

H.264 동영상 표준 부호화 방식 성능 향상을 위한 잔여 신호 전처리 필터

*김도령, **송원선, ***홍민철

*숭실대학교 정보통신전자공학부

*kdoreyng@vipl.ssu.ac.kr, **won@vipl.ssu.ac.kr, ***mhong@e.ssu.ac.kr

Residual Pre-filter for Improving Performance of H.264 Video Coder

*Do-Ryung Kim, **Won-Seon Song, ***Min-Cheol Hong

*School of electronic engineering, college of engineering, Soongsil university

요약

본 논문은 H.264 동영상 표준 부호화 방식의 성능 향상을 위한 잔여 신호 전처리 필터에 대해 제안한다. 영상의 화질에 대한 최종적인 판단은 인간의 시각에 의하므로 HVS(Human Visual System)을 영상 압축에 적용할 수 있다[1,2,3]. 잔여 신호 전처리 필터는 원 영상으로부터 중요하지 않은 부분을 제거하여 영상의 화질을 주어진 비트율로 최대화 시켜 부호화율을 향상 시키는데 목적을 둔다[4]. 부호화 방식의 형태에 따라 잔여신호(DFD: Displaced Frame Difference)에 잔여 신호 전처리 필터 알고리즘을 적용하여 노이즈를 제거하고 전송 비트율을 감소시킬 수 있다. 제안된 방식의 성능을 실험 결과로부터 확인할 수 있다.

1. 서론

일반적인 디지털 동영상 정보는 정보량이 방대한 이유로 부호화 방식에 수많은 연구가 진행되어 왔으며, 특히 1988년에 디지털 정보의 부호화 및 저장에 대한 표준 규격의 필요성이 대두되면서 ITU에서는 유무선 통신망 환경에서 동영상 서비스를 위한 표준 규격 제정을 위해 노력해 왔다. 이동망과 같은 새로운 통신채널의 급속한 보급에 따라 기존 압축 방법에 비해 압축율이 더욱 향상된 동영상 표준 부호화 방식의 필요성이 대두되었다. H.264 동영상 표준 부호화 압축 방식은 기존의 압축율 보다 2배 이상의 압축율을 제공하므로 동일 전송 채널 폭을 사용할 경우 개선된 화질을 제공한다.

H.264 동영상 표준 부호화 방식에서는 4x4 블록 크기의 정수 DCT변환 방식과 함께, 부호화 시 많은 계산량을 요구하는 가변 블록 크기의 움직임 벡터 및 예측, 다양한 방향성을 갖는 인트라 부호화 방식 및 다중 참고 프레임 움직임 추정 방식 등을 사용한다.

다른 동영상 코딩의 표준화와 마찬가지로, 양자화는 다른 입력 영상을 나타내는 정보를 줄이기 위한 주요 과정으로 사용되고 있다. 그러나 비트율이 낮은 경우에는, 현재의 표준화처럼 양자화 과정에 의해 발생하는 주된 정보의 손실로 인해 복원된 영상에서 시각적인 격감이 높아진다[5]. 양자화 스텝 사이즈가 커지면, 더 높은 압축율을 나타낼 수가 있지만 그만큼 정보 손실의 양도 커지게 되므로 더

시각적으로 불편한 압축 현상들이 나타나게 된다.

제안된 잔여 신호 전처리 필터는 이러한 저 계산량의 노이즈 제거 및 압축효율을 개선하기 위해서 구현된 필터이다. 본 논문에서는 H.264 동영상 표준 부호화 방식에서 양자화로 인해 생긴 노이즈와 잔여 신호 전처리 필터 및 실험을 서술하고 결론을 맺는다.

2. H.264 동영상 표준 부호화 방식의 양자화 노이즈

H.264 이전의 동영상 표준 부호화 방식에서는 변환과 양자화 과정이 분리되어 수행되었다[6]. 그러나 H.264 동영상 표준 부호화 방식은 4x4 정수 DCT 변환을 위해서 양자화가 함께 수행된다. 따라서 H.264 동영상 표준 부호화 방식에서 양자화 노이즈를 구하기 위해서는 결합된 변환과 양자화 과정을 분리해야 한다.

$$X \rightarrow Y = AXA^T \rightarrow Z = (Y * QP) / (2^{15 + qp/6}) \quad (1)$$

식(1)은 4x4 행렬이 변환과 양자화 되는 식이다. X 는 입력 4x4 행렬이고 Y 는 변환된 4x4 행렬, Z 는 변환과 양자화가 수행된 4x4 행렬이다. $*$ 은 element product, $qp/6$ 은 양자화 스텝 사이즈를 6으로 나눈 몫이다. 식 (1)은 $Z = (AXA^T * QP) / (2^{15 + qp/6})$ 으로 표현된다. H.264 동영상 표준 부호화 방식의 변환 과정을 살펴보면 다음과

같다.

$$Y = AXA^T$$

$$= \begin{bmatrix} a & a & a & a \\ b & c & -c & b \\ a & -a & -a & a \\ c & -b & b & c \end{bmatrix} X \begin{bmatrix} a & b & a & c \\ a & c & -a & b \\ a & -c & -a & b \\ a & -b & a & -c \end{bmatrix} \quad (2)$$

여기서 $a = 1/2$, $b = \sqrt{1/2} \cos(\pi/8)$, $c = 1/2 \cos(3\pi/8)$

$$= \begin{bmatrix} a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & a & 0 \\ 0 & 0 & 0 & b \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & d & -d & 1 \\ 1 & -1 & -1 & 1 \\ d & -1 & 1 & d \end{bmatrix} X \begin{bmatrix} 1 & 1 & 1 & d \\ 1 & d & -1 & -1 \\ 1 & -d & -1 & 1 \\ 1 & 1 & 1 & d \end{bmatrix} \begin{bmatrix} a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & a & 0 \\ 0 & 0 & 0 & b \end{bmatrix}$$

여기서 $d = c/b$ (3)

$$= \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & d & -d & 1 \\ 1 & -1 & -1 & 1 \\ d & -1 & 1 & -d \end{bmatrix} X \begin{bmatrix} 1 & 1 & 1 & d \\ 1 & d & -1 & -1 \\ 1 & -d & -1 & 1 \\ 1 & -1 & 1 & -d \end{bmatrix} * \begin{bmatrix} a^2 & ab & a^2 & ab \\ ab & b^2 & ab & b^2 \\ a^2 & ab & a^2 & ab \\ ab & b^2 & ab & b^2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & 2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix} X \begin{bmatrix} 1 & 2 & 1 & 1 \\ 1 & 1 & -1 & -2 \\ 1 & -1 & -1 & 2 \\ 1 & -2 & 1 & -2 \end{bmatrix} * \begin{bmatrix} a^2 & ab/2 & a^2 & ab/2 \\ ab/2 & b^2/4 & ab/2 & b^2/4 \\ a^2 & ab/2 & a^2 & ab/2 \\ ab/2 & b^2/4 & ab/2 & b^2/4 \end{bmatrix}$$

$$Y = (CX C^T) * S$$

$$= [(CX C^T) * S * QP] / (2^{15+qp/6})$$

$$= [W * S * QP] / (2^{15+qp/6})$$

$$= [Y * QP] / (2^{15+qp/6}) \quad (4)$$

H.264 동영상 표준 부호화 방식에서 부호화 예서는 4x4 정수 DCT 변환을 위해서 식 (4)와 같이 양자화 과정에서 $Y = (CX C^T) * S$ 의 S 가 양자화 부분과 결합하여 새로운 양자화 행렬이 된다.

이것을 식으로 표현하면,

$$QE = S * QP$$

$$QP = QE * S^{-1} \quad (5)$$

식 (5)에서 S^{-1} 는 다음과 같다.

$$S^{-1} = \begin{bmatrix} 4 & 2\sqrt{10} & 4 & 2\sqrt{10} \\ 2\sqrt{10} & 10 & 2\sqrt{10} & 10 \\ 4 & 2\sqrt{10} & 4 & 2\sqrt{10} \\ 2\sqrt{10} & 10 & 2\sqrt{10} & 10 \end{bmatrix}$$

식 (5)에서 QE 는 H.264 동영상 표준 부호화 방식의 양자화 스텝 사이즈에 따른 4x4 양자화 행렬, S 는 식 (4)에서 구한 정수 변환을 위해 양자화 과정에 합쳐진 스케일링 행렬, QP 는 원 양자화 테이블 값이다.

H.264 동영상 표준 부호화 방식에서 양자화 식은 다음과 같다.

$$Z = \frac{W * QE}{2^{15+qp/6}} \quad (6)$$

W 는 H.264 동영상 표준 부호화 방식에서 CXC^T 된 4x4 행렬, Z 는 W 를 양자화 한 행렬이다. 식 (5)를 식 (6)에 대입하면,

$$Z = \frac{W * QE}{2^{15+qp/6}} = \frac{W * S * QP}{2^{15+qp/6}} \quad (7)$$

변환된 행렬은 $W * S$ 이므로 식 (7)은

$$Z = \frac{Y * QP}{2^{15+qp/6}}$$

$$= \frac{Y}{(2^{15+qp/6})/QP} \quad (8)$$

식 (8)에서 양자화 노이즈는 다음과 같다.

$$Error = (2^{15+qp/6})/QP \quad (9)$$

H.264 동영상 표준 부호화 방식에서 양자화 스텝 사이즈는 0-51 이므로 식 (9)에 대입하면 표 1과 같은 양자화 스텝 사이즈에 따른 최대 양자화 노이즈를 얻을 수 있다 [7]. $qp\%6$ 은 양자화 스텝 사이즈를 6으로 나눈 나머지이고, $qp/6$ 은 양자화 스텝 사이즈를 6으로 나눈 몫이다.

qp%6 \ qp/6	0	1	2	3	4	5
0	0.64	0.7	0.81	0.9	1	1.14
1	1.28	1.14	1.63	1.8	2	2.28
2	2.56	2.82	3.26	3.6	4	4.56
3	5.12	5.64	6.52	7.2	8	9.12
4	10.24	11.28	13.04	14.4	16	18.24
5	20.48	22.56	26.08	28.8	32	36.48
6	40.96	45.12	52.16	57.6	64	72.96
7	81.92	90.24	104.32	115.2	128	145.92
8	163.84	180.48	208.64	230.4		

표 1. 양자화 스텝 사이즈에 따른 최대 양자화 노이즈

3. 잔여 신호 전처리 필터

그림 1은 본 논문에서 제안된 잔여 신호 전처리 필터의 위치를 보여준다.

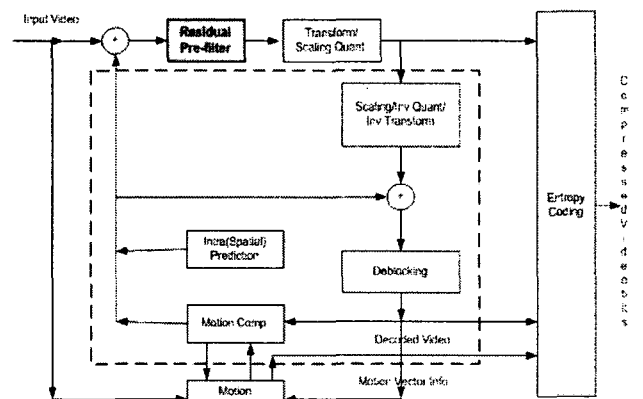


그림 1. 인코더에서 잔여 신호 전처리 필터 위치도

잔여 신호 전처리 필터는 DFD(Displaced Frame Difference)를 필터링한 후 변환과 양자화 과정을 수행한다.

동영상 표준 부호화 방식은 부호화 형태에 따라 intra coded image 와 inter coded image 로 구분되어 진다. intra 영상은 같은 사이즈의 블록으로 분할되어 각 블록이 DCT를 거쳐 양자화를 수행하고 inter 영상은 같은 사이즈의 블록으로 분할되어 움직임 추정(motion estimation) 을 거쳐 DFD를 DCT하고 양자화 한다. intra 영상과 inter 영상은 부호화 과정이 유사하지만 움직임 추정 및 보상을 하는 블록에서는 잔여 신호 값이 블록의 민감도를 결정하게 된다. 더불어, 양자화 과정은 동영상 부호화 기법의 압축 효율을 결정하며 양자화 스텝 사이즈가 커질수록 비트율은 감소되고 압축율은 높아지지만 시각적인 오류가 발생한다. 양자화 과정에 의한 손실은 다음과 같이 구분될 수 있다.

1) 블록화 현상 (Blocking Arifact)

: bit-rate가 낮으면 고주파 정보가 손실되고 블록의 DC값 전송으로 인해 발생하게 된다.

2)링 현상 (ring Artifact)

: 양자화 과정에 의한 고주파 성분의 손실로 인한 에지 영역의 진동 현상을 일컫는다.

본 논문에서 제안된 잔여 신호 전처리 필터는 intra 블록일 때는 4x4 입력 영상의 intensity에 필터링을 수행하고 inter 블록일 때는 4x4 DFD 에 필터링을 수행하게 된다. 그림 2에서는 잔여신호 전처리 필터의 흐름도 이다.

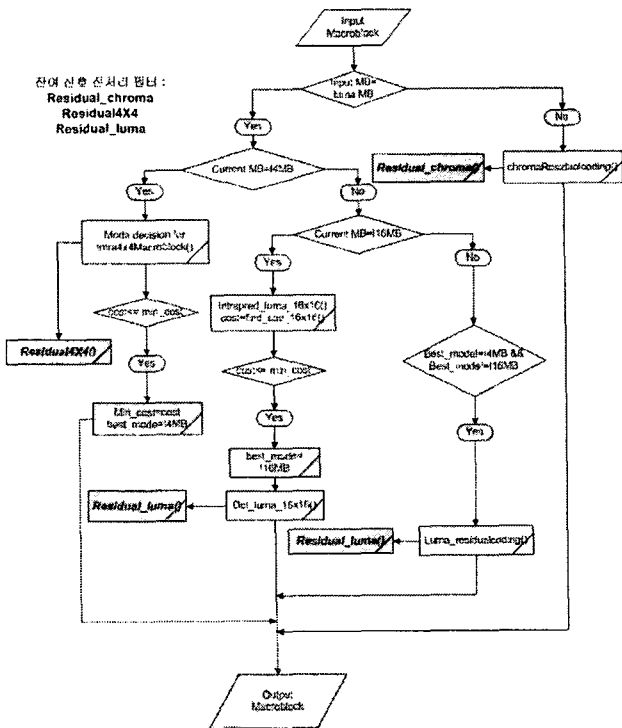


그림 2. 잔여 신호 전처리 필터의 흐름도

획득된 영상은 기본적으로 영상 획득 장치에 따라 정도의 차이는 있으나 침부 노이즈가 존재하며 양자화 과정에 의한 화질 저하 현상은 더욱 심각하게 된다. 본 논문에서 수행한 필터는 영상의 통계적인 분포가 가우시안 분포를 갖는다는 가정 하에 필터의 특성이 다음과 같은 가우시안 모델을 사용하였다.

$$H(i, j) = \frac{1}{Z} \exp \left\{ -\frac{(i^2 + j^2)}{\frac{\sigma_B^2}{\sigma_N^2}} \right\} \quad (10)$$

식(10)에서 Z는 정규화 상수, σ_B^2 는 국부지역 분산, σ_N^2 은 양자화 노이즈에 대한 분산, σ_Z^2 는 smoothing 파라미터이다. 필터 구현을 위해 식 (10)을 가우시안 분리 특성에 따라 1차원 모델로 표현하여 수평방향 필터와 수직 방향 필터를 분리하여 필터링 하였다.

$$H(m) = \frac{1}{\sqrt{Z}} \exp \left\{ \frac{k}{\frac{\sigma_B^2}{\sigma_Z^2}} m^2 \right\} \quad (11)$$

식 (11)에서 k 는 $\frac{1}{\sigma_Z^2}$ 를 의미한다. σ_B 값이 커질수록 필터 영상의 노이즈 제거는 효과적인 반면, 영상의 열화 정도는 더욱 심각하게 된다. k는 보정 계수에 해당되는 변수로 2장에서 구한 양자화 노이즈 특성을 반영한다.

4. 실험 결과

본 논문에서는 수행된 계산량 절감을 위한 H.264의 전처리 필터와 주관적, 객관적 성능을 비교하기 위해, 여러 동영상의 다양한 비트율 및 프레임 울에서 실험하였으며, JM6.1e 환경에서 실험하였다. 제안된 방식의 성능을 측정하기 위해, PSNR이 사용되었으며, 이는 화소 당 8비트로 구성된 $M \times N$ 영상 크기에 대해 다음과 같이 정의된다.

$$PSNR = 10 \log \frac{MN \times 255^2}{\|f - \hat{f}\|^2} \quad (12)$$

식(8)의 $\|\cdot\|$ 은 Euclidean norm을 의미하며, f 는 및 \hat{f} 은 각각 원 영상과 복원 영상을 의미한다. 본 논문에서 제안된 방식의 성능 비교를 위해 다양한 QP에 따른 PSNR 및 비트율을 사용하였다.

표 2,3은 QCIF Foreman, Container 영상을 잔여 신호 전처리 필터를 사용하였을 때, 부호화 및 복호화 한 결과이다. 그림 3,4 는 QCIF Foreman 2번째 영상을 QP=20 으로 잔여 신호 전처리 필터를 사용하지 않은 영상, 잔여 신호 전처리 필터를 사용한 결과 영상을 나타낸다. 표 5은 동일 QP에 대한 QCIF Foreman 영상에 대한 비트율 성능 비교를 25dB 노이즈를 첨부한 영상에 대해 나타낸 그래프 이다.

아래 결과로부터 제안된 방식은 동일 QP에 대해 비트율 절감이 약 10% 정도 나타났으며, 노이즈가 심할수록 그리고 QP가 낮고 움직임이 적은 영상일수록 효과적임을 알 수 있었다.

		QP	Without		With	
			Error Residual		Error Residual	
			Filter		Filter	
			SNR Y	Bitrates	SNR Y	Bitrates
Foreman QCIF 10 fps	without Noise	16	45.04	329.85	44.05	314.21
		20	41.70	200.71	40.86	190.90
		24	83.59	122.35	38.01	118.58
		28	35.88	74.11	35.32	71.28
		32	33.07	44.85	32.67	43.03
	25 dB Noise	16	39.44	830.28	39.57	776.26
		20	39.08	465.96	39.22	383.14
		24	38.17	176.90	37.87	158.46
		28	35.89	83.44	35.40	78.92
		32	33.10	46.41	32.69	44.65
	30dB Noise	16	42.83	532.17	42.62	453.96
		20	41.19	251.21	40.55	225.43
		24	38.58	133.38	38.01	127.10
		28	35.86	76.33	35.36	73.48
		32	33.13	45.30	32.64	43.14

표 2. Foreman 영상의 QP에 따른 잔여 신호 전처리 필터 성능 비교

		QP	Without		With	
			Error Residual		Error Residual	
			Filter		Filter	
			SNR Y	Bitrates	SNR Y	Bitrates
Container QCIF 10 fps	without Noise	16	44.98	201.03	43.37	167.73
		20	41.58	105.52	40.36	84.09
		24	38.56	54.45	37.92	46.53
		28	35.96	27.09	35.46	24.55
		32	33.14	14.38	32.79	13.60
	25 dB Noise	16	40.53	696.76	40.68	620.15
		20	40.04	338.77	40.08	215.77
		24	38.72	88.93	38.22	63.21
		28	35.25	31.99	35.73	27.60
		32	33.38	15.32	33.00	14.31
	30dB Noise	16	43.66	389.47	43.05	282.55
		20	41.54	143.52	40.41	102.42
		24	38.80	61.59	38.10	50.28
		28	36.13	27.99	35.65	25.39
		32	33.28	14.63	32.94	13.76

표 3. Container 영상의 QP에 따른 잔여 신호 전처리 필터 성능 비교



그림 3. QCIF Foreman 부호화 2 번째 영상(25dB 노이즈, QP = 20)



그림 4. QCIF Foreman 부호화 2 번째 영상(25dB 노이즈, 잔여 신호 전처리 필터, QP = 20)

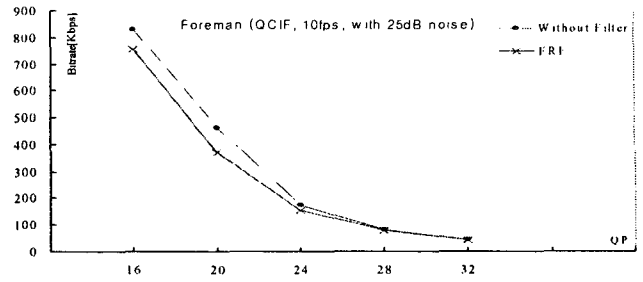


표 4. QP에 따른 비트율 성능 비교 (25dB 노이즈, QCIF Foreman, 10fps)

5. 결 론

본 논문은 양자화 노이즈를 제거하여 주어진 비트율로 최대화 시키는 잔여 신호 전처리 필터에 관하여 서술하였다. 블록의 특성에 따라 잔여 신호 전처리 필터를 적용함으로써 동일 QP에 대해 비트율 절감을 할 수 있었으며, 노이즈가 심할 수록 QP가 낮고 움직임이 적은 영상 전회기와 같은 경우 더욱 효과적인 결과를 얻을 수 있었다.

참 고 문 헌

- [1] A.J. Ahumada and H.A. Peterson, "Luminance-Model-Based DCT Quantization for Color Image Compression", *Human Visual Processing and Digital Display, Proceedings SPIE*, vol. 1913, pp191-201, 1993.
- [2] Andrew B. Watson, Joshua A. Solomon, Albert J. Ahumada, Jr, "Visibility of DCT basis functions : effects of display resolution" *Proceedings SPIE*, 1994
- [3] Joshua A. Solomon, Andrew B. Watson, Albert Ahumada, "Visibility of DCT basis functions : effects of viewing distance and contrast masking", In B.E Rogowitz (Ed), *Human Vision, Visual Processing and Digital Display IV, pp.99-108, Proceedings of SPIE*, 1994.
- [4] C. Andrew Segall, Passant Karunaratne and Aggelos, K. Katsaggelos, "Preprocessing of compressed digital video", *Processings SPIE*, 2001.
- [5] ITU-T SG16/Q6 JVT-G050r1 Draft ITU-T Recommendation on Final Draft International Standard of Joint Video Specification, May, 2003.
- [6] A. Hallapuro, M. Karczewicz and H. Malvar, "Low complexity Transform and Quantization - Part I", *JVT document JVT-B038, Geneva*, February, 2002.