

특집논문 (Special Paper)

방송공학회논문지 제23권 제2호, 2018년 3월 (JBE Vol. 23, No. 2, March 2018)

<https://doi.org/10.5909/JBE.2018.23.2.206>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

인간 시각의 인지 특성을 이용한 영상 화질 향상 방법

방 성 배^{a)}, 김 원 하^{a)†}

Image Enhancement Using Human Visual Perception

Seangbae Bang^{a)} and Wonha Kim^{a)†}

요 약

본 논문은 영상 신호의 방향을 고려하여 기존의 multiband energy scaling 방법의 문제점을 보완하면서 human visual system(HVS)에 적합한 영상 local contrast 향상 방법을 개발하였다. 기존의 multiband energy scaling 방법은 신호 방향에 대한 고려 없이 화질을 향상시켜 ringing artifact가 발생하였으나 본 논문에서는 block gradient를 사용하여 신호의 방향을 측정하고 측정된 신호 방향에 따라 주파수 신호를 향상시켜 ringing artifact의 발생 없이 화질을 향상시켰다. 또한 본 논문은 human visual system(HVS)은 각 신호의 값 하나하나 보다는 각 신호가 가지는 주파수에 성분에 민감하게 반응한다는 것을 이용하여 주파수 성분에 대한 인간 시각의 민감도를 모델링한 contrast sensitivity function(CSF)에 따라 영상의 화질을 향상시켰다. 결국 본 논문에서 제안하는 방법은 신호의 특성과 인간 시각의 특성을 모두 고려하여 영상의 화질을 향상시키기 때문에 기존의 화질 향상 방법들에 비해 영상 신호와 인간 시각 특성에 더욱 적합하게 화질을 향상시킬 수 있다.

Abstract

We develop the signal processing method for adaptive implementing direction of signal and the frequency sensitivity of human visual system(HVS). Existing multiband energy scaling method makes ringing artifact because it does not consider signal direction. To solve this problem, we use block gradient for signal direction in addition to existing method. And we use the fact that frequency component of signal is more sensitive than value of signal over human eyes. we enhance the signal according to contrast sensitivity function(CSF) which is the model of frequency sensitivity of human eye. Compared that the existing analysis models only improve the efficiencies in the existing systems, the developed method can process the image signals to be more desirable and suitable to HVS.

Keyword : Contrast Sensitivity, Human Visual System(HVS), DCT, Signal Direction

a) 경희대학교 전자정보대학(College of Electronics & Information, Kyung Hee University)

† Corresponding Author : 김원하(Wonha Kim)

E-mail: wonha@khu.ac.kr

Tel: +82-2-201-2030

ORCID: <http://orcid.org/0000-0002-4797-3526>

※ 이 논문의 연구결과 중 일부는 “한국방송·미디어공학회 2017년 추계학술대회”에서 발표한 바 있음.

※ 본 연구는 산업통상자원부의 재원으로 수행되었음 (10052967, 재난재해 대응용 특수목적기계 통합제어시스템 개발).

※ This work was supported by the Korean Minister of Industry under Grant 10052967.

· Manuscript received December 26, 2017; Revised March 8, 2018; Accepted March 8, 2018.

Copyright © 2016 Korean Institute of Broadcast and Media Engineers. All rights reserved.

“This is an Open-Access article distributed under the terms of the Creative Commons BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited and not altered.”

1. 서론

UHD와 같은 초 고화질 해상도 영상에서는 이전의 해상도에서 수치적으로 크기가 작아서 나타나지 않았던 미세한 신호들이 인간의 시각에 더욱 두드러지게 인지된다. 이는 공간해상도뿐만 아니라 주파수 해상도가 모두 증가한 초고화질 해상도 영상에서의 미세한 신호를 인간 시각이 신호의 수치적 크기보다는 주파수 성분에 따라서 인지했기 때문이다. 이전 해상도 영상은 주파수 해상도가 충분하지 않아 인간 시각의 인지 특성을 고려하기 어려웠기 때문이다. 이전 해상도 영상은 주파수 해상도가 충분하지 않아 인간 시각의 인지 특성을 고려하기 어려웠기 때문에 주파수 영역에서 인간 시각의 인지 특성을 고려하지 않고 신호의 수치적 특성을 이용한 신호처리만으로도 비교적 충분한 성능을 얻을 수 있었다. 그러나 주파수 해상도가 충분히 높은 초 고화질 해상도 영상에서는 영상 신호에 대한 신호적 특성뿐만 아니라 인간 시각의 인지 특성을 고려하여 영상의 화질을 개선해야 만족스러운 성능을 얻을 수 있다.

인간은 일반적으로 민감하게 인지되는 영상 신호 성분이 많은 영상일수록 더욱 선호한다^[1]. 그러므로 인간 시각에 적합하게 영상의 화질을 향상시키기 위해서는 영상 신호의 특성에 따라서 인간의 시각이 인지하는 정도 즉, 시각 민감도를 파악하여 민감도가 높은 영상 신호의 성분을 강조해야 한다. 그리고 영상의 화질을 향상시킬 때 발생하는 *ringing artifact*와 *block artifact*를 방지하기 위해서는 영상의 신호적 특징을 고려하여야 한다. 그러나 기존의 영상 화질 향상 방법들은 영상에 대한 인간 시각의 인지 특성을 고려하지 않거나 인간 시각의 인지 특성을 고려하였지만 신호적 특성을 고려하지 않아 *ringing artifact*와 *block artifact*를 발생시켰다.

공간 영역 기법은 영상의 신호적 특성을 이용하는 방법으로 크게 영상 신호 값들의 분포를 이용하는 *histogram* 기법들과 영상 신호의 특성을 분석하여 적합한 *spatial filter*를 적용하는 기법들로 나누어진다. *Histogram* 기법들은 *pixel*들의 분포를 이용하여 영상의 화질을 향상시킨다. 그러나 지역별로 나타나는 신호의 특성을 이용하지 않기 때문에 부자연스럽게 영상을 변형시킬 수 있다. *spatial filter* 기반

기법들은 경계선의 강도, 영상 신호의 변화율, 영상 신호의 방향등의 신호 특성에 적합한 *filter*를 적용하여 영상의 화질을 향상시켰다. 그러나 공간 영역에서는 초고화질에서 나타나는 미세한 변화를 처리하기 어렵기 때문에 *spatial filter* 기반의 영상 화질 향상 기법은 UHD와 같은 초고화질에서는 만족한 만한 성능을 내지 못한다.

Multi-scaling 방법은 영상 전체를 해상도 별로 분리하여 각 해상도에 해당하는 신호의 성분을 인간 시각적 특성에 따라 조절하여 영상의 화질을 향상시킨다. Choi는 인간 시각의 *retinex* 이론을 이용하여 영상을 *illumination*과 몇 개의 *reflectance layer*들로 나누었고 영상의 *contrast*를 이용하여 각 *reflectance layer*의 신호 에너지를 조절하여 영상의 화질을 향상시켰다^[12]. 다른 방법으로는 DWT를 이용하는 방법이 있다. 이 방법은 영상 신호를 해상도별로 분리하고 해상도별로 인간 시각이 인지하는 *contrast* 정도와 일치하도록 해상도별 영상 신호의 에너지를 조절하여 영상의 화질을 향상시켰다^[13]. *Multi-scaling* 방법은 영상 신호를 전체적으로 처리해야하기 때문에 지역별 영상 신호의 특성에 따라 다르게 신호를 처리하기 어렵다. 또한 많은 메모리 사용과 높은 복잡도를 요구한다.

Multi-band 방법은 인간의 시각이 영상 신호를 주파수 영역에서 받아들이며, 영상 신호의 특성이 국부적으로 변한다는 것을 근거로 DCT 영역에서 영상 신호의 에너지를 주파수 별로 *recursive*하게 증가시켜 영상의 화질을 향상시킨다^[14]. 그러나 이 방법은 영상의 신호적 특성을 고려하지 않아 *ringing artifact*와 *block artifact*를 발생시킨다. Lee는 *block artifact*를 방지하기 위하여 DC를 포함한 일정범위의 저주파수 밴드의 주파수 에너지를 고정하였다^[15]. 그러나 이 방법은 화질 향상 효과 또한 감소시켰다.

본 논문은 인간 시각의 인지 특성과 영상의 신호적 특성을 모두 이용하여 *artifact* 없이 인간 시각에 적합하게 영상의 화질을 향상시키는 새로운 기법을 제안한다. 본 논문은 인간 시각의 인지 특성에 대한 모델을 이용하여 새로운 *multi-band* 방법을 제시하고 기존의 *multi-band* 방법에서 발생시키는 *ringing artifact*와 *block artifact*를 영상의 신호적 특성을 이용하여 화질 향상 효과를 유지하면서 감소시

켰다. 본 논문은 다음과 같이 구성되어 있다. 두 번째 장에는 새로운 multi-band 방법을 소개하고 세 번째 장에는 본 논문에서 제안하는 초고화질 영상 화질 향상 기법을 소개한다. 네 번째 장에서는 본 논문에서 제시하는 방법에 대해 주관적, 객관적으로 평가하고 다섯 번째 장에서는 제시하는 방법에 대한 결론을 내린다.

II. 인간 시각의 인지 특성

많은 심리학적, 생물학적 실험들을 통하여 영상 신호를 받아들이는 인간의 시각 세포는 영상 신호를 주파수 성분별로 인식한다는 것이 검증되어 있다. 그러므로 인간 시각 인지 특성에 적합하도록 영상 신호를 향상시키기 위해서는 주파수 영역에서 처리해야 한다. 또한, 인간의 시세포 밀도는 초점을 중심으로 지수적으로 감소하기 때문에 인간은 영상을 전체적으로 인지하기 보다는 초점을 중심으로 국부적으로 인지한다. 그러므로 국부적이면서 주파수 영역에서 영상 신호를 처리하기 위하여 DCT가 널리 사용되어오고 있다.

$N \times N$ 블록의 (i, j) 위치에 있는 픽셀을 $f(i, j)$ 라고 하고 $F(u, v)$ 를 DCT 계수라고 하면 $F(u, v)$ 는 아래의 식으로부터 구할 수 있다.

$$F(u, v) = C_u C_v \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} f(i, j) \cos \frac{(2i+1)u\pi}{2N} \cos \frac{(2j+1)v\pi}{2N}, \quad (1)$$

$$0 \leq u, v \leq N-1$$

여기서 $C_u = \begin{cases} \sqrt{1/N} & \text{for } u=0 \\ \sqrt{2/N} & \text{otherwise} \end{cases}$ $C_v = \begin{cases} \sqrt{1/N} & \text{for } v=0 \\ \sqrt{2/N} & \text{otherwise} \end{cases}$ 이다.

기존의 DCT를 사용하는 영상 향상 기법은 영상 신호의 수치적으로 측정된 local contrast를 향상시켰다. 그러나 인간 시각이 인지하는 contrast와 수치적으로 계산되는 contrast는 차이가 있다. 기존 연구에 의하면 인간 시각은 주파수 성분별로 다른 민감도를 가지고 있으며, 민감도가 높은 신호일수록 인간은 두드러지게 인지한다. 또한, 인간 시각

이 특정 주파수 성분을 인지할 때 배경 신호의 에너지 정도에 영향을 받는다.

Contrast sensitivity function (CSF)는 주파수별 인간 시각의 민감도를 나타내는 모델이다. CSF에 따르면 인간 시각은 3~5cycles/degree의 신호를 가장 민감하게 인지한다. 그리고 인간 시각의 oblique 효과에 따라 같은 주파수 성분이라도 방향에 따라 민감도가 다르다. 인간 시각은 대각선 신호 성분보다 수직, 수평 신호 성분에 더욱 민감하다. 그러므로 주파수 $F(u, v)$ 에 대한 인간 시각의 민감도 $CSF(u, v)$ 는 다음과 같다^[6].

$$CSF(u, v) = s \cdot \frac{1}{\phi_u \phi_v} \frac{\exp(cw(u, v))/(a + bw(u, v))}{r + (1-r)\cos^2(\psi(u, v))} \quad (2)$$

여기서 a, b, c 는 상수이고 w 는 spatial frequency (cycle/degree)이다. Wei는 경험적으로 $a = 1.33, b = 0.11, c = 0.18$ 로 하였다^[6]. 그리고 DCT block에서 (u, v) 번째의 $w(u, v)$ 는 다음과 같이 계산된다.

$$w(u, v) = \frac{1}{2N} \sqrt{(u/\theta_x)^2 + (v/\theta_y)^2} \quad (3)$$

여기서 θ_x 와 θ_y 는 한 pixel이 가지는 수평, 수직에 대한 visual angle이다. 만약 한 pixel에 대한 수평, 수직의 visual angle이 같다면 θ_x 와 θ_y 는 아래와 같다.

$$\theta_x = \theta_y = 2 \cdot \arctan\left(\frac{1}{2 \times R_{vd} \times Pic_h}\right) \quad (4)$$

여기서 R_{vd} 는 viewing distance와 영상의 높이의 비이고 Pic_h 는 영상의 수직방향의 pixel 수이다. ITU-R BT.500-11에 따르면 영상의 크기에 따라 R_{vd} 는 3에서부터 6의 값을 가진다. 식(2)에서 s 는 spatial summation effect를 고려하기 위한 파라미터이고 경험적으로 0.25이다^[6]. 식(2)에서 $1/(r + (1-r)\cos^2\psi(u, v))$ 는 인간 시각의 oblique effect를 나타낸다. 여기서 r 은 경험적으로 0.6이고 $\psi(u, v)$ 는 DCT block에서 (u, v) 번째 계수에 대한 신호의 각도이다. 주파수 (u, v) 의 $\psi(u, v)$ 는 아래와 같다.

$$\psi(u, v) = \arcsin\left(\frac{2w(u, 0)w(0, v)}{w(u, v)^2}\right) \quad (5)$$

주파수 $F(u, v)$ 에 대한 Background 신호는 주파수 $F(u, v)$ 보다 주파수가 낮은 저주파수 신호들로 이루어져 있다. 그러므로 주파수 $F(u, v)$ 의 background 신호 에너지, $B(u, v)$,는 다음과 같이 구할 수 있다.

$$B(u, v) = \sum_i \sum_j |F(i, j)|^2 \quad \text{for } \sqrt{i^2 + j^2} < \sqrt{u^2 + v^2} \quad (6)$$

그러므로 $N \times N$ DCT block에서 주파수 $F(u, v)$ 에 대한 perceived contrast (PC)는 다음과 같이 측정된다^[2,3,4,5].

$$PC(u, v) = \frac{|F(u, v)|^{2.4}}{CSF(u, v)^{-2} + B(u, v) + |F(u, v)|^2} \quad (7)$$

거의 모든 영상 신호에서는 영상 신호의 저주파수 대역의 에너지가 고주파수 대역의 에너지보다 높으므로 식(7)에서 $B(u, v)$ 은 $|F(u, v)|^p$ 에 비해 큰 값을 가진다. 따라서 $PC(u, v)$ 을 아래와 같이 계산할 수 있다.

$$PC(u, v) \approx \frac{|F(u, v)|^{2.4}}{CSF(u, v)^{-2} + B(u, v)} \quad (8)$$

III. 인간 시각의 인지 특성을 이용한 영상 화질 향상 방법

인간은 일반적으로 민감하게 인지되는 영상 신호의 성분이 많은 영상을 더욱 선호한다^[1]. 그러므로 인간 시각에 적합하게 영상의 화질을 향상시키기 위해서는 영상 신호의 특성에 따라서 영상의 시각적 민감도를 증가시켜야한다. 그러므로 본 논문은 인간 시각의 perceived contrast를 증가시켜 인간 시각에 적합한 영상 화질 향상 방법을 개발한다.

1. Perceived contrast 향상 방법

인간 시각에 적합하게 화질이 향상된 영상의 perceived

contrast는 원본 영상의 perceived contrast 보다 커야한다. 따라서 본 논문에서는 원 영상의 perceived contrast, $PC(u, v)$,에 1보다 큰 scaling factor를 곱하여 영상의 perceived contrast를 증가시킨다. 향상된 perceived contrast, $\overline{PC}(u, v)$,는 아래와 같이 계산된다.

$$\overline{PC}(u, v) = \lambda(u, v) \cdot PC(u, v) \quad (9)$$

여기서 $\lambda(u, v)$ 는 주파수 (u, v) 에서의 scaling factor이다. 그리고 식(9)에 식(8)을 대입하면 식(10)은 아래와 같다.

$$\begin{aligned} & \frac{\overline{F}(u, v)^{2.4}}{CSF(u, v)^{-2} + \overline{B}(u, v)} \\ &= \lambda(u, v) \cdot \frac{F(u, v)^{2.4}}{CSF(u, v)^{-2} + B(u, v)} \end{aligned} \quad (10)$$

위의 식에서 본 논문은 원본 신호와 향상된 신호의 에너지 향상율을 R 로 정의하였다. R 은 아래와 같이 계산한다.

$$R(u, v) = \frac{CSF(u, v)^{-2} + \overline{B}(u, v)}{CSF(u, v)^{-2} + B(u, v)} \quad (11)$$

그러므로 R 을 이용하여 perceived contrast가 향상된 DCT 계수 $\overline{F}(u, v)$ 는 아래와 같이 구할 수 있다.

$$\overline{F}(u, v) = \{\lambda(u, v) \cdot R(u, v)\}^{1/2.4} \cdot F(u, v) \quad (12)$$

본 논문에서 제안하는 영상 화질 향상 기법은 인간 시각에 적합하게 영상 신호를 향상시키므로 초고화질 영상에서 인간 시각에 두드러지게 인지되는 세밀한 신호들을 향상시킨다. 또한, 더욱 인간 시각에 적합하게 영상의 화질을 향상시키기 위하여 본 논문은 scaling factor, λ ,을 CSF에 따라 아래와 같이 각 주파수별로 다르게 조절하였다.

$$\lambda(u, v) = \alpha \cdot CSF(u, v) + 1 \quad (13)$$

식(13)의 scaling factor는 CSF를 따르므로 인간 시각이 가장 민감하게 반응하는 중간 주파수 대역에서 가장 값이

높고 노이즈를 발생시키는 고주파수 대역에서는 값이 낮다. 그러므로 본 논문은 식(12)를 이용하여 노이즈를 발생시키지 않으면서 인간 시각에 적합하게 영상의 화질을 향상시킨다.

2. 신호 특성에 적응적인 영상 화질 향상 방법

방향성 신호에서 신호 방향에 관계없이 주파수 성분들의 에너지를 증가시키면 **ringing artifact**가 발생한다. 방향성 신호에서 신호 방향에 수평인 주파수 성분들을 향상시키면 그림1의 (b)와 같이 **edge** 주변으로 **ringing artifact**가 발생한다. 반대로 신호 방향에 수직인 주파수 성분들이 향상되면 그림1의 (c)와 같이 **ringing artifact**가 발생하지 않으면서 **edge**신호를 향상시킬 수 있다. 그러므로 방향성 신호에서 **artifact** 없이 신호를 향상시키기 위해서는 신호 방향을 측정하고 신호 방향에 수직인 주파수 성분들을 향상시켜야 한다. 그러나 방향성이 없는 신호에서 하나의 방향성 성분

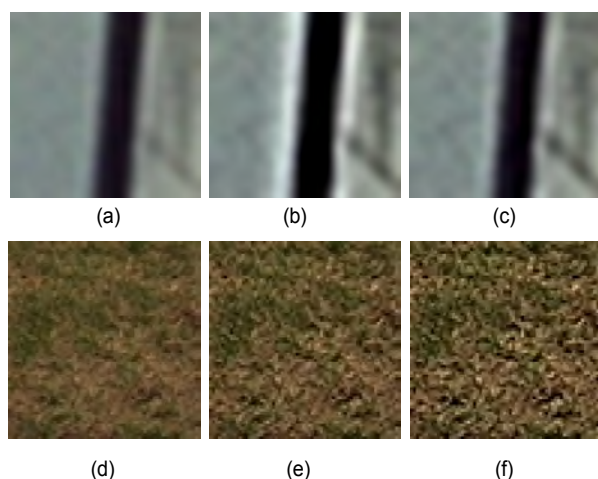


그림 1. 신호 방향에 따른 영상 화질 향상 결과. (a) 방향성이 있는 영상, (b)신호 방향에 수평인 주파수 성분을 향상시킨 영상, (c)신호 방향에 수직인 주파수 성분을 향상시킨 영상, (d) 방향성이 없는 영상, (e) 신호 방향에 수직인 주파수 성분을 향상시킨 영상, (f)전방향의 주파수 성분을 향상시킨 영상.

Fig 1. Image enhancement result by signal direction. (a) Directional image, (b) Image enhanced frequency components parallel to signal direction, (c) Image enhanced frequency components perpendicular to signal direction, (d) Non-directional image, (e) Image enhanced frequency components perpendicular to signal direction, (f) Enhanced image by all direction.

만 향상시키면 그림1의 (e)와 같이 영상을 충분히 향상시킬 수 없다. 그러므로 방향성이 없는 신호에서는 그림1의 (f)와 같이 영상 신호를 충분히 향상시키기 위하여 모든 방향의 주파수 에너지를 증가시켜야 한다.

본 논문은 DCT 영역에서 신호의 방향을 측정하기 위하여 DCT 영역에서의 에너지 분포를 이용한다. DCT의 각 계수들은 주파수와 방향에 대한 정보를 가지고 있다. 그림 2는 수직, 수평 방향의 영상 신호에 대한 DCT 영역에서의 에너지 분포를 보여준다. 완벽한 수평 방향의 **edge** 영상인 경우 수직 방향으로만 신호의 변화가 있기 때문에 DCT 영

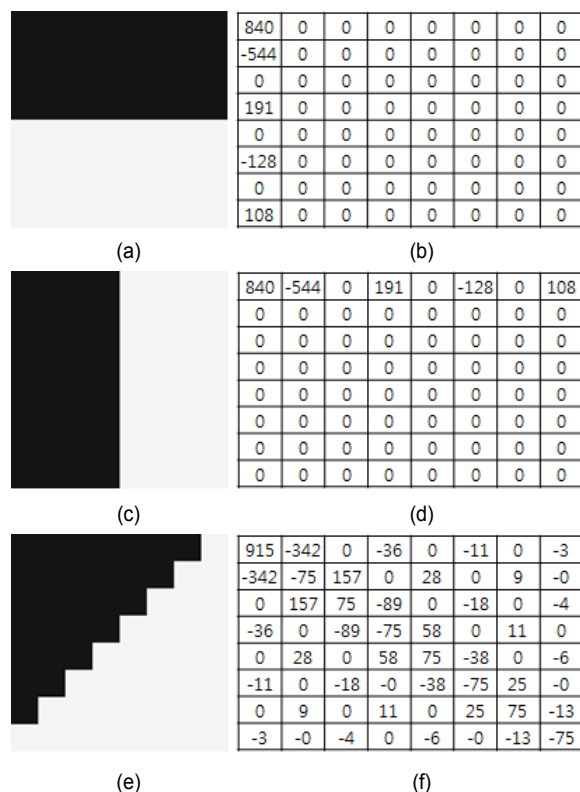


그림 2. 신호 방향에 따른 DCT 영역에서 에너지 분포. (a) 수평 edge, (b)수평 edge에 대한 DCT 영역에서의 에너지 분포, (c)수직 edge, (d)수직 edge에 대한 DCT 영역에서의 에너지 분포, (e)대각선 edge, (f)대각선 edge에 대한 DCT 영역에서의 에너지 분포.

Fig 2. Frequency energy distribution by signal direction in DCT. (a) Horizontal edge, (b) Energy distribution of horizontal edge in DCT domain, (c) Vertical edge, (d) Energy distribution of vertical edge in DCT domain, (e) Diagonal edge, (f) Energy distribution of diagonal edge in DCT domain.

역에서는 수평 방향의 성분들은 0이 되고 수직 방향의 성분들만 값을 가진다. 마찬가지로 완벽한 수직 방향의 edge 영상의 경우 DCT 영역에서 첫 번째 행의 주파수 성분들만 값을 가진다. 그리고 대각선 edge 영상의 경우 주파수 성분들이 대각선을 기준으로 대칭을 이룬다. 그러므로 신호의 방향은 수직 수평 성분에 해당하는 주파수 성분의 에너지 비율로 측정할 수 있다^[10,11]. 수평, 수직 방향의 신호 에너지 비율을 각각 ∇_{hor} , ∇_{ver} 라고 하면, ∇_{hor} , ∇_{ver} 은 다음과 같이 측정된다.

$$\nabla_{hor} = \frac{\sum_{i=1}^{N-1} |F(0,i)|}{\sum_{i=1}^{N-1} \{|F(i,0)| + |F(0,i)|\}}, \quad (14)$$

$$\nabla_{ver} = \frac{\sum_{i=1}^{N-1} |F(i,0)|}{\sum_{i=1}^{N-1} \{|F(i,0)| + |F(0,i)|\}}$$

본 논문은 식(14)의 ∇_{hor} , ∇_{ver} 을 이용하여 영상의 모든 방향의 측정한다. $\nabla_{hor}=1$, $\nabla_{ver}=0$ 이면 영상은 수평방향이며, $\nabla_{hor}=0$, $\nabla_{ver}=1$ 이면 영상은 수직방향이다. 그리고 $\nabla_{hor}=\nabla_{ver}=0.5$ 이면 영상은 대각선 방향이다. 본 논문은 수평, 수직 방향에 적합하게 perceived contrast를 향상시키기 위하여 \bar{F}_{ver} 와 \bar{F}_{hor} 을 각각 계산한 후 방향에 따라 \bar{F}_{ver} 와 \bar{F}_{hor} 을 합성하여 신호 방향에 적응적인 화질 향상 방법을 개발하였다. 식(13)의 scaling factor는 전 방향으로 에너지를 향상시키므로 각 방향에 적합하게 에너지를 향상시키기 위해서 수평, 수직에 대한 scaling factor, λ_{hor} , λ_{ver} ,를 아래와 같이 구한다.

$$\lambda_{hor}(v) = \alpha \cdot CSF(0,v) + 1, \quad (15)$$

$$\lambda_{ver}(u) = \alpha \cdot CSF(u,0) + 1$$

그리고 수평, 수직에 대한 background 신호 에너지, $B_{hor}(v)$, $B_{ver}(u)$,는 다음과 같다.

$$B_{hor}(v) = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{v-1} |F(i,j)|^2 \quad (16)$$

$$B_{ver}(u) = \frac{1}{N^2} \sum_{i=0}^{u-1} \sum_{j=0}^{N-1} |F(i,j)|^2$$

그리고 수평, 수직에 대한 에너지 향상율, R_{hor} , R_{ver} , 은 다음과 같이 계산된다.

$$R_{hor}(u,v) = \frac{CSF(u,v)^{-2} + \bar{B}_{hor}(v)}{CSF(u,v)^{-2} + B_{hor}(v)}, \quad (17)$$

$$R_{ver}(u,v) = \frac{CSF(u,v)^{-2} + \bar{B}_{ver}(u)}{CSF(u,v)^{-2} + B_{ver}(u)}$$

그러므로 수평, 수직 방향에 적합하게 향상된 DCT 계수, \bar{F}_{hor} , \bar{F}_{ver} , 는 아래와 같다.

$$\bar{F}_{ver}(u,v) = (\lambda_{ver}(u) \cdot R_{ver}(u,v))^{1/2.4} \cdot F(u,v)$$

$$\bar{F}_{hor}(u,v) = (\lambda_{hor}(v) \cdot R_{hor}(u,v))^{1/2.4} \cdot F(u,v) \quad (18)$$

그리고 신호 방향에 따라 \bar{F}_{hor} 와 \bar{F}_{ver} 을 합성하기 위하여 본 논문은 방향별 신호 에너지 비율을 아래와 같이 사용하였다.

$$\bar{F}_D(u,v) = \nabla_{ver} \cdot \bar{F}_{ver}(u,v) + \nabla_{hor} \cdot \bar{F}_{hor}(u,v)$$

$$= \left\{ \begin{array}{l} \nabla_{ver} \cdot (\lambda_{ver}(u,v) \cdot R_{ver}(u,v))^{1/2.4} \\ + \nabla_{hor} \cdot (\lambda_{hor}(u,v) \cdot R_{hor}(u,v))^{1/2.4} \end{array} \right\} \cdot F(u,v) \quad (19)$$

여기서 \bar{F}_D 는 방향성 신호에 대한 향상된 주파수 계수이다. 그러므로 \bar{F}_D 는 방향별 신호 에너지 비율에 따라 신호에 적응적으로 변환다. 수평 방향에 영상 신호에서는 $\nabla_{hor} \gg \nabla_{ver}$ 이므로 $\bar{F}_D \approx \bar{F}_{hor}$ 이고, 수직 방향에 영상 신호에서는 $\nabla_{hor} \ll \nabla_{ver}$ 이므로 $\bar{F}_D \approx \bar{F}_{ver}$ 이다. 그리고 대각선 방향의 영상 신호에서는 $\nabla_{hor} \approx \nabla_{ver}$ 이므로 \bar{F}_D 는 \bar{F}_{hor} 와 \bar{F}_{ver} 의 합성으로 결정된다.

방향성 신호와 달리 방향성이 없는 신호에서는 영상 신호를 충분히 향상시키기 위하여 모든 방향으로 주파수 성분을 향상시켜야 한다. 영상 신호를 모든 방향으로 향상시키기 위하여 본 논문은 식 (13)의 scaling factor와 식 (11)의 R 을 사용하여 방향성이 없는 신호에 대한 향상된 주파수 계수 \bar{F}_{ND} 를 아래와 같이 구한다.

$$\bar{F}_{ND}(u,v) = \{\lambda(u,v) \cdot R(u,v)\}^{1/2.4} \cdot F(u,v) \quad (20)$$

일반적인 영상 신호는 방향성 신호와 방향성이 없는 신호를 복합적으로 포함하고 있다. 그러므로 신호 특성에 적합하게 영상의 화질을 향상시키기 위해서는 신호의 방향성에 따라 \bar{F}_N 와 \bar{F}_{ND} 을 조합해야 한다. 방향성이 없는 신호에서는 주파수 영역에서 에너지가 모든 주파수에 퍼져있고 방향성이 있는 신호에서는 저주파수 대역에 에너지가 집중된다. 그러므로 본 논문은 신호의 방향성 정도를 측정하기 위하여 저주파수 대역의 에너지와 고주파수 대역의 에너지 비율, ER ,을 아래와 같이 사용한다.

$$ER = \frac{\sum_{i^2+j^2=1}^{\lfloor (2N^2-1)/3 \rfloor} B(i,j)}{\sum_{i^2+j^2=\lfloor (2N^2-1)/3 \rfloor+1}^{2N^2-1} B(i,j)} \quad (21)$$

그리고 ER 을 fuzzy thresholding 하여 아래와 같이 \bar{F}_N 과 \bar{F}_{ND} 을 결합하여 enhanced된 영상의 주파수 계수 \bar{F} 을 구한다.

$$\bar{F}(u,v) = \frac{ER}{ER+\varepsilon} \cdot \bar{F}_D(u,v) + (1 - \frac{ER}{ER+\varepsilon}) \cdot \bar{F}_{ND}(u,v) \quad (22)$$

여기서 ε 는 fuzzy threshold이다. 본 논문은 실험적으로 ε 을 3으로 사용하였다.

3. Block artifact 방지 기법

Ringings artifact와 더불어 DCT기반의 영상 처리 방법들은 block 단위로 신호를 처리하기 때문에 block artifact를 발생시킨다. 주파수 영역에서 DC와 저주파수 성분들은 영상의 전체적인 밝기에 영향을 미친다. 그리고 저주파수 성분들의 에너지가 block별로 변하면 block 간의 밝기가 불연속하게 되어 block artifact가 발생된다. 그러므로 block artifact를 방지하기 위해서는 DC와 저주파수 성분들의 에너지를 유지시켜야 한다.

기존의 block artifact 방지 방법은 8×8 DCT의 경우 DC를 포함한 처음 3번째 대역까지의 주파수 성분들의 에너지를 유지시켰다^[15]. 그러나 기존의 방법은 신호 특성에 상관없이 고정된 범위의 주파수 성분들의 에너지를 유지시키므로 화질 향상 효과를 감소시키면서 특정 신호에서는 block

artifact를 여전히 발생시켰다. 그러므로 본 논문은 신호 특성에 적응적으로 에너지를 유지시켜야 할 주파수 성분들의 범위를 조절하여 화질 향상 효과를 감소시키지 않으면서 block artifact를 감소시키는 방법을 제안한다.

인간 시각은 상대적으로 저주파수 대역에 에너지가 많은 신호에 block artifact를 쉽게 인지하지만, 에너지가 전 주파수 대역 퍼져있는 신호에서 block artifact를 쉽게 인지하지 못한다. 그러므로 본 논문은 저주파수 대역의 에너지와 고주파수 대역의 에너지 비율인 ER 에 따라 에너지를 유지시켜야 할 주파수 범위, PR ,을 경험적으로 다음과 같이 설정한다.

$$PR = \begin{cases} \text{round}(N/7) & ER < 5 \\ \text{round}(N/5) & 5 \leq ER < 15 \\ \text{round}(N/3) & ER \geq 15 \end{cases} \quad (23)$$

그리고 본 논문은 식(23)의 PR 에 따라 아래와 같이 λ , λ_{ver} , λ_{hor} 을 변경한다.

$$\begin{aligned} \lambda'_{ND}(u,v) &= \begin{cases} 1 & u^2+v^2 \leq 2PR^2 \\ \lambda(u,v) & u^2+v^2 > 2PR^2 \end{cases} \\ \lambda'_{ver}(u) &= \begin{cases} 1 & u \leq PR \\ \lambda_{ver}(u) & u > PR \end{cases} \\ \lambda'_{hor}(v) &= \begin{cases} 1 & v \leq PR \\ \lambda_{hor}(v) & v > PR \end{cases} \end{aligned} \quad (23)$$

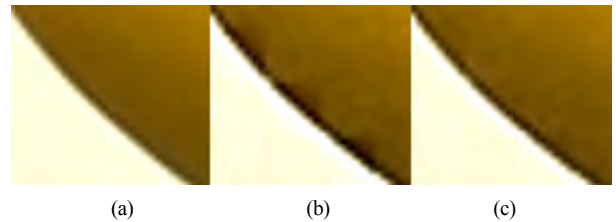


그림 3. Block artifact 방지 기법. (a) 원본 영상, (b) 기존 block artifact 방지 기법으로 고정된 주파수 범위의 에너지를 유지시켰을 때 block artifact가 발생한 영상, (c)제안하는 block artifact 방지 기법을 이용해 block artifact를 감소시킨 영상.

Fig 3. Block artifact preventing method. (a) Original image, (b) Block artifact image by existing method, (c) Enhanced image without block artifact by proposed method.

그림 3은 기존의 block artifact 방지 기법과 본 논문에서

제안하는 기법으로 영상의 화질을 향상시켰을 때 block artifact의 발생 정도를 보여준다. 그림 3의 (a)와 같은 신호는 저주파수 성분들에 상대적으로 에너지가 몰려 있기 때문에 기존의 기법으로는 그림3의 (b)와 같이 block artifact가 발생하지만 본 논문에서 제안하는 기법은 신호에 적응적으로 에너지를 유지시켜야할 주파수 범위를 조절하였기 때문에 그림3의 (c)와 같이 block artifact가 거의 발생하지 않을 것을 볼 수 있다.

IV. 실험 및 논의

본 논문에서는 DCT block 크기를 HD(1280×720)보다 해상도가 높은 경우는 16×16 block을 사용하였고 해상도

가 HD 이거나 그보다 작은 경우 8×8 block을 사용하였다. 본 논문은 영상에 관계없이 parameter들을 고정하였다. α 는 7로 사용하였다. 본 논문은 개발한 방법의 화질 향상 정도를 증명하기 위해 객관적인 평가지표로 CPBD를 사용하였고 주관적인 평가지표로는 MOS를 사용하였다. 실험에 사용된 영상은 [18]에서 제공하는 초고화질 영상들이다.

표 1은 CPBD 실험 결과를 보여준다. CPBD 실험을 통해 본 논문이 제시하는 영상 화질 향상 방법은 기존의 방법들에 비해 대체적으로 우수한 것을 볼 수 있다. 그러나 flower_foveon과 spider_web과 같은 texture가 적은 영상에서는 히스토그램 기법인 HE와 CLAHE와 영상 전체에 DCT 수행하는 Celik의 기법이 높은 CPBD 값을 가진다.

본 논문에서는 MOS 실험하기 위해 동일한 TV 두 개를 나란히 하였고 처리된 영상과 원본 영상을 무작위로 배치

표 1. 기존의 방법들과 제안하는 방법에 대한 CPBD 점수

Table 1. CPBD scores between the conventional methods and proposed method

Image	HE	CLAHE	S. Lee ^[15]	T. Celik ^[16]	Proposed
artificial	0.8048	0.7128	0.7435	0.7685	0.7731
bridge	0.7467	0.6678	0.7028	0.7636	0.7753
cathedral	0.7228	0.6610	0.7550	0.8070	0.7891
deer	0.8778	0.8443	0.8799	0.8993	0.9002
fireworks	0.9337	0.7162	0.6750	0.5677	0.6540
flower_foveon	0.7969	0.6690	0.4557	0.6591	0.4666
hdr	0.7436	0.6921	0.7034	0.7942	0.7388
leaves_iso_200	0.5351	0.4481	0.5371	0.6547	0.6242
nightshot_iso_100	0.8497	0.5683	0.6320	0.6996	0.6512
spider_web	0.8719	0.7911	0.7381	0.8595	0.7498

표 2. 기존의 방법들과 제안하는 방법에 대한 MOS 실험

Table 2. MOS test between the conventional methods and proposed method

Image	HE	CLAHE	S. Lee ^[15]	T. Celik ^[16]	Proposed
artificial	0.075	0.065	0.062	0.068	0.071
bridge	0.885	0.726	0.728	0.798	0.829
cathedral	0.441	0.586	0.641	0.682	0.675
deer	0.156	0.231	0.259	0.236	0.287
fireworks	0.956	0.844	0.592	0.645	0.671
flower_foveon	0.464	0.353	0.331	0.355	0.323
hdr	0.375	0.256	0.254	0.263	0.265
leaves_iso_200	0.445	0.586	0.711	0.688	0.735
nightshot_iso_100	0.875	0.675	0.568	0.613	0.684
spider_web	0.187	0.155	0.210	0.235	0.215

표 3. 제안하는 방법의 α 값에 따른 MOS 변화Table 3. Different MOS value of several α in proposed method

α	3.0	5.0	7.0	9.0	11.0
artificial	0.065	0.069	0.071	0.070	0.061
bridge	0.793	0.805	0.829	0.832	0.820
cathedral	0.633	0.659	0.675	0.673	0.663
deer	0.210	0.268	0.287	0.291	0.276
fireworks	0.629	0.657	0.671	0.664	0.653
flower_foveon	0.319	0.322	0.323	0.322	0.320
hdr	0.220	0.255	0.265	0.263	0.257
leaves_iso_200	0.674	0.713	0.735	0.736	0.708
nightshot_iso_100	0.643	0.663	0.684	0.679	0.671
spider_web	0.165	0.199	0.215	0.212	0.198

하였다. 그리고 실험자와 TV와의 거리를 약 2m로 하였다. 총 25명의 실험자에게 실험을 진행하였고 ITU-R BT.500-11에서 제안하는 방식에 따라 MOS 실험을 하였다^[17]. MOS의 신뢰도는 95%로 정하였다. [-3, 3]의 범위에서 화질 평가를 하였으며 0은 처리된 영상과 원본 영상과 같다는 의미이고 -3은 처리된 영상이 원본 영상보다 화질이 상당히 저하되었다는 의미이다. 그리고 3은 처리된 영상이 원본 영상보다 상당히 좋아졌다는 의미이다. 표 2는 실험한 MOS 결과를 보여준다. 표 1에서는 edge나 texture가 거의 없는 flower_foveon과 spider_web영상에서 제안하는 방법과 기존의 방법들의 CPBD값이 큰 차이를 보였지만 향상시킬 수 있는 신호 성분이 적으므로 모든 화질 향상 방법들이 대체적으로 낮은 MOS 값을 가지며 서로 큰 차이를 보이지 않는다. 반면에 bridge, cathedral, fireworks 그리고 leaves_iso_200와 같이 edge나 texture가 많은 영상에서는 본 논문에서 제안하는 영상 화질 향상 기법이 기존의 방법들에 비해 높은 MOS 값을 가지므로 본 논문에서 제안하는 영상 화질 향상 방법이 기존의 방법들에 비해 인간 시각에 적합하게 영상의 화질을 향상시키는 것을 수치적으로 알 수 있다. 그림 4와 5는 각 화질 향상 방법에 대한 결과 영상을 보여준다. 그림 4, 5의 (b), (c)는 원본 영상에 비해 과도하게 향상되어 일부 실험자들은 부자연스럽게 인식하였다. 반면에 그림 4와 5의 (d), (e), (f)는 모든 실험자들이 자연스럽게 향상되었다고 느꼈다. 그러나 그림 4의 (d)에서의 기둥 줄무늬, 천장 무늬의 edge 부근 그리고 그림 5의 (d)에서의 나무줄기 부분에서 ringing artifact와 block artifact가 보

인다. 그림 4의 (e)는 본 논문에서 제안하는 방법의 결과 영상과 큰 차이가 없이 보이나 [16]의 방법은 영상을 전체적으로 처리해야하기 때문에 본 논문에서 제안하는 방법이 실제 알고리즘 구현측면에서 좀 더 효율적이다. 그리고 그림 5에서는 본 논문에서 제안하는 방법이 그림 5의 (e)보다 자연스러우면서 contrast를 더욱 향상시켰다. 표 3은 제안하는 방법에서 여러 α 에 따른 MOS 실험 결과를 보여준다. 대부분의 영상들에서 α 가 7일 때 가장 높은 MOS 수치를 보였다. 그러나 대부분의 영상에서 α 가 7과 9일 때 MOS 수치에 큰 차이가 없으며, texture가 많은 영상에서는 α 가 9일 때 가장 높은 MOS 수치를 갖는다. 그리고 α 가 3, 5일 때는 낮은 enhancement 효과로 MOS 수치가 낮으며, α 가 11일 때는 과도한 enhancement 효과로 오히려 MOS 수치가 떨어졌다.

제안하는 방법은 기존의 multiband 방법에 비하여 수직, 수평 gradient를 구하는 과정이 추가되었고, 기존의 방법인 한 방향으로만 주파수 대역별 에너지를 구하는 것에 비해 수직, 수평의 두 방향에 대해 주파수 대역별 에너지를 구하는 것이 추가 되었다. 수직, 수평 gradient를 측정하는 데 필요한 계산량은 화소 당 덧셈 연산이 0.094개 그리고 곱셈 연산이 0.125개이다. 수직, 수평 방향에 대해 주파수 대역별 에너지를 구하는 계산 과정이 기존의 방법에 비하여 추가되는 계산량은 화소 당 덧셈 연산이 0.984개 곱셈 연산이 2.906개, 그리고 root 연산이 0.047개이다. 그리고 CSF에 따라 λ 를 구하는 과정은 입력 영상의 해상도에 따라 이미 결정되어있는 값이므로 계산량에 추가되지 않는다. 결과적

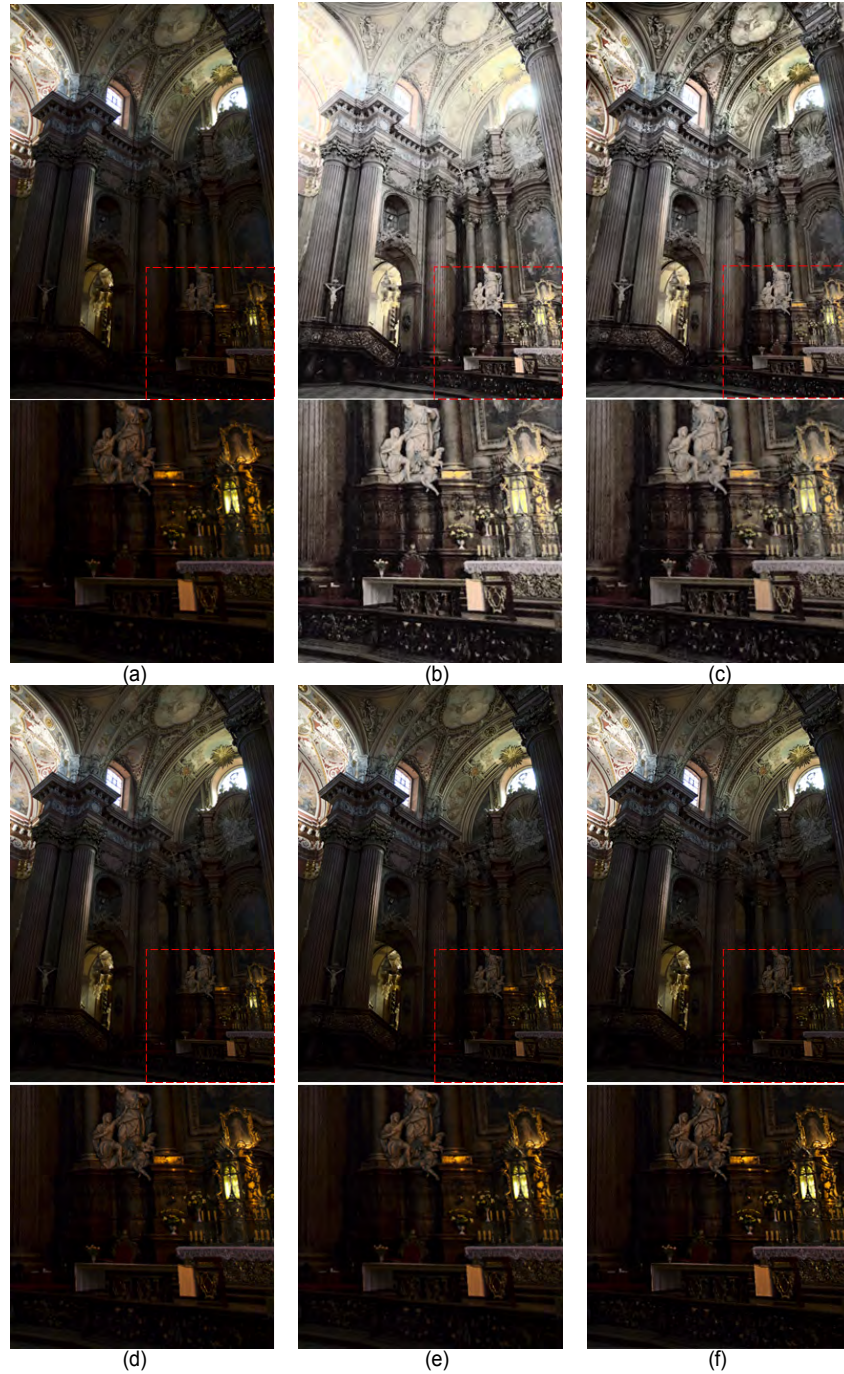


그림 4. 영상 화질 향상 방법 비교. (a) 원본 영상, (b)Histogram Equalization 방법으로 향상시킨 결과 영상, (c)CLAHE 방법으로 향상시킨 결과 영상, (d) Lee의 방법으로 향상시킨 결과 영상[15], (e)Celik의 방법으로 향상시킨 결과 영상[16]. (f) 제안하는 방법으로 향상시킨 결과 영상
Fig 4. Comparison of image enhancement result. (a) Original image, (b) Image enhanced by histogram equalization, (c) Image enhanced by CLAHE, (d)Image enhanced by Lee[15], (e)Image enhanced by Celik[16], (f)Image enhanced by proposed method.

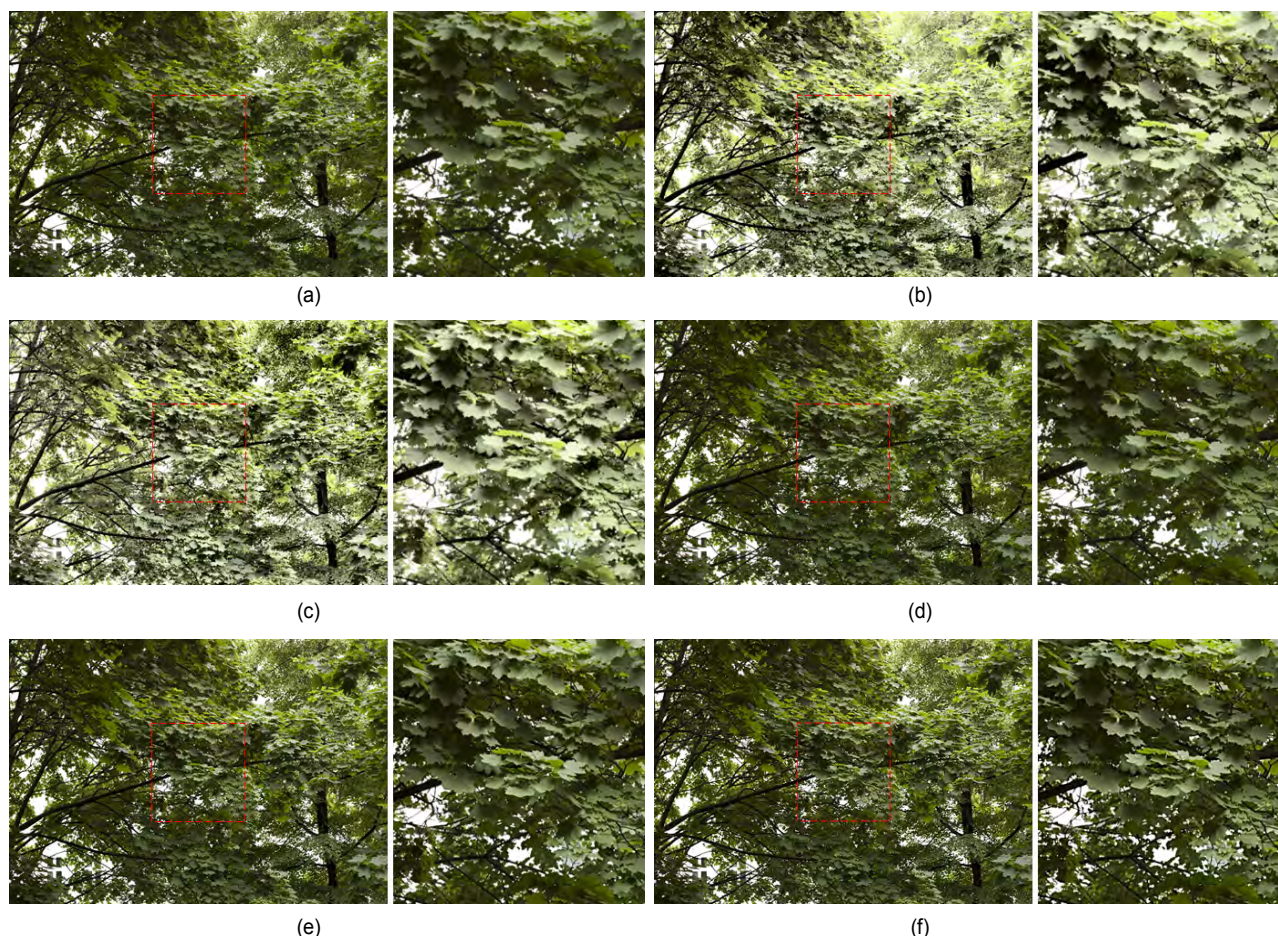


그림 5. 영상 화질 향상 방법 비교. (a) 원본 영상, (b)Histogram Equalization 방법으로 향상시킨 결과 영상, (c)CLAHE 방법으로 향상시킨 결과 영상, (d) Lee의 방법으로 향상시킨 결과 영상[15], (e)Celik의 방법으로 향상시킨 결과 영상[16]. (f) 제안하는 방법으로 향상시킨 결과 영상

Fig 5. Comparison of image enhancement result. (a) Original image, (b) Image enhanced by histogram equalization, (c) Image enhanced by CLAHE, (d)Image enhanced by Lee[15], (e)Image enhanced by Celik[16], (f)Image enhanced by proposed method.

으로 기존의 방법에 비해 추가되는 총 계산량은 화소 당 덧셈 연산 1.078개 곱셈 연산 3.03개, 그리고 root 연산 0.047개이므로 계산량 증가는 매우 미미하다. 실제로 제안하는 방법을 Intel(R) Core(TM) i5-2500 CPU의 컴퓨터에서 실행하였을 때 HD영상을 0.01sec이내에 처리한다.

V. 결 론

본 논문은 영상 신호에 대한 인간 시각의 인지 특성을 이용하여 인간 시각에 적합하게 영상의 local contrast를 증

가시켰다. 그리고 신호처리 과정에서 발생하는 ringing artifact와 block artifact를 발생시키지 않기 위하여 신호의 방향과 주파수 영역에서의 신호 에너지 분포를 이용하였다. 그러므로 본 논문에서 제안하는 방법은 영상의 신호의 물리적 특성에 따라, 인간 눈의 생리적 특성에 따라 영상 신호를 향상시킬 수 있다.

참 고 문 헌 (References)

- [1] B. Spehar, S. Wong, S. Klundert, J. Lui, C. W. G. Clifford, and R. P. Tatlör, "Beauty and the beholder: the role of visual sensitivity in visual

- preference,” *Original Research*, Sep 2015.
- [2] A. M. Haun, and E. Peil, “Perceived contrast in complex images,” *Journal of Vision*, Nov 2013.
 - [3] A. M. Haun, and E. A. Essock, “Contrast Sensitivity for Oriented Patterns in 1/f Noise : Contrast Response and The Horizontal effect,” *Journal of Vision*, Aug 2010.
 - [4] J. M. Foley, “Human Luminance Pattern-vision Mechanisms : Masking Experiments Require a New Model,” *Optical Society of America*, vol.11, no.6, June 1994.
 - [5] M. W. Cannon, and S. C. Fullenkamp, “A Transducer Model fore Contrast Perception,” *Vision Res.* vol.31, no.6, 1991.
 - [6] Z. Wei, and K. N. Ngan, “Spatio-Temporal Just Noticeable Distortion Profile for Grey Image/Video in DCT Domain,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol.19, no.3, Mar 2009.
 - [7] S. H. Bas, and M. Kim, “A DCT-based Total JND Profile for Spatio-Temporal and Foveated Masking Effects,” *IEEE Transaction, on Circuits and Systems for Video Technology*, 2015.
 - [8] S. H. Base, and M. Kim, “A Novel-DCT based JND Model for Luminance Adaptation Effect in DCT Frequency,” *IEEE Signal Processing Letters*, vol.20, no.9, Sep 2013.
 - [9] S. H. Bae, and M. Kim, “A Novel Generalized DCT-Based JND Profile Based on an Elaborate CM-JND Model for Variable Block-Sized Transforms in Monochrome Images,” *IEEE Transactions on Image Processing*, vol.23, no.8, Aug 2014.
 - [10] M. Eom, and Y. Choe, “Fast Extraction of Edge Histogram in DCT Domain based on MPEG7,” *Proceedings of World Academy of Science Engineering and Technology*, vol.9, Nov 2005.
 - [11] H. S. Chang, and K. Kang, “A Compressed Domain Scheme for Classifying Block Edge Patterns,” *IEEE Transactions on Image Processing*, vol.14, no.2, Feb 2005.
 - [12] D. K. Choi, L. H. Jang, M. H. Kim, and N. C. Kim, “Color Image Enhancement Based on Single-Scale Retinex With a JND-Based Nonlinear Filter,” *Circuits and Systems, IEEE International Symposium on*, 2007.
 - [13] S. C. Nercessian, K. A. Panetta, and S. S. Agaian, “Non-Linear Direct Multi-Scale Image Enhancement Based on the Luminance and Contrast Masking Characteristics of the Human Visual System,” *IEEE Transactions on Image Processing*, vol.22, no.9, Sep 2013.
 - [14] J. Tang, E. Peli, and S. Acton, “Image Enhancement Using A Contrast Measure in the Compressed Domain,” *IEEE Signal Processing Letters*, vol.10, no.10, Oct 2003.
 - [15] S. Lee, “An Efficient Content-Based Image Enhancement in the Compressed Domain Using Retinex Theory,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol.17, no.2, Feb 2007.
 - [16] T. Celik, “Spatial Entropy-Based Global and Local Image Contrast Enhancement,” *IEEE Transactions on Image Processing*, vol.23, no.12, Dec 2014.
 - [17] ITU-R Recommendation BT.500-11, *M. for the Subjective Assessment of the Quality of Television Pictures 2002*, 2002.
 - [18] The New Image Compression Test Set - Jan 2008, http://imagecompression.info/test_images/ (accessed Mar. 13, 2018).

저 자 소 개



방 성 배

- 2013년 2월 : 경희대학교 전자전파공학과 공학사
- 2015년 2월 : 경희대학교 전자전파공학과 석사
- 2015년 2월 ~ 현재 : 경희대학교 전자전파공학과 박사과정
- ORCID : <http://orcid.org/0000-0001-7696-5974>
- 관심분야 : 영상/동영상 신호처리, 컴퓨터 비전



김 원 하

- 1985년 2월 : 연세대학교 전자공학과 공학사
- 1988년 5월 : University of Wisconsin-Madison 전기공학과 석사
- 1996년 1월 ~ 7월 : (미) Motorola 연구원
- 1997년 5월 : University of Wisconsin-Madison 전기공학과 박사
- 1997년 8월 ~ 2000년 2월 : (미) Los Alamos National Lab. 연구원
- 2000년 3월 ~ 2003년 8월 : 명지대학교 정보통신공학과 조교수
- 2009년 8월 ~ 2010년 8월 : University of California San Diego (UCSD) 교환 교수
- 2003년 9월 ~ 현재 : 경희대학교 전자정보대학 전자공학과 교수
- ORCID : <http://orcid.org/0000-0002-4797-3526>
- 관심분야 : 영상/동영상 신호처리, 컴퓨터 비전, 동영상 부호화