

다중기술 변환기법을 이용한 비디오 부호화 알고리즘

류상욱, 양창모, 호요성
광주과학기술원 정보통신공학과
광주광역시 북구 오룡동 1번지

Video Coding using Multiple Description Transform Coding

Sang-Uk Ryu, Changmo Yang and Yo-Sung Ho
Kwangju Institute of Science and Technology (K-JIST)
1 Oryong-dong, Puk-gu, Kwang-ju, Korea

E-mail: suryu@geguri.kjist.ac.kr, changmo@geguri.kjist.ac.kr, hoyo@kjist.ac.kr

요약

IP 네트워크를 통해 실시간 비디오를 전송할 때 네트워크 특성을 고려하여 비디오 데이터를 부호화하면 패킷 손실로 인한 품질 열화를 최소화하여 더 좋은 품질의 비디오를 얻을 수 있다. 이를 위해 현재 네트워크의 정보를 빠르고 정확하게 얻어내는 메커니즘과 부호화 변수를 네트워크 상황에 적응적으로 조절하여 패킷 손실에 강인한 압축 및 전송 메커니즘의 두 가지 기능이 요구된다. 첫번째 메커니즘은 RTP(Real Time Transport Protocol)을 통해 구현될 수 있으며, 두번째 메커니즘을 위해 본 연구에서는 다중기술 변환부호화(Multiple Description Transform Coding) 기법을 적용한 비디오 부호화 알고리즘을 제안한다. RTP에서 제공하는 RTCP(Real Time Control Protocol) 정보를 이용하여 현재 네트워크 정보를 얻을 수 있으며, 다중기술 변환부호화 기법을 이용하여 현재의 패킷 손실률에서 최적의 품질을 보장하도록 부호화 변수를 조절할 수 있다. 본 논문에서는 다중기술 변환부호화 기법을 비디오 부호화에 적용하여 순수 비디오 정보에 추가되는 잉여 정보량과 패킷 손실에 대한 강인성 사이의 관계를 도출함으로써 다중기술 변환부호화 기법이 네트워크 적응적 부호화 방식에 적합한 방식임을 제시한다.

1. 서론

멀티미디어 실시간 전송을 위해 대부분의 응용이 UDP 계층에서 설계되며, UDP에 패킷 전송의 순차성을 추가하기 위해 RTP(Real Time Transport Protocol)를 채용한다. RTP는 UDP 기반의 프로토콜로서 이를 이용하여 수신단에서는 수신되는 데이터 패킷을 부호화 과정에서 RTP 헤더에 지정되는 순차번호를 이용하여 적절한 순서로 재정렬할 수 있다[1]. 또한 RTP는 순차적 전송이나 서비스 질(QoS)에 관한 보장이 없기 때문에 데이터 전송의 품질을 보장하기 위해 RTCP(Real Time Control Protocol)를 이용한다. RTCP는 RTP전송에 대한 제어 기능과 확인(Identification) 메커니즘을 제공한다. 순차적 전송에 관한 문제가 RTP를 이용하여 해

결되더라도 네트워크 혼잡 등으로 발생하는 패킷 손실에 대해서 RTP는 자체적으로 해결을 하지 못한다. 패킷 손실로 인해 발생하는 영상 왜곡을 최소화하기 위해 Joint Source/Channel Coding 기법에 기반을 둔 여러 방법들이 제시되었다[2]. 네트워크 대역폭의 시변 특성을 고려하여 멀티미디어 트래픽의 비트율을 제어하는 방식, 전방향 오류 보정(FEC)에 기반한 오류 제어 기법, 그리고 패킷 손실 효과를 최소화하는 오류 은폐 기법이 대표적인 경우이다. 최근 이론적 측면에서 주로 논의되었던 다중기술(Multiple Description) 기법이 멀티미디어 정보에 대해 구현되어 IP 네트워크에서 멀티미디어 전송에 적합한 기술로 주목받고 있다.

본 논문에서는 다중기술 변환부호화(Multiple Description Transform Coding) 기법을 이용한 비디오 부호화를 구현하였다. 제2절에서 정지영상에 대해 다중기술 변환부호화 기법이 어떻게 적용되는지에 대한 개요를 기술하고 정지영상에 대한 실험 결과를 제시한다. 제3절에서는 다중기술 변환부호화 기법이 적용된 비디오 부호화에 대해 기술하고, 제4절에서는 이에 대한 모의 실험을 통해 추가된 잉여 정보량과 패킷 손실에 대한 강인성과의 관계를 도출하고, 마지막 제5절에서 본 연구에 대한 결론을 맺는다.

2. 다중기술 변환 부호화(MDTC)

주어진 정보를 두 개의 기술(Description)로 표현하여 이를 각각 독립된 채널로 전송한 뒤 수신단에서는 모든 기술을 수신했을 경우 가장 좋은 품질의 데이터를 얻고, 독립된 각 채널로부터 하나의 기술만을 수신할 경우 비교적 만족할 만한 품질을 얻을 수 있도록 부호화하는 방식을 다중기술(Multiple Description) 기법이라 한다. 1980년대 초반 정보이론 측면에서 활발히 연구된 다중기술 기법은 1993년 V. Vaishampayan에 의해 제안된 MDSQ(Multiple Description Scalar Quantizer)로 구체적인 구현이 이루어졌으며, 이후 다중기술 변환부호화 방식 등 여러 방법들이 제시되었다[3]. 본 논문에서는 A. Leibman 등에 의해 제시된 다중기술 변

환부호화(MDTC) 방식[4]을 기반으로 보다 간단하면서도 패킷 손실에 강한 MDTC를 구현하였다.

2.1 MDTC 정지영상 부호화기

MD 부호화의 핵심은 주어진 정보를 두개의 기술로 표현하는 데에 있다. 8x8 영상 블록은 먼저 DCT 변환되고 그 결과 64개의 계수들은 각 계수의 분산값에 의해 32x2의 열로 재정렬된다. 각 32개의 계수는 독립적인 양자화기를 거쳐 양자화된다. 양자화된 각 계수 쌍은 상관 변환(Correlating Transform)과 정수-정수(Integer to Integer)변환을 거쳐 허프만 부호화되어 독립된 2개의 채널로 전송된다. 그림 1은 MDTC 부호화 과정을 나타낸다.

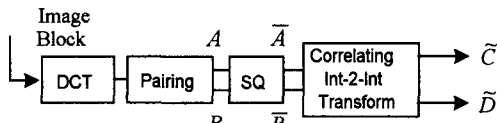


그림 1. MDTC 블록도

DCT 계수는 서로 독립적인 통계적 특성을 가지는 반면 상관변환은 각 계수가 서로 상관 관계를 가지게 하여 수신단에서 잃어버린 기술을 수신한 기술로부터 추정하는 것을 가능하게 한다. 여기서 상관 변환과 정수-정수 변환은 다음과 같다.

$$T = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, a = -c = \sqrt{\frac{\cot \theta}{2}}, b = d = \frac{1}{2a} \quad \dots (1)$$

$$\begin{aligned} 1. \bar{A} &= \begin{bmatrix} A \\ Q \end{bmatrix}, \bar{B} = \begin{bmatrix} B \\ Q \end{bmatrix} & 2. W &= \bar{B} + \begin{bmatrix} 1+c \\ d \end{bmatrix} \bar{A} \\ 3. \bar{D} &= [dW] - \bar{A} & 4. \bar{C} &= W - \begin{bmatrix} 1-b \\ d \end{bmatrix} \bar{D} \quad \dots (2) \end{aligned}$$

2.2 MDTC 정지영상 복호기

수신단에서 두개의 기술이 모두 유효할 경우 부호화의 역방식으로 전송된 데이터를 복호한다. 그림 2는 수신단에서 두개의 기술이 유효할 경우 복호기의 구조를 나타내며 여기서 역 정수-정수 변환은 다음과 같이 구해진다.

$$\begin{aligned} 1. W &= \bar{C} + \begin{bmatrix} 1-b \\ d \end{bmatrix} \bar{D} & 2. \bar{A} &= [dW] - \bar{D} \\ 3. \bar{B} &= W - \begin{bmatrix} 1+c \\ d \end{bmatrix} \bar{A} & 4. \hat{A} &= \bar{A}Q, \hat{B} = \bar{B}Q \quad \dots (3) \end{aligned}$$

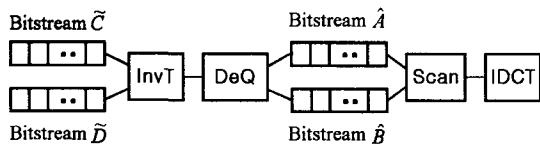


그림 2. 모든 기술이 유효할 경우 복호기

그러나 수신단에서 하나의 기술만이 유효할 경우,

수신된 하나의 기술로부터 잃어버린 기술을 추정하고 복원된 계수 쌍으로부터 역상관 변환, 역양자화 그리고 IDCT를 수행한 후 전송된 영상을 복원한다. 그림 3은 수신단에서 하나의 기술이 유효할 경우 복호기의 구조를 나타내며, 수신한 기술 \bar{C} 로부터 $\hat{D}(\bar{C})$ 를 추정하는 과정과 이때 발생하는 왜곡은 다음과 같이 구한다.

$$\hat{D}(\bar{C}) = \gamma \bar{C} \quad \& \quad \gamma = \frac{\frac{ac}{Q^2} \left(\sigma_A^2 + \frac{Q^2}{12} \right) + \frac{bd}{Q^2} \left(\sigma_B^2 + \frac{Q^2}{12} \right)}{\frac{a^2}{Q^2} \left(\sigma_A^2 + \frac{Q^2}{12} \right) + \frac{b^2}{Q^2} \left(\sigma_B^2 + \frac{Q^2}{12} \right)} \quad \dots (4)$$

$$\begin{aligned} D_1 &= (a^2\gamma - ac)^2 \sigma_A^2 + (1 + ab\gamma - bc)^2 \sigma_B^2 + \frac{Q^2}{12} (ac - a^2\gamma)^2 + \\ &\frac{Q^2}{12} (bc - ab\gamma)^2 + (1 - ad + ba\gamma)^2 \sigma_A^2 + (b^2\gamma - bd)^2 \sigma_B^2 + \\ &\frac{Q^2}{12} (da - ba\gamma)^2 + \frac{Q^2}{12} (bd - b^2\gamma)^2 \quad \dots (5) \end{aligned}$$

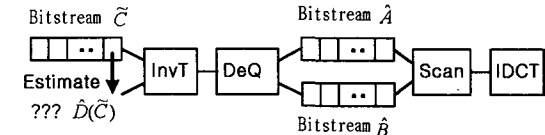


그림 3. 하나의 기술이 유효할 경우 복호기

식 (4)에서 구한 추정값에 따라 식(5)의 왜곡이 결정되며, $\theta = 0^\circ$ 일 때 최소의 왜곡을 가지며, $\theta = 90^\circ$ 일 때 최대의 왜곡을 가진다. 한편 그림 4에서 알 수 있듯이, θ 의 값이 0° 에서 90° 까지 변환에 따라 두 기술의 비트량이 감소하게 된다. 따라서 θ 값을 변동시켜 많은 잉여 정보(작은 왜곡)를 혹은 적은 잉여 정보(큰 왜곡)를 추가할 수 있으며, 이에 따라 복원되는 영상의 화질이 결정된다.

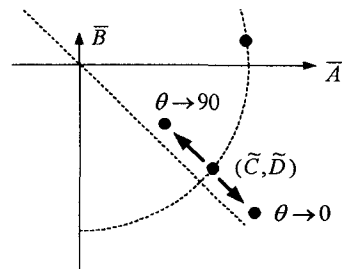


그림 4. 잉여 정보 추가 원리

2.3 정지영상에 대한 실험결과

본 절에서는 앞 절에서 설명한 다중기술 변환부호화(MDTC) 기법을 정지영상에 적용한 모의 실험 결과를 제시한다. 512x512 LENA 영상에 대해 θ 값을 변화하면서 복원된 영상의 PSNR을 비교하였다. 두개

의 기술이 모두 유효한 경우, 첫번째 기술만 유효한 경우, 그리고 두번째 기술만 유효한 경우에 대해 실험하여 세가지 경우에 대해 왜곡 정도(D_0, D_1 & D_2)를 PSNR을 이용하여 비교하였다. 그림 5는 θ 와 PSNR과의 관계를 나타낸다. 두개의 기술이 모두 유효할 경우 같은 양자화기를 이용하는 JPEG과 동일한 화질을 얻을 수 있다. θ 가 커질수록 예측 오류가 커져서 더 큰 열화가 발생함을 알 수 있으며 θ 가 작을수록 추가된 많은 잉여 정보로 인해 하나의 기술만으로도 JPEG 화질을 얻을 수 있다. 일반적으로 하나의 기술만이 유효하더라도 JPEG과 비교해서 2~3 dB 화질 차이의 좋은 품질의 영상을 복원할 수 있다.

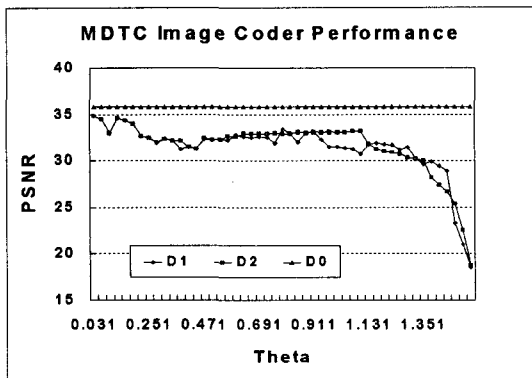


그림 5. MDTC: θ 와 PSNR과의 관계

3. MDTC 비디오 부호화

본 연구에서 구현된 MDTC 비디오 부호화기는 H.263 부호화기에서 매크로블록 잉여 정보(Residual Error)가 Inter 혹은 Intra 모드로 부호화될 때, 이를 다중기술 기법으로 부호화하여 두 개의 기술을 만든다. 즉 하나의 매크로 블록 당 두 개의 기술이 생성된다. 패킷화 단위가 프레임 단위일 경우 각 프레임 당 두 개의 패킷이 생성된다. 픽처, GOB, MB 헤더 정보 및 움직임 벡터는 각 패킷에 동일한 정보로 복사되고, 잉여 정보에 대한 각 기술이 각 패킷에 수송된다. 이때, 잉여 정보에 대한 다중 기술 기법은 앞 절의 정지영상에 대한 MDTC를 따른다.

수신단에서 각 기술을 수송한 모든 패킷이 수신될 경우, 복호기는 각 패킷에 동일하게 복사된 헤더와 움직임 정보 그리고 이전 프레임을 이용하여 움직임 보상된 영상을 복원한다. 각 패킷에서 수송된 잉여 정보(Residual Error)는 정지영상에 대한 MDTC 복호 과정을 거쳐 움직임 보상된 영상에 더해진다. 두 개의 기술 중 어느 하나의 기술만을 수신할 경우에, 먼저 수신된 패킷에서 수송된 헤더 정보와 움직임 벡터 정보로부터 움직임 보상의 과정을 수행한다. 다음 단계로써 수신한 패킷이 수송한 잉여 정보(Residual Error)로부터 잃어버린 기술을 추정하고, 그 결과를 이용하여 잉여 정보(Residual Error)를 복원하여 움직임 보상된 영상에 더해져 한 프레임의 영상을 복원한다. 모든 기술이 수신단에서 모두 유효하지 않을 경우, 수신단

은 복원된 이전 프레임을 현재 프레임으로 대체하는 오류 은폐 과정을 수행한다.

4. 실험 결과

본 실험에서는 순수 비디오 정보에 잉여 정보를 추가하여 잉여 정보량 대 패킷 손실에 대한 강인성과의 관계를 도출하였다. Carphone 시퀀스에 대해 양자화 단계를 7로 고정한 VBR과 발생 비트율을 128kbps로 고정한 CBR에 대해 실험하였다. MDTC 비디오 부호화기에서 발생하는 잉여 정보량은 θ 값으로 조절할 수 있다. 본 모의 실험에서 VBR 비디오의 경우 일정한 양의 비디오 정보에 가변량의 잉여 정보를 추가하였으며, CBR 비디오의 경우 순수한 비디오 정보량과 잉여 정보량의 비를 조절하여 잉여 정보량을 변화시켰다. VBR, CBR 각각에 대해 가변의 잉여 정보량과 패킷 손실률에서 수신된 비디오의 평균 PSNR을 측정하여 잉여 정보량과 패킷 손실에 대한 강인성을 비교하였다. 표 1은 다양한 패킷 손실률(Packet Loss Rate)에서 VBR의 잉여 정보량(Redundancy)과 평균 PSNR을 비교하였다.

표 1. 잉여 정보량, 패킷 손실률 & PSNR 관계 (VBR)

PLR R	0%	1%	5%	10%	20%
85.2%	34.57	34.49	34.17	33.77	33.28
73.8%	34.57	34.49	34.16	33.78	33.28
33.8%	34.57	34.11	32.49	30.85	29.13
29.3%	34.57	34.07	32.27	30.58	28.74

일정한 순수 비디오 정보량에 가변의 잉여 정보를 추가하였으므로 패킷 손실이 없는 상황에서는 H.263 부호화기와 같은 품질을 얻을 수 있다. 한편 추가된 더 많은 잉여 정보량은 더 높은 패킷 손실률에서도 좋은 품질을 보장한다. 즉, 비트율의 회생을 통해 패킷 손실의 강인성을 얻을 수 있다. 표 2는 CBR의 경우 패킷 손실에 대한 강인성과 잉여 정보량과의 관계를 보인다. 여기서 PLR은 패킷 손실률을 나타내며 R은 순수 비디오 정보와 잉여 정보의 비율을 나타낸다. CBR의 경우 잉여 정보의 추가는 순수 비디오 정보의 감소를 의미하므로 패킷 손실이 없는 경우 같은 비트량을 가지는 H.263 부호화기에 비해 다소 떨어지는 품질이 예상된다. 그러나 패킷 손실이 빈번한 환경에서는 같은 비트량을 가지는 H.263 부호화기보다 패킷 손실에 더 강인하다.

표 2. 잉여 정보량, 패킷 손실률 & PSNR 관계 (CBR)

PLR R	0%	1%	5%	10%	20%
54:46	29.96	29.94	29.88	29.80	29.73
58:42	30.07	30.05	29.99	29.9	29.83
75:25	31.12	30.09	29.98	28.69	27.51
78:22	31.21	31.01	29.96	28.91	27.61

다음 실험은 다중기술 기법과 [5]에서 제시한 오류 은폐 기법을 이용한 비디오 부호화기의 성능을 다양한 패킷 손실 환경에서 비교하였다. 본 실험에서 하나의 비디오 프레임은 두 개의 패킷으로 나누어진다. 오류 은폐 기법을 사용한 부호화기는 프레임의 짝수 번째 GOB와 홀수번째 GOB를 각각 다른 패킷에 수송한다. 즉 한 프레임에 대해 두 패킷이 발생하고, 하나의 패킷이 손실되었을 때 수신한 나머지 패킷으로부터 움직임 벡터를 구한 뒤 움직임 보상으로 패킷 손실로 인한 오류를 은폐한다. 이러한 방식을 SDEC (Single Description with Error Concealment)이라 하자. VBR과 CBR의 경우 SDEC와 MD의 잉여 정보 대 패킷 손실에 대한 강인성을 가변의 패킷 손실 환경에서 비교하였다.

VBR의 경우, 일정한 순수 비디오 정보에 85%의 잉여 정보를 추가한 MDhigh, 29% 잉여 정보의 MDlow 그리고 순수 비디오 정보만을 가지는 SDEC에 대해 실험하였다. 그림 6에서 VBR-SDEC는 작은 패킷 손실률에서도 급격한 화질 열화가 나타나는 반면, MDhigh는 모든 패킷 손실 환경에서도 일정한 품질이 보장됨을 알 수 있다. 또한, MDlow는 패킷 손실률이 증가함에 따라 선형적인 화질 열화가 발생하나, SDEC에 비해 3-5dB이상의 좋은 화질을 얻을 수 있다.

CBR의 경우, 순수 비디오 정보와 잉여 정보 비가 54:46인 MDhigh와 78:22의 MDlow 그리고 같은 비트율을 가지는 SDEC에 대해 실험하였다. MDhigh는 모든 패킷 손실률에서 일정한 화질을 보장하며, MDlow는 패킷 손실률에 따른 선형적인 화질 열화를 보인다. MDlow는 잉여 정보의 추가, 즉 순수 비디오 정보량 감소로 인해 패킷 손실이 없는 경우, SDEC에 비해 약 1dB 정도의 화질 열화를 초래하게 된다. 그러나 패킷 손실률이 2% 이상일 경우, 잉여 정보로 얻어지는 패킷 손실에 대한 강인성으로 인해 MD가 SDEC보다 약 2~3 dB의 화질 향상을 가진다.

5. 결론

본 연구에서는 다중기술 변환부호화 기법을 비디오 부호화에 적용하여 순수 비디오 정보에 추가되는 잉여 정보량과 패킷 손실에 대한 강인성 사이의 관계를 도출하였다. 각 매크로 블록의 부호화시 잉여 정보 (Residual Error)를 다중기술 변환부호화 방식을 통해 부호화 하였고, 패킷 손실에 대한 강인성을 위해 추가되는 잉여 정보는 변환 변수 θ 의 값을 변화시켜 그 양을 조절하였다. VBR의 경우 일정한 순수 비디오 정보 외에 추가적인 잉여 정보를 추가하여 비트율에서 효율성을 희생하여 패킷 손실에 대한 강인성을 획득할 수 있었다. 한편 CBR의 경우 잉여 정보의 증가(순수정보량의 감소)로 패킷 손실이 없는 경우 1dB 정도의 화질 열화를 희생하여 2% 이상의 모든 패킷 손실 환경에서 오류 은폐 기법보다 3dB 이상 좋은 화질을 얻을 수 있었다.

감사의 글

본 연구는 광주과학기술원(K-JIST) 초고속망네트워킹 연구센터(UFON)를 통한 한국과학재단 우수연구센터

(ERC)와 교육부 두뇌한국21(BK21) 정보기술사업단의 지원에 의한 것입니다.

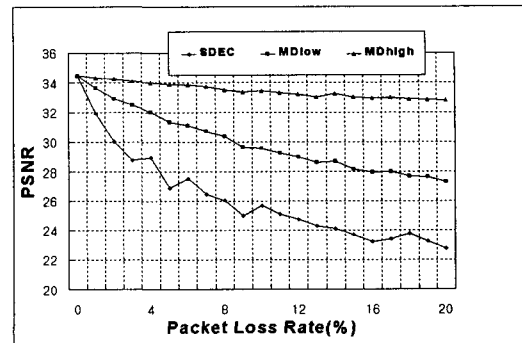


그림 6. MDTC와 오류은폐기법 성능비교(VBR)

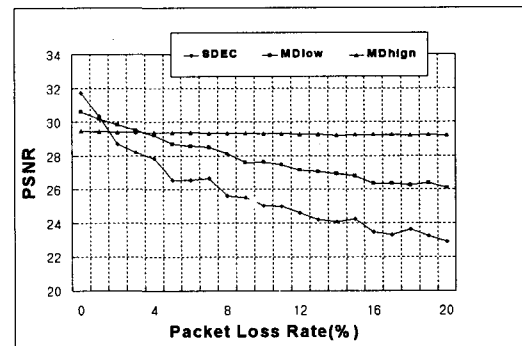


그림 7. MDTC와 오류은폐기법 성능비교(CBR)

참고 문헌

- [1] R. Frederick H. Schulzrinne, S. Casner and V. Jacobson, "RTP: A Transport Protocol for Real-time Applications," RFC 1889, Jan. 1996.
- [2] Y. Wang and Q.F. Zhu, "Error Control and Concealment for Video Communications: A Review," *Proceeding of IEEE*, vol. 86, no. 5, pp. 974-997, May 1998.
- [3] Y. Wang, M. Orchard and A. Reibman, "Optimal Pairwise Correlating Transforms for Multiple Description Coding", *1998 IEEE International Conference on Image Processing*, pp. 679-683, October 1998.
- [4] A. Reibman, H. Jafarkhani, Y. Wang, M. Orchard, and R. Puri, "Multiple description coding for video using motion compensated prediction", *1999 IEEE International Conference on Image Processing*, pp. 837-841, October 1999.
- [5] ITU Telecom. Standardization Sector of ITU, "Video Codec Test model near-term, Version 10 (TMN10)," *ITU-T Study Group 16, Video Experts Group, Document Q15-D-65*, Apr. 1998.