

# FD/CD 트랜스코딩 기법에서 CD에 의한 전파/누적왜곡에 대한 모델링

\*김진수 · \*\*김재곤 · \*\*강경욱 · \*\*김진웅

\*한밭대학교 정보통신컴퓨터공학부 · \*\*한국전자통신연구원 방송미디어연구부

\*E-mail : [jskim67@hanbat.ac.kr](mailto:jskim67@hanbat.ac.kr)

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

\*Jin-soo Kim · \*\*Jae-Gon Kim · \*\*Kyeongok Kang · \*\*Jinwoong Kim

\*Hanbat National University, \*\*Electronics and Telecommunications Research Institute

### 요 약

다양한 종류의 사용자 단말과 더불어 이종 망의 도입에 따라, 디지털 영상컨텐츠를 효과적으로 제공하기 위해 트랜스코딩에 대한 연구가 활발히 진행 중에 있다. 특히, FD(Frame Dropping)/CD(Coefficient Dropping) 트랜스코딩기법은 계산량을 줄이고, 손쉽게 구현할 수 있다는 측면에서 주목을 받고 있다. 그러나, 기존의 연구에서는 CD로 인하여 발생하는 왜곡 특성이 GOP내의 다른 복호화 영상 품질에 전파 및 누적되는 성질을 고려하지 않고 있다. 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 모델링될 수 있음을 보이고, 모의실험을 통하여, 제안된 모델링은 CD로 인해 발생하는 전체 복호화 왜곡 크기를 잘 예측하는데 사용할 수 있음을 보인다.

### 1. 서 론<sup>1)</sup>

다양한 사용자 단말의 사용과 더불어 이종의 망간에는 망자원 특성이 매우 유동적이다. 이와 같은 환경에 적응적으로 디지털 비디오 컨텐츠를 효과적으로 수용하는 기술은 필수적이며, 이를 해결하기 위한 한 방법으로서 트랜스코딩 기법에 대한 연구가 활발히 진행되고 있다. 압축 부호화된 비트열을 망 자원 또는 사용자 단말 등의 제한적 특성에 맞추기 위한 트랜스코딩 기법으로는 크게 세가지 기법으로 분류된다<sup>1),2),3),4)</sup>. 첫 번째 기법으로는 압축 부호화된 비트열을 완전히 복호한 후에 화소 영역에서 움직임 추정/DCT 변환, 양자화, VLC부호화를 수행하는 방법(CPDT: Cascaded Pixel Domain Transcoder)이 있다. 이

기법은 다른 기법에 비해 성능이 뛰어나다는 장점이 있으나, 계산량의 절감 및 구현 측면에 있어서 매우 불리한 접근 방법이다<sup>1)</sup>. 두 번째 기법으로는 VLC부호화를 수행한 후에 DCT영역에서 양자화를 다시 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 응용 환경으로 확장하는데 여전히 단점이 있다<sup>2)</sup>. 마지막으로 VLC 영역에서 수행하는 FD(Frame Dropping)와 CD(Coefficient Dropping)를 수행하는 방법이 있다. 이 방법은 시간축상에 위치하는 프레임울 소거 기법에 의한 FD와 부호화된 DCT 영역의 고주파 성분을 제거하는 CD에 의해 수행된다. 이 기법은 성능 측면에서 앞에서 논한 방법에 비해 우수하지 않지만, 계산량을 매우 절감시킬 수 있

1) 본 연구는 정보통신부 과제의 일부로서 수행되었습니다.

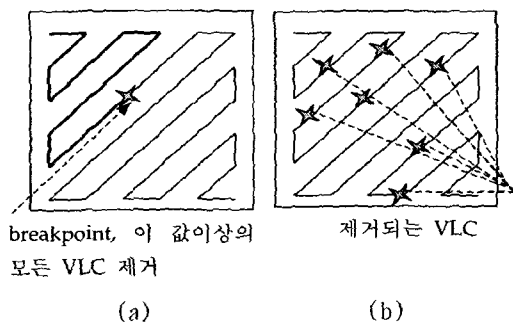
고, 손쉽게 구현하여 사용할 수 있다는 장점이 있다<sup>13,14,51</sup>.

그러나, 기존의 FD-CD기법에 의한 트랜스코딩 기법은 전파에러와 CD에 의한 에러 효과를 적절히 분석하지 못하고 있다<sup>14,51</sup>. 실제, 각 프레임별로 CD를 수행함에 따라 예측 부호화를 사용하는 영상 압축 원리에 따라, 왜곡은 전파되고, 또한, 누적되게 된다. 본 논문에서는 CD에 의해 초래되는 전파/누적 왜곡을 수식적으로 분석하고, 그 효과를 모의실험을 통하여 분석한다. 또한, 다양한 응용처를 위하여 CD를 수행함에 따라 발생하는 왜곡의 크기를 효과적으로 예측할 수 있는 모델을 제시한다.

## II. FD/CD 트랜스코딩 기법

FD는 압축 부호화된 몇 개의 프레임을 제거함으로써 많은 양의 비트율을 제어할 수 있다. 가장 중요도가 떨어지거나 또는 다른 부호화 프레임이 의존하지 않는 프레임을 제거하면서 가능한 우수한 품질을 유지하기 위한 방법으로 고려된다. MPEG에 의한 압축 부호화된 비디오 시퀀스에서는 B프레임은 다른 프레임에 대해 부호화 의존성을 제공하지 않으므로 가장 먼저 제거된다. B프레임이 제거된 후에는 P프레임이 제거대상이 되며, GOP(Group of Pictures)내에서 가장 나중에 나오는 P프레임 순으로 제거될 수 있다.

CD에 의한 트랜스코딩은 FD에 비해 적은 양의 데이터를 제어할 수 있는 수단이 된다. CD기법은 중요하지 않는 고주파 DCT계수에 대응하는 VLC를 제거함으로써 비트율을 제어하는 기법이다. VLC 코드를 제거할 수 있는 기법으로는 ECB코드에서부터 순차적으로 제거하는 방법(그림1(a))과 단일 블록내에서 위치에 상관없이 중요도가 낮은 계수를 없애는 방법(그림1(b))이 있을 수 있다. 참고문헌[4,5]에서는 그림1의 두 가지 방식에 대한 화질 변화는 매우 미미한 수준으로 보고 하고 있다. 실제 구현에 있어서는 가능한 한 비트열 문법 구조의 변화를 주지 않는 방법을 사용하는 것이 구현의 간단함에 기여할 것이다.



(a) (b)  
그림1. CD 제어 방식

한 프레임 내에서 CD를 이용한 비트율 제어 방식은 크게 두 가지 방식으로 접근할 수 있다. 각각의 블록에 초래되는 왜곡을 고려하지 않고, 목표 비트율을 만족할 수 있도록 동일한 비율로 고주파 계수들에 대한 VLC계수를 제거하는 기법(URCD: Uniform Rate-based CD)과 한 프레임의 최적화 윈도우 내에서 각 블록내의 CD계수에 대한 최적의 Breakpoint를 찾는 기법(LOCD: Lagrange Optimization CD)이 있다<sup>14,51</sup>. LOCD는 CD로 초래되는 왜곡을 최소화하기 위해 Lagrangian 탐색을 사용한다. 본 논문에서는 성능 비교를 위하여 LOCD에 의한 기법에 의한 비트율 제어를 수행하고, CD제어 방식으로 그림1(a)을 사용한다.

한편, 참고문헌 [4,5]에서는 비트율 제어에 있어, 기억 소자를 갖지 않는 모델을 사용하여 부호화된 프레임들간에 CD에 의한 왜곡은 매우 작다고 가정하여 고려하지 않았다. 그리고, 실험 결과에 따르면 약 0.3dB정도 이내의 품질 저하를 갖게 됨으로 인하여, CD에 의한 전파 왜곡과 누적 왜곡은 화질에 거의 영향을 주지 않는다고 가정을 하였다. 모든 프레임은 독립적인 영상내 부호화 계수와 동일하게 취급하였다.

## III. CD에 의한 전파 왜곡

본 논문에서는 전개의 편의성을 위해 CD에 의한 비트율 제어 방식은 그림1.(a)에 나타난 방법을 사용한 것으로 가정한다(일반성을 잃지 않음). 이를 바탕으로, 부호화된 j번째 프레임의 i번째 블록에 대하여 CD를 행함으로써 초래되는 왜곡을 표현해보자. 이를 위해, 먼저, 트랜스코딩 왜곡없이 복호화된 형태의 표현( $y_j(i)$ )은 다음과 같이 표현한다.

$$y_j(i) = MC_j(y_{j-1}) + e_j \quad (1)$$

여기서,  $MC_j(y_{j-1})$ 는 j-1번째 프레임을 기준으로 움직임 보상된 j번째 프레임의 i번째 블록 성분 크기이고,  $e_j$ 는 움직임 예측 오차신호 성분이다. 만일, j-1번째 프레임과 j번째 프레임의 i번째 블록이 CD를 수행하였다면, 다음과 같이 표현할 수 있다.

$$\hat{y}_j(i) = MC_j(\hat{y}_{j-1}) + \hat{e}_j \quad (2)$$

식(1)과 (2)를 바탕으로, j번째 프레임의 i번째 블록에 대한 트랜스코딩 왜곡  $D_{j,i}(b_i)$ 에 대한 정의는 그림1.(a)에 나타난 DCT블록의 breakpoint 위치  $b_i$ 에 대한 함수로 아래와 같이 표현된다.

$$D_{j,i}(b_i) = \|y_j(i) - \hat{y}_j(i)\| = \|a_i + e_i - \hat{e}_i\| \quad (3)$$

여기서, zig-zag부호화 순서 형태의 의미에서  $a(i) = MC_j(y_j(i)) - MC_j(\hat{y}_{j-1}(i))$ 이다. 식(3)에 대한 DCT표현으로  $A(k) = DCT_k(a(i))$ 로 나타낸

다. 또한,  $E(k) = DCT_k(e(i))$  과  $\hat{E}(k) = DCT_k(\hat{e}(i))$  로 각각 표현하면, 왜곡은 아래와 같은 형태로 표현된다.

$$D_{j,i}(b_i) = \sum_{k=0}^{N-1} \{A(k) - E(k) - \hat{E}(k)\} \\ = \sum_{k=0}^{N-1} A^2(k) + 2 \sum_{k=b_i}^{N-1} A(k)(E(k) - \hat{E}(k)) + \\ \sum_{k=b_i}^{N-1} (E(k))^2 \quad (4)$$

식(4) 우측의 세 번째 항목은 j번째 프레임의 i 번째 블록에서 CD를 수행함으로써 발생하는 왜곡 성분이다. 즉,  $b_i$  위치에 따라 크기가 결정되는 왜곡이다. 반면에 식(4) 우측의 첫 번째 항은  $a(i) = MC_i(y_j(i)) - MC_i(\hat{y}_{j-1}(i))$  로 측정되는 왜곡으로서 j-1번째 프레임에서 예측 부호화기법을 사용함에 따라 전해져 오는 왜곡이다. 즉, j-1 번째 프레임의 트랜스코딩으로 인해 초래되는 왜곡 성분이다. 그리고, 식(4) 우측의 두 번째 항은 왜곡의 측정단위가 평균자승오차(Mean Square Error: MSE)임으로 인해 관측되는 것으로서, j-1 번째 프레임에서 움직임 보상에 의한 CD성분에 의한 왜곡과 j번째 프레임의 CD에 의한 왜곡 신호 성분간의 상관관계로 인해 발생하는 왜곡 성분이다.

한편, 식(4)의 첫 번째 항은 다시 회귀적인 형태의 분석이 가능하다. 즉,  $\hat{y}_{j-1}(i)$  는 j-1번째 프레임의 CD에 의한 왜곡, j-2번째 프레임에서 전해져 오는 왜곡성분, 그리고 이 두 왜곡 성분의 상관관계에서 발생하는 왜곡으로 나눌 수 있다. 이와 같은 분석을 연속적으로 확장하여 j번째 프레임까지 확대 해석하면, j-2번째 프레임에서 CD에 의한 왜곡이 j-1번째 프레임에 전파되고, 그것이 움직임 보상의 방법으로 j번째 프레임까지 누적 왜곡(accumulated error)으로 전달되기도 하고, 경우에 따라서는 j-1번째 프레임까지도 영향을 줄 수도 있다.

#### IV. CD왜곡과 모델링

CD왜곡의 특성을 모의실험을 통해 분석하기 위해, 1.5Mbps로 부호화된 CIF 해상도의 Forman 시퀀스를 사용한다. 비트스트림의 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.

각 GOP내에 비트율 할당에 따른 영향이 다를 수 있으므로 모든 프레임별로 일정 비율로 CD를 수행하는 방식으로 모의실험을 한다. 트랜스코딩 전/후의 비트율을 각각 R과 R'으로 나타낼 때, R-R'만큼의 양을 B프레임을 제거와 그 후에,

IPPPPP시퀀스에 일정하게 CD하도록 제어한다. 그림2의 결과는 R=1.5Mbps, R'=800kbps으로 트랜스코딩한 후의 IPPPP시퀀스에서 발생하는 평균자승오차를 나타낸 것이다. 그림에서 알 수 있듯이, 앞선 프레임에 의한 CD로 인한 과급효과가 매우 큼을 알 수 있다. 특히, 움직임이 많은 영역에서는 현재 프레임의 CD왜곡 보다도 이전프레임에서 전해져오는 왜곡이 더 크게 나타나는 것을 볼 수 있다. 또 한가지 더 주목해야할 실험적 결과는 현재 프레임에서 수행하는 CD 왜곡과 이전 프레임에서 전달되어 오는 왜곡과의 상관관계에 의한 왜곡은 두 방식 모두에서 무시할 만하다는 것을 알 수 있다. 이상에서 같은 GOP내에서 앞선 프레임에 비해 상대적으로 뒤쪽에 위치하는 프레임에서 전파왜곡 특성이 크게 나타남을 알 수 있다.

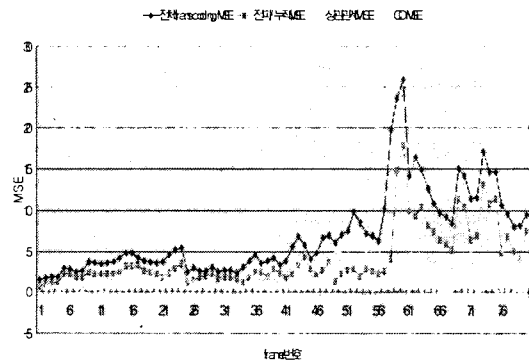
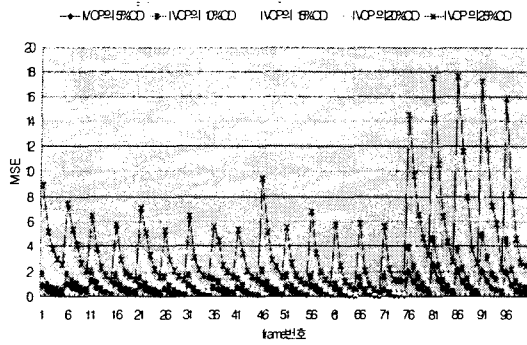


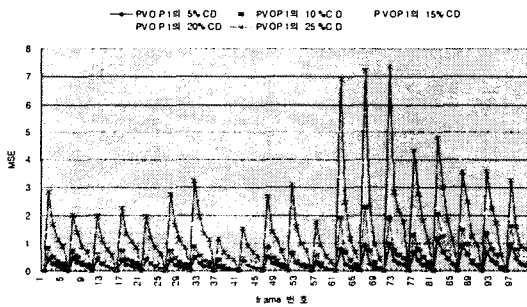
그림2. 일정을 CD제어 방식에 따른 왜곡 특성

그림3(a)에서는 I프레임에서 5%, 10%, 15%, 20%, 25%로 각각 CD를 하였을 경우 같은 GOP내에 전파되는 왜곡특성을 나타낸 것이다. 패턴은 동일한 형태가 나오지만 I프레임에서 CD를 많이 할수록 그 전파효과는 크게 나오며, 왜곡의 전파는 지수함수적으로 감소하는 형태를 보인다. 그림3(b)에서는 GOP내의 첫 번째 P프레임에서 CD를 행한 경우에 패턴은 동일한 형태가 나오지만, I프레임의 CD효과에 비해 상대적으로 작은 크기로 왜곡이 전파됨을 알 수 있다. 또한, 그림3(a)와 (b)에서, CD율이 클수록, 급한 감소기울기를 갖는다.

이상의 결과를 바탕으로 GOP내에 모든 프레임에 대해 모두 독립적으로 15%의 일정률로 CD한 결과를 나타낸 것이 그림4에 나타내었다. 여기서, 주목할 것은 독립적으로 수행하여 발생하는 왜곡에 대한 전체 합은 동시에 모든 프레임에 대해 15%를 수행한 것과 복호화 왜곡은 거의 동일하다는 것을 알 수 있다. 즉, 개개의 프레임단위로 수행되는 CD에 의한 왜곡과 이전 프레임에서 전파/누적되어 오는 왜곡과의 상관관계는 무시할 수 있다. 이와 같은 실험적인 결과를 바탕으로 개개의 프레임에 대해 첨가되는 CD 왜곡에 따른 전파/누적 왜곡은 지수 감소함수의 형태로 다음과 같이 모델링될 수 있다.



(a) I프레임의 CD



(b) GOV내의 첫 번째 P프레임에 대한 CD

그림3. CD 왜곡의 전파 특성

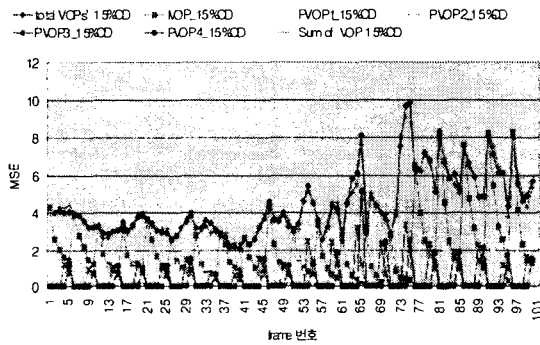


그림4. 프레임별 독립적 15%CD의 경우와 비교

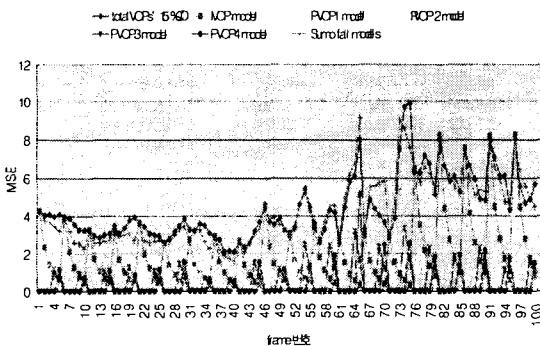


그림5. 프레임별 독립적 15%CD에 대한 모델링 결과와 비교

$$D_i(n) = D_i e^{-\alpha_i n - \beta_i}, n=1,2,3,4, \text{ where } D_i = \text{MSE of IVOP}$$

$$D_{P1}(n) = D_{P1} e^{-\alpha_{P1} n - \beta_{P1}}, n=1,2,3, \text{ where } D_{P1} = \text{MSE of PVOP1}$$

$$D_{P2}(n) = D_{P2} e^{-\alpha_{P2} n - \beta_{P2}}, n=1,2, \text{ where } D_{P2} = \text{MSE of PVOP2}$$

$$D_{P3}(n) = D_{P3} e^{-\alpha_{P3} n}, n=1, \text{ where } D_{P3} = \text{MSE of PVOP3}$$

(5)

식(5)에 따른 모델링 결과를 그림5에 나타내었으며, 모델링에 의한 결과와 실제 전체 복호화 왜곡은 근사적으로 일치함을 보인다. 따라서, 실제로 지수 감소함수 형태로 모델링하기 위한 파라미터 값만 충실히 얻어질 수 있다면, 전파/누적 왜곡에 의한 복호화 왜곡은 적절히 제어될 수 있을 것이다.

## V. 결론

본 논문에서는 FD/CD트랜스코딩 기법에서 CD로 인한 전파/누적 왜곡은 무시할 수 없을 정도의 매우 큰 왜곡을 발생시킴을 알 수 있었다. 그리고 CD로 인한 프레임의 품질 왜곡은 이전 프레임에서 CD로 인해 전파되어 오는 왜곡과 상관성이 없음을 확인하였다. 이와 같은 실험적 결과를 바탕으로 개개의 프레임 단위로 전파/누적 되는 왜곡특성에 대해 모델링을 수행하였고, 제안된 모델은 전체의 복호화된 왜곡을 잘 예측할 수 있음을 보였다. 본 논문에서 분석되고, 제안된 기법은 FD/CD 트랜스코딩 기법에서 일정화질 유지 등의 복호화된 화질을 제어하는데 효과적으로 사용될 수 있을 것이다.

## 참고문헌

- [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.Zdopski, "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.