

FD/CD 트랜스코딩기법에서 CD에 의한 왜곡 분석

김진수* · 김재곤** · 강경옥** · 김진웅**

*한밭대학교 정보통신컴퓨터공학부 · **한국전자통신연구원 방송미디어연구부

An Analysis of CD Distortions in Simple FD/CD Transcoding Algorithm

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

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

E-mail : jskim67@hanbat.ac.kr

요 약

이기종의 망과 다양한 종류의 사용자 단말 보급으로 디지털 영상 콘텐츠를 효과적으로 제공하는 기술이 요구되고 있고, 이를 위한 한 가지 접근 방법으로서 트랜스코딩 기법에 대한 연구가 활발히 진행되고 있다. 특히, FD(Frame Dropping)/CD(Coefficient Dropping)기법은 VLC영역에서 트랜스코딩을 수행하여 계산량을 절감하고, 가장 손쉽게 구현할 수 있는 기법으로 고려되고 있다. 그러나, 이 기법에 근거한 기존의 연구 결과는 이전 프레임의 CD로 인해 현재 프레임에 전파되어 오는 왜곡 특성을 적절히 분석하지 않고 적용되고 있다. 본 논문에서는 VLC영역에서 수행되는 CD에 따른 왜곡의 특성 및 관계를 수식적으로 유도하고, 또한, 각 왜곡의 특성을 실험적으로 분석한다. 모의실험을 통하여 본 논문에서 분석된 결과들은 VLC영역에서의 효과적인 FD/CD기법을 구하는데 효과적으로 사용될 수 있을 것이다.

ABSTRACT

In the framework of universal multimedia access, one challenge for video transmission (communication) is to deliver video content through heterogeneous network channels matching the diversity of client devices. As one of the many adaptation methods, media transcoding is commonly considered. Particularly, FD (Frame Dropping)/CD Coefficient Dropping) transcoding is used due to the low computational complexity and simple implementation. In this scheme, CD is simply implemented by truncating the high frequency AC DCT coefficient bits. But, the CD error tends to be propagated within single GOV. In this paper, we derive the distortion relationships between CD error and propagation error, and investigate the error characteristics by computer simulations. The CD error characteristics can be effectively used in the FD/CD transcoding scheme.

키워드

트랜스코딩, 전파에러, CD(Coefficient Dropping), FD(Frame Dropping)

1. 서 론)

UMA(Universal Multimedia Access)의 환경에서 이기종의 망자원과 다양한 사용자 단말 특성에 맞추어 비디오 콘텐츠를 효과적으로 수용하는 기술은 필수적이다. 이를 구현하기 위한 한 방법으로서 트랜스코딩 기법에 대한 연구가 활발히 진행되고 있다.

압축 부호화된 비트열을 망자원 또는 사용자 단말 등의 제한적 특성에 맞추기 위한 트랜스코딩 기법으로는 크게 세가지 기법으로 분류된다^{1,2,3,4}. 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 절감에 있어서는 매우 불리한 접근 방법이다¹. 두 번째

1) 본 연구는 정보통신부 과제의 일부로서 수행되었습니다.

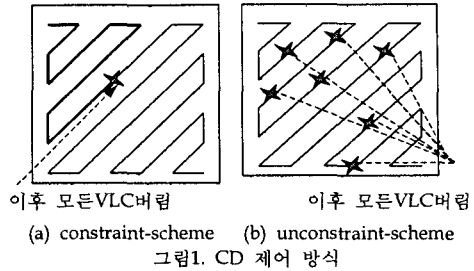
기법으로는 DVLC를 수행한 후에 DCT영역에서 양자화를 다시 수행하는 방법(DDT: DCT-Domain Transcoder)이 있다. DCT영역에서 움직임 추정을 하거나 또는 재양자화를 선택적으로 수행하는 방법(OLT: Open-Loop Transcoder)은, 계산량을 CPDT에 비해 줄일 수 있으나, 다양한 응용으로 확장하는데 여전히 단점이 있다^[2]. 마지막으로 VLC 영역에서 FD(Frame Dropping)와 CD(Coefficient Dropping)방법을 수행하는 방법이 있다. 이 방법은 시간축상에 위치하는 프레임율을 줄이는 소거 기법에 의한 FD와 고주파 성분을 제거하는 CD에 의해 수행된다. 이 기법은 성능 측면에서 앞에서 논한 방법에 비해 우수하지 않지만, 계산량을 매우 절감시킬 수 있고, 다양한 응용처에 확장하여 적용할 수 있다^[3,4,5].

기존의 VLC영역에서 수행하는 FD/CD기법에 의한 트랜스코딩 기법은 전파에러와 CD에 의한 에러 효과를 적절히 분석하지 못하고 있다^[4,5]. 본 논문에서는 CD에 의해 초래되는 전파 에러를 수식적으로 분석하고, 그 효과를 모의실험을 통하여 분석한다.

II. FD/CD 트랜스코딩 기법

FD(Frame Dropping)는 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 우선 제거 대상이 되며, GOP(Group of Pictures)내에서 가장 나중에 나오는 P프레임 순으로 제거될 수 있다.

CD(Coefficient Dropping)에 의한 트랜스코딩은 FD에 비해 적은 양의 데이터를 제어할 수 있는 수단이 된다. CD기법은 중요하지 않는 고주파 DCT 계수에 대응하는 VLC를 제거함으로써 비트율을 제어하는 기법이다. VLC코드를 제거할 수 있는 기법으로는 EOB코드에서부터 순차적으로 제거하는 방법(constraint-scheme)과 단일 블록내에 위치에 상관없이 중요도가 낮은 계수를 없애는 방법(unconstraint-scheme)이 있을 수 있다. 그림1은 CD를 수행하기 위한 두가지 방식에 대한 것을 그림으로 설명하고 있다. [4,5]에서는 그림1의 두가지 방식에 대한 화질 변화는 매우 미미한 수준으로 보고 하고 있다. 이와 같은 모의 실험결과를 바탕으로 가능한한 비트열 문법 구조의 변화를 주지 않는 방법(constraint-scheme)을 사용할 것을 권고하고 있다.



CD를 이용한 비트율 제어 방식은 크게 두가지 방식으로 접근할 수 있다. 각각의 블록에 초래되는 왜곡을 고려하지 않고, 일정을 이상의 VLC계수를 제거하는 기법(URCD: Uniform Rate-based CD)과 한 프레임 기간 동안의 프레임 최적화 윈도우 내에서 각 블록내의 CD계수에 대한 최적의 EOB 지점을 찾는 기법(LOCD: Lagrange Optimiztion CD)이 있다^[4,5]. LOCD는 CD로 초래되는 왜곡을 최소화하기 위해 Lagrangian 탐색을 사용한다. 본 논문에서는 성능 비교를 위하여 LOCD에 의한 기법에 의한 비트율 제어를 수행하고, CD제어 방식으로 constraint-scheme을 사용한다.

이상에서 논의된 FD/CD에 의한 트랜스코딩 기법을 개관적으로 설명하는 것이 그림2에 나타난다.

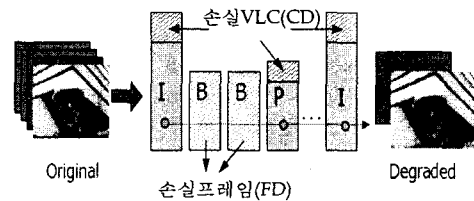


그림2. FD/CD에 의한 transcoding 개념도

한편, 참고문헌 [4,5]에서는 비트율 제어에 있어, 기억 소자를 갖지 않는 모델(memoryless-model)을 사용하여, 부호화된 프레임들간에 CD에 의한 왜곡은 매우 작다고 가정하였다. 그리고, CD에 의한 전파 왜곡과 누적 왜곡은 화질에 거의 영향을 주지 않는다고 가정을 하였고, 실험 결과에 따르면, 약 0.3dB정도의 품질 저하를 갖는다고 하였다.

본 논문에서는 [4,5]에서 제안한 알고리즘을 바탕으로 CD에 의한 전파 왜곡을 수식적으로 유도하여 보고, 실험적으로 분석한다.

III. CD에 의한 전파 왜곡

본 논문에서는 전개의 편의성을 위해 CD의 레이트 세이핑 방식은 그림1.(a)에 나타난 방법을 사용하는 것으로 가정한다(일반성을 잃지 않음). 이를 바탕으로, 부호화된 j번째 프레임의 i번째 블록에 대하여 트랜스코딩을 행함으로써 초래되는 왜

곡을 나타낸다. 먼저, 트랜스코딩 왜곡없이 복호화된 형태의 표현($y_j(i)$)은 다음과 같이 표현한다.

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

여기서, $MC_i(y_{j-1})$ 는 j-1번째 프레임을 기초로 움직임 보상된 성분 크기이고, e_i 는 움직임 예측 오차 신호 성분이다.

만일, j-1번째 프레임이 CD를 수행하였다면, i번째 블록에 대한 표현을 다음과 같이 각각 표현할 수 있다.

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

j번째 프레임의 i번째 블록에 대한 트랜스코딩 왜곡 $D_{j,i}(b_i)$ 에 대한 정의는 그림1.(a)에 나타난 DCT블록의 breakpoint위치 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 $a(i) = MC_i(y_j(i)) - MC_i(\hat{y}_{j-1}(i))$ 이며, 이 식에 대한 DCT표현으로 $A(k) = DCT_k(a(i))$ 로 나타낸다. 또한, $E(k) = DCT_k(e(i))$ 과 $\hat{E}(k) = DCT_k(\hat{e}(i))$ 로 각각 표현하면, 왜곡은 아래와 같은 형태로 표현된다.

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

식(4)에서 $E(k)$ 와 $\hat{E}(k)$ 사이의 관계는 b_i 이후에 있는 DCT계수 성분만큼의 CD에 의해 왜곡이 발생한다.

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

이 된다. 식(5) 우측의 세 번째 항목은 j번째 프레임의 i번째 블록에서 CD를 수행함으로써 발생하는 왜곡 성분이다. 즉, b_i 위치에 따라 크기가 결정

되는 왜곡이다. 반면에 식(5) 우측의 첫 번째항은 $a(i) = MC_i(y_j(i)) - MC_i(\hat{y}_{j-1}(i))$ 로 측정되는 왜곡으로서 j번째 프레임의 CD에 의한 왜곡과 무관하게 j-1번째 프레임에서 전해져 오는 왜곡이다. 즉, j-1번째 프레임의 트랜스코딩으로 인해 초래되는 왜곡이 성분이다. 그리고, 식(5) 우측의 두 번째항은 왜곡의 측정단위가 MSE(Mean Square Error)임으로 인해 관측되는 것으로 j-1번째 프레임에서 움직임 보상에 의한 차신호와 j번째 프레임의 CD에 의한 차신호 성분간의 상관관계로 인해 발생하는 왜곡 성분이다.

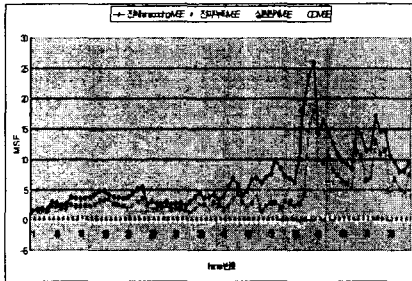
한편, 식(5)의 첫 번째 항은 다시 회귀적인 형태의 분석이 가능하다. 즉, $\hat{y}_{j-1}(i)$ 는 j-1번째 프레임의 CD에 의한 왜곡, j-2번째 프레임에서 전해져 오는 왜곡 성분, 그리고 이 두 가지 차신호 성분의 상관관계에서 발생하는 왜곡으로 나눌 수 있다. 이와 같은 분석을 연속적으로 확장하여 j번째 프레임까지 확대 해석하면, j-2번째 프레임에서 CD에 의한 왜곡이 j-1번째 프레임에 전파되고, 그것이 움직임 보상의 방법으로 j번째 프레임까지 누적 왜곡(accumulated error)으로 전달되기도 하고, 경우에 따라서는 j-1번째 프레임까지만 영향을 줄 수도 있다. 기존의 연구^[4,5]에서는 이 이전 프레임에서 전해져 오는 왜곡을 무시하였지만, 본 논문에서는 이전 프레임(들)에서 넘어오는 왜곡 특성을 시뮬레이션을 통하여 확인한다.

IV. 모의 실험 및 결과

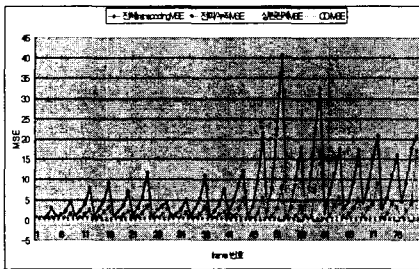
모의 실험을 위해 1.5Mbps로 부호화된 CIF해상도의 forman시퀀스를 사용한다. 비트스트림의 부호화 특성은 MPEG-4 simple profile, 1~300 frames, IBBPBBPBB...형태(GOV=15, SubGOV=3)이다. 본 논문에서는 I/P프레임만을 CD의 대상으로 고려하고, B프레임은 FD의 대상으로 한다.

각 GOV내에 비트율 할당에 따른 영향이 다르므로, 두가지 방식으로 각각 나누어 모의 실험을 한다. 트랜스코딩전/후의 비트율을 각각 R과 R'으로 나타낼 때, R-R'만큼의 양을 IPPPPP시퀀스에 균등하게 할당하는 일정을 할당 CD제어방식과 GOV내의 앞선 프레임에 많은 비트를 주는 차등률 할당 CD제어 방식을 사용한다. 그림3의 결과는 R=1.5 Mbps, R'=800kbps으로 트랜스코딩한 후의 IPPPP시퀀스에서 발생하는 평균자승오차(MSE: mean square error)를 나타낸 것이다. 그림에서 알 수 있듯이, 일정을로 비트율을 할당하는 방식에서 앞선 프레임에 의한 CD로 인한 파급효과가 매우 큼을 알 수 있다. 특히, 움직임이 많은 영역에서는 현재 프레임의 CD왜곡 보다도 이전프레임에서 전해져 오는 왜곡이 더 크게 나타나는 것을 볼 수 있다. 또 한가지 더 주목해야 할 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 GOV내에서 앞선 프레임에 비해 상대적으로 뒤쪽에 위치하는 프레임에서 전파 왜곡 특성이 크게 나타남을 알 수 있다.

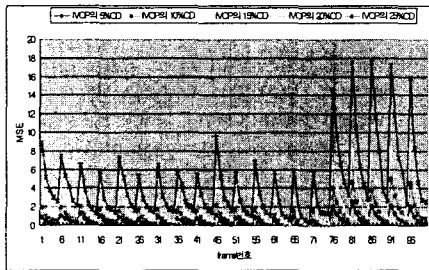


(a) 일정틀할당 CD제어 방식

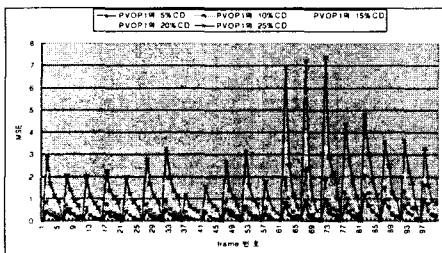


(b) 차등틀 할당 CD제어 방식

그림3. 두가지 제어 방식에 따른 왜곡 특성



(a) I프레임의 CD



(b) GOV내의 첫 번째 P프레임에 대한 CD

그림4. CD 왜곡의 전파 특성

그림4(a)에서는 I프레임에서 5%, 10%, 15%, 20%, 25%로 각각 CD를 하였을 경우에 같은 GOV내에 전파되는 왜곡특성을 모의실험을 통하여 나타낸 것이다. 패턴은 동일한 형태가 나오지만, I프레임에서 CD를 많이 할수록 그 전파 효과는 크게 나오며, 왜곡의 전파는 지수함수적으로 감소하는 형태이다. 그림4(b)에서는 첫 번째 P프레임에서 CD를 행한 경우에 패턴은 동일한 형태가 나오지만, I프레임에서 CD효과에 비해 작은 크기로 왜곡이 전파됨을 알 수 있다.

V. 결론

본 논문에서는 FD/CD트랜스코딩 기법에서 CD 왜곡과 이에 따른 전파 왜곡 발생을 수식적으로 논하고, 그 결과를 모의실험을 통하여 분석하였다. 모의실험을 통하여, 기존의 다른 연구결과와 달리, 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. 본 논문에서 논의되고, 분석된 결과들은 FD/CD 트랜스코딩 기법을 더욱 효과적으로 구현하고, 또한, 성능 개선을 위한 제어 모델에 사용될 수 있을 것이다.

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.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.