

웨이블릿 화상 코딩에서의 시각적 엔트로피 이득

박진철*, 이형극*, 이상훈*
*연세대학교 전기전자공학과
e-mail:dewofdawn@yonsei.ac.kr

Visual Entropy gain for Wavelet Image Coding

Jin-Cheol Pa가*, Hyung-Keuk Lee*, Sang-Hoon Lee*
*Dept of Electrical and Electronic Engineering, Yonsei University

요 약

웨이블릿 화상 코딩 기법은 자연적으로 레이어드된 비트스트림을 생성해 내기 때문에, 주파수 제한적인 채널 상황에서 에러에 강한 성능을 나타내고 있다. 본 논문에서는 Progressive Image Coder의 성능을 비교하고 평가하는 새로운 기법인 시각적 엔트로피를 이용해, 웨이블릿 영역에서 시각적인 가중치를 이용해 정량화하려고 한다. 이 시각적인 가중치는 주파수 영역과 공간 영역에서, 인간의 시각 체계(HVS, human visual system)에 기반 하여 만들어진 것으로, 웨이블릿 계수들의 코딩 순서를 결정하는 기준으로 사용되고, 이렇게 해서 시각적인 화질을 개선할 수가 있다. 정규화된 채널 용량이 0.3일때, 전송 이득은 시각적 엔트로피로 측정해 보았을 때 23% 이상 얻을 수 있다.

1. 서론

최근 몇년 간, EZW [1], SPIHT [2], LPPRFB [3] 그리고 EFIC [4] 등 과 같은 많은 웨이블릿 기반의 알고리즘들이 제안되어 왔다. 웨이블릿 영역에서의 공간과 주파수 영역의 정보를 동시에 해석할 수 있는 계층적 구조는 한 Subband의 정보로부터 전체적인 영상의 특성을 파악하는데 좋은 구조로 이루어져 있다. 그 뿐만 아니라, Multi-Resolution의 특성은 Progressive Image Coder의 데이터 전송에 있어서 매우 유용하게 사용된다. 이 점진적인 영상 전송의 기법에서는 왜곡에 심한 영향을 주는 중요한 웨이블릿 영역에서의 계수를 먼저 전송하게 된다.

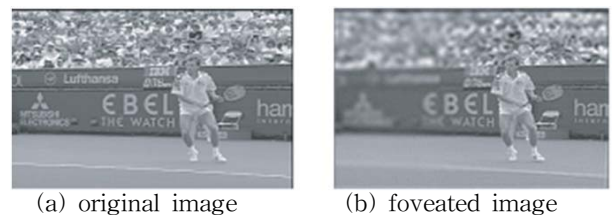
또한 사람의 망막에서는 시신경의 공간적인 분포가 비선형적이라는 특성이 있다. 즉, 포비아를 중심으로 시신경이 가장 밀집되어 있고, 이 후에는 포비아로부터 멀어질수록 급격히 감소한다. 그러므로 이 시신경으로부터 감지되는 시각 주파수 대역폭은 포비아로부터 멀어질수록 급격히 감소하게 된다. 따라서 공간 영역에서의 인간 시각 체계(HVS)의 특성을 고려하여 영상 부호 정보의 전송에서 우선권을 결정하는데 유용하게 사용할 수 있을 것이다 [4][5][6].

1) 이 연구에서 우리는 웨이블릿 영역에서 인간 시각 체계의 특성을 고려한 시각적 엔트로피를 정의하고, Progressive Image Coder에 이를 활용하여 시각적 엔트로피의 효율성을 검증한다. 그리고 새로운 화질 가중치를 주파수와 공간적인 시각적 분해능을 고려하여 웨이블릿 영역에서 다시 정의하였다. 이 가중치를 기반으로 해서 점진

적 전송의 순서를 결정하였고, 해당 시각적 엔트로피 이득을 채널 전송 용량에 의해 측정하였다. 마지막으로, 시각적 엔트로피의 새로운 관점에서 점진적 영상 부호기의 이득을 정량화 하였다.

2. 공간 영역에서 시각적 엔트로피 이득

사람의 눈은 망막에 존재하는 비선형적인 시신경의 분포를 통해서 정보를 샘플링하기 때문에, 비선형적으로 시각적인 데이터를 얻는다. (그림 1)에서 처럼 사람의 시신경에 의해서 비선형적으로 인식되는 영상을 포비에이티드 영상(foveated image)이라 한다 [6].



(그림 1) Original and foveated images over the cartesian coordinates

사람의 눈에서 Display Resolution은 망막을 중심으로 대칭적 패턴을 가지며 지수적으로 감소한다. 이와 같은 비선형적인 매핑 구조에서 얻어지는 새로운 좌표계를 Curvilinear Coordinates $\Phi(x)$ 로 정의 하겠다. 원 영상의 공간적 영역을 $S_o \subset R^2$ 으로 표기하고, Cartesian Coordinates에서 원영상에 해당하는 면적으로 A_o 로 표시 하자. 그러면 이에 대응하는 Curvilinear Coordinates에서

본 연구는 서울시 산학연 협력사업 중 (과제번호:11136) 보유기술 사업화 지원 사업의 연구 결과로 수행되었음.

의 면적은 $A_c = \int_{S_o} J_\Phi(x) dx$ 과 같이 나타낼 수 있다. 여

기서 $J_\Phi(x)$ 는 좌표변환을 나타내는 Jacobian 이다.

원영상에 대한 총 Differential 엔트로피를 $H_d^T(X) = A_o H_d(X)$ 로 Curvilinear Coordinates 상에서 Foveated Image의 총 시각적 엔트로피를 $H_d^T(\Phi) = A_c H_d(\Phi)$ 로 표현하면 $H_d(X) = H_d(\Phi)$, $A_o \geq A_c$ 이기 때문에 Foveation을 사용하여 얻어지는 정규화 이득은 다음과 같이 표현 된다. $G^m = \frac{A_o - A_c}{A_o}$

3. 웨이블릿 영역에서의 시각적 엔트로피 이득

주어진 시각 가중치 w_m^t 에 대해서, 시각적 엔트로피는 다음과 같이 표현 할 수 있다.

$$H_d^w(a[m]) = H_d^w(b[m]) = w_m^t H_d(a[m]), \quad (3.1)$$

$a[m]$ 과 $b[m]$ 은 각각 원영상과 Foveated Image의 웨이블릿 계수이다. 여기서 w_m^t 은 두 시각적 요소로 구성되어 있다. 하나는 공간적 영역에 관한 것이고 다른 하나는 주파수 영역에 관한 것이다. f_m 을 [5]에서 정의한 것처럼 웨이블릿 영역에서의 Local Frequency라고 정의한다면, 이는 다음과 같이 표현할 수 있다.

$$f_m = \min(f_c, f_d) \text{ (cycles/degree)}. \quad (3.2)$$

여기서 m 은 웨이블릿 계수 $a[m]$ 의 지표이며, f_c 는 임계 주파수, f_d 는 Display Nyquist 주파수이다. 2-D 공간적 영역에서는 정규화된 f_m 의 제곱값이 공간 영역에서의 가중치가 된다.

$$w_m^s = \left(\frac{f_m}{\max(f_m)} \right)^2 \quad (3.3)$$

공간영역에서의 가중치 f_m 과 더불어, 주파수 상에서의 가중치는 각 웨이블릿의 Subband에 의해 결정된다. 참고 문헌 [7]에서는 시각적으로 감지 가능한 노이즈 임계값을 실험을 통하여 구하였다. 여기에서 구한 해당 Subband에서의 Error Sensitivity S_w 는 정규화 함으로써 주파수 영역에서의 가중치가 되고 다음과 같이 표현된다.

$$w_m^f = \frac{S_w}{\max(S_w)} \quad (3.4)$$

식 (3.3)와 (3.4)로부터 공간 및 주파수 영역의 민감도를 종합 하여고려한 정규화된 가중치는 다음과 같다.

$$w_m^t = w_m^s \cdot w_m^f \quad (3.5)$$

4. 시각적 엔트로피의 전송

채널용량이 제한되어 있을 때 시각적 전송량을 최대화 하기위해서, 상대적으로 높은 시각 정보를 담고 있는 계수 값 들을 먼저 보내는 것이 필수적일 것이다. 한 Bit에 담

겨있는 시각 정보는 주파수와 공간 영역에서 특성을 고려 하여 설계되어 있는 가중치에 의해 평가할 수가 있다. 그러므로 웨이블릿 계수들의 전송 순위는 시각적 가중치의 크기에 의해 결정된다. 주어진 채널 용량 C를 가지고 전달할 수 있는 계수의 총 개수를 M을 이용하면 다음처럼 표현 할 수 있다

$$\sum_{m=0}^{M-1} H_d(a[m]) = C \quad (4.1)$$

이러한 Random Variable은 시각 가중치에 따라서 다시 재배치할 수가 있다. 지표 k를 w_m^t 의 크기를 기준으로 해서 재배치한 변수들의 순서라고 하자. 그렇다면, 이 때에 전송할 수 있는 시각적 엔트로피는 다음과 같이 구해진다.

$$\sum_{k=0}^{K-1} H_d^w(a[k]) = \sum_{k=0}^{K-1} w_k^t H_d(a[k]) = C_w \quad (4.2)$$

여기서 C_w 는 주어진 채널용량에 전송된 시각적 엔트로피의 합을 나타낸다. 만약 w_m^t 를 기준 가중치로서 사용한다면, w_m^t 에 의해 상대적인 시각적 엔트로피 이득은 다음과 같은 방법으로 구할 수 있다.

$$G_t = \frac{1}{C_w^T} \left(\sum_{k=0}^{K-1} H_d^w(a[k]) - \sum_{m=0}^{M-1} H_d^w(a[m]) \right), 0 \leq G_t \leq 1 \quad (4.3)$$

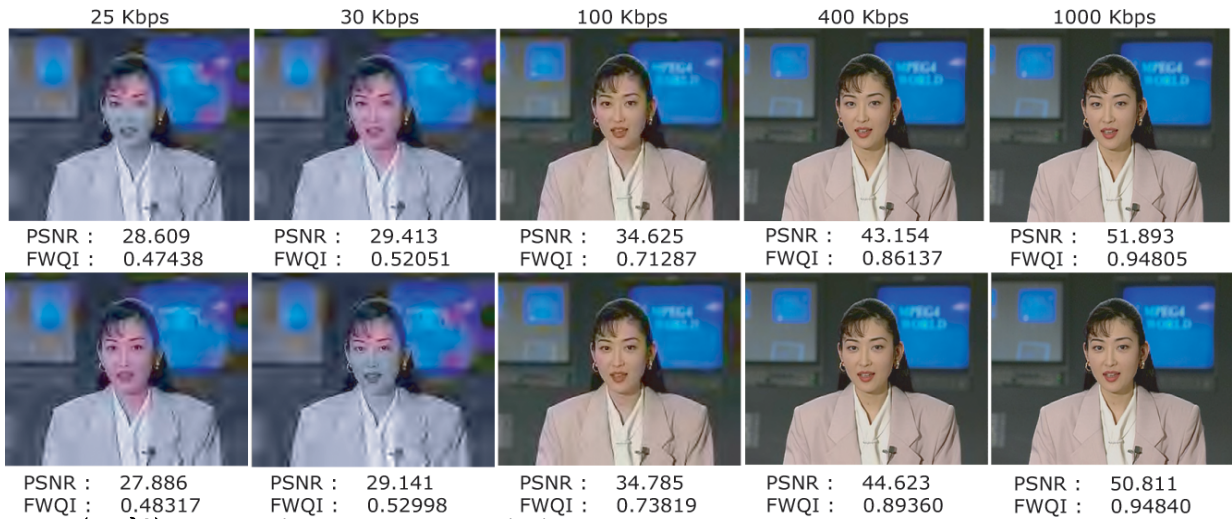
여기서 $\sum_{k=0}^{K-1} H_d(a[k]) = \sum_{m=0}^{M-1} H_d(a[m]) = C$ 를 만족한다.

G_t 는 가중치의 차이에 의해 얻을 수 있는 시각적 엔트로피의 이득이고, C_w^T 는 웨이블릿 계수들의 엔트로피에 시각적 가중치를 고려하여 계산한 전체 시각적 엔트로피이다: $C_w^T = \sum_{m=0}^{M^T-1} H_d^w(a[m])$ 여기서 M^T 는 총 웨이블릿 계수들의 개수이다.

4. 실험결과

Subband의 데이터분포는 각 웨이블릿의 Decomposition 레벨에 따라 다르며, Laplacian 분포의 형태를 갖는다. 그리고 각 Subband 마다 서로 다른 분산을 가지고 있다. 이와 같은 Laplacian 분포를 이용하여 시각적 가중치를 식 (4.1)와 같이 $w_m = w_m^t$ 로 갖는 시각적 엔트로피를 웨이블릿 계수들의 전송 순위를 결정하는데 기준으로 사용하였다.

식 (3.5)에서 정의한 가중치의 유효성을 판단하기 위해서, 참고문헌 [2]의 SPIHT과 본 연구에서 제안한 VWIC 부호화 방법을 수행, 비교하였다. 복원한 영상의 화질은 Target Bit Rate에 따라 측정하였다. 화질 측정 방법으로는 PSNR과 참고문헌 [4]에 정의된 FWQI를 이용하였다. (그림2)는 낮은 비트율에서 VWIC의 알고리즘으로 복원한 영상 화질이 SPIHT의 방법에 비해서 뚜렷하게 우수함을



(그림2) Progressive Image Transmission

Using an SPIHT (above) and a VWIC (below)

보여준다. 비트율이 증가함으로써, 전송되는 웨이블릿 계수들의 양이 증가하므로, 복원한 영상의 화질은 점차로 비슷해진다. 또한 여기서 주목할 것은 PSNR은 일관된 화질 평가하기에 적당하지 않다는 것이다. FWQI의 화질 평가 방식을 이용하여 평가한 결과가, VWIC을 이용해서 얻어진 화질의 우수함과 더욱 잘 들어맞는 것을 확인할 수가 있다.

5. 결론

본 연구에서는 시각적 엔트로피를 주파수와 공간 영역에서의 시각적 가중치를 고려하여 웨이블릿 영역에서 정의하였다. 공간적 영역에서는 사람 눈의 시각축을 기준으로 하여 정규화된 Local Bandwidth를 적용하였고, 주파수 영역에서는 각 웨이블릿 영역에서 Subband의 Error Sensitivity를 구하여, 이 두 가지의 요소의 정규화된 곱을 웨이블릿 영역에서의 시각적 가중치로 정의하였다. 이 가중치를 이용하여 논문에서 정의한 VWIC 기법을 수행하였으며, 이를 통해 기존 웨이블릿 영상 부호화 방식과 비교하였을 때, 시각적으로 더 좋은 화질을 얻을 수 있었다. 이 시각적 엔트로피는 주어진 채널 전송 용량 아래에서, VWIC에 의해 얻어지는 시각적 효과를 정량화하는데 활용하였으며, 또한 전송 시의 우선순위는 가중치를 곱한 웨이블릿 계수들의 크기로 결정하였다. 총 전송된 시각적 엔트로피는 시각 엔트로피 정의에 의해 측정되었는데, 주관적인 화질 평가에 근거해, 채널 용량이 0.1에서 0.45가 되는 Low Channel Capacity 영역에서 0.2 이상의 상대적으로 높은 시각적인 효과를 얻을 수가 있었다.

참고문헌

- [1] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. Signal Processing*, vol. 41, pp. 3445-3462, Dec. 1993
- [2] 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
- [3] T. D. Tran and T. Q. Nguyen, "A progressive transmission image coder using linear phase uniform filterbanks as block transforms," *IEEE Trans., Image Processing*, vol. 8, pp. 1493-1507, Nov. 1996
- [4] Z. Wang and A.C. Bovik, "Embedded foveation image coding," *IEEE Trans. Image Processing*, vol. 10, pp. 1397-1410, Oct. 2001
- [5] S. Lee, M. S. Pattichis and A. C. Bovik, "Foveated video compression with optimal rate control," *IEEE Trans. Image Processing*, vol. 10, pp.977-992, Jul. 2001.
- [6] S. Lee, M. S. Pattichis, and A. C. Bovik, "Foveated video quality assessment," *IEEE Trans. Multimedia*, vol. 4, pp.129-132, Mar. 2002
- [7] A. B. Watson, G. Y. Yang, J. A. Solomon, and J. Villasenor, "Visibility of wavelet quantization noise," *IEEE Trans. Image Processing*, vol.6, pp.1164-1175, Aug. 1997