

Human Visual System에 적응적인 영상 화질 향상 기법

*방성배 **김원하

경희대학교

*sungbae9023@hanmail.net

Human Visual System Adaptive Image Enhancement

*Bang, Seang-Bae **kim, won-ha

Kyung Hee University

요약

본 논문은 human visual system(HVS)에 따른 주파수 민감도와 공간에서 다양한 특성들을 구현하기 위한 신호처리 방법을 개발하였다. 인간의 눈은 주파수 성분에 따라 민감도가 다르며 초점에서 멀수록 인지 가능한 해상도가 떨어진다. 주파수 민감도를 구현하기 위해서 본 논문은 영상 신호의 에너지 스펙트럼 모양이 contrast sensitivity function(CSF)의 모양이 되도록 하여 영상 신호의 에너지를 증가시켰으며 신호 방향에 적응적인 multiband energy scaling 방법을 개발하였다. 기존의 시스템에서 능률만을 향상시키는 기존의 분석 모델과 비교하면 개발한 방법은 HVS에 좀 더 적절하고 선호되게 영상 신호를 처리할 수 있다.

1. 서론

인간의 눈은 시각 정보를 망막에 분포되어 있는 광수용기를 통해 인식한다. 광수용기의 생리적 특성에 의해 human visual system(HVS)의 감각적 민감도는 신호의 주파수 성분에 따라 다르게 나타나며 중간 주파수에서 가장 민감하게 반응한다[1]. 또한 광수용기의 밀도는 눈의 초점인 망막의 중심에서 가장 높으며, 망막의 중심에서 멀어질수록 밀도가 급격히 감소한다. 그러므로 human visual system(HVS)이 인식할 수 있는 지역적 시각 정보량은 초점으로부터 멀어질수록 감소한다.[2] 이러한 특성들 때문에 HVS의 감각적 민감도는 위치와 주파수 성분에 따라 다양하게 나타난다. 인간은 눈의 감각적 민감도에 의해 처리된 신호를 영상 신호로 인식한다.

광수용기의 생리적 주파수 민감도는 contrast sensitivity function(CSF)에 의해 모델화 되어 있다[3]. 영상 신호의 주파수 성분은 낮은, 중간, 높은 주파수로 구성되어 있다[4]. 낮은 주파수 성분들은 영상의 가시성에 해당하는 global contrast와 전체적 밝기를 결정한다. 중간 주파수 성분들은 영상의 local contrast를 결정하는 edge와 세밀한 texture를 나타낸다. 사람은 local contrast가 선명하게 나타나는 영상일수록 선호한다. 높은 주파수 성분들은 noise와 관련되어 있으며 human visual system에 의해 잘 인지할 수 없다[5].

기존의 시스템의 효율을 향상시키기 위해 CSF 모델을 사용해야 하지만 비디오 코딩 시스템과 같은 경우에[6] HVS의 주파수에 대한 다양한 인지 특성들이 아직 충분히 연구되어지거나, 개발되어지지 않아 사용하기가 어렵다. 효율만을 향상시키는 기존 시스템들의 분석 모델과 비교하면 이러한 모델의 구현은 HVS에 더욱 적절하고 선호되게 영상 신호를 처리할 수 있고 다양한 분야에서 널리 확대되어 사용될 수 있으며 이러한 필요성을 만족시키기 위해서 본 논문은 인간 인식 스펙트럼을 구현하기 위한 신호처리 방법을 개발하였다.

DCT는 주파수 영역에서 작동할 뿐 아니라 공간적으로 국부적 처리가 가능하기 때문에 본 논문은 DCT영역에서 특성들을 구현하였다. 신호의 에너지를 증가시켰을 때 blocking artifact와 ringing artifact를 발생시키지 않기 위하여 우리는 신호의 방향에 적응적인 multiband energy scaling 방법을 고안하였다. 그리고 본 논문은 energy scaling factor의 모양을 결정하여 CSF에 따라 신호 에너지의 스펙트럼의 형태를 만들었다.

2. Multiband energy scaling 방법

DCT 영역에서 같은 주파수를 가지는 계수들의 밴드는 다이아몬드 모양을 이루고 있다. Ω_n 을 n번째 주파수 밴드의 DCT 계수라고 하면 아래와 같이 Ω_n 을 표현할 수 있다.

$$\Omega_n = [F(0, N), \dots, F(u, n-u), \dots, F(n, 0)] \quad (1)$$

여기서 $l(\Omega_n)$ 을 Ω_n 에서 DCT 계수의 개수라고 하면 식은 아래와 같다.

$$l(\Omega_n) = \begin{cases} n+1 & \text{for } 0 \leq n \leq N-1 \\ 2N-n-1 & \text{for } N \leq n \leq 2N-2 \end{cases} \quad (2)$$

DCT 영역에서 신호의 에너지를 조절하는 방법으로 Multiband Energy Scaling방법이 있다[7],[8]. 이 방법은 n번째 주파수 밴드에 대한 local contrast를 n번째 주파수 밴드의 에너지와 그 이전의 주파수 밴드들의 평균 에너지의 비율로 계산할 수 있는 것을 이용한다. 즉, local contrast는 아래와 같이 측정할 수 있다.

$$\frac{\|\Omega_n\|^2 / l(\Omega_n)}{\sum_{k=0}^{n-1} \|\Omega_k\|^2 / \sum_{k=0}^{n-1} l(\Omega_k)} \quad \text{for } n \geq 1 \quad (3)$$

여기서 $\|\cdot\|$ 는 L^2 -norm을 의미한다.

$\bar{F}(u,v)$ 을 향상된 영상의 DCT 계수라고 하며 n번째 주파수 밴드들의 이러한 DCT 계수들의 집합을 $\bar{\Omega}_n$ ($n = u + v$)라고 하자. 그러면 향상된 local contrast와 원본 local contrast의 관계는 아래와 같다.

$$\frac{\|\bar{\Omega}\|^2/l(\bar{\Omega}_n)}{\sum_{k=0}^{n-1} \|\bar{\Omega}\|^2/\sum_{k=0}^{n-1} l(\bar{\Omega}_k)} = \lambda \cdot \frac{\|\Omega_n\|^2/l(\Omega_n)}{\sum_{k=0}^{n-1} \|\Omega\|^2/\sum_{k=0}^{n-1} l(\Omega_k)} \quad (4)$$

여기서 λ 는 energy scaling factor이며 λ 가 1보다 클수록 local contrast가 향상된다. 위의 식을 간단히 정리하면 다음과 같이 될 수 있다.

$$\|\bar{\Omega}_n\|^2 = \lambda \cdot R_n \cdot \|\Omega_n\|^2 \quad (5)$$

여기서 $R_n = \frac{\sum_{k=0}^{n-1} \|\bar{\Omega}_k\|^2}{\sum_{k=0}^{n-1} \|\Omega_k\|^2}$ 이다. R_n 은 향상된 DCT 계수들의 에너지 합과 원본 DCT 계수들의 에너지 합의 비율이다. [8]에 따라 향상된 local contrast 영상의 DCT 계수는 원본 DCT 계수들로부터 아래와 같이 유도될 수 있다.

$$\bar{F}(u,v) = \sqrt{\lambda} \cdot \sqrt{R_n} \cdot F(u,v) \text{ for } n = u + v \quad (6)$$

낮은 주파수 성분은 블록의 전체적인 밝기에 영향을 주므로 낮은 주파수 성분에 갑작스러운 변화를 주면 block artifact라고 불리는 부자연스러운 block artifact가 생기게 된다[7]. 이러한 block artifact를 줄이기 위해서 Lee는 세 번째 이하의 주파수 밴드에 대해서 에너지의 변화를 주지 않았다. 그러므로 block artifact가 없는 multiband energy scaling 방법에 대한 식은 다음과 같이 된다.

$$\bar{F}(u,v) = \begin{cases} F(u,v) & \text{for } u+v \leq \lfloor N/3 \rfloor \\ \sqrt{\lambda} \cdot \sqrt{R_n} \cdot F(u,v) & \text{otherwise} \end{cases} \quad (7)$$

기존의 방법은 에너지를 조절할 때 대각선 방향으로 조절하고 있다. 그러나 원본 신호의 방향을 고려하지 않고 에너지를 조절하면 edge 부근에서 값이 지나치게 변화하여 ringing 현상이 발생한다. ringing artifact는 원본 edge방향의 수직 성분이 강조되어 나타나는 것으로 강한 edge보다 가시성이 높다.

3. 신호방향에 적응적인 multiband energy scaling

여기서는 원본 신호의 방향에 따라 주파수 에너지 성분을 조절하여 ringing artifact를 발생시키지 않는 multiband energy scaling 방법에 대해서 소개하겠다. 이전의 연구들에 의해 block gradient는 block 내의 이미지 신호의 방향을 예측하는 데에 충분한 정확도를 가지고 있는 것이 입증되었다[9]. 그래서 본 논문은 block내의 이미지 신호의 방향을 예측하기 위하여 block gradient를 사용한다. NxN block에서 신호의 방향을 예측하기 위하여 아래와 같이 block gradient를 사용할 수 있다.

$$\vec{\nabla} = \nabla^{ver} \cdot \hat{i} + \nabla^{hor} \cdot \hat{j} \quad (8)$$

∇^{ver} , ∇^{hor} 을 각각 수직, 수평 방향의 block gradient라고 하고 \hat{i} , \hat{j} 을 각각 수직, 수평 방향의 단위벡터라고 하면 block gradient는 block을 반으로 나누어 각각의 합의 변화로부터 구할 수 있다. 즉, 수직 block gradient는 아래와 같이 구할 수 있다.

$$\begin{aligned} \nabla^{ver} &= \text{Sum of upper pixels} - \text{Sum of lower pixels} \\ &= \sum_{i=0}^{N/2-1} \sum_{j=0}^{N-1} [f(i,j) - f(N-i-1,j)] \end{aligned}$$

DCT 계산은 각 방향에 대해서 분리할 수 있으므로, DCT 계수를 수직과 수평방향으로 나눈다. 각 방향에 대한 향상된 영상 신호의 DCT 계수는

$$\begin{aligned} \bar{F}^{hor}(u,v) &= \sqrt{\lambda} \cdot \sqrt{R_v^{hor}} \cdot F(u,v) \\ \bar{F}^{ver}(u,v) &= \sqrt{\lambda} \cdot \sqrt{R_u^{ver}} \cdot F(u,v) \end{aligned} \quad (9)$$

이다. 여기서 R_v^{hor} , R_u^{ver} 는 각 방향에 대한 에너지 비율이다.

신호의 방향에 따라 에너지를 조절하기 위하여 각 방향에 대한 block gradient의 비율을 가중치로 사용한다. DCT에서는 신호의 방향과 수직으로, block gradient의 방향과는 수평으로 방향 성분들이 나타난다. 그래서 신호 방향에 적응적으로 에너지가 조절된 DCT계수는 다음과 같이 표현될 수 있다.

$$\begin{aligned} \bar{F}(u,v) &= \frac{|\nabla^{hor}|}{|\vec{\nabla}|} \cdot \bar{F}^{hor}(u,v) + \frac{|\nabla^{ver}|}{|\vec{\nabla}|} \cdot \bar{F}^{ver}(u,v) \\ &= \frac{\sqrt{\lambda}}{|\vec{\nabla}|} \cdot \{ |\nabla^{hor}| \cdot \sqrt{R_v^{hor}} \\ &\quad + |\nabla^{ver}| \cdot \sqrt{R_u^{ver}} \} \cdot F(u,v) \end{aligned} \quad (10)$$

여기서 $|\vec{\nabla}| = |\nabla^{hor}| + |\nabla^{ver}|$ 이다. 원본 신호가 완전한 수직방향의 성분만을 가졌을 때 방향에 대한 가중치는 $|\nabla^{hor}|/|\vec{\nabla}| = 1$ 이고 $|\nabla^{ver}|/|\vec{\nabla}| = 0$ 이며 수평방향의 신호에 대해서는 반대의 가중치를 가지게 된다. 만약 이미지 신호의 방향이 대각선이면 두 방향에 대한 가중치는 모두 1/2이 된다.

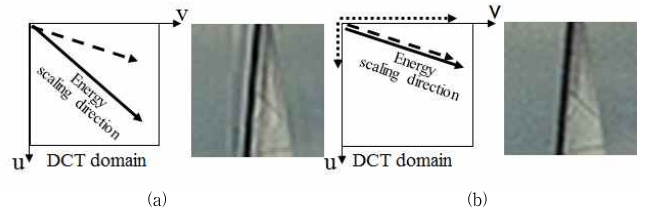


그림 1. 신호 방향에 따른 화질 향상 비교

그림 2는 기존의 multiband method과 제안한 방법으로 화질이 향상된 영상을 비교해 놓았다. 기존의 방법은 ringing artifact를 발생시키지만 제안한 방법은 신호의 방향에 따라 에너지를 증가시켰기 때문에 ringing artifact를 시각적으로 감지하기 어렵다.

4. 주파수 민감도의 구현 방법

개발한 신호의 방향을 고려하는 multiband energy scaling 방법에 더하여 본 논문은 HVS의 주파수의 민감도를 알고리즘에 구현하였다. 기존의 연구들은 주파수 민감도를 표현한 contrast sensitivity function(CSF)이 개발하였다[2]. 주파수의 민감도를 구현하기 위해서는 모든 주파수 밴드에 energy scaling factor를 상수 값으로 유지하기 보다는 CSF에 따라 영상의 스펙트럼을 조절하는 방식으로 주파수 밴드마다 energy scaling factor를 조절해야 한다.

CSF는 연속적인 공간적 주파수 f (cycle/degree)로 측정된다. CSF는 아래와 같이 표현될 수 있다.

$$CSF(f) = (0.2 + 0.45f) \cdot \exp(-0.18f) \quad (11)$$

v_d (pixel)가 시각 거리이고 d (pixel)가 거리일 때 연속적인 공간 주파수 f 을 이산 정규화 주파수 f_d 로 다음과 같은 방법으로 변환할 수 있

다[4].

$$f_d = \min [0.5, \beta \cdot f] \quad (\text{cycle/pixel}) \quad (12)$$

여기서 conversion factor β 는

$$\beta = \tan^{-1} \left(\frac{d}{v_d} + \frac{1}{2v_d} \right) - \tan^{-1} \left(\frac{d}{v_d} - \frac{1}{2v_d} \right) \quad (13)$$

이다. 거리에 따른 특성을 제한하면 $d=0$ 이 되고 $\beta = 2 \tan^{-1}(1/2v_d)$ 된다. $N \times N$ DCT에서 CSF는

$$\begin{aligned} \widehat{CSF}(k) &= CSF(f = k/2N\beta) \\ &= \left\{ 0.2 + \frac{0.1125k}{N \cdot \tan^{-1}(1/v_d)} \right\} \\ &\quad \times \exp \left\{ \frac{-0.045k}{N \cdot \tan^{-1}(1/2v_d)} \right\} \end{aligned} \quad (14)$$

이 된다.

주파수 민감도를 구현하기 위해서 energy scaling factor를 CSF에 따라 조절하였다. blocking artifact를 막기 위해서 낮은 주파수 밴드에서는 energy를 유지해야 하므로 낮은 주파수 밴드에 대해서는 energy scaling factor를 1로 정하였다. 그래서 energy scaling factor는 아래와 같이 될 수 있다.

$$\lambda_k = \begin{cases} 1 & \text{for } k < [N/3] \\ \alpha_k \cdot \widehat{CSF}(k) & \text{for } k \geq [N/3] \end{cases} \quad (15)$$

여기서 α_k 는 k번째 주파수 밴드에 대한 가중치이다.

α_k 는 사람의 시각 인지에 맞추어 값이 정해져야한다. local contrast를 향상시키기 위해 우리는 중간 주파수에 대해서 강조를 하였다. 반면에 높은 주파수에 있는 noise는 감소시켰다. 그래서 $\{\alpha_k\}$ 는 중간 주파수 대역에서 높은 값을 가지며 높은 주파수 대역에서는 작은 값을 가진다.

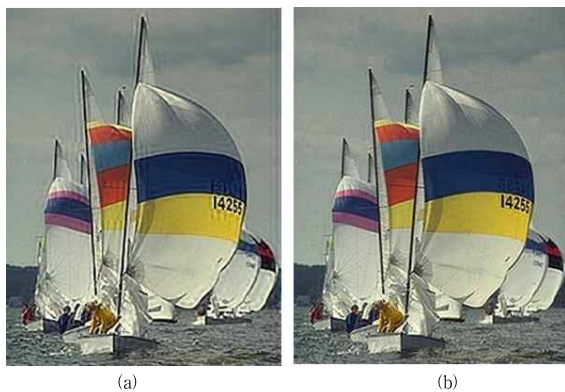


그림 1. DCT 영역에서 에너지 향상 방법 비교 (a) 기존의 에너지 향상 방법, (b) 제안하는 방법

5. 결론

본 논문은 DCT 영역에서 신호의 에너지를 조절하여 ringing artifact를 발생시키지 않는 multiband energy scaling 방법과 원본 이미지의 스펙트럼을 주파수 민감도인 CSF의 모양으로 만드는 방법을 개발하였다. 그러므로 본 논문에서 제안하는 방법은 영상의 신호의 물리적 특성에 따라, 사람 눈의 생리적 특성에 따라 영상을 향상시킬 수 있다. 그림 1.을 보면 기존의 multiband energy scaling 방법에 비해 제안하는 방법은 ringing artifact가 적고 영상 신호의 중간 주파수 대

역이 높은 주파수 대역보다 상대적으로 강조되어 noise가 강조되지 않으면서 영상신호가 향상되는 것을 볼 수 있다.

감사의 글

본 연구는 산업통상지원부의 재원으로 수행되었음 (10052967, 재난 재해 대응용 특수목적기계 통합제어시스템 개발)

참고문헌

- [1] D. H. Kelly, "Visual contrast sensitivity," *Optica Acta: International Journal of Optics*, vol. 24, no. 2, pp. 107-129, 1977.
- [2] B. Wandell, "Foundations of Vision". Sinauer Associates Inc, May 1995.
- [3] J. Mannos and D. Sakrison, "The effects of a visual fidelity criterion of the encoding of images," *IEEE Trans. Information Theory*, vol. 20, no. 4, pp. 525-536, 1974.
- [4] L. Itti, "Automatic foveation for video compression using a neurobiological model of visual attention," *IEEE Trans. Image Processing*, vol. 13, no. 10, pp. 1304 - 1318, 2004.
- [4] Z. Wei and K. Ngan, "Spatio-temporal just noticeable distortion profile for grey scale image/video in dct domain," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 19, no. 3, pp. 337-346, 2009
- [5] H. A. W. Schober and R. Hilz, "Contrast sensitivity of the human eye for square-wave gratings," *Journal of the Optical Society of America*, vol. 55, no. 9, pp. 1086-1090, 1965.
- [6] Z. Chen and C. Guillemot, "Perceptually-friendly h.264/avc video coding based on foveated just-noticeable-distortion model," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 20, no. 6, pp. 806-819, 2010.
- [7] J. Tang, E. Peli, and S. Acton, "Image enhancement using a contrast measure in the compressed domain," *IEEE Signal Process. Lett*, vol. 20, pp. 289-292, 2003.
- [8] S. Lee. "An efficient content-based image enhancement in the compressed domain using retinex theory," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 17, no. 2, p. 199-213, feb. 2007.
- [9] L.-K. Liu and E. Feig, "A block-based gradient descent search algorithm for block motion estimation in video coding" *IEEE Trans. Circuits and Systems for Video Technology*, vol. 6, no. 4, p. 419-422, aug 1996.