

R-D 최적화를 이용한 정지영상 부호화 방법

이 병 기, 호 요 성
광주과학기술원 정보통신공학과
전화 : 062-970-2263 / 핸드폰 : 016-802-8411

R-D Optimized Image Coding Algorithm

Byong-Ki Lee, Yo-Sung Ho
Dept. of Information & Communications, K-JIST
E-mail : {bklee, hoyo}@kjist.ac.kr

Abstract

In this paper, we propose the efficient algorithm for wavelet-based still image coding method that uses the rate-distortion theory. Main coding algorithm is the SPIHT(Set Partitioning In Hierarchical Trees). In this algorithm, we use the R-D(Rate-Distortion) slope as the criterion for the coding-order of wavelet coefficient in SPIHT's lists.

The proposed method performs better than the previous SPIHT algorithm and the RDE(Rate-Distortion optimized Embedding) algorithm in the aspect of PSNR(Peak Signal to Noise Ratio) performance.

I. 서론

최근 들어 정지영상에 대한 웨이브렛(wavelet) 변환의 적용이 점차 증가하고 있다. 이러한 경향은 JPEG(Joint Picture Experts Group) 2000 표준에서 기본 변환 방식을 기존 JPEG에서 사용하던 DCT(Discrete Cosine Transform) 에서 웨이브렛 변환으로 바꾼 것에서도 분명히 드러나고 있다. DCT를 기본으로 한 블록 단위의 부호화는 낮은 비트율에서 블록간의 경계가 나타나는 현상(blocking effects)이 나타나는 단점을 가지는 반면, 웨이브렛 변환은 화면 전체

를 단위로 하여 변환하므로 위와 같은 블로킹 효과가 없을 뿐만 아니라, 성능 면에서도 DCT에 비해 높은 결과를 보이고 있다.

이러한 웨이브렛 변환을 이용한 정지영상의 부호화는 이미 많은 연구가 진행되었고, 이들 중 대표적인 것들은 다음과 같다. Shapiro는 각 부대역 간의 상관관계를 이용하여 부호화 효율을 향상시키기 위해 EZW(Embedded Zerotree Wavelet) 방식을 제안하였고[2], Said와 Pearlman은 부호화 과정에 여러 가지 리스트를 생성하는 개념을 사용하여, 이러한 EZW 방식을 더욱 개선하는 SPIHT(Set Partition In Hierarchical Trees) 방식을 제안하였다[3]. 이들 방식과 독립적으로, Li와 Lei는 각 웨이브렛 계수의 R-D(Rate-Distortion) 경사를 추정하여 계수들의 스캔 순서를 최적화 하는 RDE(Rate-Distortion optimized Embedding) 방식을 제안하였다[4].

RDE 방식은 SPIHT 방식에 비해 높은 성능을 보이지만, 전체 웨이브렛 계수를 RD 경사 임계값(threshold)의 개수만큼 스캔해야 하므로, 복잡도가 크게 증가한다. 반면, SPIHT 방식은 각 리스트에 있는 좌표들, 일반적인 래스터(raster) 방식으로 스캔하기 때문에, 본 논문에서는 여기에 RDE 방식을 적용하여 리스트 내의 좌표들의 스캔 순서를 최적화 하여 동일 bpp에서 부호화 효율을 최적화 할 수 있는 알고리즘을 제안한다.

II. SPIHT 부호화 알고리즘

웨이브렛 변환은 일반적으로 QMF(Quadrature Mirror Filter)를 사용하여 구현된다 [1]. 영상의 행과 열 방향으로 반복적으로 웨이브렛 변환을 수행하면, 그림 1과 같은 부대역(Subband) 분할 결과를 얻을 수 있다. 또한 이들 각 부대역 사이에는 그림 1과 같은 부모-자식(Parent-Children) 상관관계가 존재한다.

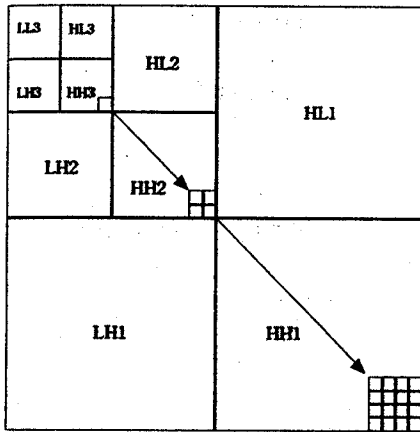


그림 1. 웨이브렛 계수의 부대역간 상관관계

SPIHT 부호화에는 세 개의 리스트가 쓰인다. 먼저, LSP(List of Significant Pixels)는 이전 비트 평면에서 중요하다고 판단된 웨이브렛 계수들의 좌표를 포함하고, LIP(List of Insignificant Pixels)는 현재 비트 평면까지 중요하다고 판단되지 않은 웨이브렛 계수들의 좌표를 포함한다. LIS(List of Insignificant Sets)는 현재 비트 평면까지 중요성 여부가 판단되지 않은 Tree 구조의 가장 상위에서 있는 계수들의 좌표를 Tree의 형식과 함께 포함한다.

여기서 Tree 구조의 형식은 A와 B의 두 가지 형식이 존재하고, A형식은 자식부터 모든 자손(Descendants)의 중요성을 판단하고, B형식은 자식을 제외한 모든 자손의 중요성을 판단한다.

SPIHT 방법의 부호화 과정은 다음과 같다.

먼저, 초기화 과정은 최상위 비트 평면에서 시작하며, 각 리스트를 초기화한다. LSP는 비워두고, LIP에는 LL 대역에 있는 모든 계수들의 좌표를 넣는다. 마지막으로 LIS에는 LIP에 있는 계수들 중 자식을 지니는 모든 계수들의 좌표를 A형식과 함께 저장한다.

다음으로 LIP 정렬 과정(Sorting Pass)에서는 LIP에 있는 각 계수에 대해, 현재 비트 평면에서의 중요성

판단 여부를 부호와 함께 출력한다. LIS 정렬 과정에서는 LIS에 있는 각 Tree 구조에 대해, 현재 비트 평면에서의 중요성 판단 여부를 출력하고, 만일 중요한 계수가 Tree 구조 내에 존재하면, Tree를 분할하고, 위의 작업을 반복한다.

세분화(Refinement) 과정에서는 이전까지의 비트 평면에서 중요하다고 판단된 계수들을 세분화하는 과정으로, 여기서는 LIP 정렬 과정에 필요한 부호 비트의 부호화가 존재하지 않는다.

임계값 갱신 과정은 비트 평면을 최상위 비트부터 하나씩 제거하는 작업과 동일하다. 이 과정을 거친 후 다시 LIP 정렬 과정으로 돌아가서 이전의 작업을 반복한다[3].

III. RDE 부호화 알고리즘

RDE 부호화의 기본 개념은 가장 R-D 경사가 큰 부호를 가장 먼저 부호화 한다는 것이다. 이러한 개념은 그림 2에 쉽게 표현된다. 그림 2에서 실선은 일반적인 스캔 순서로 5개의 부호를 부호화한 비트율과 왜곡의 변화율을 나타내고, 점선은 R-D 경사에 따라 최적화되어 정렬된 순서로 5개의 부호를 부호화한 비트율과 왜곡의 변화율을 나타낸다. 그림 2에서 볼 수 있듯이 최적적인 비트율과 왜곡은 두 경우에 동일하지만, 중간 단계에서의 비트율과 왜곡은 차이가 있게 된다. 이러한 특성은 임베디드(embedded) 비트열(bitstream)을 만들때 있어서 주어진 비트율에서 최적의 영상 품질을 지니도록 할 수 있다.

이러한 RDE 부호화의 성능을 최적화하기 위해서는 모든 웨이브렛 계수를 R-D 경사에 따라 정렬하여 그 순서대로 부호화 하는 것이다. 하지만, 이러한 방식은 계수들의 부호화 순서를 위해 좌표도 함께 전송해야 하는 부담을 지니게 되므로 오히려 비트율이 증가하게 된다. 그리고, 실제 계수를 이용하여 R-D 경사를 계산한다면, 복호기에서는 실제 계수에 의한 R-D 경사를 알 수 없으므로, 이러한 방식은 적합치 않다. 이러한 단점을 보완하기 위해, RDE 방식에서는 R-D 경사 임계값을 정의한다. 모든 웨이브렛 계수들의 기대 R-D 경사를 계산하여, 그 값이 현재 R-D 경사 임계값보다 크면, 그 계수를 부호화 하고, 아니면, 임계값을 조금씩 줄여가면서 다음 임계값으로 부호화 결정 여부를 넘기는 방식이다.

이러한 기대 R-D 경사는 주변 계수의 상태에 따른 중요성 확률을 추정하고, 그에 따른 왜곡 감소와 비트율 증가의 비로 계산된다. 즉, i번째 부호화 단계에서, 평균 왜곡 감소량을 $E[AD_i]$ 라고 하고, 평균 비트율

증가량을 $E[\Delta R_i]$ 라고 하면, 기대 R-D 경사는 다음과 같이 구해진다[4].

$$\lambda_i = \frac{E[\Delta D_i]}{E[\Delta R_i]} \quad (1)$$

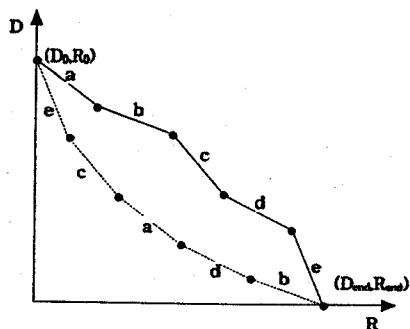


그림 2. 비트율-왜곡 최적화의 예시

IV. R-D 최적화 SPIHT 부호화

EZW나 SPIHT과 같은 대표적인 웨이브렛 기반 정지영상 부호화 알고리즘은 높은 압축효율을 지니면서, 점진적 전송에 유리한 임베디드 특성도 지니는 장점을 가지고 있지만, 비트율-왜곡 이론의 관점에서는 주어진 비트율에서 최적의 왜곡 성능을 제공하지는 못한다. 이러한 문제점을 보완하기 위해 본 논문에서는 SPIHT 부호화 알고리즘의 LIP 정렬 과정과 세분화 과정에 RDE 기법을 결합하였다.

먼저, $M_{n,b}$ 는 부호화 이후의 구간, $M_{n,b}$ 는 부호화 이전의 구간이고, n_k 는 부호화 이후의 복원값, r_k 는 부호화 이전의 복원값이며, $p(x)$ 는 구간 내에서의 부호의 확률분포일 때, i 번째 부호화 단계에서 평균 왜곡 감소량은 다음과 같이 계산된다.

$$E[\Delta D] = \sum_{k=0}^{K-1} \int_{M_{k,b}}^{M_{k+1,b}} [(x-r_k)^2 - (x-r_{k+1})^2] p(x) dx \quad (2)$$

그리고, 평균 부호율은 다음과 같이 계산된다.

$$E[\Delta R] = \sum_{k=0}^{K-1} -n_k \log_2 n_k \quad (3)$$

여기서, 부호화 이후의 구간별 확률은 다음과 같이 정의된다.

$$n_k = \int_{M_{k,b}}^{M_{k+1,b}} p(x) dx \quad (4)$$

LIP 정렬 과정에서는 해당 계수의 중요성 여부가 주

변의 중요성 여부에 영향을 받을 확률이 높지만, 세분화 과정은 해당 계수의 중요성 여부가 불규칙적이라는 특성을 이해하면, 두 과정의 R-D 경사 계산이 동일하지 않음을 알 수 있다. 먼저 LIP 정렬과정에 사용되는 평균 왜곡 감소값과 평균 부호율 및 기대 R-D 경사는 T_n 가 i 번째 부호화 단계에서의 양자화 계단 크기일 때, 다음과 같이 도출될 수 있다 [4].

$$E[\Delta D] = p 2.25 T_n^2 \quad (5)$$

$$E[\Delta R] = p_i + H(p) \quad (6)$$

$$\lambda_{i,avg} = \frac{E[\Delta D]}{E[\Delta R]} = \frac{2.25 T_n^2}{1 + \frac{H(p)}{p_i}} \quad (7)$$

여기서 p 의 확률이 D 일 때, 이전 부호의 엔트로피 $H(p)$ 는 다음과 같이 정의된다.

$$H(p) = -p \log_2 p - (1-p) \log_2 (1-p) \quad (8)$$

다음으로, 세분화 과정에 사용되는 평균 왜곡 감소값과 평균 부호율 및 기대 R-D 경사는 다음과 같이 도출될 수 있다 [4].

$$E[\Delta D] = 0.25 T_n^2 \quad (9)$$

$$E[\Delta R] = 1 \quad (10)$$

$$\lambda_{i,avg} = \frac{E[\Delta D]}{E[\Delta R]} = 0.25 T_n \quad (11)$$

본 논문에서 제안한 RDO-SPIHT(Rate-Distortion Optimized SPIHT)의 기본적인 부호화 과정은 다음과 같다. 초기화 과정에서는 SPIHT 초기화 과정과 동일한 초기 리스트 상태를 지니고, 초기 R-D 경사 임계값을 $\gamma_0 = T_0^2/16$ 으로 설정한다. 그리고, R-D 경사 임계값의 갱신이 일어날 때마다, LIP와 LSP에 있는 계수들을 대상으로 기대 R-D 경사를 계산하여 저장해 둔다. 다음으로 LIP 정렬 과정에서는 LIP중에서 R-D 경사 임계값 γ 보다 큰 R-D 경사 $\lambda_{i,avg}$ 를 지니는 계수를 부호화하고, 보다 작은 R-D 경사를 지니는 계수는 건너뛴다. LIS 정렬 과정은 현재 비트 평면에 대한 부호화가 끝날 때마다 한번씩 수행되어, Tree 분할을 수행하고, LIP 및 LSP에 좌표들을 추가하게 된다. LSP 정렬 과정에서는 LSP중에서 R-D 경사 임계값 γ 보다 큰 R-D 경사 $\lambda_{i,avg}$ 를 지니는 계수를 부호화한다. 이런 식으로 3개의 리스트에 있는 모든 좌표를 스캔한 후에 R-D 경사 임계값은 일정 비율(여기서는 1.25 사용)로 감소되어, 다시 각 리스트에 있는 좌표를 대상으로 기대 R-D 경사가 계산되고, LIP 정렬 과정 부터 반복된다.

SPIHT 부호화 알고리즘은 비트 평면 단위로 부호화 이루어지지만, RDOSPIHT에서는 R-D 경사에 따라 비트 평면이 부호화 될지 미리 알 수 없으므로, 때마다 현재 몇 번째 비트 평면을 부호화하고 있는 지 나타내는 지도(map)가 부가적으로 필요하다. 또 각 리스트에 있는 계수들의 확률 모델은 베이시안(Bayesian) 추정을 사용하고 확률 추정의 복잡도를 낮기 위해 QM-coder의 상태 전이표를 사용하였다.

V. 실험 및 결과

본 논문에서는 JPEG 2000에서 손실 압축의 표준 필드 사용되는 9/7 Daubechies 필터를 사용하여 512 크기의 영상들을 5 단계로 분해한 웨이블릿 계수를 대상으로 실험을 수행하였다. 복원된 영상의 품질을 측정하기 위해, PSNR(Peak Signal-to-Noise Ratio)을 사용하였다.

1은 본 논문에서 제안한 방법을 512x512 크기의 Lena, Barbara, Goldhill 영상에 적용하여 얻은 실험 결과이다. 실험에 쓰인 압축율은 64:1(0.125 bpp), 32:1(0.25 bpp), 16:1(0.5 bpp), 8:1(1.0 bpp)이다. SPIHT와 RDOSPIHT 모두 임베디드 부호화기이므로, 확률은 정확한 비트율에서 정지될 수 있다. 표 1에 나타낸 것처럼, 본 논문에서 제안한 방법은 Lena나 Goldhill 처럼 복잡도가 낮은 영상보다는 Barbara 처럼 복잡도가 높은 영상에서 높은 성능을 보여줄 수 있다. 이는 복잡도가 높은 영상일수록, 비트율-왜곡의 현상이 자주 일어나는데, 제안된 방법은 이 부분에서 선택했기 때문이다.

VI. 결론

본 논문에서 제안한 SPIHT와 같은 알고리즘은 높은 압축율과 비트율 특성 때문에 웨이블릿 기반 정지영상이나 동영상 부호화에 많이 쓰였다. 하지만, 이들 방식은 비트율-왜곡의 관점에서는 최적의 방법이 아니다. 반면에, 비트율-왜곡의 관점에서는 최적의 효율을 지니지만, 임계값을 갱신할 때마다, 전체 웨이블릿 계수를 검색해야 하는 복잡도 상의 문제점을 지니고 있다. 본 논문에서는 이러한 두 가지 방법의 단점을 해결하기 위해, SPIHT 방법의 각 리스트에 존재하는 계수 대상으로 비트율-왜곡 최적화를 수행하는 방식을 제안하였다. 이로 인해, 복잡도는 RDE 방식보다 떨어졌고, 성능은 보다 높은 부호화 이득을 얻을 수 있었으며, 복잡도가 높은 영상에 대해서 기존의 알고리즘보다 높은 부호화 이득을 얻을 수 있었다.

표 1. 제안된 방법과 SPIHT 방식의 성능 비교

Image	Rate (bpp)	SPIHT (1) (dB)	RDO-SPIHT (2) (dB)	(2)-(1) (dB)
Lena	0.125	31.10	31.22	0.12
	0.25	34.14	34.20	0.06
	0.5	37.25	37.21	-0.04
	1.0	40.45	40.35	-0.10
Goldhill	0.125	28.92	29.16	0.24
	0.25	31.26	31.37	0.11
	0.5	34.23	34.00	-0.23
	1.0	38.13	38.04	-0.09
Barbara	0.125	24.86	25.03	0.17
	0.25	27.58	27.86	0.28
	0.5	31.40	31.76	0.36
	1.0	36.41	36.74	0.33

감사의 글

본 연구는 광주과학기술원(K-JIST) 초고속망네트워킹연구센터(UFON)를 통한 한국과학재단 우수연구센터(ERC)와 교육부 두뇌한국21(BK21) 정보기술사업단의 지원에 의한 것입니다.

참고 문헌

- [1] S. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," IEEE Trans. Pattern Anal. Mach. Intell., vol. 11, no. 7, pp. 674-693, July 1989.
- [2] J. M. Shapiro, "Embedded image coding using zero tree of wavelet coefficient," IEEE Trans. on Signal Processing, vol. 41, no. 12, pp. 3445-3462, December 1993.
- [3] A. Said and W. A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," IEEE Trans. on Circuits and Systems for Video Tech., vol. 6, no. 3, pp. 243-250, June 1996.
- [4] J. Li and S. Lei, "An embedded still image coder with rate-distortion optimization," IEEE Trans. on Image Processing, vol. 8, no. 7, pp. 913-924, July 1999.