

MPEG2에서 H.264로의 트랜스코딩 기법

정회원 김 동 형*, 정 제 창**

An MPEG2-to-H.264 Transcoding Method

Donghyung Kim*, Jechang Jeong** *Regular Members*

요 약

본 논문에서는 MPEG-2로 부호화된 비디오 스트림에서 H.264 부호화 표준으로의 트랜스코딩 기법을 제안한다. 제안하는 MPEG2에서 H.264로의 트랜스코더는 MPEG2 복호화부와 H.264 부호화부로 구성되어 있으며, MPEG2 복호화부가 입력되는 MPEG2 비디오 스트림의 전복호화(全復號化)를 수행하는 공간영역에서의 트랜스코딩 기법이다. 반면 MPEG2 복호화 과정에서 사용되는 매크로블록 타입과 부호화블록패턴은 H.264의 매크로블록 모드와 상당한 연관성을 가지는데, 이를 이용하여 H.264 부호화부에서는 MPEG2의 매크로블록 타입, 부호화블록패턴 정보를 이용하여 H.264 부호화과정에서의 매크로블록 모드를 적응적으로 선택하고 추가적으로 움직임 벡터를 활용함으로써 전체 트랜스코더의 복잡도를 감소시킨다. 제안하는 트랜스코더를 시퀀스에 적용한 결과 전복호화 후 전부호화의 결과와 비교하여 평균 0.03dB의 PSNR 차이만을 보이는 반면 전체 트랜스코딩 시간은 최대 61% 감소된 것으로 나타난다.

Key Words : transcoding, MPEG2, H.264, macroblock type, macroblock mode, coded block pattern.

ABSTRACT

In this paper, we present a transcoding algorithm for converting an MPEG-2 video bitstream to an H.264 bitstream. The proposed transcoder consists of two parts. One is MPEG2 decoding part and the other is H.264 encoding part. Because our algorithm is for transcoding in the spatial domain, MPEG2 decoding part carries out full decoding of MPEG2 bitstream. While, because macroblock type and coded block pattern in MPEG2 are significantly related to macroblock mode in H.264, macroblock mode is selected adaptively according to macroblock type and coded block pattern in H.264 decoding part. Furthermore, motion vector is also used as side-information for 16x16 macroblock mode. Simulation results show that the proposed transcoder yields high reduction of total transcoding time at comparable PSNR.

I. 서 론

비디오 트랜스코딩이란 부호화된 비디오 스트림의 특성을 다른 특성으로 변환하는 것을 지칭한다 [1]. 이러한 비디오 스트림의 특성에는 비트율, 프레임율, 공간적 해상도 및 부호화 표준 등이 포함될

수 있다. 트랜스코딩을 입출력 비디오 스트림의 비디오 표준에 따라 분류하면 크게 동일한 비디오 표준 간의 트랜스코딩과 상이(相異)한 비디오 표준 간의 트랜스코딩으로 나눌 수 있다. 동일한 비디오 표준 간의 트랜스코딩은 일반적으로 비트율 감소를 주(主) 목적으로 하며 이에 양자화 계수의 변환을

* 한양대학교 전자통신전파공학과 영상처리 및 신호처리 연구실(kimdh@ece.hanyang.ac.kr)

** 한양대학교 전자전기컴퓨터 공학부(jjeong@ece.hanyang.ac.kr)

논문번호 : KICS2005-05-182, 접수일자 : 2005년 5월 2일

※ 본 논문은 정보통신부의 출연금으로 수행한 IT SoC 핵심설계인력양성사업의 수행결과입니다.

통하여 비트율을 감소시키는 SNR 트랜스코딩^[2]과 프레임율의 변환을 수행하는 프레임 트랜스코딩^[3], 그리고 프레임의 공간적 크기 변환을 수행하는 공간적 해상도 트랜스코딩으로 나눌 수 있다^[4]. 서로 상이한 비디오 표준간의 트랜스코딩은 서로 다른 영상 표준을 지원하는 단말기간의 전송 등과 같이 주로 영상 포맷 자체의 변환을 목적으로 하며, MPEG2에서 MPEG4 Part2로의 트랜스코딩^[5] 등이 이에 속한다.

네트워크 단말기 분야의 발달과 함께 여러 단말기에서의 영상 콘텐츠 소비가 늘어나면서 이러한 트랜스코딩의 중요성은 더욱 증대되고 있다. 특히 현존하는 비디오 표준들 중 가장 높은 부호화 효율을 가지는 H.264 부호화 표준을 대상으로 하는 트랜스코딩 기법의 개발은 많은 이점(利點)을 가져올 수 있다. 하지만 부호화의 요소 기술 측면에서 이전의 비디오 영상표준들(MPEG1, MPEG2, MPEG4 Part2 등)과 많이 상이점을 보이는 H.264 비디오 표준은 높은 부호화 효율만큼이나 높은 복잡도를 가진다. 따라서 H.264 비디오 표준으로의 영상 포맷 변환의 경우 전부호화 후 전부호화(全復號化 後全符號化)의 트랜스코더 구조는 최적의 성능을 나타내지만 높은 복잡도로 인한 많은 문제점을 갖는다.

H.264 비디오 표준으로의 트랜스코더 복잡도 문제를 해결하기 위해 많은 연구들이 이루어져 왔다. Bialkowski 등은 H.263 인트라(Intra) 프레임과 H.264 인트라 프레임 사이의 고속 트랜스코딩 기법을 제안하였으며, 이는 H.263의 세 가지의 인트라 블록 예측방향(DC, Vertical, Horizontal)과 H.264에서의 인트라 16x16에서의 4가지 예측방향 및 인트라 4x4에서의 9가지 예측 방향과의 상관성을 이용하여 인트라 프레임간의 트랜스코딩을 수행하는 방법으로 H.264의 모든 인트라 예측방향을 검색하는 것에 비해 약 4.2%의 복잡도만을 갖는다^[6]. 하지만 비트율에 따라 최대 1dB까지의 화질열화가 발생하는 단점을 가진다. 또한 Bialkowski 등은 움직임 벡터 재추정을 통한 H.263 및 H.264의 예측(Predicted) 프레임간의 고속 트랜스코딩 기법도 제안하였으며, 이는 움직임 추정시의 검색 지점의 수를 현저히 줄임으로써 0.1~0.5dB의 화질열화가 발생하지만 약 90%의 트랜스코딩 복잡도 감소효과를 보인다^[7]. H.263에 비하여 H.264 부호화 표준과 부호화 요소기술면에서 더욱 많은 차이를 보이는 MPEG 비디오 표준으로부터 H.264 비디오 표준으로의 트랜스코딩 또한 활발히 진행되고 있다. Kalva는 그의

논문에서 H.264와 MPEG2 사이의 트랜스코딩 효율성과 함께 MPEG2에서 H.264로의 개략적(概略的) 트랜스코더를 소개하였으며^[8], Chen 등은 변환영역(Transform domain)에서 MPEG2에서 H.264로의 트랜스코딩 기법을 제안하였다^[9]. 이는 MPEG2로 부호화된 비디오 스트림을 입력으로 받아 변환영역까지 복호화 한 후 변환영역에서 8x8 DCT 기저(Basis)와 4x4 DCT 기저 사이의 변환과 해상도의 축소(1/2)를 수행함으로써 트랜스코딩을 수행한다. 이러한 Chen의 트랜스코딩 기법은 서로 다른 변환 기법을 사용하는 MPEG2와 H.264 비디오 표준사이의 트랜스코딩을 변환영역에서 수행함으로써 복잡도 측면에서 강점을 가지지만 공간영역에서 수행되는 루프필터를 고려하지 못하는 단점이 있다.

MPEG2로 부호화된 비트스트림 내에는 움직임 벡터와 같은 복호화에 필요한 모든 정보가 포함되어 있다. 이중 매크로블록 타입은 매크로블록이 참조프레임으로부터 예측을 사용하였는지의 여부(Intra/Inter)와 움직임 벡터의 값이 (0,0)인지의 여부(MC/No-MC) 또는 전방향(Forward) 및 역방향(Backward) 예측 또는 양방향(Interpolated) 예측인지의 여부를 나타내는 정보이며, 프레임 구조(I, P, B)에 따라 서로 다른 가지 수의 매크로블록 타입을 가진다. 또한 매크로블록 내(內)의 8x8 블록에 대한 양자화 이후의 값에 따라 각 8x8블록의 부호화 여부(Coded/Not-Coded)를 나타내는 부호화블록패턴(Coded Block Pattern) 정보 또한 MPEG2 비트스트림 내에 포함되어 있다. 이러한 MPEG2의 매크로블록 타입과 부호화블록패턴은 매크로블록 예측 오차의 크기와 관련이 있으며 이는 H.264의 매크로블록 모드(16x16, 16x8, 8x16 등)를 결정하는 율-왜곡 비용(Rate-Distortion Cost) 함수에서의 왜곡 값과 상당한 연관성을 가진다.

이러한 상관성을 이용하여 본 논문에서는 MPEG2 비트스트림 내에 포함되어 있는 움직임벡터 정보, 매크로블록 타입 및 부호화블록패턴을 H.264 부호화시에 부가정보로서 사용함으로써 화질열화를 최소화함과 동시에 트랜스코더의 복잡도를 감소시키는 MPEG2 비디오 표준^[10]에서 H.264 비디오 표준^[11]으로의 트랜스코딩 기법을 제안한다. 제안하는 트랜스코더는 MPEG2의 매크로블록 타입과 부호화블록패턴을 사용하여 H.264 부호화 과정에서 선택할 수 있는 매크로블록 모드를 적응적으로 선택함으로써 트랜스코더 내의 매크로블록 모드 선택과정의 복잡도를 현저히 줄일 수 있다.

II. MPEG2 및 H.264 표준 비교

MPEG2 비디오 표준과 H.264 비디오 표준은 요소기술면에서 상당한 차이점을 가진다. 특히 공간영역에서 수행되는 루프필터 및 인트라 예측과정, 그리고 4x4 정수변환은 MPEG2와 비교되는 H.264의 부호화 특징들이다. 본 장에서는 두 표준의 개략적 차이점과 함께 제안하는 알고리즘에서 사용하게 될 MPEG2에서의 매크로블록 타입과 H.264에서의 매크로블록 모드에 관하여 기술한다.

2.1 MPEG2 및 H.264의 개략적 차이점

2.1.1 엔트로피 부호화

MPEG2에서와는 달리 H.264 비디오 표준에서는 프로파일에 따라 두 가지의 엔트로피 부호화를 사용한다. 첫 번째 엔트로피 부호화 기법은 변환 계수를 제외한 대부분의 신택스(Syntax)를 데이터의 특성에 따라 하나의 코드워드(Codeword) 테이블로 변환 후 EGC(Exponential Golomb Code)로 부호화하고, 변환계수는 이미 부호화 된 주변 블록들의 특성을 이용하는 CAVLC(Context Adaptive Variable Length Coding)로 부호화 하는 방법이다. 두 번째의 엔트로피 부호화 기법은 기존의 산술 부호화 방식에 컨텍스트 모델링(Context Modeling)을 추가함으로써 추정된 심볼(Symbol) 확률 분포 모델들을 적용적으로 선택하는 방식으로 EGC와 CAVLC를 이용하는 엔트로피 부호화 방식에 비해 높은 복잡도를 가지지만 5~15%의 비트율 감소효과가 있다.

2.1.2 인트라 예측 부호화

MPEG2에서는 주파수 영역에서 DC 계수만을 가지고 인트라 예측을 사용하는 반면 H.264 비디오 표준은 공간영역에서의 인트라 예측을 수행하며, 3가지의 인트라 모드에 따라 다양한 예측방향을 제공한다. 인트라 4x4의 경우 모두 9가지의 예측 모드를 가지며, 인트라 16x16은 경우 4가지의 예측 모드를 가진다. 또한 H.264 버전 3에서는 인트라 8x8 모드가 추가되었으며 인트라 4x4와 동일한 9가지의 예측 방향을 가진다. 공간영역에서의 인트라 예측은 영상의 평탄한 영역뿐만 아니라 세밀한 영역에서도 높은 압축 효율을 갖게 한다.

2.1.3 참조 프레임 및 움직임 예측

MPEG2에서는 한 장(전방향 및 역방향 예측의 경우) 또는 두 장(양방향 예측의 경우)의 I 또는 P 프레임을 참조프레임으로 사용하는 반면 H.264 비

디오 표준에서는 참조 영상으로서 이전의 재구성된 여러 장의 프레임을 사용한다. 이러한 다중 참조 영상의 사용은 단일 참조 영상만을 사용하는 경우에 비해 보다 높은 압축 효율을 가지게 하며, 특히 반복적인 영상이나 다른 객체에 의해 가려져 있다가 나타난 영역의 부호화 경우에 더욱 두드러진 성능을 보인다. 또한 MPEG2의 움직임 예측은 1/2 화소 단위까지 수행되며 이때 보간(Interpolation)에는 2탭 쌍선형(Bilinear) 필터를 사용한다. 반면 H.264 비디오 표준에서는 1/4 화소 단위의 움직임 예측 정밀도를 가지며, 1/2 화소 위치의 보간에는 6탭 필터, 1/4 화소 단위 보간에는 2탭의 쌍선형 필터를 사용함으로써 움직임 추정의 정확도를 증가시킨다.

2.1.4 움직임 추정에 사용되는 블록 크기

MPEG2에서는 16x16 및 16x8의 크기로 움직임 추정이 수행되는 반면 H.264에서는 블록의 크기를 다양화하여 16x16, 16x8, 8x16, 8x8, 8x4, 4x8, 4x4의 크기로 움직임을 추정한다. 이러한 다양한 크기의 블록 단위, 특히 8x8 부-매크로블록(Sub Macroblock) 이하의 크기인 8x4, 4x8, 4x4 블록 크기의 움직임 추정은 보다 세밀한 영상 영역에서도 높은 압축 효율을 가질 수 있게 한다^[2].

이외에도 슬라이스 구조 및 변환(Transform) 단위 등에서도 많은 차이점을 보인다.

2.2 MPEG2에서의 매크로블록 타입

MPEG2의 부호화 과정에서의 매크로블록 타입이란 매크로블록이 참조프레임으로부터 예측을 사용하였는지의 여부(Intra/Inter)와 움직임 벡터의 값이 (0,0) 인지의 여부(MC/No-MC) 또는 전방향 및 역방향 예측 또는 양방향 예측인지의 여부(Forward/Backward/Interpolated)등을 나타내는 정보이며 프레임의 구조(I, P, B)에 따라 서로 다른 가지 수의 매크로블록 타입을 갖는다.

먼저 I 프레임내의 매크로블록은 참조프레임으로부터의 예측과정 없이 이전 매크로블록의 양자화 계수를 그대로 사용하여 인트라로 부호화되는 경우와 새로운 양자화계수를 사용하여 인트라로 부호화되는 경우가 있으며, 이 두 가지의 매크로블록 타입 중 하나로 선택될 수 있다.

반면 P 프레임 또는 B 프레임의 경우에는 참조프레임으로부터의 예측을 사용함으로써 보다 많은 선택 가능한 매크로블록 타입을 가지는데 먼저 P 프레임을 살펴보면 움직임벡터 정보의 유무에 따라 움직임추정(MC)과 움직임비추정(No-MC)으로 나뉘

고 움직임추정의 경우 다시 참조프레임으로부터의 예측 사용 여부에 따라 인트라(Intra)와 인터(Inter)로 나뉜다. 또한 매크로블록 내의 양자화 이후의 값이 모두 0값을 갖는 경우는 부호화 되지 않으므로 이러한 부호화의 여부에 따라 부호화(Coded)와 비부호화(Not-Coded)로 나뉜다. 마지막으로 I 프레임에서와 동일하게 사용되는 새로운 양자화 계수를 사용하는 경우 변환된 양자화 계수(Change QP)를 포함하는 매크로블록 타입이 선택된다. 특히 움직임 비추정 인터 비부호화의 매크로블록 타입은 부호화 시 어떠한 정보도 포함하지 않는 경우를 나타낸다. 이러한 스킵타입을 포함하여 P 프레임내의 매크로블록은 8가지의 매크로블록 타입 중 하나로 선택되며 그림 1은 이를 도시한다.

B 프레임은 P 프레임과 유사한 형태의 매크로블록 타입을 갖지만 양방향 예측을 사용함으로써 보다 많은 매크로블록 타입을 가진다. 또한 B 프레임에서의 스킵타입은 P 프레임의 스킵타입과 다른 의미를 가진다. P 프레임에서는 움직임 벡터 값이 (0, 0)이며 매크로블록내의 모든 블록이 부호화 되지 않은 경우 선택되나 B 프레임의 경우 이전의 매크로블록과 같은 움직임 벡터 값을 가지면서 움직임 벡

터의 추정방향(Forward/ Backward/Interpolated)이 같은 경우 스킵되어진다. 이러한 매크로블록들은 추가의 DCT 계수에 대한 정정이 없기 때문에 전송하는데 아주 적은 비트만을 필요로 한다. B프레임내의 매크로블록 타입은 스킵모드 이외에 11개의 매크로블록 타입을 가지며 그림 2는 이를 도시한다.

2.3 H.264에서의 매크로블록 모드

매크로블록 모드는 매크로블록의 공간 영역 내에서의 예측 및 시방향(時方向) 예측의 형태를 지칭하며, H.264 부호화 과정에서의 이러한 매크로블록 모드는 프레임의 구조(I, P, B)에 따라 서로 다른 모드 집합을 가진다. 먼저 I 프레임 내의 매크로블록은 공간 영역 내에서의 예측만을 사용하며 이러한 공간 내 예측을 사용하는 매크로블록 모드에는 인트라 16x16, 인트라 8x8 및 인트라 4x4 매크로블록 모드가 있다. 즉, I 프레임내의 매크로블록은 이제 가지 매크로블록 모드 중 하나로 선택될 수 있다. 그림 3은 I 프레임 내에서 선택 가능한 세 가지의 매크로블록 모드의 블록 모양을 나타낸다.

위의 세 가지 매크로블록 모드 중에서 인트라 16x16모드는 4가지의 공간 내 예측 방향을 가지며 인트라 8x8 및 인트라 4x4는 9가지의 공간 내 예측 방향을 가진다. 표 1은 I 프레임에서 선택 가능한 세 가지 매크로블록 모드에서의 예측 방향을 나타낸다.

반면 P, B 프레임내의 매크로블록은 다양한 크기의 움직임 추정을 사용하는 H.264의 특징으로 인하여 보다 다양한 매크로블록 모드를 가지는데, 선택 가능한 매크로블록 모드로는 16x16, 16x8, 8x16, 8x8 모드가 있으며 8x8 크기의 부-매크로블록은 다시 8x8, 8x4, 4x8, 4x4의 크기로 나누어질 수 있다.

또한 P, B 프레임내의 매크로블록 모드는 공간 내 예측을 사용하는 매크로블록 모드도 가질 수 있으며, MPEG2에서와 같이 SKIP 모드도 선택 대상이 된다. 따라서 P, B 프레임내의 매크로블록은 16x16, 16x8, 8x16, 8x8, 인트라 16x16, 인트라 8x

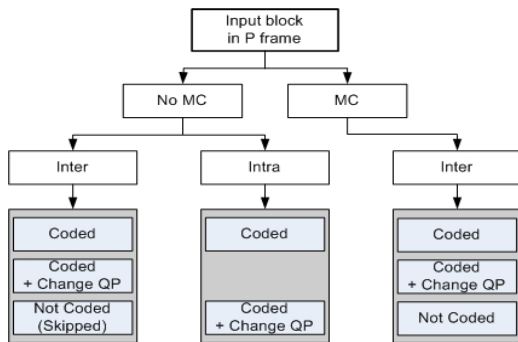


그림 1. P 프레임에서의 매크로블록 타입
Fig. 1. Macroblock types in P frame

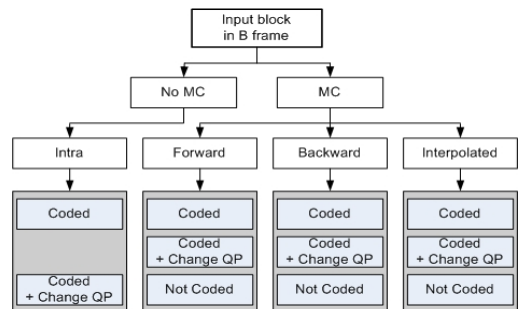


그림 2. B 프레임에서의 매크로블록 타입
Fig. 2. Macroblock types in B frame

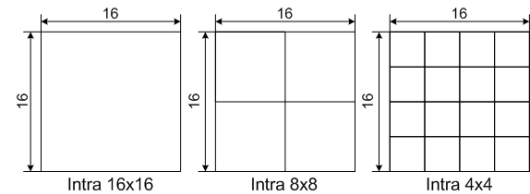


그림 3. I 프레임에서의 매크로블록 모드
Fig. 3. Macroblock modes in I frame

표 1. 인트라 예측 모드의 예측 방향
Table 1. Prediction directions of intra prediction mode

Prediction Mode	Intra4x4 Intra8x8	Intra16x16
0	Vertical	Vertical
1	Horizontal	Horizontal
2	DC	DC
3	Diagonal Down Left	Plane
4	Diagonal Down Right	
5	Vertical Right	
6	Horizontal Down	
7	Vertical Left	
8	Horizontal Up	

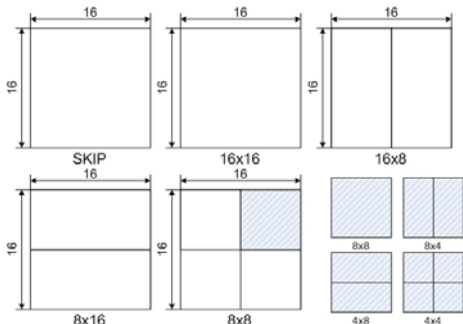


그림 4. 시방향 예측을 위한 매크로블록 모드
Fig. 4. Macroblock modes for temporal prediction

8, 인트라 4x4, 그리고 SKIP 모드 중에서 선택될 수 있다. 그림 4는 P, B 프레임 내에서 선택 가능한 매크로블록 모드 중 인트라 모드를 제외한 매크로블록 모드를 나타낸다.

그림 4에서 보듯이와 같이 P, B 프레임내의 매크로블록은 세 개의 인트라 매크로블록 모드 이외에 다섯 가지의 추가적인 매크로블록 모드를 포함하며 특히 8x8 매크로블록의 경우 각 블록은 네 가지의 부-매크로블록 모드 중에서 선택 가능하다.

2.4 MPEG2 매크로블록 타입과 H.264 매크로블록 모드와의 상관성

H.264 부호화 과정에서 매크로블록 모드의 선택은 각 매크로블록 모드에 대한 비용값들을 구하고 그 비용값을 최소로 하는 모드를 선택함으로써 수행되어진다. 율-왜곡 최적화(Rate-Distortion Optimization) 기법이 사용되는 경우 비용값 산정에 사용되는 함수가 율-왜곡 함수이며 식 (1)과 같다.

$$RDcost = Distortion + \lambda \cdot Rate \quad (1)$$

율-왜곡 비용함수내의 왜곡(Distortion) 값은 각 매크로블록 모드에서의 SNR을 구함으로써 계산할 수 있으며 율(Rate) 값은 매크로블록의 부호화가 끝나는 시점에서의 압축된 비트열(bitstream)의 비트율을 나타낸다.

H.264에서 매크로블록 모드선택에 사용되는 율-왜곡 비용함수의 파라미터 중 왜곡의 크기는 MPEG2의 매크로블록 타입과 많은 상관성을 가지고 있다. 이는 MPEG2의 매크로블록 타입 및 부호화블록패턴이 움직임 예측 오차의 크기에 따라 결정되기 때문이다. 가령 MPEG2의 매크로블록 타입이 스킵타입이었다면 이는 해당 매크로블록의 움직임 추정 이후의 변환 및 양자화 과정에서 모든 값이 0값이 됨을 의미하며 이는 아주 적은 왜곡만을 갖는다고 볼 수 있다. 또한 P, B 프레임내의 매크로블록이 인트라 타입을 갖는다면 해당 매크로블록이 상대적으로 큰 왜곡 값을 갖는 경우라 볼 수 있다.

이러한 MPEG2에서의 매크로블록 타입과 움직임 예측오차와의 상관성은 MPEG2에서의 매크로블록 타입과 H.264에서의 매크로블록 모드와의 상관성으로 연결될 수 있으며, 이와 같은 상관성은 MPEG2에서 H.264 비디오 표준으로의 트랜스코딩에 효율적으로 사용될 수 있다.

다음 장에서 MPEG2에서의 움직임 벡터뿐만 아니라 앞서 기술한 상관성을 이용하여 매크로블록 타입 및 부호화블록패턴을 이용하는 저복잡도 트랜스코딩 알고리즘을 기술한다.

III. 제안하는 알고리즘

전술(前述)한 바와 같이 H.264 비디오 부호화 표준은 MPEG2 비디오 부호화 표준과 비교하여 많은 상이점을 가지고 있다. 특히 MPEG2의 8x8 DCT 변환에 비교되는 H.264의 4x4 정수 DCT 변환의 사용은 변환 영역에서의 트랜스코딩에 많은 어려움을 가져온다. 또한 다양한 블록크기로 움직임을 추정하는 H.264의 특징은 기존의 변환영역에서의 트랜스코딩에서 많이 사용되어 왔던 MC-DCT의 사용마저 제한하고 있다. 또한 MPEG2의 복호기부와 H.264 부호기부를 포함하고 있는 전체 트랜스코더의 복잡도중에서 MPEG2의 복호기부는 H.264의 부호기부와 비교하여 상당히 낮으며 H.264 부호기부 중에서도 4x4 DCT 변환은 움직임 추정 및 매크로블록 모드선택과 비교하여 상대적으로 매우 낮은 복잡도만을 갖는다.

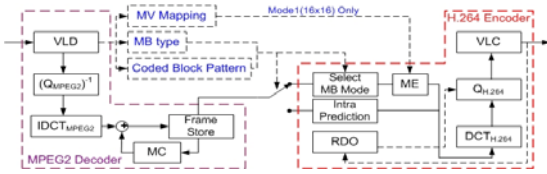


그림 5. 제안하는 트랜스코더의 구조
Fig. 5. Proposed transcoder structure

따라서 제안하는 MPEG2에서 H.264로의 트랜스코딩 기법은 변환영역이 아닌 공간영역에서의 트랜스코딩 즉, MPEG2 전복호화 구조를 갖는다. 그림 5는 제안하는 MPEG2에서 H.264 비디오 부호화 표준으로의 트랜스코딩을 위해 제안하는 트랜스코더의 구조를 도시한다.

그림 5에서 보는바와 같이 제안하는 트랜스코더는 MPEG2의 복호화 과정에서 사용되는 매크로블록 타입 및 부호화블록패턴을 이용하여 H.264 부호화 과정에서의 공간 내 예측(Intra) 및 시방향 예측(Inter)의 선택과 함께 시방향 예측 시 H.264에서의 매크로블록 모드를 적응적으로 선택한다. 또한 적응적 모드 선택 이후 16x16 모드가 선택된 경우에는 정수화소 위치의 MPEG2에서의 움직임 벡터 정보를 이용하며 이 경우 부-화소(Sub-pixel) 움직임을 추정만을 수행한다.

3.1 MPEG2 매크로블록 타입 및 부호화 블록 패턴 정보의 사용

MPEG2 스트림내의 매크로블록 타입 및 부호화 블록패턴의 정보는 해당 매크로블록의 특징을 표현하고 있다. 가령 P프레임에서 한 매크로블록의 타입이 스킵모드라면 이는 해당 매크로블록이 배경 또는 움직임이 없는 객체를 나타내고 있음을 유추할 수 있다. 또한 하나의 매크로블록에 포함되어 있는 네 개의 블록 각각에 대한 부호화 여부를 나타내는 부호화블록패턴은 움직임 추정과정후의 각 블록에 대한 예측오류를 나타내고 있다. 부호화된 블록이 많을수록 16x16 크기의 움직임 추정이 부적합하다고 유추할 수 있으며 반대로 부호화된 블록이 적을수록 16x16 크기의 움직임 추정이 유리하다고 할 수 있다.

H.264 부호화 과정에서도 MPEG2와 동일하게 16x16 크기 매크로블록 기반으로 부호화 과정을 수행하기 때문에 각 매크로블록의 특징을 표현하고 있는 이 두 가지의 정보는 H.264 부호화 과정에 포함되는 매크로블록 모드 선택과정에 효율적으로 사용될 수 있다.

제안하는 알고리즘은 각 I, P, B 프레임 내의 매크로블록 타입을 다섯 가지로 분류하고 그에 따라 H.264 부호화과정 중에 선택 가능한 매크로블록 모드를 적응적으로 선택한다. 이 중 (0,0)가 아닌 움직임 벡터를 가지며 매크로블록내의 하나이상의 블록이 부호화 되는 움직임추정 부호화 매크로블록 타입의 경우에는 부호화블록패턴에 따라 다시 세 가지의 서로 다른 선택 가능한 매크로블록 모드 집합을 갖는다.

3.1.1 스킵 매크로블록 (Skipped Macroblock)

MPEG2의 스킵 매크로블록은 P프레임 또는 B프레임 내에서 선택될 수 있는 매크로블록 타입이며 주로 화면내의 배경이나 움직임이 없는 객체를 포함하는 매크로블록에서 선택된다. 이러한 스킵 매크로블록 타입은 매크로블록 내의 모든 블록이 부호화되지 않았음을 나타내며 이는 H.264 부호화 표준에서의 울-왜곡 관점에서 매우 적은 왜곡 값을 가지는 경우에 해당되기 때문에 H.264 부호화 과정에서도 스킵모드로 선택될 가능성이 크다고 볼 수 있다. 따라서 MPEG2의 부호화과정에서 P 프레임 또는 B 프레임에서의 스킵 매크로블록으로 선택된 매크로블록은 H.264내의 매크로블록 모드들 중 SKIP 매크로블록 모드만을 고려하여 선택한다.

3.1.2 움직임 비추정 인트라 부호화 매크로블록 (No-MC Intra Coded Macroblock)

MPEG2의 움직임 비추정 인트라 매크로블록은 모든 프레임 구조 (I, P, B 프레임)에서 선택가능한 매크로블록 타입이며, P프레임 또는 B프레임에서의 이러한 매크로블록 타입은 움직임 추정 이후의 예측오차에 대한 분산값이 일정값(9*256) 이상이고 원 영상의 분산값보다 클 때 이전프레임으로부터의 예측을 사용하지 않고 인트라로 부호화 하는 경우를 나타낸다. 이와 같은 매크로블록 타입은 주로 이전의 프레임에서는 없던 새로운 객체 및 배경을 포함하는 매크로블록에서 주로 선택된다. 따라서 이와 같은 타입을 갖는 매크로블록의 경우에 대해서는 H.264 부호화 과정중의 매크로블록 모드를 인트라 4x4 및 인트라 16x16만을 고려하여 모드를 선택한다.

3.1.3 움직임 비추정 인트라 부호화 매크로블록 (No-MC Intra Coded Macroblock)

MPEG2의 움직임 비추정 인트라 부호화 매크로블록은 P프레임에만 선택 가능한 매크로블록 타입이며 이는 움직임 추정이후의 예측오차의 분산값이 움직

임 벡터가 (0,0)일 때 예측오차의 분산값의 0.8배 이상인 경우 움직임 벡터 (0,0) 위치에서의 예측오차를 부호화 한 경우를 나타내며 이는 주로 아주 미세한 움직임만을 갖는 객체가 매크로블록에 포함된 경우에 해당한다. 따라서 이와 같은 매크로블록은 16x16크기의 움직임 추정 시 (0,0) 위치에서도 어느 정도 작은 예측오차만을 가지는 경우라 생각될 수 있고 이와 같은 경우 H.264의 매크로블록 모드 선택 시 SKIP과 16x16 모드만을 고려하여 선택한다.

3.1.4 움직임 추정 비부호화 매크로블록 (MC Not-Coded Macroblock)

MPEG2의 움직임 추정 비부호화 매크로블록은 P 프레임과 B프레임에서 선택 가능한 매크로블록 타입으로 이는 움직임 추정이후 추정된 움직임 벡터 위치에서의 예측 오차가 양자화 과정 이후 모두 0 값을 가짐으로서 매크로블록내의 모든 블록이 부호화 되지 않은 경우를 나타낸다. 이는 해당 매크로블록을 16x16 블록크기의 움직임 추정을 통해서 참조 프레임 내에서 정확히 찾은 경우에 해당되며 이와 같은 타입을 갖는 매크로블록의 경우 움직임 비추정 인터부호화 매크로블록의 경우와 마찬가지로 H.264의 매크로블록 모드 선택 시 SKIP과 16x16 모드만을 고려한다.

3.1.5 움직임 추정 부호화 매크로블록 (MC coded Macroblock)

MPEG2에서의 움직임 추정 부호화 매크로블록의 경우 일반적으로 가장 많은 부분을 차지하는 매크로블록 타입으로 P 프레임 및 B 프레임에서 선택될 수 있다. 이는 움직임 추정이후 추정된 움직임 벡터 위치에서의 예측오차를 부호화 한 경우이며, 매크로블록이 이러한 타입으로 선택된 경우 부호화블록패턴을 사용하여 H.264에서의 매크로블록 모드를 적응적으로 선택하는데 만일 매크로블록내의 두개이하의 블록이 부호화되었다면 이는 16x16 크기로 움직임 추정 시 적어도 16x8 크기는 정확한 움직임 추정이 이루어졌다고 볼 수 있다. 따라서 이와 같은 경우 H.264에서의 매크로블록 모드는 SKIP, 16x16, 16x8, 8x16 중에서 선택한다. 만일 매크로블록내의 블록 중 3개의 블록이 부호화 되었다면 하나의 블록은 정확한 움직임 추정이 이루어졌다고 볼 수 있으며 이때에는 SKIP, 16x16, 16x8, 8x16, 8x8 중에서 H.264의 매크로블록 모드를 선택하며, 8x8 블록의 부-매크로블록 모드는 고려하지 않는다. 마치

표 2. MPEG2 매크로블록타입 및 부호화블록패턴에 따른 선택 가능한 H.264 매크로블록 모드 집합
Table 2. The selectable macroblock mode sets according to MPEG2 macroblock types

MPEG2 MB type	H.264 MB mode set
SKIP	SKIP
No-MC Intra Coded	Intra 4x4, Intra 16x16
No-MC Inter Coded	SKIP, 16x16
MC Not-Coded	SKIP, 16x16
MC Coded	
one/two coded block	SKIP, 16x16, 16x8, 8x16
three coded block	SKIP, 16x16, 16x8, 8x16, 8x8
four coded block	SKIP, 16x16, 16x8, 8x16, 8x8, 8x4, 4x8, 4x4

막으로 매크로블록내의 4개의 블록이 모두 부호화 되었다면 모든 8x8크기의 블록에 대해서도 예측오차가 크다는 것을 나타내며 따라서 이와 같은 경우에는 SKIP, 16x16, 16x8, 8x16, 8x8 뿐만 아니라 부-매크로블록 모드인 8x4, 4x8, 4x4 모드 즉, 모든 시방향 예측을 사용하는 매크로블록 모드를 고려하여 선택한다.

표 2는 앞서 기술한 MPEG2의 복호화 과정 중에 사용되는 매크로블록타입 및 부호화 블록패턴에 따른 H.264에서의 선택 가능한 매크로블록 모드 집합을 나타낸다.

3.2 움직임 벡터 정보의 사용

MPEG2로 부호화된 압축 비트스트림 내에 포함된 움직임 벡터 정보는 16x16 블록크기로 추정된 벡터 정보를 나타내며 1/2 화소단위의 정확도를 가지고 있다. 반면 H.264에서의 움직임 벡터 정보는 4x4크기의 블록에서부터 16x16까지의 다양한 크기에 따라 움직임 벡터를 추정하며 1/4 화소단위의 정확도를 요구한다. 따라서 MPEG2의 움직임 벡터 정보는 H.264에서의 16x16 매크로블록 모드에서의 움직임 추정 시에만 사용 가능하며 정수화소의 움직임벡터를 받아 1/2 및 1/4 화소단위의 움직임은 재 추정한다.

이때 MPEG2의 움직임벡터가 1/2 화소 단위의 움직임 벡터 정확도를 가짐에도 불구하고 정수단위의 움직임 벡터 정보만을 사용하는 이유는 MPEG2와 H.264에서의 1/2화소 위치의 보간기법에 서로 다른 필터를 사용하기 때문이다. MPEG2의 경우 단순히 2탭 쌍선형필터(bilinear filter)를 사용하는 반면 H.264에서는 이후 1/4 화소 위치를 비교적 정확히 보간하기 위해 6탭 필터를 사용하기 때문이다. 따라서 제안하는 트랜스코더는 비록 MPEG2의 움

직입벡터정보가 1/2 화소 위치의 정보까지 포함하지만 사용되는 정보는 정수화소단위의 움직임 벡터정보만을 사용하며 16x16 매크로블록 모드 선택 시에만 한정적으로 사용한다.

IV. 실험 및 결과

실험은 MPEG2 메인 프로파일의 메인 레벨을 사용하여 500Kbits/sec의 비트율로 부호화된 MPEG2 비트스트림을 입력으로 사용하였으며 트랜스코딩내의 H.264 부호기는 베이스라인 프로파일의 3.0 레벨을 사용하여 부호화하였고 다양한 비트율로의 트랜스코딩을 고려하기 위하여 다섯 가지(100~500 Kbits/sec)의 서로 다른 비트율로 실험하였다. 또한 실험에서는 다양한 움직임 크기를 고려하기 위하여 비교적 큰 움직임을 포함하는 Coastguard, Foreman, Stefan, Table Tennis 시퀀스의 처음 100 프레임과 상대적으로 적은 움직임만을 포함하는 Hall Monitor, Mobile 시퀀스의 처음 100 프레임을 테스트 시퀀스로서 사용하였다.

본 논문에서 제안하는 MPEG2에서 H.264 비디오 표준으로 트랜스코딩 기법은 P 프레임 및 B 프레임 모두에 적용가능하나 H.264의 베이스라인 프로파일로의 트랜스코딩을 위하여 입력으로 사용된 MPEG2 압축 스트림에서도 P 프레임만을 고려하였으며, 따라서 MPEG2 코덱 내의 M(I 프레임과 P 프레임 사이의 거리)은 1값을 사용하였다. 표 3은 제안한 트랜스코더의 성능평가를 위해 사용된 실험환경을 나타낸다.

표 3. 실험 환경
Table 3. Experimental Condition

	MPEG2	H.264
CPU	Intel P4 2.8GHz	
Memory	512 MHz	
Operating System	Windows XP Pro.	
CODEC	TM 1.1	JM 9.3
Profile	Main	Baseline
Level	main	3.0
Structure	IPPP	IPPP
GOP	10	10
Bitrates	500Kbits/sec	100Kbits/sec
		200Kbits/sec
		300Kbits/sec
		400Kbits/sec
		500Kbits/sec
Frame rates	30frames/sec	30frames/sec
RDO		used

프레임율은 MPEG2 및 H.264에서 모두 30 frames/sec를 사용하였으며 H.264 부호화부에서는 울-왜곡 최적화 기법을 사용하였고, MPEG2에서 H.264로의 전복호화 후 전부호화의 경우와 제안하는 트랜스코더를 이용한 경우에 대해서 수행하였다.

그림 6은 Foreman 시퀀스에 대해서 트랜스코더의 출력 비트스트림의 비트율이 300Kbits/sec일 때의 PSNR 값을 프레임별로 비교한 결과이다.

그림 6의 결과에서 보는바와 같이 전복호화 후 전 부호화의 결과와 제안하는 알고리즘 결과의 차이는 상당히 미세한 것으로 나타난다. 프레임에 따라 오히려 제안하는 알고리즘이 전복호화 후 전부호화 방법에 비하여 높게 나타나는 것은 비트율 제어기법의 사용으로 인해 프레임에 따라 보다 작은 양자화 계수를 사용할 수 있기 때문이다.

표 4는 실험에 사용된 다섯 가지의 트랜스코더 출력 비트율에 대해서 각 시퀀스 100프레임에서의 평균 PSNR을 나타낸다. 결과에서 볼 수 있는바와 같이 제안하는 기법은 전복호화 후 전부호화에 비하여 실험에 사용된 모든 비트율에서 0.1dB 미만의 차이를 나타내며, 평균 0.03 dB의 매우 작은 차이만을 보인다.

그림 7는 각 시퀀스에 대해서 전복호화 후 전부호화의 방법과 제안하는 트랜스코딩 기법의 전체 트랜스코딩 시간을 나타낸다.

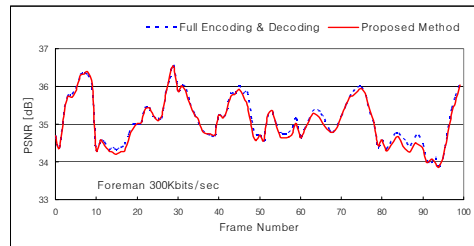


그림 6. Foreman 시퀀스에 대한 프레임별 PSNR 비교
Fig. 6. PSNR comparison of Foreman sequence according to frame number

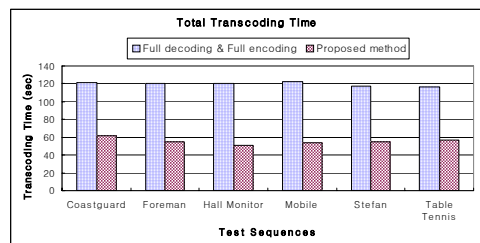


그림 7. 전체 트랜스코딩 시간의 비교
Fig. 7. The comparison of total transcoding time

표 4. 전복호화 후 전부호화 방법과 제안하는 알고리즘의 실험 결과
Table 4. Experimental results of full decoding and full encoding vs. proposed algorithm

Test Sequences	Comparison	Bitrate Ratios (H.264 Kbits/sec/ MPEG2 Kbits/sec)				
		100/500	200/500	300/500	400/500	500/500
Coastguard	Full De.&En.	29.12	30.72	31.49	31.97	32.26
	Proposed	29.12	30.70	31.48	31.95	32.25
	Difference	0.00	0.02	0.01	0.02	0.01
Foreman	Full De.&En.	31.99	34.34	35.11	35.44	35.63
	Proposed	31.97	34.29	35.06	35.40	35.60
	Difference	0.02	0.05	0.05	0.04	0.03
Hall Monitor	Full De.&En.	34.18	36.10	36.68	36.97	37.10
	Proposed	34.09	36.06	36.66	36.95	37.08
	Difference	0.09	0.04	0.02	0.02	0.02
Mobile	Full De.&En.	22.52	24.28	25.10	25.47	25.65
	Proposed	22.50	24.28	25.10	25.45	25.64
	Difference	0.02	0.00	0.00	0.02	0.01
Stefan	Full De.&En.	23.92	26.83	27.93	28.42	28.72
	Proposed	23.87	26.80	27.87	28.38	28.70
	Difference	0.05	0.03	0.06	0.04	0.02
Table Tennis	Full De.&En.	30.82	32.51	33.19	33.56	33.78
	Proposed	30.77	32.45	33.16	33.52	33.74
	Difference	0.05	0.06	0.03	0.04	0.04

그림 7에서 보는 바와 같이 제안하는 트랜스코더의 사용 시 전체 트랜스코딩 시간은 최대 61.9% 감소하였으며 평균 55.4% 감소하는 것으로 나타난다.

에서 H.264로의 실시간 트랜스코딩에 더욱 다가갈 수 있을 것으로 생각된다.

V. 결 론

제안한 MPEG2에서 H.264로의 트랜스코더는 공간영역에서의 움직임 재추정 즉, 공간영역에서의 트랜스코딩 기법이며, 트랜스코더의 구조는 MPEG2 복호화부와 H.264 부호화부로 나뉠 수 있다. MPEG2 복호화부는 입력된 MPEG2 압축 비트스트림의 전복호화를 수행하며 H.264 부호화부는 MPEG2 압축 비트스트림에 포함된 움직임 벡터 뿐 아니라 매크로블록 타입과 부호화블록패턴의 정보를 사용하여 매크로블록 모드를 적응적으로 선택함으로써 보다 낮은 복잡도의 트랜스코더를 구현케 한다.

제안하는 트랜스코딩 기법을 사용하는 경우 MPEG2 비트스트림의 전복호화 후 H.264로의 전부호화와 비교하여 단지 평균 0.03 dB의 화질열화가 발생하는 반면 전체 트랜스코딩 시간은 최대 61.9% 단축된 것으로 나타났다.

본 논문에서 제안하는 MPEG2에서 H.264로의 트랜스코딩 알고리즘은 MPEG2 비트스트림 내에 포함된 부가정보를 이용함으로써 최소한의 화질열화만을 갖는 저복잡도 트랜스코딩 기법이며, 이는 고속 움직임 추정기법등과 같이 사용될 경우 MPEG2

참 고 문 헌

- [1] A. Vetro, C. Christopoulos, and H. Sun, "Video transcoding architectures and Techniques: An Overview," IEEE Signal Processing Magazine, pp.18-29, Mar. 2003.
- [2] A. Eleftheriadis and D. Anastassiou, "Constrained and general dynamic rate shaping of compressed digital video," in Proc. IEEE Int. Conf. Image Processing, vol. 3, pp.396-399, 1995.
- [3] J. Youn, M. T. Sun, and C. W. Lin, "Motion vector refinement for high-performance transcoding," IEEE Trans. Multimedia, vol. 1, pp. 30-40, Mar. 1999.
- [4] R. Dugad and N. Ahuja, "A fast scheme for image size change in the compressed domain," IEEE Trans. Circuit Syst. Video Technol., vol. 11, pp.461-474, Apr. 2001.
- [5] J. Xin, M. T. Sun, and K. Chun, "Motion re-estimation for MPEG-2 to MPEG-4 simple profile transcoding," Packet Video Workshop, Pittsburgh, PA, 2002.

[6] J. Bialkowski, A. Kaup, and K. Illgner, "Fast transcoding of intra frames between H.263 and H.264," in Proc. IEEE Int. Conf. Image Processing, vol. 4, pp.2785-2788, Oct. 2004.

[7] J. Bialkowski, M. Menden, M. Barkowsky, K. Illgner, and A. Kaup, "A fast H.263 to H.264 inter-frame transcoder with motion vector refinement," Picture Coding Symposium 2004 (PCS 2004), Dec. 2004.

[8] H. Kalva, "Issues in H.264/MPEG-2 video transcoding," Consumer Communications and Networking Conference, pp.657-659, 2004.

[9] C. Chen, P. H. Wu, and H. Chen, "MPEG-2 to H.264 Transcoding," Picture Coding Symposium 2004, pp.15-17, 2004.

[10] ISO/IEC 13818-2, "Information technology Generic coding of moving pictures and associated audio information: Video (Second edition)," Dec. 2000.

[11] ISO/IEC 14496-10, "Version 3 of H.264/AVC," Jun. 2004.

[12] 정인철, 한중기, "H.264를 위한 효율적인 움직임 벡터 추정 알고리즘," 한국통신학회논문지, Vol 29, Dec. 2004.

김 동 형 (Donghyung Kim)

정회원



1999년 2월 충북대학교 전자공학과 졸업
 2001년 8월 충북대학교 전자공학과(석사)
 2002년 3월~현재 한양대학교 전자통신전공학과 박사과정
 <관심분야> 영상처리 및 영상압축

정 제 창 (Jechang Jeong)

정회원



1980년 2월 서울대학교 전자공학과 졸업
 1982년 2월 KAIST 전기전자공학과(석사)
 1990년 미국 미시간대학 전기공학과(공학박사)
 1980년~1986년 KBS 기술연구소 연구원 (디지털 TV 및 뉴미디어 연구)
 1990년~1991년 미국 미시간대학 전기공학과 연구교수(영상 및 신호처리 연구)
 1991년~1995년 삼성전자 멀티미디어 연구소(MPEG, HDTV, 멀티미디어 연구)
 1995년~현재 한양대학교 전자전기컴퓨터공학부 교수 (영상통신 및 신호처리 연구실)
 1998년 11월 27일 과학기술자상 수상
 1998년 12월 31일 정보통신부장관상 표창
 <관심분야> 영상처리 및 영상압축