

특집논문-07-12-6-02

DCT 계수 제거 트랜스코딩에서의 전파 왜곡의 통계적 특성 분석

김진수^{a)}, 김재곤^{b)†}

Analysis of Statistical Properties of Propagation Errors in DCT Coefficient-Dropping Transcoder

Jin-soo Kim^{a)}, and Jae-Gon Kim^{b)†}

요 약

DCT(Discrete Cosine Transform) 계수 제거 기법은 MC(Motion Compensated)-DCT 기반의 MPEG 비디오에서의 효율적인 움직임 트랜스코딩 기법이다. 그러나, 이들 기법에서는 DCT 계수 제거로 인한 왜곡이 전파되게 되고 종종 심각한 화질 열화를 유발하게 된다. 본 논문에서는 왜곡 전파 특성에 대한 두 가지의 통계적 성질을 제시하고 수식적으로 분석한다. 즉, 현재 프레임의 DCT 계수 제거 왜곡과 이전 프레임에서 전파되어 오는 왜곡 간에 상관성이 없음을 보이고, 각 프레임의 DCT 제거로 발생하는 전파 왜곡의 누적과 현 프레임의 DCT 계수 제거 왜곡의 합으로 전체 왜곡을 근사할 수 있음을 보인다. 시뮬레이션을 통하여, 본 논문에서 수식적으로 제시한 통계적 특성이 실제 비디오 시퀀스에서 유효함을 실험적으로 증명한다.

ABSTRACT

Discrete cosine transform (DCT) coefficient dropping is well recognized as an efficient rate adaptation transcoding in motion-compensated (MC)-DCT based MPEG-compressed videos. However, in this scheme, the errors incurred by the DCT coefficient-dropping are propagated and often result in significant visual quality degradation. This paper presents two propositions describing well the statistical properties of propagated errors. That is, we propose that the DCT error of the current frame is not correlated to the propagated errors of the previous frames. We also propose that the overall distortions in a given frame can be approximated as the sum of the DCT error of the current frame and the propagated errors from the previous frames. Then, it is shown that several computer simulations with different video sequences verify the effectiveness of the proposed statistical analyses.

Keywords: Transcoding, Rate adaptation, Propagated errors, DCT coefficient-dropping

1. 서론

UMA(Universal Multimedia Access)의 환경에서는 이중

의 망특성과 다양한 성능의 단말에 적응적으로 비디오를 제공하기 위한 비디오 적응 기법이 필수적으로 요구된다^[1]. 이러한 이중의 소비 환경에서 비디오를 적응적으로 제공하는 방법으로 트랜스코딩(transcoding) 및 스케일러블 비디오(Scalable Video Coding: SVC) 적응(adaptation) 등에 대한 연구가 활발히 진행되고 있다. 압축 부호화된 비트열을 망 자원 또는 사용자 단말 등에 의해 제한되는 특성에 적응시키기 위한 트랜스코딩 기법은 기존의 표준 부호화 기법

a) 한밭대학교 멀티미디어공학과
Dept. of Multimedia Engineering, Hanbat National University
b) 한국항공대학교 항공전자 및 정보통신공학부
School of Electronics, Telecommunication and Computer Engineering,
Korea Aerospace University
† 교신저자 : 김재곤(jgkim@kau.ac.kr)

을 그대로 사용할 수 있다는 장점이 있으며, 그 접근 방식으로는 크게 다음의 세 가지 방법으로 나눌 수 있다^{[2],[3],[4]}.

첫 번째 접근 방법으로는 압축 부호화된 비트열을 완전히 복호화한 후에 화소 영역에서 움직임 추정/DCT 변환, 양자화, 가변길이부호화(VLC: Variable Length Coding)를 수행하는 방법(CPDT: Cascaded Pixel Domain Transcoder)이다. 이 방법은 다른 접근 방법에 비해 성능이 뛰어나다는 장점이 있으나, 많은 연산량을 필요로 한다. 두 번째 접근 방법으로는 VLC 복호화를 수행한 후에 DCT 영역에서 재양자화를 수행하는 방법(DDT: DCT-Domain Transcoder)이 있다. DCT 영역에서 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.

마지막으로는 VLC 영역에서 프레임 제거(FD: Frame Dropping)와 DCT 계수 제거(CD: Coefficient Dropping)를 수행하는 방법이 있다^{[5],[6],[7]}. 이 방법은 시간 영역에서 프레임율을 줄이는 방식과 더불어 DCT 계수 중에 고주파 성분을 제거하는 방법으로 구현된다. 이 기법은 성능 측면에서 앞에서 논한 방법들에 비해 우수하지 않지만, 압축 비트열의 일부를 제거하는 것으로 매우 적은 연산량만으로 매우 간단하게 구현이 가능하므로 실시간 응용에 적합한 접근 방법이다.

본 논문에서는 압축 부호화된 비트열에서 손쉽게 구현할 수 있고, UMA 환경에서 실시간 적용에 적용할 수 있는 세 번째 접근 방식인 프레임 제거 및 DCT 계수 제거 방법을 결합한 FD-CD 트랜스코딩 기법에 대해 다룬다. 기존의 연구에서 FD는 다른 프레임의 예측부호화를 위한 기준 프레임으로 사용되지 않는 B프레임을 가장 먼저 제거하고, 다음으로 GOP내에서 가장 뒤 쪽에 위치한 P프레임 순으로 제거하는 기법을 취한다^{[5],[7],[8],[9]}. 그리고 CD를 수행함에 있어서는 고주파 성분을 구성하는 DCT 계수를 먼저 제거한다. 이때, CD의 적용은 FD와 달리 나중에 복호화되는 프레임에 왜곡을 전파/누적시키게 된다. 기존의 FD-CD 기법에서는 이러한 전파/누적 왜곡이 무시할만한 수준으로 가정하고, 각 계수의 제거는 영상내 부호화(Intra Coding) 방식과 동일하게 취급하였다^{[1],[5],[7]}. 그리고 UMA 환경에 적

용하기 위한 비트율 제어 등에 관심을 두었으며, DCT 계수 제거로 인한 왜곡 전파 특성에 대한 수학적 분석이나 모의 실험 결과를 제시하지 않고 있다.

본 논문에서는 DCT 계수 제거로 인하여 발생하는 왜곡 전파의 특성을 크게 두 가지의 통계적 성질로 분석한다. 즉, 첫째, 현재 프레임의 DCT 계수 제거 왜곡과 이전 프레임에서 전파되어 오는 왜곡 간에 상관성이 없음을 제안하고, 둘째, 각 프레임의 DCT 제거로 발생하는 전파 왜곡의 누적과 현 프레임의 DCT 계수 제거 왜곡의 합으로 전체 왜곡을 근사화할 수 있음을 제안한다. 또한 시뮬레이션을 통하여, 본 논문에서 수식적으로 제안한 통계적 성질이 실험적으로 유효함을 증명한다.

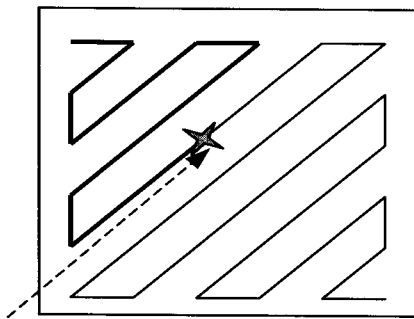
본 논문의 구성은 다음과 같다. 우선 II장에서는 기존의 FD-CD 트랜스코딩 기법에 대해 간단히 살펴보고, III장에서는 DCT 계수 제거(CD)에 의해 발생하는 왜곡의 전파 특성을 DCT 계수가 갖는 통계적인 특성에 기초하여, 두 가지의 특성을 분석한다. 그리고 IV장에서는 모의실험을 통하여 III절에서 제안된 수학적 분석 특성이 유효함을 모의실험을 통하여 검증한다. 끝으로, VI장에서는 결론을 맺는다.

II. FD-CD 트랜스코딩 기법

프레임 제거(FD)는 압축 부호화된 몇 개의 프레임을 제거함으로써 많은 양의 비트율을 제어할 수 있다. 가장 중요도가 떨어지거나 또는 다른 부호화 프레임에 기준 프레임으로 사용되지 않는 프레임을 제거함으로써, 가능한 우수한 품질을 유지할 수가 있다^{[5],[7]}. MPEG으로 압축 부호화된 비디오 시퀀스에서 B 프레임은 다른 프레임에 대해 부호화 의존성을 제공하지 않으므로 가장 먼저 제거된다. B 프레임이 제거된 후에는 P프레임이 우선 제거 대상이 되는데, 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.

압축된 비트열에서 DCT 계수 제거(CD)에 의한 트랜스코딩은 FD에 비해 적은 양의 데이터를 세밀하게 제어할 수

있는 기법이 된다^{[5],[12]}. CD 기법은 중요하지 않은 고주파 DCT 계수의 VLC를 제거함으로써 비트율을 제어하는 기법이다. VLC코드를 제거하는 방법으로는 EOB(End of Block) 코드에서부터 순차적으로 제거하는 방법(constraint-scheme)과 단일 블록내에 위치에 상관없이 중요도가 낮은 계수를 없애는 방법(unconstraint-scheme)이 있다. 이러한 두 가지 방식에 대한 성능 차이는 매우 미미한 수준으로 보고하고 있다^[5]. 또한 모의 실험결과를 바탕으로 압축된 비트열에 대한 문법 구조 변화를 최소로 유지하므로, 더욱 쉽게 구현될 수 있으므로, 본 논문에서는 [12]에 따라 그림 1에 의한 방식으로 DCT 계수 제거를 수행한다. 그림 2는 이상에서 설명한 FD-CD에 의한 트랜스코딩 기법에 대한 개념도를 나타내고 있다^[1].



k_i 이상의 DCT 계수 제거

그림 1. EOB에서 순차적으로 DCT 계수를 제거하는 방식
Fig. 1. CD operation of the constrained scheme

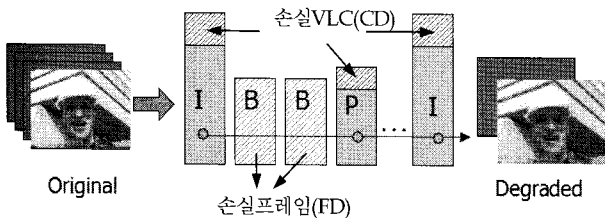


그림 2. FD-CD에 의한 트랜스코딩에 대한 개념도
Fig. 2. FD-CD combined transcoding

기존에 CD를 적용하는 프레임에 대한 기존의 비트율 제어 기법은 크게 두 가지 방식이 제안되어 있다. 먼저, 각각의 블록에 초래되는 왜곡 크기를 고려하지 않고, CD 대상

의 프레임에 할당된 비트율을 맞추기 위해, 일정을 이상의 VLC 계수를 제거하는 기법(URCD: Uniform Rate-based CD)과 한 프레임 기간 동안의 윈도우 내에서 각 블록의 CD 계수에 대한 최적의 EOB 지점을 찾는 기법(LOCD: Lagrange Optimization CD)이 있다^{[1],[12]}. LOCD는 CD로 초래되는 왜곡을 최소화하기 위해 Lagrangian 탐색을 사용한다. 즉, 주어진 비트량하에서 최소의 왜곡이 발생하도록 하는 블록별 독립적인 EOB를 찾는 방식이다. 본 논문에서는 CD제어 방식으로 constraint-scheme을 사용하는 것으로 제한 한다. LOCD 기법은 주어진 j-번째 프레임에 대한 예산 비트량 B_{keep}^j 이 주어질 때, j-번째 프레임에 속하는 모든 블록에 대해 EOB 지점의 집합 즉, $\vec{k} = (k_1, k_2, \dots, k_M)$ 을 다음과 같이 구한다.

$$\vec{k} = \underset{(k_1, k_2, \dots, k_M)}{\text{minimize}} \sum_{i=1}^M \sum_{k=k_i+1}^{N-1} z_{i,k}^2 \text{ such that} \tag{1}$$

$$\sum_{i=1}^M \sum_{k=1}^{k_i} b_{coeff}^{i,k} \leq B_{keep}^j$$

여기서, z_i 은 i-번째 블록의 k-번째 DCT 계수 역양자화된 DCT 계수를 나타내고, $b_{coeff}^{i,k}$ 은 i-번째 블록의 k-번째 DCT 계수의 부호화된 비트량을 나타낸다. N(흔히, N=64)은 블록 내의 DCT 계수의 수를 나타내고, M은 프레임내의 전체 블록의 개수를 나타낸다. 이와 같이 구현되는 LOCD는 주어진 예산 비트량을 모든 블록이 같은 비트율 증가량에 대해 같은 화질의 증가량을 갖도록 분배하는 방식이다.

III. I-프레임과 P-프레임에서 DCT 계수 제거에 의한 전파 왜곡 특성

1. 심볼과 기호

M개의 블록을 갖는 복호화된 프레임 j-번째의 영상을 $y_j = y_{j,i} \ i=1, \dots, M$ 와 같이 표현한다. 여기서, $y_{j,i}$ 는 j-번째 프레임내에 있는 i-번째 블록의 복호화된 표현을 나타낸다. 약, $y_{j,i}$ 가 화면간 부호화 모드(inter-coded)로 부호화된다면,

다음과 같이 복원된다.

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

여기서, $C_{j,i}(y_{j-1})$ 는 j-번째 프레임의 i-번째 블록이 (j-1)-번째 프레임을 참조 프레임으로 하여 움직임 보상으로 얻어지는 성분을 나타내고, $e_{j,i}$ 는 부호화된 움직임 예측 잔차 신호의 복호화된 신호 성분을 나타낸다.

만약, $y_{j,i}$ 가 화면내 부호화 모드(intra-coded)로 부호화된다면, 이 블록은 다음과 같이 복원된다.

$$y_{j,i} = e_{j,i}, j \geq 1 \quad (3)$$

만약, 화면내 부호화 모드로 부호화된 DCT 계수들의 일부가 제거된다면, 이 블록은 부호화 왜곡 외에도 제거된 DCT 계수 만큼에 의해 추가적인 화질 열화가 발생하게 된다. 이때, 기호는 다음과 같이 표현한다.

$$\hat{y}_{j,i} = \hat{e}_{j,i}, j \geq 1 \quad (4)$$

유사하게, 화면간 부호화 모드로 부호화된 블록에서 움직임 예측 잔차 신호의 DCT 계수 일부가 제거된다면, 그 기호는 다음과 같이 표현된다.

$$\hat{y}_{j,i} = C_{j,i}(y_{j-1}) + \hat{e}_{j,i}, j \geq 1 \quad (5)$$

이때, j-번째 프레임내에 있는 i-번째 블록을 복원할 때, (j-1)번째 프레임에서 DCT 계수의 일부가 제거되었다면, 움직임 보상에 의해 (j-1)번째 프레임의 왜곡이 영향을 받게 된다. 식 (5)는 다음과 같이 표현된다.

$$\hat{y}_{j,i} = C_{j,i}(\hat{y}_{j-1}) + \hat{e}_{j,i}, j \geq 1 \quad (6)$$

2. DCT 계수 제거로 인해 발생하는 복호화 왜곡

그림 1에서 나타내었듯이, 스캔 순서가 k_i 보다 큰 DCT

계수들이 화면내 부호화 모드에서 제거되었다면, j-번째 프레임의 i-번째 블록의 평균 왜곡(MSE: Mean Square Error)은 다음과 같이 정의되고 구해진다.

$$D_{j,i}(k_i) = \frac{1}{N} \|y_{j,i} - \hat{y}_{j,i}\|^2 = \frac{1}{N} \|e_{j,i} - \hat{e}_{j,i}\|^2 \quad (7)$$

DCT는 정규직교변환이므로, 에너지 보존 성질^[10]이 만족되며, 이를 식 (7)에 적용하면, 다음과 같이 DCT 영역에서 쉽게 구해질 수 있다.

$$\begin{aligned} D_{j,i}(k_i) &= \frac{1}{N} \sum_{k=0}^{N-1} \{E_{j,i}(k) - \hat{E}_{j,i}(k)\}^2 \\ &= \frac{1}{N} \sum_{k=k_i}^{N-1} E_{j,i}^2(k) \end{aligned} \quad (8)$$

여기서, $E_{j,i}(k) = DCT\{e_{j,i}\}$, $\hat{E}_{j,i}(k) = DCT\{\hat{e}_{j,i}\}$, $k=0, \dots, N-1$ 이다.

한편, DCT 계수 제거 연산이 (j-1)-번째 프레임에 적용되어, j-번째 프레임에 왜곡이 전파되는 화면 간 부호화 모드를 고려해보자. 즉, 식 (5)에서 식 (2)에 의한 차이 $C_{j,i}(y_{j-1}) + e_{j,i} - C_{j,i}(\hat{y}_{j-1}) + \hat{e}_{j,i}$ 가 복원 영상의 품질 차이가 된다. 이를 이용하여 j-번째 프레임의 i-번째 블록의 평균 왜곡을 구하면 다음과 같이 구해진다.

$$\begin{aligned} D_{j,i}(k_i) &= \frac{1}{N} \|y_{j,i} - \hat{y}_{j,i}\|^2 \\ &= \frac{1}{N} \|C_{j,i}(y_{j-1}) - C_{j,i}(\hat{y}_{j-1}) + e_{j,i} - \hat{e}_{j,i}\|^2 \end{aligned} \quad (9)$$

DCT는 앞에서 언급한 에너지 보존 성질외에도 다음과 같이 압축 영역에서 연산을 간단히 수행할 수 있는 성질을 갖는다^[10]. 즉,

$$\begin{aligned} DCT[A+B] &= DCT[A] + DCT[B] \\ DCT[A*B] &= DCT[A]*DCT[B] \end{aligned} \quad (10)$$

와 같은 성질이 있다. 이를 식 (9)에 적용하면,

$$\begin{aligned}
 D_{j,i}(k_i) &= \frac{1}{N} \|DCT\{C_{j,i}(y_{j-1}) - C_{j,i}(\hat{y}_{j-1}) + e_{j,i} - \hat{e}_{j,i}\}\|^2 \\
 &= \frac{1}{N} \|DCT\{C_{j,i}(y_{j-1}) - C_{j,i}(\hat{y}_{j-1})\} + DCT\{e_{j,i} - \hat{e}_{j,i}\}\|^2 \\
 &= \frac{1}{N} \|DCT\{C_{j,i}(y_{j-1}) - C_{j,i}(\hat{y}_{j-1})\}\|^2 + \frac{2}{N} \|DCT\{C_{j,i}(y_{j-1}) - C_{j,i}(\hat{y}_{j-1})\} DCT\{e_{j,i} - \hat{e}_{j,i}\}\| \\
 &\quad + \frac{1}{N} \|DCT\{e_{j,i} - \hat{e}_{j,i}\}\|^2
 \end{aligned} \tag{11}$$

와 같다.

따라서, 화면간 부호화 모드로 부호화된 블록은 DCT 계수 제거로 인하여 초래되는 왜곡은 크게 세 가지 종류의 왜곡에 의해 초래됨을 알 수 있다. 즉, 식 (11)의 첫 번째 항은 (j-1)번째 프레임에서 전파되는 왜곡을 나타낸다. 그리고 식 (11)의 세 번째 항은 현 블록(j-번째 프레임의 i-번째 블록)의 DCT 계수 제거에 의해 발생하는 왜곡이다. 이러한 두 가지 왜곡 성분 외에도 식 (11)의 두 번째 항은 이러한 두 가지 왜곡 성분들의 상관성에 의해 발생하는 성분이다.

식 (11)의 세 번째 항은 그림 1에 나타난 것 처럼 DCT 계수들의 크기를 관찰함으로써 쉽게 계산될 수 있다. 그러나, 식 (11)의 첫 번째 항에서 $C_{j,i}(y_{j-1}) - C_{j,i}(\hat{y}_{j-1})$ 은 채택된 움직임 보상 방식에 의해 결정되는 움직임 보상 함수 $C(\cdot)$ 에 의존하기 때문에 압축 영역에서 $C_{j,i}(y_{j-1}) - C_{j,i}(\hat{y}_{j-1})$ 을 직접 구하는 것은 매우 복잡하다. 그렇지만, 흔히 사용하는 움직임 보상 방식은 동영상내 움직임 객체가 평형 이동을 가정하여 블록정합방식(BMA: Block Matching Algorithm)을 사용하기 때문에, 본 논문에서는 이러한 응용 환경을 고려함으로써 문제 분석을 간략화 한다. 즉, 움직임 보상 함수 $C(\cdot)$ 를 다음 식과 같이 선형이라고 가정한다.

$$C(A+B) = C(A) + C(B) \tag{12}$$

이러한 가정하에서, j-번째 프레임에서 평균화된 왜곡은 다음과 같이 구해질 수 있음을 제안한다.

성질 1. 이전 프레임에서의 DCT 제거 왜곡으로 인한 전과 왜곡과 현재 프레임의 DCT 제거 왜곡과는 상관성이 없다.

증명) I-프레임에서는 모든 블록들이 화면내 부호화모

드로 부호화되기 때문에 움직임 보상이 이루어지지 않는다. 따라서, 식 (11)에서 첫 번째 항과 두 번째 항은 존재하지 않게 되므로, 평균 왜곡 D_j 는 다음과 같이 표현된다.

$$D_j = \frac{1}{M} \sum_{i=1}^M D_{j,i}(k_i) = \frac{1}{MN} \sum_{i=1}^M \sum_{k=k_i}^{N-1} E_{j,i}^2(k) \tag{13}$$

P-프레임은 식 (11)에서 세 가지 종류의 왜곡이 모두 존재하게 된다. i-번째 블록에서 스캔 순서가 k_i 보다 큰 DCT 계수들이 제거되기 때문에, 식 (11)의 두 번째 항과 세 번째 항은 다음과 같이 표현될 수 있다.

$$\begin{aligned}
 D_j &= \frac{1}{M} \sum_{i=1}^M D_{j,i}(k_i) \\
 &= \frac{1}{MN} \sum_{i=1}^M \|DCT\{C_{j,i}(y_{j-1}) - C_{j,i}(\hat{y}_{j-1})\}\|^2 \\
 &\quad + \frac{1}{MN} \sum_{i=1}^M \sum_{k=k_i}^{N-1} DCT\{C_{j,i}(y_{j-1}) - C_{j,i}(\hat{y}_{j-1})\} E_{j,i}(k) \\
 &\quad + \frac{1}{MN} \sum_{i=1}^M \sum_{k=k_i}^{N-1} E_{j,i}^2(k)
 \end{aligned} \tag{14}$$

일반적으로 $C_{j,i}(y_{j-1}) - C_{j,i}(\hat{y}_{j-1})$ 과 $e_{j,i}$ 는 독립적으로 결정된다. 게다가 DCT는 변환 계수들이 상관성이 없도록 (uncorrelated) 하는 특성을 갖는다. $DCT\{C_{j,i}(y_{j-1}) - C_{j,i}(\hat{y}_{j-1})\}$ 과 $E_{j,i}(k)$ 는 통계적인 평균 값이 0이 된다^[13]. 따라서, $NM \gg 1$ 에 대해 식(11)은 다음과 같이 간략화될 수 있다.

$$\begin{aligned}
 D_j &= \frac{1}{M} \sum_{i=1}^M D_{j,i}(k_i) \\
 &\approx \frac{1}{MN} \sum_{i=1}^M \|DCT\{C_{j,i}(y_{j-1}) - \hat{y}_{j-1}\}\|^2 + \frac{1}{MN} \sum_{i=1}^M \sum_{k=k_i}^{N-1} E_{j,i}^2(k)
 \end{aligned}$$

3. 연속적인 프레임에서 DCT 계수 제거에 의한 왜곡 누적

앞에서 논하였던 DCT 계수 제거를 연속적인 두 프레임 즉, (j-2)-번째 프레임과 (j-1)-번째 프레임에 적용한다고 가정한다. (j-2)-번째 프레임에서 DCT 계수를 제거하고 그 영향이 (j-1)-번째의 복호화된 영상에 영향을 주게 된다. 이와 더불어, (j-1)-번째 프레임에서도 DCT 계수를 제거하면, 이들 두 프레임에서의 DCT 계수 제거는 j-번째 프레임의 복호화된 영상의 품질로 영향을 주게 된다. 이러한 관계를 식 (6)에 적용하면 다음과 같이 표현된다.

$$\begin{aligned}\hat{y}_{j,i} &= C_{j,i}(\hat{y}_{j-1}) + \hat{e}_{j,i} \\ &= C_{j,i} \sum_{i=1}^M \hat{y}_{j-1,i} + \hat{e}_{j,i}, \quad j \geq 2 \quad (15) \\ &= C_{j,i} \sum_{i=1}^M \{C_{j-1,i}(\hat{y}_{j-2}) + \hat{e}_{j-1,i}\} + \hat{e}_{j,i}\end{aligned}$$

식 (12)에서 나타낸 움직임 보상 함수의 선형성을 식 (15)에 적용하면, 식 (15)는 다음과 같이 표현된다.

$$\hat{y}_{j,i} = C_{j,i} \left\{ \sum_{i=1}^M C_{j-1,i}(\hat{y}_{j-2}) \right\} + C_{j,i} \left\{ \sum_{i=1}^M \hat{e}_{j-1,i} \right\} + \hat{e}_{j,i}, \quad j \geq 2 \quad (16)$$

j-번째 프레임의 i-번째 블록의 평균 왜곡은 다음과 같이 나타낼 수 있다.

$$\begin{aligned}D_{j,i} &= \frac{1}{N} \| y_{j,i} - \hat{y}_{j,i} \|^2 \\ &= \frac{1}{N} \left\| C_{j-1} \sum_{i=1}^M C_{j-1,i}(y_{j-2}) + C_{j-1} \sum_{i=1}^M e_{j-1,i} + e_{j,i} \right. \\ &\quad \left. - C_{j,i} \sum_{i=1}^M C_{j-1,i}(\hat{y}_{j-2}) - C_{j,i} \sum_{i=1}^M \hat{e}_{j-1,i} - \hat{e}_{j,i} \right\|^2\end{aligned} \quad (17)$$

식 (10)을 사용함으로써 식 (17)은

$$\begin{aligned}D_{j,i} &= \frac{1}{N} \left\| DCT C_{j,i} \sum_{i=1}^M C_{j-1,i}(y_{j-2} - \hat{y}_{j-2}) \right. \\ &\quad \left. + DCT C_{j,i} \sum_{i=1}^M (e_{j-1,i} - \hat{e}_{j-1,i}) \right. \\ &\quad \left. + DCT e_{j,i} - \hat{e}_{j,i} \right\|^2\end{aligned} \quad (18)$$

로 표현된다.

j-번째 프레임의 i-번째 블록의 전체 왜곡은 $C(\cdot)$ 과 $DCT(\cdot)$ 의 연산 순서를 달리함으로써 다음과 같이 표현된다.

$$\begin{aligned}D_{j,i} &= \frac{1}{N} \left\| C_{j,i} \sum_{i=1}^M C_{j-1,i}(DCT y_{j-1} - \hat{y}_{j-2}) \right. \\ &\quad \left. + C_{j,i} \sum_{i=1}^M DCT(e_{j-1,i} - \hat{e}_{j-1,i}) \right. \\ &\quad \left. + DCT e_{j,i} - \hat{e}_{j,i} \right\|^2\end{aligned} \quad (19)$$

이상의 전개에서 DCT 계수 제거를 연속적으로 두 개의 프레임에 적용하였을 경우에, 이들 프레임에 기초하여 복호화되는 프레임에 대한 왜곡은 다음과 같이 제안된다.

성질 2. 두 개 이상의 이전 프레임에서 DCT 계수 제거를 수행한 경우에 각각에서 발생된 전파 왜곡들과 현재 프레임의 DCT 계수 제거 왜곡의 합으로 현재 프레임의 전체 왜곡을 간단히 근사화할 수 있다.

증명) 성질 1과 거의 유사한 방법으로 증명된다. 즉, 세 가지 종류의 DCT 계수 블록

$$\begin{aligned}C_{j,i} \sum_{i=1}^M \left\{ \sum_{i=1}^M C_{j-1,i}(DCT\{y_{j-1,i} - \hat{y}_{j-1,i}\}) \right\}, \\ C_{j,i} \sum_{i=1}^M DCT\{e_{j-1,i} - \hat{e}_{j-1,i}\} \text{과 } DCT\{e_{j,i} - \hat{e}_{j,i}\}\end{aligned}$$

들은 상호 독립적으로 결정된다. 더구나,

$NM \gg 1$ 에 대해 DCT 계수들의 통계적인 평균치는 0이므로, 다음의 통계적인 성질들이 있다.

$$\begin{aligned}\frac{1}{MN} \sum_{i=1}^M \left\| C_{j,i} \sum_{i=1}^M C_{j-1,i}(DCT\{y_{j-2} - \hat{y}_{j-2}\}) \right\| \times \\ C_{j,i} \sum_{i=1}^M DCT\{e_{j-1,i} - \hat{e}_{j-1,i}\} \approx 0\end{aligned}$$

$$\frac{1}{MN} \sum_{i=1}^M \left\| C_{j,i} \sum_{i=1}^M C_{j-1,i} (DCT\{y_{j-2} - \hat{y}_{j-2}\}) \times \right. \\ \left. C_{j,i} \sum_{i=1}^M DCT\{e_{j-1,i} - \hat{e}_{j-1,i}\} \right\| \approx 0$$

$$\frac{1}{MN} \sum_{i=1}^M \left\| C_{j,i} \sum_{i=1}^M DCT\{e_{j-1,i} - \hat{e}_{j-1,i}\} DCT\{e_{j,i} - \hat{e}_{j,i}\} \right\| \approx 0$$

이다. 따라서, (j-1)번째 프레임과 j-번째 프레임에서 DCT 계수들을 제거함으로써, j-번째 프레임에서 관찰되는 평균 왜곡은 다음과 같이 간략화 된다.

$$D_j \approx \frac{1}{MN} \sum_{i=1}^M \left\| C_{j,i} \sum_{i=1}^M C_{j-1,i} (DCT\{y_{j-2} - \hat{y}_{j-2}\}) \right\|^2 \\ + \frac{1}{MN} \sum_{i=1}^M \left\| C_{j,i} \sum_{i=1}^M E_{j-1,i} \right\|^2 \\ + \frac{1}{MN} \sum_{i=1}^M \sum_{k=k_0}^{N-1} E_{j,i}^2(k)$$

성질 2는 j-번째 프레임의 왜곡은 (j-2)-번째 프레임과 (j-1)-번째 프레임의 DCT 계수 제거에 의해 전파되는 왜곡 성분과 j-번째 프레임에서의 DCT 계수 제거에 의한 왜곡 성분의 합으로 구해질 수 있음을 알 수 있다. 그러나, 이전 프레임에서 전파되는 왜곡들은 움직임 보상 함수 $C(\cdot)$ 가 갖는 성질에 의존적으로 결정되기 때문에 완전하게 구하는 것은 매우 복잡하며, 본 논문에서는 더 이상의 전개는 생략한다.

이상에서의 전개를 (j-4), (j-3) 등의 이전 프레임에 대해 확장하면, 주어진 프레임의 전체 왜곡은 현재 프레임에서 제거한 DCT 계수 제거 왜곡과 이전 프레임들에서 전파되어오는 왜곡들의 합으로 표현될 수 있는 것으로 표현할 수 있다.

IV. 모의실험을 통한 검증 및 검토

1. 성질 1의 실험을 통한 검증

CIF 형식의 Foreman 시퀀스 300 프레임을 MPEG-4 심

플 프로파일(Simple Profile)에 의해 GOP=15, sub-GOP=3의 조건으로 1.5Mbps로 압축하였다. 이렇게 부호화된 비트열을 800kbps로 트랜스코딩하기 위해, 모든 B-프레임들은 프레임 제거 대상으로 하였으며, 추가적인 비트율 감소는 I-프레임과 P-프레임에서 DCT 계수 제거 기법을 통하여 구현되었다. 이때, B-프레임들은 모두 프레임 제거로 고려한 이유는 B-프레임에서의 DCT 계수 제거는 왜곡 전파가 없으므로 본 논문에서 고려대상에서 제외하였다. 그림 3은 부호화된 300 프레임 중에서 B-프레임들을 모두 제거하였고, 남은 I-프레임과 P-프레임에 대해 LOCD 기법을 적용하여 DCT 계수 제거를 적용하였을 때 발생하는 왜곡을 그린 것이다. "current error"는 식 (11)에서 세 번째 항에 해당되는 왜곡으로서 각 프레임에서 DCT 계수 제거를 함으로써 관찰되므로 쉽게 얻어질 수 있다. 그림에서 "accumulated error"는 식 (11)에서 첫 번째 항에 해당되는 왜곡이며, 실험적으로 얻기 위해 부호화된 각각의 프레임에 대해 독립적으로 DCT 계수 제거를 적용한다. 즉, 단일 GOP 내에 I-프레임에만 DCT 계수 제거를 수행하고, GOP 내의 다른 P-프레임들은 DCT 계수 제거를 수행하지 않는다. 이렇게 하면, I-프레임에서 DCT 계수 제거로 인해 발생된 왜곡이 다음의 P-프레임에 전파하게 된다. 이때, I-프레임에 DCT 계수 제거를 적용하지 않은 경우와 적용한 경우에 대해, 각각 첫 번째 P-프레임을 복호화하여 그 화질 변화를 측정함으로써 "accumulated error"가 얻어진다. 이와 유사하게 GOP 내에 I-프레임과 첫 번째 P-프레임에 대해 DCT 계수를 제거하고, 다른 P-프레임들은 DCT 계수 제거를 적용하지 않음으로써, GOP내의 두 번째 P-프레임에 대한 왜곡 전파 특성을 관찰할 수 있다. 이와 같은 방식으로 확장하면, GOP내의 다른 모든 P-프레임들에 대해서도 전파되는 왜곡을 측정할 수 있게 된다. 이때, 주의할 것은 "current error"를 구할 때와 같이 모든 블록에 대해 동일한 DCT 계수 제거 방식을 적용해야 한다.

한편, 그림 3에서는 "correlated error"로 표시한 왜곡 성분은 식 (11)의 두 번째 항을 의미하며, 이 왜곡 성분을 수학적으로나 또는 실험적으로 구하는 것은 매우 복잡하다. 본 논문에서는 식 (11)의 LHS 항, RHS의 첫 번째 항("accumulated error"), RHS의 세 번째 항("current error")을 이용

하여 간접적으로 구한다. 식 (11)의 LHS 항은 모든 I-프레임과 P-프레임에 대해 DCT 계수를 적용하여, 800kbps로 트랜스코딩하여 그 결과를 복호화하였을 때 관찰되는 왜곡이므로, 실험적으로 쉽게 구해진다. 따라서, 본 논문에서는 식 (11)의 두 번째 항을 구하기 위해, 식 (11)의 LHS 항에서 "current error" 성분과 "accumulated error" 성분을 뺄으로써 구할 수 있게 된다. 그림에 나타난 바와 같이 "correlated error"는 "current error"와 "accumulated error"에 비해 매우 미미하며, 무시 가능한 수준임을 알 수 있다. 따라서, 본 논문에서 제안한 성질 1은 타당한 제안임을 알 수 있다.

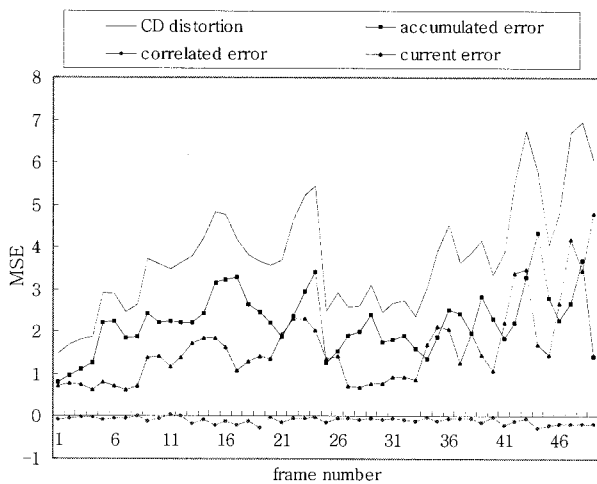


그림 3. DCT 계수 제거에 따른 왜곡 성분 (사용영상: Foreman)
Fig. 3. Each component of CD distortion for Foreman sequence

2. 성질 2의 실험을 통한 검증

본 실험을 위해 압축된 프레임에 대해 URCD(Uniform Ratio of Coefficient-Dropping) 즉, 일정 비율의 DCT 계수를 제거한다. 먼저, 앞의 실험과 동일하게 B-프레임들은 프레임 제거로 비트량을 제거하며, I-프레임과 P-프레임에 대해서는 압축된 전체 비트량중에 일정 비율만큼만 제거하도록 하여 왜곡 전과특성을 관찰한다. 그림 4는 1.5Mbps로 압축된 Foreman 시퀀스에 대해 모든 B-프레임을 제거하고, 남은 프레임들에 대해 모두 각각의 프레임에서 15%의 비율로 비트량을 제거하여 얻은 결과를 나타낸 것이다. 예로써,

"I_15%"는 I-프레임에 대해서만 압축 비트량에 비해 15%를 DCT 계수 제거를 통하여 비트량을 줄이고, 다른 모든 프레임들에 대해서는 DCT 계수 제거를 하지 않음으로써 발생하는 왜곡을 나타낸 것이다. 이때, I-프레임에 대해 왜곡을 삽입하였으므로 I-프레임을 참조하여 P-프레임들은 복호화되므로, 다른 P-프레임들은 DCT 계수를 제거하지 않았다하더라도 왜곡이 발생하게 된다. 이와 유사하게 "Px_15%"라고 표시한 것은 GOP내에 x-번째 위치하는 P-프레임에만 DCT 계수를 제거하고 다른 모든 프레임은 DCT 계수 제거를 하지 않았을 때, 복호화된 영상에서 관찰되는 왜곡을 나타낸 것이다. 그림4에서 알 수 있듯이, 이러한 왜곡 전과는 모두 단조 감소하는 형태로 전과됨을 보여주고 있다. 이러한 개개의 독립적인 실험 결과에 의한 왜곡 성분들의 합("sum_15%")은 모든 프레임에 대해 동시에 15%의 DCT 계수 제거를 수행한 경우("all_15%")와 거의 동일함을 알 수 있다. 따라서, 본 논문에서 제안한 성질 2가 매우 타당한 제안임을 알 수 있다.

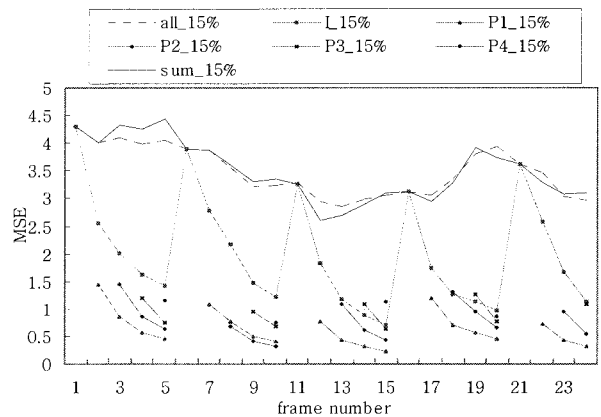


그림 4. Foreman 시퀀스에 대해 DCT 계수 제거 왜곡의 전과 특성과 왜곡 합
Fig. 4. Error propagation behavior of the CD distortion (Foreman)

그림 5는 CIF 형식의 Akiyo 시퀀스 300프레임을 GOP 크기=15, sub-GOP 크기=3으로 MPEG-4 심플 프로파일에 의해 1.2Mbps로 압축된 비트열에 대해 25%비트율 감소를 수행한 실험 결과이다. 위 실험과 동일하게 모든 B-프레임은 프레임 제거 대상이 되고, 남은 프레임들은 모두 25%의 비트율을 DCT 계수 제거를 통하여 구현하였다. 실험 방식

은 위와 동일하며, 그 결과를 통하여 알 수 있듯이 개개의 독립적인 실험 결과에 의한 왜곡 성분들의 합("sum_25%")은 모든 프레임에 대해 동시에 25%의 DCT 계수 제거를 수행한 경우("all_25%")와 거의 동일함을 알 수 있다. 따라서, 현재 프레임의 DCT 계수 제거 왜곡과 이전 프레임에서 전파되어 오는 독립적인 왜곡들의 합으로 전체 왜곡은 근사화되어짐을 실험적으로 확인할 수 있다.

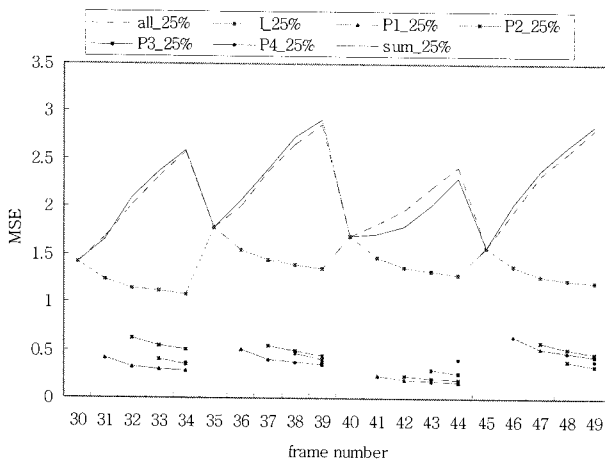


그림 5. Akiyo 시퀀스에 대해 DCT 계수 제거 왜곡의 전파 특성과 왜곡 합
Fig. 5. Error propagation behavior of the CD distortion (Foreman)

3. 성질 1과 성질 2의 실험 결과 검토

그림 4와 그림 5에서 I-프레임에서 동일한 비율의 DCT 계수 제거는 다른 프레임에 비해 매우 높은 왜곡 전파 특성을 가짐을 알 수 있다. 또한, Akiyo 시퀀스의 경우에는 Foreman 시퀀스에 비해 I-프레임에 의한 왜곡 특성이 매우 높음을 알 수 있으며, Foreman 시퀀스의 왜곡 전파 특성은 급격히 떨어지는 경향을 보이지만, Akiyo시퀀는 매우 완만하게 감소하는 특성을 갖는다. 이러한 특성은 압축된 동영상상의 콘텐츠가 갖는 특성으로서, 움직임이 많고, 공간적 복잡도가 높은 영상일수록 급격한 감쇄 특성을 보이며, 움직임이 적고 공간적 복잡도가 낮은 영상은 감쇄 특성이 완만한 것을 확인할 수 있다. 그리고 또한 각 GOP내에서 DCT 계수 제거에 의한 감쇄 기울기는 각 실험 영상별로 거의 일정하게 유지됨을 알 수 있다. 이러한 특성을 압축된

DCT 계수 영역에서 구할 수 있는 경우에는 계산량을 절감하면서 효과적인 비트열 트랜스코딩 기법이 제안될 수 있을 것이다.

V. 결론

본 논문은 압축된 동영상의 비트열에서 DCT 계수 제거를 통하여 비트열을 간단히 바꿀 수 있는 유효한 트랜스코딩 기법에 대해 다루었다. 이러한 기법은 압축 비트열의 일부를 제거하는 것으로 매우 낮은 계산량으로 구현될 수 있는 반면, 왜곡이 발생하게 되며 특히, 그 왜곡이 전파되는 특성을 갖게 되는데, 본 논문에서는 그 왜곡들이 갖는 통계적인 전파 특성을 제시하고 수식적인 간략화를 유도하였다. 그리고 모의실험을 통하여 전파 왜곡과 현 프레임의 DCT 계수 제거 왜곡은 서로 독립적으로 결정되며, 또한, 전체 왜곡은 전파 왜곡의 누적과 현 프레임의 DCT 계수 제거 왜곡의 합으로 간략화될 수 있음을 검증하였다.

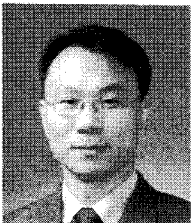
DCT 계수 제거로 인하여 발생하는 왜곡에 대하여 본 논문에서 제안한 두 가지의 통계적 분석 제안은 트랜스코딩 또는 전송오류로 인한 전체 왜곡을 예측하는 모델 개발에 사용될 수 있을 것이며, 나아가서는 압축된 비트열에서 왜곡을 고려한 트랜스코딩 기법 구현에 이용될 수 있을 것이다. 추가적인 연구로서는 화면내 예측 부호화를 사용하고 있는 H.264/AVC 기법에도 적용하여 화면내에서의 왜곡 전파 특성을 분석할 필요가 있으며, 또한, I-프레임과 P-프레임에서의 DCT 계수 제거에 따른 B-프레임에서의 관찰되는 왜곡에 대한 연구를 계속적으로 진행할 필요가 있다.

참고 문헌

- [1] J. -G. Kim, Y. Wang, S.-F. Chang, and H.-M. Kim, "An optimal framework of video adaptation and its application to rate adaptation transcoding," ETRI J., vol. 27, no. 4, pp. 341-354, Aug. 2005.
- [2] A. Vetro, C. Christopoulos, and H. Sun, "Video Transcoding Architectures and Techniques: An overview", IEEE Signal Processing Mag., vol. 20, pp. 18-29, Mar. 2003.
- [3] Huifang Sun; Kwok, W.; Zdepski, J.W., "Architectures for MPEG

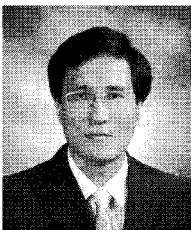
- compressed bitstream scaling," IEEE Trans. Circuits Syst. Video Technol., vol. 6, no. 2, pp. 191-199, April 1996.
- [4] J.Youn, J.Xin, and M.T.Sun, "Fast Video Transcoding Architectures for Networked Multimedia," in Proc. IEEE Int. Symp. Circuits and Systems, Geneva, Switzerland, vol. 4, May 2000, pp. 25-28.
- [5] A. Eleftheriadis, "Dynamic Rate Shaping of Compressed Digital Video," Ph.D. dissertation, Dept. Elec. Eng., Columbia University, June 1995.
- [6] J.Wang and S.Yu, "Dynamic Rate Scaling of Coded Digital Video for IVOD Applications," IEEE Trans. Consumer Electronics, vol. 44, pp. 743-749, Aug. 1998.
- [7] J.G. Kim, J.S.Kim, and H.M.Kim, "Simple Method for Estimating Distortion Caused by DCT Coefficient Drop in Compressed Domain," IEE Electronics Letters, vol. 40, no. 25, pp. 1581-1582, Dec. 2004.
- [8] P. Batra and A. Eleftheriadis, "Analysis of Optimal Dynamic Rate Shaping of Markov-1 Sources," in Proc., ICIP-2003, Barcelona, Spain, Sep. 2003.
- [9] Y. Liang and Y.-P. Tan, "A New Content-based Hybrid Video Transcoding Method," in Proc. ICIP'01, Oct. 2001, pp. 429-432.
- [10] K.R.Rao and P.Yip, Discrete Cosine Transform Algorithms, Advantages, Applications, Academic Press, INC., San Diego, 1990.
- [11] Jin-soo Kim, Jae-gon Kim, Kyeong-Ok Kang and Jin-woong Kim "A Distortion Control Scheme for Allocating Constant Distortion in FD-CD Video Transcoder," in Proc. IEEE Conf. Multimedia Expo (ICME'04), Taipei, Taiwan, June 2004.
- [12] A. Eleftheriadis and P. Batra, "Dynamic Rate Shaping of Compressed Digital Video," IEEE Trans. Multimedia, vol. 8, no. 2, pp. 297-314, April 2006.
- [13] H. Lee, Y. Kim, A.H. Rowberg and E.A. Riskin, "Statistical Distributions of DCT Coefficients and Their Application to an Interframe Compression Algorithm for 3-D Medical Images," IEEE Trans. Medical Image, vol. 12, no. 3, pp. 478-485, Sept. 1993.

저 자 소 개



김진수

- 1991년 : 경북대학교 전자공학과 학사
- 1993년 : 한국과학기술원 전기 및 전자공학과 석사
- 1998년 : 한국과학기술원 전기 및 전자공학과 박사
- 1997년~2000년 : 삼성전자 선임연구원
- 2000년~현재 : 한밭대학교 정보통신컴퓨터공학부 부교수
- 주관심분야 : 멀티미디어스트리밍, 영상통신, SVC, 멀티미디어 컨버전스, 디지털방송



김재곤

- 1990년 2월 : 경북대학교 전자공학과 학사
- 1992년 2월 : KAIST전기 및 전자공학과 석사
- 2005년 2월 : KAIST전기 및 전자공학과 박사
- 1992년 3월 ~ 2007년 2월 : 한국전자통신연구원 선임연구원/팀장
- 2001년 9월 ~ 2002년 11월 : 뉴욕 콜롬비아대학교 연구원
- 2007년 3월 ~ 2007년 8월 : 한밭대학교 전기전자제어공학부 전임강사
- 2007년 9월 ~ 현재 : 한국항공대학교 항공전자 및 정보통신공학부 조교수
- 주관심분야 : 비디오 신호처리/코딩, 디지털방송 미디어, 미디어 컨버전스, 멀티미디어 통신