

저 전송률 동영상에서 양자화 잡음의 평균값 보상을 사용한 부호화기의 성능 개선

정회원 신정환*, 백성학**, 김재호***

Performance Improvement Using Mean Compensation of Quantization Noise in Low Bit-rate Video Encoder

Jeong-Hwan Shin*, Seong-Hak Baek**, Jae-Ho Kim*** *Regular Members*

요약

손실압축 기법을 사용하는 대부분의 표준 동영상 압축 방법에서는 이미지 변환 후, 변환된 계수들의 양자화를 수행하게 된다. 양자화 과정에서 발생하는 양자화 잡음으로 인한 영상정보의 손실은 복구 영상의 화질을 저하시킨다. 본 논문에서는 H.26L의 테스트 모델인 TML-2를 기반으로 양자화 잡음의 평균값을 추가 보상함으로써 부호화기의 성능을 개선하는 방법을 제안한다. 실험 결과 비트율에서 평균 1.29% 개선과 PSNR에서 평균 0.93% 개선이 있었으며, 특히 움직임이 적은 영상에서 더 큰 개선 효과가 있었다.

ABSTRACT

In lossy compression method, the transformed coefficients are quantized. This results in the quantization noise. The video image quality and bit rate is closely related with the quantization step. In this paper, we proposed a new quantization function for the improved performance. The DC value of each macroblock is compensated depending on the magnitude of DC quantization error. It is implemented very low bit-rate video coding, i.e., H.26L. The experimental result is useful when the object motion is not severe.

1. 서론

최근에 저장기술과 전송기술이 비약적으로 발전하면서 멀티미디어에 대한 관심이 증가하고 있다. 멀티미디어는 여러 가지 형태의 정보를 하나로 묶어서 저장하거나 전송하는 것을 말한다. 멀티미디어의 요소 중 정보전달 측면에서 가장 효율적인 것은 영상 정보이다. 따라서 대부분의 경우 멀티미디어는 영상에 음향이나 문자 등의 정보가 결합된 형태를 취한다. 이러한 정보들을 저장하거나 전송하기 위해서는 많은 양의 데이터가 필요 하다. 따라서 세계 각 국에서는 이러한 정보들을 효율적으로 압축하기 위한 표준안을 제정하기 위해서 국제 표준

기구들을 통하여 공동의 노력을 기울여 왔다. 그 결과 정지영상의 경우에는 JBIG(Joint Bilevel Imaging Group), JPEG(Joint Photographic Experts Group)등이, 그리고 동영상의 경우에는 MPEG(Moving Picture Expert Group)-1, 2, 4, H.261, H.263 등이 표준안으로 제정되었다.^[1-7]

이러한 영상이나 음향 정보를 압축하는 기본 개념으로는 정보 내에 내재되어 있는 중복성(redundancy)을 제거하는 것이다. 영상의 경우는 시간적(temporal), 공간적(spatial), 통계적(statistical) 중복성이 존재한다. 우선 시간적 중복성이란 연속된 화면들 사이에 움직임이 존재하더라도 움직임이 있는 부분을 제외한 배경부분은 거의 같다는 것으로 음

* 부산대학교 전자공학과 ([jhwshin,baeksh,jhkim]@pusan.ac.kr)

논문번호 : K01033-0127, 접수일자 : 2001년 1월 27일

※ 본 연구는 정보통신부 논문연구과제 지원 및 부산대학교 ITS 센터 관리로 수행되었습니다.

직접 보상(Motion Compensation)을 통해서 제거된다. 다음으로 공간적 중복성이란 한 화면 내에서 인접한 화소(pixel)들 사이에 값이 비슷하다는 것으로 화소값의 공간적 예측을 통해서 제거된다. 마지막으로 통계적 중복성이란 양자화된 DCT(Discrete Cosine Transform)^[8] 계수들이 통계적으로 발생 확률이 높은 값과 낮은 값을 갖는다는 것으로 가변길이 부호화(VLC, Variable Length Coding)를 통해서 제거된다.

본 논문에서는 동영상 압축 표준안으로 현재 표준화가 진행 중인 H.26L의 테스트 모델인 TML-2 (Test Model Long Term Number 1)^[9]를 기반으로 양자화로 인한 잡음의 평균값을 보상함으로써 부호화의 성능을 개선시키는 방법을 제안한다.

본 논문의 구성은 다음과 같다. II장에서 실험에 사용된 TML-2에 대해서 H.263과 크게 달라진 점을 중심으로 간단히 설명하고, III장에서는 양자화 잡음과 제안하는 방법인 양자화 잡음의 보상방법에 대해서 설명한다. IV장에서는 모의실험을 통해서 제안한 방법의 성능을 평가해보고, V장에서 결론을 맺는다.

II. H.26L

1998년 ITU(International Telecommunication Union)-T에서 저 전송률 비디오 부호화 표준안인 H.263에 대하여 부호화 효율 개선뿐만 아니라 추가 기능들을 확장하여 H.263 권고안을 개정 발표하였다. 그러나 ITU-T에서는 부호화 성능뿐만 아니라 다른 표준들에서 제공하지 않는 기능들을 포함하는 새로운 부호화 알고리즘에 대한 잠재적 필요성이 여전히 남아있다고 여기고, 부호화 성능과 기능성 측면에서 기존의 비디오 표준안(H.261, H.262, H.263)과 당시에 연구 중에 있었던 표준안(MPEG-4)보다 더 뛰어난 차세대 비디오 부호화 표준안으로 H.26L에 대한 표준화 활동을 시작하였다. 당시 2002년 표준화 제정을 목표로 삼았던 H.26L은 실시간 시스템과 비 실시간 시스템 모두에 사용될 수 있으며, 대표적인 응용 분야를 몇 가지 살펴보면 다음과 같다.

- 실시간 회의 시스템 (비디오 폰).
- 이동 망에서의 시청각 통신.
- 인터넷 비디오 응용 분야.
- 비디오 저장과 전송 서비스 (주문형 비디오, 비

디오 메일).

1. H.26L TML-2

1998년에 시작된 표준화 활동의 결과로 1999년 11월에 TML-2가 발표되었다. 기본적인 부호화 알고리즘은 움직임 보상, 이미지변환, VLC 등을 사용하는 H.263 과 비슷하지만 TML-2에서 크게 달라진 부분들을 살펴보면 다음과 같다.

- 부호화를 위해서 하나의 VLC를 사용 : 모든 구문 요소(syntax element)들의 부호화에 동일한 VLC를 사용한다.

- 움직임 추정에 1/2, 1/3 그리고 1/6 화소 위치를 이용 : H.263에서는 움직임 벡터의 정밀도가 1/2 화소단위까지 사용되었으나, TML-2에서는 1/2 화소 단위 외에 추가로 1/3, 1/6 화소 단위까지 사용된다.

- 움직임 추정(Motion Estimation)에 여러 개의 참조 프레임을 사용 : H.263에서는 가장 이전에 복구된 프레임만을 참조하여 움직임 추정을 수행하였지만 TML-2에서는 최대 이전 5프레임까지를 참조하여 움직임 추정을 수행한다.

- 4x4의 블록 크기와 정수변환(integer transform) 사용 : 움직임 보상을 하고 남은 블록에 대해서 H.263에서는 16x16 크기의 매크로 블록을 4개의 8x8 블록으로 나누어서 8x8 2D-DCT를 사용하였지만, TML-2에서는 16개의 4x4 블록으로 나누어서 4x4 2D-정수변환을 사용한다.

- 움직임 추정을 몇 가지 블록 크기에 대해서 수행 : H.263의 고급 예측 모드에서는 16x16 크기의 매크로 블록에 4개의 움직임 벡터가 사용될 수 있었다. TML-2에서는 이와 비슷하게 1개에서 최대 16개까지의 움직임 벡터가 사용될 수 있다.

III. 양자화 잡음의 평균값 보상

대부분의 영상 압축 표준안에서는 움직임 보상을 수행하고 남은 블록에 대해서 이미지 변환을 수행하고, 변환된 계수들에 대해서 양자화를 수행한다. 이렇게 양자화를 거치면서 양자화 잡음 이 발생한다. 본 장에서는 먼저 양자화와 양자화 잡음에 대해서 살펴보고, 다음으로 제안하는 방법인 양자화 잡음의 평균값 보상에 대해서 설명한다.

1. 양자화와 양자화 잡음

일반적인 양자화 과정은 아래의 수식으로 수행되

며, 여기서 C_q 는 양자된 계수 값, C 는 양자화 될 계수 값, 그리고 Q 는 양자화 값이다.

$$C^q = \frac{C}{Q} \quad (1)$$

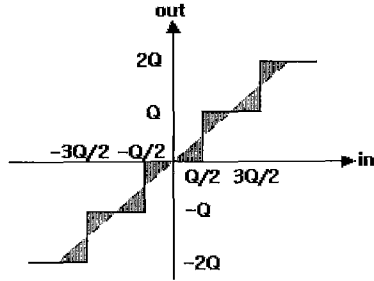


그림 1. 일반적인 양자화기의 특성

그림 1에는 일반적인 양자화기의 특징이 나타나 있다. 그림에서 가로축의 in값이 양자화기의 입력으로 위의 수식에서 C 에 해당하고, 세로축의 out값이 양자화기의 출력 값으로 C_q 에 해당한다.

TML-2에서 사용되는 양자화기는 그림 1의 일반적인 양자화기와 비슷한 특징을 가지고, 수식은 다음과 같다. 여기서 K 는 양자화 될 계수이고, LEVEL은 양자화 될 값이며, K 는 역 양자화 될 계수 값이다. 그리고 f 는 $(-0.5, 0.5)$ 범위의 K 와 부호가 같은 상수이며, $A(QP)$, $B(QP)$ 는 양자화 값 (QP , Quantization Parameter)에 따른 상수배열로 표1과 같은 값들을 갖는다.

$$\begin{aligned} \text{Quantizatin} & : \text{LEVEL} = (K \times A(QP) + f \times 2^{20}) / 2^{20} \\ \text{InveseQuantizatin} & : K' = \text{LEVEL} \times B(QP) \end{aligned} \quad (2)$$

위의 수식을 통해서 양자화 과정을 거친 계수들은 역 양자화 과정을 거치더라도 원래의 값으로 완전하게 복원되지 않는다. 이렇게 원래의 계수와 복원된 계수의 차이 값을 양자화 잡음이라고 하며, 그림 1에서 어둡게 보이는 삼각형 부분이 양자화 잡음이다.

그림 2에서는 실제 영상의 경우 양자화 잡음이 어떻게 나타나는지를 양자화 값 21에 대해서 보여 주고 있다. 그림 2에서 첫 번째 영상이 원 영상이며, 두 번째 영상은 양자화와 역 양자화를 거친 계수를 역 변환한 복원 영상이다. 그리고 마지막 영상은 원 영상과 복원 영상의 차이인 양자화 잡음 영상이다.

표 1. 양자화기의 상수값

A(QP=0, ... ,31)	620, 553, 429, 439, 391, 348, 310, 276, 246, 219, 195, 174, 155, 138, 123, 110, 98, 87, 78, 69, 62, 55, 49, 44, 39, 35, 31, 27, 24, 22, 19, 17
B(QP=0, ... ,31)	3881, 4351, 4890, 5481, 6154, 6914, 7761, 8718, 9781, 10987, 12339, 13828, 15523, 17435, 19561, 21873, 24552, 27656, 30847, 34870, 38807, 43747, 49103, 54683, 61694, 68745, 77615, 89113, 100253, 109366, 126635, 141533



원영상 복원된 영상 양자화 잡음
그림 2. 복원된 영상과 양자화 잡음

양자화 잡음은 양자화 값이 커질수록 커지게 된다. 그림 3에서는 양자화 값 21, 25, 29에 대해서 양자화 잡음의 평균값이 어떠한 분포를 가지는지를 보여 준다. 그림에서 볼 수 있듯이 양자화 값이 커질수록 양자화 잡음의 평균값이 커 진다는 것을 알 수 있다.

2. 양자화 잡음의 평균값 보상

영상 부호화기의 성능은 PSNR(Peak Signal-to-Noise Ratio)과 비트율로 측정할 수 있다. 그 중에서 화질을 나타내는 PSNR은 아래의 수식으로 주어지며, 수식에서 M 은 비교하려는 영상의 화소 수를 나타내고, o_n 은 원 영상의 화소 값, r_n 은 복원된 영상의 화소 값을 나타낸다.

$$PSNR = 10 \log M \sum_{n=1}^M \frac{255^2}{(o_n - r_n)^2} \quad (3)$$

위의 수식에서 볼 수 있듯이 PSNR은 원 영상과 복원 영상의 차이, 즉 양자화 잡음의 영향을 받는다.

본 논문에서는 양자화 잡음을 줄이기 위해서 양자화 잡음의 평균값을 복원된 영상에 추가로 보상한다. 우선 보상의 효과를 보기 위해서 그림 2에서 복원 영상의 전 화소에 '1'을 보상 하였다. 그 결과 그림4에서 보는 바와 같이 PSNR 이 0.4dB이상씩 높아지는 것을 알 수 있다.

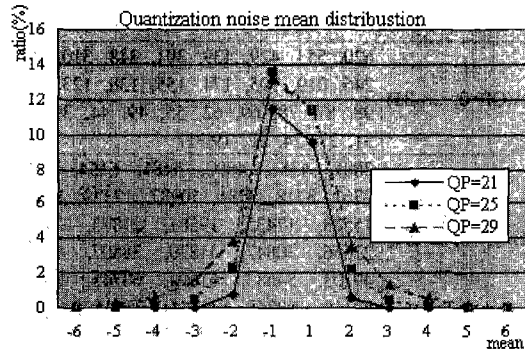


그림 3. 양자화 잡음의 평균값 분포

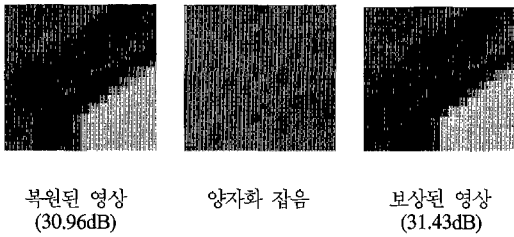


그림 4. 평균값 보상의 효과

실제 TML-2에서의 보상 과정을 살펴보면, 우선 보상 여부를 결정하기 위해서 원 영상과 복구 영상의 차이 값의 평균값을 구하고, 구해진 평균값의 절대값이 문턱 값보다 크게 되면 보상을 하게 된다. 문턱 값은 양자화 값이 25 미만일 때는 0.5이고, 양자화 값이 25 이상일 때는 1.5로 설정하였다. 그리고 보상 값은 양자화 값이 25미만일 때는 '+1', '-1'이고, 양자화 값이 25이상일 때는 '+2', '-2'가 된다. 양자화 값이 커질수록 양자화 잡음의 평균값이 커지기 때문에 보다 큰 보상 값을 가진다. 위에서 설명한 과정을 순서도로 나타내면 그림 5와 같다. 그림 5에서 $mean$ 이 양자화 잡음의 평균값이고 $org(.)$ 가 원 영상내의 화소 값, $recon(.)$ 이 복구 영상내의 화소 값이다.

송신 측에서는 원 영상과 복구 영상을 모두 가지고 있기 때문에 보상 여부를 결정할 수 있지만, 수신 측에서는 원 영상을 가지고 있지 않기 때문에 보상 여부를 알기 위해서는 송신 측에서 정보를 보내 주어야 한다. 본 논문에서는 보상 정보를 표시해 주기 위해서 보상 여부를 나타내는 비트인 C_y 와 보상 부호를 나타내는 비트인 C_v 를 추가로 사용하였다. 그림 6에 추가적인 TML-2의 구문 요소도와 새로운 비트를 추가시킨 구문 요소도가 나타나 있다. 그림에서 보는 바와 같이 C_v 비트는 C_y 비트에 따라서 존재할 수도 있고, 존재하지 않을 수도

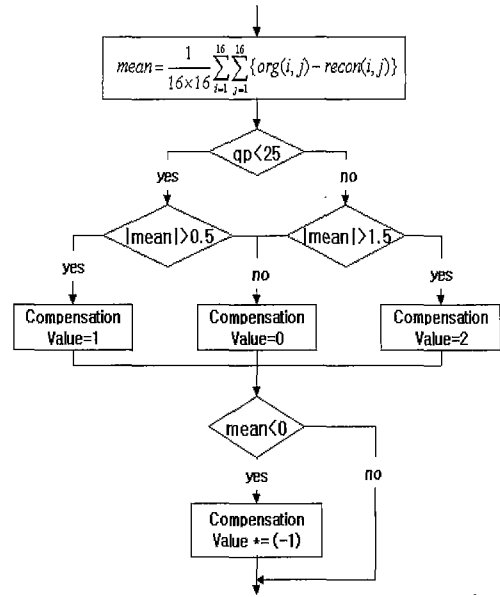


그림 5. 제안하는 방법의 순서도

있다. 즉 C_y 가 '1'일 때만 보상을 수행하기 때문에 보상 값의 부호를 나타내는 C_v 비트가 필요하고, C_y 가 '0'일 때는 보상을 수행하지 않기 때문에 보상 값의 부호를 표시할 필요가 없기 때문이다.

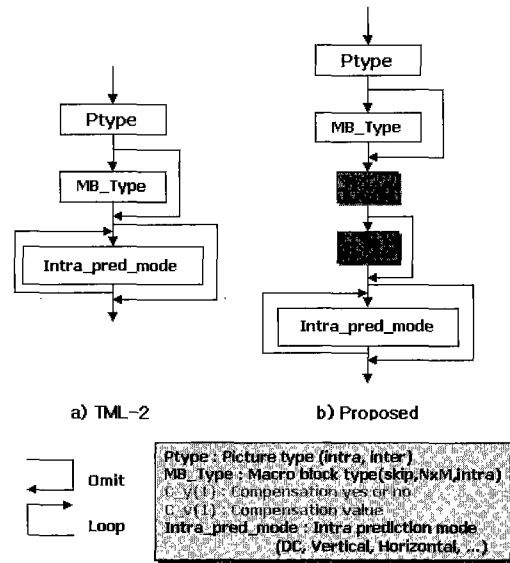


그림 6. 구문 요소도

제안하는 방법을 양자화 측면에서 살펴보면, 일반

적인 양자화기와 비교해 보았을 때 양자화기는 그림 7과 같은 특성을 가진다. 보상 여부에 따라서 기존의 레벨에 아주 작은 추가적인 레벨이 존재하게 된다.

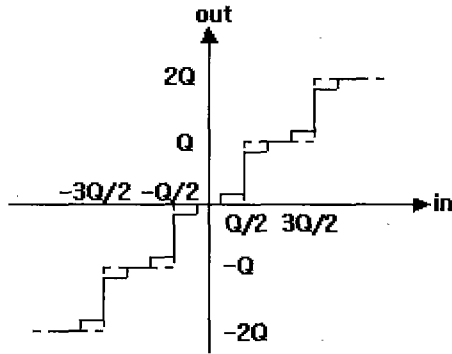
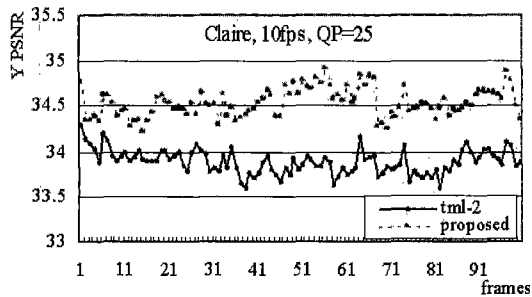


그림 7. 제안하는 방법의 양자화기의 특성

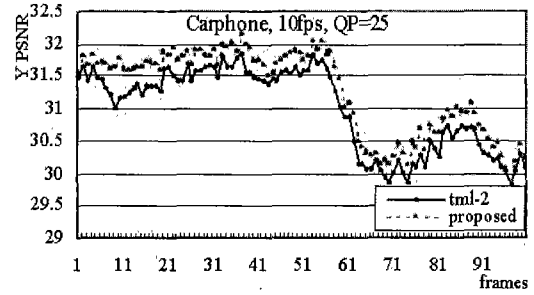
IV. 실험 및 결과 고찰

IV본 논문에서는 해상도가 176x144인 QCIF형식의 carphone, claire, container, foreman, grandmother, hallmonitor의 6가지 영상에 대하여 제안한 방법의 성능을 실험하였다. 양자화 값은 18, 21, 24, 25, 27, 29의 6가지를 선택하였고, 사용된 총 프레임 수는 300 프레임이며 프레임 율은 2프레임씩 건너뛰게 하여 10fps(frame per second)로 통일하였다.

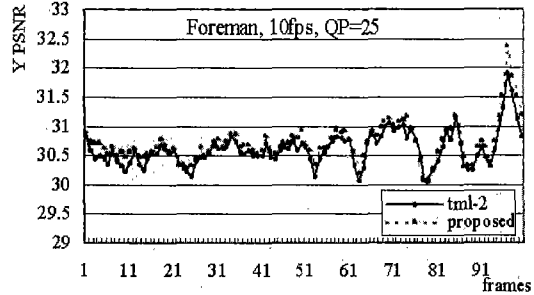
우선 그림 8에서는 claire, carphone, foreman영상에 대해서 양자화 값을 25로 선택하였을 때, 프레임에 따른 PSNR을 보여주고 있다. 그림에서 보듯이 세 개의 영상 중 움직임이 가장 적은 claire영상에 대해서 PSNR 개선이 크게 나타났으며, 움직임이 가장 크고 영상 내에 강한 에지가 많이 존재하는 foreman영상에 대해서 PSNR 개선이 작게 나타났다.



(a) claire



(b) carphone



(c) foreman

그림 8. 프레임별 PSNR 비교

그리고 표 2와 그림 9에서는 선택한 양자화 값에 따른 비트율과 PSNR을 비교하였다. 표 2의 result 항목은 '+'일 경우 그 숫자만큼의 성능개선을, '-'일 경우 그 숫자만큼의 성능 저하를 나타낸다. 앞의 그림 8에서 살펴본 바와 마찬가지로 claire와 container영상에 대해서 비트율과 PSNR 개선이 모두 크게 나타났으며, foreman영상에 대해서 비트율과 PSNR개선이 가장 작게 나타났다. 그리고 전체적으로 평균 1.29%의 비트율 개선과 평균 0.93%의 PSNR개선이 있었다. 제안하는 방법이 추가보상을 통해서 PSNR의 개선을 시키고자 하는 방법이지만 동영상의 경우 움직임 보상을 수행하기 때문에 참조 프레임의 PSNR이 좋아지면 움직임 보상을 더 정확하게 수행할 수 있고, 따라서 보상을 하고 남은 오류가 적어지기 때문에 비트율에서의 개선이 오히려 더 크게 나타났다.

V. 결론

대부분의 동영상 부호화기의 경우 양자화 과정을 수행하기 때문에 그 과정에서 양자화 잡음이 발생한다. 그리고 양자화 잡음으로 인해서 복구 영상의 PSNR이 저하된다. 양자화 잡음을 줄이기 위해서는

표 2. 제안하는 방법과 TML-2의 성능 비교

QP		Sequence	Carphone	Claire	Container	Foreman	Grand mother	Hall monitor
18	Bitrate (kbit/s)	TML-2	71.59	19.55	28.78	81.63	21.26	29.84
		Proposed	71.28	19.45	27.32	81.52	21.43	29.95
		Result(%)	+0.43	+0.51	+5.07	+0.13	-0.8	-0.37
	PSNR (dB)	TML-2	35.97	38.73	35.10	35.22	35.87	36.19
		Proposed	36.05	39.04	35.24	35.31	35.99	36.38
		Result(%)	+0.22	+0.8	+0.4	+0.26	+0.33	+0.53
24	Bitrate (kbit/s)	TML-2	31.46	9.11	12.34	37.96	7.97	13.7
		Proposed	31.32	8.94	11.39	38.74	8.16	13.72
		Result(%)	+0.45	+1.87	+7.7	-2.05	-2.38	-0.15
	PSNR (dB)	TML-2	31.63	34.24	30.74	31.12	31.98	31.6
		Proposed	31.83	35.15	31.06	31.32	32.29	31.74
		Result(%)	+0.63	+2.66	+1.04	+0.64	+0.97	+0.44
25	Bitrate (kbit/s)	TML-2	27.95	7.99	11.07	34.28	7.04	12.24
		Proposed	27.79	7.99	10.21	34.78	7.11	12.25
		Result(%)	+0.68		+7.77	-1.46	-0.99	-0.08
	PSNR (dB)	TML-2	31	33.89	30.21	30.62	31.63	30.93
		Proposed	31.27	34.54	30.54	30.72	31.78	31.03
		Result(%)	+0.87	+1.92	+1.09	+0.33	0.47	+0.32
29	Bitrate (kbit/s)	TML-2	16.97	5.38	6.89	22.55	4.45	7.66
		Proposed	16.98	5.22	6.5	22.88	4.18	7.73
		Result(%)	-0.06	+2.97	+5.66	-1.46	+6.06	-0.91
	PSNR (dB)	TML-2	28.27	30.19	27.28	28.13	29.24	27.9
		Proposed	28.56	31.21	27.75	28.29	29.66	28.09
		Result(%)	+1.03	+3.38	+1.72	+0.57	+1.43	+0.68

양자화를 작은 값으로 수행해야 하는데 그렇게 되면 그만큼 많은 수의 비트가 필요로 하게 된다.

본 논문에서는 추가적으로 많은 수의 비트를 사용하지 않고 양자화 잡음의 평균값을 보상함으로써 양자화 잡음을 줄이는 방법을 제안하였다. 보상 여부와 보상 값은 양자화 잡음의 평균값과 양자화 값

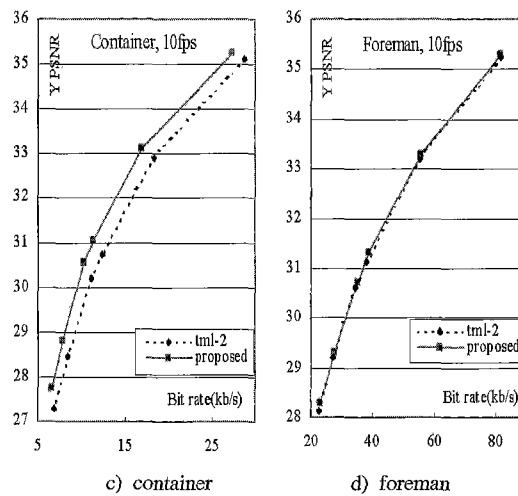
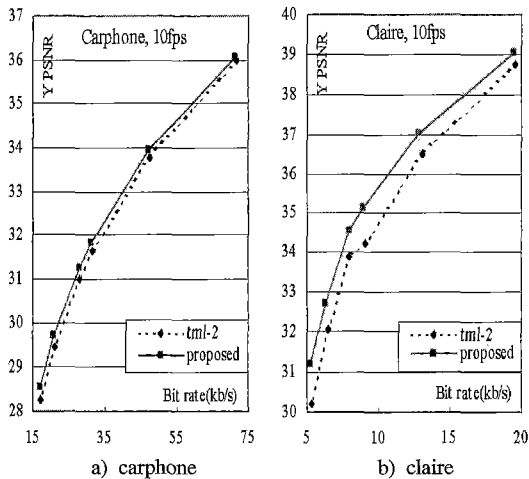


그림 9. 제안하는 방법과 TML-2의 비교

에 따라서 결정하고, 보상 정보를 수신 측으로 전송 해주기 위해서 새로운 비트를 할당하였다.

제안하는 방법을 H.26L TML-2를 기반으로 여러 가지 영상에 대해서 실험해 본 결과 상대적으로

움직임이 적어서 움직임 예측이 잘되는 영상에 대해서 비트율과 PSNR에서 큰 개선이 있었고, 움직임이 많아서 움직임 예측을 한 후에도 많은 오류가 남아 있는 영상에 대해서는 적은 개선이 있었다. 전체적으로는 평균 1.29%의 비트율 개선과, 평균 0.93%의 PSNR 개선이 있었다.

참고 문헌

[1] ISO/IEC International Standard 11544, “*JBIG, Progressive Bi-level Image Compression*”, ITU Recommendation T.82, 1993.

[2] William B. Pennebaker and Joan L. Mitchell, “*JPEG Still Image Data Compression Standard*”, Van Nostrand Reinhold, New York, 1993.

[3] ISO-IEC/IS 11172, “*Coding of Moving Pictures and Associated Audio For storage at up to about 1.5 Mbit/s*”, 5 Nov. 1992.

[4] ISO/IEC 13818-2, “*MPEG-2 video*”, ITU-T Recommendation H.262, Jan. 1995.

[5] ISO/IEC JTC1/SC29/WG11 N2459, “*MPEG-4 Overview*”, Oct. 1998.

[6] ITU-T(CCITT), “*Video Codec for Audiovisual Services at p*64 kbits*”, Recommendation H.261, Mar. 1993.

[7] ITU-T(CCITT), “*Video Coding for Low Bitrate Communication*”, Draft Recommendation H.263, Dec. 1995.

[8] K.R.Rao and P.Yip, *Discrete Cosine Transform*, New York, Academic, 1990.

[9] ITU-T, “*H.26L Test Model Long Term Number 1(TML-2) draft 0*”, Nov. 1999.

신정환(Jeong Hwan Shin) 정회원
 1999년 2월: 부산대학교 전자 공학과 (공학사)
 2001년 2월: 부산대학교 전자 공학과 (공학석사)
 <주관심 분야> 비디오 코딩

백성학(Seong Hak Baek) 정회원
 1992년 2월: 동아대학교 전자 공학과 (공학사)
 1994년 2월: 아주대학교 전자 공학과 (공학석사)
 2000년 3월~현재: 부산대학교 대학원 전자공학과
 박사과정 재학중
 <주관심 분야> 비디오 코딩

김재호(Jae Ho Kim) 정회원
 한국통신학회논문지 제20권 제4호 참조