Halo Artifacts Reduction Method for Variational based Realtime Retinex Image Enhancement

Hiroshi Tsutsui*, Satoshi Yoshikawa†, Hiroyuki Okuhata‡, and Takao Onoye†
*Dept. of Communications and Computer Engineering,
Graduate School of Informatics, Kyoto University
Yoshida-Honmachi, Sakyo, Kyoto 606-8501, Japan
†Dept. of Information Systems Engineering,
Graduate School of Information Science and Technology, Osaka University
1-5 Yamadaoka, Suita, Osaka 565-0871, Japan
‡Synthesis Corporation
Awajimachi TC Bldg. 4th Floor, Awajimachi 2-6-9, Chuo-ku, Osaka 541-0047 Japan

***Abstract*—In this paper, we propose a novel halo reduction method for variational based Retinex image enhancement. In variational based Retinex image enhancement, a cost function is designed based on the illumination characteristics. The enhanced image is obtained by extracting the illumination component, which gives minimum cost, from the given input image. Although this approach gives good enhancement quality with less computational cost, a problem that dark regions near edges remain dark after image enhancement, known as halo artifact, still exists. In order to suppress such artifacts effectively, the proposed method adaptively adjusts the parameter of the cost function, which influences the trade-off relation between reducing halo artifacts and preserving image contrast. The proposed method is applicable to an existing realtime Retinex image enhancement hardware implementation.**

I. Introduction

Recently, consumer digital imaging devices such as digital cameras and color liquid crystal displays have been gaining much popularity. Digital images taken by digital still cameras and digital camcorders are sometimes deteriorated due to under- or over-exposure, blur, out of focus, noise, inappropriate illumination, and so forth. Thus, digital image enhancement in terms of color tone and/or contrast is indispensable for these kinds of consumer imaging devices.

General image enhancement schemes such as gamma correction, masking, and histogram equalization, are used for long years [1], which compensate each pixel value uniformly based on given equations. In contrast, adaptive image enhancement schemes may compensate each pixel value referring its surrounding pixels in order to reproduce a high quality image.

As one of the well-known adaptive image enhancement schemes, the Retinex theory, conceived by Land and McCann [2], introduces a model of the lightness and color perception of human visual system. The Retinex theory, which is based on a property of the color constancy phenomenon, decomposes a given image into two different images: the reflectance image and the illumination image as depicted in Fig. 1.

A number of illumination models, i.e. algorithms, are proposed for the Retinex theory [2]–[8]. However, due to enormous iteration processes incurred in the Retinex algorithm, applications have been only limited to low resolution videos. Motivated by this, we proposed architectures for real-time Retinex video image enhancement [9]–[11]. The architecture of [11] supports WUXGA (1,920×1,200) 60 fps realtime processing. However, a problem, known as halo artifact, still exists, which dark regions near edges remain dark after image enhancement. In Retinex based image enhancement schemes, this halo artifact is always a problem.

In this paper, we propose a novel halo reduction method for Retinex image enhancement, which can be applicable to the existing implementation [11]. In variational method, a cost function is designed based on the illumination characteristics. The proposed method adaptively adjusts the parameter of the cost function, which influences the trade-off relation between reducing halo artifacts and preserving image contrast. In addition, we also propose a generic post processing to reduce halo artifacts, which is applicable to other illumination estimation based image enhancement methods.

The remainder of this paper is organized as follows. In Section II, we briefly review the Retinex theory and its variational model. The mechanism of halo artifacts generation is also explained in this section with brief review of existing works. In Section III, we present the proposed halo reduction method. In Section IV, we present experimental results of the proposed method. Finally, the paper is summarized in Section V.

II. Retinex Theory and Halo Artifacts

Color constancy phenomenon stands for the fact that human visual system is able to practically recognize and match colors under a wide range of different illuminations. The Retinex theory [2] utilizes this property to extract the illumination image. In the Retinex theory, an input image $S(x, y)$ can be expressed by the following equation

$$S(x, y) = L(x, y)R(x, y), \tag{1}$$

where $L(x, y)$ and $R(x, y)$ indicate the illumination image and the reflectance image, respectively. Image enhancement can

(a) Input image $S(x, y)$.

(b) Illumination image $L(x, y)$.

(c) Reflectance image $R'(x, y)$.

Fig. 1. Example of Retinex image enhancement.

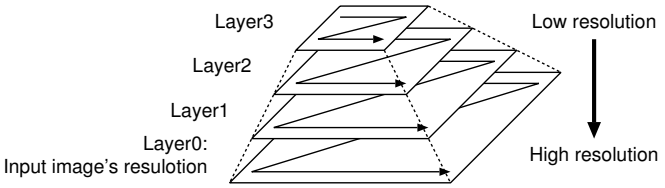


Fig. 2. Layer hierarchy for illumination estimation.

be achieved by extracting $L(x, y)$ from $S(x, y)$ in order to generate $R(x, y)$ as the illumination-independent image.

The original Retinex theory does not explicitly specify a scheme to estimate the illumination image, which is in general an ill-posed problem. Several illumination models are proposed so far for the Retinex theory [2]–[8]. Among them, we focus on the variational model as the most viable for practical applications in terms of computational cost and image quality.

A. Variational Model

Variational model algorithm is constructed to minimize the following cost function,

$$F[l] = \int_{\Omega} (|\nabla l|^2 + \alpha(s-l)^2 + \beta|\nabla(s-l)|^2)dxdy, \quad (2)$$

where, α and β are parameters, $s(x, y)$, $r(x, y)$, and $l(x, y)$ represent logarithmic expression of input image, reflectance image, and illumination image, respectively. Cost terms, such as $|\nabla l|^2$, $(s-l)^2$, and $|\nabla(s-l)|^2$ represent spatial smoothness of illumination image, closeness between l and s, and spatial smoothness of the reflectance image r, respectively.

To find illumination image l which minimizes the cost $F[l]$, a projected normalized steepest descent (PNSD) algorithm is used, in which the following steps are iteratively executed from low resolution layer to high resolution layer as illustrated in Fig. 2. Note that G is the gradient of $F[l]$, and the layer k represents k-times decimated image. Hereafter, we use images in logarithmic space, such as s, r, and l, instead of S, R, and L.

1) Input image $s(x, y)$, parameters α and β, and number of iterations T_k for each layer k are set.

2) Gaussian pyramid of image s, as shown in Fig. 2, is constructed by decimation with the following low-pass filter. Here, k-times decimated image is denoted by $s^{(k)}$. Note that $s^{(0)}$=s.

$$\kappa_{\mathrm{PYR}} = \frac{1}{16}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \quad (3)$$

3) The following steps are iteratively executed for each layer k from the lowest resolution layer N. Before the following iteration, the initial illumination in layer N, which is denoted by $l_0^{(N)}$, is initialized by $l_0^{(N)} = \max\{s^{(N)}\}$. The iteration index is denoted by j, and j=$1, 2, \ldots, T_k$.

a) Calculation of gradients for the illumination image and input image.

$$G_{\mathrm{A}}(x, y) = \nabla l_{j-1}^{(k)}(x, y), \quad (4)$$

$$G_{\mathrm{B}}(x, y) = \nabla s^{(k)}(x, y), \quad (5)$$

$$\begin{aligned} G(x, y) = & -G_{\mathrm{A}}(x, y) \\ & +\alpha\left(l_{j-1}^{(k)}(x, y) - s^{(k)}(x, y)\right) \\ & -\beta\left(G_{\mathrm{A}}(x, y) + G_{\mathrm{B}}(x, y)\right), \end{aligned} \quad (6)$$

where the Laplacian ∇ is defined as follows,

$$\nabla l = l * k_{\mathrm{LAP}} 2^{-2k}, \quad (7)$$

$$k_{\mathrm{LAP}} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}. \quad (8)$$

b) Calculation of μ_{NSD}.

$$\mu_{\mathrm{A}} = \sum_x \sum_y G(x, y)^2, \quad (9)$$

$$\mu_{\mathrm{B}} = -\sum_x \sum_y G(x, y)\nabla G(x, y), \quad (10)$$

$$\mu_{\mathrm{NSD}} = \frac{\mu_{\mathrm{A}}}{\alpha\mu_{\mathrm{A}} + (1+\beta)\mu_{\mathrm{B}}}. \quad (11)$$

c) Update of the illumination image with the constraint $l \leq s$.

$$l_j^{(k)}(x, y) = \max\left(l_{j-1}^{(k)}(x, y) - \mu_{\mathrm{NSD}} G(x, y), s^{(k)}(x, y)\right) \quad (12)$$

4) After T_k iterations, initialization of the next resolution layer is carried out. The initial illumination image of the next layer k-1, which is $l^{(k-1)}$, is determined by up-scaling resulted illumination image $l^{(k)}_{T_k}$
5) The illumination image of the layer 0, which would be $l^0_{T_0}$, whose resolution is the same as the input image, is finally output.

After the above mentioned steps, the reflectance image R can be obtained by

$$R = \exp(s - l). \tag{13}$$

The reflectance image is sometimes over-enhanced. To avoid this, in [6], applying a gamma corrected illumination back to the reconstructed reflectance image. The flow of this Retinex image enhancement with illumination correction is shown in Fig. 3. The gamma correction is expressed by $L' = L^{1/\gamma}$.

B. Halo Artifacts

In general, halo artifacts are observed along with edge regions that have large intensity gradient. Examples of halo artifacts are shown in Fig. 4. In Fig. 4(a), the regions around the pine trees and the statue are over-enhanced and made relatively bright. This type of artifact is generally called halo. On the other hand, in Fig. 4(b), the inside areas of the building and the roof regions are not enhanced sufficiently and kept relatively dark. This type of artifact is also halo. To distinguish these two kinds of halo, in this paper, the former one (Fig. 4(a)) is called *positive halo* and the latter (Fig. 4(b)) *negative halo*.

Halo artifacts can be seen in images enhanced by using adaptive image enhancement methods. This is because that many of such methods utilize the constraint that illumination image should be spatially smooth. When the illumination is estimated in the regions around edge with this constraint, the regions in reluctance image tend to be over-enhanced or not to be sufficiently enhanced. This halo artifact generation mechanism is illustrated in Fig. 5 using 1-d expression. In this figure, the pixel values of a line of a image, which is likely to be shown in Fig.8(a), is displayed. Note that halo effect becomes large as the intensity gradient along with edges becomes large.

In the variational model, positive halo artifacts are suppressed by using the constraint $l \geq s$, which is realized in Eq. (12), as shown in Fig. 6. In Fig. 5, positive halo artifacts can be found in the area where $l < s$, which is located in the right side of the edge. On the other hand, at the same position in Fig. 6, the constraint $l \geq s$ makes estimated l equal to s. Therefore positive halo artifacts are suppressed. In contrast, negative halo still remains in the area at the left side of the edge.

In order to suppress such halo artifacts, in [12, 13], the cost function Eq. (2) is modified. In these methods, the first term of Eq. (2) is adaptively weakened according to the edge presence in input images. In [14, 15], Retinex based image enhancement filters designed to suppress halo artifacts are proposed. However these existing methods require large sized filters or many multiplications and divisions. The proposed method belongs

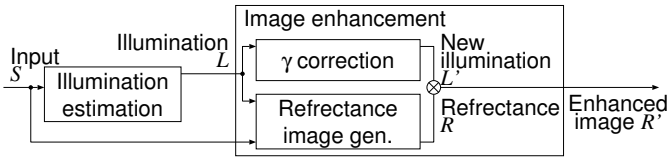


Fig. 3. Retinex based image enhancement flow.

(a) Positive halo. (b) Negative halo.

Fig. 4. Examples of halo artifacts.

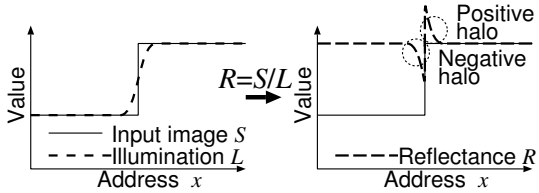


Fig. 5. Halo artifacts generation mechanism (1-d expression).

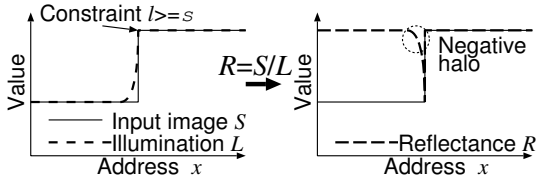


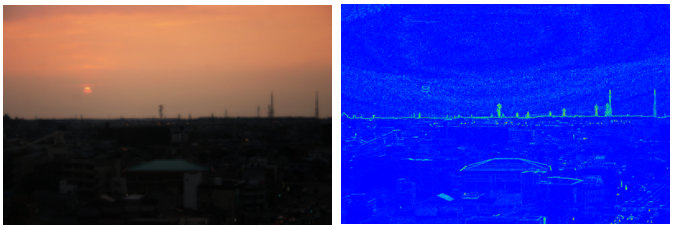
Fig. 6. Positive halo artifacts suppression by the constraint $l \geq s$ (1-d expression).

to same category of [12, 13]. However the proposed method is applicable to the existing hardware implementation [11]. In addition, we also propose a generic post processing to reduce halo artifacts.

III. Proposed Halo Artifacts Reduction Method

In the proposed halo artifacts reduction method, we modify the cost function Eq. (2). In order to make estimated illumination around edges similar to the original image, we change parameter α adaptively according to the intensity gradient. The first term of Eq. (2) makes the estimated illumination spatially smooth while the second term makes that nearly equal to the original image. Using parameter α, the weight of these two constraints, which are contrary to each other, can be adjusted. When α is large, halo artifacts are suppressed but the contrast of the detail part are weakened. In contrast, Using small α keeps the contrast detail but suffers from halo artifacts. Therefore, in the proposed method, we set α large at the area where the intensity gradient is large and keep α small at the detailed part.

In addition to this adaptive α approach, we adopt a post process for the estimated illumination to suppress negative



(a) Input image. (b) Values of $\alpha(S')$.

(c) Color map to display α values in this paper.

Fig. 7. Value of $\alpha(S')$ used in the proposed method.

halo artifacts. The fundamental cause of negative halo is that the illumination l is larger than the input image s around edges as shown in Fig. 6. Hence, in the proposed post processing, we erode high illumination area.

A. Illumination Estimation using Adaptive α

In the proposed method, large value of α is used for the area, which have large intensity gradient such as edges, to make the weight of the constraint large, which makes the illumination nearly equal to the original image. Therefore we adaptively change α by a function of $\alpha(S'(x,y))$, where $S'(x,y)$ is the intensity gradient of input image S and given by the Laplacian filter Eq. (7).

When we replace α of Eq. (2) by $\alpha(S')$, the procedure described in Sec. II must be modified. First, Eq. (6) is replaced by

$$G = -G_{\mathrm{A}} + \alpha(S')(l-s) - \beta(G_{\mathrm{A}} - G_{\mathrm{B}}), \tag{14}$$

and Eq. (11) is replaced by

$$\mu_{\mathrm{NSD}} = \frac{\mu_{\mathrm{A}}}{\mu_{\mathrm{C}} + (1+\beta)\mu_{\mathrm{B}}}, \tag{15}$$

where μ_{C} corresponds to $\alpha\mu_{\mathrm{A}}$ in Eq. (11) and is given by

$$\mu_{\mathrm{C}} = \sum_x \sum_y \alpha(S'(x,y)) G(x,y) \nabla G(x,y). \tag{16}$$

These modifications have only small impact to the original procedure in terms of data flow. Therefore the proposed method can be implemented based on the existing architecture.

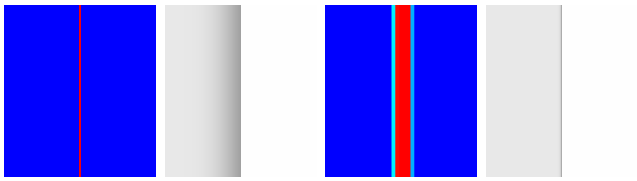
As for function $\alpha(S')$, it is reasonable to use $\alpha(S')$ that becomes large when S' is large. In this paper, we use $\alpha(S')$ given by

$$\alpha(S') = \exp\left(-\frac{a}{1+bS'(x,y)}\right), \tag{17}$$

where a and b are parameters. In this paper, we use a=10 and b=256. The value of this function becomes rapidly large as S' becomes large. Note that the intensity gradient S' is normalized as $0 \le S' \le 1$. An example of α values used in the proposed method is shown in Fig. 7. Hereafter, we use the color map shown in Fig. 7(c) to display α values.

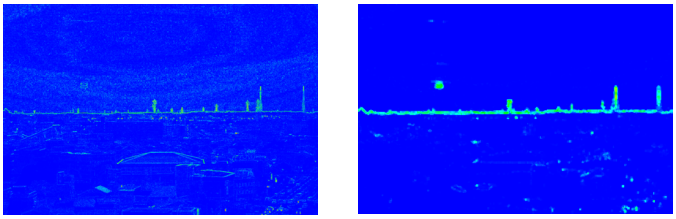
Here, there are two issues to be considered as follows.

(a) Input image.



(b) α (2 pixels) (c) Result of (b) (d) α (16 pixels) (e) Result of (d)

Fig. 8. Relation between the area size of large α values and halo artifacts suppression performance.



(a) $\alpha(S')$ calculated in layer 0. (b) $\alpha(S')$ calculated in layer 2.

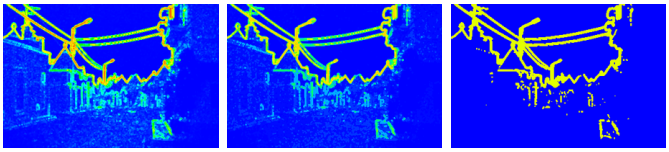
Fig. 9. Comparison of $\alpha(S')$ calculated in layer 0 (original resolution) and 2.

1) α becomes large even in regions around noise or details, since α is sensitive to them, and
2) the size of large α area is relatively small in halo artifacts regions around true edges between bright and dark regions.

The first issue is obvious. As for the second issue, an example is shown in Fig. 8, where a test input image are processed using two sets of α values; the area width of large α values is 2 or 16 pixels as shown in Figs. 8(b) and (d), respectively. As we can see from the figure, when the area size of large α value is small, the halo artifacts suppression effect is not observable. Therefore, the large α value area should be made properly large.

In order to overcome the above mentioned two issues, in the proposed method, we calculate $\alpha(S')$ in low resolution layers. Figure 9 shows the comparison of $\alpha(S')$ calculated in layer 0 (original resolution) and layer 2. From this figure, it can be seen that, in the proposed method, the area size of large α values at the regions around the horizon and the sun, where halo artifacts tend to appear, is larger than that calculated in the original resolution. In addition, α becomes small at regions around noise or details. Therefore the proposed approach is effective to both two issues. In addition, in the variational model, such scaled images are calculated in its procedure as described in Sec. II, so it is reasonable to calculate $\alpha^k(S')$ in each layer k.

In [6], the parameters in Eq. (2) are set α=0.0001 and β=0.1. Considering hardware implementation, it is better to use $\alpha =$



(a) No α quantization. (b) 8-bit α quantization. (c) 1-bit α quantization.

Fig. 10. Values of α after quantization.

2^{-13} and $\beta = 2^{-3}$ in terms of hardware cost. In this case, the multiplications in Eq. (6) can be implemented by only shifters. It is also key issue how to implement Eq. (17). We propose to use only two values of 2^{-4} and 2^{-11} for $\alpha(S')$. We confirmed that such drastic simplification has no impact to the resulted images. In this case, $\alpha(S')$ is calculated by just thresholding S'. Fig. 10 shows the $\alpha(S')$ values after 8-bit quantization and the proposed 1-bit quantization.

B. Erosion of Estimated Illumination

Figure 11 shows the mechanism of the proposed halo artifacts suppression by eroding estimated illumination. Erosion is one of two fundamental operations in morphological image processing. It is originally defined for binary images but it can be extended to gray scale images. In this paper, we implement erosion for the estimated illumination image as a min filter as follows

$$l_{\mathrm{er}}(x,y) = \min_{(h_x,h_y)\in S_{\mathrm{er}}} l(x+h_x, y+h_y), \qquad (18)$$

where l_{er} is erosion of estimated illumination l, and S_{er} is 3×3 pixel window. By eroding illumination image, the high l value area is shrunk as we can see in Fig. 11 as the right shift of l curve. This operation enlarges the area where positive halo occurs, but this positive halo can be suppressed using the constraint of $l \geq s$. In the proposed method, this constraint will be also applied after the erosion process. Note that this erosion approach can be applicable to other Retinex based image enhancement schemes that utilized the constraint of $l \geq s$.

Figure 12 shows the effect of the proposed erosion approach for two types of α values. According to this figure, the effect of the erosion is small when α is small, where halo sensitive edges are not detected, while the effect is large when α is large, where halo artifacts tend to appear.

C. Discussion about Hardware Implementation

Figure 13 shows the block diagram of the proposed illumination estimation module based on the previous architecture [11]. The data flow and parameters are same to the previous ones, where the lowest illumination layer N=5 and the iteration numbers are T_5=30, T_4=20, and T_3=10. In this implementation, estimated illumination in layer 3, $l^{(3)}_{T_3}$, will be transferred to the external SRAM. The next module reads illumination data from this SRAM and scales it to the original image resolution. Image enhancement is performed using this scaled illumination with Gamma correction conforming the constraint of $l \geq s$. The modules newly added to the

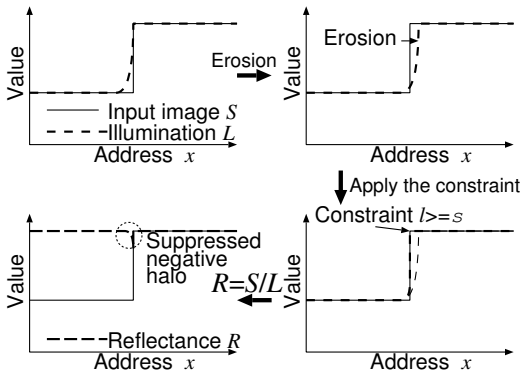


Fig. 11. Mechanism of the proposed halo artifacts suppression by eroding estimated illumination.

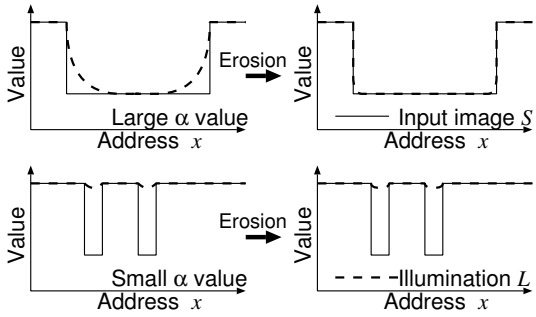


Fig. 12. Effect of the proposed erosion approach for different value of α.

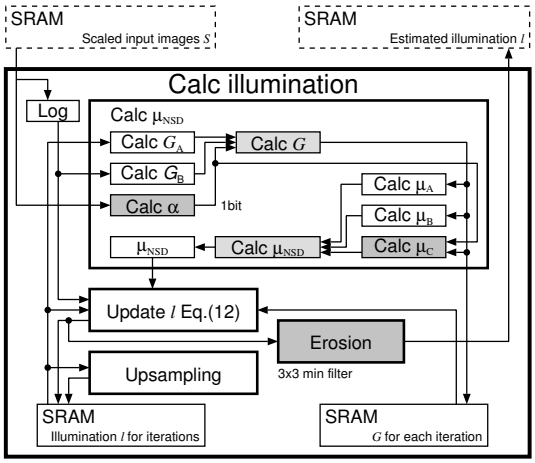


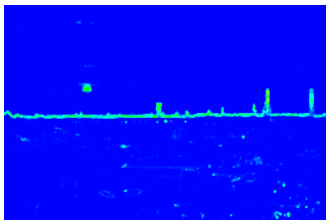
Fig. 13. The block diagram of the illumination estimation module.

previous work are only three modules: α and calculation, μ_{C} calculation, and erosion post processing modules. Calculation of G and μ_{NSD} is modified based on Eqs. (14) and (15). Therefore the proposed halo artifacts reduction method can be implemented to the existing architecture with small area overhead.

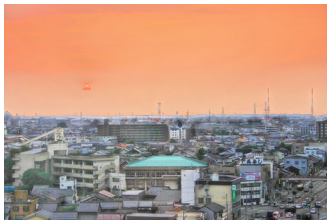
IV. Experimental Results

Figure 14 shows enhanced images by using proposed method. The resulted α and enhanced image of the input image

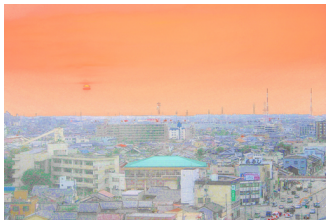
(a) Input image.



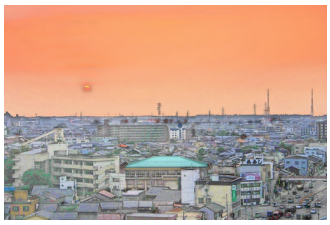
(b) Obtained α values from (a).



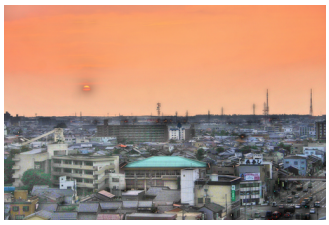
(c) Enhanced image (proposed).



(d) Enhanced image (α=1).



(e) Enhanced image (α=0.01).

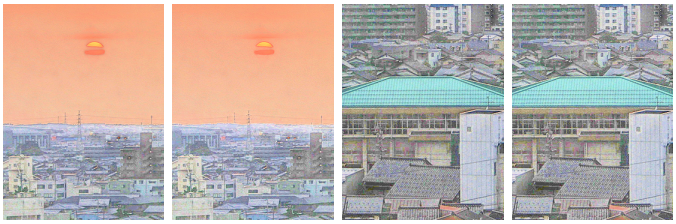


(f) Enhanced image (α=0.0001).

Fig. 14. Comparison result. (a) Input image. (b) α values obtained from (a) by using the proposed method. (c) Enhanced image by using the proposed method. (d)–(f) Enhanced images by using fixed α values of 1, 0.01, and 0.001, respectively.

Fig. 14(a) are presented in Figs. 14(b) and 14(c), respectively. For comparison purpose, we present enhanced images by using fixed α values of 1, 0.01, and 0.001, in Figs. 14(d), (e), and (f), respectively. According to Fig. 14(b), it can be seen that α values around the horizon and the sun, where halo artifacts tend to appear, are successfully made large while they are kept small in the other regions. Therefore, the halo artifacts around the horizon and the sun is successfully suppressed while the contrast details of the other regions are kept as same as the case of α=0.01.

In order to evaluate the proposed erosion post processing, enhanced images with or without the proposed erosion are shown in Fig. 15. In this figure, we focus on areas where halo artifacts tend to appear (Figs. 15(a) and (b)), and where we want to keep the contrast details (Figs. 15(c) and (d)). As we can see from this figure, the negative halo artifacts which occurred in the case without the proposed erosion (Fig. 15(a)) around the horizon are successfully suppressed with the proposed method (Fig. 15(b)). On the other hand, as for the contrast details, there are almost no differences with or without the proposed erosion (Figs. 15(c) and 15(d)).



(a) No erosion. (b) With erosion. (c) No erosion. (d) With erosion.

Fig. 15. Effect of the proposed erosion post processing for suppressing halo artifacts.

V. Conclusion

In this paper, we explained the halo artifacts generation mechanism and proposed a halo reduction method, which is applicable to the existing hardware implementation. The proposed method consists of two approaches: (1) illumination estimation using edge adaptive parameter and (2) erosion of estimated illumination. The detailed evaluation of the proposed method and implementation remains as a future work.

References

[1] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 3rd ed. Prentice Hall, 2008.

[2] E. H. Land and J. J. McCann, "Lightness and Retinex theory," *Journal of the Optical Society of America*, vol. 61, Dec. 1964.

[3] D. J. Jobson, Z. Rahman, and A. Woodell, "A multiscale Retinex for bridging the gap between color images and the human observation of scenes," *IEEE Trans. Image Processing*, vol. 6, no. 7, pp. 965–976, Jul. 1997.

[4] D. J. Jobson, Z. Rahman, and G. A. Woodell, "Properties and performance of a center/surround Retinex," *IEEE Trans. Image Processing*, vol. 6, no. 3, pp. 451–462, Mar. 1997.

[5] D. Terzopoulos, "Image analysis using multiglid relaxation methods," *IEEE Trans. PAMI*, vol. 8, no. 2, pp. 129–139, Mar. 1986.

[6] R. Kimmel, M. Elad, D. Shaked, R. Keshet, and I. Soble, "A variational framework for Retinex," *International Journal of Computer Vision*, vol. 52, no. 1, pp. 7–23, Apr. 2003.

[7] S. Carrato, "A pseudo-Retinex approach for the visualisation of high dynamic range images," in *Proc. 5th European Cooperation in the Field of Scientific and Technical Research (COST) 276 Workshop*, Oct. 2003, pp. 1–6.

[8] C. T. Lin and C. H. Huang, "CNN-based Retinex technology," in *Proc. of the European Conf. on Circuit Theory and Design (ECCTD '03)*, vol. 2, Sep. 2003, pp. 69–72.

[9] K. Takahashi, Y. Nozato, H. Okuhata, and T. Onoye, "VLSI architecture for real-time Retinex video image enhancement," in *Proc. Workshop on Synthesis and System Integration of Mixed Information Technologies (SASIMI 2007)*, Oct. 2007, pp. 81–86.

[10] H. Okuhata, K. Takahashi, Y. Nozato, T. Onoye, and I. Shirakawa, "Video image enhancement scheme for high resolution consumer devices," in *Proc. Intl. Symp. on Comm., Control and Signal Processing (ISCCSP)*, Mar. 2008, pp. 639–644.

[11] H. Tsutsui, H. Nakamura, R. Hashimoto, H. Okuhata, and T. Onoye, "An FPGA implementation of real-time Retinex video image enhancement," in *Proc. World Automation Congress (WAC)*, Sep. 2010, pp. 1–6.

[12] C. Shen and W. Hwang, "Color image enhancement using Retinex with robust envelope," in *Proc. ICIP*, 2009, pp. 3141–3144.

[13] G. Guarnieri, S. Marsi, and G. Ramponi, "High dynamic range image display with halo and clipping prevention," *IEEE Trans. Image Processing*, vol. 20, no. 5, pp. 1351–1362, 2011.

[14] M. Elad, "Retinex by two bilateral filters," *Scale Space and PDE Methods in Computer Vision*, pp. 217–229, 2005.

[15] S. Marsi and G. Ramponi, "A flexible FPGA implementation for illuminance–reflectance video enhancement," *Journal of Real-Time Image Processing*, pp. 1–13, 2011.