

H.263 Annex I 기반 화면내 부호화 기법의 성능개선

유 국 열

영남대학교 공과대학 전자정보공학부

Improved Intraframe Coding Method based on H.263 Annex I

K.-y. Yoo

School of Electrical Engineering and Computer Science, Yeungnam University

요약문

H.263부호기의 화면내 부호화(intraframe coding)를 위한 부가 부호기법인 Annex I는 기존의 JPEG, MPEG-1, -2 부호기들의 화면내 부호방법들과는 달리 부호화하고자 하는 현재 블록의 DCT 계수를 기부호화된 블록들의 DCT 계수들을 이용하여 예측부호화 (Differential Pulse Coded Modulation; DPCM)를 행하고, 이를 통한 부호화 이득의 향상을 얻고 있다. 본 논문에서는 이런 H.263 Annex I의 예측기법의 공간영역에서의 물리적인 의미를 유도한다. 이를 통해서 H.263 Annex I의 예측기법의 비효율성을 지적하고, 영상신호의 통계적 특성에 맞게 예측방식을 수정한다. 제안된 DCT 계수 예측 방식과 기존의 H.263 Annex I 방식의 이론적인 예측성능을 평가하고, H.263 부호기에 적용하여 제안방식의 효율성을 검증한다

Abstract

The H.263 Annex I method for the intraframe coding is based on the prediction in DCT domain, unlike JPEG, MPEG-1, and MPEG-2 where the intraframe coding uses block DCT, independent of the neighboring blocks. In this paper, we show the ineffectiveness of H.263 Annex I prediction method by mathematically deriving the spatial domain meaning of H.263 Annex I prediction method. Based on the derivation, we propose a prediction method which is based on the spatial correlation property of image signals. From the experiment and derivation, we verified the proposed method.

Keywords: H.263 Annex I, DCT Coefficient Prediction, VLC, Correlation

I. 서론

인터넷의 광범위한 보급과 더불어 이동 통신망의 광대역화 추세로 인해 방송과 같은 대규모의 공중망을 통한 영상신호의 전송에서 개인화되고, 휴대화된 영상통신 시스템에 대한 관심이 높아지고 있다. 이들 통신망을 통한 영상 신호의 전송문제는 예전의 H.261과 같이 종합통신망(Integrated Service Digital Network; ISDN)과 같이 보장된 대역폭을 가진 통신망을 통한 영상신호의 전송문

제와는 확연히 구별된다[1]. 즉, 인터넷의 경우 이용 가능한 대역폭이 가변적이기 때문에 망폭주(network congestion)에 의해서 전송된 패킷의 손실이 발생하고 [1,2], 이동 통신망의 경우 비트오류나 군집오류로 인해서 전송된 영상 비트스트림의 손상은 필연적이다[3]. 이런 망의 오류는 시간방향의 예측부호화기법에 기반을 둔 영상신호의 경우, 시간방향의 오류전파(error propagation)에 의해서 장시간에 걸친 영상 신호의 손상을 불러 일으킨다[3]. 그러므로 불안정한 통신망을 통한 영상신호의 전송문제는 효율적인 오류의 복원을 필수적으로 포함하고 있다. 손상된 영상을 복원하는데 가장 효과적인 방식이 화면내 부호화 기법(Intraframe coding method)이다[4,5]. 그러나, 화면내 부호화기법의 경우 예측 부호화를 사용하는 화면간 부호화(Interframe coding method)에 비해서 화면당 발생 비트율의 과다로 인해 네트워크에 상당한 부담을 줄 수 있기 때문에, 오류 복원시에 빈번한 사용은 제한적이다[6]. 그러므로, 효율적인 화면내 부호화 기법은 효과적인 인터넷 또는 이동망을 통한 영상신호의 전송에 있어서 필수적인 기법이다.

H.263++ 부호기의 부가 부호기법인 Annex I는 기존의 JPEG, MPEG-1, -2[7,8,9]와는 달리 부호화하고자 하는 현재 블록의 DCT 계수를 기부호화된 블록들의 DCT 계수들을 이용하여 예측부호화 (Differential Pulse Coded Modulation; DPCM)를 행함으로써 부호화 이득(coding efficiency)의 향상을 얻고 있다[10].

본 논문에서는 H.263 Annex I 방식의 효율성을 검증하기 위해서, H.263 Annex I에서 사용한 DCT영역의 예

측기법을 수학적 유도를 통해서 공간영역에서의 의미를 보인다. 이를 통해서, 현재의 H.263 Annex I의 DCT 계수 예측방식이 공간영역에서 8화소(pixel)의 거리에 있는 화소로부터의 예측하는 방식임을 보인다. 이는 영상신호의 경우 공간적으로 거리가 멀어질수록 상관성(correlation)이 급격히 감소하는 특성을 감안하면 H.263 Annex I의 예측방식이 매우 비효율적임을 의미한다. 이런 비효율적인 예측성능을 개선하기 위해서 본 논문에서는 공간적으로 가장 가까운 거리에 있는 화소로부터의 예측을 할 수 있도록 DCT 영역에서의 예측방식을 변형한다. 또한 제안한 DCT 계수 예측방식을 H.263 부호화에 맞게 변형하여 부호화 방식을 제안하고, 실험을 통해서 제안방식의 성능을 검증한다.

본 논문은 다음과 같이 구성되어 있다. II장에서 기존의 H.263 Annex I 부호화 방식을 설명한다. III장에서 H.263 Annex I에서 사용한 DCT 영역의 예측 부호화 기법을 분석 및 예측방식의 비효율성을 제시한다. 또한, 분석 결과를 바탕으로 제안 방식을 유도한다. IV장에서는 실험을 통해서 제안방식과 기존의 H.263 Annex I 방식의 성능을 평가하고, V장에서 결론을 서술한다.

II. H.263 Annex I 부호화 기법 [10, 11]

그림 1은 H.263 Annex I 부호화 기법의 전체 블록선도를 보여준다.

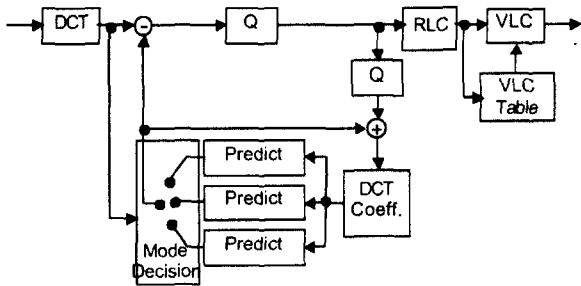


그림 1: H.263 Annex I 화면내 부호화기

입력된 영상은 MB(Macroblock) 단위로 영역화(segmentation)를 한다. 각 매크로 블록은 4개의 휘도 블록(luminance block)들과 2개의 색차 블록(chrominance block)으로 구성된다. 이때 각 블록은 8x8의 화소들로 구성된다. 부호화시에 부호화 모드 결정의 기본 단위는 MB 단위로 이루어진다. MB내의 각 블록의 화소값들은 DCT를 통해서 DCT 계수로 변환된다. 이때 DCT 변환은 아래 식으로 정의된다.

$$C(u, v) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} c(i, j) \cos\left(\frac{2\pi}{N} ui\right) \cos\left(\frac{2\pi}{N} vj\right) \quad (1)$$

여기에서 $c(i, j)$ 는 $[i, j]^T, i, j \in [0, N-1]$ $C(u, v)$ 는 $[u, v]^T, u, v \in [0, N-1]$ 에서의 DCT 계수를 의미한다. 또한 u 와 v 는 가로축 및 세로축 방향들의 주파수들을 의미한다. H.263의 경우 블록은 8x8의 화소들로 구성되어 있으므로, N 은 8에 해당한다. 이때 현재 블록의 DCT 계수들은 인접한 블록들의 DCT 계수값들을 이용해서 예측을 하고, 예측오차를 양자화한다. 이때 예측하는 방법은 다음과 같은 세 가지의 예측 방법들이 있고, 각 예측 방법들을 부호화 모드라고 부른다. 현재 블록을 c 로 나타내었고, 블록들 a 와 b 는 현재블록의 상측 블록 및 좌측 블록을 의미한다. 그리고, 본 논문에서는 블록 a, b, c 의 DCT 변환 계수들을 각기 대문자 A, B, C 로 표기한다.

Mode 0: DC 예측 모드

$$\hat{C}(0,0) = (A(0,0) + B(0,0)) / 2 \quad (2)$$

$$\hat{C}(u,v) = 0, \text{ for } u, v \neq 0 \quad (3)$$

Mode 0는 $\hat{C}(u,v)$ 는 $C(u,v)$ 의 예측치를 의미한다. 현재 블록의 DCT 계수, DC값만 예측오차를 부호화하고, DC 계수를 제외하고 다른 DCT 계수들은 예측되지 않고, DCT계수값 자체로 부호화를 행한다.

Mode 1: 상측 블록을 이용한 DC 및 AC계수 예측 모드

$$\hat{C}(u,0) = A(u,0), u = 0, \dots, N-1 \quad (4)$$

$$\hat{C}(u,v) = 0, \text{ for } v \neq 0 \quad (5)$$

Mode 2: 좌측 블록을 이용한 DC 및 AC계수 예측 모드

$$\hat{C}(0,v) = B(0,v), v = 0, \dots, N-1 \quad (6)$$

$$\hat{C}(u,v) = 0, \text{ for } u \neq 0 \quad (7)$$

위의 각 예측 방식들은 평가함수로 SAD(Sum of Absolute Difference)를 사용하여 가장 작은 예측 오차치를 가지는 모드를 현재 MB내의 모든 블록들의 예측 방식으로 선택한다[11].

예측된 오차계수 및 예측이 적용되지 않은 DCT계수들은 양자화(quantization)를 행한 후, 스캐닝을 통해서 일차원 신호로 변환된다. 스캐닝을 통해서 일차원 신호로 변형된 DCT 계수들은 RLC (Run-Length Coding)을 이용해서 기호(symbol)로 사상(mapping)되고, 각 기호별로 VLC표에 지정된 비트패턴으로 이진 부호화된다. 이때 주의할 점은 Annex I를 사용할 경우에 VLC표는 Annex I를 사용하지 않을 때와는 다른 VLC표를 사용한다[10]. 이는 예측자 DCT계수의 통계적인 특성이 예측을 행하지 않은 DCT계수와 상이하기 때문이다.

III. 제안한 DCT 계수 예측 방법

1 H.263 Annex I DCT 계수 예측 방식의 분석 및 제안 DCT계수 예측 방법의 유도

2장에서 설명한 H.263 Annex I DCT 계수 예측방식의 물리적인 의미를 살펴보고자 한다. 우선 Mode 1의 물리적인 의미를 살펴본다. 식 (1)과 식 (4)를 이용해서 DCT 예측차 계수를 표현하면 식 (10)으로 표현된다.

$$E(u,0) = C(u,0) - \hat{C}(u,0) \tag{8a}$$

$$= \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \{c(i,j) - a(i,j)\} \cos\left(\frac{2\pi}{N} ui\right) \cos\left(\frac{2\pi}{N} vj\right) \tag{8b}$$

$$= \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \{c(i,j) - a(i,j)\} \cos\left(\frac{2\pi}{N} ui\right) \tag{8c}$$

여기에서 $E(u,0)$ 는 예측차 DCT 계수를 나타낸다. 현재 블록 위치를 나타내는 위치벡터를 영상내의 위치로 변환하여 식 (9)로 나타내면, 상측 블록의 위치는 식 (10)으로 나타낼 수 있다.

$$l(x,y) = c(i,j) \tag{9}$$

$$l(x,y-N) = a(i,j) \tag{10}$$

여기에서 $[x,y]^T$ 는 $[i,j]^T$ 에 해당하는 위치벡터를 의미한다. 식 (10)을 살펴보면, 현재의 H.263 Annex I의 DCT 계수 예측방법은 N화소, 즉 8화소의 거리에 떨어져 있는 화소로부터 공간예측을 한 것임을 알 수 있다. 이는 영상 신호의 상관성은 화소간의 거리가 멀어질수록 급격히 감소하는 성질에 비추어 볼 때 상관성이 매우 낮고, 예측 효율이 낮음을 의미한다 [12, 13].

본 논문에서 $c(i,j)$ 와 $a(i,j)$ 를 $a(i,N-1)$ 로 대체해서 예측하는 것을 제안한다. 식 (11)는 본 논문에서 제안한 예측 방식을 보여준다. 예측 효율을 증대시키기 위해서 상측 블록의 가장 하단에 위치한 화소인 가장 가까운 거리에 있는 상측 블록의 화소를 이용해서 예측하도록 예측 방식을 변형한다.

$$\begin{aligned} E(u,0) &= \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (c(i,j) - a(i,N-1)) \cos\left(\frac{2\pi}{N} ui\right) \\ &= C(u,0) - N \sum_{i=0}^{N-1} a(i,N-1) \cos\left(\frac{2\pi}{N} ui\right) \end{aligned} \tag{11}$$

식 (11)에 따르면, 현재 블록의 각 화소와 상측 블록 내에 위치한 예측 화소간의 거리는 1에서 8까지의 거리에 변형되었음을 알 수 있다.

Mode 2의 경우는 Mode 1의 해석과 동일하므로, 본 논문에서 유도를 생략하고자 한다. 또한, DC 성분의 경우에도 앞에서 설명한 상관성을 고려하여 현재 블록과 가장 가까운 화소줄(pixel line)의 DC값으로 대체한다. 제안 방식을 정리하여 3.2절에서 나타내었다.

2 제안한 DCT 예측 기법

앞 절에서 유도한 변형된 예측 방식을 기반으로 DCT 계수 예측방식을 제안한다.

Mode 0: DC 예측 모드

$$A(0,0) = \frac{1}{N} \sum_{i=0}^{N-1} a(i, N-1) \tag{12}$$

$$B(0,0) = \frac{1}{N} \sum_{j=0}^{N-1} b(N-1, j) \tag{13}$$

Mode 1: 상측 블록을 이용한 DC 및 AC계수 예측 모드

$$\hat{C}(u,0) = N \sum_{i=0}^{N-1} a(i, N-1) \cos\left(\frac{2\pi}{N} ui\right), u=0, \dots, N-1 \tag{14}$$

Mode 2: 좌측 블록을 이용한 DC 및 AC계수 예측 모드

$$\hat{C}(0,v) = N \sum_{j=0}^{N-1} b(N-1, j) \cos\left(\frac{2\pi}{N} vj\right), v=0, \dots, N-1 \tag{15}$$

제안 예측 방식의 모드 선정 방식은 H.263 Annex I 방식과 동일한 방식을 사용하였다.

IV. 모의 실험

제안 방식의 성능의 평가는 ITU-T SG16/Q.15 H.263++ 표준화 회의에서 제시한 공통 실험 환경에 따라 수행하였다[14]. 실제적인 부호화에서 제안방식의 성능을 평가하기 위해서 실험에 사용한 영상은 News (QCIF, 10Hz, 100 frames)과 Paris (CIF, 10Hz, 100 frames)이다. H.263 Annex I 부호기에서 DCT 계수 예측 부분을 제안 방식으로 대체한 부호기와 H.263 Annex I 부호기의 부호화 성능을 비교하였다. ITU-T SG16/Q.15 H.263++ 표준화 회의에서 제시한 공통 실험 환경에 따라 양자화 간격(quantization stepsize)를 고정시키고, 발생한 비트량을 비교하였다. 공통 실험 환경에서 제시된 양자화 간격은 QP = 4, 5, 7, 10, 15, 25이다. 또한, 제안 방식은 화면내 부호화 기법을 다루므로, 모든 영상 sequence 를 화면내 부호화를 한 후 평균 발생 비트율과 평균 PSNR (Peak Signal to Noise Ratio)를 구하였다. 표 1과 2에 따르자면, 제안 방식이 H.263 Annex I 방식에 비해서 약간의 PSNR 개선환경에서 약 3%정도의 비트량 감소효과가 있음을 알 수 있다. 양자화 간격이 감소할수록 제안 방식의 개선효과가 증대됨을 알 수 있다. 이는 제안 방식이 예측시에 좌측 및 상측 블록의 한 화소줄만을 이용해서 예측을 하기 때문에, 상대적으로 부호화 잡음의 영향에 좀 더 민감함에 기인한다.

제안 방식의 상관성 개선 효과에 비해서 부호화 이득의 개선 효과가 낮은 이유는 본 논문에서 제시된 예측 방식에 의한 RUN-LEVEL쌍의 통계적인 특성과는 상이한 VLC표인 H.263 Annex I의 VLC표를 비교에 이용했기 때문으로 판단된다. 그러므로, 본 연구에서는 제안 방식에 적합한 VLC표를 작성하는 것을 계속연구로 남겨두고 있다.

표 1: 평균 PSNR을 통한 성능비교

Q P	News		Paris	
	H.263	제안방식	H.263	제안방식
4	41.07	41.07	40.12	40.15
5	38.94	38.96	37.97	37.98
7	36.82	36.85	35.76	35.78
10	34.83	34.83	33.72	33.73
15	32.02	32.07	30.89	30.91
25	28.92	29.00	27.81	27.85
평균	35.43	35.46	34.38	34.40

표2: 발생비트량을 통한 성능비교; H.263열 (kbits/frame), 제안방식: H.263에 대한 비트감축율(%)

Q P	News		Paris	
	H.263	제안방식	H.263	제안방식
4	50.07	3.20	255.8	2.94
5	42.1	3.40	212.9	2.95
7	32.9	3.21	167.0	2.88
10	25.3	2.71	128.7	2.82
15	17.6	2.52	89.5	2.74
25	10.9	2.19	54.5	1.92
평균	35.43	2.87	151.4	2.83

V. 결론

본 논문에서는 H.263 Annex I에서 사용되는 DCT 계수 예측방식을 분석하여 예측 효율이 비효율적임을 지적하였다. 또한 분석결과를 바탕으로 변형된 DCT 계수 예측방식을 제시하였다. 제안 방식은 영상 신호의 상관성이 근접한 화소일수록 높다는 통계적인 특성에 근거하여 DCT 계수 예측시에 현재 블록에 가장 상관성이 높은 화소값들을 이용해서 DCT 계수 예측이 가능하도록 변형하였다. 제안 방식의 근거를 제시하기 위해서 제안 방식과 H.263 Annex I 방식에 의한 예측 방식들의 상관성을 비교하였고, 실제적인 부호기에서의 성능을 평가하였다. 모의 실험 결과를 통해서 제안 방식이 기존의 H.263 Annex I 방식에 비해서 부호화 효율성(coding efficiency) 개선효과가 있음을 검증하였다. 향후과제로는 DCT 계수의 부호화에 직접적인 영향을 미치는 VLC표를 제안 방식에 적합하도록 설계하는 것을 남겨놓고 있다.

참고 문헌

[1] J. W. Kim, C.-C. Jay Kuo, et. al., "TCP-friendly Internet video streaming employing variable frame-rate encoding and interpolation," *IEEE*

Trans. on Circ. Syst. Video Technol., vol. 10, no. 7, pp. 1164-1177, Oct. 2000.

[2] G.J. Conklin, G.S. Greenbaum, et. al., "Streaming video over the Internet: approaches and directions," *IEEE Trans. on Circ. Syst. Video Technol.*, vol. Vol. 11, no. 3, pp. 282-300, March 2001.

[3] K.-y. Yoo, "Adaptive resynchronisation marker positioning method for error resilient video transmission," *IEE Electronics Letters*, vol. 34, no. 22, pp. 2084-2085, Nov. 1998.

[4] Judy.Y. Liao and John Villasenor, "Adaptive intra block update for robust transmission of H.263," *IEEE Trans. on Circ. Syst. Video Technol.*, vol. 10, no. 1, pp. 30-35, Jan. 2000.

[5] G. Cote and F. Kossentini, "Optimal intra coding of blocks for robust video communication over Internet," *Signal Processing:Image Communication*, vol. 15, nos. 1-2, pp. 25-34, Sept. 1999.

[6] Byeong-hee Roh and Jae-kyoon Kim, "Starting time selection and scheduling methods for minimum cell loss ratio of superposed VBR MPEG video traffic," *IEEE Trans. on Circ. Syst. Video Technol.*, vol. 9, no. 6, pp. 920-928, Sept. 1999.

[7] W.B. Pennebaker and J.L. Mitchell, *JPEG: still image data compression standard*, New York: Van Nostrand, 1993.

[8] MPEG Editorial Group, "Coding of moving pictures and associated audio for DSM at upto about 1.5Mbps/s: ISO/IEC 11172-2 Video IS," *ISO/IEC JTC1/SC29/WG11*, April 1993.

[9] MPEG Editorial Group, "Generic coding of moving pictures and associated audio: ISO/IEC 13818-2 Video IS," *ISO/IEC JTC1/SC29/WG11*, April 1995.

[10] ITU-T, "ITU-T Recommendation H.263, Video coding for low bit rate communication," *ITU-T*, Feb. 1998.

[11] Signal Process. Multimed. Lab., "Univ. British Columbia, TMN 8 (H.263+) encoder/decoder," version 3.1.3, TMN 8 (H.263+) Codec, Feb. 1998.

[12] 이상희, 유국열, 김재균, "미분 추정 기법에 의한 2차원 이동 파라미터 추정의 수렴 특성 분석," *한국통신학회논문집*, 제23 권, 제7호, pp. 1869-1882, 1998년 7월

[13] Anil K. Jain, *Fundamentals of digital image processing*, Prentice Hall Inc., 1989

[14] Gary Sullivan, "Draft Meeting Report of the Eighth Meeting (Meeting H) of the ITU-T Q.15/16," *ITU-T Q.15/16*, Doc. # Q15-H37d1, Berlin, Germany, July 1999.