

Shuffling 및 상관도를 이용한 웨이블릿 영상 압축

김승중^o, 민병석, 정제창
한양대학교 전자통신공학과

Wavelet Transform Image Compression Using Shuffling and Correlation

Seungjong Kim, Byongseok Min and Jechang Jeong

Department of Communications Engineering, Hanyang University

E-mail : ksaj@icsp2.hanyang.ac.kr

Abstract

In this paper, we propose wavelet transform image compression method such that an image is decomposed into multiresolutions using biorthogonal wavelet transform with linear phase response property and decomposed subbands are classified by maximum classification gain. The classified data is quantized by allocating bits in accordance with classified class informations within subbands through arbitrary set bit allocation algorithm. And then, quantized data in each subband are entropy coded.

The proposed coding method is that the quantized data perform shuffling before entropy coding in order to remove sign bit plane. And the context is assigned by maximum correlation direction for bit plane coding.

1. 서 론

웨이블릿 변환은 입력 신호를 다중 해상도로 분해하여 서로 다른 해상도로 표현하기 위한 부 신호들과 나머지 신호들을 표현하기 위한 세부 신호들로 나누어진다. 웨이블릿 변환된 영상 신호는 서로 다른 주파수 특성을 가지며, 이들간에는 어느 정도의 상관 관계가 존재한다. 신호를 다중 해상도로 분해하여 각각의 대역에 맞는 부호화 방법을 적용할 수 있으므로 기존의 이산여현변환을 이용한 방법보다 압축률을 높일 수 있고 각 대역의 특성을 충분히 부호화에 반영할 수 있으므로 블록킹 현상이 나타나지 않아 인간의 시각에 보다 적절한 압축 방법이라고 할 수 있다.

웨이블릿 변환된 영상의 부밴드는 국부적으로 비정상(nonstationary)적인 특성을 가지며, 이러한 특성을 충분히 이용하기 위해서는 계수 분할(classification) 기법이 요구된다. 한편, 웨이블릿 변환된 영상에서 저주파 성분을 갖는 LL 밴드는 원 영상과 유사한 통계적 특성을 가지는 반면 고주파 성분을 갖는 밴드들은 밴드내 상관성이 작다. 따라서, LL 밴드는 변환이나 예측(prediction)을 통해 상관성을 제거하고 고주파 성분을 갖는 밴드들은 비기억 신호원(memoryless source)으로 모델링하면 효율적인 압축 성능을 얻을 수 있다. 그러나, 분해된 영상의 에너지 분포를 보면 원 영상의 에지 성분에 일치하는 영역 주변에 상당히 집중되어 있음을 알 수 있다. 이와 같은 비정상적 특성을 이용하는 것은 영상 압축에 있어서 중요한 요소이다. Chen[1] 등은 영상의 DCT 변환에 있어서 공간 적응적 양자화(spatially adaptive quantization) 기법을 제안하였으며, 이는 DCT 블록들을 AC 에너지에 따라 분할하여 적응적 양자화를 수행하는 방법이다. Woods[2] 등은 부밴드 부호화에 분할 방법을 적용하였고 Joshi[3] 등은 분할

이득(classification gain)이라는 관점에서 분할로 인해 얻을 수 있는 이점들을 제시하였다.

본 논문에서는 최대 분할 이득(maximum classification gain)을 이용해 웨이블릿 변환 계수들을 부 밴드별로 분할하여 밴드내의 클래스 정보를 이용, 각각의 클래스 집합에 최적 비트 할당을 통해 양자화 한 후, 엔트로피 부호화하는 방법을 제안한다. 엔트로피 부호화는 내용기반(context based) bit-plane 부호화 방법을 이용하며, 각각의 bit-plane에 대하여 부 밴드별, 클래스별 상관계수를 구하고 가장 큰 상관계수 값을 갖는 방향으로 context를 부여하여 bit-plane 부호화 시 이용한다. 또한 context 기반 bit-plane 부호화 시 MSB 쪽의 bit-plane 데이터는 데이터들 간에 상관도가 높지만 LSB 및 sign bit-plane은 랜덤한 특성을 갖는다. 이러한 특성은 context 기반 부호화 시, 부호화 성능을 저해하는 요인이 된다. 따라서, 본 논문에서는 bit-plane 부호화에 앞서 데이터들을 shuffling한 후, 엔트로피 부호화를 수행하는 방법을 제안한다. 본 논문에서 제안한 방법은 bit-plane 부호화 시 bit-plane 간의 상관관계를 이용한 방법으로서 밴드별로 같은 context를 부여하는 기존의 방법 보다 성능면에서 우수함을 입증하고자 한다.

2. 계수 분할 방법

계수 분할을 통한 양자화 기법에서는 분할 이득이 최대가 되도록 계수들을 분할하게 되는데, 웨이블릿 변환된 계수들의 분산을 이용한 분할 이득은 비정상 신호원에 대해서 식(1)과 같이 표현된다.

$$G_c = \frac{\epsilon_x^2 \sigma_x^2 2^{-2R_r}}{\prod_{j=1}^J (\epsilon_j^2; \sigma_j^2)^{b_j} 2^{-2R}} \quad (1)$$

식(1)에서 J 는 분할한 클래스 수, G_c 는 분할 이득, 분할되는 신호원 X 와 X_i ($i \leq J$)에 대해 평균을 0,

$\epsilon_x^2 = \epsilon_i^2$; 으로 가정하여 이득을 구하는 방법이다. 이와 같은 조건하에 블록들을 J 개의 클래스로 분할한다고 가정하면 분할 이득은 식(2)와 같이 간략화되고 계수 분할은 식(2)가 최대가 되도록 하는 문제로 귀착된다. 따라서, 최대 분할 이득(maximum classification gain)을 위한 분할 문제는 식(3)과 같이 모델링 할 수 있다.

$$\frac{\sigma_x^2}{\prod_{j=1}^J (\sigma_j^2)^{b_j}} \quad (2)$$

$$\min_{N_1, N_2, \dots, N_J} \prod_{j=1}^J (\sigma_j^2)^{b_j} \quad (3)$$

식(3)에서 J 는 클래스 수, N 은 분할하고자 하는 전체 블록수를 의미한다. 결국 최대 분할이득은 $\sum_{i=1}^J N_i = N$ 하에 식(3)을 만족하도록 분할한다. 그림 1은 최대 분할 이득을 통해 4개의 클래스로 분할된 결과를 보여주고 있으며, 그림에서도 알 수 있듯이 영상의 국부적인 특성에 따라 효율적으로 분할됨을 볼 수 있다. 그러나, 영상의 국부적 통계 특성을 고려하여 분할했기 때문에 계수 분할 정보인 맵(map)을 전송해야 한다. 이는 영상 부호화에 있어서 오버헤드(overhead)이며, 웨이블릿 변환된 부 밴드 전체에 대해 맵 정보를 전송하는 것은 압축률에 있어서 상당히 많은 부담을 주게 된다. 이러한 문제점을 해결하기 위해서는 웨이블릿 변환된 부 밴드간의 상관성을 이용해서 맵 정보 전송에 따른 오버헤드를 줄일 수 있다. 즉, 각 방향별로 그림 2와 같이 같은 맵 정보를 적용하여, 실제로 전송해야 하는 맵 정보는 가장 상위 레벨의 맵 정보만이 전송되며, 오버헤드 정보는 무시할 만큼 매우 작다.

그림 2와 같이 방향별로 같은 맵 정보를 적용하는 방법에는 가장 낮은 레벨로부터 상위 레벨로 맵 정보를 적용하는 방법, 또는 레벨 2로부터 레벨 1, 레벨 3, 4 등으로 적용하는 방법 등이 있다. 성능면에서 입력 영상에 따라 맵 정보 적용 방법을 달리하는 것이 가장 낮은 레벨로부터 상위 레벨로 적용하는 것보다 효율적이다.



그림 1. 최대 분할 이득을 이용한 분할 결과($J = 4$)

3. 양자화

분할을 통해 계수들을 통계적 특성에 따라 분류한 다음 분할된 맵에 따라 클래스별로 적절한 비트를 할당해야 한다. 최적 비트 할당은 변환된 영상의 밴드별로 비트를 할당하는 방법이지만 이를 맵 정보에 따라 밴드내 클래스별로 변경할 경우 비트 할당에 상당한 오류를 보임을 실험으로 확인할 수 있었다. 따라서 보다 효과적인 비트 할당이 맵 정보에 따라 이루어지게 하려면, 각 클래스 집합별로 GGD(General Gaussian Distribution) 신호원을 특징짓는 모양 정보(shape parameter)를 추출하여 이에 사용하는 정규 비트율-왜곡 곡선(R-D curve)을 이용하면 된다.

본 논문에서 사용한 비트 할당 모델은 [8-9]에서와 같이 계수들을 양자화 하는 데에 있어 비트율-왜곡 성능을 예측하기 위한 통계적 모델이며, 비트율-왜곡 모델은 일반화된 가우시안 확률 밀도 함수에 따라 각기 다른 형태로 모델링되며, 이는 주어진 제한 비트율에서 예측된 MSE(Mean Square Error)를 최소화하는 모델이다. 비트 할당 모델은 높은 비트율의 경우와 낮은 비트율의 경우로 나뉘어 모델링되며, 높은 비트율일 경우 비트율-왜곡 관계는 식(4)와 같이 표현된다.

0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

그림 2. 방향별 맵 정보 이용 방법

$$d_i(R_i) = \epsilon^2 2^{-2R_i} = c_i \frac{\Delta_i^2}{12} \quad (4)$$

c_i 는 신호원의 밀도와 독립적이고 1보다 작다고 가정하고 식(4)를 미분하면 식(5)와 같이 표현된다.

$$\begin{aligned} \log(-d'_i(R_i)) &= -2R_i \log(2) + \log(2 \epsilon^2_i \ln 2) \\ &= m_1 R_i + b_1 \quad (5) \\ &= 2 \log(\Delta_i) + \log\left(\frac{c_i \ln 2}{6}\right) \\ &= m_2 \log(\Delta_i) + b_2 \end{aligned}$$

$-d'_i$ 는 음이 아닌 값을 가지며 식(4)와 식(5)는 높은 비트율에서 R_i 와 $\log(\Delta_i)$ 가 $\log(-d'_i)$ 의 선형 함수로 모델링 됨을 보여준다. 식(5)와 같은 모델은 그림 3과 같이 높은 비트율에서 정확한 반면 낮은 비트율에서는 선형방정식이 주어진 조건을 만족하기에 불충분하다. 따라서 낮은 비트율에서도 효과적으로 적용되는 모델을 구하기 위해 식(6)과 같이 선형근사식에 쌍곡선 항을 추가한다.

$$\begin{aligned} R_h &= \frac{-1}{2 \log 2} \log(-d') + \frac{1}{2 \log 2} \log(2 \epsilon^2 \ln 2) \quad (6) \\ &= m_h \log(-d') + b_h \end{aligned}$$

R_h 에 추가될 쌍곡선 항은 근사적으로 식(6)과 같은 모델이 낮은 비트율에서 0으로 접근하도록 하기 위함이며, y 축에 초점을 가지고 있는 쌍곡선 식(7)과 같이 주어진다.

$$\frac{y^2}{a^2} - \frac{x^2}{b^2} = 1 \quad (7)$$

쌍곡선의 점근선 기울기는 a/b , $-a/b$ 가 된다. 높은 비트율에서는 R_h 는 $\zeta = \log(2 \epsilon^2 \ln 2)$ 에서 x 와 만나게 되는데, 그림 4와 같이 그 점에 보상항인 R_c 를 더하면 선형 근사식은 휘어지게 되어 x 축 위에 놓이게 된다. 즉 $\log(-d') < \zeta$ 에 대해서 R_h , $\log(-d') \geq \zeta$ 에서는 $R_h + R_c$ 모델을 이용한다. 쌍곡선 점근선이 ζ 를 통과하

고 기율기 $m_c = 1/2\log 2$ 를 갖는 것과 $a = m_c b$ 에서 $(-b + \zeta, -a)$ 점을 지나는 것을 고려하면 쌍곡선은 식 (8)과 같고 이에 따른 보상항은 식(9)와 같이 표현된다.

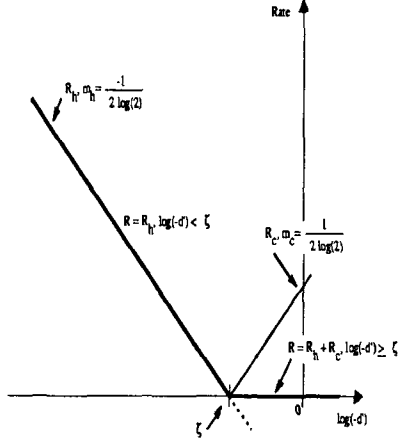


그림 3. 비트 할당 모델의 직선 근사

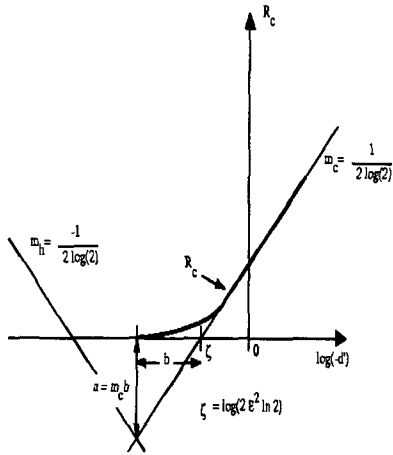


그림 4. 비트 할당 모델의 쌍곡선 보정항 근사

$$\frac{(R_c + m_c b)^p}{(m_c b)^p} - \left(\frac{\log(-d) + b - \zeta}{b}\right)^p = 1 \quad (8)$$

$$R_c = m_c b \sqrt[p]{1 + \left(\frac{\log(-d) + b - \zeta}{b}\right)^p} - m_c b \quad (9)$$

식(9)에서 b 와 p 는 모델과 실제 데이터간의 에러가 최소가 되도록 선택되게 되는데 이는 GGD 파라미터에 따른 모델링과 같은 맥락으로 간주될 수 있다.

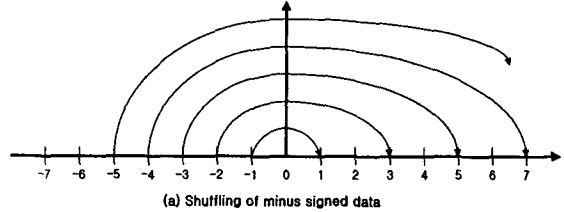
4. 엔트로피 부호화

4-1. Shuffling

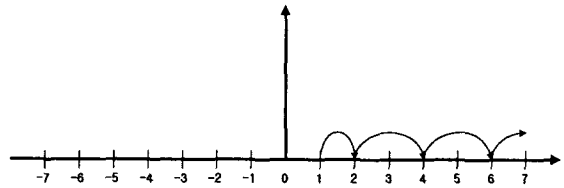
context 기반 bit-plane 부호화 시 MSB 쪽의 bit-plane 데이터는 데이터들 간에 상관도가 높지만 LSB 및 sign bit plane은 거의 랜덤한 특성을 갖는다. 이러한 특성은 context 기반 부호화 시, 부호화 성능을 저해하는 요인이

된다. 따라서, 본 논문에서는 bit-plane 부호화에 앞서 데이터들을 shuffling 한 후, 엔트로피 부호화를 수행하는 방법을 제안한다.

Shuffling 방법은 그림 5와 같이 수행된다. 즉, 음수 부분의 데이터를 양수쪽으로 shuffling 하여, 결국 sign bit plane이 발생하지 않도록 한다.



(a) Shuffling of minus signed data



(b) Shuffling of plus signed data

그림 5. Shuffling 방법

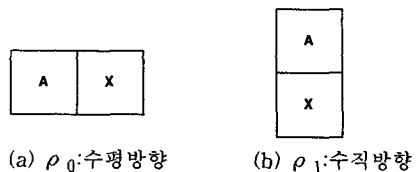
4-2. 상관도를 이용한 내용기반 bit-plane 부호화

양자화된 계수들에 대한 내용기반 bit-plane 부호화 시, 기존의 방법과 같이 같은 context를 주어 부호화 하는 방법은 웨이블릿 부 밴드의 특성을 제대로 반영하지 못한다. 따라서, 본 논문에서는 분할된 각 클래스별로 양자화된 계수들에 대한 엔트로피 부호화 시, 밴드별로 분할된 맵 정보를 이용하여 각 클래스 데이터를 방향별로 스캔한 후, bit-plane으로 분할한다. 밴드별, 클래스별로 분할된 bit-plane에 대하여 식(10)과 같이 상관도를 수직, 수평, 대각방향으로 측정하여 가장 큰 상관도를 갖는 방향으로 그림 6과 같이 context를 주어 bit-plane 적응적 산술 부호화를 수행한다.

$$\rho_k = \frac{\langle p_k^0, p_k^1 \rangle}{\|p_k^0\| \|p_k^1\|}, \quad k=0,1,2,3 \quad (10)$$

식(10)에서 분자 부분은 내적(inner product)을 의미하며, 분모는 각 벡터의 norm을 의미한다. 또한, ρ_0 는 수평방향, ρ_1 는 수직방향, ρ_2 는 대각방향(/), ρ_3 는 대각방향(\)의 상관계수를 각각 의미한다. 식(10)에 의해 구해진 각 방향별 상관계수에 대하여 식(11)과 같이 가장 큰 상관계수를 갖는 방향을 구하고 그림 6과 같이 해당하는 방향별로 context를 주어 bit-plane 부호화시 이용한다.

$$\rho = \max\{\rho_0, \rho_1, \rho_2, \rho_3\} \quad (11)$$



(a) ρ_0 :수평방향

(b) ρ_1 :수직방향

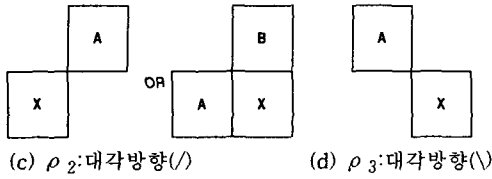


그림 6. 방향별 context

5. 실험결과 및 고찰

실험에 사용한 영상은 512×512 “Lena” 영상을 사용하였고 쌍직교 웨이블릿 변환을 위한 저대역 및 고대역 필터는 Daubechies의 9/7 탭 필터를 사용하였다.

LL 밴드 부호화는 JPEG의 무손실 압축 모드에서 사용되는 방법을 변형하여 사용하였다. 이 방법은 적절한 양자화 크기로 미리 양자화 된 LL 밴드를 DPCM 한 후, 예측 오차를 오차의 비트 크기 정보와 차이값의 정보로 나누어 각각 부호화 하는 방식이다. 여러 가지 예측 모드(prediction mode)를 사용하여 차이값의 정보가 가장 적게 발생하는 예측 모드를 선택하고 선택된 예측 모드로 LL 밴드를 부호화 하게 된다.

표 1은 최대 분할 이득을 이용하여 웨이블릿 변환된 계수들을 분할하고 분할된 각각의 클래스별로 비트 할당 후, 할당된 비트를 이용하여 양자화 테이블을 설계하여 양자화 하고 본 논문에서 제안한 shuffling 및 상관도를 고려한 context 부여 방식에 의해 부호화를 수행한 결과를 보이고 있다. 실험에서 클래스의 개수는 2개로 하였으며, 클래스 1은 내용기반 bit-plane 부호화를 수행하였고 클래스 2는 bit-plane으로 분할하지 않고 양자화된 계수들을 적응적 산술부호화를 수행하였다(방법 I). 왜냐하면, 클래스 1은 데이터들의 분산이 매우 적고 양자화된 계수값도 상당히 작으므로 적은 bit-plane으로 표현할 수 있다. 그러나, 클래스 2는 데이터들의 분산이 비교적 크고 값들 또한 크기 때문에 bit-plane 부호화 방식을 이용하면 오히려 압축율이 저하되는 경향을 보인다. 모든 클래스에 대하여 bit-plane 부호화를 적용한 방식(방법 II)과 비교하여 약 0.1dB의 화질 향상을 보였으며, 밴드별, 레벨별로 균일한 context를 부여하는 방법(방법 III)과 비교하여 1dB 이상의 화질 향상을 보였다.

표 2는 계수분할에 이용된 맵 정보를 어떻게 이용하는 지에 대한 이용 방법별 결과를 보이고 있다. 실험결과, 4레벨 dyadic 분해의 경우 레벨 3의 맵 정보를 이용하는 것이 가장 좋은 효율을 보였다.

표 1. 제안한 방법의 PSNR 결과(dB)

압축률	방법 I	방법 II	방법 III	JPEG	EZW	SFQ	SPIHT
8:1	40.04	39.98	39.09	34.79	36.28	40.52	40.46
16:1	36.93	36.87	35.95	31.55	33.17	37.34	37.21
32:1	34.01	33.90	32.95	27.92	30.23	34.33	34.11
64:1	31.06	30.96	30.19	-	27.54	-	-

본 논문에서 제안한 방법의 실험 결과 SFQ[6] 및 SPIHT[7] 보다는 압축률이 다소 떨어지지만 JPEG[5] 보다는 PSNR이 6dB 이상 우수한 결과를 나타냈으며 웨이블릿 변환을 사용하였기 때문에 DCT와 같은 블록 기반 부호화의 단점인 블록킹 현상이 나타나지 않았다. 또한 EZW(Embedded Zerotree Wavelet)[4] 보다 약 4dB 이상 우수한 결과를 얻을 수 있었으며, 특히 스칼라 양자화에 분할과 임의 클래스별 비트 할당 기법, shuffling 및 상관도를 고려한 내용기반 bit-plane 부호화를 적용하여 구현함으로써 간단하면서도 높은 PSNR을 얻을 수 있었다.

표 2. 맵 정보의 이용 방법에 따른 결과(dB)

적용 방법	압축률
레벨 1로부터 적용	33.79
레벨 2로부터 적용	33.95
레벨 3로부터 적용	34.01
레벨 4로부터 적용	33.79

6. 결 론

본 논문에서는 선형 위상 응답 특성을 가지는 쌍직교 웨이블릿 변환(Biorthogonal Wavelet Transform)을 이용하여 영상을 다중 해상도로 분해하고 분해된 부밴드들을 최대 분할 이득에 따라 분할(classification)하여 밴드내의 클래스 정보를 이용, 각각의 클래스 집합에 최적의 비트 할당을 통해 양자화 한 후, shuffling 및 각 밴드별, 클래스별 상관계수를 구하여 가장 큰 상관계수를 갖는 방향으로 context를 부여하여 bit-plane을 부호화 하는 방법을 제안하였다. 평가결과 기존의 밴드별로 양자화 하는 방법에 비해 주관적, 객관적 화질의 우수성이 입증되었다.

향후 연구과제로는 적응적 산술 부호화 시, 데이터 시퀀스에 따른 가장 최적의 윈도우의 크기를 구하는 방법 및 context의 부여 방식에 대한 연구가 필요하다.

7. 참고문헌

- [1] W. H. Chen and C. H. Smith, "Adaptive coding of monochrome and color images," IEEE Trans. Comm., vol. 25, pp.1285-1292, Nov. 1977.
- [2] J.W. Woods and T. Naveen, "A filter based bit allocation scheme for subband compression of HDTV," IEEE Trans. Image Proc., vol. IP-1, pp.436-440, July 1992.
- [3] Rajan L. Joshi et al, "Comparison of Different Methods of Classification in Subband Coding of Images," IEEE Trans. Image Proc., vol.6, no.11, pp.1473-1485, November 1997.
- [4] J. M. Shapiro, "Embedded Image Coding Using Zerotrees of Wavelet Coefficient," IEEE Trans. Signal Proc., vol.41, no.12, pp.3445-3462, December 1993.
- [5] W. M. Pennebaker, and J. L. Mitchell, JPEG STILL IMAGE DATA COMPRESSION STANDARD, Van Nostrand Reinhold, New York, Ch.11, pp.189-191, 1993.
- [6] 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.
- [7] 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 technology, vol. 6, no. 3, pp. 243-250, June 1996.
- [8] Y. Shoham and A. Gersho, "Efficient bit allocation for an arbitrary set of quantizers," IEEE Trans. Acoust., Speech, and Signal Proc., vol. ASSP-36, pp.1445-1453, September 1988.
- [9] James H. Kasner, Michael W. Marcellin, and Bobby R. Hunt, "Universal Trellis Coded Quantization," submitted to IEEE Transactions on Image Processing.