

논문 2010-47SP-4-10

JPEG2000에서 저 전송 부호화를 위한 비주얼 양자화기 설계

(Design of Visual Quantizer for very low Bit-rate Coding on JPEG2000)

김 동 혁*, 전 준 현**

(Donghyeok Kim and Joonhyeon Jeon)

요 약

대역분할 부호화방식의 표준인 비가역 9/7 JPEG2000의 경우 0.15bpp 이하에서 양자화 손실로 배경과 에지 부분에서 급격한 화질 왜곡이 발생하는 문제점을 갖고 있다. 본 논문에서는 이러한 문제를 해결하기 위해 고주파대역들의 L -pdf(Laplace probability density function) 통계특성을 이용한 VQ(Visual Quantizer)를 제안하였다. 제안된 VQ는 재생영상의 화질왜곡을 최소화하기 위한 비주얼파라메타와 효율적인 비트율 감축을 위한 가중치 파라메타를 각각 사용하여 설계되었다. 실험결과 제안된 VQ는 0.15bpp이하에서도 재생영상의 시각적 손실을 최소화하면서 재생영상의 화질왜곡 기준척도인 MSE(Mean-Squared Error)도 최소화 할 수 있었다.

Abstract

The irreversible 9/7 JPEG2000, which is one of sub-band coding techniques, has a problem of severe picture quality distortion at the edge and the background caused by the quantization error below 0.15bpp. In this paper, to solve such problems we propose a VQ(Visual Quantizer) based on L-pdf(Laplace probability density function) statistical characteristics of high frequency sub-bands. The proposed VQ is designed by visual parameter for improving the subjective quality and weighting parameter for increasing the compression ratio. A proposed method, based on 9/7 JPEG2000 scheme, gives the high subjective quality to reconstructed images below 0.15bpp and provides minimum MSE(Mean-Squared Error) regardless of the compression ratio.

Keywords : JPEG2000; DWT; SBC; Quantizer; Visual Quantizer

I. 서 론

오늘날 멀티미디어 기술의 발전은 고화질 영상(high resolution image)의 디지털화와 더불어 방대한 양의 디지털 영상 정보를 효율적으로 저장·전송하기 위한 고화질 정지영상압축을 필요로 하게 되었으며, 최근 국제표준 및 이에 관한 연구가 지속적으로 이루어지고 있다.

1992년 ISO/IEC JTC1/SC29/WG1 국제표준그룹에서는 최초의 정지영상 압축 방식인 JPEG(Joint Photographic Expert Group)이 국제 표준으로 채택되어 다양한 멀티미디어 응용분야에 사용되고 있다^[1]. 그러나 JPEG은 구현이 쉽다는 장점에도 불구하고 0.15bpp(bit per pel)이하에서 양자화 손실로 인한 뚜렷한 화질 왜곡(distortion)을 보이고 있다. 최근 이러한 문제점들을 해결하기 위해서 DWT(Discrete Wavelet Transform)을 사용하는 대역분할 부호화 방식인 JPEG2000이 국제표준으로 채택되었다^[2~3]. JPEG2000의 특징은 손실/무손실 압축을 동시에 지원이 가능하고, 손실 압축의 경우 ROI(Region of Interest) 마스크(mask)를 통하여 배경영역과 관심영역을 구분하여 부호화하므로써 0.15bpp

* 학생회원, ** 평생회원-교신저자, 동국대학교 정보통신공학전공

(Department of Information and Communication Engineering, Dongguk University)

※ 본 과제(결과물)는 교육과학기술부·지식경제부의 출연금으로 수행한 산학협력중심대학육성사업의 연구결과입니다.

접수일자: 2010년4월2일, 수정완료일: 2010년4월27일

이하에서 배경의 화질은 떨어지나 영상의 가장 중요한 부분의 에지부분에 대한 화질 왜곡을 최소화 한다는 점이다^[3]. 그럼에도 불구하고 높은 압축률을 갖는 9/7 JPEG2000(손실압축)의 경우 양자화기가 양자화 손실에 따른 시각적(visual) 왜곡을 고려하지 않고 단지 최소 MSE(Mean-Squared Error)을 갖도록 설계되어 있기 때문에 0.15bpp 이하에서 PSNR (Peak Signal-to Noise Ratio) 감소량에 비해 배경과 에지 부분에서 급격한 화질 왜곡이 발생하는 문제점을 갖고 있다.

JPEG2000과 같은 대역분할 부호화 방식에서 대역분할시 저주파대역(low frequency subband)을 제외한 고주파대역(high frequency subband)에 속한 부 대역 영상(subband image) 신호들은 L -pdf(Laplace probability density function)의 통계특성을 갖는다. 따라서 이러한 고주파대역들의 통계특성을 바탕으로 본 논문에서는 기존의 9/7 JPEG2000에서 고압축에 따른 양자화 손실로 인한 급격한 화질 저하의 문제점을 해결하고자 VQ(Visual Quantizer)를 제안하였다. 제안된 VQ의 각 대역 양자화기는 압축효율을 높이기 위하여 시각적무손실 임계값(visual-lossless threshold value)에 의한 데드존(dead zone)을 갖고 있으며, 양자화 손실에 따른 재생영상에서의 화질열화 개선을 위하여 양자화 스텝사이즈는 L -pdf 통계특성으로부터 추출된 비주얼 파라미터(visual parameter)를 사용하여 설계되었다. 여기서 시각적 무손실이란 인간의 시각적 특성을 이용한 것으로 데드존에서 임계필터링(thresholding)으로 발생하는 원영상과 재생영상 간의 MSE(Mean Squared Error)는 존재하나 시각적으로 화질의 열화가 보이지 않는 것을 의미한다^[4]. 제안된 VQ를 9/7 JPEG2000 (ref.4의 JASPER, version 1.9)에 적용하여 실험한 결과 0.15bpp 이하에서 VQ를 적용한 방식이 기존의 스칼라 양자화기를 사용한 방식보다 PSNR이 0.5dB이상 향상된 것을 확인하였으며, 재생영상의 주관적 화질 평가에서는 매우 우수한 것으로 나타났다.

본 논문의 II장에서는 기존의 9/7 JPEG2000과 고주파대역들의 통계특성 및 양자화 손실에 대한 영향등에 대하여 분석을 하였으며, III장에서는 II장의 분석을 바탕으로 L -pdf 기반의 VQ를 제안하였다. IV장에서는 시뮬레이션을 통하여 제안된 VQ를 9/7 JPEG2000 표준에 적용하여 성능을 평가하고 기존 표준인 JASPER v1.9^[5]과도 성능을 비교 검토하였다. 끝으로 V장에서는 결론을 맺었다.

II. JPEG2000 및 고주파대역 신호 분석

2.1 9/7 JPEG2000 시스템^[2-3, 6-7]의 양자화 문제점
대역분할 부호화 방식을 사용하는 비가역 9/7 JPEG2000(손실 압축)은 그림 1에 나타난 바와 같으며, 인코딩 과정을 살펴보면 비가역 9/7 DWT를 통해 대역분할된 각 부 대역 영상신호들은 양자화를 한 후 엔트로피 코딩 과정을 거친 다음, 최종적으로 JPEG2000 code-stream을 생성 한다^[1]. 2차원 DWT는 1차원 수평 및 수직 필터링에 의해 이루어진다.^[8-11]

그림 2는 9/7 JPEG2000(JASPER, version 1.9)과 본 논문에서 사용한 13채널 대역분할을 하기위한 계층적(hierarchical) DWT 트리 구조(4 계층 구조)를 나타낸 것으로 저주파 대역만을 반복적으로 대역분할 한 것을 알 수가 있다.

여기서 $S_i, i=0,1,2, \dots, 12$ 는 i 번째 부 대역 영상을 나타낸 것으로 0번째 계층에는 S_0 , 1번째 계층에는 $S_i, i=1,2,3$, 2번째 계층에는 $S_i, i=4,5,6$, 3번째 계층에는

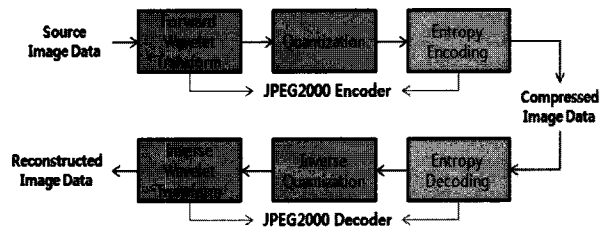


그림 1. 손실 JPEG2000의 인코딩과 디코딩 과정
Fig. 1. Process of Lossy JPEG2000 for Encoding and Decoding.

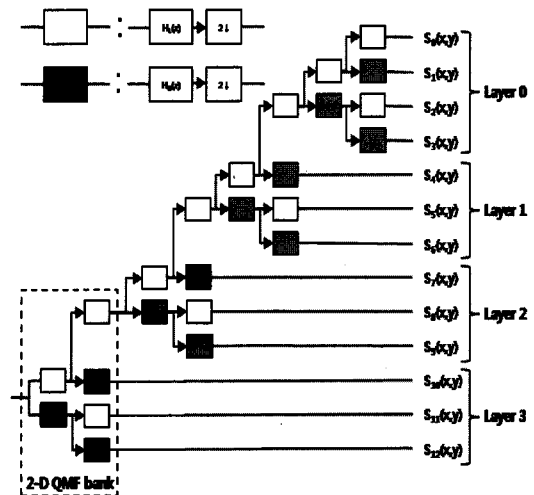


그림 2 13채널 대역 분할을 위한 QMF बैं크 트리
Fig. 2. QMF bank tree for 13 channel band division.

$S_i, i=7,8,9$ 그리고 4번째 계층에는 $S_i, i=10,11,12$ 부대역 영상들이 각각 존재한다.

9/7 JPEG2000을 사용하여 영상을 부호화할 때 대역 분할된 각 부대역 영상신호들은 양자화 과정을 거치며, 엔트로피 부호화 전에 각 양자화 값은 ROI 마스크 유도로 결정된 관심영역과 배경영역으로 분류되어 다운스케일링(down-scaling)과 업스케일링(up-scaling)을 거친 후 비트변환이 된다. 이와 같이 ROI 영역에 따른 스케일링은 낮은 비트율에서 배경의 화질은 떨어지나 영상의 가장 중요한 부분의 에지부분에 대한 화질 왜곡을 최소화하기 위함이다. 그럼에도 불구하고 ROI 기법 적용 전에 각 부대역 영상신호들은 주파수대역 특성의 구분 없이 단지 전체 평균 손실이 MMSE(Minimum Mean-Squared Error)을 갖도록 양자화 되기 때문에 만일 비트율이 낮을 경우 각 주파수대역의 에지와 배경 성분들은 양자화 손실로 인한 심각한 왜곡이 이미 발생한 상태로서 ROI 영역 구분에 따른 비트변환 준위 및 배경 성분의 스케일링 등은 재생영상의 화질의 개선에 크게 영향을 끼치지 못한다. 결론적으로 9/7 JPEG2000은 ROI 영역 구분에 따른 코딩으로 JPEG 보다 낮은 비트율에서도 좋은 성능을 갖을 수는 있으나 0.15bpp 이하의 고압축률에서는 양자화기의 한계로 인한 ROI 효과가 그다지 크지 않다는 것이다.

2.2 부대역 영상신호들의 주파수 특성 분석

DWT를 사용하여 영상을 대역 분할하였을 때, 각 주파수 대역의 부대역 영상신호들은 일반 영상과는 다른 통계특성을 갖고 있기 때문에 대역별로 양자화 손실에 따른 시각적 왜곡이 재생 영상에서 다르게 나타난다. 그림 3은 512x512 크기의 8비트 해상도를 갖는 그림 4의 4가지 실험 영상 “LENNA”, “COUPLE”, “TANK”, “ZELDA” 들을 1024x1024 영상으로 사용하여 그림 2와 같이 4 계층의 13 채널로 대역분할 하였을 때 저주파대역의 영상 S_0 을 제외한 고주파대역 영상들 $S_i, i=1,2,\dots,12$ 의 통계 특성들을 나타낸 것으로 모든 고주파대역에서 98~99%의 부대역 화소 값이 ± 20 범위 내에 분포되어 있으며, 이중 85~90%가 ± 7 범위 내에 집중되어 있음을 알 수가 있다. 이와 같은 데이터 통계 특성은 각 고주파 대역별 주파수 성분의 중요도에 대한 구분 없이 나타나고 있어 낮은 비트율에서의 양자화로 인한 화질 손상은 심각하게 나타난다. 예를 들어 그림 2의 1번째 계층에 속한 $S_i, i=1,2,3$ 은 입력 영상의 에

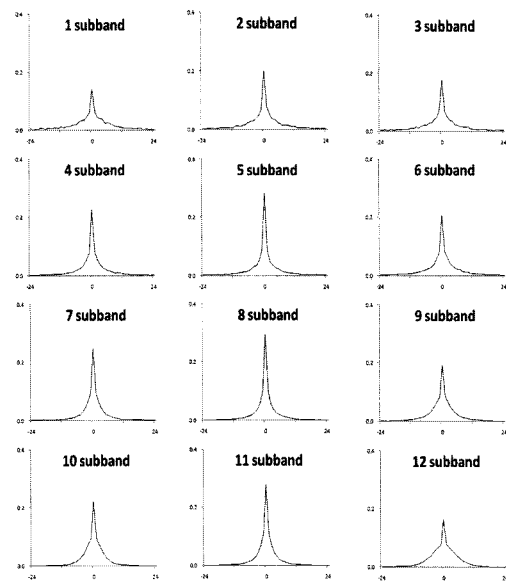


그림 3. 13채널 대역분할에 따른 고주파대역들의 통계 특성
(세로축 : 확률, 가로축 : 대역화소 값)
Fig. 3. Characteristics of high frequency statistics according to 13 channel band division.
(Horizontal axis : Probability, Vertical axis : Picture element value of subband)

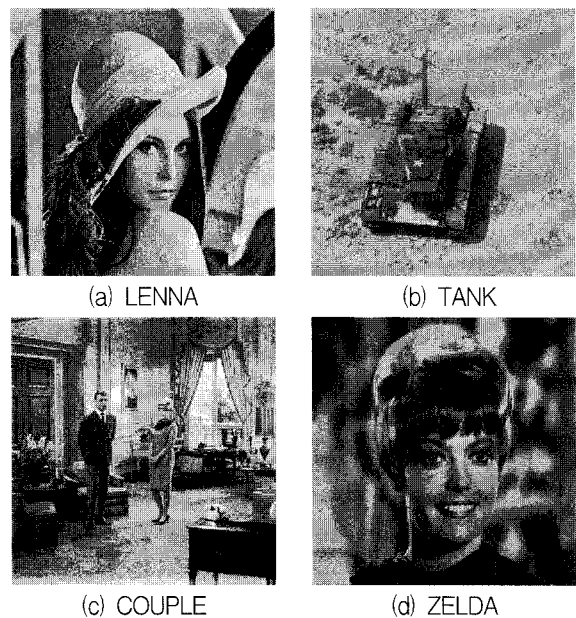


그림 4. 트레이닝 영상
Fig. 4. Training Image.

지에 대한 매우 중요한 고주파 성분을 가지고 있는 부대역 영상들로서 아주 적은 양자화 오차가 발생하더라도 다른 계층에 속한 고주파 대역에서 발생한 양자화 오차보다 재생영상 화질에 영향을 미치는 정도가 매우 크다. 또한 각 계층에서 대각 방향의 부대역 영상들

$S_i, i=3,6,9,12$ 에서 발생하는 양자화오차는 같은 계층에 속한 수평 및 수직 부 대역 영상들에서 발생하는 양자화오차에 비해 재생영상 화질에 미치는 영향이 상대적으로 적다. 따라서 각 고주파대역별로 발생하는 양자화오차가 재생영상에서의 주관적 화질에 미치는 영향이 다르기 때문에 각 대역별 양자화기 설계 시 해당 고주파대역의 정확한 통계 특성에 대한 고려가 필요하다.

그림 3의 각 고주파대역의 화소 값들에 대한 통계 분포를 살펴보면 라플라스-확률밀도함수(L -pdf)의 분포 특성과 매우 유사한 것을 알 수가 있다^[4, 12]. 그림 3에서 0번째 저대역을 제외한 i 번째 고주파대역을 $S_{i \neq 0}$ 라 하자. 만일 $S_{i \neq 0}$ 에 속한 신호들이 정상적 랜덤 프로세서(stationary random process)의 불규칙변수(random variable)로서 통계 함수가 좌우 대칭인 L -pdf 통계특성을 갖는다면 아래 식과 같이 표현될 수가 있다.

$$f_i(x) = \frac{1}{2b_i} e^{-\frac{|x-\mu_i|}{b_i}} \quad (1)$$

여기서 x 와 μ_i 는 $S_{i \neq 0}$ 의 부 대역 화소 값과 평균값이며, b_i 는 $S_{i \neq 0}$ 의 특성 파라미터로서 양수 값을 갖는다. 이때 그림 3에서 알 수가 있듯이 저주파대역을 제외한 모든 고주파대역 신호들의 평균값은 $\mu_i \cong 0$ 을 가지므로, 식(1)은 아래의

$$f_i(x) = \frac{1}{2b_i} e^{-\frac{|x|}{b_i}} \quad (2)$$

와 같이 표현된다. 이때 식(2)에서 b_i 를 $S_{i \neq 0}$ 의 표준편차 σ_i 값으로 하여 실제 영상의 고주파대역 특성에 근사화 된 L -pdf를 나타낼 수가 있다. 그럼에도 불구하고 원영상마다 통계 특성(고주파 성분)이 다르므로 σ_i 는 원영상의 표준편차에 의존하는 유동적인 값을 갖기 때문에 모든 영상에서 통용될 수 있는 $S_{i \neq 0}$ 의 확률밀도함수(pdf)를 식(2)의 L -pdf로 정확하게 표현하는 것은 불가능하다. 따라서 본 논문에서는 통계특성이 다른 위 4가지 실험영상의 σ_i 을 b_i 값으로 사용하여 근사화 된 L -pdf를 사용하였다.

2.3 대역별 양자화오차 영향 분석

본 절에서는 제 II장 2절에서 논의된 바 있는 대역별 발생하는 양자화 오차가 재생영상의 MSE에 미치는 영향을 분석하였다. 정확한 분석을 위하여 연산오차(반올

표 1. 대역별 오차에 따른 재생영상에서의 MSE
Table 1. MSE of reconstructed image according to each band error.

	$e_i = 2$	$e_i = 3$	$e_i = 4$	$\rho_i(\xi_i/\epsilon_i)$ *
ξ_0	4.00000	9.00000	16.00000	1.00000
ξ_1	0.86937	1.95609	3.47750	0.21734
ξ_2	0.86937	1.95609	3.47750	0.21734
ξ_3	0.18895	0.42514	0.75581	0.04724
ξ_4	0.82293	1.85159	3.29172	0.20573
ξ_5	0.82293	1.85159	3.29172	0.20573
ξ_6	0.16930	0.38093	0.67721	0.04233
ξ_7	0.75666	1.70249	3.02666	0.18917
ξ_8	0.75666	1.70249	3.02666	0.18917
ξ_9	0.14314	0.32205	0.57254	0.03578
ξ_{10}	1.00000	2.25000	4.00000	0.25000
ξ_{11}	1.00000	2.25000	4.00000	0.25000
ξ_{12}	0.25000	0.56250	1.00000	0.06250

* $\epsilon_i = e_i^2$

림이나 버림)가 발생하지 않는 5/3필터를 사용하였으며, 그림 3에서 나타낸 바와 같이 13채널로 대역을 분리하였다. 실험 방법으로는 앞의 실험영상을 DWT로 분리한 13채널 대역중 12채널을 제외한 1채널 대역에 속한 부 대역 화소에 균일한 오차(e_i)를 발생시킨 후 다시 13채널을 IDWT하여 해당 대역의 MSE가 재생영상의 MSE에 미치는 영향을 구하였다. 표 1은 대역화소 오차(error/subband pel)가 각각 2, 3, 4일 때의 재생영상에서의 MSE를 나타낸 것이다. 만일 ξ_i 를 i 번째 대역의 오차(e_i)로 인해 재생영상에서 발생한 MSE 라고 하였을 때, e_i 가 재생영상에서의 MSE에 끼치는 영향을 표 1의 실험 결과에 따라 나타내면 아래 식과 같다.

$$\xi_0 > \xi_{10} = \xi_{11} > \xi_7 = \xi_8 > \xi_4 = \xi_6 > \xi_1 = \xi_2 > \xi_{12} > \xi_9 > \xi_5 > \xi_3 \quad (3)$$

위의 결과에서 알 수가 있듯이 S_0 대역에서 발생하는 오차가 재생영상에서의 MSE 값에 가장 치명적으로 나타나며, 수직에지 성분을 갖고 있는 S_1, S_4, S_7, S_{10} 대역과 수평에지 성분을 포함하는 S_2, S_5, S_8, S_{11} 에 비해 대각선에지 성분을 갖고 있는 S_3, S_6, S_9, S_{12} 대역의 오차가 상대적으로 재생영상 MSE에 미치는 영향이 적다. 그럼에도 불구하고 압축률 향상을 위한 데이터 압축효율은 계층 3에 속한 S_{10}, S_{11}, S_{12} 이 가장 높고 계층이 낮을수록 압축효율은 급속히 떨어진다. 이것은 고주파

대역 일수록 부 대역 영상의 사이즈가 크고 “0”에 가까운 화소 값들이 많아 양자화를 하더라도 MSE가 적고 재생화질에 크게 영향을 미치지 않기 때문이다. 그 밖에 높은 계층일수록 압축효과가 크므로 재생영상에서의 MSE를 최소화하고 효율적인 전송 비트율 감축을 위해서는 계층3의 양자화기 설계가 무엇보다도 중요하다. 이와 같은 사실은 무엇보다도 앞서 II장 2절에서 설명한 각 대역의 통계특성에 기인한 것으로 판단된다.

위의 분석을 바탕으로 본 논문에서는 VQ 설계시 다음과 같은 사항들이 고려되었다.

- ① 고주파대역 신호들은 *L-pdf* 통계특성을 갖고 있으며, 높은 고주파대역일 수록 재생영상 MSE 영향은 크나 대부분이 “0”에 가까운 화소 값들이거나 재생영상 화질에 직접적인 영향이 적은 고주파 노이즈로서 전송 비트율 감축에 효과적이다.
- ② S_0 의 경우 압축효율(비트율)은 매우 저조하나 재생영상의 MSE에 미치는 영향이 매우 치명적이다
- ③ S_3, S_6, S_9, S_{12} 의 경우 같은 계층의 부 대역보다 재생영상의 MSE에 미치는 영향이 적다.
- ④ 각 계층의 수평 및 수직대역 오차는 재생영상의 MSE에 미치는 영향이 같다.
- ⑤ 높은 계층일수록 비트율 감축효율이 향상된다. 즉, 비트율 감축효율이 계층0 < 계층1 < 계층2 < 계층3 되도록 양자화기를 설계해야 한다. (대역 가중치 함수 사용)

III. VQ

인간의 시력은 적은 량의 데이터 손실로 발생하는 재생영상에서의 화질 왜곡을 감지하고 인식하는 것이 민감하지 않으며, 특히 높은 주파수 성분의 손실은 저주파 성분에 비해 매우 덜 민감하기 때문에 대부분의 압축 방식에서 이를 이용한다. 따라서 본 논문에서는 DWT에 따른 각 고주파대역 신호들의 *L-pdf* 통계특성을 바탕으로 이러한 인간의 시각적 효과(visual effect)를 갖는 양자화기인 VQ를 설계하였다. 제안된 VQ의 각 대역은 시각적무손실 임계필터링을 위한 데드존과 비주얼 파라메타 값이 적용된 양자화 스텝사이즈에 의해 양자화 된다. 여기서 시각적무손실이란 인간의 시각적 특성을 이용한 것으로 데드존에서 임계필터링(thresholding)으로 발생하는 원영상과 재생영상 간의 MSE는 존재하나 시각적으로 화질의 열화가 보이지 않는 것을 의미한다.^[4]

3.1 비주얼 임계값과 비주얼 파라메타

제안된 VQ의 각 대역 양자화기의 양자화 스텝사이

즈(step size)는 그림 5(a)와 같이 시각적 무손실을 갖는 임계값(threshold value) t_i 을 구한 후 이를 바탕으로 결정된 비주얼 파라메타를 이용하여 설계된다. 이 때 t_i 는 그림 5(b)에서 나타난 바와 같이 각 고주파 대역 양자화기의 데드존이 된다. 본 절에서는 VQ 설계를 위한 비주얼 임계값과 비주얼 파라메타의 최적화 과정에 대하여 서술하였다.

그림 4의 4가지 영상을 하나의 입력영상(1024×1024)으로 하여 그림 2와 같이 13채널로 대역분할 하였을 때, 저주파대역을 제외한 모든 각 고주파대역 신호들이 식 (2)에서 언급한 바와 같이 *L-pdf* 통계특성을 갖는다고 가정하고 이때 i 번째 고주파대역 $S_{i \neq 0}$ 의 특성 파라메타 b_i 가 $S_{i \neq 0}$ 의 표준편차 σ_i 값으로 표현된다면 식(2)은 아래와 같이 나타낼 수가 있다.

$$f_i(x) = \frac{1}{2\sigma_i} e^{-\frac{|x|}{\sigma_i}} \tag{4}$$

그림 3에서 나타낸 실제 고주파대역에서 부 대역 신호들은 평균값($\mu_i \approx 0$)을 중심으로 약 85%이상이 ± 7 범위의 값에 분포되어 있는 반면 이들 신호들이 재생화질에

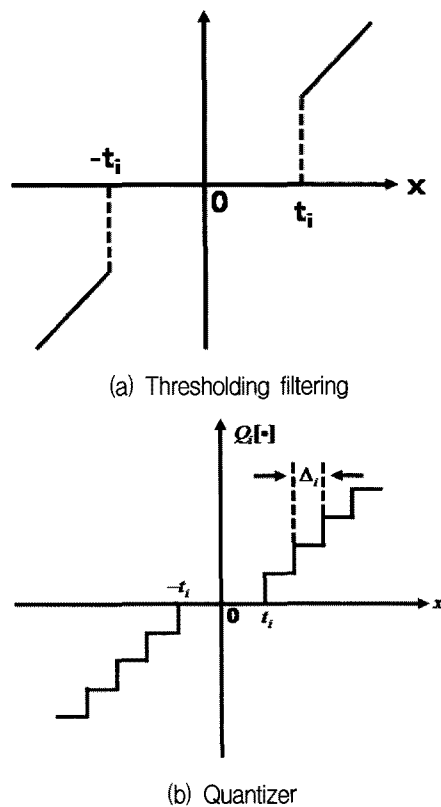


그림 5. 임계 필터링과 양자화기

Fig. 5. Thresholding filtering and Quantizer.

미치는 영향이 매우 미약하다. 따라서 이들 신호들에 대하여 시각적 효과를 고려한 그림 5(a)의 임계필터링을 적용할 경우 압축 효율을 높일 수가 있다. 이때 임계필터링을 위한 임계값 t_i 는 그림 5(b)의 i 번째 고주파 대역 양자화기의 데드존 임계값(dead-zone threshold value)이 된다^[11, 13]. 이때 $\pm t_i$ 데드존에서 발생하는 각 대역에서의 MSE를 $\epsilon_i(t_i)$ 라고 가정하면 다음과 같으며

$$\epsilon_i(t_i) = \int_{-t_i}^{t_i} x^2 f_i(x) dx \tag{5}$$

이때 식(4)을 식(5)에 대입하여 정리하면

$$\epsilon_i(t_i) = 2\sigma_i^2 - e^{-t_i/\sigma_i} (t_i^2 + 2t_i\sigma_i + 2\sigma_i^2) \tag{6}$$

이 된다. 결과 식에서 나타난 바와 같이 $\epsilon_i(t_i)$ 는 t_i 값이 증가할수록 급격히 증가함을 알 수가 있으며, 만일 $\epsilon_i(t_i)$ 가 주어질 경우 t_i 을 식(6)의 관계식으로부터 구할 수가 있다. 따라서 시각적 무손실을 갖는 t_i (한계치)는 재생영상에서 시각적 왜곡이 나타나기 시작하는 $\epsilon_i(t_i)$ 의 최소값이 주어진다면 얻을 수가 있다. 따라서 t_i 의 한계치를 구하기 위해서는 t_i 의 증가에 따른 $\epsilon_i(t_i)$ 의 면밀한 분석이 필요하다. 이를 위하여 각 대역의 비주얼 파라메타(visual parameter) k_i 을

$$k_i = t_i/\sigma_i, \quad 0 < k_i < 1 \tag{7}$$

와 같이 정의하여 식(6)에 적용하면

$$\epsilon_i(k_i) = \sigma_i^2 \{ 2 - e^{-k_i} (k_i^2 + 2k_i + 2) \} \tag{8}$$

와 같이 k_i 의 함수로 간단히 나타낼 수가 있다. 그림 6은 k_i 증가에 따른 $\epsilon_i(k_i)/\sigma_i^2$ 을 나타낸 것으로 다음과 같이

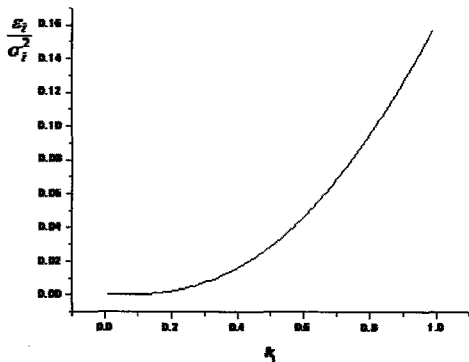


그림 6. k_i 에 따른 $\epsilon_i(k_i)/\sigma_i^2$
 Fig. 6. $\epsilon_i(k_i)/\sigma_i^2$ according to k_i .

$\epsilon_i(k_i)$ 의 범위를 구할 수가 있다.

$$0 < \epsilon_i(k_i) < 0.1606\sigma_i^2, \quad 0 < k_i < 1 \tag{9}$$

식(9)에서 $k_i=1$ 의 경우 $t_i=\sigma_i$ 을 의미하며, 이 경우 90% 이상이 데드존 양자화로 "0"이 된다는 것으로 이때 $\epsilon_i(\sigma_i) = 0.1606\sigma_i^2$ 로서 $S_{i \neq 0}$ 에 속한 모든 신호들은 약 $\bar{e}_i = 0.4\sigma_i$ 의 평균오차를 갖게 된다.

그림 6에서 무엇보다 중요한 사실을 알 수가 있는데, $\epsilon_i(k_i)$ 는 k_i 가 0.2보다 적을 경우 거의 증가가 없다가 k_i 가 0.4보다 커질 경우 $\epsilon_i(k_i)$ 는 급격히 증가함을 알 수가 있다. 따라서 시각적 무손실을 갖는 k_i 와 t_i 는 각각

$$0.3 < k_i < 0.5 \quad \text{and} \quad 0.3\sigma_i < t_i < 0.5\sigma_i \tag{10}$$

범위 내에서 존재함을 알 수가 있다. 식(10)의 범위는 13 채널 대역 분할된 각각의 영상들에 다양한 t_i 을 사용하여 그림 5와 같이 임계필터링을 적용한 재생영상 화질을 주관적 테스트를 하였을 때, 배경과 예지에서 왜곡이 나타나기 시작하는 범위이다. 또한 객관성을 위하여 그림 4의 트레이닝 영상들 외의 다른 영상들을 사용하여 k_i 의 범위를 추정하였을 때도 식(10)과 같은 결과를 얻었다. 이것은 대역분할을 할 경우 입력 영상의 종류에 따라 각 고주파 성분들은 약간의 차이는 거의 유사한 *L-pdf* 통계특성을 갖는 사실을 입증한 것이다. 이와 더불어 앞의 제 2.3 절에서 설명한 바와 같이 S_3, S_6, S_9, S_{12} 대역은 다른 고주파대역들에 비하여 상대적으로 재생영상의 MSE에 미치는 영향의 매우 적기 때문에 압축 효율을 더 높이기 위하여 $0.5 < k_i$ ($: 0.5\sigma_i < t_i$)의 범위를 가질 수가 있다. 그림에도 불구하고 제안된 식(8)에 의존하여 k_i 와 t_i 을 결정할 경우 위의 값이 각 대역의 표준편차 값에 의존하고 있다는 문제점이 존재한다. 그러나 다양한 영상들에 대한 각각 대역의 표준 편차들을 분석한 결과, 고주파 대역일수록 모든 영상들의 해당 σ_i 값들이 작고 동일하게 갖는다는 것을 알 수가 있었으며 σ_i 값에 대한 식(4)의 *L-pdf* 통계특성이 매우 유사하게 나타남을 확인하였다. 이것은 고주파 대역일수록 대부분이 원영상의 잡음(noise) 성격의 고주파수 성분들이기 때문이며, 실제 양자화시 재생 영상에 영향을 크게 끼치지 않는다.

3.2 VQ 설계와 구현

본 절에서는 제 III장 1절에서 분석한 결과를 바탕으로 VQ를 구현하기 위하여 각 고주파 대역의 양자화 스

템사이즈를 설계하고 구현에 대하여 서술하였다.

그림 5(b)는 i 번째 고주파대역의 양자화기를 나타낸 것으로 각 대역의 양자화 스텝사이즈를

$$\Delta_i = k_i w_i \sigma_i^2, \quad 1 \leq i \leq 12 \quad (11)$$

와 같이 정의 하였다. 여기서 w_i 는 효율적인 비트율 감축을 위한 대역 가중치(weighting factor)로서 계층별 부 대역 영상 크기에 따라 다음 식과 같다.

$$w_i = 1/2^{8-2n}, \quad n = 0, 1, 2, 3 \quad (12)$$

여기서 계층 n 에 속한 부 대역의 가중치 값들은 $\{w_{3n+1}, w_{3n+2}, w_{3n+3}\}$ 가 된다. 표 2는 그림 4의 4가지 실험영상을 이용하여 앞서 앞에서 제안된 (k_i, w_i) 관계식을 이용하여 설계된 Δ_i 을 나타낸 것이다. 그림 7은 최적 k_i 을 얻기 위한 흐름도를 나타낸 것으로 먼저 시각적 무손실을 갖는 재생영상에서의 MSE 최소값인 $\xi_i = 2$ (:실험 값)에 대해 표1의 $\rho_i = \xi_i / \epsilon_i$ 관계식으로부터 구한 ϵ_i 을 이용하여 식(8)을 만족하는 최적 k_i 을 구하였다. 또한 표2에서 $\widehat{\Delta}_i$ 는 아래 식에서 나타낸 바와 같이 Δ_i 의 최소값인 Δ_2 로 정규화 하였는데,

$$\widehat{\Delta}_i = \Delta_i / \Delta_2, \quad 1 \leq i \leq 12 \quad (13)$$

이것은 $\Delta_i, 1 \leq i \leq 5$ 값이 1보다 적은 값으로 나타나 실제 양자화 시, 양자화 된 값이 양자화 전의 값보다 커

표 2. VQ 스텝사이즈 파라미터
Table 2. Step size Parameter of VQ.

S_i	σ_i	k_i	w_i	Δ_i	$\widehat{\Delta}_i$
S_1	14.2084	0.2943	0.0039	0.2321	1.25821
S_2	11.8932	0.3339	0.0039	0.1845	1.00000
S_3	12.6030	0.5486	0.0039	0.3404	1.84521
S_4	11.7700	0.3467	0.0156	0.7506	4.06848
S_5	9.92193	0.4049	0.0156	0.6228	3.37612
S_6	11.1280	0.6413	0.0156	1.2408	6.72564
S_7	10.1176	0.3989	0.0625	2.5523	13.8341
S_8	8.43154	0.4749	0.0625	2.1104	11.4386
S_9	8.45010	0.9520	0.0625	4.2488	23.0292
S_{10}	6.98238	0.4332	0.2500	5.2805	28.6208
S_{11}	6.20815	0.4810	0.2500	4.6349	25.1215
S_{12}	6.09788	0.9925	0.2500	9.2268	50.0103

지는 것을 방지하기 위함이다.

만약 $S_i(x, y)$ 을 $i(\neq 0)$ 번째 대역에 속한 부 대역 화소 값을 $S_i(x, y)$ 라고 하였을 때, 만일 $|S_i(x, y)| \geq |t_i|$ 일 경우 식(13)을 이용한 양자화는 다음과 같다.

$$Q_i[S_i(x, y)] = \left\lfloor \frac{S_i(x, y)}{d \times \widehat{\Delta}_i} \right\rfloor, \quad 1 \leq i \leq 12 \quad (14)$$

여기서 $[a]$ 는 실수 a 의 반올림된 정수 값을 의미하며, d 는 양자화 스케일 값이다. 식(15)의 역양자화(inverse quantization)의 경우

$$Q_i^{-1}[\widehat{S}_i(x, y)] = [\widehat{S}_i(x, y) \times d \times \widehat{\Delta}_i], \quad 1 \leq i \leq 12 \quad (15)$$

여기서 $\widehat{S}_i(x, y)$ 는 식(14)에 의해 양자화 된 부 대역 화소 값으로 정수형이다.

IV. 실험 방법 및 결과

본 장에서는 III장에서 설계된 VQ를 “Daubechies 9/7 taps” 비가역 필터를 사용한 JPEG2000(:그림 1)에 적용하여 시뮬레이션을 한 결과를 나타냈으며, 또한 기존의 9/7 JPEG2000^[5](:JASPER v1.9)과 비교 분석하였다. 시뮬레이션 방법으로는 앞서 설명된 바와 같이 그림 2와 같이 4개 계층의 13채널 DWT을 적용하였으며, S_0 대역을 제외한 고주파 12개 대역에 대해서만 설계된 VQ를 적용한 후 엔트로피 부호화를 하였다. 이때 양자화와 역양자화는 표 2에서 구한 $\widehat{\Delta}_i$ 을 사용하여 식(14)과 식(15)에 따라 구현되었으며, 다양한 비트율(average bit per pel :

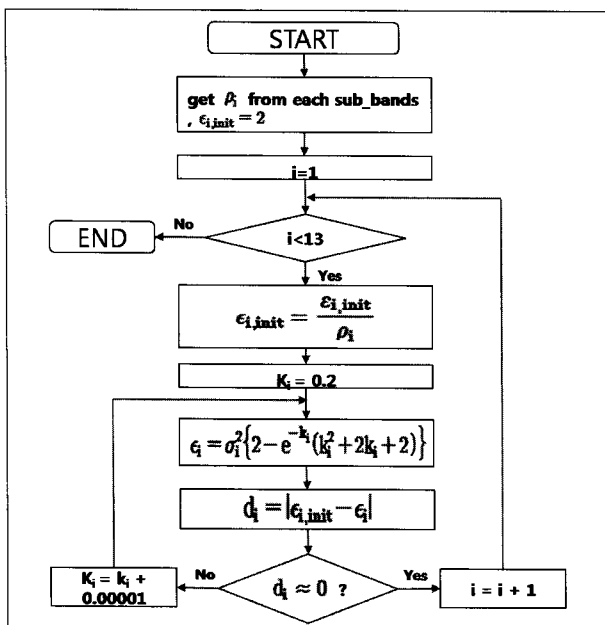


그림 7. 최적 k_i 을 구하기 위한 트레이닝 흐름도
Fig. 7. Training flowchart for finding the optimal k_i .

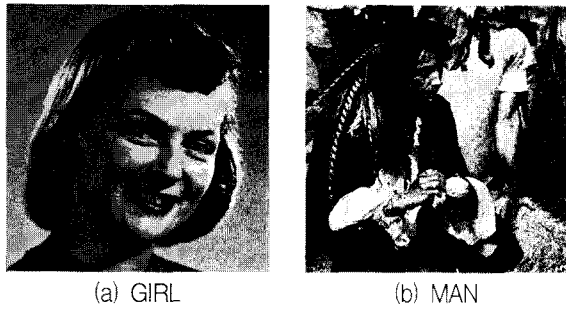


그림 8. 실험 영상
Fig. 8. Test Image.

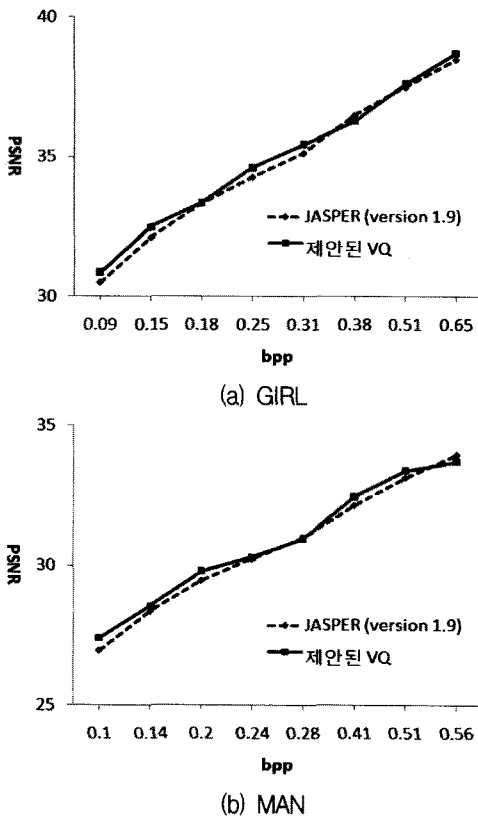


그림 9. 비트율에 따른 PSNR 성능 비교
Fig. 9. Comparison of PSNR performance according to bit rate.

bpp) 실험을 위하여 여러 가지 d 값을 사용하였다. 또한 제안된 VQ의 객관성을 위해 성능평가에 사용된 입력영상으로는 그림 8의 "GIRL"과 "MAN"을 사용하였는데, 이들 영상은 k_0 을 최적화하기 위해 사용된 그림 1의 트레이닝 영상으로 사용하지 않은 영상이다. 그리고 성능 비교를 위하여 PSNR과 주관적(subjective) 화질 평가를 하였다.

그림 9는 "GIRL"과 "MAN"에 대하여 VQ가 적용된 9/7 JPEG2000과 기존의 9/7 JPEG2000 표준의 양자화기를 사용한 JASPER v1.9와의 평균 비트율에 따른 PSNR



JASPER(Version 1.9)



제안 방식
(a) 0.15 bpp



JASPER(Version 1.9)

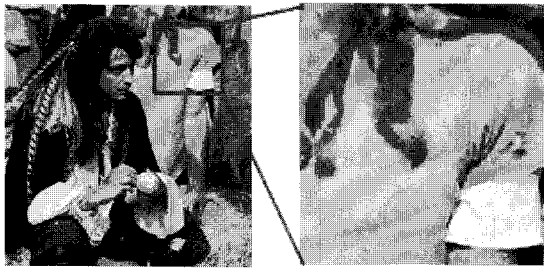


제안 방식
(b) 0.1 bpp

그림 10. 표준 JPEG2000의 압축 후 복원 영상과 고안한 통계적 양자화기를 통해 복원한 영상간의 화질 비교 (GIRL)

Fig. 10. Comparison between the reconstructed image compressed by JPEG2000 standard and the image reconstructed using the proposed statistical Quantizer (GIQL).

을 비교한 것이다. 그림 9에 따르면 0.15bpp 이하에서 제안된 VQ를 사용한 방식이 기존의 JASPER에 비하여 PSNR이 항상된 사실을 알 수가 있다. 또한 그림 10과 그림 11은 입력영상들에 대한 주관적 화질평가를 나타낸



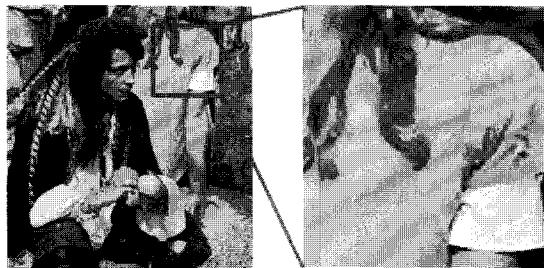
JASPER(Version 1.9)



제안 방식
(a) 0.15 bpp



JASPER(Version 1.9)



제안 방식
(b) 0.1 bpp

그림 11. 표준 JPEG2000의 압축 후 복원 영상과 고안한 통계적 양자화기를 통해 복원한 영상간의 화질 비교 (MAN)

Fig. 11. Comparison between the reconstructed image compressed by JPEG2000 standard and the image reconstructed using the proposed statistical Quantizer(MAN).

것으로 0.15bpp 이하에서 제안된 VQ를 사용한 방식이 기존의 JASPER에 비하여 화질 왜곡이 많이 개선된 것을 알 수가 있다. 결론적으로 본 논문에서 제안한 VQ는 0.15bpp 이상에서는 기존의 방식과 유사한 성능을 갖는

반면, 0.15bpp 이하의 낮은 비트율에서는 시각적 효과를 갖는 고화질의 재생영상을 얻을 수 있음이 증명되었다.

V. 결 론

본 논문에서는 영상의 시각적 고품질을 보장하면서 압축률을 향상 시킬 수 있는 VQ를 제안하였다. VQ는 대역분할 부호화 방식에서 고주파 대역의 L -pdf 통계특성에 맞게 화질 왜곡을 최소화할 수 있는 비주얼 파라메타와 효율적인 비트율 감축에 필요한 가중치 파라메타를 기반으로 설계되었다. 실험결과 제안된 VQ는 기존 JPEG2000표준에서 사용한 양자화기보다 0.15bpp이하에서도 PSNR 향상뿐 만아니라 주관적 화질 개선을 나타냈다.

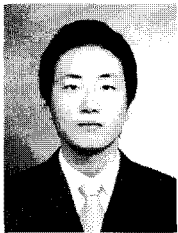
결론적으로 본 논문에서 제안한 VQ는 대역분할 부호화 방식에 적합한 양자화기로서 압축 성능이 매우 우수하다는 것을 알 수가 있다.

참 고 문 헌

- [1] W.B. Pennebaker and J.L. Mitchell, "JPEG: Still Image Data Compression Standard," Van Nostrand Reinhold, 1993.
- [2] T. Acharya, P. Tsai, "JPEG2000 standard for Image Compression (Concepts, algorithms and VLSI architectures)," John Wiley & Sons Inc, 2004.
- [3] ISO/IEC, "ISO/IEC 15444-1:2000: Information technology - JPEG2000 Image Coding System - Part 1 final draft international standard," 2000
- [4] 봉정식, 전준현, 양기주, "JPEG2000에서 시각적 무손실 임계값을 이용한 진단의료영상 압축기법," 한국통신학회논문지, Vol. 34, no. 7, 2009년 7월
- [5] Michae ID.Adams, Assistant Professor Dept. of Electrical and Computer Engineering University of Victoria, "JasPer Software Reference Manual (Version 1.900.0)," Dec 2006.
- [6] A. S. Lewis and G. Knowles, "Image compression using the 2-D wavelet transform," IEEE Trans. on Image Processing, Vol. 1, no. 2, pp.244~250, Apr 1992.
- [7] 문형진, 정갑천, 박성모, "JPEG2000 영상 압축을 위한 효율적인 비율-왜곡 최적화 알고리즘 구현," 전자공학회논문지, 제43권 CI편, 제3호, 2006년 5월
- [8] 전준현, "2차원 영상 필터링 효율 향상을 위한 기술연구," 전자공학회논문지, 제42권 SP편, 제6호, 2005년 11월

- [9] 전준현, "근접 완전재생 Pseudo-QMF बैं크를 위한 선형위상 프로토타입 저역통과 필터의 효율적인 설계 방법," 한국통신학회논문지, Vol.33, no.3, 2008년 3월
- [10] 이성목, 조성대, 강봉순, "JPEG2000 Hard-wired Encoder를 위한 칼라 2-D DWT processor의 구현," 신호처리·시스템학회논문지, 제4권, 제9호, 2008년 10월
- [11] S. H. Lee and D. H. Yoon, "Introduction to the Wavelet Transform," Jinhan Books. 2002.
- [12] Han Oh, Ali Bilgin, Michael W. Marcellin, "VISIBILITY THRESHOLDS FOR QUANTIZATION DISTORTION IN JPEG2000," Quality of Multimedia Experience, 2009. QoMEX 2009. International Workshop on, San Diego, U.S.A, July 2009.
- [13] Scoot Daly, "The visible differences predictor: an algorithm for the assessment of image fidelity," In Digital Images and Human Vision, Andrew. B. Watson, Ed. chapter 14, pp.179-206, Cambridge MIT Press, UK, May 1993.

 저 자 소 개



김 동 혁(학생회원)
 2008년 2월 동국대학교 정보통신
 공학과 학사졸업
 2010년 2월 동국대학교 정보통신
 공학과 공학석사
 2010년~현재 동국대학교 정보통
 신공학전공 박사과정

<주관심분야 : 영상통신, 영상압축>



전 준 현(평생회원)-교신저자
 1984년 동국대학교 전자공학과
 학사졸업.
 1986년 한국과학기술원 전기
 및 전자 공학석사 졸업.
 1991년 한국과학기술원 전기
 및 전자 공학박사 졸업.

1991년~1999년 한국통신 인터넷 연구실장 및
 사업부장

2000년~2001년 한누리투자증권 리서치팀 이사

2000년~2001년 드림라인 전무

2001년~현재 동국대학교 정보통신공학전공
 부교수

<주관심분야 : 영상통신, 영상압축, 에너지변환
 등>