

# FD-CD 트랜스코딩기법에서 CD에 의한 전파 왜곡의 적응적 예측 모델

김진수<sup>†</sup>, 김재곤<sup>\*\*</sup>

## 요 약

비디오를 다양한 단말과 이종의 망에 효과적으로 제공하기 위한 방법 중에, FD(Frame Dropping)-CD(Coefficient Dropping) 트랜스코딩 기법은 계산량을 줄이고, 손쉽게 구현할 수 있다는 측면에서 주목을 받고 있다. 그러나, FD-CD 트랜스코딩 기법은 CD에 의한 왜곡이 GOP내의 부호화 의존성을 갖는 프레임의 복호화 영상 품질에 많은 영향을 미친다. 본 논문에서는, CD로 인하여 각 프레임별로 전파되는 왜곡의 크기를 적응적으로 예측하는 모델링 기법을 제안한다. 제안된 기법은 각 프레임별로 부호화된 매크로블록의 부호화 모드 비율에 대해 적응적으로 구해진다. 모의실험을 통하여, 제안된 모델은 CD로 인해 초래되는 전체 왜곡 크기를 잘 예측할 수 있음을 보이고, 또한, 부호화된 비트스트림에서 CD에 의해 초래되는 프레임 단위로 일정한 전체 왜곡을 삽입하는데 효과적임을 보인다.

## An Adaptive Estimation Model for Propagation Errors Incurred by CD in FD-CD Transcoding

Jin-soo Kim<sup>†</sup>, Jae-Gon Kim<sup>\*\*</sup>

## ABSTRACT

Recently, FD (Frame Dropping)-CD (Coefficient Dropping) transcoding is considered mainly due to the low computational complexity and simple implementation. But, CD errors in the FD-CD transcoding scheme tend to be propagated and they have a significant effect on the qualities of decoded images. In this paper, we derive the error characteristics incurred by the CD operations and propose an effective estimation model that adaptively describes well the characteristics of propagation/accumulation errors in compressed domain. Furthermore, we apply the proposed model to distortion control achieving nearly constant distortion allocation among frames. Simulation results show that the proposed model is quite accurate in estimating the overall distortions and is effectively applied to distortion control over a range of sequences with varying scene types.

**Key words:** Coefficient Dropping(계수제거), UMA(Universal Multimedia Access), Video Transcoding (비디오트랜스코딩)

※ 교신저자(Corresponding Author) : 김진수, 주소 : 대전시 유성구 덕명동 산16-1(300-719), 전화 : 042)821-1147, FAX : 042)821-1595, E-mail : jskim67@hanbat.ac.kr  
접수일 : 2004년 1월 27일, 완료일 : 2004년 6월 4일  
<sup>†</sup> 정회원, 한밭대학교 정보통신컴퓨터공학부 부교수  
<sup>\*\*</sup> 한국전자통신연구원 방송컨텐츠연구팀장  
(E-mail : jgkim@ctri.re.kr)  
※ 본 연구는 정보통신부 과제의 일부로서 수행되었습니다.

## 1. 서 론

이종의 망과 다양한 사용자 단말 특성에 맞추어 비디오 콘텐츠를 효과적으로 수용하는 기법에 대한 연구로서, 트랜스코딩 기법에 대한 연구가 활발히 진행되고 있다[1-3]. 다양한 종류의 트랜스코딩 기법 중에서 압축 영역에서 수행하는 FD(Frame Drop-

ping)-CD(Coefficient Dropping) 트랜스코딩 기법은 연산량을 매우 절감시킬 수 있고, 구현이 간단하여, 쉽게 사용할 수 있는 장점으로 인하여 관심을 끌고 있다[4-7].

FD-CD 트랜스코딩 기법에서 FD는 압축 부호화된 몇 개의 프레임은 제거함으로써 많은 양의 비트율을 제어할 수 있다. 가장 중요도가 떨어지거나 또는 다른 부호화 프레임에 의존하지 않는 프레임을 제거하면서, 우수한 품질을 유지한다[4,5]. 일반적으로, MPEG으로 압축 부호화된 비디오 시퀀스에서 B프레임은 다른 프레임에 대해 부호화 의존성을 제공하지 않으므로 가장 먼저 제거된다. B프레임들이 제거된 후에는 P프레임이 우선 제거대상이 되며, 부호화 의존성을 고려하여 GOP(Group of Pictures)내에서 가장 나중에 나오는 P프레임 순으로 제거될 수 있다. 이렇게 함으로써, 트랜스코딩된 영상의 품질을 최대한 우수한 결과가 귀결되도록 제어할 수 있다. CD에 의한 트랜스코딩은 FD에 비해 적은 양의 데이터를 세밀하게 제어할 수 있는 수단이 된다. CD기법은 중요하지 않은 고주파 DCT계수의 VLC를 제거함으로써 비트율을 제어하는 기법이다. VLC코드를 제거하는 방법으로는 그림 1에 나타낸 것과 같이, EOB(End of Block)코드에서부터 순차적으로 제거하는 방법과 단일 블록내에 위치에 상관없이 중요도가 낮은 계수를 없애는 방법이 있다. [4,5]에서 두 방식에 대한 성능 차이는 매우 미미한 수준으로 보고하고 있다. 이와 같은 모의 실험결과를 바탕으로, 그림 1.(a)에 의한 방식은 CD를 수행할 때, 압축된 비트열에 대한 문법 구조 변화를 최소로 유지하므로, 더욱 쉽게 구현될 수 있다.

한편, CD가 적용되는 프레임에 대한 기존의 비트

율 제어기법은 크게 두 가지로 나뉜다. 할당된 목표 프레임 비트량을 맞추기 위해, 각 DCT블록에서 버려지는 DCT계수 양을 균등하게 하여 CD하는 기법(URCD: Uniform Rate-based CD)이 있다. 그러나, 이 기법은 초래되는 왜곡의 크기를 최소화하지 않는 불합리한 접근 방법이다. 이에 비해 한 프레임 기간 동안의 전체 DCT블록에 최소의 왜곡이 초래되도록 각 블록에 대해 최적의 EOB 지점을 찾는 기법(LOCD: Lagrange Optimization CD)이 있다[4,5]. 즉, 프레임내에  $N$ 개의 DCT블록이 존재한다고 할 때,  $i$  번째 DCT블록이  $K_i$ 개의 non-zero DCT계수를 포함한다면, LOCD 기법은 다음과 같다.

$$\min \left\{ \sum_{i=1}^N \left( \sum_{k=1}^{K_i} x_{i,k}^2 + \lambda \sum_{k=1}^{b_i} r_{i,k}^{coeff} \right) \right\} \quad (1)$$

여기서,  $x_{i,k}$ 는  $i$ 번째 블록의  $k$ 번째 non-zero계수 크기이고,  $r_{i,k}^{coeff}$ 는  $i$ 번째 블록의  $k$ 번째 non-zero계수를 부호화하기 사용되는 비트량을 나타낸다. LOCD에서 사용되는 식(1)은 주어진 목표 프레임 비트량에서 왜곡을 최소화하도록 하는  $\lambda$ 를 구함으로써, 임의의 블록에 최적의 왜곡을 주는 breakpoint위치(EOB)로서  $b_i$ 를 찾게 된다[6].

그러나, GOP내의 부호화 의존성을 갖는 프레임은 복호화 영상 품질에 많은 영향을 미치지만, 기존의 연구 결과에서는 이에 대한 영향을 무시하고 있다 [4-6]. 본 논문에서는 CD로 인하여 각 프레임별로 전파되는 왜곡의 크기가 무시할 정도로 작지 않음을 보이고, CD왜곡의 전파특성을 적용적으로 예측하는 모델을 제안한다. 제안된 예측 모델은 각 프레임별로 부호화된 매크로블록의 부호화 모드 비율에 대해 적

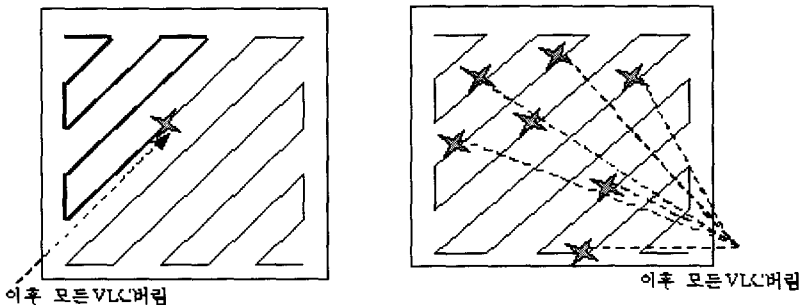


그림 1. CD에 대한 두 가지 제어 방식

응적으로 구해진다. 모의실험을 통하여, 제안된 모델은 CD로 인해 초래되는 전체 왜곡 크기를 잘 예측할 수 있음을 보이고, 또한, 부호화된 비트스트림에서 CD에 의해 초래되는 프레임 단위로 일정한 전체 왜곡을 삽입하는데 효과적임을 보인다.

## 2. CD에 의한 전파 왜곡 분석

우선, 본 논문에서는 전개의 편리함을 위해 CD 제어 방식은 그림 1.(a)에 나타난 방법을 사용하는 것으로 가정한다. 트랜스코딩을 수행하지 않았을 경우에,  $j$ 번째 프레임의  $i$ 번째 블록에 대한 복호화된 표현( $y_j(i)$ )은 다음과 같이 나타낸다.

$$y_j(i) = MC_i(y_{j-1}) + e_{j,i} \tag{2}$$

여기서,  $MC_i(y_{j-1})$ 는  $j-1$ 번째 프레임을 기초로 움직임 보상된(Motion Compensated) 성분이고,  $e_{j,i}$ 는 움직임 예측 오차신호 성분의 복호화된 표현이다. 만일,  $j-1$ 번째와  $j$ 번째 프레임이 각각 독립적으로 CD를 수행한다면,  $j$ 번째 프레임의  $i$ 번째 블록에 대한 복호화된 표현( $\hat{y}_j(i)$ )은 식(2)에 기초하여 다음과 같이 나타낼 수 있다.

$$\hat{y}_j(i) = MC_i(\hat{y}_{j-1}) + \hat{e}_{j,i} \tag{3}$$

여기서,  $MC_i(\hat{y}_{j-1})$ 는 CD를 적용한  $j-1$ 번째 프레임을 기초로 움직임 보상된 성분이고,  $MC_i(\cdot)$ 는 FD-CD 트랜스코딩 기법과 무관한 함수이다. 그리고,  $\hat{e}_{j,i}$ 는  $e_{j,i}$ 에 대해 CD를 적용한 후의 복호화된 표현이다. 식 (1)과 (2)로부터, CD에 의한 트랜스코딩을 수행함으로써 초래되는  $j$ 번째 프레임의  $i$ 번째 블록에 대해 평균자승오차(MSE: Mean Square Error)로 측정되는 왜곡( $D_{j,i}(b_i)$ )은 다음과 같이 표현할 수 있다.

$$D_{j,i}(b_i) = \frac{1}{N} |y_j(i) - \hat{y}_j(i)|^2 = \frac{1}{N} |a_{j,i} + e_{j,i} - \hat{e}_{j,i}|^2 \tag{4}$$

여기서,  $b_i(0 \leq b_i \leq N-1, N=64)$ 는 CD를 수행함에 따라 새로이 결정되는  $i$ 번째 DCT블럭의 EOB위치를 나타낸다. 그리고,  $a_{j,i}$ 는  $j-1$ 번째 프레임에서 CD를 수행함으로써,  $j$ 번째 프레임으로 전파되는 움

직임 보상 오차 즉,  $MC_i(y_{j-1}) - MC_i(\hat{y}_{j-1})$ 를 부호화 스캔(예, zig-zag 스캔) 순서에 의해 2차원 데이터 표현에서 1차원 데이터 표현으로 나타낸 것이다. 식 (4)를 DCT영역의 표현으로 나타내기 위해,  $A(k) = DCT_k a_{j,i}$ ,  $E(k) = DCT_k e_{j,i}$ , 그리고  $\hat{E}(k) = DCT_k \hat{e}_{j,i}$ 으로 각각 표기하면, 왜곡은 다음과 같이 측정된다.

$$\begin{aligned} D_{j,i}(b_i) &= \frac{1}{N} \sum_{k=0}^{N-1} \{A(k) + E(k) - \hat{E}(k)\}^2 \\ &= \frac{1}{N} \sum_{k=0}^{N-1} A^2(k) + 2 \sum_{k=0}^{N-1} A(k)E(k) + \sum_{k=0}^{N-1} E^2(k) \end{aligned} \tag{5}$$

식(5)에서  $E(k)$ 에서  $b_i$ 이후에 있는 DCT계수 성분들은 모두 CD되므로, 이 계수들에 의해 차이에 의해 왜곡이 발생한다. 반면에 식(5) 우측의 첫 번째 항은  $MC_i(y_{j-1}) - MC_i(\hat{y}_{j-1})$ 에 의해 발생하는 왜곡이며,  $j-1$ 번째 프레임에서 전해져 오는 왜곡이다. 그리고, 식(5) 우측의 두 번째 항은 왜곡의 측정단위가 블록단위의 평균 자승 오차(MSE: Mean Square Error)임으로 인해 측정되는 것으로  $j-1$ 번째 프레임에서 움직임 보상 오차 신호 성분과  $j$ 번째 프레임의 CD로 인해 발생하는 오차신호 성분간의 상관관계로 인한 왜곡성분이다.

이상에서 논의된 CD에 의한 전파 왜곡 분석을 위해, 모의실험을 수행한다. 비트 스트림의 부호화 특성은 MPEG-4 simple profile, Foreman시퀀스 CIF 1~300 frames, IBBPBBPBB... 형태(1.5Mbps, GOP=15/SubGOP=3)이다. 본 논문에서는 I/P프레임만을 CD의 대상으로 고려하고, B프레임은 FD의 대상으로 제한하며, 목표 프레임 비트량 제어는 I절에서 논한 LOCD제어기법을 사용한다. 트랜스코딩 이전과 이후의 비트율을 각각  $R=1.5Mbps$ 와  $R'=800kbps$ 으로 하여 트랜스코딩한다. 이때 발생한 평균자승오차(MSE: mean square error)를 그림 2에서 보여주고 있다. 그림에서 알 수 있듯이, CD로 인하여 왜곡의 전파가 매우 크게 발생함을 알 수 있으며, 전파 및 누적 왜곡의 크기는 각 프레임의 CD 왜곡과 거의 비슷한 크기로 발생함을 확인할 수 있다. 또한, 같은 GOP내에서 앞선 프레임에 비해 뒤쪽에 위치하는 프레임에서 전파 왜곡이 누적되는 특성으로 인해 상대적으로 크게 나타남을 알 수 있다. 여기서 한 가지 주목해야할 실험적 결과는 이전 프레임에서 전해져

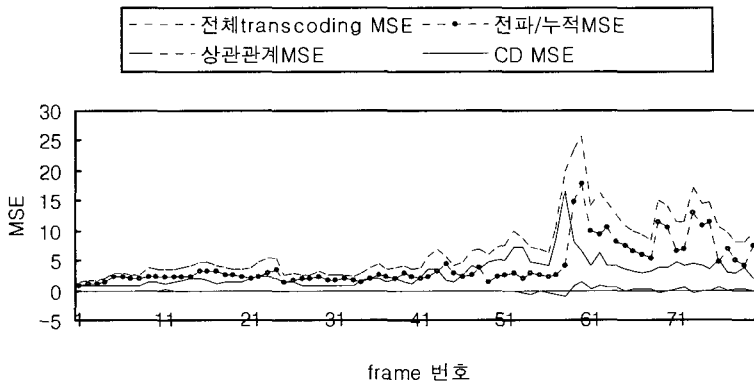


그림 2. 일정률 CD 할당 실험

오는 전파 및 누적 왜곡과 현재 프레임에 적용되는 CD 왜곡사이에는 상관관계가 존재하지 않고, 독립적인 경우로 가정할 수 있다는 것이다.

### 3. 프레임단위의 CD 적용에 따른 전파왜곡의 분석

부호화된 각 프레임의 비트스트림에 대해 독립적으로 일정율의 CD를 적용함으로써 발생하는 왜곡의 전파 및 누적 특성을 알아보기 위해 모의실험을 수행한다. 이를 위해, 그림 2의 모의실험에서 사용된 Foreman 시퀀스(R=1.5Mbps로 압축, GOP=15, SubGOP=3)를 B프레임들은 모두 FD하고, 모든 P프레임과 I프레임은 부호화된 non-zero DCT AC계수를 식(1)로 나타낸 LOCD방식에 의해 균등하게 15%를 CD하도록 한다. 이 결과를 복호하여 CD로 인하

여 초래된 전체 왜곡을 조사하여 나타낸 것이 'total VOP's 15%CD'로 주어진다. 그림 3에서 'IVOP\_15%CD'로 표시한 것은 각 GOP내의 I프레임에 대해서만 15%CD를 수행하고, 다른 프레임들은 CD하지 않은 경우에 복호화한 후의 왜곡 특성을 나타낸 것이다. 이와 유사하게 각 GOP내의 모든 P프레임에 대해서도 이와 같은 방식으로 독립적으로 15%CD를 수행하고, 다른 프레임들은 트랜스코딩을 하지 않은 경우에 대한 복호화후의 왜곡 특성을 나타낸다. (그림 3에서 각각 'PVOP1\_15%CD', 'PVOP2\_15%CD', 'PVOP3\_15%CD', 'PVOP4\_15%CD'로 표시함). 그림 3에서 'Sum of VOP 15%CD'로 표시한 것은 이상에서 각 프레임 별로 독립적으로 15% CD를 수행하고 복호화한 후의 왜곡들의 합을 나타낸 것이다. 여기서 알 수 있듯이, I프레임에 대한 CD가 제일 큰 왜곡을 발생시키고, 왜곡의 전파 특성이 가장 크게 나타난

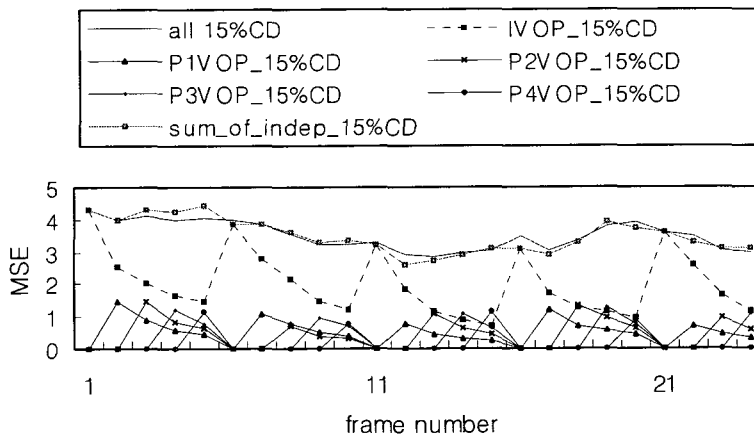
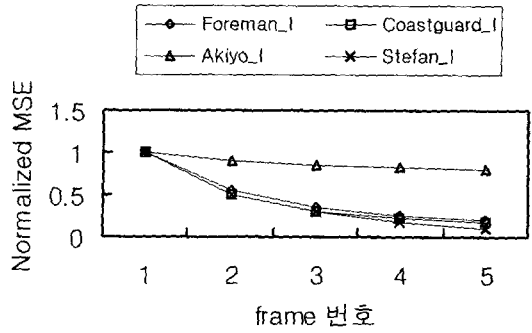


그림 3. 통합 15%CD의 경우와 개개 프레임별 15%CD 경우의 왜곡 특성 비교

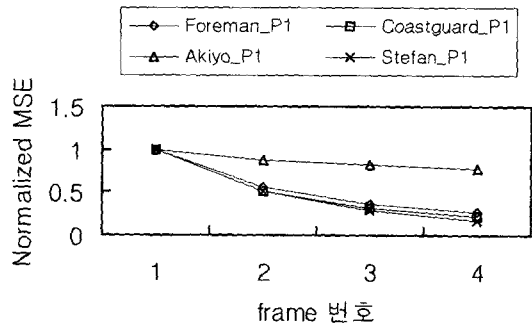
다. 그리고 I프레임과 다른 P프레임들의 CD왜곡 특성은 지수 감소함수형태로 전파되고, 각 프레임의 전체 왜곡으로 누적됨을 보이고 있다. 또한, 그림 3에서 주목할 특징은 개개의 프레임별로 독립적으로 CD를 수행하고, 복호화된 결과의 합('Sum of VOP 15%CD')에 대한 특성은 통합적으로 15%CD를 수행한 경우('total VOP's 15%CD')의 특성과 매우 유사한 특성을 보임을 알 수 있다. 이상에서 논의된 그림 2와 그림 3으로부터, 각 프레임단위로 수행되는 CD 왜곡에 대한 전파 및 누적되는 특성은 개개의 프레임 단위로 독립적인 특성을 나타내고 있다. 즉, 각 프레임에 첨가되는 CD왜곡[식 (5) 우측의 세 번째 항]과 이전 프레임에서 전파 및 누적되는 왜곡[식 (5) 우측의 첫 번째 항]사이에는 독립적인 관계를 갖는다고 볼 수 있다. 그러므로, 개개의 프레임 단위로 수행되는 CD 특성을 모델링하여 전체 왜곡을 근사적으로 예측 가능해진다.

본 논문에서는 다른 영상 시퀀스의 압축 부호화된 비트스트림에 대한 특성을 조사하기 위해, Akiyo, Stefan, Coastguard 시퀀스에 대해서도 앞에서 사용한 Foreman시퀀스와 동일한 모의실험을 수행한다. 표1에 나타낸 바와 같이, Akiyo, Stefan, Coastguard 시퀀스에 대해서 각각 1.2Mbps, 1.6Mbps, 1.4Mbps로 압축(부호화구조: GOP=15, SubGOP=3, 모두 300 프레임)한 후에, IPPPP...구조의 트랜스코딩을 위해 B프레임은 FD를 적용하고, I/P프레임은 5%~30% 범위내에서 일정률의 CD를 적용한다.

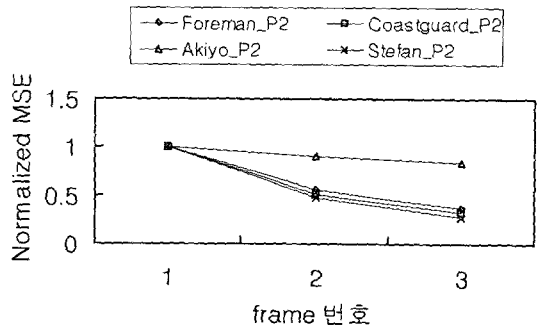
그림 4의 (a), (b)와 (c)는 각각 I프레임, GOP내의 첫 번째 P프레임, GOP내의 두 번째 P프레임 그리고 GOP내의 세 번째 P프레임에서 독립적으로 5%~30%범위의 CD로 인한 왜곡의 평균적인 전파특성을 정규화하여 나타낸 것이다. 이 그림에서 Akiyo시퀀스와 같이 skipped-MB로 부호화된 매크로블록의 개수가 많은 경우에는 CD에 의한 왜곡 전파 효과가 매우 크게 발생함을 알 수 있다. 즉, 움직임이 없는 배경과 같은 영역은 CD에 의해 왜곡 크기가 그대로 전파되게 된다. 반면에 Foreman, Stefan, Coastguard 시퀀스들의 경우에는 비슷한 특성을 보이고 있다. 단, Stefan시퀀스와 같이 움직임이 많거나 복잡한 장면들로 구성된 경우에는 intracoded-MB로 선택된 매크로블록이 다소 많아서 CD에 의한 왜곡의 전파효과를 차단시키는 결과를 얻게 된다. 이상에



(a) I프레임



(b) GOP내 첫 P프레임



(c) GOP내 두 번째 P프레임

그림 4. 정규화된 CD 왜곡의 평균적인 전파 특성

서 얻은 모의실험 결과를 정리하면 다음과 같이 요약된다.

정리. CD로 인해 초래되는 전파 왜곡은 다음과 같은 특성을 갖는다.

i). 각 프레임별 적용되는 CD 왜곡과 이전 프레임에서 전파되어 오는 왜곡에 의한 상관 왜곡은 무시 가능하다.

ii). 각 프레임 구간별 왜곡의 감쇄 특성은 CD 왜곡

값의 절대 크기에 의존적인 지수 감소함수형태로 전파된다.

iii). 부호화되지 않은 매크로블록(skipped-MB)은 CD 왜곡의 전파를 매우 크게 한다.

iv). 화면내 부호화 매크로블록(intra-coded-MB)은 CD 왜곡의 전파를 차단하는 특성을 갖는다.

#### 4. CD에 의한 전파 왜곡의 적응적 예측 모델

본 절에서는 CD에 의한 전파 왜곡을 각 부호화된 프레임단위에 대해 매크로블록의 통계적 특성에 기초하여 모델링한다. 본 논문에서 제안되는 모델링은 그림 4의 정규화된 왜곡의 전파 특성을 직접 매칭하여 얻어진다. 먼저, 이를 위해, CD가 적용되는 프레임임을 0이라는 기호를 사용하고, 이 프레임에 CD를 적용하여 GOP내의 전파 왜곡 크기를 구하려는 프레임 기호는  $i(i \geq 1)$ 로 표시한다. 단일 프레임내의 전체 매크로블록의 개수를  $N_{MB}$ 라고 하고,  $i$ 번째 프레임 내에서 skipped 매크로블록의 개수와 영상내 부호화 매크로블록의 개수를 각각  $N_s$ 와  $N_a$ 로 표현하고,  $D_x(i-1)$ 은 인덱스  $i-1$ 프레임에서 실제 CD된 왜곡 값 또는 예측된 왜곡 값이라고 할 때, 인덱스  $i$ 에서의 전파 왜곡 값  $D_x(i)$ 는 다음과 같은 모델을 이용하여 예측한다.

$$D_x(i) = D_x(i-1)\delta_s(i) + D_x(i-1)1 - \delta_s(i) - \delta_a(i)e^{-\alpha \cdot i/2.6^{i-1}} \\ = D_x(i-1)\delta_s(i) + D_x(i-1)1 - \delta_a(i)e^{-\alpha \cdot i/2.6^{i-1}}, i \geq 1 \quad (6)$$

$$\alpha(i) = 0.6 + \delta_a(i) + \delta_s(i) = 1.6 - \delta_a(i) \quad (7)$$

여기서,  $\delta_s(i) = N_s(i)/N_{MB}$ ,  $\delta_a(i) = N_a(i)/N_{MB}$ 와  $\delta_a(i) = 1 - \delta_a(i) - \delta_s(i)$ 를 각각 나타낸다. 식(6)에서 지수 감소 함수의 기울기  $\alpha(i)$ 는 intra mode로 부호화된 매크로블록의 비율이 많은 프레임에 대해 비례하여 크게 감쇄하게 된다. 이는 intra mode 매크로블록은 CD에 의한 왜곡 전파를 차단하고, refresh 시키는 효과를 갖기 때문이다. intra mode 매크로블록 비율을 inter mode 매크로블록의 비율  $\delta_a(i)$ 로 표현하면,  $\delta_a(i)$ 에 반비례하는 특성을 유지한다.  $\delta_s(i)$ 의 관계는 식(6) RHS 첫 번째 항으로 보상을 하므로, 식(7)에서 이 성분에 의한 감쇄특성은 제외되도록 한다. 감쇄 상수에서  $i/2.6^{i-1}$ 은  $i \geq 1$ 에 대해 1차 지수

감소 함수를 보상하기 위해 도입된 계수이다. 즉,  $i=2$ 에서  $2.0/2.6$ ,  $i=3$ 에서  $3/2.6^2$  ...등으로 왜곡의 전파 특성 기울기를 다소 완만하게 도입하기 위해 도입된다.

그리고, X는 GOP내에 있는 순서화된 영상의 부호화 형태를 나타내는 기호로서, X는 I, P1, P2, P3, P4 등을 나타낸다.  $D_x(0)$ 는 X-VOP에 직접 CD에 의해 발생하는 왜곡 크기를 나타내고,  $D_x(0)$ 에 기초하여, 인덱스  $i$ 는 전파왜곡이 적응적으로 예측되는 프레임 길이를 표시한다. 따라서, 각 부호화된 VOP의 전체 왜곡에 대한 예측값은 다음과 같이 구해진다.

IVOP:  $D_f(0)$

GOP내의 첫 번째 PVOP:  $D_f(1) + D_{P1}(0)$

GOP내의 두 번째 PVOP:  $D_f(2) + D_{P1}(1) + D_{P2}(0)$

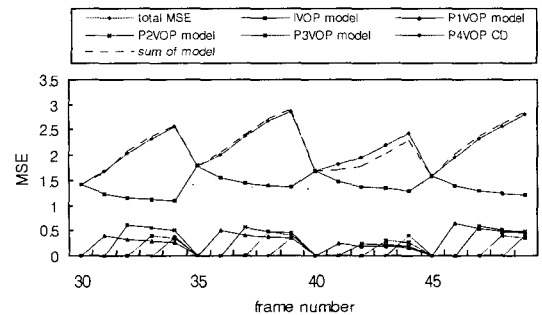
GOP내의 세 번째 PVOP:

$$D_f(3) + D_{P1}(2) + D_{P2}(1) + D_{P3}(0)$$

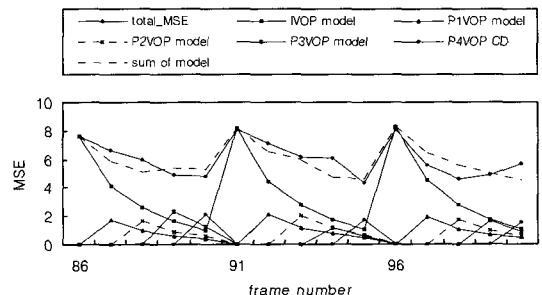
GOP내의 네 번째 PVOP:

$$D_f(4) + D_{P1}(3) + D_{P2}(2) + D_{P3}(1) + D_{P4}(0)...(8)$$

그림 5는 제안된 적응적 예측 모델에 대한 모의실험 결과를 나타낸다. 그림에서 "total\_MSE"는 CD비



(a) "Akiyo" 시퀀스



(b) "Foreman" 시퀀스

그림 5. 적응적 예측 모델과 시뮬레이션 결과

율을 전체 AC DCT계수에서 15% 또는 25%의 비율로 식(1)에 의한 CD제어를 수행한 후에 복호하여 프레임별로 측정된 실제 왜곡 크기를 나타낸 것이다. IVOPmodel, PIVOPmodel, P2VOPmodel 그리고 P3VOPmodel은 GOP내의 해당 VOP에 CD왜곡을 삽입후, 식(7)에 의한 왜곡의 전파 특성을 예측하여 나타낸 것이고, P4VOPCD는 BVOP는 FD대상이어서 실제 전파되는 왜곡이 없는 값이다. 그리고, "sum of model"은 이들 예측된 각 전파왜곡에 대한 결과들의 합 즉, 식(8)에 의한 결과를 나타낸 것이다. 이 비교를 통하여 알 수 있듯이, 제안된 적응적 예측 모델은 전파 왜곡의 특성을 적절히 잘 근사화시킬 수 있음을 보여준다.

Akiyo시퀀스에 대해서는 각 VOP별로 25%의 비율로 CD를 수행한 경우이다. IVOP의 CD 왜곡은 다른 PVOP들에 비해 매우 큰 영향을 미치는 것을 알 수 있는데, 그 이유는 skipped-MB로 선택되는 매크로블록이 많으며, 그로 인해 대부분의 CD 왜곡은 전파되게 된다. 이와 같은 각 CD왜곡에 따른 전파 기울기 특성은 GOP내의 다른 PVOP에서 거의 동일한 기울기로 발생한다. 반면에, Foreman시퀀스는 각 부호화된 프레임에서 평균적으로 15%율의 CD를 수행한 경우이지만, 발생하는 왜곡의 크기는 Akiyo시퀀스보다 더욱 큰 값으로 발생한다. 그러나, 왜곡 크기

의 값은 상대적으로 크게 나타나지만, 왜곡의 전파 기울기 특성은 Akiyo에 비해 매우 급하게 나타남을 확인할 수 있다. 이와 같은 특성은 intra mode의 매크로블록 비율이 Foreman시퀀스에서는 다소 차지하며(표 1 참조), 또한, skipped macroblock이 매우 작게 발생하여, 왜곡 전파를 차단하는 효과를 낳는다.

그림 6에서는 그림 5에 나타낸 실제 측정된 값("total\_MSE")과 예측 값("sum of model")의 차이를 백분율로 환산하여 나타낸 것이다. 이 그림에서 알 수 있는 것처럼, Akiyo시퀀스는 약 ±3%범위내에서 오차를 구할 수 있으며, Foreman시퀀스는 약 ±17%이내에서 오차를 유지함을 확인할 수 있을 정도로 효과적임을 확인할 수 있다. 특히, 여기서 확인할 수 있는 것은 Foreman시퀀스는 Akiyo시퀀스에 비해 다소 왜곡의 전파 특성이 다소 불규칙하며, 그 백분율에 의한 크기도 다소 크게 나타난다.

본 논문에서 제안된 적응적 예측모델의 효용성을 이용한 한가지의 응용으로서, CD에 의한 왜곡을 프레임별로 일정하게 할당하는 기법을 설계할 수 있다. 그 방식으로서 식(1)과 유사하게 설계된다. 즉, k-번째 프레임에 대해 전체 왜곡량을  $D_{keep}$  만큼 할당하고자 할 때, 모든  $8 \times 8$  DCT블럭들에 대한 새로운 EOB세트  $\{\vec{k} = (k_1, k_2, \dots, k_N), N = \text{DCT블럭의}$

표 1. 모의실험에 사용된 영상시퀀스의 부호화모드 특성

| 영상시퀀스      | 트랜스코딩전의 압축 비트율 | avg_intra MB개수 | avg_inter MB개수 | avg_skipped MB개수 |
|------------|----------------|----------------|----------------|------------------|
| Foreman    | 1.5Mbps        | 7.3            | 384.3          | 2.7              |
| Akiyo      | 800kbps        | 0              | 270.6          | 125.5            |
| Coastguard | 1.4Mbps        | 0.26           | 395.7          | 0.04             |
| Stefan     | 1.6Mbps        | 8.6            | 376.8          | 9.6              |

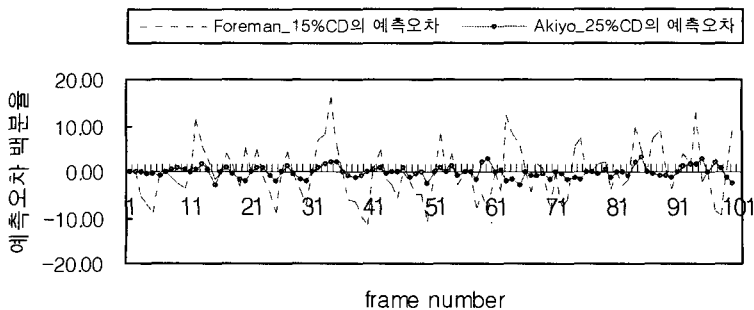


그림 6. 그림 5의 실험의 백분율 예측오차

개수}을 구하는 것이 요구된다.

$$\vec{k} = (k_1, k_2, \dots, k_N) \left\{ \sum_{i=1}^N \left( \sum_{k=b_i}^{K_i} x_{i,k}^2 + \lambda \sum_{k=1}^{b_i} r_{i,k}^{coeff} \right) \right\}$$

such that  $\left| D_{keep} - \sum_{j=1}^N \sum_{i=k+1}^{k_i} x_i^2 \right| \leq \varepsilon$  (9)

여기서,  $\varepsilon$ 는 허용되는 왜곡의 한계치이다.  $D_{keep}$ 가 모든 부호화된 프레임에 대해 거의 일정하게 할당되어야 하므로, 각 PVOP에서 CD에 의한 왜곡은  $D_{keep}$ 로부터 예측된 전과왜곡 만큼을 줄인 양이 더해져야 한다. 한가지 예로서, GOP내의 두 번째 PVOP는  $D_{I(2)} + D_{P(1)}$ 가 예측되어지고,  $D_{keep} - D_{I(2)} + D_{P(1)}$ 만큼이 실제로 두 번째 PVOP에 CD되어야 하는 왜곡의 양이 된다. 그림 7은 제안된 제어 방식에 대한 모의실험 결과의 일부분을 보여주고 있다. 모의 실험결과 테스트 시퀀스("Foreman"과 "Akiyo"시퀀스)에 무관하게 각 VOP에 대해 거의 일정한 왜곡이 더해질 수 있음을 보여주고 있다. 즉, 제안된 모델을 이용하면, 트랜스코딩으로 인하여 초래될 수 있는 왜곡의 크기를 일정하게 유지하여, 일정 화질 제어에 효과적으로 사용될 수 있음을 보여준다.

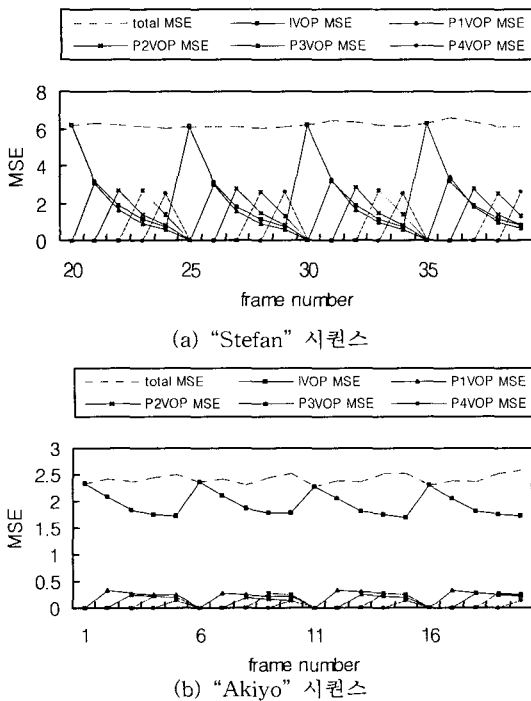


그림 7. CD에 의한 일정 왜곡 삽입 제어

## 5. 결 론

본 논문에서는 FD-CD트랜스코딩 기법에서 CD에 따른 전과 왜곡 발생을 수식적으로 전개하였고, 그 결과를 모의실험을 통하여 분석하였다. 모의실험을 통하여 CD로 인한 전과/누적 왜곡은 무시할 수 없을 정도의 매우 큰 왜곡을 발생시킴을 보였고, 또한, CD로 인한 왜곡은 이전 프레임에서 전과되어 오는 왜곡과 상관성이 없음을 확인하였다. 이 결과를 바탕으로 독립적인 지수 감소함수형태로 개개의 프레임별로 발생하는 CD왜곡의 전과 특성을 적응적으로 모델링하였고, 제안된 모델은 CD로 인한 전체 왜곡을 잘 예측할 수 있음을 보였다. 또한, 제안된 모델에 대한 한 응용으로서, 각 프레임별로 CD로 인한 전체 왜곡을 거의 일정하게 할당하는데 효과적으로 사용할 수 있음을 보였다.

앞으로는 B프레임을 포함하는 다양한 종류의 부호화 형식, 장면 전환 등을 포함하는 많은 종류의 시퀀스에 대해서도 효과적인 FD-CD 트랜스코딩 기법에 대한 연구를 수행할 필요가 있다. 그리고 나아가서는 최적의 FD-CD 트랜스코딩 기법을 찾기 위한 연구를 수행하는 것이 필요하다.

## 참 고 문 헌

- [1] J.Youn, and M.-T Sun, "Video Transcoding with H.263 Bit Streams," *Journal of Visual Communication and Image Representation*, Vol.11, December 2000.
- [2] Q.-F.Zhu, L.Keofsky, and M.B.Garrison, "Low-Delay Low-Complexity Rate Reduction and Continuous Presence for Multipoint Videoconferencing," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol.9, No.4, June 1999.
- [3] H.Sun, W.Kwok, and J.W.Zdowski, "Architecture for MPEG Compressed Bitstream," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol.6, Apr. 1996.
- [4] J.-G.Kim, Y.Wang, S.-F.Chang, K.Kang, J.Kim, "Description of utility function based optimum transcoding," *ISO/IEC JTCl/*



SC29/WG11 M8319, Fairfax, May 2002.

- [5] A. Eleftheriadis, "Dynamic Rate Shaping of Compressed Digital Video, Doctoral Dissertation," *Graduate School of Arts and Sciences*, Columbia University, June 1995.
- [6] J.-G. Kim, Y.Wang, and S.-F. Chang, "Content-adaptive Utility-based Video Adaptation," in *Proceedings of IEEE Conference on Multimedia and Expo(ICME 2003)*, Baltimore, Maryland, June 2003.
- [7] J.-s.Kim, J-G. Kim, K.Kang, and J.Kim, "A Distortion Control Scheme for Allocating Constant Distortion in FD-CD Video Transcoder," in *Proceedings of IEEE Conference on Multimedia and Expo(ICME 2004)*, Taipei, Taiwan, June 2004.



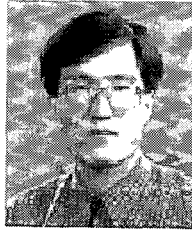
김진수

1991년 2월 경북대학교 전자공학과(공학사)  
 1993년 2월 한국과학기술원 전기및전자공학과(공학석사)  
 1998년 8월 한국과학기술원 전기및전자공학과(공학박사)  
 1995년 5월~2000년 3월 삼성전자

선임연구원

2000년 4월~현재 한밭대학교 정보통신컴퓨터공학부 부교수

2003년 3월~현재 한국전자통신연구원 초빙연구원  
 관심분야: 멀티미디어 신호처리, Video Streaming, Scalable Video Coding



김재곤

1990년 2월 경북대학교 전자공학과(공학사)  
 1992년 2월 한국과학기술원 전기 및 전자공학과(공학석사)  
 2001년 9월~2002년 11월 미국 Columbia 대학 방문연구원  
 1992년 3월~현재 한국전자통신

연구원(선임연구원)

2004년 6월~현재 한국전자통신연구원 방송컨텐츠연구팀장

관심분야: 비디오신호처리, 멀티미디어 프레임워크, 디지털방송