

FD-CD 트랜스코딩기법에서 CD에 의한 전파/누적 왜곡의 분석과 모델링

정회원 김진수*, 정회원 김재곤**, 정회원 김형명***, 정회원 홍진우**

An Analysis and Modeling of Propagation/Accumulation Errors Incurred by CD in the FD-CD Transcoding

Jin-soo Kim* *Regular Members*, Jae-Gon Kim**, Hyung-Myung Kim***, Jin-Woo Hong**
Regular Members

요약

최근에 FD(Frame Dropping)-CD(Coefficient Dropping) 트랜스코딩기법은 계산량을 줄이고, 손쉽게 구현할 수 있다는 측면에서 주목을 받고 있다. 그러나, 기존의 기법에서는 CD로 인하여 발생하는 왜곡 특성이 GOP내의 부호화 의존성을 갖는 프레임의 복호화 영상 품질에 전파 및 누적되는 성질을 고려하고 있지 않다. 본 논문에서는 CD로 인하여 발생하는 왜곡을 수식적으로 분석하고, CD왜곡이 전파/누적됨을 모의실험을 통하여 분석한다. 또한, 실험적인 결과에 기초하여, CD에 의한 전파/누적 왜곡 특성을 지수 감소함수 형태로 모델링한다. 모의실험을 통하여, 제안된 모델은 CD로 인해 초래되는 전체 누적왜곡 크기를 잘 예측할 수 있음을 보인다.

Key Words : FD(Frame Drop), CD(Coefficient Drop), Transcoding

ABSTRACT

Recently, FD (Frame Dropping)-CD (Coefficient Dropping) transcoding is considered due to the low computational complexity and simple implementation. The conventional FD-CD transcoding schemes have not considered the CD errors that tend to be propagated and accumulated. In this paper, we derive the error characteristics incurred by the CD errors and, through computer simulations, we show that the CD errors are propagated and not negligible to the decoded qualities of the next frames within single GOP. Then, we propose an exponential-decaying model that describes well the characteristics of propagation/accumulation errors. Finally, it is shown that the proposed model can be effectively used for estimating the overall distortions incurred by the CD errors.

I. 서론

UMA(Universal Multimedia Access)의 환경에서는 이종의 망자원과 다양한 사용자 단말 특성에 맞추어 비디오 콘텐츠를 효과적으로 수용하는 기법은 필수적인 기술이다. 이를 위한 한 가지 방법으로서, 트랜스코딩 기법에 대한 연구가 활발히 진행되고 있다. 압

축 부호화된 비트열을 망자원 또는 사용자 단말 등에 의해 제한되는 특성에 적응시키기 위한 트랜스코딩 기법으로는 크게 세 가지 접근 방법으로 나눌 수 있다[1,2,3,4]. 첫 번째 접근 방법으로는 압축 부호화된 비트열을 완전히 복호화한 후에 화소 영역에서 움직임 추정/DCT 변환, 양자화, VLC부호화를 수행하는 방법(CPDT, Cascaded Pixel Domain Transcoder)이

* 한밭대학교 정보통신컴퓨터공학부(jskim67@hanbat.ac.kr) ** 한국전자통신연구원 방송미디어연구그룹(jgkim@etri.re.kr)

*** 한국과학기술원 전기및전자공학전공(hmkum@kaist.ac.kr)

** 한국전자통신연구원 방송미디어연구그룹(jwhong@etri.re.kr)

논문번호 030523-1125, 접수일자 2003년 11월 25일

다 이 방법은 다른 접근 방법에 비해 성능이 뛰어나다는 장점이 있으나, 연산량 절감에 있어서는 매우 불리한 접근 방법이다[1]. 두 번째 접근 방법으로는 VLC 복호화를 수행한 후에 DCT영역에서 재양자화를 수행하는 방법(DDT: DCT-Domain Transcoder)이 있다. DCT영역에서 움직임 추정을 하거나 또는 재양자화를 선택적으로 수행하는 방법(OLT: Open-Loop Transcoder)은, 계산량을 CPDT에 비해 줄일 수 있으나, 다양한 응용으로 확장하는데 여전히 제한적이라는 점에서 단점이다[2] 마지막으로 VLC(Variable Length Code) 영역에서 FD(Frame Dropping)와 CD(Coefficient Dropping)을 수행하는 방법이 있다. 이 방법은 시간축 상에 위치하는 프레임들을 줄이는 소거 기법에 의한 FD와 DCT의 고주파 성분을 제거하는 CD에 의해 수행된다 이 기법은 성능 측면에서 앞에서 논한 방법들에 비해 우수하지 않지만, 연산량을 매우 절감시킬 수 있고, 구현이 간단하여 쉽게 사용할 수 있는 장점이 있다[3,4,5]

본 논문에서는 손쉽게 구현할 수 있고, UMA 환경에 적용할 수 있는 세 번째 접근 방식인 FD-CD 트랜스코딩 기법에 대해 다룬다. 기존의 연구에서 FD는 다른 프레임에게 예측 부호화를 제공하지 않는 B프레임을 가장 먼저 제거하고, 다음으로는 GOP내에서 가장 뒤 쪽에 위치한 P프레임 순으로 제거하는 기법을 취한다 그리고, CD를 수행함에 있어서는 고주파 성분을 구성하는 DCT 계수를 먼저 제거한다. 이때, CD의 적용은 FD와 달리 나중에 복호화되는 프레임에 왜곡을 전파/누적시키게 된다. 기존의 기법에서는 이러한 전파/누적 왜곡이 무시할만한 수준이고, 각 계수의 제거는 영상내 부호화 방식과 동일하게 취급하였다 그러나, 본 논문에서는, 실제 모의실험을 수행하여, CD로 인한 전파/누적 왜곡은 매우 큰 값이고, GOP내의 복호화된 영상 화질에 매우 큰 영향을 미치게 됨을 보인다. 또한, 모의실험 결과에 기초하여, 전파/누적 왜곡 특성을 지수 감소함수 형태로 모델링한다. 결론으로, 제안된 모델은 FD-CD 트랜스코딩 기법을 적용함에 있어, CD로 인해 각 프레임에 초래되는 전파 왜곡의 크기 및 누적된 전체 왜곡 크기를 매우 잘 예측할 수 있음을 보인다. 본 논문의 구성은 다음과 같다 우선 II절에서는 기존의 FD-CD 트랜스코딩 기법에 대해 간단히 살펴보고, III절에서는 CD에 의해 발생하는 왜곡 특성을 수식적으로 유도한다. 그리고, IV절에서는 모의실험을 통하여 CD왜곡이 전파/누적되는 현상을 모의실험을 통하여 확인하고, 이를 바탕으로, 전파/누적 왜곡에 대한 모델을 제시한다. V절

에서는 모의실험을 통하여 제안된 모델이 CD로 인해 초래되는 전체 왜곡 특성을 잘 예측할 수 있음을 보인다. 결론으로, VI절에서는 결론을 맺는다.

II. FD-CD 트랜스코딩 기법

FD는 압축 부호화된 몇 개의 프레임들을 제거함으로써 많은 양의 비트율을 제어할 수 있다 가장 중요도가 떨어지거나 또는 다른 부호화 프레임에 의존하지 않는 프레임들을 제거하면서, 우수한 품질을 유지하고자 한다. MPEG으로 압축 부호화된 비디오 시퀀스에서 B프레임은 다른 프레임에 대해 부호화 의존성을 제공하지 않으므로 가장 먼저 제거된다. B프레임이 제거된 후에는 P프레임이 우선 제거대상이 되며, GOP(Group of Pictures)내에서 가장 나중에 나오는 P프레임 순으로 제거될 수 있다 이렇게 함으로써, 부호화 의존성을 제공하지 않는 프레임이 제거되어, 트랜스코딩된 영상의 품질을 최대한 우수하게 할 수 있다.

CD에 의한 트랜스코딩은 FD에 비해 적은 양의 데이터를 세밀하게 제어할 수 있는 수단이 된다 CD 기법은 중요하지 않은 고주파 DCT계수의 VLC를 제거함으로써 비트율을 제어하는 기법이다. VLC코드를 제거하는 방법으로는 그림1에 나타난 바와 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, 그림1.(a)에 의한 방식은 CD를 수행할 때, 압축된 비트열에 대한 문법 구조 변화를 최소로 유지하므로, 더욱 쉽게 구현될 수 있다

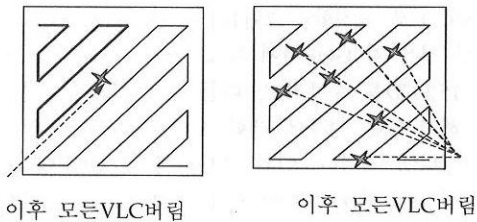
다음은 CD가 적용되는 프레임에 대한 기존의 비트율 제어기법에 대해 간단히 알아본다. 할당된 목표 프레임 비트량을 맞추기 위해, 각 DCT블록에서 버려지는 DCT계수 양을 균등하게 하여 CD하는 기법(URCD: Uniform Rate-based CD)이 있다. 그러나, 이 기법은 초래되는 왜곡의 크기를 최소화하지 않는 불합리한 접근 방법이다. 이에 비해 한 프레임 기간 동안의 전체 DCT블록에 최소의 왜곡이 초래되도록 각 블록에 대해 최적의 EOB 지점을 찾는 기법(LOCD Lagrange Optimization CD)이 있다[4,5] 즉, 프레임내에 N개의 DCT블록이 존재한다고 할 때, i번째 DCT블록이 K개의 non-zero DCT계수를 포함한

다면, LOCD 기법은 다음과 같다.

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

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

한편, 그림2는 이상에서 논의된 FD-CD에 의한 트랜스코딩 기법을 개괄적으로 설명하는 예를 나타내고 있다. 그림2의 예에서는 B프레임들을 FD하여 프레임 율을 줄이고, I/P프레임을 CD하여 비트량을 제어하는 경우를 나타낸다.



(a) 순차적 제거 기법 (b) 비중요도순 제거기법
그림 1. CD에 대한 두 가지 제어 방식

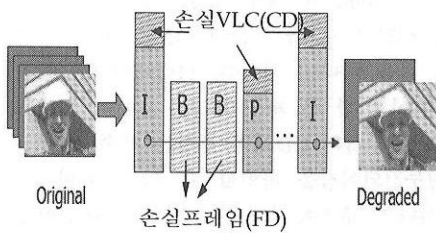


그림 2. FD-CD에 의한 트랜스코딩에 대한 개념도

III. CD에 의한 전파 왜곡 분석

우선, 본 논문에서는 전개의 편리함을 위해 CD 제어 방식은 그림1.(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.

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

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, $j-1$ 번째 프레임의 i 번째 블록에 대한 복호화된 표현($\hat{y}_j(i)$)은 식(2)에 기초하여 다음과 같이 나타낼 수 있다.

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

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)로 부터, CD에 의한 트랜스코딩을 수행함으로써 초래되는 j 번째 프레임의 i 번째 블록에 대한 왜곡($D_{j,i}(b_i)$)은 다음과 같이 표현할 수 있다.

$$D_{j,i}(b_i) = \|y_j(i) - \hat{y}_j(i)\| = \|a_j(i) + e_{j,i} - \hat{e}_{j,i}\| \quad (4)$$

여기서, $b_i(0 \leq b_i \leq N-1, N=64)$ 는 CD를 수행함에 따라 새로이 결정되는 i 번째 DCT블록의 EOB위치를 나타낸다. 그리고, $a_j(i)$ 는 $j-1$ 번째 프레임에서 CD를 수행함으로써, j 번째 프레임의 i 번째 블록으로 전파되는 움직임 보상 오차인 $MC_i(y_{j-1}) - MC_i(\hat{y}_{j-1})$ 를 부호화스캔(예, zig-zag 스캔) 순서에 의해 2차원 데이터 표현에서 1차원 데이터 표현으로 나타낸 것이다. 식 (4)를 DCT영역의 표현으로 $A(k) = DCT_k\{a_j(i)\}$, $E(k) = DCT_k\{e_{j,i}\}$, 그리고 $\hat{E}(k) = DCT_k\{\hat{e}_{j,i}\}$ 으로 각각 표기하면, 식(4)에 나타낸 왜곡은 아래와 같이 표현된다.

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

식(5)에서 $E(k)$ 에서 b_i 이후에 있는 DCT계수 성분들은 모두 CD되므로, 이 계수들에 의해 차이에 의해 왜곡이 발생한다. 따라서, 식 (5) 우측의 세 번째 항은 j 번째 프레임의 i 번째 블록에서 CD를 수행함으로써 발생하는 왜곡 성분이다 즉, b_i 의 위치에 따라 크기가 결정되는 왜곡 성분이다. 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 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

한편, 식(5)의 첫 번째 항과 두 번째 항은 다시 회귀적인 형태의 분석이 가능하다 즉, $\hat{y}_{j-1}(i)$ 는 $j-1$ 번째 프레임의 CD에 의한 왜곡, $j-2$ 번째 프레임에서 전해져 오는 왜곡성분, 그리고 이 두 가지 오차신호 성분의 상관관계에서 발생하는 왜곡으로 나눌 수 있다 이와 같은 분석을 연속적으로 j 번째 프레임까지 확장하면, $j-2$ 번째 프레임에서 CD에 의한 왜곡이 $j-1$ 번째 프레임에 전파되고, 그것이 움직임 보상의 방법으로 j 번째 프레임까지 누적 왜곡(Accumulation Error)으로 전달될 수 있다

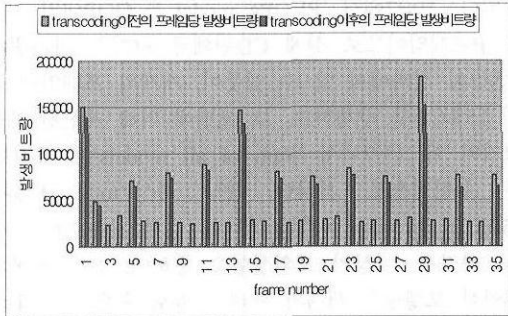
IV. CD에 의한 전파/누적 왜곡 모델링

III절에서 논의된 CD에 의한 전파/누적 왜곡 분석을 위한 모의실험을 수행한다 1.5Mbps로 부호화된 CIF해상도의 Foreman 시퀀스를 사용한다 비트 스트림의 부호화 특성은 MPEG-4 simple profile, 1~300 frames, IBBPBBPBB 형태(GOP=15, SubGOP=3)이다 본 논문에서는 I/P프레임만을 CD의 대상으로 고려하고, B프레임은 FD의 대상으로 한다. 트랜스코딩 이전과 이후의 비트율을 각각 R과 R'으로 나타낼 때, 그림3의 결과는 R=1.5Mbps, R'=800kbps으로 트랜스코딩한 후의 IPPPP 시퀀스 순서로 P프레임에 초래된 트랜스코딩후의 평균자승오차(MSE mean square error)를 나타낸 것이다 그림3은 R-R'만큼의 양을 IPPPPP 시퀀스에 균등하게 할당하는 일정률 CD를 할당하고 제어는 식(1)의 LOCD를 사용한다.(그림3.(a)) 그림3 (b)에서 알 수 있듯이, CD로 인하여 왜곡의 전파가 매우 크게 발생함을 알 수 있으며, 전파/누적 왜곡의 크기는 각 프레임의 CD 왜곡과 거의 비슷한 크

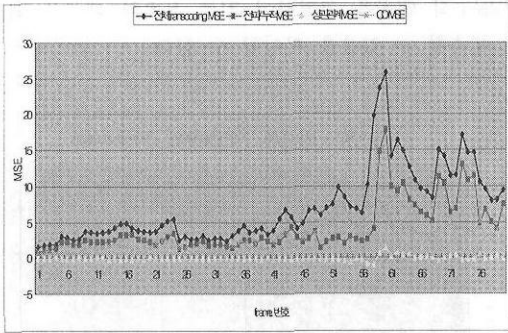
기로 발생함을 확인할 수 있다. 또한, 같은 GOP내에서 앞선 프레임에 비해 상대적으로 뒤쪽에 위치하는 프레임에서 전파 왜곡이 누적되는 특성으로 인해 상대적으로 크게 나타남을 알 수 있다. 그림3에서 한 가지 주목해야할 실험적 결과는 현재 프레임에서 수행하는 CD 왜곡과 이전 프레임에서 전달되어 오는 왜곡사이의 상관관계에 의한 왜곡[식 (5) 우측의 두 번째 항]은 무시할 수 있을 정도임을 알 수 있다. 따라서, 이전 프레임에서 전해져오는 전파/누적 왜곡과 현재 프레임에 적용되는 CD 왜곡사이에는 상관관계가 존재하지 않고, 독립적인 경우로 가정할 수 있다

개개 프레임 별 CD 왜곡 성분이 이전 프레임(들)에서의 CD로 인해 전해져오는 전파/누적 왜곡 특성과의 관계를 알아보기 위해 모의실험을 수행하여 그 결과를 나타낸 것이 그림4이다. 그림4에서 먼저, R=1.5Mbps로 압축된 Foreman 시퀀스를 IPPPP...구조로 트랜스코딩할 때 모든 프레임에 대해 균등하게 non-zero DCT계수의 15%를 CD한다('total VOP's 15%CD'로 표시함). 그림4에서 'IVOP_15%CD'로 표시한 것은 각 GOP내에 독립적으로 I프레임에 대해서만 15%CD를 수행하고, 다른 프레임들은 트랜스코딩을 하지 않은 경우에 대한 모의실험 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 왜곡 특성을 나타낸다 (그림4에서 각각 'PVOP1_15%CD', 'PVOP2_15%CD', 'PVOP3_15%CD', 'PVOP4_15%CD'로 표시함). 그림4에서 'Sum of VOP 15%CD'로 표시한 것은 개개의 프레임 별로 독립적으로 15% CD를 수행한 결과들의 합을 나타낸 것이다. 그림에서 알 수 있듯이, I프레임이 가장 큰 왜곡 특성을 보이며, I프레임과 다른 P프레임들의 CD왜곡 특성은 지수 감소함수형태로 전파/누적되어짐을 확인할 수 있다 또한, 그림4에서 또 하나의 주목할 특성은 개개의 프레임별로 독립적으로 CD를 수행한 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 이상의 모의실험 결과인 그림3과 그림4을 통하여, CD 왜곡의 전파/누적 특성은 개개의 프레임 단위로 독립적인 전파 특성을 나타내고, 각 프레임에 첨가되는 CD왜곡[식 (5) 우측의 세 번째 항]과 이전 프레임에서 전파/누적되는 왜곡 [식 (5) 우측의 첫 번째 항]사이에는 독립적인 관계를 갖는다고 가정할 수 있다. 그러므로, 개개의 프레임

단위로 수행되는 CD 특성을 모델링하여 전체 왜곡을 근사적으로 예측 가능하게 한다.



(a) 일정률 할당 CD 제어방식



(b) 전체 왜곡 및 각 왜곡 성분

그림 3. 일정률할당 CD제어 방식과 트랜스코딩 왜곡

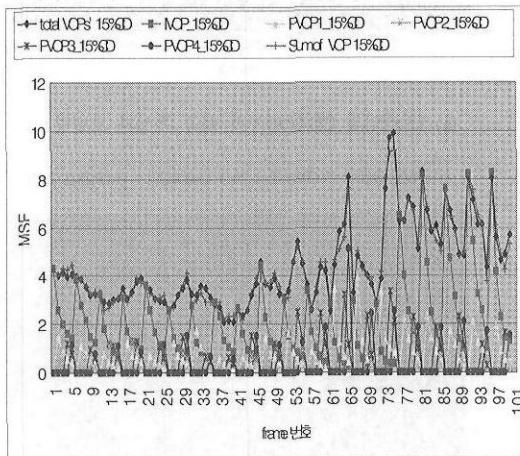


그림 4. 통합 15%CD의 경우와 개개 프레임별 15%CD 경우의 왜곡 특성 비교

이상에서 관찰된 실험적 특성을 바탕으로, 본 논문에서는 개개의 프레임에 대한 CD왜곡의 전과 특성은 모두 지수 감소함수 형태로 표현하고, 각각의 독립된 구간으로 분리하여 접근한다. 즉, 그림5에 나타낸 것

과 같이 각 구간별로 전과 왜곡의 지수함수의 감쇄상수는 동일하고, 또한, 일정하다고 가정한다. 먼저, 그림5에 나타낸 I프레임의 CD에 의해 D_I 만큼의 왜곡이 발생한 경우에 이후에 나오는 P프레임에 전과 왜곡은 연쇄적으로 다음과 같이 표현할 수 있다.

$$\begin{aligned} \hat{D}_I(0) &= D_I e^{-\alpha_1} \\ \hat{D}_I(2) &= \hat{D}_I(1) e^{-\alpha_2} = D_I e^{-\alpha_1 - \alpha_2} \\ &\dots \\ \hat{D}_I(n) &= \hat{D}_I(n-1) e^{-\alpha_n} = D_I e^{-\sum_{i=1}^n \alpha_i} \end{aligned} \quad (6)$$

그리고, GOP내의 첫 번째 P프레임에 CD에 의해 D_{P1} 만큼의 왜곡이 발생한 경우에 이후에 나오는 P프레임에 대한 전과 왜곡은 다음과 같이 표현된다.

$$\hat{D}_{P1}(n) = \hat{D}_{P1}(n-1) e^{-\alpha_n} = D_{P1} e^{-\sum_{i=2}^n \alpha_i}, n=2,3,4 \quad (7)$$

이와 유사하게 진행하면, GOP내의 두 번째와 그 이후의 P프레임에 대한 전과 왜곡 특성은 다음과 같이 표현된다.

$$\hat{D}_{P2}(n) = \hat{D}_{P2}(n-1) e^{-\alpha_n} = D_{P2} e^{-\sum_{i=3}^n \alpha_i}, n=3,4,\dots \quad (8)$$

$$\hat{D}_{P3}(n) = \hat{D}_{P3}(n-1) e^{-\alpha_n} = D_{P3} e^{-\sum_{i=3}^n \alpha_i}, n=4,\dots \quad (9)$$

식(6)-(9)의 정의식에 의해 GOP내의 복호화 순서에 따른 CD왜곡과 전과왜곡을 고려한 전체 누적왜곡은 다음과 같이 각각 표현된다.

I프레임: D_I

GOP내 1번째 P프레임: $D_I e^{-\alpha_1} + D_{P1}$

GOP내 2번째 P프레임:

$$D_I e^{-\alpha_1 - \alpha_2} + D_{P1} e^{-\alpha_2} + D_{P2}$$

GOP내 3번째 P프레임:

$$D_I e^{-\sum_{i=1}^3 \alpha_i} + D_{P1} e^{-\sum_{i=2}^3 \alpha_i} + D_{P2} e^{-\sum_{i=3}^3 \alpha_i} + D_{P3}$$

GOP내 4번째 P프레임:

$$\begin{aligned} D_I e^{-\sum_{i=1}^4 \alpha_i} + D_{P1} e^{-\sum_{i=2}^4 \alpha_i} + D_{P2} e^{-\sum_{i=3}^4 \alpha_i} + D_{P3} e^{-\sum_{i=4}^4 \alpha_i} + D_{P4} \\ \dots \end{aligned} \quad (10)$$

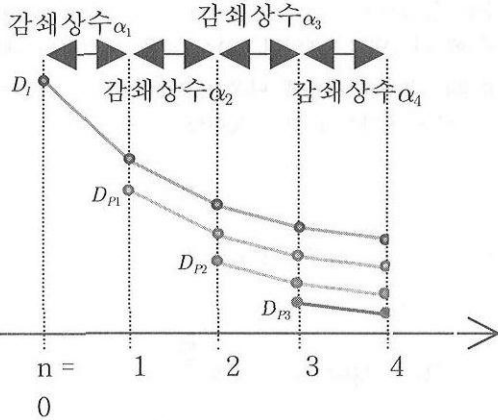


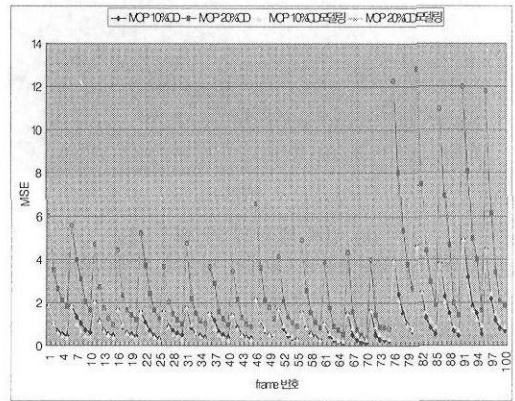
그림 5. CD 전파 왜곡의 모델링의 개념적인 그래프

그림6은 I프레임과 GOP내의 첫 번째 P프레임에 대하여 일정률 CD[10%,20%의 경우를 모의실험 수행한 경우에 대한 왜곡 특성을 나타낸다. CD에 의한 왜곡이 상대적으로 큰 경우(20%CD)에 감쇄상수는 미세하게 큰 값을 나타내지만, 10%의 CD를 적용한 경우와 유사한 지수 감소함수 형태를 보이고 있다. 본 논문에서 식(6)~(9)로 제안한 지수 감소함수에 대한 모델링 파라미터를 $\alpha_1=0.6$, $\alpha_2=0.47$, $\alpha_3=0.31$, $\alpha_4=0.27$ 로 하여 적용한 경우에 대한 CD왜곡의 전파 특성을 나타내었다. 이 그림에서 알 수 있듯이, 본 논문에서 제안된 모델은 각 프레임별로 CD 왜곡 크기에 대해 이로 인해 전파되는 왜곡 특성을 잘 예측하고 있음을 볼 수 있다.

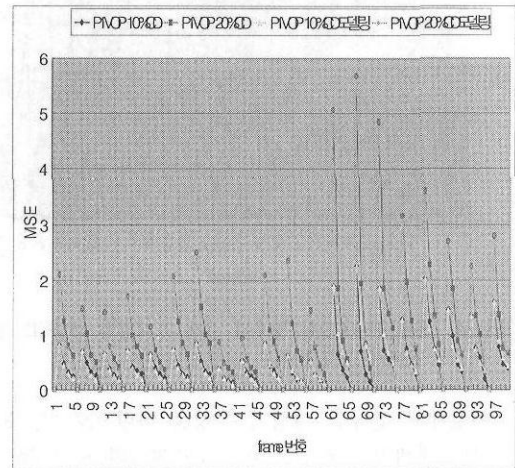
Foreman 시퀀스에 대해 5% ~ 30%CD를 모든 프레임에 대해 균일하게 적용한 경우에 대해 CD 왜곡 크기의 전파 특성을 조사하면, 표1과 같은 결과를 얻는다. I프레임에 의한 CD 왜곡이 GOP내의 첫 번째 P프레임으로 전파 효과가 작게 나타나지만, 전반적으로 거의 일정한 감쇄 상수를 갖는다는 사실을 확인할 수 있다. 또한, GOP내의 끝에 위치하는 P프레임에 가까울수록 감쇄상수는 작은 값이어서, 전파 효과가 상대적으로 크게 나타난다.

이상의 논의를 바탕으로, 전체 전파/누적 왜곡을 예측하기 위해, 먼저, 모든 프레임에 대해 15%CD를 통합적으로 수행한 경우에 대해 복호화하여 트랜스코딩으로 발생한 전체 왜곡을 조사한다. 이 결과를 그림7의 'total VOP's 15%CD'로 표시한다. 그리고, 개개의 프레임에 대해 15%CD를 수행하고, 이로 인해 발생하는 전파/누적 왜곡은 각 구간별 모델링 파라미터로서 $\alpha_1=0.6$, $\alpha_2=0.47$, $\alpha_3=0.31$, $\alpha_4=0.27$ 을 사용하

여 전파/누적 왜곡의 특성값을 얻는다. 그림7의 'IVOP model', 'PIVOP model', 'P2VOP model', 'P3VOP model', 그리고 'P4VOP model'로 각각 표시한다. 그림7에서 'PVOP4 model'은 GOP내의 마지막 P프레임이므로 실제 CD왜곡 크기를 나타낸다. 그림7의 결과에서 알 수 있듯이, 개개의 프레임에 대해 15%CD를 수행하고, 그 모델에 의해 계산된 왜곡 크기의 합(그림7에서 'Sum of all models'로 표시함)은 통합적으로 처리한 왜곡 특성('total VOP's 15%CD')과 매우 유사한 결과를 얻게 됨을 알 수 있다. 따라서, 이상의 모의실험을 통하여, 본 논문에서 제안한 모델링은 개개의 프레임 별로 수행되는 CD왜곡의 전파 특성과 더불어 트랜스코딩으로 초래되는 전파/누적 왜곡을 고려하여 전체 복호화 왜곡을 적절히 예측할 수 있음을 확인할 수 있다.



(a) I프레임의 CD전파/누적 왜곡 특성과 모델링



(b) GOP내 첫 P프레임의 CD전파/누적 왜곡 특성과 모델링
그림 6. CD 왜곡의 전파/누적 왜곡과 모델링 결과

이상의 모의실험은 모두 Foreman 시퀀스에 대해 수행한 결과이다. Akiyo, Stefan, Coastguard 시퀀스에 대해서도 Foreman 시퀀스에 대한 것과 같이 모의실험을 수행한다. Akiyo, Stefan, Coastguard 시퀀스에 대해서 각각 1.2Mbps, 1.4Mbps, 1.5Mbps로 압축한 후에, IPPPP... 구조로 B프레임은 FD를 적용하고, I/P 프레임은 5%/10%/15%/20%/25%/30%로 일정률의 CD를 적용하여 모의실험을 한다. 그림8은 정규화된 CD 왜곡의 전파 특성을 나타낸 그래프이다. Stefan, Coastguard 시퀀스는 Foreman 시퀀스와 유사한 왜곡 전파특성을 보이고 있으나, Akiyo 시퀀스의 경우에는 CD에 의한 왜곡 특성이 매우 크게 나타남이 주목된다. 그 이유는 Akiyo 시퀀스의 경우에는 skipped-MB로 부호화된 매크로블럭의 개수가 많이 발생하여, 다른 시퀀스에 비해 왜곡의 전파가 크게 나타난다. 반면에, Stefan 시퀀스는 움직임이 매우 많은 시퀀스이며, 이것은 intra-coded-MB가 상대적으로 많이 관측되며, 이 경우에는 CD로 인해 발생하는 왜곡의 전파 효과가 다소 낮게 되는 이유가 된다.

표 1. Foreman 시퀀스에 대한 FD-CD 적용에 따른 파라미터의 변동 범위

α_1	α_2	α_3	α_4
0.61~0.55	0.47~0.45	0.31~0.29	0.27~0.26

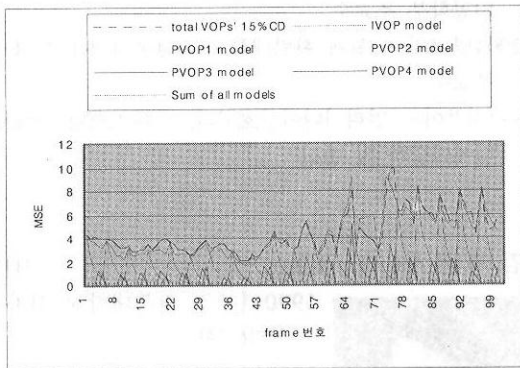


그림 7. 각 프레임에 대한 일정률 15%CD 할당과 모델링 비교

그림9는 skipped-MB로 부호화된 매크로블럭의 전파왜곡 성분을 제거한 경우에 전파 왜곡 특성을 나타낸 것이다. 본 논문에서는 다음과 같이 간단한 예측 기법을 적용한다.

$$\hat{C}_M(n-1) = \hat{D}_M(n-1) \left(1 - \frac{N_{skippedMB}}{N_{MB}}\right) \quad (11)$$

$$\hat{C}_M(n) = \hat{C}_M(n-1) e^{-\alpha_n} \quad (12)$$

여기서, M은 식(6)~(9)에서의 인덱스 I, P1, P2, P3, ... 등을 지칭하는 것이고, N_{MB} 과 $N_{skippedMB}$ 는 단일 프레임내의 전체 매크로블럭의 개수와 skipped-MB의 개수를 각각 나타낸다. 그림9에서는 식(11)과 (12)에서의 절차를 거쳐 skipped-MB에 의한 전파왜곡을 제거한 경우에 대한 왜곡을 나타낸 것이다. 이 경우에 4개의 비디오 시퀀스별로 유사한 왜곡의 전파특성을 보인다.

표2에서는 skipped-MB 성분을 제거한 후의 평균적인 모델링 파라미터를 보여주고 있다. 표2에서 avg_intraMB, avg_interMB, avg_skippedMB는 각 압축 시퀀스별로 P프레임당 매크로블럭의 평균적인 부호화모드 개수 즉, intra-coded, inter-coded, skipped 매크로블럭의 개수를 나타낸다.(본 논문에서는 모두 CIF 해상도를 사용함.) 4개의 테스트 시퀀스에서 파라미터 값의 차이는 intra-coded-MB의 개수가 많은 경우에 감쇄상수가 다소 큰 값을 갖게 된다는 것을 확인할 수 있다. 매크로블럭의 부호화모드의 비율에 적절히 파라미터를 선택할 수 있을 것이다.

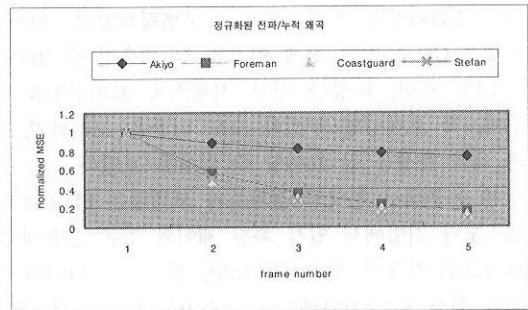


그림 8. 정규화된 CD 왜곡의 전파 특성

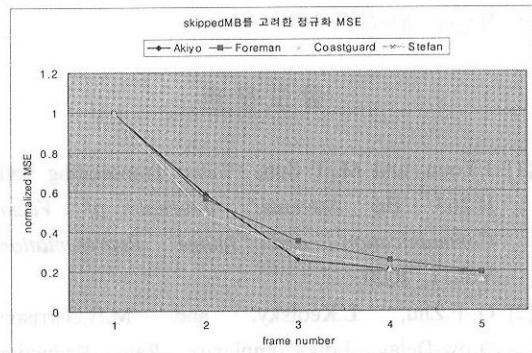


그림 9. skipped MB 성분을 제거한 경우의 정규화된 CD 왜곡의 전파 특성

표 2. skipped-MB성분 제거 후의 평균적인 모델링 파라미터

	avg_in traMB	avg_in terMB	avg_skip pedMB	α_1	α_2	α_3	α_4
Foreman	7.3	384.3	2.7	0.61~ 0.55	0.47~ 0.44	0.31~ 0.29	0.27~ 0.26
Akiyo	0	270.6	125.5	0.59~ 0.54	0.67~ 0.44	0.29~ 0.27	0.1~ 0.09
Coastguard	0.26	395.7	0.04	0.72~ 0.57	0.49~ 0.45	0.33~ 0.28	0.28~ 0.26
Stefan	5.0	376.8	14.2	0.76~ 0.54	0.64~ 0.51	0.58~ 0.35	0.52~ 0.34

VI. 결론

본 논문에서는 FD-CD트랜스코딩 기법에서 CD에 따른 전파 왜곡 발생을 수식적으로 논하였고, 그 결과를 모의실험을 통하여 분석하였다. 모의실험을 통하여 CD로 인한 전파/누적 왜곡은 무시할 수 없을 정도의 매우 큰 왜곡을 발생시킴을 보였고, 또한, CD로 인한 왜곡은 이전 프레임에서 전파되어 오는 왜곡과 상관성이 없음을 확인하였다. 이 결과를 바탕으로 독립적인 지수 감소함수형태로 개개의 프레임별로 발생되는 CD왜곡의 전파 특성을 모델링하였고, 제안된 모델은 CD로 인한 전체 왜곡을 잘 예측할 수 있음을 보였다. 또한, 다양한 영상 시퀀스에 모의실험을 한 결과, 각 프레임별 부호화 모드 비율을 고려하면, 전파 왜곡 특성을 잘 묘사할 수 있음을 보였다.

본 논문에서 분석된 결과들과 모델링은 FD-CD 트랜스코딩 기법에서 일정 화질 제어와 같은 응용에 효과적으로 사용될 수 있다. 또한, 앞으로는 다양한 부호화 형식 또는 B프레임을 포함하는 CD 트랜스코딩 기법에 대하여 연구를 수행하여 파라미터의 특성에 대해 조사할 필요가 있으며, 프레임별 적응적으로 전파 특성을 계산하는 연구가 필요하다.

참고 문헌

[1] J.Youn, and M.-T Sun, "Video Transcoding with H.263 Bit Streams," *Journal of Visual Communication and Image Representation*, Vol.11, 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 JTC1/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.

김진수 (Jin-soo Kim)

정회원



1991년 2월 경북대학교 전자공학(공학사)
1993년 2월 한국과학기술원 전기 및 전자공학(공학석사)
1998년 8월 한국과학기술원 전기 및 전자공학(공학박사)
1995년 3월~2000년 3월 삼성전자 시스템LSI사업부 선임연구원

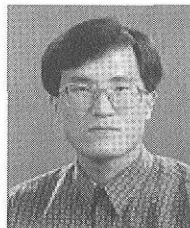
2000년 4월 ~ 2004년 3월 한밭대학교 정보통신컴퓨터공학부 조교수

2004년 4월 ~ 현재 한밭대학교 정보통신컴퓨터공학부 부교수

<관심분야> 멀티미디어신호처리, 트랜스코딩, 멀티미디어다중화 및 통신

김재곤 (Jae-Gon Kim)

정회원



1990년 2월 경북대학교 전자공학(공학사)
1992년 2월 한국과학기술원 전기 및 전자공학(공학석사)
2001년 9월~2002년 11월 미국 Columbia 대학 방문 연구원
1992년 3월~현재 한국전자통신연구원(선임연구원)

2004년 6월~현재 한국전자통신연구원 방송컨텐츠연구팀장

<관심분야> 비디오신호처리, 멀티미디어 프레임워크, 디지털방송

