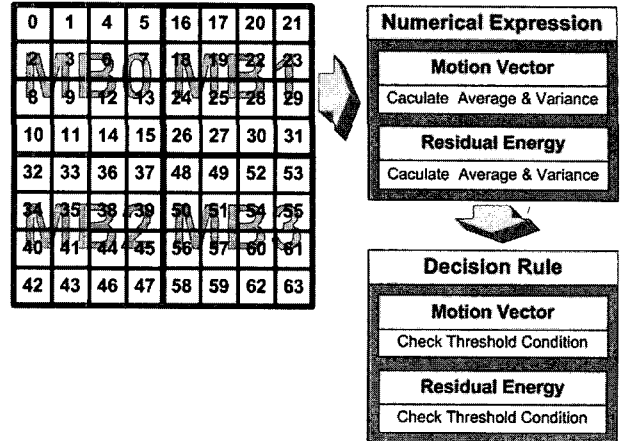


그림 4. 입력 움직임 벡터와 차-영상 에너지 정보 분석
 Fig. 4. Analysis of input motion vectors and input residual energy information.

영상과 9가지의 모드 수는 한 매크로블록에 최적화된 모드 하나를 찾기 위해 복잡한 계산이 필요함을 의미한다. 따라서 트랜스코딩 시스템을 구성하기 위하여 입력 정보를 이용하여 어떻게 효율적으로 모드를 결정하여 계산 복잡도를 감소시킬 것인가는 매우 중요한 문제이다.

III. 입력 정보의 확률 분포 특성 분석

본 논문이 제안하는 바와 같이, 재-부호화 과정에서 매크로블록의 모드를 결정하기 위해, 입력된 모드뿐만 아니라, 움직임 벡터와 차-영상 에너지 분포를 이용하기 위한 방법을 찾아내기 위해서는 식 (1) 및 식 (2)와 같이 움직임 벡터와 차-영상 에너지를 수식적으로 표현할 필요가 있다.

$$MV_{Ave.} = \frac{\sum_{i=0}^{63} mv_i}{16 \times Num_{InterMB}} \quad (1)$$

$$MV_{Var.} = \sum_{i=0}^{63} \frac{\| mv_i - MV_{Ave.} \|^2}{mode \neq Intra} \quad (2)$$

식 (1), (2)에서 입력 움직임 벡터 mv_i 는 기본 블록 단위를 4x4블록으로 한 것으로, 하나의 인터 매크로블록에는 16개의 움직임 벡터로 구성 된다. 여기에서 i 는 그림 4에서 0~63의 4x4블록을 의미하고, $Num_{InterMB}$ 는 입력된 4개의 매크로블록 중에서 인터 모드로 부호화된 매크로블록의 수를 의미한다. 만약 입력된 4개의

매크로블록이 모두 인터 모드이면 64개의 움직임 벡터를 가지게 된다. 먼저 입력된 4개 매크로블록에 존재하는 4x4블록 단위의 움직임 벡터의 평균값에 해당하는 $MV_{Ave.}$ 을 식 (1)과 같이 구한다. 그 후, 입력된 움직임 벡터들 간에 분포 집중도를 표현하기 위해 $MV_{Var.}$ 을 식 (2)와 같이 분산으로 표현한다. $MV_{Var.}$ 의 값이 클수록 4개 매크로블록 내에 존재하는 움직임 벡터들 간의 차이가 커져 이웃한 움직임 벡터들 간의 유사성이 미미함을 의미한다.

$$Residual_{Ave.} = \frac{1}{4} \sum_{j=0}^3 Residual_{MB_j} \quad (3)$$

$$Residual_{Var.} = \sum_{j=0}^3 |Residual_{MB_j} - Residual_{Ave.}|^2 \quad (4)$$

식 (3)의 MB_j 는 그림 4와 같이 입력 매크로블록 0~3을 의미하고, 식 (3), (4)에서 차-영상 에너지 $Residual_{MB_j}$ 는 j 번째 매크로블록 내에 존재하는 16개의 4x4블록에 해당하는 Non-zero DCT 계수의 총 개수를 나타낸다. 입력된 4개 매크로블록의 차-영상 에너지의 평균값에 해당하는 $Residual_{Ave.}$ 을 식 (3)과 같이 구하고, 차-영상 에너지들 간에 분포 집중도를 표현하기 위해 $Residual_{Var.}$ 을 식 (4)와 같이 분산으로 표현한다. $Residual_{Var.}$ 의 값이 작으면 4개 매크로블록 간의 영상 특성이 유사함을 의미한다. 이는 트랜스코딩 모드 결정시 간단히 영상의 특성을 확인하는데 이용할 수 있다.

$$MV_{Var.} \leq MV_{Threshold} \quad (5)$$

$$Residual_{Var.} \leq Residual_{Threshold} \quad (6)$$

4개 매크로블록들의 움직임 벡터와 차-영상 에너지의 유사성을 판단하기 위하여 식 (5), (6)와 같이 경계치 조건을 정하였다. $MV_{Var.}$ 이 $MV_{Threshold}$ 보다 같거나 작으면 입력 움직임 벡터들 간에 유사성이 있다고 판단하고, 그렇지 않은 경우에는 유사성이 없다고 판단한다. 그리고 차-영상 에너지에 대해서도 동일하게, $Residual_{Var.}$ 이 경계치 값보다 같거나 작으면 매크로블록의 에너지들 간에 유사성이 있다고 판단하고, 그렇지 않은 경우에는 유사성이 없다고 판단한다.

표 1. 인트라 슬라이스에서의 통계특성
Table 1. Statistical characteristics of intra slice.

[%]

Decoded 4MBs' modes	Transcoded MB's mode	
	Intra4x4	Intra16x16
All 4 MBs' modes are Intra16x16	6%	94%
Else	97%	3%

여기에서 $MV_{Threshold}$ 와 $Residual_{Threshold}$ 을 각각 128과 36으로 결정하였다. 이는 4개 매크로블록 간의 움직임 벡터의 수평, 수직 성분 값들이 약 $\pm 1/2$ 화소 근접 범위 안에 존재하여 이웃한 움직임 벡터 값들이 아주 유사함을 띄고 있기 때문이다. 그리고 차-영상 에너지에 대해서도 4개 매크로블록 $Residual_{MB_j}$ 의 값들이 약 ± 6 근접 범위 안에 존재하여 이웃한 차-영상 에너지 간에 상관성이 크기 때문이다.

1. 입력 정보와 출력 매크로블록의 최적 모드간의 관계

전체 영상을 구성하는 단위인 슬라이스는 크게 인트라 슬라이스와 인터 슬라이스로 구분할 수 있다. 인터 슬라이스에서의 인트라 모드는 아주 작은 부분을 차지하지만 모든 매크로블록에서 최적의 모드를 결정하기 위해 인트라 모드 또한 계산해야 한다. 이에 따라 어떤 경우에 인트라 모드가 나올 가능성이 없는지를 파악하여 인트라 모드를 제외시킨다면 계산량을 감소시킬 수 있다. 입력된 4개의 매크로블록 중에 인트라 매크로블록이 존재하지 않는 경우가 그런 경우보다 재-부호화된 매크로블록이 인트라 모드가 발생할 가능성이 적다고 예상할 수 있다.

입력 매크로블록과 출력 매크로블록 정보간의 상관성을 파악하기 위해, 입력 비트스트림으로부터 도출된 모드, $MV_{Var.}$, 그리고 $Residual_{Var.}$ 정보를 토대로 재-부호화된 매크로블록의 최적화된 모드와의 확률적 상관성에 대한 결과는 표 1~3과 같다.

실험은 V장의 실험 조건과 동일한 조건에서 'Foreman', 'Mobile', 'Mother and Daughter', 'Paris' 4개 영상의 CIF급 영상과 QCIF급 영상에 대하여 QP 값을 각각 28로 부호화하여 얻은 결과이다.

먼저 인트라 슬라이스인 경우를 알아보면, 표 1에서 볼 수 있듯이, 인트라 슬라이스에서 매크로블록이 가질 수 있는 모드는 Intra4x4와 Intra16x16인데, 입력된 4개 매크로블록의 모드가 모두 Intra16x16인 경우에는 재-부호화 매크로블록의 모드가 Intra16x16이 될 가능성이

표 2. 인터 슬라이스에서 인트라 매크로블록이 존재하는 경우

Table 2. Case of inter slices having intra MBs.

[%]

Decoded 4MBs' information		Transcoded MB's mode				
		P16x16	P16x8 P8x16	P8x8	Intra4x4	Intra16x16
3 MBs are P16x16	Both of equations 5 & 6 are satisfied	85	4	1	0	10
	Else	65	22	6	1	6
2 MBs are P16x16	One MB is Intra MB	46	30	21	1	2
	2 MBs are Intra MB	41	20	11	10	18
Less than 2 MBs are P16x16	One MB is Intra MB	25	29	44	2	0
	2 MBs are Intra MB	19	29	33	17	2
	3 MBs are Intra MB	25	8	11	34	23
	All 4 MBs are Intra MB	2	3	4	92	0
	Less than 2 MBs are Intra16x16 Eles	11	4	0	34	51

표 3. 인터 슬라이스에서 인트라 매크로블록이 존재하지 않는 경우

Table 3. Case of inter slice having no intra MB.

[%]

Decoded 4MBs' information		Transcoded MB's mode			
		P16x16	P16x8 P8x16	P8x8	Intra MB
All 4 MBs are P16x16	At least, one of equations 5 & 6 are satisfied	95	5	0	0
	Else	65	25	10	0
3 MBs are P16x16	Both of equations 5 & 6 are satisfied	83	14	3	0
	Else	61	26	13	0
1 or 2 MBs are P16x16	Both of equations 5 & 6 are satisfied	73	21	6	0
	Else	42	28	30	0
None of all 4 MBs is P16x16		25	20	55	0

매우 높음을 알 수 있다.

인터 슬라이스 경우에는 입력된 4개 매크로블록 중에 인트라 매크로블록의 존재여부에 따라 크게 2가지로 분류하였다.

인터 슬라이스 안에 인트라 매크로블록이 존재할 경우의 확률적 상관성을 알아보면, 표 2에서 관찰할 수 있듯이, 우선 입력된 매크로블록 내에 인트라 매크로블록이 존재하더라도, 아주 많은 부분이 인터 매크로블록으로 재-부호화되며, 입력 매크로블록 중, P16x16 모드의 수가 많을수록 재-부호화된 매크로블록의 모드가 P16x16일 확률이 높다. 또한 P16x16 모드의 수가 3인 경우에는 움직임 벡터와 차-영상 에너지 특성에 의해서도 영향을 받는다. 그리고 입력 매크로블록 중, P16x16 모드의 수가 감소함에 따라, 출력 매크로블록 모드에서 P16x16 모드가 차지하는 비중은 감소하고, 그 외의 인터와 인트라 모드의 비중은 증가함을 관찰할 수 있다. 또한 입력된 4개 매크로블록이 모두 인트라 모드 일지라도, 대부분의 경우에는 인트라 모드로 재-부호화할 확률이 높지만, 경우에 따라서는 P16x16 모드로 재-부호화할 확률도 있음을 알 수 있다.

마지막으로 입력된 매크로블록 중 인트라 매크로블록이 존재하지 않을 경우, 입력된 영상과 재-부호화할 영상의 확률적 상관성을 관찰해 본 결과에 대해 살펴본다. 이 경우에는 재-부호화할 매크로블록의 모드가 P16x16인 경우가 많이 발생하고, 그 다음으로 P16x8과, P8x16이 많이 발생함을 알 수 있다. 또한 인트라 모드는 거의 발생하지 않음을 알 수 있다. 입력 P16x16 모드 수에 따라 재-부호화할 매크로블록 모드의 확률 상관성이 달라짐을 표 3을 통해 알 수 있다. 예로 입력된 P16x16 모드 수가 4인 경우에 재-부호화할 매크로블록의 모드가 P16x16이 될 확률이 높고, P16x16 모드 수가 감소함에 따라 다른 인터 모드 또한 최적모드로 발생할 확률이 증가한다. 하지만 입력 모드의 조건은 같을지라도 움직임 벡터와 차-영상 에너지의 경계치 조건의 만족 여부에 따라 재-부호화할 매크로블록의 최적 모드의 확률적 상관성이 변하게 된다. 표 3에서 보여주듯이, 입력된 P16x16 모드 수가 4이더라도 움직임 벡터와 차-영상 에너지의 경계치 조건 중 하나라도 만족할 경우에는 재-부호화할 매크로블록 모드로 P16x16이 나올 확률이 95%이지만, 두 개의 경계치 조건을 모두 만족 못

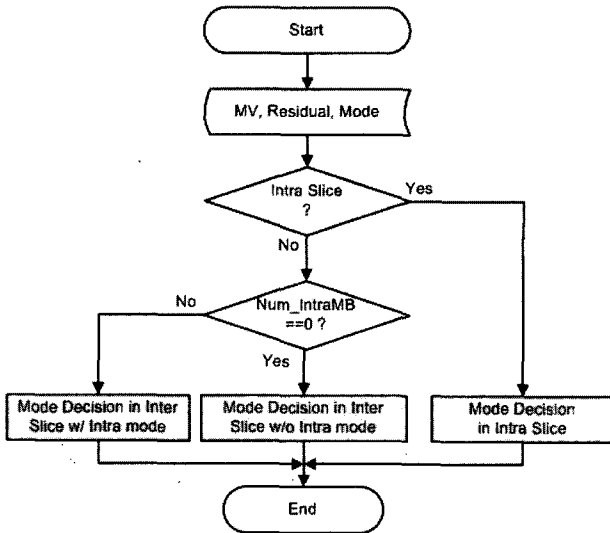


그림 5. 슬라이스 종류에 따른 모드 결정 방법
Fig. 5. Mode decision method by slice type.

할 시에는 P16x8이나 P8x16이 나올 확률이 25%로 높아짐이 관찰된다.

IV. 제안한 트랜스코딩을 위한 고속모드 결정 방법

이번 장에서는 III장에서 얻어진 입력정보와 출력된 매크로블록 모드의 확률적 상관성을 바탕으로 트랜스코딩을 위한 고속 모드 결정 방법을 제안한다. 전체적인 구조는 그림 5와 같이 크게 슬라이스 종류에 따라 2종류로 나누고, 인터 슬라이스에서는 인터 매크로블록의 존재 여부에 따라 다시 2가지 경우로 세분화한다.

1. 인트라 슬라이스 모드 결정 방법

그림 6은 인트라 슬라이스에서의 모드 결정 방법을 나타낸다. 입력된 4개의 매크로블록 모드가 모두 Intra16x16인 경우에는 Intra16x16을 재-부호화할 매크로블록의 모드로 결정하고, 그 이외에는 Intra4x4로 결정한다.

2. 인트라 모드가 존재하는 경우의 인터 슬라이스 모드 결정 방법

입력 인터 슬라이스 내에 인트라 모드가 존재할 때의 결정 방법을 다음과 같이 P16x16 매크로블록의 수에 따라 크게 3가지로 세분화 하였다.

(1) P16x16 매크로블록이 3개인 경우

표 2에서 재-부호화 매크로블록은 P16x16, P16x8, P8x16, 그리고 Intra16x16 모드로 되고, 그 중 P16x16이

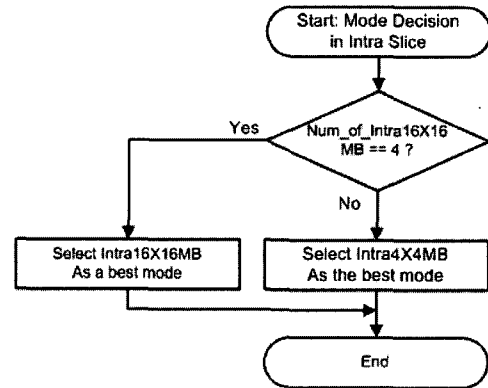


그림 6. 인트라 슬라이스에서의 모드 결정 방법
Fig. 6. Mode decision method in intra slice.

발생할 확률이 높게 나타났다. 그래서 모드 결정 과정을 좀 더 세분화하여, 식 (2), (4)에서 얻은 MV_{Var} 와 $Residual_{Var}$ 의 값이 식 (5), (6) 조건을 둘 다 만족시킬 경우, 재-부호화할 모드를 P16x16, 그리고 Intra16x16으로 제한한다. 그 외의 경우에는 재-부호화할 모드를 P16x16, P16x8, P8x16, 그리고 Intra16x16으로 제한한다.

(2) P16x16 매크로블록이 2개인 경우

인트라 매크로블록이 1개인 경우에는 재-부호화될 모드를 인터 모드로만 제한하고, 2개인 경우에는 아무런 제한을 두지 않고 모든 모드를 사용한다.

(3) P16x16 매크로블록이 1개 이하인 경우

이 경우에는, 표 2를 통해 알 수 있듯이 인트라 매크로블록의 수뿐만 아니라, 인트라 모드 형태에 따라서 재-부호화 매크로블록 모드를 결정한다. 위와 같은 사실을 토대로 하여 다음과 같이 재-부호화 매크로블록의 가능한 모드를 결정하였다.

- 인트라 매크로블록이 1개:
 - ✓ 모든 인터 모드
- 인트라 매크로블록이 2개:
 - ✓ Intra16x16을 제외한 모든 모드
- 인트라 매크로블록이 3개:
 - ✓ P16x8, P8x16을 제외한 모든 모드
- 인트라 매크로블록이 4개:
 - ✓ Intra16x16 매크로블록이 2개 이하
 - Intra4x4 모드
 - ✓ Intra16x16 매크로블록이 3개 이상
 - P16x16, Intra16x16, Intra4x4 모드

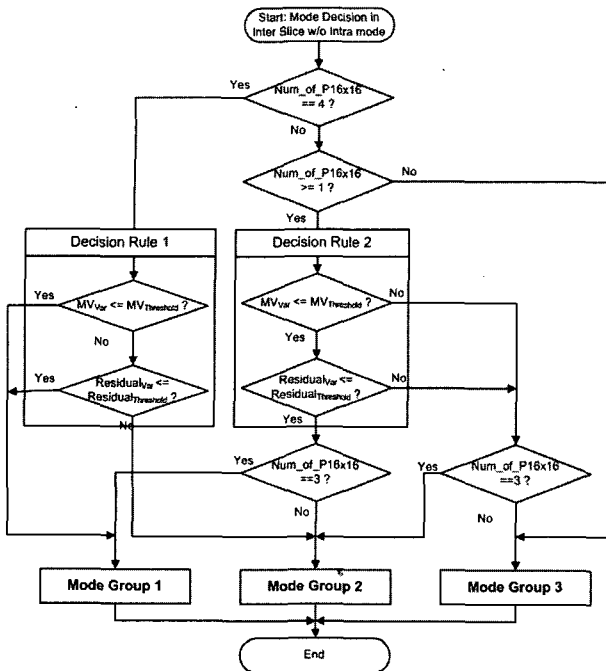


그림 7. 인터 모드 결정 방법
Fig. 7. Inter mode decision method.

3. 인트라 모드가 존재하지 않는 경우의 인터 슬라이스 모드 결정 방법

인터 슬라이스에서 입력된 4개의 매크로블록이 모두 인터 매크로블록인 확률은 약 95% 이상을 차지하여, 전체 모드 결정 과정 중에서 이 방법이 차지하는 비중은 아주 크다.

III장에서 도출해낸 실험결과를 바탕으로 발생 가능한 모드를 다음과 같이 3개의 그룹으로 분류하였다.

- 모드 그룹 1: P16x16
- 모드 그룹 2: P16x16, P16x8, P8x16
- 모드 그룹 3: 모든 인터 모드

이를 기반으로 하여, 인터 모드 결정 방법에 대한 전체적인 과정을 그림 7에 도식하였다. 인터 모드 결정 방법에서의 주된 분류는 1차적으로 입력된 매크로블록의 모드 정보를 가지고 하는데, 특히 P16x16 모드의 수에 따라서 3가지로 세분화 한 다음, MV_{Var} 와 $Residual_{Var}$ 의 값을 고려하여 재-부호화할 매크로블록이 속할 모드 그룹을 정한다.

(1) P16x16 매크로블록이 4개인 경우

재-부호화할 매크로블록의 모드가 대부분 모드 그룹 1에 포함되지만, 움직임 벡터와 차-영상 에너지의 확률 분포에 따라서 모드 그룹 2에 포함되기도 한다. 따라서 그림 7에서 보는바와 같이 결정 규칙 1 (Decision Rule 1)에 의해 재-부호화할 매크로블록이 속할 모드 그룹을

표 4. 실험조건
Table 4. Experimental conditions.

	Input Stream	Output Stream
Profile	Baseline	Baseline
Sequence Size	352x288	176x144
Frame number	300	300
Frame-rate	30 frames/sec	30 frames/sec
Intra Period	15	15
Reference Frame #	1	1
QP	28	28, 32, 36, 40
Rate-Distortion Option	ON	ON

정한다. 식 (5), (6) 두 조건을 모두 만족하지 못할 경우에는 모드 그룹 2를 선택하고, 그 외의 경우에는 모드 그룹 1을 선택한다.

(2) P16x16 매크로블록이 없는 경우

이 경우에는 모든 인터 모드의 확률분포가 골고루 존재한다. 그리고 이 경우에는 MV_{Var} 와 $Residual_{Var}$ 의 값 보다는 모드에 의해서 주로 결정이 되기 때문에, 아무런 조건 없이 재-부호화할 매크로블록을 모드 그룹 3으로 선택한다.

(3) 그 외 경우

P16x16이 3개인 경우에 결정 규칙 2 (Decision Rule 2)의 식 (5), (6) 두 조건을 모두 만족하면, 재-부호화할 매크로블록을 모드 그룹 1로 선택하고, 그렇지 않으면 모드 그룹 2로 선택한다. P16x16이 2개나 1개인 경우에는 결정 규칙 2를, 모두 만족하면 재-부호화할 매크로블록을 모드 그룹 2로 선택하고, 그렇지 않으면 모드 그룹 3을 선택한다.

V. 실험

본 논문에서 제안된 방법의 성능을 평가하기 위해 H.264/AVC JM (Joint Model) 9.3 참조 소프트웨어를 기반으로 하는 직렬형 (cascade) 트랜스코딩 시뮬레이터를 구성하였다. 실험용 영상으로는 'Foreman', 'Mobile', 'Mother and Daughter', 'Paris' 등 4개의 영상을 이용하였으며, 자세한 실험 조건은 표 4와 같다.

제안 방법의 성능 평가를 위해 제안한 고속 모드 결정법과 기존 방법 [9]과의 PSNR, 비트율, 그리고 소요되는 속도를 비교하였다. 여기에서 속도는 재-부호화에 소요되는 속도만을 고려하였는데, 이는 복호화에 소요되는 시간은 모두 동일하기 때문이다.

그림 8은 각각의 영상에 대한 비트율 대비 PSNR 값

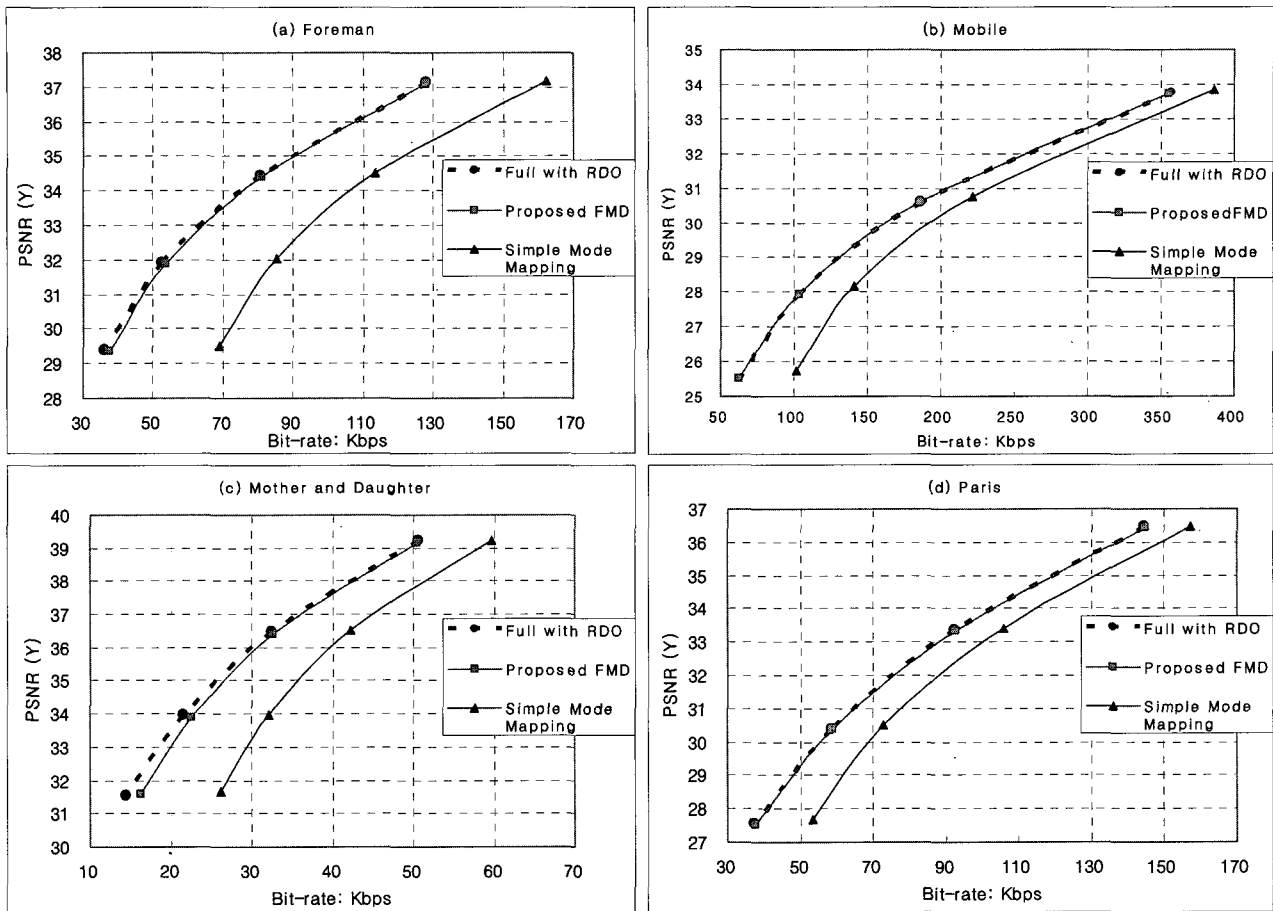


그림 8. PSNR과 비트율 비교
 Fig. 8. Comparison of the PSNR and bit-rate.

을 도표화 한 것으로, 실험 균으로는 본 논문에서 제안한 방법(Proposed FMD), 비교 균으로는 Zhang이 제안한 간단한 모드 결정법 (Simple Mode Mapping)^[9], 그리고 일반적인 JM 9.3 참조 소프트웨어에서의 모드 결정방법이 (Full with RDO) 있다.

그림 8에서 볼 수 있듯이, Proposed FMD (Fast Mode Decision)는 PSNR과 비트율에 있어서 Full with RDO에 비하여 성능 차이가 거의 없는 반면에, Simple Mode Mapping은 Full with RDO와 비교하여 큰 차이가 있음을 관찰할 수 있다.

표 5는 실험 영상 4개에 대한 PSNR, 비트율, 그리고 계산속도 향상에 있어서, Full with RDO를 기준으로 하여 Proposed FMD와 Simple Mode Mapping의 성능을 수치화한 것이다. 그리고 직렬형 (cascade) 트랜스코더의 성능에 대비하여 Δ PSNR은 PSNR 값의 증가치를, bit-rate(%)는 비트율의 증가치를, Computational improvement는 재-부호화 속도의 증가치를 의미한다. Δ PSNR의 값이 +이면 성능이 향상되었고, Δ bit-rate

경우는 값이 -이면 성능이 향상되었음을 의미한다.

QP 값이 28일 경우, Proposed FMD를 Full with RDO와 비교하면, 비트율은 0.15% 감소하였고, PSNR 값은 0.05dB 감소하여 성능에 있어서 차이가 거의 없다. 재-부호화하는 QP 값이 증가하면서 Proposed FMD의 비트율이 Full with RDO의 비트율 보다 약간 높아진다. 재-부호화하는 QP 값이 40일 경우, 평균적으로 약 4.69% 증가 한다. 이러한 현상의 원인은 입력과 출력 비스트림에서 이용하는 QP 값이 다르기 때문이다. QP 값의 차이가 클수록 모드의 확률 분포가 약간씩 달라진다. 이로 인해, 최적 모드 확률 분포의 신뢰도가 떨어져 비트율이 약간씩 증가된다. 하지만 이러한 성능 저하는 계산속도 면에서 얻어지는 이점이 훨씬 큰 응용에서는 무시할 수 있다.

Proposed FMD의 성능은 전체적으로 PSNR은 약 0.04dB 감소하고, 비트율은 약 1.6%로 약간 증가하지만, 재-부호화 속도는 2.66배 증가한다. Simple Mode Mapping의 성능은 전체적으로 PSNR은 약 0.09dB 향

표 5. PSNR, 비트율 차이 및 계산적인 향상치

Table 5. Differences of PSNR and bit-rate and computational increasement.

Sequence	Comparison QP	Δ PSNR (dB)		Δ bit-rate (%)		Computational improvement (times)	
		Proposed FMD	Simple	Proposed FMD	Simple	Proposed FMD	Simple
Foreman	28	-0.06	0.02	-0.05	26.56	2.47	2.97
	32	-0.04	0.08	0.72	40.81	2.41	2.91
	36	-0.06	0.13	1.81	61.89	2.36	2.86
	40	-0.07	0.10	4.01	90.58	2.31	2.81
Mobile	28	-0.02	0.07	-0.39	8.32	2.44	3.00
	32	0.00	0.15	-0.04	18.87	2.34	2.89
	36	0.00	0.23	-0.03	36.70	2.26	2.79
	40	0.00	0.25	0.58	62.41	2.16	2.68
Mother and Daughter	28	-0.07	0.01	-0.40	17.98	3.12	3.26
	32	-0.08	0.03	0.67	30.08	3.05	3.21
	36	-0.08	0.02	4.02	48.84	3.00	3.17
	40	0.01	0.12	12.87	81.62	2.96	3.15
Paris	28	-0.05	0.00	0.24	9.03	3.05	3.31
	32	-0.03	0.05	0.44	14.78	2.96	3.22
	36	-0.02	0.11	0.46	24.64	2.87	3.15
	40	-0.05	0.11	1.29	43.18	2.80	3.09

상 되었고 속도는 약 3배 증가하였으나, 비트율은 오히려 40%정도 증가하여, 전체적인 성능 면에서는 Proposed FMD가 보다 효율적임을 알 수 있다.

VI. 결 론

본 논문에서는 H.264/AVC 공간해상도 축소 트랜스코딩을 위한 고속 모드 결정 방법을 제안하였다. 고속 모드 결정 방법은 입력된 매크로블록의 모드, 움직임 벡터, 그리고 차-영상 에너지 정보를 이용하여 재-부호화할 매크로블록의 모드 결정에 필요한 계산상의 복잡도를 감소시키는데 그 목적이 있다. 제안된 고속 모드 결정법은 일반적인 부호화 방법에 비교하여 영상의 화질이나 압축률의 저하가 크게 없이 재-부호화하는 과정에서 2.16~3.12배 속도의 개선이 있음을 실험을 통해 확인하였다.

본 논문에서는 입력 비트스트림의 정보를 이용하여 단지 재-부호화할 매크로블록의 모드 결정에 대해서만 알아보았다. 제안된 고속 모드 결정 방법과 더불어 입력된 정보를 이용하여 움직임 벡터를 찾는 데 드는 시간을 단축하는 기존의 방법과 병행하여 사용한다면, 더 많은 속도의 개선이 있을 것으로 예상된다.

참 고 문 헌

- [1] J. Xin, C. W. Lin, and M. T. Sun, "Digital Video Transcoding", Proceedings of the IEEE, Vol. 93, Issue 1, pp. 84-97, January 2005.
- [2] 권혁균, 이영렬, "H.264에서 MPEG-4로 빠른 트랜스코딩," 전자공학회 논문지 제 41권 CI편 제6호, 91-99쪽, 2004년 11월.
- [3] 이상선, 이영렬, "MPEG-4에서 H.264로 빠른 트랜스코딩," 전자공학회 논문지 제 41권 SP편 제5호, 275-282쪽, 2004년 9월.
- [4] B. Shen, I. K. Sethi, and V. Bhaskaran, "Adaptive Motion Vector Resampling for Compressed Video Down-Sampling," IEEE Transactions on Circuits and Systems for Video Technology, vol. 9, pp. 929-936, September 1999.
- [5] Y. Tan, H. Sun, and Y. Liang, "On the Methods And Applications of Arbitrarily Downsizing Video Coding," Proceedings of 2002 IEEE International Conference on Multimedia and Expo, vol. 1, pp. 609-612, August 2002.
- [6] J. Xin, M. Sun, B. Choi, and K. Chun, "An HDTV-to-SDTV Spatial Transcoder," IEEE Transactions on Circuits and Systems for video Technology, vol. 12, no. 11, pp. 998-1008, November 2002.
- [7] S. H. Jang and N. Jayant, "An Adaptive Non-

- linear Motion Vector Resampling Algorithm For Down-scaling Video Transcoding," Proceedings of 2003 International Conference on Multimedia and Expo, vol. 2, pp. 229-232, July 2003.
- [8] Y. P. Tan and H. Sun, "Fast Motion Re-Estimation for Arbitrary Downsizing Video Transcoding using H.264/AVC Standard," IEEE Transactions on Consumer Electronics, vol. 50, pp 887-894, August 2004.
- [9] Zhang, Y. Lu, Q. Huang, and W. Gao, "Mode Mapping Method For H.264/AVC Spatial Dwnscaling Transcoding," Proceedings of 2004 IEEE International Conference on Image Processing, vol. 4, pp 2781-2784, October 2004.
- [10] C. Li, C. Wang, and T. Chiang, "A Fast Downsizing Video Transcoder Based on H.264/AVC Standard," Lecture Notes in Computer Science, vol. 3333, pp. 215 - 223, November 2004.
- [11] Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, "Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H.264 & ISO/IEC 14496-10 AVC)," Doc. JVT-G050r1, March 2003.
- [12] 이제윤, 전병우, "H.264 동영상 압축을 위한 고속 부호화 모드 결정 방법," 전자공학회 논문지 제 41 권 SP편 제6호, 165-173쪽, 2004년 11월.

저 자 소 개



권 상 구(정회원)
 2004년 성균관대학교 정보통신 공학부 졸업 (학사)
 2006년 성균관대학교 전자전기 공학과 졸업 (석사)
 2006년 2월~현재 현대기아자동차 연구개발본부 연구원
 <주관심분야: 컴퓨터, 신호처리, 영상 압축, 비디오 트랜스코딩>



전 병 우(정회원)
 1985년 2월 서울대학교 전자공학과 졸업 (학사)
 1987년 2월 서울대학교 전자공학과 졸업 (석사)
 1992년 12월 Purdue Univ, School of Elec. 졸업 (공학박사)
 1993년~1997년 8월 삼성전자 신호처리연구소 수석연구원
 1997년 9월~현재 성균관대학교 정보통신공학부 부교수
 <주관심분야: 멀티미디어 영상 압축, 영상인식, 신호처리>



정 봉 수(학생회원)
 2002년 2월 성균관대학교 전기 전자 및 컴퓨터공학부 졸업(학사)
 2004년 8월 성균관대학교 정보통신공학부 졸업 (석사)
 2004년 9월~현재 성균관대학교 전자전기공학과 박사과정
 <주관심분야: 영상압축, 멀티미디어 통신, 에러 강인 디지털 비디오, 멀티미디어 VLSI, Resource-Aware 비디오 처리>