Dynamic range compression and contrast enhancement for digital images in the compressed domain

Sangkeun Lee
Victor H. S. Ha, MEMBER SPIE
Samsung Information Systems America
Digital Media Solutions Laboratory
3345 Michelson Dr. #250
Irvine, California 92612
E-mail: sangkny@sisa.samsung.com

Yeong-Hwa Kim
Chung-Ang University
Heukseok-Dong
Dongjak-Gu, Seoul, Korea

Abstract. We develop a simple and efficient algorithm for dynamic range compression and contrast enhancement of digital images in the compressed domain. The basic idea of our approach is to separate illumination and reflectance components of an image in the compressed domain. We adjust the amount of contribution of the illumination component to effectively compress the dynamic range of the image. For contrast enhancement, we modify the reflectance component based on a new measure of the spectral contents of the image. The spectral content measure is computed from the energy distribution across different spectral bands in a discrete cosine transform (DCT) block. The advantages of the proposed algorithm are (1) high dynamic range scenes are effectively mapped to the smaller dynamic range of the image, (2) the details in very dark or bright areas become clearly visible, (3) the computational cost is low, and (4) the compressibility of the original image is not affected by the algorithm. We evaluate the performance of the proposed algorithm with well-known existing methods, such as histogram equalization and α-rooting algorithm, using a few different enhancement quality metrics.

© 2006 Society of Photo-Optical Instrumentation Engineers. [DOI: 10.1117/1.2174802]

Subject terms: retinex; illumination; reflectance; dynamic range compression; contrast enhancement; image enhancement quality metrics; image enhancement quality factor.

Paper 050189R received Mar. 15, 2005; revised manuscript received Jun. 7, 2005; accepted for publication Jul. 12, 2005; published online Feb. 22, 2006.

1 Introduction

Image enhancement processes improve the appearance of images to human viewers. Dynamic range compression and contrast enhancement are important aspects of image enhancement with interesting applications in the areas of image analysis and interpretation, automatic target detection, classification, and identification.

The purpose of dynamic range compression is to map the natural dynamic range of a signal to a smaller range. A scene often contains a large dynamic range that cannot be adequately captured by the imaging devices. In the dynamic range compression process, we apply image processing techniques to recover the effects of the full dynamic range from the captured image. After the dynamic range compression is achieved, however, some details may become clustered together within a small intensity range, especially in very dark or bright regions of the image. A contrast enhancement process adjusts the local contrast in different regions of the image so that the details in dark or bright regions are again brought out and revealed to the human viewers. Dynamic range compression and contrast enhancement together improve the quality of the image. The enhanced image looks closer to real natural scenes, clearer with more details, and more visually pleasing.

Many techniques have been developed for dynamic range compression and contrast enhancement.1-10 According to the survey of available techniques found in Refs. 1, 2, and 11, the existing techniques can be categorized into two classes:12 spatial domain processing and compressed-domain processing. The spatial domain methods operate directly on image pixels. Many of these methods are based on gray-level histogram modifications,2 while others are based on local contrast measures and edge information.3,4 In particular, the retinex approach has been introduced and successfully applied to image dynamic range compression.8 The compressed domain methods7 operate directly on the transform coefficients of the images that are compressed by, for example, Fourier, wavelet, or discrete cosine transforms (DCTs). Some of the examples in this category include α-rooting,6 modified unsharp masking,2,13 and multiscale contrast measure algorithms. The advantages of compressed domain processing techniques are (1) low complexity of computations and (2) ease of viewing and manipulating the frequency composition of the image.12 However, it is reported that the compressed domain methods often introduce block artifacts, i.e., superfluous edges near block boundaries. They cannot simultaneously enhance all parts of the image very well, either.1,2,11

In this paper, we present a simple and efficient algorithm to compress image dynamic range and to enhance image contrast in the compressed domain. We pay special attention to block artifacts and other common side effects to overcome the weakness of compressed domain approaches.
The main idea of the proposed algorithm is to separate the DCT coefficients into illumination (dc coefficients) and reflectance (ac coefficients) components. The dc coefficients are adjusted based on the retinex theory to compress the image dynamic range. To enhance the contrast, the ac coefficients are modified according to a newly defined measure of spectral content of the image. To reduce block artifacts, several ac coefficients in low-frequency bands receive a special treatment during the contrast enhancement process. The proposed algorithm has the following advantages: (1) the algorithm is fast because it operates directly in the compressed domain; (2) the algorithm reduces storage requirements and computational costs as the majority of the DCT coefficients in the compressed domain are zero after quantization; (3) the algorithm improves the visibility of image details in very dark or bright areas; (4) the algorithm is noniterative, whereas other previous approaches including the original retinex theory require many iterations; (5) the algorithm does not affect the compressibility of the original image; and (6) the algorithm can be applied to images that are compressed by any DCT-based compression standards, such as JPEG, MPEG, and H.26X, without any significant modification.

This paper is organized as follows. Section 2 reviews basic concepts of the DCT-based image compression scheme, the retinex theory, and the new measure of spectral content of images. Section 3 presents the proposed algorithm for dynamic range compression and contrast enhancement. In Sec. 4, we apply the proposed method to the luminance component of different test images to demonstrate and compare its performance with several existing algorithms. For quantitative comparison of performance, several quality metrics based on the perceptual properties of the human visual system are used. Section 5 concludes the paper with a brief summary and a discussion on future works.

2 Background

We review the basic concepts used in this paper. First, we provide a brief introduction to image compression based on the DCT. The retinex theory is described next to set the foundation for the dynamic range compression problem. Finally, a new measure of spectral contents of images is defined and discussed in relation to the contrast enhancement problem.

2.1 DCT-Based Image Compression

We provide a simple synopsis of the JPEG image coding standard as a DCT-based image compression scheme. The basic idea of JPEG image coding is easily extended to other DCT-based image/video compression schemes including MPEG and H.26X.

A baseline image compression scheme employing DCT consists of the following steps:

1. Partition an image into blocks of 8 × 8 pixels
2. Perform the DCT on each block
3. Quantize the DCT coefficients
4. Zigzag scan and entropy code the quantized DCT coefficients

In the first step, an image is divided into a set or “tiles” of blocks where a block is an array of 8 × 8 pixels. Each block is then transformed into the spatial frequency domain via a forward DCT, defined for an 8 × 8 block $I_{ij}$ as

$$d_{u,v} = \frac{1}{4} C_0 C_u \sum_{i=0}^{7} \sum_{j=0}^{7} \cos \left( \frac{(2i+1)u\pi}{16} \right) \times \cos \left( \frac{(2j+1)v\pi}{16} \right) I_{ij}, \quad \text{for } u,v = 0, \ldots, 7,$$

where

$$C_\eta = \begin{cases} \frac{1}{\sqrt{2}}, & \text{for } \eta = 0, \\ 1, & \text{otherwise}. \end{cases}$$

The element $d_{0,0}$ in the upper left corner of an 8 × 8 DCT encoded block is the dc coefficient, whereas the 63 other elements are the ac coefficients. Figure 1 shows an array of 64 DCT bases. These basis functions are arranged in order of increasing spatial frequencies from the upper left corner to the lower right corner. Each coefficient $d_{u,v}$ in Eq. (1) represents the contribution from the DCT basis function located at the $u$'th column and $v$'th row in Fig. 1, for $(u,v) = 0, \ldots, 7$ (Refs. 14 and 15).

After DCT, the 64 DCT coefficients in the 8 x 8 block are quantized, i.e., each coefficient is divided by a corresponding quantization parameter (quantization step) and then rounded to the nearest integer. The default values of the quantization parameters are usually determined by the compression standard. The quantization process eliminates small coefficient values and maps the rest of coefficients to a set of predetermined values. This is the only step in which information is lost.

In the final step, the quantized DCT coefficients are aligned in a zigzag scan order and entropy coded. The entropy coding scheme that is often used is the run-length coding, where a long string of zeros is effectively compressed. The encoding process with a block DCT followed by quantization results in a very few nonzero coefficients remaining in each 8 × 8 block. The zigzag scan and run-
length entropy coding takes advantage of this property. Any compressed domain algorithm for image enhancement also benefits from this property when improving the computation speed and complexity.

To reconstruct the original image in the decoding process, the compressed image is first entropy decoded, dequantized by point-to-point multiplication with the quantization parameters, and inverse transformed where the inverse DCT (IDCT) is defined as

\[
L_{i,j} = \frac{1}{4} \sum_{u=0}^{7} \sum_{v=0}^{7} C_u C_v \cos \left( \frac{2(i + 1)u \pi}{16} \right) \cos \left( \frac{(2j + 1)v \pi}{16} \right) d_{u,v}, \quad \text{for } u,v = 0, \ldots, 7. \tag{3}
\]

Each block of an image is reconstructed from the weighted sum of the DCT coefficients that correspond to the specific spatial frequency contributions. Thus, the distribution of the DCT coefficients provides a natural way to define a spectral content measure of the image in the DCT domain, as will be discussed in Section 2.3.

### 2.2 Retinex Theory for Dynamic Range Compression

The retinex theory has been known for more than 30 years as a simple and effective model of the human vision. The name retinex, which comes from the contraction of two words “retina” and “cortex,” indicates the intention to take into account the biological elements that influence our visual perception. The retinex theory is designed to emulate the specific human visual ability, i.e., the ability to see the same objects under different illumination conditions such as in direct sunlight, in shadow, or in the presence of artificial illuminations of different types. This psychophysical phenomenon is often called “brightness/lightness constancy,” or more generally, “color constancy.”

The basic concept in the retinex theory is to separate the illumination and reflectance components of an image. It is assumed that the available luminance data in the image is the product between illumination and reflectance. This means that the reflectance component can be estimated as the ratio between the luminance and an estimate of the illumination. Let the image \(L_{x,y}\) be the result of the point-by-point product of the illumination \(I_{x,y}\) of the scene and the reflectance \(R_{x,y}\) of the objects in the scene:

\[
L_{x,y} = I_{x,y}R_{x,y}. \tag{4}
\]

In the retinex theory, as depicted in Fig. 2, the illumination component is first estimated as \(\tilde{L}_{x,y}\). Then, an estimate of reflectance \(\tilde{R}_{x,y}\) is determined as the ratio between the input luminance \(I_{x,y}\) and the estimated illumination \(\tilde{L}_{x,y}\). That is,

\[
\tilde{R}_{x,y} = I_{x,y}/\tilde{L}_{x,y}. \tag{5}
\]

The estimated illumination \(\tilde{L}_{x,y}\) is next modified by a non-linear operator \(G(\cdot)\) to compress the dynamic range of the image. Finally, the output \(O_{x,y}\) is formed as the product between the estimated reflectance \(\tilde{R}_{x,y}\) and the modified illumination \(G(\tilde{L}_{x,y})\), i.e.,

\[
O_{x,y} = \tilde{R}_{x,y}G(\tilde{L}_{x,y}). \tag{6}
\]

The output image \(O_{x,y}\) is an enhanced version of the input image via dynamic range compression.

At the core of the retinex method is the estimation of the illumination component \(\tilde{L}_{x,y}\). In our algorithm, \(\tilde{L}_{x,y}\) is estimated by applying a low-pass filter to the luminance data \(I_{x,y}\). This is because the illumination is assumed to vary smoothly over the entire scene. The details of dynamic range compression using the retinex theory are presented in Sec. 3.

### 2.3 Spectral Content Measure for Contrast Enhancement

We perform contrast enhancement by modifying the reflectance component \(R_{x,y}\) of the image. Since we are only given an estimate \(\tilde{R}_{x,y}\) of the reflectance component under the retinex framework, we apply the function \(F(\cdot)\) to the estimated reflectance component and generate \(F(\tilde{R})\). Finally, the contrast of the output image \(O_{x,y}\) is enhanced by replacing the original estimate of reflectance \(\tilde{R}\) with \(F(\tilde{R})\). That means, the contrast enhancement function is implemented here as a non-linear operator \(F(\cdot)\) applied to the reflectance component. The function \(F(\cdot)\) adjusts the image details in different frequency bands based on a spectral content measure \(\Phi_{u,v}\) computed adaptively for each frequency band. In this section, we introduce this new measure of spectral contents of images.

The human visual system detects the details of the scene based on the ratio between high-frequency and low-frequency contents. The spatial frequency characteristics of the scene can thus be used to define a relative measure of scene contents. The 64 DCT coefficients in an \(8 \times 8\) block are clustered into 15 frequency bands according to their spatial frequency properties as shown in Fig. 1. Let the 15 frequency bands be expressed as a set \(\phi\)

\[
\phi = [\varphi_0, \varphi_1, \ldots, \varphi_{n-1}, \varphi_n], \quad \text{for } 0 \leq n \leq 14. \tag{7}
\]

Then, each frequency band vector \(\varphi_n\) consists of the quantized DCT coefficients \(d_{u,v}\) where \(n = u + v\). That is,

\[
\varphi_n = (d_{u,v}) \quad \text{for } n = u + v, \quad 0 \leq u,v \leq 7. \tag{8}
\]

The DCT coefficients in each frequency band vector have similar spatial frequency properties. The number of elements in each spectral band vector is computed as
For instance, the spectral band vector \( \varphi_i \) is composed of the elements \( d_{i,1} \) and \( d_{i,0} \), and the number of elements in the vector is computed as \( N(\varphi_i) = 1 + 1 = 2 \). The spectral band vector \( \varphi_n \) with \( n = u + v \) indicates a diamond-shaped approximation to a circle, and its DCT elements are selected approximately in equal radial frequencies.\(^6\) If only one frequency band \( \varphi_n \) is selected in Eq. (3), and an image block is reconstructed using it, the reconstructed block can be thought of as the bandpass version of the original block with a radial frequency \( n \). As the spectral band number increases, the reconstructed block contains a higher frequency content of the input block.

In this paper, the relative spectral content measure \( \Phi_{w,n} \) of the \( w \)th block at the \( n \)th band is defined by the energy ratio of the high spectral content energy to the low spectral content energy in the bands of the DCT coefficients. That is,\(^9\)

\[
\Phi_{w,n} = \frac{\| \varphi_{w,n} \|^2 / \sum_{n=0} \| \varphi_{w,n} \|^2}{1 - n / \sum_{n=0} \| \varphi_{w,n} \|^2}, \quad \text{for } n \geq 1, \tag{10}
\]

where \( \| \cdot \|^2 \) denotes the square of Frobenius norm,\(^16\) the numerator is an average energy of the \( n \)th spectral band, and the denominator is an average energy in the spectral bands lower than the \( n \)th band. The definition of the relative content measure \( \Phi_{w,n} \) is similar to that of the multiscale contrast measure in Refs. 6 and 7. However, we extend and improve their definition in terms of processing speed and image quality especially in dark and bright areas. In Refs. 6 and 7, the authors successfully enhance an image contrast by emphasizing the high-frequency components of the image with their multiscale contrast measure. However, their methods do not consider details hidden inside dark or bright regions of the image. In addition to that, their methods manipulate DCT coefficients using a single measure without considering the sensitivity of human visual system in low-frequency bands, and thus easily create block artifacts. Our new definition in Eq. (10) pays closer attention to the details in dark and bright regions. We also avoid block artifacts in our approach, as discussed in detail in the next section.

3 Dynamic Range Compression and Contrast Enhancement

The proposed algorithm for image enhancement is based on dynamic range compression and contrast enhancement introduced in Secs. 2.2 and 2.3, respectively. The basic idea of the proposed scheme is to manipulate the DCT coefficients directly in the compressed domain where a modification of the dc coefficients by the function \( G(\cdot) \) performs the dynamic range compression while the modification of the ac coefficients by the function \( F(\cdot) \) performs the contrast enhancement. See Fig. 3 for the block diagram of the final proposed algorithm. The modified illumination component \( G(\tilde{L}) \) and reflectance component \( F(\tilde{R}) \) are combined in a multiplication to generate the final enhanced output image \( O \). In this section, we provide the details of the functions \( G(\cdot) \) and \( F(\cdot) \).

First, for dynamic range compression, we identify the dc component of the DCT coefficients as the illumination component of the image. Then, we apply the retinex theory in the compressed domain by adjusting the dc components to compress the dynamic range of the image. Adjustment of the dc components is equivalent to the modification of the illumination component by the function \( G(\cdot) \).

For improving image contrast, we identify the ac coefficients as the reflectance component of the image. We define a new measure \( \Phi_{w,n} \) that computes the spectral content of an image in the compressed domain. Based on the computed value of \( \Phi_{w,n} \), we modify the ac coefficients to enhance the contrast in the image. Adjustment of the ac coefficients is equivalent to the modification of the reflectance component by the function \( F(\cdot) \).

3.1 Dynamic Range Compression

The purpose of dynamic range compression is to map the natural dynamic range of a signal to a smaller range. This is achieved by modifying the illumination component of the image. In this paper, we observe that the illumination of the image varies smoothly over the entire image and thus the dc coefficients, the average value of each block, in the compressed domain are considered to represent the illumination component \( \tilde{I}_{x,y} \) in Eq. (6). To compress the dynamic range, the proposed algorithm manipulates the dc coefficients of the image and improves the details in very dark or bright regions of the input image in the compressed domain.

First, we compute the average value \( D_{0,0} \) of an \( 8 \times 8 \) block in the compressed domain as

\[
D_{0,0} = \frac{1}{8} \sum_{i=0}^{7} \sum_{j=0}^{7} I_{i,j},
\]

where \( d_{0,0} \) is the dc coefficient. Then, the illumination component \( \tilde{I}_{x,y} \) is modified to \( G(\tilde{L}) \) in dark regions of the image as

\[
G(\tilde{L}) = \frac{I_{\max}}{\left( \frac{D_{0,0}}{I_{\max}} \right)^{1/\gamma}},
\]

where \( I_{\max} \) is a maximum value of an input image (for example, 255 for an 8-bit input image) and \( \gamma \) is a user controlled variable that reflects the degree of enhancement desired by the user. The enhanced output \( O'_{x,y} \) in dark regions is then expressed as
where \( Q^{-1}[-] \) denotes the dequantization process and \( Q_{0,0} \) is the quantization parameter for the quantized dc coefficient \( \hat{d}_{0,0} \). Note that the illumination component is normalized to enhance details in dark area. Similarly, the enhanced output in bright regions is defined as

\[
O_{x,y}^b = \bar{R}_{x,y} I_{\max} \left[ 1 - \left( 1 - \frac{1}{8} \frac{Q_{0,0}}{I_{\max}} \right)^{1/\gamma} \right]. \tag{14}
\]

This computation is the same procedure for the inverse image of the input image. To consider details in both dark and bright regions, we combine the above two equations and generate the final output \( O_{x,y} \) as

\[
O_{x,y} = \frac{1}{2} (O_{x,y}^d + O_{x,y}^b) = \bar{R}_{x,y} Q^{-1}(\hat{d}_{0,0}). \tag{15}
\]

Here, \( \hat{d}_{0,0} \) is the enhanced dc coefficient.

An intermediate result of the proposed algorithm is shown in Fig. 4. For an explanatory convenience from now on, each subimage in, if any, a figure will be labeled from left to right and top to bottom directions. Figure 4(a) is the original image with a high dynamic range. Figure 4(b) is the resultant image after adjusting the dc coefficient with an enhancement value \( \gamma = 2 \). We can notice that details in dark and bright areas are brought out, especially around the river and on the right side of forest. However, Fig. 4(a) is still better than Fig. 4(b) in terms of contrast. Figure 4(b) also contains some block artifacts.

The characteristics of the mapping function \( G(\cdot) \) of the dc coefficient is shown in Fig. 5. Each curve indicates that the proposed algorithm compresses the image dynamic range by emphasizing the lower intensity values and suppressing the higher intensity values. The user-controlled enhancement value \( \gamma \) adjusts the degree of enhancement. As the enhancement value \( \gamma \) is increased, the pixels of the input image are saturated toward the center value, for example, 128 in an 8-bit input image. According to the characteristics of the mapping function, the range of the suitable enhancement value \( \gamma \) should be centered around 2. The value should be further adjusted according to the contents and brightness of the input image.

3.2 Contrast Enhancement

The image contrast is enhanced by modifying the ac coefficients of the image in the compressed domain. It is well known that the human visual system (HVS) is more sensitive to variation in low frequency components than high frequency components. To incorporate this property of HVS, we separate the DCT coefficients into two parts, as
shown in Fig. 1. The coefficients enclosed within the dotted lines are the low frequency components that are used for bringing out details in both dark and bright areas. These coefficients require a special weighting factor during the contrast enhancement process to reduce the block artifacts. The rest of the ac coefficients enclosed within the solid lines are manipulated using the relative spectral content measure $\Phi_{w,n}$.

A dc enhancement factor $\lambda$ is defined by using the ratio of the enhanced dc coefficient to the original dc coefficient as

$$\lambda = \frac{\delta + \tilde{d}_{0,0}}{\delta + \tilde{d}_{0,0}},$$  \hspace{1cm} (16)

where $\delta = 0.01$ is a small value to avoid a division by 0.

An ac contrast enhancement factor $\lambda_{w,n}$ is defined separately for different frequency bands as

$$\lambda_{w,n} = \begin{cases} 
\lambda_1 & \text{for } 1 \leq n \leq 2 \\
\lambda_2 & \text{for } 2 < n \leq 14.
\end{cases}$$  \hspace{1cm} (17)

The preceding definition enables different weighting factors to be used in the contrast enhancement for the ac coefficients in the low frequency bands, $n < 2$.

The enhanced measure of spectral content $\Phi_{w,n}$ is defined for the ac coefficients in the $n$th block of an image as

$$\Phi_{w,n} = \frac{\|\tilde{\Phi}_{w,n}\|^2}{\sum_{r=0}^{N} \tilde{N}(\Phi_{w,n})}$$
for $n > 1$,

(18)

with the ac enhancement factor $\lambda_{w,n}$ multiplied to the measure of spectral content $\Phi_{w,n}$ introduced in Sec. 2.3. By substituting the expression from Eq. (16), we write

$$\Phi_{w,n} = \frac{\|\tilde{\Phi}_{w,n}\|^2}{\|\tilde{\Phi}_{w,n}\|^2}$$

$$= \lambda_{w,n} \frac{\|\tilde{\Phi}_{w,n}\|^2}{\sum_{r=0}^{N} \tilde{N}(\Phi_{w,r})}$$
for $n > 1$.

(19)

After some manipulations of the terms, we obtain

$$\frac{\|\tilde{\Phi}_{w,n}\|^2}{\sum_{r=0}^{N} \tilde{N}(\Phi_{w,n})} = \lambda_{w,n} R_{w,n}$$
for $n > 1$.

(20)

The energy ratio $R_{w,n}$ of the enhanced ac coefficients is defined as

$$R_{w,n} = \frac{\|\tilde{\Phi}_{w,n}\|^2}{\|\Phi_{w,n}\|^2}$$
for $n > 1$.

(21)

The energy ratio $R_{w,n}$ is obtained by a recursive process since its value is the ratio between the enhanced energy sum and the original energy sum computed at all of the spectral bands below the $n$th spectral band.

Finally, the contrast enhanced DCT coefficients $\tilde{d}_{u,v}$ are computed from the original quantized coefficient $d_{u,v}$, the energy ratio $R_{w,n}$, and the ac enhancement factor $\lambda_{w,n}$ as

$$O_{x,y} = \text{DCT}^{-1}[Q^{-1}(\tilde{d}_{u,v})].$$

That is,
where \( Q_{w,n} \) represents the quantization parameter for the DCT coefficient \( d_{w,n} \).

We now discuss how to choose the value of the enhancement factor \( \lambda_{w,n} \). Figure 6 illustrates some comparative results using different values of the enhancement factor \( \lambda_{w,n} \). Figure 6(a) is an original input image compressed by JPEG using the quality factor \( Q_{75} \). In the figure, the right bottom part of the image is the enlarged area of the sold box in the middle of the image. The quality factor adjusts the quantization parameters so that the small value results in more compression. That is, \( Q_{x} = \begin{cases} 100 - x & \text{if } x > 50 \\ 50 & \text{otherwise} \end{cases} \) (27)

The commonly used quantization table \( Q_{50} \) is defined as

\[
Q_{50} = \begin{bmatrix}
16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\
12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\
14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\
14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\
18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\
24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\
49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\
72 & 92 & 95 & 98 & 112 & 100 & 103 & 99
\end{bmatrix}
\] (28)

We now discuss the choice of the ac enhancement value \( \lambda_{w,n} \). We first consider the case in which the ac coefficients are not separated into two partitions according to the frequency bands. Therefore, a single enhancement value \( \lambda_{w,n} = (\lambda_1, \lambda_2) = (\gamma, \gamma) \) is applied to the entire image. The result obtained with the enhancement value \( \gamma = 1.95 \) is shown in Fig. 6(b). It is clear that the resulting image contains a lot of block artifacts. This is because the low-frequency band coefficients are modified without considering the spectral contents of the image.

Figure 6(c) shows the result obtained from substituting a constant 1 into \( \lambda_1 \), i.e., \( \lambda_{w,n} = (1, 1) \). This means that the coefficients in the low-frequency bands, first and second bands, remain the same as the corresponding original DCT coefficients. Only the rest of the ac coefficients in higher frequency bands are modified. Figure 6(c) visually looks better than Fig. 6(b) since it keeps the low-frequency components unmodified, avoiding the generation of block artifacts. However, the result exhibits poor contrast since it is not adaptive to the variation of the brightness in the input image.

Figure 6(d) shows the result employing the dc enhancement factor \( \lambda \) defined in Eq. (16), i.e., \( \lambda_{w,n} = (1, \lambda) \). This makes sense with respect to reducing the block artifacts because the dc enhancement factor is smaller than \( \gamma \) and a larger enhancement factor \( (\lambda_{w,n}) \) tends to cause more block artifacts when applied to all of the ac coefficients. Some block artifacts still remain around the strong edges. Further information about the spectral content of a block must be taken into account to reduce these block effects while enhancing the details.

We study the energy distribution of the DCT coefficients to determine the proper value of the enhancement value \( \gamma \). The updated enhancement value \( \gamma_w \) for the \( w \)th block in the image is

\[
\gamma_w = 1 + (\gamma - 1)f(\alpha_w),
\] (29)

where \( \alpha_w \) is the energy ratio between the high-frequency bands and the low-frequency bands in the \( w \)th DCT block of an image defined as

\[
\alpha_w = \frac{\sum_{k=3}^{K} \|\psi_k\|^2}{\sum_{k=1}^{2} \sum_{k=3}^{K} N(\psi_k)},
\] (30)

and \( f(\cdot) \) is the weighting function defined as

\[
f(\chi) = \begin{cases} 0 & \text{for } \chi < \tau_1 \\ \frac{1}{\tau_2 - \tau_1} (\chi - \tau_1) & \text{for } \tau_1 \leq \chi \leq \tau_2 \\ 1 & \text{for } \chi > \tau_2 \end{cases}
\] (31)

with the first and second thresholds \( \tau_1 \) and \( \tau_2 \) as shown in Fig. 7. The new enhancement value \( \gamma_w \) enables us to enhance the details adaptively in high-frequency bands without affecting the strong edges in the low-frequency bands. Compared to Fig. 6(c), Fig. 9(a) in Sec. 3.3 contains a significantly reduced amount of block artifacts, especially around the clouds and near the top of a mountain.

![Fig. 7 Thresholds for computing the weighting function f(·).](image)

![Fig. 8 Four neighboring blocks and their enhancement values of the wth block.](image)
3.3 Block Artifacts Removal

The contrast enhancement process described so far operates on each $8 \times 8$ block of DCT coefficients independently of other neighboring blocks. This discontinuity of processing at the block boundaries becomes the source of block artifacts in the enhanced images. To deal with the block artifacts, we smooth the enhancement value $\gamma_w$ over the four neighboring blocks (left, upper right, upper, and upper left) of the image as shown in Fig. 8. Let the current enhancement value of the current block and its four neighbors be collected in $\mathbf{R}_w=[\gamma_w \, \gamma_{w-1} \, \gamma_{w-p+1} \, \gamma_{w-p} \, \gamma_{w-p-1}]^T$. Then, the average value $\overline{\gamma}_w$ is obtained by

$$\overline{\gamma}_w = \frac{\mathbf{1} \cdot \mathbf{R}_w}{N(\mathbf{R}_w)}, \quad (32)$$

where $\mathbf{1}$ is a $1 \times 5$ row vector of $1$ and $N(\mathbf{R}_w)$ is the number of elements in $\mathbf{R}_w$. The improvement in reducing block artifacts is shown in Fig. 9. Figure 9(a) shows the result using $\gamma_w$. A better result is shown in Fig. 9(b) using $\overline{\gamma}_w$. The enhanced image in Fig. 9(b) is more visually natural with less block artifacts. The modified enhancement values using the energy ratio ($\gamma_w$) and low-pass filter ($\overline{\gamma}_w$) are plotted in Fig. 9(c) and Fig. 9(d), respectively.
3.4 Computation Speed

The energy \( |\varphi_n|_F^2 \) at each frequency band indicates the amount of contribution from the corresponding frequency band. It is known that most of DCT coefficients in high-frequency bands go to zero after the quantization step. It is empirically verified that the energy of natural images is concentrated between the zeroth and eighth frequency bands, as indicated in Fig. 10. The horizontal axis shows the energy band \( n \), \( n = 0, \ldots, 14 \) and the vertical axis indicates the percentage ratio of the accumulated energy to the total energy. The compression domain methods, such as our proposed algorithm, are fast in terms of the computation speed because most of DCT coefficients in high frequency bands \( (n > 8) \) are zero. This property will be used to save computations and speed the proposed algorithm up in the following section.

We provide a simple example to verify this. An image \( I_{x,y} \) is split into \( k_1 \times k_2 \) DCT blocks of size \( 8 \times 8 \) and the energy of the \( n \)’th band in the \( w \)’th block is defined by

\[
E_{w,n} = \sum_{m+n=0} (\tilde{d}_{n,m,Q_{w,v}})^2.
\]

Then, a normalized and accumulated ratio up to the \( n \)’th frequency band over the entire image is expressed by

\[
\bar{\xi}_n = \frac{\sum_{t=0}^{n} \sum_{r=1}^{k_1} \sum_{w=1}^{k_2} E_{w,r}}{\sum_{t=0}^{n} \sum_{r=1}^{k_1} \sum_{w=1}^{k_2} E_{w,r}}.
\]

Each curve in Fig. 10 is the average ratio of 100 natural images with different quality factors \( Q_{35}, Q_{50}, Q_{75}, \) and \( Q_{95} \). DCT-based compression uses lossy compression techniques and allows control over the degree of “lossiness.” Lossy compression causes loss of image information. For JPEG, this usually means that the image becomes blurred.

The degree of blurring is controlled by the quality factor \( Q \), which is normally in the range from 50 to 95%. This is a measure of how much frequency domain information is preserved. The value of 75% is considered adequate for most images. Compression ratios of 30:1 to 50:1 are not unusual for JPEG although increasing compression corresponds to more image defects. It is important to remember that JPEG compression works well with natural images and badly with, for example, cartoon images. This is because the absence of sharp edges in the former tends to hide the effect of moderate JPEG compression. This should be considered when evaluating the performance of the enhancement techniques.

3.5 Proposed Algorithm

We present the steps of the proposed algorithm to compress dynamic range and enhance image contrast based on Retinex theory and a new measure of image spectral content.

1. We compute the averaged enhancement value \( \tilde{\gamma}_e \) using Eq. (32) for the \( w \)’th block of an image.
2. In the first band \( (n=0) \) of the \( w \)’th block, we compute the enhanced dc coefficient \( \tilde{d}_{0,0} \) using \( \tilde{\gamma}_e \) instead of the user input \( \gamma \) with Eq. (13) through Eq. (15). The ac enhancement value is obtained as \( \lambda_{w,n} = \{\lambda_1, \lambda_2\} \) = \{1, \tilde{\gamma}_e\} for the rest of DCT coefficients.
3. We use Eq. (21) to compute \( R_{w,n} \) in the next band \( (n=n+1) \).
4. We use Eq. (22) to obtain \( \tilde{d}_{u,v} \) for \( n = u + v \). Then, the enhanced set of spectral bands \( \tilde{\phi} \) corresponding to the original set of spectral bands \( \phi \) is updated by using Eq. (8) and Eq. (7), consecutively.
5. If \( R_{w,n} = R_{w,n-1} \) and \( n > 3 \), or \( n = 14 \), the algorithm goes to step 6. Otherwise, go to step 3.
6. If the current \( w \)’th block reaches the last block of the

Fig. 10 Accumulated energy distribution at frequency bands \( n = 0, \ldots, 14 \).
input image, terminate. Otherwise, go to step 1 with \( w = w + 1 \).

In step 2, \( \gamma \) is an image enhancement value that is chosen by the user. In practice, to enhance the details in dark and bright regions of the image, the value \( \gamma > 1 \) should be used. In step 5, the constrained energy ratio \( R_{w,n} \) is used to terminate the current block processing if the value is equivalent to the previous ratio \( R_{w,n-1} \).

Step 5, which compares the energy ratios at each band, is used to save the computation time of the proposed algorithm. That is, we don’t decode and compute all DCT coefficients. For instance, we can save about 50% of computations if the algorithm terminates at the frequency band \( n = 8 \).

4 Performance

In this section, we perform a series of experiments to demonstrate the performance of our proposed algorithm. We also compare the results with those generated by two other algorithms, histogram equalization algorithm and \( \alpha \)-rooting algorithm. A set of JPEG compressed images with high dynamic ranges are used in the experiments. These JPEG compressed images are available at (http://dragon.larc.nasa.gov/retinex/pao/news/). By default, the contrast enhancement variable \( \gamma \), the thresholds \( \tau_1 \) and \( \tau_2 \) in our proposed algorithm are set to 1.95, 0.10, and 1.95, respectively, for consistency over all of the tested images. Only the luminance component was used in this experiment to enhance the images.

4.1 Existing Methods

To evaluate the performance of the proposed algorithm, two other existing algorithms, histogram equalization, and \( \alpha \)-rooting algorithm, are implemented and compared against the proposed algorithm.

The histogram equalization technique provides a sophisticated method for modifying the dynamic range and contrast of an image. It alters the intensity histogram of the image such that the desired effects are obtained. In this paper, we modify the intensity histogram into a flat shape by setting all entries of the histogram to the same value. The histogram equalization technique is used in many image comparison processes because it performs well for detail enhancement and the correction of non-linear effects introduced by, e.g., a digitizer or display system.

The \( \alpha \)-rooting algorithm modifies the magnitude of each DCT coefficient by multiplying with a scaling parameter raised to the power of \( \alpha \), where \( \alpha \) is a positive real number. Let \( \hat{d}_{u,v} \), \( d_{u,v} \), and \( I_{\text{max}} \) be a DCT coefficient, a quantized DCT coefficient, and the maximum value of an input image, respectively. Then the modified DCT coefficient \( \tilde{d}_{u,v} \) is expressed as

\[
\tilde{d}_{u,v} = \hat{d}_{u,v} \left( \frac{d_{u,v}}{I_{\text{max}}} \right)^{-\alpha}.
\]

When \( \alpha < 1 \), the magnitude of smaller transform coefficients are enhanced by a greater degree than the larger transform coefficients. Since the coefficients are usually smaller in the high frequency bands, the \( \alpha \)-rooting algorithm effectively enhances the edges and details in the image. The major differences between \( \alpha \)-rooting algorithm and the proposed algorithm are that (1) \( \alpha \)-rooting algorithm does not modify dc coefficients, i.e., it does not enhance the details in dark and bright areas in the image; (2) the \( \alpha \)-rooting algorithm does not employ any content measure that examines the spectral content of the image while the proposed algorithm does. In our implementation, the value of \( \alpha \) is set to 0.9 and the value of \( I_{\text{max}} \) is set to 255, the maximum value for an 8-bit pixel.

4.2 Enhancement Quality Metrics

In this paper, we use three different quality metrics for quantitative evaluations of the image enhancement techniques. The amount of improvement after enhancement is often very difficult to measure objectively. In a subjective test, we ask the readers to determine if the enhanced images display more details and convey other useful information that are not clearly visible in the original images. However, for an objective evaluation, there is no universal quality measure that can effectively compare different enhancement methods. We introduce three different quality metrics. We provide a brief description of these metrics and we discuss the suitability of these metrics in interpreting our experimental results.

In Ref. 12, a quantitative measure of image enhancement, denoted as EME, is introduced. This measure is a modification of Weber’s law \( \frac{1}{2} \) and Fechner’s law. An image \( I_{xy} \) is first split into \( n_1 \times n_2 \) blocks \( w_{k,l} \), \( k = 1, \ldots, n_1 \) and \( l = 1, \ldots, n_2 \). Then, the difference between the maximum and minimum pixel values is computed within each block as a log ratio. Finally, the EME metric is computed for the entire image as the average of the differences. That is,

\[
\text{EME} = \frac{1}{n_1n_2} \sum_{k=1}^{n_1} \sum_{l=1}^{n_2} 20 \log \left( \frac{\max(w_{k,l})}{\min(w_{k,l})} \right),
\]

where \( \max(w_{k,l}) \) and \( \min(w_{k,l}) \) are the maximum and minimum pixel values in each block \( w_{k,l} \), respectively. The size of each block \( w_{k,l} \) in our experiments is set to \( 8 \times 8 \). The EME metric only gives an indication of the average degree of contrast present in the image. This metric is not an effective measure of the dynamic range compression or contrast enhancement in terms of localized details and human perception factors.

Watson’s metric uses a DCT for video decomposition into spatial channels. It is related to the DCTime metric that is developed for optimization of still image compression. Watson’s metric has lower computation requirements and has good correspondence with subjective quality tests. Watson’s metric computes the visibility of artifacts expressed in the DCT domain. Watson computes the metric as a product of temporal, spatial, and orientation functions. That is,

\[
T(u,v,w) = T_0 T_w(w) T_f(u,v) T_o(u,v),
\]

where \( T_0 \) is a global scaling factor (the remaining three functions have a unit peak gain), the temporal function \( T_w(w) \) is the inverse of the magnitude response of a first-order discrete infinite impulse response (IIR) low-pass filter, the spatial function \( T_f(u,v) \) is the inverse of Gaussian,
and the orientation function $T_o(u, v)$ takes into account a higher threshold for oblique frequencies and the imperfect visual summation between two component frequencies. Readers are referred to Ref. 23 for details. In our implementation, the threshold $T_0$ is assigned a value of 50.0 to properly scale the measured values in comparison to other metrics. Watson’s metric requires a reference image, the original, not-yet-enhanced image such as the peak signal-to-noise ratio (PSNR). A higher visual quality, measured by Watson’s metric, is indicated by a smaller number, whereas other metrics use a larger number to represent a higher visual quality. The score of Watson’s metric depends only on the difference between the original (reference) image and the enhanced image. A large metric value only indicates that there is a large difference between the source image and the target image. Again, Watson’s metric is not an accurate measure of the performance of the enhancement algorithms.

Wang and Bovic$^{18}$ introduce an objective quality measurement scheme that does not involve any reference images. To develop this metric, the authors have established a JPEG image database and conducted subjective experiments based on this database.

Wang’s metric considers block artifacts and blurring effects in the compressed images when measuring the enhancement quality. First, the block artifact $B$ is estimated as the average differences across all block boundaries. Second, the blurring effects are estimated by two activity measures $A$ and $Z$. Each of these factors is calculated separately in horizontal and vertical directions, and then averaged:

$$B = \frac{B_h + B_v}{2}, \quad A = \frac{A_h + A_v}{2}, \quad Z = \frac{Z_h + Z_v}{2},$$

Wang’s metric is computed as

$$S = \alpha + \beta B^{\gamma_1} A^{\gamma_2} Z^{\gamma_3},$$

where $\alpha = -245.9$, $\beta = 261.9$, $\gamma_1 = -0.0240$, $\gamma_2 = 0.0160$, and $\gamma_3 = 0.0064$ are the parameters with the values assigned by the authors. Wang’s quality metric reflects the characteristics of the image based on the human visual perception factors, such as block artifacts and blurring effects, and is the metric of our choice in this paper.

---

Fig. 11 Results of the proposed algorithm: the first column indicates original input images and the second column shows the processed results. From left to right and top to bottom: input images consist of (a) a high contrast image, (c) example under unbalanced illumination, (e) outdoor image taken in heavy sunlight, and (g) another outdoor image with high contrast under strong sunlight. (b), (d), (f), and (h) are the results of the corresponding input images.

Fig. 12 Some comparative results: first row original images; second row the result of histogram algorithm; third row the result image of $\alpha$-rooting algorithm; and fourth row the results of the proposed scheme.
4.3 Experimental Results

We present some experimental results of the proposed image enhancement algorithm. In Fig. 11, the images in the left column are the input images and the corresponding images in the right column are the results of the proposed algorithm. Figure 11(b) is the enhanced image bringing out details that are otherwise lost in dark areas of the high contrast original image in Fig. 11(a). The details of the trees on the top-left part of the original image are more clearly revealed. Also, many details on the tinted car window are brought out after the proposed processing. Figure 11(c) shows a large room with a light falloff from the left side to the right side of the image. A flashlight would have only made this image worse by increasing the brightness in the foreground only. Our proposed algorithm again improves the details in the dark side of the room as shown in Fig. 11(d). In Fig. 11(e), the original image is taken in a strong sunlight, causing many details to be lost in the shadows. After the proposed processing in Fig. 11(f), these details are brought out and clearly visible. In particular, the car parked under the trees, shown on the lower left corner of the image, is now much more recognizable. Finally, in Figs. 11(g) and 11(h), the details of the building, both the exterior parts and the interior part hidden in the shadows are greatly enhanced.

For quantitative performance evaluation of different algorithms, we perform a set of experiments using two images with high dynamic ranges. The original images are shown in Figs. 12(a) and 12(b). These images of indoor scenes suffer from a heavy light falloff [Fig. 12(a)] and hazing from the canopy reflections [Fig. 12(b)]. First, the histogram equalization algorithm is applied to these images. The results are shown in Figs. 12(c) and 12(d). The results of the $\alpha$-rooting algorithm are shown in Figs. 12(e) and 12(f). The value of $\alpha$ used in these experiments is 0.9. Finally, the results of our proposed algorithm are shown in Figs. 12(g) and 12(h). We now measure numerically the quality of enhancement by the three different algorithms, histogram equalization, $\alpha$-rooting, and the proposed algorithm. Table 1 shows the performance of these algorithms measured by the first two quality metrics, EME and Watson’s metrics. Figure 13 illustrates the results graphically. The scores listed in the first columns of the tables are computed for the enhanced images of the first test image in Fig. 12(a). The scores in the second columns are computed for the second test image in Fig. 12(b). The third columns of the tables are computed as the average scores of all test images used in this paper.

The EME metric gives higher scores to the proposed algorithm on average as shown in the third column of Table 1 (left). On the other hand, Watson’s metric gives the highest score to the $\alpha$-rooting algorithm. The proposed algorithm performs relatively well compared to the $\alpha$-rooting algorithm and superbly compared to the histogram equalization algorithm, based on Watson’s metric. As discussed earlier, the EME metric depends only on the brightness difference between minimum and maximum values in each block, and is not concerned with the overcontrasted or un-

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Image 1</th>
<th>Image 2</th>
<th>Avg. of all test images</th>
<th>Image 1</th>
<th>Image 2</th>
<th>Avg. of all test images</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$-rooting</td>
<td>15.3287</td>
<td>11.7237</td>
<td>18.9372</td>
<td>2.8343</td>
<td>2.8782</td>
<td>3.1342</td>
</tr>
<tr>
<td>The proposed</td>
<td>18.7310</td>
<td>12.6176</td>
<td>21.7235</td>
<td>5.4754</td>
<td>4.2247</td>
<td>8.7446</td>
</tr>
</tbody>
</table>

Fig. 13 Quality values versus quality metrics with different image enhancement methods; plots of (a) EME and (b) Watson’s metric.
decompressed images. Watson’s metric depends only on the difference between the original reference image and the enhanced image, and does not measure whether the enhanced version contains more visual information or not.

The scores by Wang’s metric are shown in Table 2 and graphically in Fig. 14. The scores by Wang’s metric indicate that our proposed algorithm outperforms the other two algorithms in all cases. These scores also agree with subjective evaluations of human observers. This result is expected because Wang’s metric considers human visual perception factors such as block artifacts and blurring effects when evaluating the quality of the enhancements.

5 Summary and Discussion

We presented an efficient image enhancement algorithm for improving details in dark and bright areas in the compressed domain. The proposed algorithm employed a basic concept of retinex theory and operated on the dc coefficients to compress the dynamic range of the image. To enhance contrast, the proposed algorithm modified the ac coefficients according to the new measure of spectral image content. The user-controlled enhancement variable adjusts the degree of enhancement to be achieved. To improve the processing speed, the distributional characteristics of the DCT coefficients were exploited as well. The experimental results showed that the proposed algorithm improved the image dynamic range and contrast effectively without causing block artifacts.

The performance of the proposed algorithm may be affected by the amount of noise present in the image. The contrast enhancement process that emphasizes the details in high frequency bands may boost the magnitudes of noise as well. If the noise level were known a priori in the compressed domain, our proposed algorithm could adjust the enhancement gain appropriately and avoid enhancing noise components. Unfortunately, there are no well-known methods to accurately estimate the noise level in the compression domain. To develop an algorithm that is more robust to noise, it is necessary to adjust the enhancement control variable $\gamma_w$ according to the noise level and the compression quality of the image. This will be the topic of our future work. Even though we described the $8 \times 8$ DCT block-based scheme, the basic ideas of the proposed approach, that use low-frequency bands for image dynamics and high-frequency bands for image contrast, can be easily extended to the different size kernels and transforms such as a fast Fourier transform (FFT) and even wavelets.

### Table 2 Wang’s metric.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Image 1</th>
<th>Image 2</th>
<th>Avg. of all test images</th>
</tr>
</thead>
<tbody>
<tr>
<td>Histogram</td>
<td>18.6339</td>
<td>18.1819</td>
<td>18.2116</td>
</tr>
<tr>
<td>$\alpha$-rooting</td>
<td>20.3156</td>
<td>20.3589</td>
<td>19.8161</td>
</tr>
<tr>
<td>The proposed</td>
<td>20.6273</td>
<td>21.0772</td>
<td>20.1350</td>
</tr>
</tbody>
</table>

**References**

Sangkeun Lee received his BS and MS degrees in electronic engineering from Chung-Ang University, Seoul, Korea, in 1996 and 1999, respectively, and his PhD degree in electrical and computer engineering from Georgia Institute of Technology, Atlanta, in 2003. In 2003 became senior research engineer in the Digital Media Solutions Lab, Samsung Information Systems America, Irvine, California, where he has been involved in the development of video processing and enhancement algorithms (DNLe) for Samsung's high-definition television (HDTV). His current research and development interests include digital video/image processing, especially video analysis/synthesis, denoising, compression for HDTV and multimedia applications, and content-based video representation and abstraction. He is a member of the IEEE.

Victor H. S. Ha received his BASc degree in computer engineering from the University of Toronto in 1995, his MEng degree in electrical engineering from Cornell University in 1996, and his PhD degree in electrical and computer engineering from Carnegie Mellon University in 2002. In 2002, he joined the Digital Media R&D Center of Samsung Electronics, Suwon, Korea, where he worked on software and hardware implementations of an MPEG-4 Part 10 AVC/H.264 video codec. Since 2004 he has been working in the Digital Media Solutions Lab., Samsung Information Systems America (Irvine, California). His research interests include image and video compression, image enhancement, pattern recognition, and shape theory. He is a member of IEEE, SPIE, and SID.

Yeong-Hwa Kim received his BS and MS degrees in statistics from Chung-Ang University, Seoul, Korea, in 1990 and 1992, respectively, and his PhD degree in statistics at the University of Florida, Gainesville, in 2000. He is currently an assistant professor with the Department of Statistics, Chung-Ang University, Seoul, Korea. His research interests are statistical image processing, including Bayesian estimations for various image processing applications including denoising, signal and noise classification, Bayesian edge detection, and statistical feature extraction.