Saliency guided naturalness enhancement in color images

Yuecheng Li\textsuperscript{a,b}, Hong Zhang\textsuperscript{a,∗}, Wenyan Jia\textsuperscript{b}, Ding Yuan\textsuperscript{a}, Feiyang Cheng\textsuperscript{a}, Ruiming Jia\textsuperscript{c},
Lu Li\textsuperscript{d}, Mingui Sun\textsuperscript{b}

\textsuperscript{a} Image Processing Center, Beihang University, Beijing 100191, China
\textsuperscript{b} Department of Neurosurgery, University of Pittsburgh, PA 15213, USA
\textsuperscript{c} School of Information Engineering, North China University of Technology, Beijing, China
\textsuperscript{d} Department of Information Technology, Agricultural Bank of China, Beijing, China

\section*{ARTICLE INFO}

\textbf{Article history:}
Received 7 July 2014
Accepted 27 July 2015

\textbf{Keywords:}
Image enhancement
Human visual system
Naturalness
Perceptual contrast

\section*{ABSTRACT}

Recent technological advances have enabled smartphones, tablet computers and wearable devices to acquire digital images conveniently. Thus, taking pictures, enjoying them, and using them in a variety of applications (e.g., social networking) have become an important part of technology based lifestyle for many people. As a result, efficient algorithms for image enhancement play an increasing role for improving image quality for both human visualization and computer image analysis. In this paper, we learn from the human visual system (HVS) which has a powerful biological mechanism to optimize scene perception in complex and varying illumination conditions. Using this knowledge, we combine the concept of popular Retinex and the histogram equalization (HE) and propose an efficient image naturalness enhancement algorithm for both non-uniform and low light images. Our algorithm emphasizes the perceptual contrast while reducing halo artifacts, detail-clipping effects, and over-enhancements. An input lightness is computed based on the white-patch assumption. A Retinex based lightness correction is utilized to preserve and improve the contrast for bright regions. The halo artifacts are suppressed based on Bilateral Filtering (BF). The detail-clipping effect is reduced by imposing certain constraints. For dark regions in input images, a perceptual contrast enhancement method with dynamic range adjustment is adopted which enhances perceptual contrast and adjusts dynamic range without over-enhancements. Finally, a weighting combination guided by a simplified saliency map is employed to blend results of enhancements. Our method produces more natural image appearance and better perceptual contrast in comparison with a number of state-of-the-art algorithms. Besides, the proposed algorithm is relatively simple and efficient in computation, suitable for implementation in mobile and wearable devices to support a variety of practical applications.

© 2015 Published by Elsevier GmbH.

\section{1. Introduction}

With the recent advances in smartphones, tablet computers and wearable devices, digital images have been widely utilized for documenting life experiences, enjoyments as a form of visual arts, and sharing with people in a variety of social networks. In addition, the computer image processing technology has attracted growing attentions for applications in different fields. However, despite the improved quality of digital cameras in mobile devices, the man-made imaging system still cannot compete with the human visual system (HVS) which is well known to have outstanding abilities in lightness/color adaptation and visual content discrimination under complex natural environments. The man-made imaging system is affected significantly by a number of degrading factors, such as a limited dynamic range, an undesirable or insufficient ambient lighting, and improper weather conditions. Degraded images may have a low contrast, a dim color and dark/bright regions that hide details of important contents. For example, Fig. 1(a) shows an example of image degradation due to unfavorable lighting which reduces both visual satisfaction in observation and the effectiveness of feature extraction in computer analysis. Therefore, image enhancement plays a critical role in improving image quality for modern ubiquitous mobile devices and facilitating higher-level image analysis and interpretation.

In general, it is a difficult problem to both effectively enhance an image and preserve its naturalness without producing noticeable artifacts and amplifying noise. Different from the image restoration algorithms, where a degradation model is fully considered [1],
image enhancement algorithms usually only takes into account the visual characteristics of degraded images regardless of the specific causes of degradation. For non-uniform images (e.g., Fig. 1(a)), multiple visual characteristics can be observed. Because of under-exposure, dark regions often hide details and also contain artifacts and noise; mid-tone gray occupies very narrow dynamic range and lacks sharpness though they are relatively exposed well; bright regions are composed of obvious details against high-intensity background and some saturated regions; the overall color is dim due to improper exposure.

In biological perception systems, such as the one in humans, the problems stated previously usually produce less visual negative effects. The human visual system (HVS) is a complex sensory system dominated by a neural muscular ocular control within the eye, a sophisticated retinal perceptive organization, and parallel processing, feedback, feedback and lateral connections in the visual cortex [2]. Its dynamic neural network selects attended locations in order of decreasing saliency so as to break down the complex problem of scene understanding by rapidly selecting conspicuous locations to be analyzed in detail in a computationally efficient manner [3]. This represents the dynamic routing of attention model and means that intermediate and higher visual processes appear to select a subset of the available sensory information before further processing to reduce the complexity of scene analysis [4]. All those features enable the HVS to interpret complex scenes in real time and achieve optimal visual control in varying contrast regions and simultaneously enhances the hidden details.

So, we believe that the HVS has a powerful adaptability of separately processing salient regions with a good contrast before adjusting itself to differentiate perceptual contrast in degraded regions. The most salient regions in the scene attract the attention of HVS first because they usually contain the information that is easier to process. Then, the HVS performs perceptual contrast extraction for non-salient regions (e.g. dark regions that possibly contain useful details). Motivated by this, for non-uniform and low-light images, we believe it is necessary and more consistent to separately enhance different regions in images according to saliency.

Thus, we propose a naturalness enhancement algorithm to effectively enhance image contrast with a low computational load while avoiding halo artifacts, loss of details, and over-enhancement. The proposed method learns from the HVS by applying different emphasis for dark and bright regions guided by image content saliency. Specifically, a lightness correction is utilized in bright regions to preserve details. And, a perceptual contrast enhancement with a dynamic range adjustment is adopted for dark regions to improve perceptual contrast. Finally, the results are blended being weighted by a saliency map. Taking advantage of certain characteristics of degraded images, we simplify saliency detection using a by-product in improving perceptual contrast so that the saliency map is obtained without additional computation.

The rest of this paper is organized as follows. We review the related literature in Section 2. In Section 3, our algorithm is detailed in four subsections. We then present experimental results in Section 4 with both subjective and objective comparisons between our method and several forms of commonly used histogram equalization and Retinex methods. A comparison on computational cost is also presented to demonstrate the efficiency of the proposed algorithm. Finally, we conclude this paper and discuss the presented methods in Section 5.

2. Related work

Over decades, many efforts have been made to develop enhancement methods to recover visual information in degraded images. In this section, we briefly review several state-of-the-art methods. Interested readers are referred to [5] for a more extensive overview about image enhancement techniques.

The biological mechanisms of the HVS have been considered in developing an enhancement model [6]. A widely adopted concept for enhancing non-uniform images based on the HVS is the Retinex theory. It is originally introduced by Land and McCann [7] based on the color constancy phenomenon with random path [7] and Center/Surround models [8]. Center/Surround Retinex decomposes a given image into a reflectance and an illumination images and is the most effective method to calculate relative reflectance although there are scenes that violate its gray-world assumption [9]. Based on the Center/Surround model and the gray-world assumption, Jobson et al. [10,11] proposed a single-scale Retinex (SSR), multi-scale Retinex (MSR), and MSR with color correction (MSRCR) in which the reflectance is independently considered as the enhancement result. The illumination estimation on multi-scale results in a trade-off between detail enhancement and natural visual appearance. However, the adopted isotropic Gaussian kernel results in halo artifacts as observed in Fig. 1(b). For suppressing halo artifacts and obtaining a more natural visual perception, several researchers used edge preserving filters for illumination estimation, e.g. Bilateral Filtering (BF) [12,13] and Adaptive Filtering (AF) [14].

Although halo artifact suppression is considered for preserving naturalness, the non-constraint illumination estimation by BF and AF do not promise correct reflectance which should be no larger than one. Kimmel et al. [15] proposed a variational framework for Retinex where the illumination estimation was achieved through minimizing a cost function with the constraint that estimated illumination is not smaller than image lightness. This constraint results in the correct reflectance thus help preserve image details. Based on this framework, [16-18] modified the coefficients in the cost function for illumination estimation based on gradient amplitude. It enables original edges to be preserved so that halo artifacts are also prevented when image details are preserved. Besides computationally expensive, due to unpredictable lightness order error the variational based methods cannot adequately preserve visual naturalness although image details are well preserved and halo artifacts are successfully suppressed. And in some versions of these methods, a simple mapping is used for illumination correction which cannot promise natural results for images taken in complex light conditions.

Seeking naturalness, Tao and Asari [19] proposed to firstly compress the image dynamic range by using a nonlinear inverse sigmoid function to unveil details in dark and bright regions, and then adopt the center/surround model based on local contrast
enhancement on a small scale. Shin [20] proposed an integrated global and local processing method for dynamic range compression with weighted combination of global dynamic range adjustment and local contrast improvement. Chen [21] proposed a natural enhancement algorithm for color image (NECI). Comparable natural lightness appearance is achieved with this method. However, loss of details happens for non-uniform scenes which are not consistent with the HVS and this method is only applicable to offline processing due to its computational complexity [21]. Recently, Wang et al. [22] propose a balanced color contrast enhancement method which can improve the visual quality of both highlights and dark areas. Although this method can effective improve the visual quality for both bright and dark regions and realize vivid color, mid-tone are still sacrificed which results insufficient global contrast and flatten appearance. Furthermore, different color spaces are used in this method which involves multiple conversions. Wang et al. [23] proposed a naturalness preserved enhancement method for non-uniform images by defining a bright pass filter (BPF) for constraint illumination estimation. They also proposed lightness-order-error (LOE) to access naturalness preservation objectively. From their experiments, naturalness are preserved well for certain non-uniform scenes. But obvious artifacts and noises are produced when there are extremely dark regions in images according to our extensive experiments as shown in Section 4. This may be due to incorrect illumination estimation by BPF for dark regions.

These methods are similar with Retinex in spirit. Different from MSR which processing color channels simultaneously, those variations (except Wang Y.F.) usually processing the lightness only which saves tremendous computation. Also they apply different strategies for dynamic range compression and local contrast enhancement respectively. Instead of illumination estimation in large scale to preserve naturalness, certain nonlinear correction is designed to compress dynamic range preliminarily. And instead of using reflectance directly, Center/Surround concept is used in a different way to manipulate local contrast. However, complex structure and parameter tuning are the common disadvantages limiting their adaptability.

Histogram equalization (HE) is also a classical enhancement technique with low computational cost [24]. However, due to problems of brightness variation, washed-out details and over-enhancement, the global HE is not usually applied to images directly. Variations of the standard HE algorithm have been developed for practical applications. In the first class of HE algorithms focusing on preserving the input brightness, input histogram is separated into two or more sections before histogram equalization is applied to each section [25–28] or the target histogram is found before histogram projection [29]. Although brightness can be preserved by those algorithms, image contrast is poorly enhanced. The second class of HE algorithms performs equalization on input image locally to improve contrast, such as non-overlapped block histogram equalization (NOBHE) [24], block overlapped histogram equalization (BOHE) [24], and partially overlapped sub-block histogram equalization (POBHE) [30], etc. Although these algorithms can improve local contrast significantly, their over-enhancement may lead to excessive noise and result in unnatural image. The third class of HE algorithms is based on histogram modification to realize over-enhancement suppression. They can improve contrasts in specified dynamic ranges and control the amplification of noise and the enhancement rate by sacrificing a certain amount of contrast. In this class, contrast limited adaptive histogram equalization (AHE) [31] is typically proposed and processes image locally. However, the output image may contain unexpected artifacts due to sub-block processing as illustrated in Fig. 1(c). Recently, Rivera et al. propose a content-aware method based on HE which adopts contrast pairs to emphasize local structures and suppress noise [32] and results in natural look, but for this method image details are rarely enhanced.

So, with HE technique only, though simple, naturalness preservation and contrast enhancement are hard to leverage for non-uniform images. This can also be told from the commonly adopted criterions by those HE based algorithms, e.g., Absolute Mean Brightness Error (AMBE) and Entropy, which are often poorly correlated to human visual perception (HVP) [33]. In the biological visual system, regardless of the often tremendous range of brightness variation over the day or from place to place, the human visual system (HVS) can perceive objects effectively. One of the primary reasons is that the HVS is more sensitive to low brightness than brightness [34]. Therefore, in order to improve correlation to the HVS, edge sharpness is often more crucial than brightness enhancement which may cause over-enhancement and noise amplification. Recently, the adaptive height-modified histogram equalization (AHMHE) [35] is developed. It upgrades the sharpness of the backlight image for perception by emphasizing edges and gives vivid color. However, the perceptual contrast is insufficiently represented, and details in bright regions are still washed-out. Together with our work about a modified clipped histogram equalization (CHE) method [36], we previously proposed a perceptual contrast enhancement method which can effectively enhance perceptual contrast for both dark and bright regions [37]. But it also washes out small details in bright regions due to histogram combination.

In this paper, by combining the visual attention model with our perceptual contrast enhancement method, we derive a more robust algorithm for non-uniform and low light images to improve contrast and preserve naturalness. It inherits all the merits of both Retinex and HE based algorithms but avoids their demerits, i.e. halo artifacts, details clipping effects, and over enhancement, as shown in Fig. 1(d). Parameter automation and low computation cost are also the advantages of the proposed method for practical applications.

3. Methods

The flow diagram of the proposed algorithm is shown in Fig. 2. In order to enhance contrast in bright regions (upper route in Fig. 2), after an edge-preserving illumination estimation, a simple tone mapping function is proposed based on the Retinex concept. This mapping also suppresses halo artifacts and clipping effects. For dark regions (bottom route in Fig. 2), perceptual contrast extraction and noise suppression are performed by conditioning the source lightness with the perceptual contrast map. Then, the dynamic range of the conditioned image is adjusted so that details in dark regions can be unveiled with enhanced edges. Although bright regions also get enhanced, this procedure inevitably washes out small details in bright regions due to histogram merging. However, by blending the two results of routes based on a saliency map weighting, the enhanced lightness at the output provides contrast in both bright and dark regions while preserving natural appearance.

From our extensive experiments, the proposed algorithm shows great robustness and outperforms many state-of-the-art methods that are based on histogram equalization (HE) or Retinex, especially in terms of naturalness preservation and computation cost.
Compared to those involved state-of-the-art methods, our proposed algorithm has several advantages:

1. Details are well preserved for bright regions without clipping effects;
2. Perceptual contrast are improved for dark regions;
3. Natural appearance is achieved with corrected color, i.e., halo artifacts, details clipping effects, and over-enhancement are all suppressed and prevented;
4. Computation cost is low with parameters automation;

In the following, we provide details of the major function blocks in Fig. 2.

3.1. Source lightness

In existing algorithms, usually only the lightness component of the input image is processed. Then, the color image at the output is constructed based on the processed lightness component. In our method, we obtain the lightness component by

\[ L(x, y) = \max_{c \in \{r, g, b\}} X^c(x, y), \tag{1} \]

where \( X^c(x, y), c \in \{r, g, b\} \) represents the input color image. Eq. (1) is consistent with the assumption of white-patch world in color constancy. Based on our experiments, Eq. (1) produces more satisfactory white balance than other color space converting methods. In addition, it is effective and simple, preserving contrast in the original image while separating color information.

3.2. Detail-preserving enhancement (upper route)

For enhancing contrast in bright regions, a simple detail-preserving tune mapping function is proposed to enhance the brightness regions based on the Retinex concept. In natural images, most details are usually well exposed which are salient to HVS. However, it is usually difficult to retrieve useful information from saturated regions, and the remaining details are of small amplitude comparable to the noise level. Hence, the target for brightness enhancement is to improve local contrast naturally without introducing halo artifacts and sacrificing details.

In order to preserve details in bright regions, we propose a gain factor at each pixel \((i, j)\):

\[ p(i, j) = 0.1 \times \sqrt{\text{std}(i, j)_{5 \times 5}}, \tag{2} \]

where \(\text{std}(i, j)_{5 \times 5}\) is the local standard deviation of a 5 \(\times\) 5 neighbor window centered at \((i, j)\) in the lightness image \(L(x, y)\) which is produced by Eq. (1). Due to the characteristics stated above, this detail gain factor is effective in representing the density of local details. As an example, the detail gain factor for Fig. 1 is shown in Fig. 3(b).

Then, the scene reflectance in the Retinex model provides image details as

\[ R = \frac{L}{(L + c)}, \tag{3} \]

where \(L\) is the illumination component estimated by BF so that halo artifacts are suppressed and \(c\) is a small constant (e.g. 0.01) for avoiding division by zero.

A brute force implementation of BF is computational cost because BF cannot be accelerated with Fast Fourier Transform as other convolutions (e.g. Gaussian Convolution). There are many reports about accelerating BF computation. The fast BF in [39] is implemented in our experiments. In that work, BF is accelerated by using a piecewise-linear approximation in the intensity domain and appropriate subsampling in the spatial domain. Parameters are automatically determined as \(\sigma_x = \left[\min(h, w)/16\right]\) for spatial kernel and \(\sigma_r = \left[(L_{\text{max}} - L_{\text{min}})/5\right]\) for intensity influence where \(h\) and \(w\) are height and width of image, \(L_{\text{max}}\) and \(L_{\text{min}}\) are maximum and minimum of lightness, and \([x]\) means the largest integer smaller than \(x\).

With detail gain factor, details (high frequency components) get emphasized as \(D_{\text{en}} = R^p\).

However, the result with emphasized details is not taken as the final result directly otherwise halo artifacts and detail clipping effects are obvious. We propose to apply a tune mapping function with \(D_{\text{en}}\) for improving the contrast of bright regions as

\[ L_{\text{bright-en}} = L D_{\text{en}}. \tag{5} \]

Similar strategies are adopted by other Retinex variations mentioned in Section 2, the difference for our proposed method is details clipping effects are prevented with effective constraints. As we mentioned in Retinex model, estimated illumination \(L\) should be larger or equal to \(L\) so that the \(R \leq 1\), otherwise details (high values) will be clipped in the result. In other reports, a constraint term is added while estimating illumination [15] or only the brighter pixels are considered [23]. Instead, in our method, the prevention of clipping effects is accomplished simply by two strategies. The first is that the maximum value in color channels is used as image lightness which represents the strongest sub-band radiation. Secondly, we further simply restrict \(D_{\text{en}}\) under 1 which enables Eq. (5) preserve high values without modifying them, but stretch lower values away from high values with smaller \(D_{\text{en}}\) to improve local contrast. Through this two steps, fine details in bright regions are well preserved. As shown in Fig. 3(c), for regions full of high frequency components, the tune mapping factor is constrained under 1. The constrained tune mapping factor guarantees fine details in bright regions (mountains in the far) can be preserved when contrast of bright regions and mid-tone range are improved as illustrated in Fig. 3(d). The dark regions are left unaltered in this details preserving contrast enhancement procedure due to low lightness.

3.3. Auto perceptual contrast enhancement (bottom route)

A general computational principle in the retina, lateral geniculate nucleus, and primary visual cortex, is the so-called center-surround which akin to visual receptive fields [40] and can also be implemented as the difference between fine and coarse scales. In our previous work [37], we proposed a perceptual contrast map (PCM) based on the modified difference of Gaussian (DOG)
to extract edge contents in the input image for enhancement. As the DOG model functions as a band-pass filter, the amplitude of the PCM represents the strength of information in the frequency band of interest. Hence, by optimizing the choices of parameters, it becomes feasible to construct a PCM that highlights perceptually important local edges while suppressing high frequency noise in smooth regions. For a complete content here, we briefly describe our previous work in this subsection with illustration of the proposed parameters automation.

In order to construct the PCM, the central component (in a \((2r_c+1) \times (2r_c+1)\) central mask) \(R_c(x, y)\) and the surrounding component (in a \((2r_s+1) \times (2r_s+1)\) surrounding mask) \(R_s(x, y)\) with the midpoint of the receptive-field being at \((x, y)\) of the input image are respectively computed by

\[
R_c(x, y) = \sum_{i=x-r_c}^{x+r_c} \sum_{j=y-r_c}^{y+r_c} \text{Center}(i-x, j-y)I(i, j), \tag{6}
\]

\[
R_s(x, y) = \sum_{i=x-r_s}^{x+r_s} \sum_{j=y-r_s}^{y+r_s} \text{Surround}(i-x, j-y)I(i, j), \tag{7}
\]

where \(I(x, y)\) is the input image, and \(\text{Center}(x, y)\) and \(\text{Surround}(x, y)\) are two Gaussian functions:

\[
\text{Center}(x, y) = \exp \left[ -\frac{3x^2}{r_c} - \frac{3y^2}{r_c^2} \right], \tag{8}
\]

\[
\text{Surround}(x, y) = 0.85 \frac{r_c}{r_s} \exp \left[ -\frac{3x^2}{r_s} - \frac{3y^2}{r_s^2} \right]. \tag{9}
\]

The input image is conditioned by

\[
S(x, y) = \frac{I(x, y)}{\text{PCM}(x, y)}, \tag{10}
\]

where \(\text{PCM}\) is defined using the ratio form, i.e.

\[
\text{PCM}(x, y) = \frac{R_s(x, y) - R_c(x, y)}{R_c(x, y)}. \tag{11}
\]

The ratio form of Eq. (11) allows a control of DOG output by the mean luminance of the surrounding component of the receptive field. In other words, because the DOG is scaled by the mean luminance, the contrast output becomes independent from the luminance. As a result, the contrast information in both dark and bright regions can be extracted. Besides this scaling property, the relative DOG model in Eq. (11) has another important property as a normalized band-pass filter with two cut-off frequencies determined by \(r_c\) (high-pass cut off) and \(r_s\) (low pass cut-off) \((r_c < r_s)\). With suitable choices of \(r_c\) and \(r_s\) values, the pre-conditioned image \(S(x, y)\) preserves perceptual contrast while suppressing noise in smooth areas. Though \(r_c\) and \(r_s\) are related to image size, extensive experiments showed that acceptable performance can be obtained by setting \(r_c = 1\) and \(r_s = 0.01 \times \min(h, w)\) with further limited in [3, 9].

Because the calculation in Eq. (10) may result in an overflow beyond the dynamic range \([X_{\text{min}} \cdot X_{\text{max}}]\) in \(S\), before the subsequent dynamic range adjustment, \(S\) is normalized to \(S_{\text{nor}}\) with the maximum and minimum values of \(S\). Then, our modified CHE algorithm [36] is conducted on \(S_{\text{nor}}\) to realize dynamic range adjustment.

There is a parameter \(T\) in our modified CHE algorithm to control the enhancement rate for the output image. As an extremely case, \(T = 0.5\) will make the clipping threshold \(T_c \approx 0\) and give a modified histogram with most bins having nearly the same height, resulting in an output without dynamic range adjustment. And \(T = 0\) will result in an output image with a similar dynamic range adjustment provided by the standard HE. Here, we propose to automatically determine \(T\) according to the lightness distribution of original image. In general, visually optimized images are more tightly clustered about a single mean value and have much higher standard deviations [41]. It further supports the idea that visual optimization centers the data mean on the mid-point of the image dynamic range (e.g. 128 for 8-bit) and spreads the signal excursions out across the dynamic range to a maximal extent while at the same time limiting any over- and under-shoots spatially. Consequently, the automatic selection of \(T\) is simply derived as

\[
T = 0.5 \left( 1 - \frac{L_{\text{mean}} - 128}{128} \right), \tag{12}
\]

with further restriction of \(T \geq 0.05\) to avoid over-enhancement. \(L_{\text{mean}}\) is the mean value of input lightness.

With automatically selected \(T\), we can obtain a result image, \(L_{\text{perceptual.en}}\), with improved perceptual contrast. Compared to the result of previous details preserving enhancement, \(L_{\text{perceptual.en}}\) provides improved perceptual contrast for dark regions but may cause washout for small details in bright regions. To obtain an optimized result, as illustrated in next subsection, it is proposed to blend \(L_{\text{perceptual.en}}\) with \(L_{\text{bright.en}}\) under saliency guide to generate natural result image with improved contrast.

### 3.4. Saliency guided blending

The human visual system can process enormous visual data instantly. Many computational models try to achieve such capabilities in different ways. Among various visual mechanisms the visual saliency is a key component and can be used as a prior for other components to provide an efficient approximate solution to more difficult problems. Computational saliency with respect to the human fixation is a central issue for these applications.

Various algorithms have been presented to produce saliency maps of images [3,42–44]. Some saliency detection methods attempt to find rules for combining low-level information (e.g. color, texture, and contrast) in an image based on the motivation of HVS. Other learning-based methods treat this as a classification or regression problem, in which training data are necessary to learn the mapping from image features to saliency levels. We refer readers to [45] for more details about saliency detection problem.

For different applications, the production of saliency map is highly relevant to the interested features. The Graph Based Visual Saliency (GBVS) model proposed by Harel et al. [42] is applied in our proposed method to extract the saliency map, \(S_{\text{sal}}\), from original image. It is a bottom-up approach that uses graph and dissimilarly measure to construct the saliency model. We trust the salient regions are regions containing fine details in well exposed regions and propose to utilize the following saliency guided blending equation for combining \(L_{\text{bright.en}}\) and \(L_{\text{perceptual.en}}\):

\[
L_{\text{en}}(x, y) = S_{\text{sal}nor}(x, y) \cdot L_{\text{bright.en}} + (1 - S_{\text{sal}nor}(x, y)) \cdot L_{\text{perceptual.en}}, \tag{13}
\]

where \(S_{\text{sal}nor}\) is the normalized Saliency weight with range [0,1]. It implies that salient regions in \(L_{\text{en}}\) are enhanced with our details preserving enhancement and other regions are particularly enhanced with perceptual contrast.

Since color is dim and contrast is low for degraded images, we highly trust brightness plays an important role in the visual attention of HVS in this situation. Through our extensive test, it is found that a simplified saliency-weight can be replaced by the surrounding component obtained by Eq. (9). Compared to the saliency map by GBVS which puts over strong emphasis on bright regions, surrounding component performs better balance for blending our two enhanced results. Consequently, with such simplification, Eq. (13) gives promising blending result without additional computation cost.
After final enhanced lightness, \( L_{en} \), is carried out, the result color image of the proposed method is obtained with color correction in YUV space as

\[
\begin{align*}
    r &= \frac{L_{en}}{L} \\
    Y_{out} &= r \cdot (Y_{in} - 16) + 16 \\
    U_{out} &= r \cdot \alpha \cdot (U_{in} - 128) + 128 \\
    V_{out} &= r \cdot \alpha \cdot (V_{in} - 128) + 128
\end{align*}
\]

where, \( L \) is the source lightness from Eq. (1), \( \{Y_{in}, U_{in}, V_{in}\} \) and \( \{Y_{out}, U_{out}, V_{out}\} \) indicate input and output images in YUV space respectively, and \( \alpha \) is the color correction factor which is set to 1–2 with 1.8 as the default in our following experiments.

4. Experiments and analysis

To demonstrate the performance of the proposed algorithm, extensive experiments have been performed with typical images of different sizes from 720 × 480 to 3264 × 2448 in the format of BMP or JPEG under different ambient light conditions. For comparison, we select total eight methods from both HE based and Retinex based categories, i.e. AHE [31], AHMHE [39], a recently proposed algorithm, Rivera [32], our previous perceptual contrast enhancement method, Zhang [37], MSRCR [11], AINDANE [19], and recently proposed methods, Y. Wang [22] and S. Wang [23].

Shown in groups, we first subjectively compare our proposed algorithm with HE based algorithms which verifies that the proposed method can avoid over-enhancement while enhancing the perceptual contrast in dark regions. Then we subjectively compare our proposed method with the Retinex based algorithms which illustrates that our proposed method outperforms those algorithms in terms of naturalness with the ability of preserving the details in bright regions. Also a comprehensive comparison is shown with an extremely under-exposed image to illustrate the robustness of our proposed algorithm on naturalness preservation and contrast enhancement. To evaluate performance objectively, Lightness-Order-Error (LOE) [23] is adopted to assess naturalness preservation performance, and a No-Reference Image Quality Assessment (SFLNIA) [40] is further used to measure the degradations in images. Finally, computation comparison shows the efficiency of our proposed algorithm.

4.1. Subjective assessments

Figs. 4 and 5 show different results by AHE, AHMHE, Rivera, and the proposed algorithm respectively. The indoor images in Fig. 4 are under exposed (desk and shelf are hidden in dark). And the original image in Fig. 5 represents a high dynamic range and contains details in bright (castle against the sky) and dark (people in the shadow). As we can see from Fig. 4, AHE causes some artifacts on the dark wall and Venetian blinds due to over-enhancement. Although AHMHE can effectively enhance details in dark regions...
and gives vivid color, the details are washed out in bright regions (twigs in sky region). Rivera preserves good naturalness but cannot unveil image details. Zhang enhances the perceptual details in dark regions and gives sharp edges, but also causes over-enhancement for the bright region. Contrarily, our proposed method not only improves contrast in both dark and bright but also gives sharper edges and vivid color, and also does not over-enhance the bright sky region and thus preserves details. As shown in Fig. 5, the AHE still results in over-enhancement (black artifacts on the floor). For AHMHE, the over-enhancement is obvious in the bright sky region. Rivera cannot unveil details in the shadow. Zhang also results in over-enhancement for the sky. But the proposed algorithm effectively compresses the dynamic range with improved contrast and preserved naturalness, thus gives no over-enhancement or artifacts.

Other two sets of results are shown in Figs. 6 and 7 for comparing the performance on non-uniform images with MSRCR, AINDANE, Y. Wang, and S. Wang. It can be seen that the output of MSRCR shows obvious halo artifacts (the edge of trunk in Fig. 6) and detail loss in bright regions (mountains in the distance in Fig. 6 and rocks above the stairs in Fig. 7) because of Gaussian-based non-constraint estimation. AINDANE also sacrifices the details in bright regions although halo artifacts are prevented. Y. Wang avoids halo artifacts and results in natural appearance while unveiling details in both bright and dark regions, but it poorly enhances perceptual contrast and shows less sharp edges. As we mentioned previously, S. Wang suffers from even more halo artifacts due to incorrect illumination estimation by its BPF for dark regions, washes out details in bright regions, and gives uncomfortable visual appearance. However, the proposed algorithm gives more natural look with enhanced perceptual edges.

To further test the robustness of our proposed method, both those HE and Retinex based algorithms and the proposed algorithm are conducted on extremely low light images. One example is shown in Fig. 8 with results by AINDANE, Y. Wang, Rivera, and Zhang for their better performance than other methods in their own classes. As we can see, both AINDANE and Y. Wang amplify noise which contaminates enhanced details. Rivera, again, has no effects on details enhancement. And Zhang also amplifies noise and causes severe over-enhancement in sky region. However, the proposed algorithm provides sharp and clear details and gives more natural result by preventing noise amplification and suppressing over-enhancement.

4.2. Objective assessments

Lightness-Order-Error (LOE) is proposed and firstly adopted to quantitatively measure the performance of naturalness preservation in [23]. From the definition of LOE, the smaller the LOE value is, the better the lightness order is preserved which means more natural pleasing appearance is obtained [23]. The LOE values for the results of selected 34 typical images are depicted in Fig. 9 with average values listed in Table 1. We should note although other three algorithms get similar or even lower LOE value, they have

| Table 1 Average LOE, DMOS, and Execution time of different methods. |
|-----------------------|---------------------|-------------------|-----------------|
| Methods               | LOE                | DMOS              | Execution time (s) |
| Original              | 0                  | 5.1624            | 0.017            |
| MSRCR                 | 721.73             | 42.54             | 2.197            |
| AINDANE               | 428.47             | 40.332            | 0.861            |
| Y. Wang               | 229.50             | 40.728            | 3.431            |
| S. Wang               | 432.20             | 42.396            | 49.807           |
| AHE                   | 410.23             | 39.264            | 0.0519           |
| AHMHE                 | 141.33             | 32.964            | 7.117            |
| Rivera                | 47.16              | 32.724            | 9.889            |
| Zhang                 | 127.88             | 33.336            | 0.438            |
| Proposed              | 132.07             | 32.46             | 1.072            |
the problems of causing over-enhancement (AHMHE and Zhang) or being unable to enhance image details (Rivera).

We further evaluate all the results with SFLNIA [46]. This intelligent assessment is based on LIVE database where 5 kinds of degradations, including JPEG2000, JPEG, white Gaussian noise (WN), Gaussian blurring (BLUR), and fast fading channel distortion (FF), are considered [46]. The Difference Mean Opinion Score (DMOS) ranges from 0 to 120 with 0 indicates no quality degradation in images. We can see from Table 1 that our proposed method produces averagely the least distortion in results, though AHMHE, Zhang, and Rivera also obtain similar low DMOS value.

4.3. Computation cost

Table 1 also compares the average execution time for all the sample images of different sizes used in our experiments with MAT-LAB running on a laptop with configuration of 17 CPU@2.4 GHz with 8 GB RAM. Although AINDANE, AHE, and Zhang consume less time averagely, the proposed algorithm gives better enhancement performance. More importantly, the structure of proposed algorithm as shown in Fig. 2 are amenable to parallel processing on embedded real-time platform.

5. Conclusion

In this paper, we designed a new enhancement method based on the characteristics of HVS for non-uniform and low light images. Several merits of the proposed algorithm can be summarized as following. Firstly, perceptual contrast is enhanced obviously which provides sharp edges for dark details. Secondly, both halo artifacts suppression and details clipping effects prevention are also considered so that naturalness is highly preserved. Thirdly, simplified saliency guided scheme is exploited to mimic the visual attention mechanism of HVS as dealing with both dark and bright regions separately. Fourthly, all parameters in the proposed algorithm are automatically selected, thus the human factor is released for parameters selection. And lastly color correction is adopted for recovering vivid color. All those improvements make our proposed algorithm robust for various scenes and outperforms several state-of-the-art algorithms by producing more natural results with enhanced perceptual details and comparably low computation cost.

In the future, we are going to implement the proposed algorithm on a System-on-Chip (SoC) based wearable device which automatically records daily images under both indoor and outdoor environments [47].

Acknowledgements

We would like to thank Yanfang Wang and Shuhang Wang for making their code available. This work was sponsored by National Funding of China, No. 61272351 and National Institutes of Health of United States, Grant No. R01CA165255.

References


