

MPEG-2에서 H.264로 변환하기 위한 고속 모드 결정 기법

준회원 김 원 균*, 박 경 준*, 유 중 민*, 정회원 정 제 창*

Fast Mode Decision for MPEG-2 to H.264 Transcoding

Won kyun Kim*, Kyung jun Park*, Jong min You* *Associate Members*,
Je chang Jeong* *Regular Member*

요 약

본 논문에서는 MPEG-2로 부호화된 비디오 스트림에서 H.264 부호화 표준으로 변환하기 위한 저 복잡도 모드 결정 기법을 제안한다. 제안하는 MPEG-2에서 H.264로의 트랜스코더는 MPEG-2 복호화부와 H.264 부호화부로 구성되어 있으며, MPEG-2 복호화부에서 얻을 수 있는 현재 매크로블록의 정보를 이용해서 H.264에서 I프레임과 P프레임의 매크로블록 모드를 적응적으로 선택함으로써 전체 트랜스코더의 복잡도를 감소시킨다. 제안하는 트랜스코더를 시퀀스에 적용한 결과 주목할 만한 화질열화나 비트율의 증가 없이 30~60%정도 복잡도가 감소된 것으로 나타난다.

Key Words : Transcoding, MPEG-2, H.264, Mode Decision, Rate Distortion Optimization

ABSTRACT

In this paper, we present a efficient transcoding method from MPEG-2 to H.264. The proposed transcoder is the transcoding method for spatial domain which consists of MPEG-2 decoder part and H.264 encoding part. In transcoder, we can get useful information to estimate less probable modes from MPEG-2 decoder. Using this information, H.264 encoder chooses the macroblock mode of I-frame and P-frame adaptively to reduce the whole complexity of the transcoder. Our experimental result shows that the proposed algorithm can archive about 30~60% computational saving without significant degradation of visual quality and increasing of bit rate.

I. 서 론

트랜스코딩에 관한 정의는 명확히 구문으로 서술되지는 않았지만 통상 임의의 한 신호를 다른 형태의 신호로 변환시키는 부호화 기법을 일컫는다^[1]. 이러한 정의를 토대로 트랜스코딩의 분야는 현재 비디오 트랜스코딩에 초점을 맞추고 있다. 트랜스코딩을 입출력 비디오 스트림의 비디오 표준에 따라

분류하면 크게 동일한 비디오 표준 간의 트랜스코딩과 상이한 비디오 표준간의 트랜스코딩으로 나눌 수 있다. 동일한 비디오 표준간의 트랜스코딩은 일반적으로 비트율 감소를 주목적으로 하며 이에는 양자화 계수의 변환을 통하여 비트율을 감소시키는 SNR 트랜스코딩과 프레임율의 변환을 수행하는 프레임율 트랜스코딩, 그리고 프레임의 공간적 크기 변환을 수행하는 공간적 해상도 트랜스코딩으로 나

※ 본 연구는 서울시 산학연협력사업으로 구축된 서울 미래형콘텐츠컨버전스 클러스터 지원으로 수행되었습니다.

* 한양대학교 전자통신컴퓨터공학과 영상통신 및 신호처리 연구실 (wonkyun@ece.hanyang.ac.kr)

논문번호 : KICS2006-11-508, 접수일자 : 2006년 11월 24일, 최종논문접수일자 : 2007년 3월 8일

낼 수 있다. 서로 상이한 비디오 표준간의 트랜스코딩은 서로 다른 영상 표준을 지원하는 단말기간의 전송 등과 같이 주로 영상 포맷 자체의 변환을 목적으로 하며, MPEG-2에서 MPEG-4 Part2로의 트랜스코딩 등이 이에 속한다²⁾.

비디오 재생 기능을 가진 PDA(Personal Digital Assistant), 노트북, 스마트폰의 보급률이 증가하고, 높은 압축률을 자랑하는 H.264를 표준으로 채택한 휴대용 이동 수신 TV 단말기의 보급이 늘어나면서 이러한 트랜스코딩의 중요성은 더욱 증대되고 있다. 하지만 부호화의 요소 기술 측면에서 이전의 비디오 영상표준들(MPEG-1, MPEG-2, MPEG-4 Part2 등)과 많은 상이점을 보이는 H.264 비디오 표준은 높은 부호화 효율만큼이나 높은 복잡도 때문에 많은 문제점을 갖는다. 특히 MPEG-2 비디오 표준과 H.264비디오 표준은 요소 기술면에서 상당한 차이점을 가진다. 예를 들어 공간영역에서 수행되는 인트라 예측과정, 루프필터, 4×4정수변환, 다양한 크기의 화면간 예측 모드, 그리고 화면내 예측 모드는 MPEG-2와 비교되는 H.264의 부호화 특징들이다³⁾. 하지만 MPEG-2에서는 상대적으로 16×16 크기의 매크로블록 단위로만 움직임 추정을 실행하기 때문에 MPEG-2부호화부에 비해 H.264부호화부의 복잡도가 매우 높다. 특히 H.264에서 화면간, 화면내 예측 모드를 위한 매크로블록 모드 선택은 많은 계산과정을 필요로 한다⁴⁾. 화면내 예측 모드에서 휘도(luminance) 블록들에 대한 예측 부호화는 4×4 서브 블록이나 16×16 매크로블록 단위로 수행되고, 방향성에 따라 각각 9가지, 4가지의 예측모드가 존재한다. 또한 4:2:0 색차 포맷의 영상에 대해 하나의 매크로블록은 두 개의 8×8 크기의 색차(chrominance)블록이 존재하는데 이에 대해서도 4가지 방향성에 따른 예측 부호화를 수행한다. 단, 색차블록에 대해서는 두 블록에 대해 같은 예측 모드를 적용한다. 화면간 예측 모드에서는 휘도성분에 대해서 영상의 국부특성을 효율적으로 반영하기 위하여 16×16부터 4×4까지의 다양한 블록 크기의 움직임 보상을 지원한다. 각 매크로블록(16×16 화소)의 휘도성분은 16×16, 16×8, 8×16, 8×8블록의 4가지 형태로 나누어진다. 8×8블록모드가 선택되면, 매크로블록 내의 4개의 8×8블록은 그림3-6과 같이 8×8, 8×4, 4×8, 4×4의 4가지 형태로 다시 나누어진다. P 프레임 내의 매크로블록 모드들 중에서 인트라 예측을 위한 모드인 SKIP, 16×16, 16×8, 8×16, 8×8 모드를 각각 SKIP, Mode1, Mode2, Mode3,

Mode8로 정의하고, Mode8 내의 선택 가능한 부매크로블록 모드인 8×8, 8×4, 4×8, 4×4 모드를 각각 Mode4, Mode5, Mode6, Mode7 이라고 정의한다. H.264에서는 이러한 다양한 예측 모드 중에 화질과 비트율 측면에서 가장 적합한 모드를 선택하기 위해서 RDO(Rate Distortion Optimization)기법을 사용한다.

하지만 RDO기법은 많은 계산량을 필요로 하기 때문에 H.264부호화부의 복잡도 문제를 해결하기 위해 많은 연구들이 이루어져 있다. Lu등은 MPEG-2에서 H.264로의 트랜스코더에서 고속 모드 결정 기법을 제안하였다⁵⁾. 인트라16×16에 대해서 예측 모드 선택을 위한 과정을 수행하고 그것으로부터 인트라4×4모드의 수행여부를 결정함으로써 복잡도를 줄였다. 또한 Petljanski와 Kalva는 MPEG-2에서 H.264로의 트랜스코더에서 인트라 모드 결정 기법에 대해서 소개하였다⁶⁾. 이는 MPEG-2 부호화부의 I프레임에서 현재 매크로블록의 분산값을 구하고 그에 따라서 H.264부호화부에서 인트라16×16을 수행할지 인트라4×4를 수행할지의 여부를 결정한다. 이러한 Petljanski의 트랜스코딩 기법은 화질이나 복잡도 측면에서는 좋은 성능을 보이지만 비트율이 많이 증가하는 단점이 있다.

따라서 본 논문에서는 낮은 복잡도를 가지면서 우수한 성능을 나타내는 MPEG-2에서 H.264로의 트랜스코딩 기법을 제안한다. 제안하는 기법은 MPEG-2부호화부에서 현재 매크로블록의 정보를 이용해서 H.264부호화부에서 매크로블록 모드를 적절적으로 선택함으로써 모드 선택 과정의 복잡도를 줄일 수 있다.

본 논문의 구성은 다음과 같다. 2장에서는 참조 소프트웨어에서의 매크로블록 모드 선택 기법에 대해서 설명하고 3장에서는 제안하는 방법을 설명한다. 4장에서는 실험을 통하여 제안하는 트랜스코더의 성능을 평가하고, 마지막으로 4장에서 결론을 맺는다.

II. H.264에서 RDO(Rate Distortion Optimization)

참조 소프트웨어는 매크로블록의 모드를 선택하기 위해 RD 비용함수(Rate Distortion cost function)를 사용한다. RDO기법은 가능한 모드들 중에 비트율을 제한했을 때 가장 작은 왜곡을 갖는 모드를 선택하기 위한 방법이다.

$$J = D + \lambda R \quad (1)$$

모드결정을 위한 RDO기법은 식(1)에 의해서 결정되며 J값을 최소로 하는 모드가 최적의 모드로 선택된다⁷⁾. 식(1)에서 D는 왜곡 값을 나타내며, R은 그 모드에 해당하는 압축된 비트열의 길이를 나타낸다. 그리고 λ 는 양자화 계수와 각 모드에 따른 계수값(lagrange multiplier)을 나타낸다. D는 SSD(Sum of Squared Differences)나 SATD(Sum of Absolute Transformed Difference)를 사용한다. SSD는 원영상과 부호영상의 차분치의 합을 나타내며 식(2)와 같다.

$$SSD = \sum_{x=1}^{16} \sum_{y=1}^{16} (s[x,y] - c[x,y])^2 \quad (2)$$

$s[x, y]$ 는 원영상을 $c[x, y]$ 는 부호영상을 가리킨다. SATD는 원영상과 예측영상의 차분치를 하다마드 변환한 결과의 합을 나타내며 식(3)과 같다.

$$SATD = \sum_{x=1}^{16} \sum_{y=1}^{16} (H \cdot (|s[x,y] - p[x,y]|) \cdot H) \quad (3)$$

여기서 $p[x, y]$ 는 예측영상을 가리킨다. λ 는 양자화 계수(QP)에 따른 함수으로써 $\partial D / \partial R$ 로 구할 수 있으며 각 프레임, 모드, 움직임 벡터에 따라서 식(4)와 같이 나타낼 수 있다.

$$\begin{aligned} \lambda &= f(QP) \\ \lambda_{MODE,IP} &\approx 0.85 \times 2^{QP/3} \\ \lambda_{MODE,B} &\approx 4 \times 0.85 \times 2^{QP/3} \\ \lambda_{MOTION,P} &\approx \sqrt{0.85 \times 2^{QP/3}} \\ \lambda_{MOTION,B} &\approx \sqrt{4 \times 0.85 \times 2^{QP/3}} \end{aligned} \quad (4)$$

λ_{MODE} 는 모드를 결정하기 위해서 사용되는 계수이고 λ_{MOTION} 은 움직임 벡터를 결정하기 위한 계수이다. 이 값들을 이용해서 참조소프트웨어에서는 다음과 같은 순서로 매크로블록 모드를 결정한다.

단계 1. 양자화 파라미터에 따라 λ 값을 결정

단계 2. Mode1, Mode2, Mode3에 대한 참조 프레임 및 움직임 벡터 결정

$$J = SATD(MV, REF) + \lambda_{MOTION} R(MV, REF) \quad (5)$$

단계 3. 부 매크로블록에 대한 모드 선택

단계 3.1 첫 번째 부-매크로블록에 대한 참조 프레임 및 움직임 벡터 결정

$$J = SSD(s,c, MV, REF) + \lambda_{MOTION} R(MV, REF) \quad (6)$$

단계 3.2 첫 번째 부-매크로블록의 모드 선택

$$J = SSD(s,c, MODE) + \lambda_{MODE} R(s,c, MODE) \quad (7)$$

단계 3.3 잔여 부-매크로블록에 대해서 단계 3.1 ~ 단계 3.2 반복

단계 4. 인트라4x4 모드 선택

$$J = SSD(s,c, MODE) + \lambda_{MODE} R(s,c, MODE) \quad (8)$$

단계 5. 인트라16x16 모드 선택

$$J = SATD(MODE) \quad (9)$$

단계 6. 매크로블록의 모드 결정, 식(10)에서 가장 작은 J값을 갖는 모드를 선택

$$J = SSD(s,c, MODE) + \lambda_{MODE} R(s,c, MODE) \quad (10)$$

이와 같이 참조 소프트웨어에서 RDO기법을 사용할 경우에 모든 가능한 모드에 대해서 비용값을 계산해야 되기 때문에 많은 계산량을 필요로 한다.

III. 제안하는 알고리즘

3.1. I 프레임에서의 화면내 예측 모드 선택

기존의 동영상 부호화 표준에서는 인트라 화면을 부호화할 때, 변환영역에서 일부 계수(DC 또는 일부의 AC 계수)에 대해서 예측 부호화를 수행하였다. 그러나, 일반적으로 변환영역에서의 계수보다 공간영역에서의 화소간 상관도가 높으므로 이를 적절한 방법에 의해 제거할 수 있다면 부호화 효율을 크게 증가시킬 수 있다. 이를 위하여 H.264에서는 공간영역에서의 국부적 특징을 반영하여 블록의 방향성을 고려한 화면내 예측 부호화를 수행한다. 휘도성분에 대해서는 16×16매크로블록 단위의 4가지 모드와 4×4 서브블록 단위의 9가지 모드가 있다. 화면내 예측 모드의 경우에는 현재 프레임에서의 블록의 방향성을 고려하기 때문에 현재 매크로블록의 복잡한 정도에 따라서 선택되는 모드의 블록 크기가 좌우된다. 예를 들어, 부호화하고자 하는 매크로블록이 복잡한 경우에는 16×16매크로블록 단위로 부호화하게 되면 예측값이 적절하지 못하기 때문에 오차가 커지고 비트율이 증가하게 될 것이다. 따라서 복잡한 영상의 경우에는 4×4서브블록 단위 모드가 선택될 확률이 높다. 또한 평탄한 영상의 경우에는 예측값이 16×16매크로블록 단위의 모든 부분에 적용되더라도 오차가 적기 때문에 16×16매크로블록 단위 모드가 선택될 것이다. 이러한 특징을 이용해서 H.264 부호화부에서 현재 매크로블록의 모드 선택 과정을 수행하기 전에 매크로블록의 복잡한 정도를 알 수 있다면 매크로블록의 모드 크기를 예측함으로써 모드 선택 과정의 계산량을 줄일 수 있을 것이다. MPEG-2의 I프레임에서는 ME/MC(Motion Estimation/Motion Compensation)을 수행하지 않기 때문에 복호화 과정에서 매크로블록의 AC계수로부터 현재 매크로블록의 복잡한 정도를 예측할 수 있다⁸⁾.

제안하는 I프레임에서의 예측 모드 선택은 식(11)과 같이 MPEG-2복호화부에서 매크로블록의 AC계수의 절대값의 합을 $activity_I$ 로 정의하고, 이 값에 따라 적응적으로 모드를 선택하도록 한다.

$$activity_I = \sum_{i=1}^7 \sum_{j=1}^7 |c(i,j)| \quad (11)$$

임계값은 TH_{I1} 과 TH_{I2} 를 사용해서 TH_{I1} 보다 작은 매크로블록은 16×16모드에 대해서만 수행하고, TH_{I1} 과 TH_{I2} 사이의 값을 갖는 부분은 정확하게 구분되지 않는 부분이기 때문에 16×16과 4×4 두 가지 모두 수행하도록 한다. 마지막으로 TH_{I2} 보다 큰 매크로블록에 대해서는 4×4에 대해서만 모드 선택 과정을 수행한다.

3.2. P 프레임에서의 화면간 예측 모드 선택

MPEG-2에서 인트라 매크로블록은 가장 많은 부분을 차지하는 매크로블록 타입으로 P프레임과 B프레임에서 선택될 수 있다. 이는 움직임 추정 이후 추정된 움직임 벡터위치에서의 예측오차를 부호화 한 경우이다.

그림 1은 H.264부호화 과정에서 foreman시퀀스에서 P프레임에서의 RDO기법에 의해 선택된 모드를 나타낸다. 그림1에서 보는 바와 같이 평탄하거나 움직임이 적은 부분에서는 큰 블록 크기의 모드가 선택되고, 복잡하고 움직임이 많은 부분에서는 상대적으로 작은 블록 크기의 모드가 선택된다. 대부분의 영상에서는 주로 주 객체만이 움직임이 있고 그 외의 배경부분은 상대적으로 움직임이 적거나 평탄하기 때문에 대부분의 매크로블록이 16×16 크기의 모드가 선택된다. 하지만 RDO기법에 의해서 모드 선택을 수행하게 되면 모든 블록 크기의 RD 비용 함수를 계산해야 하기 때문에 상당히 많은 계산량을 낭비하게 된다⁹⁾.

따라서 현재 매크로블록 모드 선택 과정을 수행하기 전에 선택될 가능성이 낮은 모드를 예측할 수 있다면, 전체 부호기부 복잡도의 대부분을 차지하는 모드 선택 과정의 계산량을 줄임으로써 전체 트랜스코더의 복잡도를 낮출 수 있다.

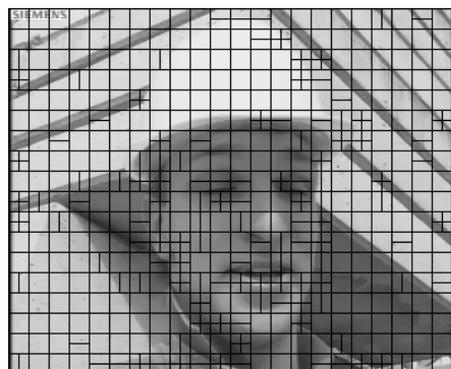


그림 1. RDO기법에 의해 선택된 매크로블록 모드

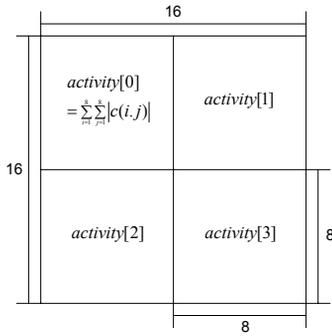


그림 2. MPEG-2 매크로블록의 activity.

MPEG-2 스트림 내의 매크로블록에서 움직임 추정 이후 예측오차는 해당 매크로블록의 특징을 잘 표현하고 있다. 움직임 추정 이후 예측오차가 충분히 작다면 H.264부호화에서도 16×16블록크기의 모드로 선택될 수 있을 것이다. 하지만 예측오차 값이 크거나 인트라로 선택된 매크로블록의 경우에는 16×16블록크기로 움직임 추정하는 것이 적합하지 않을 것이다. 이때는 이전의 프레임에서는 없던 새로운 객체 및 배경을 포함하는 매크로블록이거나 MPEG-2에서는 16×16 크기의 매크로블록만 지원하기 때문에 매크로블록의 일부분이 예측 오차가 작더라도 나머지 부분이 커지기 때문에 H.264에서도 인트라 모드가거나 더 작은 블록 크기의 모드가 선택될 확률이 높다.

제안하는 인터 모드 결정 방법은 MPEG-2에서의 이러한 정보를 사용하여 H.264부호화 과정 중에 선택 가능한 매크로블록 모드의 크기를 미리 예측함으로써 트랜스코더의 복잡도를 낮춘다. MPEG-2에서의 16×16 매크로블록을 4개의 8×8블록으로 나누고 그림2와 같이 각 블록의 움직임 추정 이후 예측오차의 합을 구한다. 이 예측오차의 합을 activity_P로 정의하고 식(12)과 같이 나타냈다.

$$activity_P = \sum_{i=0}^7 \sum_{j=0}^7 |c(i,j)| \quad (12)$$

여기서 c(i, j)는 MPEG-2에서 움직임 추정 이후 예측 오차의 양자화된 DCT(Discrete Cosine Transform)계수를 말한다. 이 activity_P값을 통해서 각 8×8블록의 예측 오차를 구할 수 있고, 나아가서 16×16크기의 블록으로 움직임 추정이 적합한 지를 판단할 수 있다.

통계적인 분석을 통해서 RDO기법을 사용하는 트랜스코더에서 activity_P값에 따라 H.264부호화 과정 중에 어떠한 모드가 선택되는지를 보고 다음과 같은 4가지 조건을 얻을 수 있었다.

- 1) 4개의 8×8블록의 activity_P가 모두 200이하 일 경우 16×16크기의 블록이 선택될 확률이 높다.
- 2) 4개의 8×8블록의 activity_P중 가장 큰 값과 가장 작은 값의 차가 작을수록 16×16크기의 블록이 선택될 확률이 높다.
- 3) activity_P가 40보다 작은 값을 갖는 블록이 많을수록 16×16크기의 블록이 선택될 확률이 높다.
- 4) MPEG-2에서의 움직임 벡터와 H.264 인터16×16의 PMV와의 차이가 작을수록 16×16크기의 블록이 선택될 확률이 높다.

앞의 조건으로부터 H.264에서 선택되는 매크로블록 모드의 크기를 미리 예측할 수 있도록 RD 비용값과 유사한 PRDcost(Predicted Rate-Distortion cost)를 다음과 같이 정의한다.

$$PRDcost = \frac{(\max_act - \min_act)}{2^{d-1}} + |PMV - MV| \quad (13)$$

식 (13)에서 max_act는 4개의 8×8블록 중에 가장 큰 activity_P값을, min_act는 가장 작은 activity_P값을 나타내고, d는 4개의 8×8블록 중에 40보다 작은 activity_P값을 갖는 블록의 개수를 가리킨다. MV는 MPEG-2에서의 움직임 벡터를 PMV는 H.264에서 인터16×16의 예측 움직임 벡터를 나타낸다. 식 (13)에서 |PMV-MV|부분은 MPEG-2에서의 움직임 벡터와 H.264의 인터16×16 예측 움직임 벡터의 차이가 작을수록 MPEG-2에서의 움직임 추정 후 예측 오차로 구한 비용값이 H.264에서 모드 선택을 위해 사용하는데 더욱 적합하다는 것을 반영하기 위한 것이다.

각 모드들을 뚜렷하게 구분할 수는 없지만 SKIP 모드(모드0)와 인터 16×16(모드1)일수록 PRDcost가 가장 작은 값을 갖는 것을 알 수 있었다. 따라서 PRDcost가 임계값 이하일 경우에는 SKIP모드와 인터 16×16만을 수행함으로써 복잡도를 낮출 수 있다.

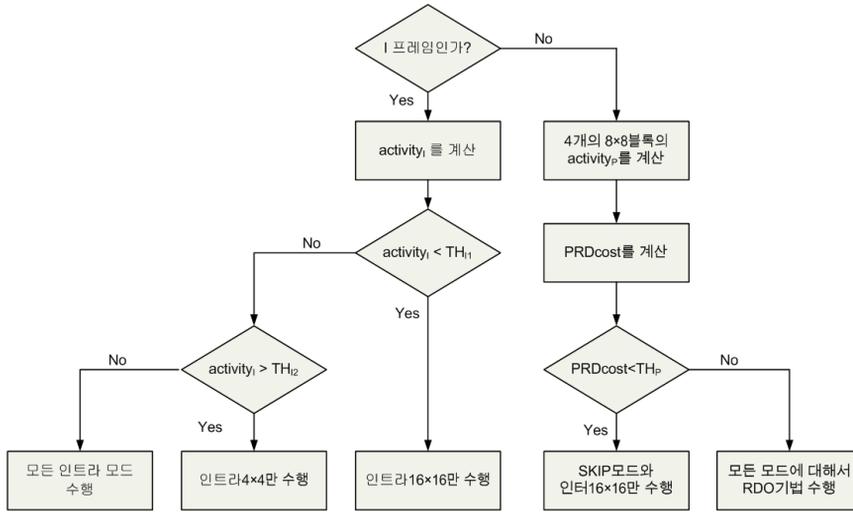


그림 3. 제안하는 알고리즘 순서도.

3.3. 전체적인 매크로블록 모드 선택 방법

제안하는 알고리즘은 MPEG-2 I프레임과 P프레임을 H.264의 I프레임과 P프레임으로 변환하는 과정에서 인트라 모드 선택에 대해서만 다룬다. 제안하는 알고리즘의 전체적인 순서도는 다음 그림 3과 같다.

제안하는 알고리즘에서 사용한 양자화 파라미터에 따른 임계값은 표1과 표2와 같다. 임계값은 실험에 의해서 화질과 비트율을 고려해서 결정했다.

표 1. activityP의 임계값

QP	Threshold	QP	Threshold
25	130	30	155
26	130	31	155
27	135	32	155
28	140	33	155
29	145		

표 2. activity_{I1}과 activity_{I2}의 임계값

QP	activity_{I1}	activity_{I2}
25	150	800
26	200	750
27	250	730
28	300	700
29	300	700
30	300	700
31	300	700
32	300	700
33	300	700

IV. 실험 및 결과

실험에는 MPEG-2부호기에서 복호한 영상을 H.264 부호기의 입력으로 사용하였고 제안한 알고리즘의 구현을 위해 움직임 벡터와 매크로블록의 activity를 H.264부호기에 전달한다. MPEG-2 복호화기는 TM5를 사용하였고, H.264 부호화기는 JM8.4베이스라인 프로파일을 사용하였다. 테스트 시퀀스는 CIF(352×288)크기로 다양한 움직임 크기를 고려하기 위하여 비교적 큰 움직임을 갖는 foreman, table tennis 시퀀스의 처음 100프레임과 상대적으로 적은 움직임을 갖는 akiyo, carphone 시퀀스의 처음 100프레임을 테스트 시퀀스로서 사용하였다. 본 논문에서 제안하는 트랜스코딩 기법은 H.264의 베이스라인 프로파일로의 변환을 위하여 입력으로 사용된 MPEG-2 스트림에서도 P프레임만을 고려했고, 인트라모드 결정을 위해서는 모두 I프레임인 영상을 인트라모드 결정을 위해서는 GOP5인 영상과 첫 장만 I프레임이고 모두 P프레임으로 된 영상(IPPP...)을 사용하였고, 마지막으로 두 가지 방법을 모두 적용하여 GOP5인 영상에 대해서 RDO 기법과 비교하였다. 양자화 파라미터는 {25, ..., 33}을 사용하였다.

그림4는 I시퀀스에 대한 비트율과 PSNR 결과를 보여준다. RDO기법을 사용했을 때와 거의 차이가 없음을 볼 수 있다. 하지만 table tennis 시퀀스의 경우에는 복잡한 영상이기 때문에 다른 시퀀스에 비해서 비트율이 늘어나는 것을 볼 수 있다.

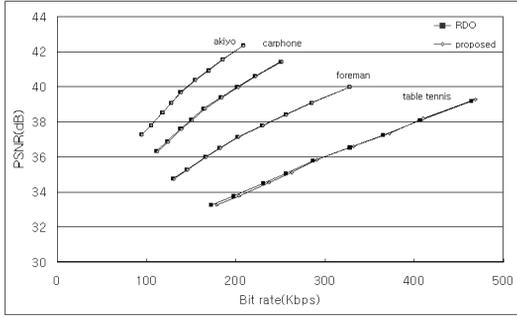


그림 4. I 시퀀스에 대한 RDO기법의 트랜스코더와 제안한 인트라모드 결정 기법 트랜스코더의 RD 결과.

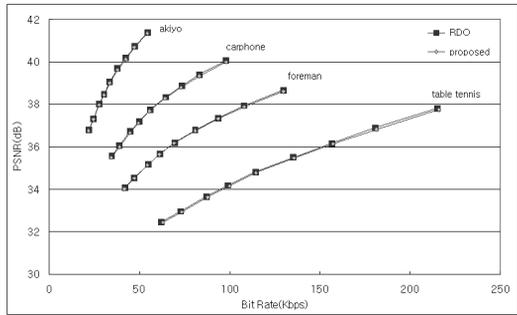


그림 5. IPPPP(N=5) 시퀀스에 대한 RDO기법의 트랜스코더와 제안한 인터모드 결정 기법 트랜스코더의 RD 결과.

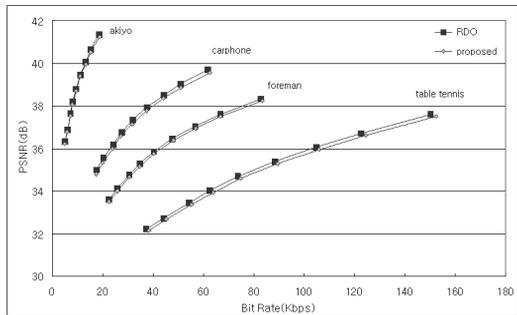


그림 6. IPPPP(첫 프레임만 I 프레임) 시퀀스에 대한 RDO 기법의 트랜스코더와 제안한 인터모드 결정 기법 트랜스코더의 RD 결과.

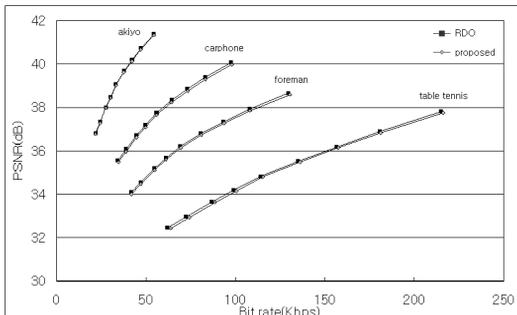


그림 7. IPPPP(N=5) 시퀀스에 대한 RDO기법의 트랜스코더와 제안한 모드 결정 기법 트랜스코더의 RD 결과.

표 3. 전체 트랜스코딩 시간의 비교

Sequence	Akiyo	Carphone	Foreman	Table tennis
Intra (OnlyI)	0.71 ~0.87	0.76 ~0.86	0.85 ~0.89	0.82 ~0.89
Inter (N=5)	0.40 ~0.43	0.59 ~0.60	0.63 ~0.68	0.65 ~0.69
Inter (IPPP...)	0.36 ~0.38	0.52 ~0.56	0.58 ~0.65	0.62 ~0.66
Intra+Inter (N=5)	0.36 ~0.37	0.48 ~0.52	0.55 ~0.58	0.60 ~0.61
RDO transcoder	1.0	1.0	1.0	1.0

그림 5는 GOP구조가 5일 때 각 시퀀스마다 비트율과 PSNR에 따른 그래프를 나타냈고, RDO기법을 사용한 트랜스코더와 제안한 방법을 적용한 트랜스코더가 큰 차이가 없음을 볼 수 있다. 또한 그림6은 IPPPP...영상에 대한 비트율과 PSNR 결과를 그래프로 나타냈다. GOP구조가 5일때 RDO기법을 사용했을 때와 더 근접함을 볼 수 있다. IPPPP...구조에서는 잘못 선택된 모드로 인한 오류가 계속 전달되어 더욱 커지게 되지만 GOP가 5인 경우에는 중간에 I프레임이 있어서 오류가 전달되지 않기 때문에 더욱 좋은 결과를 얻을 수 있다. 그림7은 IPPPP 시퀀스에 대해서 제안한 화면내 예측 모드 결정 기법과 화면간 예측 모드 결정 기법을 모두 사용했을 때의 비트율과 PSNR에 따른 그래프이다. RDO기법을 사용했을 때와 큰 차이가 없음을 볼 수 있다.

표 3은 각 시퀀스에 대해서 RDO 트랜스코더의 트랜스코딩 시간을 1로 했을 때 상대적으로 걸리는 시간을 나타낸 표이다. Intra는 제안하는 화면내 예측 모드 결정 기법을 Inter는 제안하는 화면간 예측 모드 결정 기법을 적용했을 때를 가리킨다. 화면내 예측 모드 결정 기법만을 사용했을때는 최대 30%까지 감소했고, 화면간 예측 모드 결정 기법은 최대 64%까지 줄어들었다.

또한 표 3에서 볼 수 있는 것처럼 두 가지 제안하는 방법을 모두 적용함으로써 GOP5구조일때 최대 64%까지 트랜스코딩 시간을 줄일 수 있었다. 상대적으로 복잡한 foreman, table tennis 시퀀스는 akiyo, carphone 시퀀스와 비교해서 부호화 시간의 감소가 적다. 움직임이 큰 영상은 16x16블록크기의 모드로 선택되는 개수가 적기 때문에 제안하는 방법이 적용되는 블록의 수도 감소하기 때문이다. 하

지만 비교적 평탄한 akiyo, carphone 시퀀스의 경우에는 많은 부호화 시간의 감소가 있음을 볼 수 있다.

V. 결론

제안한 MPEG-2에서 H.264로의 트랜스코더는 공간영역에서의 트랜스코딩 기법으로 MPEG-2 부호화부와 H.264 부호화부로 구성되어 있다. MPEG-2에서 얻을 수 있는 현재 매크로블록의 예측오차를 이용하여 H.264의 부호화 과정에서 많은 연산량을 필요로 하는 인터 모드 결정과정의 복잡도를 화질의 열화와 비트율의 큰 변화 없이 절감하였다. 시퀀스의 특성에 따라 다르지만 대부분 16×16 크기의 모드가 선택되는 경우가 많기 때문에 이때 모든 모드에 대해서 RDO 기법을 적용하는 것은 많은 계산량의 낭비를 초래한다. 따라서 매크로블록의 특성을 담고 있는 MPEG-2에서의 예측오차를 이용하여 선택될 모드의 크기를 예측하고 H.264 부호화 과정에서 매크로블록 모드를 적응적으로 선택함으로써 보다 낮은 복잡도의 트랜스코더를 구현할 수 있었다. 제안하는 트랜스코딩 기법을 사용하는 경우 평균 0.04dB의 화질 열화가 발생하는 반면 전체 트랜스코딩 시간은 최대 64%까지 단축되었다.

본 논문에서 제안하는 MPEG-2에서 H.264로의 트랜스코딩 알고리즘은 MPEG-2 비트스트림 내의 부가정보를 이용함으로써 최소한의 화질열화만을 갖는 저 복잡도 트랜스코딩 기법이며, 이는 고속 움직임 추정 기법 등과 같이 사용될 경우 실시간 트랜스코딩에 더욱 다가갈 수 있을 것으로 생각된다.

참고 문헌

[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] 김동형, 정제창, "MPEG2에서 H.264로의 트랜스코딩 기법," *한국통신학회 논문지* vol. 30, no. 7C, pp.706-715, Jul. 2005.

[3] H.264, Draft ITU-T Recommendation and Final Draft International Standard, Pattaya, Thailand, 2003.

[4] H. S. Malvar, A. Hallapuro, and M. Karzewicz, "Low-Complexity Transform and Quantization in H.264/AVC," *IEEE Trans. Circuits Syst.*

Video Technol., vol. 13, no. 7, pp.598 - 603, Jul. 2003.

[5] Xiaolan Lu, Tourapis, A.M., Peng Yin and Boyce, J, "Fast mode decision and motion estimation for H.264 with a focus on MPEG-2/H.264 transcoding," *in Proc. ISCAS 2005.*, vol. 2, pp.1246-1249, May 2005.

[6] Petjanski, B, Kalva, H, "DCT Domain Intra MB Mode Decision for MPEG-2 to H.264," *IEEE International Conference on Consumer Electronics*, pp.419-420, Jan. 2006.

[7] T. Wiegand, M. Lightstone, D. Mukherjee, T. G. Campbell, S. K. Mitra, "Rate-distortion optimized mode selection for very low bit rate video coding and the emerging H.263 standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, pp.182 - 190, Apr. 1996.

[8] Mester, R, Franke, U, "Spectral entropy-activity classification in adaptive transform coding Selected Areas," *IEEE Trans. Communications*, vol 10, Issue 5, pp.913 - 917, Jun. 1992.

[9] C.S. Kannagara, I.E.G. Richardson, M. Bystrom, J.R. Solera, Yafan Zhao, A. MacLennan and R. Cooney, "Low-complexity skip prediction for H.264 through Lagrangian cost estimation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, pp . 202-208, Feb. 2006.

김원균 (Won kyun Kim)

준회원



2005년 2월 : 한양대학교 전자전 기컴퓨터공학부 졸업
 2007년 2월 : 한양대학교 대학원 전자통신컴퓨터공학과 석사
 2007년 3월~현재 : 한양대학교 대학원 전자통신컴퓨터공학과 박사과정

<관심분야> 영상처리 및 영상압축, 트랜스코딩

박 경 준 (Kyung jun Park)

준회원



2005년 2월 : 홍익대학교 전자전기공학부 졸업
2007년 2월 : 한양대학교 대학원 정보디스플레이 공학과 석사
2007년 3월~현재 : 삼성전자 정보통신 총괄 사업부 <관심분야> 영상처리 및 영상압

축, 트랜스코딩

유 종 민 (Jong min You)

준회원



2003년 2월 : 한양대학교 전자전기컴퓨터공학부 졸업
2005년 2월 : 한양대학교 대학원 전자통신컴퓨터공학과 석사
2005년 3월~현재 : 한양대학교 대학원 전자통신컴퓨터공학과 박사과정

<관심분야> 영상처리 및 영상압축, 트랜스코딩

정 제 창 (Je chang Jeong)

정회원



1980년 2월 : 서울대학교 전자공학과 졸업
1982년 2월 : KAIST 전기 전자공학과 석사
1990년 : 미국 미시간대학 전기공학과 공학박사
1980~1986 : KBS 기술연구소

연구원(디지털 TV 및 뉴미디어 연구)

1990~1991 : 미국 미시간대학 전기공학과 연구교수 (영상 및 신호처리 연구)

1991~1995 : 삼성전자 멀티미디어 연구소 (MPEG, HDTV, 멀티미디어 연구)

1995~현재 : 한양대학교 전자통신컴퓨터공학과 교수 (영상통신 및 신호처리 연구실)

1998년 11월 27일 : 과학기술자상 수상

1998년 12월 31일 : 정보통신부장관상 표창

<관심분야> 영상처리 및 영상압축