

논문97-2-2-03

인간 시각 양자화기를 이용한 MPEG-4 정지영상 압축 방법의 성능 개선

김민구*, 김승종*, 정제창*

Improvement of the MPEG-4 Still Image Compression Using Visually Weighted Quantizers

Min-Gu Kim*, Seung-Jong Kim*, and Je-Chang Jeong*

요 약

본 논문에서는 최근 MPEG-4에서도 표준의 일부로 채택될 전망이다 웨이블릿 기반의 정지영상 압축 기법에 대해서 살펴보고 MPEG-4에서의 정지영상 압축 기법보다도 간단하면서, 인간 시각 특성을 고려한 양자화기를 사용하여 시각 특성을 향상시킨 효율적인 압축 기법을 제안한다. MPEG-4에서 제시된 알고리즘과 본 논문에서 제안한 알고리즘을 구현하여 성능 평가를 실시하였다. 평가 결과 제안한 방법이 MPEG-4에서의 정지영상 압축 방법보다도 훨씬 좋은 성능을 나타내었다. 또한 압축률이 매우 높을 경우를 제외하고는 Shapiro의 EZW 보다도 복잡도가 낮으면서 대부분의 경우 주관적, 객관적 화질이 우수함을 알 수 있었다. 웨이블릿 변환은 인간의 시각 특성을 잘 반영하고 있으므로 압축된 영상은 JPEG으로 압축할 때와 같이 블록간 왜곡이 거의 발생하지 않으며 전체적으로 JPEG보다 훨씬 우수한 화질의 향상을 가져왔다.

Abstract

In this paper, we investigate on the techniques for still image compression based on the wavelet transform, which will be adopted as a part of the MPEG-4 compression standards. Also we propose an effective still image compression technique, which is simpler than the MPEG-4 compression method and is improved by using a visually weighted quantizer based to HVS(Human Visual System). Simulations are carried out and compared with the algorithm proposed in MPEG-4. The simulation results show that the proposed method in this paper gives much better image quality than that of the method in MPEG-4. Also, except the case where the compression ratio is high, it shows that the proposed method has lower in complexity and provides a better subjective and objective image quality than EZW in most cases. Since wavelet transform well reflects HVS, the compressed image rarely causes blocking artifact compared with JPEG, and in most cases, it shows considerable quality improvement over JPEG.

* 한양대학교 전자통신공학과

Dept. of Electronic Communications Engineering, Hanyang University

※ 본 논문은 한국과학재단 주관의 특정기초연구과제 (96-0102-14-01-3) 연구비 지원하에 수행되었습니다.

I. 서 론

최근 영상, 비디오, 오디오 신호 처리 기술의 발달로 멀티미디어 통신 서비스가 국내외에서 상용화되고 있으며 멀티미디어 관련 기술의 상업적 이용이 커다란 관심을 끌고 있다. 통신 채널의 대역폭과 멀티미디어 응용분야에 따

라 다양한 질(Quality)의 정보를 제공하고자 하는 스케일러빌리티(Scalability)의 개념에도 불구하고, 멀티미디어 정보 중 영상 정보의 대용량은 시스템의 효율을 저해하는 요인으로 작용하고 있다. 대용량의 영상 정보를 고능률로 압축하면서 응용 분야에 따른 요구조건을 만족하기 위한 방법으로 JPEG, MPEG-1, MPEG-2, H.261과 H.263 등 영상부호화의 표준이 제정되었다. 또한, 차세대 영상부호화 방식으로 정지영상은 JPEG-2000을 중심으로 동영상은 MPEG-4를 중심으로 국제 표준의 제정이 진행되고 있으며 각국의 표준 및 기술 선점을 위한 연구가 활발히 진행되고 있다.

1980년대 후반, 신호 해석의 방법으로써 웨이블릿 변환(Wavelet Transform)이 소개된 이후 영상부호화에 있어서 웨이블릿 변환의 응용에 대한 연구가 활발히 진행되고 있다.

Lewis[1] 등은 인간 시각을 고려한 Marr[2]의 공간적 지역성(Spatial locality)에 기반한 웨이블릿 영상 압축 알고리즘을 제안하였다. 즉, 가장 주파수가 낮은 밴드의 웨이블릿 계수들은 그대로 전송하고, 나머지 부밴드는 부밴드 내에 0이 아닌 웨이블릿 계수들 중에 어느 것이 잡음에 상응하며 혹은 어느 정도 시각적으로 중요한지를 Marr의 공간적 지역화의 제약에 기반한 알고리즘을 적용하였다. 따라서, 중요한 상세(Detail) 정보가 영상에서 발생했다면 그 위치에 상응하는 웨이블릿 계수들은 한 옥타브(Octave) 이상에서 임계치를 넘을 것으로 판단할 수 있다.

Queiroz 등은 기존의 JPEG과 호환성이 있으며, 높은 압축률에서 블로킹 현상(Blocking artifacts) 및 링잉 현상(Ringing artifacts) 등으로 복원 화질이 저하되는 단점을 보완하여 JPEG 구조와 유사한 JPEG-Like 부호화 방법을 제안하였다 [3].

Xiong 등은 JPEG의 DCT 서브루틴[4]과 Shapiro가 제안한 embedded 제로 트리 양자화[5]를 결합하여 embedded 비트 스트림을 발생시키는 DCT 기반의 embedded 영상 압축 방법을 제안하였다 [6].

DCT 기반의 embedded 영상 압축 방법은 Shapiro의 EZW(Embedded zerotree wavelet)에 의해 얻어진 좋은 결과가 부분적으로는 웨이블릿 변환의 공헌 때문인 반면, DCT 또한 embedded 제로 트리 양자화와 결합될 수 있음을 제시하였고 복원된 영상은 JPEG과 Shapiro의 EZW보다 향상된 PSNR을 가져왔다.

Pennebaker 등은 부분집합 분할(Subset partitioning) 알고리즘을 기반으로 웨이블릿 계수들의 크기에 따른 부분 오더링(Ordering), 재정렬된 비트들의 비트 플레인 전송, 웨이블릿 변환 영상의 서로 다른 스케일의 자기 유사성(Self-similarity) 등을 이용한 EZW를 제안하였다 [7]. Xiong 등은 최근의 웨이블릿 부호화 알고리즘[1], [2]을 기반으로 자연 영상이 주파수와 공간 모두에서 에너지 집중도의 선형 결합으로서 잘 특징 지워질 수 있다는 사실

(영상내의 에너지 대부분이 저주파 영역에 집중되어 있으며 남아있는 고주파성분의 대부분 에너지가 에지 주위에 몰려있음)을 이용한 매우 효율적인 영상 압축 방법을 제안하였다. 즉, 매우 간단한 양자화를 통해 웨이블릿 변환의 주파수와 공간상 모두에서의 에너지 집중도 특성을 이용하고 웨이블릿의 공간상의 집중도 특성을 이용하기 위해 고주파 성분 계수의 공간적 영역은 0의 값을 갖도록 제로트리 양자화 하였다 [8].

두 채널(Dyadic) 웨이블릿 변환에 의해서 얻어진 계층적 신호들은 영상의 통계적 특성을 이용하였고 또한, HVS(Human Visual System)에 잘 맞는 양자화를 설계하는 데 편리한 구조를 제공하였다. 웨이블릿 변환은 신호를 다해상도(Multiresolution)로 분해하며, 입력 신호는 서로 다른 해상도로 표현하기 위한 부 신호들과 나머지 신호들을 표현하기 위한 세부 신호들로 나누어진다. 웨이블릿으로 변환된 영상신호는 서로 다른 주파수 특징을 갖고 이들간에는 어느 정도의 상관관계가 존재한다. 웨이블릿 변환을 이용하여 영상을 다해상도로 분해한 후 영상을 압축하는 것은 기존의 DCT를 이용한 방법보다 압축률을 높일 수 있고 블로킹 현상이 일어나지 않아 인간의 시각에 보다 적절한 압축방법이다 [9], [10].

인간의 시각은 세밀한 부분보다 불연속성이나 갑작스런 밝기의 변화와 같이 해상도가 낮은 부분의 영상에 대해 더 민감한 반응을 보인다.

웨이블릿 변환으로 분해된 부대역 영상들은 이러한 인간의 시각적 특성을 잘 반영하고 있어 영상을 효율적으로 압축할 수 있다. 즉, 영상 신호를 웨이블릿 변환을 이용하여 다해상도로 분해하고 인간시각 특성을 고려하여 추출된 에지정보를 이용하여 분해된 부밴드 영상을 양자화 한 후 양자화 된 계수들의 통계적 특성과 상관관계를 이용하면 효율적으로 압축할 수 있다.

이와 같은 바탕에서 최근 웨이블릿을 이용한 효과적인 영상 압축 기법이 많은 관심을 끌고 있으며, 본 논문에서는 최근 MPEG-4에서도 표준의 일부로 채택될 전망이다 웨이블릿 기반의 정지영상 압축기법에 대해서 개관하고 인간 시각 가중 양자화를 이용한 효율적인 부호화 방법을 제시한다. 또한 제안한 웨이블릿 기반 정지영상 압축 기법과 MPEG-4에 제시된 압축 기법에 대한 성능 평가를 토대로 제안한 알고리즘의 효율성을 입증한다.

본 논문의 구성은 다음과 같다.

2장에서는 쌍직교(Biorthogonal) 웨이블릿 변환과 MPEG-4에 제시된 웨이블릿 기반의 정지영상 압축 알고리즘에 대해 다루고 3장에서는 인간 시각 가중 양자화 및 제안하는 부호화 방법에 대해 다룬다.

4장에서는 2, 3장에서 설명한 압축 기법의 성능평가 및 비교, 분석을 토대로 본 논문에서 제시한 알고리즘의 우수성을 입증하고 5장에서 결론을 맺는다.

II. 쌍직교 웨이블릿 변환과 MPEG-4에서의 정지영상 압축 방법

1. 쌍직교 웨이블릿 변환

영상을 웨이블릿 변환하게 되면 서로 다른 단계의 다중 해상도 부밴드 영상들을 얻을 수 있다. QMF를 통과한 영상은 저주파 성분(LL)과 방향성을 가지는 고주파 성분(LH, HL, HH)으로 분해되며, LL 밴드의 영상은 원 영상에 비하여 해상도가 반으로 줄어든다. LH, HL, HH 밴드들은 각각 수직, 수평, 대각 방향의 에지 성분을 가지는 고주파 대역의 부밴드 영상이다.

그림 1은 실험에 사용된 표준영상인 Lena 영상의 3단계 웨이블릿 변환된 영상이다. 그림 1에서 볼 수 있는 바와 같이 저주파 대역의 영상은 원 영상의 저대역 통과된 성분이므로 닭은꼴을 가지고 있다. 고주파 대역의 영상은 각각 수직, 수평, 대각 방향의 에지 성분을 가지고 있으며, 저주파 성분과 마찬가지로 각 대역별로 닭은 꼴을 가지고 있다.

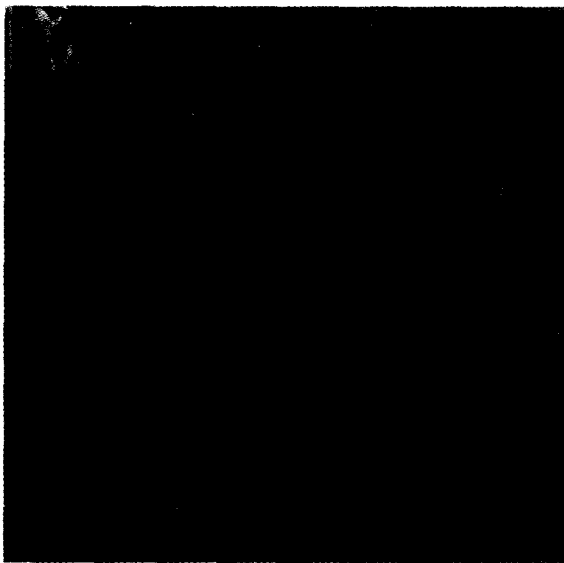


그림 1. Lena 영상을 3단계 웨이블릿 분해한 영상
Fig. 1. 3 level wavelet analysis for Lena image

직교 웨이블릿 변환(Orthogonal Wavelet Transform)에서는 완전복원을 위한 웨이블릿 필터는 항상 짝수 탭의 필터이다. 따라서 직교 웨이블릿 변환에서는 비선형 위상 응답을 갖는다. 반면 쌍직교 웨이블릿 변환에 사용되는 웨이블릿 필터는 직교성(Orthogonality)보다는 선형위상과 완전복원을 만족하도록 필터를 설계하였다.

본 논문에서 소개하는 정지영상 압축 기법은 웨이블릿 변환 고유의 자기 상관성을 이용하여 가장 주파수가 낮은

LL 밴드의 영상으로부터 다른 부밴드 계수를 양자화하기 위한 양자화의 크기를 구한다.

MPEG-4의 정지영상 압축 방법인 ZTE(ZeroTree Entropy Coding)도 웨이블릿 변환 고유의 자기 상관성을 이용한 방법이다 [11]. 또, 웨이블릿 변환 고유의 자기 상관성을 이용한 다른 방법으로 Shapiro의 EZW 방법이 있다 [5]. EZW에서는 홀수 탭의 길이를 가지는 직교 웨이블릿 변환을 사용하였다. 이 변환은 웨이블릿 변환시에 사용되는 QMF 필터가 홀수 탭의 길이를 가지므로 완전복원을 만족하지는 못하지만 선형 위상 특성을 가진다는 장점을 사용하였다. 즉, 완전복원을 약간 희생하면서 선형 위상 응답 특성을 중시한 것이다. 하지만 쌍직교 웨이블릿 변환은 선형 위상 응답의 특성을 만족하면서도 완전복원을 만족한다. 본 논문에서는 자기 상관성을 이용하여 양자화를 하기 때문에 선형 위상 특징을 갖는 쌍직교 웨이블릿 변환을 사용하는 것이 비선형 위상 응답을 갖는 직교 웨이블릿 변환을 사용하는 것 보다 유리하다.

2. MPEG-4에서의 정지영상 압축 방법

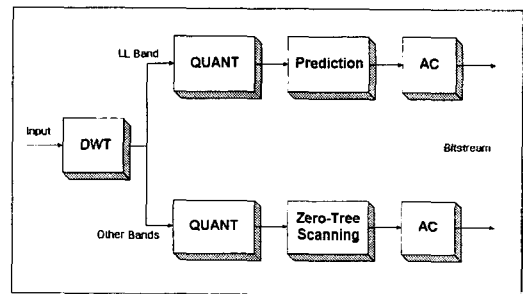


그림 2. MPEG-4 정지영상 부호화기의 구조
Fig. 2. Block diagram of the MPEG-4 still image encoder

MPEG-4에서 제시된 제로트리 웨이블릿 기반의 기본적인 부호화기는 그림 2와 같은 구조로 구성되어 있다. 즉, 입력 영상에 대하여 웨이블릿 변환을 수행하고 변환된 웨이블릿 계수들에 대하여 가장 주파수가 낮은 LL 밴드와 다른 부밴드들에 대하여 서로 다른 부호화 방법을 사용한다. 먼저 가장 주파수가 낮은 LL 밴드와 다른 부밴드들을 독립적으로 양자화를 수행한다. 다음에는 양자화된 계수들에 대하여 가장 낮은 주파수 LL 밴드는 예측 부호화를 수행한 후 예측 에러를 적응 산술 부호화[12]하고 나머지 부밴드들에 대해서는 제로트리 스캔을 수행하여 엔트로피 부호화 한다. MPEG-4의 정지영상 압축에 쓰이는 웨이블릿 필터는 Daubechies의 9/3 탭 쌍직교 웨이블릿 필터를 사용하고 있다 [13].

(1) LL 밴드 부호화

가장 주파수가 낮은 LL 밴드는 다른 부밴드들과는 다

르게 부호화 되며, LL 밴드의 웨이블릿 계수들은 균일 양자화기(Uniform midriser quantizer)를 이용하여 양자화한 후, 적응적 예측방법을 이용하여 DPCM을 수행하고 예측 에러를 적응 산술부호화(Adaptive arithmetic coding) 한다.

(2) 다른 부밴드의 부호화

MPEG-4에서의 정지영상 압축을 위한 양자화는 다양한 응용 분야에 대한 스케일러빌리티(scalability)를 제공하기 위해서 다중 레벨 양자화 방법을 사용한다. 양자화 레벨은 부호기에 의해 정의되며, 양자화 스텝사이즈의 2배에 해당하는 데드존(Dead zone)을 갖는 균일 양자화기(Uniform midriser quantizer)를 사용한다.

Shapiro는 분해된 부밴드 영상의 상위 단계를 Parent, 하위단계를 Children이라 정의하고, Parent-Children간의 자기 상관성을 이용하여 연속적인 근사 양자화(Successive approximation quantization)를 사용하였다. 이 방법이 바로 EZW이다. EZW에서는 positive significant, negative significant, zerotree root, isolated zero의 4가지 심볼(Symbol)을 이용하여 부호화 하였다. EZW에서는 연속 근사 양자화를 사용하기 때문에 일단 양자화된 계수들은 부호화 할 수 없다.

MPEG-4의 정지영상 압축에 사용되는 방법은 ZTE이다 [11]. 이 방법은 Shapiro의 EZW에서 사용된 심볼과는 조금 다른 3가지의 심볼을 사용한다. zerotree root, valued zerotree root, value의 심볼을 사용한다. valued zerotree root는 4개의 모든 Children이 모두 zerotree root이고 자기 자신의 계수는 0이 아닌 경우이고, value는 0이 아닌 Children들이 있는 계수이다. 따라서 ZTE를 사용하면 일단 양자화된 계수를 제로트리 구조를 이용하여 부호화를 할 수 있게 된다. 이 점이 EZW와는 다른 점이다. Significant map과 valued zerotree root의 계수값, value의 계수값들은 적응 산술부호화 한다.

III. 양자화기와 부호화

본 장에서는 앞장에서 설명한 MPEG-4에서의 정지영상 압축 기법보다도 간단하면서도 인간 시각 가중 양자화기를 사용하여 시각 특성을 향상시킨 새로운 방법의 영상 압축 기법을 소개한다.

영상의 압축을 위해서는 자연의 영상이 가지고 있는 통계적인 중복성(Statistical redundancy)을 제거하는 방법을 사용한다. 이산여현변환, DPCM, 서브밴드 영상들의 엔트로피 부호화 등이 통계적인 중복성을 제거하는 방법의 예들이다. 이러한 통계적인 중복성만을 제거하는 방법은 영상의 압축에 있어서 제한적인 양만을 사용할 뿐이며 고압축률을 얻는데 한계가 있다.

따라서 비중복성(Nonredundancy) 정보들을 제거하기

위한 양자화 과정을 거쳐야 한다. 그러나, 이러한 성질을 영상압축에 적용하면 영상에서의 에지(Edge)와 같이 시각적으로 중요한 부분에 에러를 야기함으로써 복원 영상의 화질 열화를 일으킨다. 따라서 영상의 압축에 사용하기 위한 양자화기는 인간의 시각 특성을 고려하여 설계되어야 한다. 즉, 양자화의 크기는 인간 시각이 민감한 부분에 대하여는 작게, 둔감한 부분에 대하여는 크게 설정하여야 한다. 본 논문에서는 이러한 인간 시각 특성을 고려하여 설계된 인간 시각 가중 양자화기를 사용하였다 [14], [15].

1. 인간 시각 가중 양자화기

인간 시각 가중 양자화기는 인간 시각 특성을 크게 세 가지로 고려하여 설계하였다. 첫째, 주파수 특성으로 높은 주파수에 비하여 낮은 주파수를 가지는 영상 정보는 인간 시각에 민감한 성질을 가진다. 웨이블릿 변환은 영상을 인간 시각과 유사한 주파수 특성을 가지는 부분 영상을 분해하는 특성을 가지므로 주파수 특성을 쉽게 영상 압축에 적용할 수 있다. 둘째, 배경의 밝기에 따라서 화질의 판단이 달라진다. 즉, 중간 밝기의 배경이 밝거나 어두운 배경보다 인간 시각에 보다 민감하다. 셋째, 영상에서 대비도가 낮은 부분이 대비도가 높은 부분보다 더 민감하다.

웨이블릿 변환된 영상을 주파수 대역, 배경의 밝기, 대비도에 의해 가장 주파수가 낮은 LL 밴드의 영상으로부터 인간 시각적으로 추정된 양자화기를 사용하여 다른 부밴드들을 양자화를 한다.

(1) 주파수 대역에 따른 양자화의 크기

높은 주파수 대역에 존재하는 양자화 오차는 낮은 주파수 대역에 존재하는 양자화 오차보다 잘 인지되지 않는다고 알려져 있다. 따라서 높은 주파수 대역에 대하여는 양자화를 크게 하고, 낮은 주파수 대역에 대하여는 양자화의 크기를 작게 설정한다. 각 주파수 대역에 대하여 일정한 값으로 Lewis가 사용한 값을 사용하였다 [1].

$$Q_{FS}(L, O) = \begin{cases} \sqrt{2}, & \text{if } O=HH \\ 1, & \text{otherwise} \end{cases} \times \begin{cases} 10.00, & L=1 \\ 3.20, & L=2 \\ 1.60, & L=3 \\ 1.00, & L=4 \end{cases} \quad (1)$$

일반적으로 인간은 수평, 수직 성분보다 대각 성분에 둔감하다고 알려져 있다. 식 (1)에서 보는 바와 같이 대각 성분의 에지를 가지는 HH 밴드의 영상에 대해서는 수평, 수직 방향보다 오차에 둔감하기 때문에 $\sqrt{2}$ 의 가중치를 두었다.

(2) 배경의 밝기에 따른 양자화의 크기

인간 시각은 주위 배경이 아주 밝거나 어두울 경우에는 오차 인지도가 떨어지고, 중간 밝기일 경우에 오차 인지도

가 가장 높다고 알려져 있다. 각 LH, HL, HH 대역의 계수에 대한 배경의 밝기는 LL 대역의 대응하는 계수 및 이웃하는 세 개의 웨이블릿 계수의 평균으로 계산될 수 있다.

$$BL(L, O, x, y) = \frac{1}{4} \sum_{i=0}^1 \sum_{j=0}^1 W(L_{max}, LL, \frac{x}{2^{L_{max}-L}} + i, \frac{y}{2^{L_{max}-L}} + j) \quad (2)$$

배경의 밝기에 따른 양자화의 크기는 배경의 밝기가 중간 값을 가질 경우 가장 작게, 밝거나 어두운 정도가 많아질수록 커지도록 결정한다.

$$Q_{BL}(L, O, x, y) = \begin{cases} Q_{min} + \frac{Q_{max} - Q_{min}}{127 - G_1} |127 - BL(L, O, x, y)|, & G_1 < BL(L, O, x, y) \leq 127 \\ Q_{min} + \frac{Q_{max} - Q_{min}}{G_2 - 127} |127 - BL(L, O, x, y)|, & 127 < BL(L, O, x, y) \leq G_2 \\ Q_{max}, & otherwise \end{cases} \quad (3)$$

여기서 Q_{min} 및 Q_{max} 는 최소 양자화 및 최대 양자화 크기를 나타내고, G_1 , G_2 는 최대 양자화 크기를 가지고 있는 평균 밝기 값의 문턱값이다.

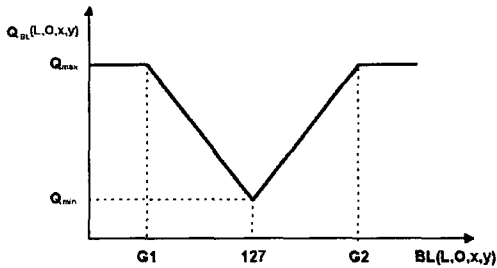


그림 3. 배경의 밝기에 따른 양자화의 크기
Fig. 3. Quantization step size by background luminance

(3) 대비도에 의한 양자화의 크기

인간 시각은 밝기의 변화가 큰 영역에서는 둔감하고 밝기의 변화가 작은 곳에서는 민감하다. 대역의 각 계수에 대한 대비도는 LL 대역의 대응 계수와 이웃하는 세 개의 계수의 차이를 이용하여 다음과 같이 결정된다.

$$MS(L, O, x, y) = \frac{1}{3} \sum_{i=0}^1 \sum_{j=0}^1 |W(L_{max}, LL, \frac{x}{2^{L_{max}-L}} + i, \frac{y}{2^{L_{max}-L}} + j) - W(L_{max}, LL, \frac{x}{2^{L_{max}-L}} + i, \frac{y}{2^{L_{max}-L}} + j)| \quad (4)$$

식 (4)에 의해서 계산된 밝기의 변화는 대비도에 의한 양자화의 크기를 결정하는데 사용된다.

$$Q_{MS}(L, O, x, y) = \begin{cases} Q_{min} & if MS(L, O, x, y) < G_3 \\ Q_{max} & if MS(L, O, x, y) > G_4 \\ Q_{min} + \frac{Q_{max} - Q_{min}}{G_4 - G_3} (MS(L, O, x, y) - G_3) & otherwise \end{cases} \quad (5)$$

여기서 G_3 및 G_4 는 최소 및 최대 양자화의 크기를 나타내는 밝기의 변화에 대한 문턱값이다.

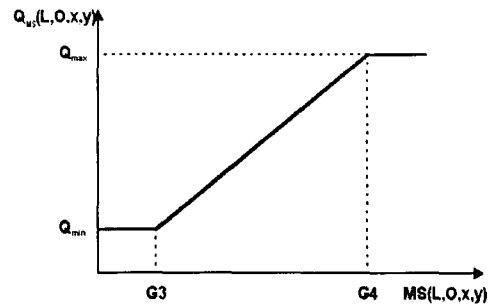


그림 4. 대비도에 따른 양자화의 크기
Fig. 4. Quantization step size by contrast

(4) 양자화의 크기

주파수 대역, 배경의 밝기 및 대비도에 의해 결정된 양자화의 크기를 조합하여 양자화기를 설계한다.

$$Q(L, O, x, y) = q \times Q_{FS}(L, O) \times Q_{BL}(L, O, x, y) \times Q_{MS}(L, O, x, y) \quad (6)$$

여기서 q 는 압축률을 조절하기 위한 값이다.

2. 부호화

영상을 웨이블릿 변환하고 변환 계수들을 적절한 양자화기로 양자화 한 후 양자화 된 계수들은 적절한 방법으로 무손실 압축하여야 한다.

가장 주파수가 낮은 LL 밴드의 영상은 다른 부밴드들과 다른 통계적인 특징과 그 중요성 때문에 독립적으로 DPCM 하였고, 다른 부밴드들은 적절한 스캔 방향에 따라 스캔하면서 산술부호화 하였다.

(1) LL 밴드의 부호화

가장 주파수가 낮은 LL 밴드의 영상은 다른 부밴드의 영상들과는 다른 통계적 특징을 가지고 있고, 다중 해상도 압축에서 상대적으로 중요한 정보를 가지고 있으므로 압축할 때 많은 비트를 할당하여야 한다. 이것은 LL 밴드의

영상과 통계적 특성이 비슷한 이산여현변환에서 DC계수를 부호화 할 때 JPEG, MPEG등의 국제 표준에서 DC계수를 AC계수에 비해 상대적으로 적은 일정한 Q값으로 나누어줌으로써 더 많은 비트를 할당하는 것과 비슷하다. 또한 LL 밴드의 영상은 다른 부밴드의 양자화를 위한 양자화기의 설계에 이용되므로 독립적으로 DPCM을 사용하여 압축하였다. LL 밴드를 부호화할 때 양자화를 하지 않는 무손실 부호화와 선형 양자화를 하는 손실 부호화의 두 가지 방법을 사용하여 부호화 하였다.

LL 밴드의 영상을 양자화를 하지 않는 무손실 압축 하면 정보의 손실 없이 양자화기를 설계할 수 있다. 그러나, LL 밴드의 영상을 양자화 없이 무손실 압축하게 되면 압축률과는 관계없이 LL 밴드에는 일정한 비트가 할당되게 된다. 압축률이 높지 않을 경우에는 문제가 되지 않지만 압축률이 높아지면 상대적으로 다른 부밴드에는 더 적은 비트가 할당되게 되는 문제점이 있다. 이러한 문제점을 해결하기 위해 LL 밴드 영상을 DPCM하기 전에 적절하게 선형 양자화를 하여 손실 부호화를 하게 되면 무손실 부호화에 비해 더 많은 비트가 다른 부밴드에 할당된다.

(2) LL 밴드의 무손실 부호화

LL 밴드의 부호화에는 JPEG의 무손실 압축 모드에서 사용되는 방법을 변형하여 사용하였다 [15]. DPCM에서 사용되는 예측 모드는 JPEG의 무손실 압축 모드에서 사용되어 지는 것과 MPEG-4의 적응적 DC 계수 예측에 쓰이는 방법을 사용하였다 [13].

8가지의 예측모드를 모두 사용하여 예측 에러가 가장 적은 예측모드를 선택하게 된다. 하지만 MPEG-4에서는 적응적 DC 계수 예측 모드 하나만을 사용한다. 실제로 Lena 영상에 대해 LL 밴드를 DPCM 하면 MPEG-4의 적응적 DC 계수 예측모드가 선택되지 않고 표 1에서의 모드 1이 선택된다.

(3) 선형 양자화를 이용한 LL 밴드의 손실 부호화

LL 밴드의 영상을 무손실 부호화 하면 압축률과 관계 없이 일정한 양의 비트가 LL 밴드에 할당된다. 압축률이 높아지면 LL 밴드에 상대적으로 많은 비트가 할당되고 다른 부밴드에는 상대적으로 적은 비트가 할당되게 된다. 따라서 압축률이 높아지게 되면 LL 밴드에 할당되는 비트를 줄이고 다른 부밴드에 할당되는 비트를 늘려야 한다. 이렇게 하기 위해서 LL 밴드를 선형 양자화 하여 DPCM 함으로써 LL 밴드에 할당되는 비트의 양을 줄였다. LL 밴드의 영상을 선형 양자화를 한 후에는 (2)의 방법과 같은 방법으로 부호화 하였다.

시각 가중 양자화기는 q 값을 가지고 압축률을 조절한다. LL 밴드에서의 선형 양자화를 위한 양자화의 크기는 실험적으로 구하였다.

표 2에 q 값에 따른 LL 밴드에서의 양자화의 크기를 나타내었다.

표 1. DPCM 부호화를 위한 여러 가지 예측 방법
Table 1. Predictors for DPCM

	c	b	
	a	x	
Mode	예측방법		
0	a		
1	b		
2	c		
3	a+b-c		
4	a+(b-c)/2		
5	b+(a-c)/2		
6	(a+b)/2		
7	if(a-b < b-c) c else a		

표 2. q 값에 따른 LL 밴드에서의 양자화의 크기
Table 2. LL band quantization step size by q value

q	LL 밴드 영상의 선형 양자화 크기
$q < 0.5$	1
$0.5 \leq q < 1.3$	2
$1.3 \leq q < 2.0$	4
$2.0 \leq q$	8

(4) 다른 부밴드의 부호화

다른 부밴드들은 인간 시각 특성을 고려하여 LL 밴드로부터 설계된 인간 시각 가중 양자화기를 이용하여 각 부밴드의 계수들을 양자화 한 후 산술 부호화 하였다.

웨이블릿 변환으로 영상을 다중해상도로 분해하면 분해된 각 부밴드 영상들은 수직, 수평, 대각 방향의 서로 다른 방향 성분을 나타낸다. 이러한 성질을 이용하여 분해된 각 부밴드 영상 별로 스캔 방향을 다르게 하면 산술 부호

화시에 더 효과적으로 압축할 수 있다.

각 부밴드 영상의 특성에 맞는 스캔 방향을 따라 양자화 된 웨이블릿 계수들을 부호화 하였다. 실제로 부밴드 영상 별로 스캔 방향을 달리 할 경우 효율이 증가하는 것을 확인하였다.

그림 5에 각 부밴드 영상에 따른 스캔 방향을 나타내었다.

그림 6과 그림 7에 본 논문에서 제안하는 영상 압축 알고리즘의 부호기와 복호기의 구조를 나타내었다.

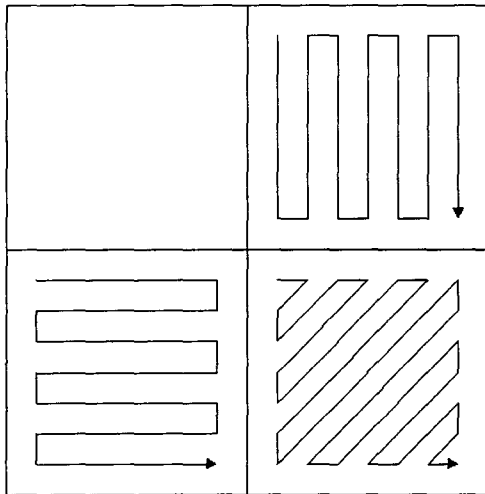


그림 5. 각 부밴드의 영상에 따른 스캔 방향
Fig. 5. Scanning directions for the subbands

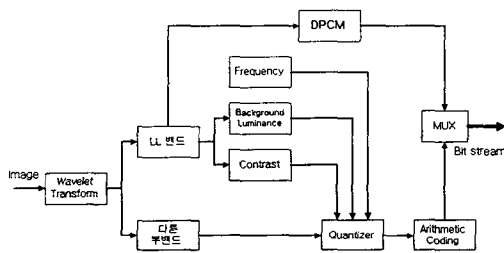


그림 6. 부호기의 블록 다이어그램
Fig. 6. Block diagram for Encoder

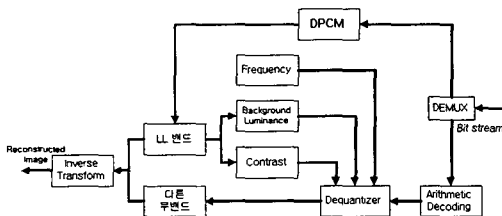


그림 7. 복호기의 블록 다이어그램
Fig. 7. Block diagram for Decoder

IV. 모의 실험 및 결과

제안한 영상 압축 방법을 이용하여 표준 영상을 가지고 모의 실험을 하였다. 제안한 기법과 MPEG-4에서의 기법, Shapiro의 EZW 기법, 그리고 현재까지의 정지영상 압축 표준인 JPEG을 서로 비교하였다.

실험에 사용된 영상은 512×512의 크기를 갖는 Lena 영상이다. 쌍직교 웨이블릿 변환을 위한 웨이블릿 필터로는 표 3의 Daubechies 9/7 탭 필터를 사용하였고 4단계의 웨이블릿 변환을 수행하였다. MPEG-4의 정지영상 압축 기법에서는 Daubechies의 9/3 탭 필터를 사용하였고 3단계의 웨이블릿 변환을 수행하였으며 가장 주파수가 낮은 LL 밴드와 다른 부밴드들은 서로 독립적으로 양자화 하였다.

방법 I은 인간 시각 가중 양자화를 사용하여 다른 부밴드를 양자화 하고 가장 주파수가 낮은 LL 밴드의 영상은 무손실 압축한 방법이고, 방법 II는 마찬가지로 인간 시각 가중 양자화를 사용하고 가장 주파수가 낮은 LL 밴드의 영상은 선형 양자화를 이용하여 손실 압축한 방법이다.

표 4에서 보는 바와 같이 실험 결과 제안한 방법 I과 방법 II가 모든 경우에 있어서 MPEG-4에서의 정지영상 압축 방법과 현재까지의 정지영상 압축 표준인 JPEG보다도 훨씬 좋은 성능을 나타내었다. Shapiro가 제안한 EZW와 비교해 보면 8:1과 128:1에서 PSNR이 약간 떨어지는 것을 볼 수 있다. 그렇지만 8:1과 같이 압축률이 높지 않은 경우에 있어서는 PSNR이 높기 때문에 0.01dB 정도의 차이는 눈에 띄지 않는다. 실제 실험 결과 Lena 영상의 경우는 37dB 이상이 되면 원영상과의 차이를 발견하기 어렵다. 그리고 방법 I에서의 128:1과 같이 압축률이 아주 높은 경우를 제외하고는 EZW 보다도 제안한 방법이 더 우수한 결과를 얻을 수 있었다. 하지만 제안한 정지영상 압축 방법은 EZW보다 부호화 및 복호화 방법이 복잡하지 않다는 장점을 가지고 있다.

그림 8, 9, 10, 11은 각각의 방법에 대한 64:1에서의 결과 영상들이다.

표 3. Daubechies 9/7 탭 웨이블릿 필터의 계수
Table 3. Daubechies 9/7 tap wavelet filter coefficients

h_0	h_1
0.026748757411	
-0.016864118443	-0.04563588155696
-0.078223266529	0.02877176311397
0.266864118443	0.29563588155671
0.602949018236	-0.55754352622844
0.266864118443	0.29563588155671
-0.078223266529	0.02877176311397
-0.016864118443	-0.04563588155696
0.026748757411	

표 4. Lena 영상에 대한 압축 결과
Table 4. Compression results for Lena image

압축률	방법 I PSNR(dB)	방법 II PSNR(dB)	MPEG-4 PSNR(dB)	EZW PSNR(dB)	JPEG PSNR(dB)
8:1	39.54	39.54	38.01	39.55	37.69
16:1	36.55	36.55	35.32	36.28	34.79
32:1	33.48	33.53	32.51	33.17	31.55
64:1	30.46	30.58	30.00	30.23	27.92
128:1	27.28	27.80	27.38	27.54	-



그림 8. 방법 I로 64:1로 압축한 Lena 영상
Fig. 8. Compression result at 64:1 by the method I



그림 9. 방법 II로 64:1로 압축한 Lena 영상
Fig. 9. Compression result at 64:1 by the method II



그림 10. MPEG-4의 ZTE 방법으로 64:1로 압축한 Lena 영상
Fig. 10. Compression result at 64:1 by the MPEG-4 ZTE



그림 11. JPEG으로 64:1로 압축한 Lena 영상
Fig. 11. Compression result at 64:1 by the JPEG

MPEG-4의 정지영상 압축 기법에서는 가장 주파수가 낮은 LL 밴드의 영상을 DPCM하기 위한 예측 방법을 하나만 사용함으로써 본 논문에서 제안하는 방법보다 효율이 낮았다. 앞서서도 언급한 바와 같이 실제 Lena 영상에서는 모드 1을 선택하여야 한다. 또한 DPCM한 후의 계수를 부호화하기 위해 산술부호화를 하였다. 압축률이 높은 경우에는 LL 밴드의 영상을 큰 양자화 스텝 사이즈로 양자화하기 때문에 DPCM한 후에 많은 수의 계수가 발생하지 않지만 압축률이 낮은 경우에 있어서는 DPCM한 후에 많은 수의 계수가 발생하여 산술 부호화를 하는 것이 효

과적인 방법이 되지 못한다. 따라서 실험결과에서 보는 것처럼 압축률이 낮을 경우가 압축률이 높은 경우보다 다른 방법들 보다 성능이 많이 떨어지는 것을 볼 수 있다.

V. 결 론

본 논문에서는 쌍직교 웨이블릿 변환의 특성과 MPEG-4에서 제시된 정지영상 압축 기법에 대하여 살펴보았다. 또한 MPEG-4에서의 정지영상 압축 기법보다도 간단하면서도 인간 시각 특성을 고려한 양자화기를 설계하여 이를 압축에 이용함으로써 시각 특성을 향상시킨 효율적인 압축 기법을 제안하였다.

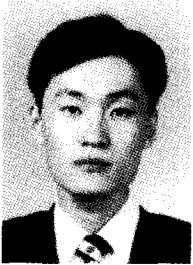
MPEG-4에서 제시된 알고리즘과 본 논문에서 제안한 알고리즘을 구현하여 성능 평가를 실시하였다. 평가 결과 제안한 방법이 모든 경우(LL 밴드의 손실 부호화, LL 밴드의 무손실 부호화)에 있어서 MPEG-4에서의 정지영상 압축 방법보다도 훨씬 좋은 성능을 나타내었다. 또한 압축률이 매우 높을 경우를 제외하고는 EZW 보다도 복잡도가 낮으면서 대부분의 경우 주관적, 객관적 화질이 우수함을 알 수 있었다.

웨이블릿 변환은 인간의 시각 특성을 잘 반영하고 있으므로 압축된 영상은 JPEG으로 압축할 때와 같이 블록간 왜곡이 거의 발생하지 않으며 전체적으로 JPEG보다 훨씬 우수한 화질의 향상을 가져왔다.

참 고 문 헌

- [1] A. Lewis and G. Knowles, "Image compression using the 2-D wavelet transform," *IEEE Trans. Image Processing*, vol. 1, pp. 244-250, Apr. 1992.
- [2] D. Marr, *Vision*, New York: Freeman, 1982.
- [3] R. Queiroz, C. K. Choi, Y. Huh, and K. R. Rao, "Wavelet transforms in a JPEG-like image coder," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 7, no. 3, pp. 419-424, 1997.
- [4] W. B. Pennebaker and J. L. Mitchell, *JPEG Still Image Data Compression Standard*, New York: Van Nostrand Reinhold, 1992.
- [5] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. Signal Processing*, vol. 41, pp. 3445-3463, Dec. 1993.
- [6] Z. Xiong, O. Guleryuz, and M. T. Orchard, "A DCT-based Embedded Image Coder," *IEEE Signal Processing Letters*, vol. 3, 11, pp. 289-290, Nov. 1996.
- [7] A. Said and W. A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, pp. 243-250, Jun. 1996.
- [8] Z. Xiong, K. Ramchandran, and M. T. Orchard, "Space-Frequency Quantization for Wavelet Image Coding," *IEEE Trans. Image Processing*, vol. 6, no. 5, pp. 677-693, 1997.
- [9] J. W. woods and S. O'Neil, "Subband coding of images," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 34, pp. 1278-1288, Oct. 1986.
- [10] A. Aldroubi and M. Unser, *Wavelets in Medicine and Biology*, CRC Press, pp. 43-46, 1996.
- [11] Stephen A. Martucci, Iraj Sodagar, Tihao Chiang, and Ya-Qin Zhang, "A Zerotree Wavelet Video Coder," *IEEE Trans. on CSVT*, vol. 7, no. 1, pp. 109-118, Feb. 1997.
- [12] I. H. Witten, R. Neal, and J. G. Cleary, "Arithmetic Coding for Data Compression," *Comm. ACM*, vol. 30, no. 6, pp. 520-540, Jun. 1987.
- [13] MPEG Video Group, "MPEG-4 Video Verification Model Version 8.0," ISO/IEC JTC1/SC29/WG11, Stockholm, Jul. 1997, draft in progress.
- [14] 김민구, 엄일규, 김재호, 정제창, "쌍직교 웨이블릿 변환과 시각 가중 양자화기를 이용한 영상압축", *신호처리 합동학술대회 논문집*, 제 10 권, pp. 217-220, 1997.
- [15] 엄일규, 김재호, "웨이브릿 영상 압축을 위한 인간 시각 가중 양자화기의 설계", *한국통신학회논문지*, 제 22 권, 제 3 호, pp. 493-505, 1997.
- [16] 민병석, 김승중, 정제창, "웨이브릿을 이용한 영상 압축", *한국통신학회지*, 제 14 권 9호, pp. 59-81, 1997

저 자 소 개



김 민 구

1996년 2월 한양대학교 공과대학 전자통신공학과 졸업(공학사)
현 재 한양대학교 대학원 전자통신공학과 석사과정
주관심분야 : 영상 통신 및 압축, 영상처리, 디지털 신호처리 등



김 승 종

1992년 2월 한양대학교 자연과학대학 수학과 졸업(이학사)
1994년 2월 한양대학교 대학원 전자통신공학과 졸업(공학석사)
현 재 한양대학교 대학원 전자통신공학과 박사과정
주관심분야 : 영상 통신 및 압축, 영상처리, 디지털 신호처리 등



정 제 창

1980년 2월 서울대학교 공과대학 전자공학과 졸업(공학사)
1982년 2월 한국 과학기술원 전기전자공학과 졸업(공학석사)
1990년 8월 미시간대학교(앤아버) 전기공학과 졸업(공학박사)
1982년 2월 ~ 1986년 7월 한국방송공사 기술연구소 연구원(뉴미디어 연구개발)
1990년 9월 ~ 1991년 1월 미시간대학교(앤아버) Post-dotorial Research Fellow
1991년 2월 ~ 1995년 2월 삼성전자 멀티미디어 연구센터 신호처리연구소 수석연구원
(HDTV 및 멀티미디어 연구개발)
현 재 한양대학교 전자통신공학과 조교수
주관심분야 : 영상 및 음성 압축, 영상처리, 디지털 신호처리, 디지털 통신, VLSI 설계 등