https://doi.org/10.2352/ISSN.2470-1173.2017.17.COIMG-423

Linear mapping based inverse tone mapping

Dae-Eun Kim, and Munchurl Kim
VIC Lab, School of Electrical Engineering, Korea Advanced Institute of Science & Technology (KAIST); Republic of Korea

Abstract

We propose an inverse tone mapping (iTM) method which can both handle the details of low dynamic range (LDR) images and expand the dynamic ranges of the LDR images. The conventional iTM algorithms often fail to precisely restore the details of the input LDR images. To deal with this problem, we take a two-layer approach where each LDR image is separated into a base layer and a detail layer by bilateral filtering. The detail layer is mapped into that of a high dynamic range (HDR) image via linear mapping while the base layer is expanded via linear stretching to the dynamic range of a target display device. Then, the two resultant base and detail layers are used to reconstruct one final HDR image. From this, the details of the reconstructed HDR image can be revived via learned linear mapping. In order to learn the mapping from the LDR detail layer to an HDR detail layer, the HDR-LDR pairs of training patches of detail layers are classified into various groups based on the features of LDR detail patches. For each group, a linear mapping is learned during a training phase, which can then be applied for HDR reconstruction in testing phases. From the experimental results, we observed that proposed method can restore much more details of HDR images than the conventional methods.

Introduction

In spite that commercial HDR TVs are appearing in the current consumer markets, there are lack of HDR videos and images. Also, legacy LDR services are mostly available, and are expected to be somehow prevailed for a certain period of time. That is, since LDR contents are more dominated compared to HDR contents for a while, the need for good inverse tone mapping (iTM) methods that can expand the dynamic ranges of LDR images to the dynamic range of HDR display devices will increase in order to enhance the visual qualities of LDR contents and to fully utilize the HDR display devices available in the current markets.

There are many works related to the iTM [1]-[14]. However, most of them have considered only the expansion of the dynamic range of the input LDR images. They are designed by focusing on how much they expand to a certain range the dynamic ranges of the LDR images so that the expanded images are to be realistic or natural when they are displayed on target HDR display devices. However, besides the compression of overall dynamic range, there is a more critical problem in LDR images. Often, the LDR image are lack of details or have low contrasts, which results from the narrower dynamic ranges. Therefore, the iTM methods should be designed by considering not only the expansion of the dynamic ranges but also the restoration of the details with high contrast. In this paper, we propose an elaborate solution to this problem.

The purpose of this work is to enhance the subjective qualities of inverse-tone-mapped images by restoring the details of the corresponding HDR images and by expanding their dynamic ranges onto target display devices. The restoration of the details is an ill-posed problem since there is little remained detail. Therefore, it is necessary to find a solution for this ill-posed problem in a more elaborated way, which motivates our learning-based linear mapping method. So the lost details of LDR images can be restored by applying learned linear mappings, which is otherwise often difficult to be reconstructed via conventional methods.

Banterle et al. [1] suggested the first work that uses the term, inverse tone mapping. They formulated the inverse of Reinhard's tone mapping algorithm [15] to expand the dynamic range of input LDR images. In addition to the inversion of the tone mapping algorithm, they calculated an expand map that indicates how much of the dynamic range the LDR pixels should be expanded depending on pixel locations. Their expand map not only expands the suppressed highlight regions of the LDR images, but also mitigates contouring artifacts that can occur due to the quantized pixel values in the LDR images during the expansion operation.

In [2], Akyüz et al. generated the HDR images from the LDR images based on a gamma curve, where the gamma value is a user-configurable parameter. In [2], they performed experiments to answer two questions: first, whether an HDR display device can give a better impression than LDR display devices; second, how we can handle the legacy LDR images. The experimental results confirmed that subjects actually turned out to prefer the HDR display to the LDR display and that it is possible to produce a plausible HDR image with a simple liner expansion with the gamma parameter set to '1'. However, with this simple linear expansion, it is impossible to reconstruct the lost details of the image.

In [3], Meylan et. al proposed an inverse tone mapping algorithm that expands the dynamic range of the input LDR image in a piece-wise-linear manner. In other words, they divide each input LDR image into two regions: the diffuse region and the specular region. Then they apply two different linear functions to stretch the dynamic range of each region. The decision on whether a pixel belongs to the diffuse region or the specular region is made based on the pixel value. If the pixel value is larger than a predefined threshold, the pixel belongs to the specular region and otherwise, to the diffuse region. This makes the suppressed pixel values in the specular region of the input LDR image get more expanded than those in the diffuse region.

In [4], Rempel et. al proposed an inverse tone mapping algorithm that expands the suppressed dynamic ranges of LDR images by finding the saturated pixel regions and by boosting the pixel values in the saturated pixel regions. In [10], an inverse tone mapping algorithm based on a human visual system is proposed. It specifically simulated the response of the retina with a sigmoid-like function which is non-linear. It strengthened the local contrast, hence preserved image details. In [11], an iTM algorithm is proposed by applying a cross-bilateral filter. The iTM algorithm is designed for a variety of exposure ranges, including under-exposure and over-exposure. In [10] and [11], DRIM [16] is used as an objective indicator to evaluate the performance of iTM algorithms, and these iTMOs improved to restore the details of the LDR images.

As mentioned above, most iTM algorithms only consider the dynamic range of the pixel values of the LDR images. That is, given a target HDR display, the main purpose of inverse tone mapping was to determine how plausibly to expand the dynamic range of pixel

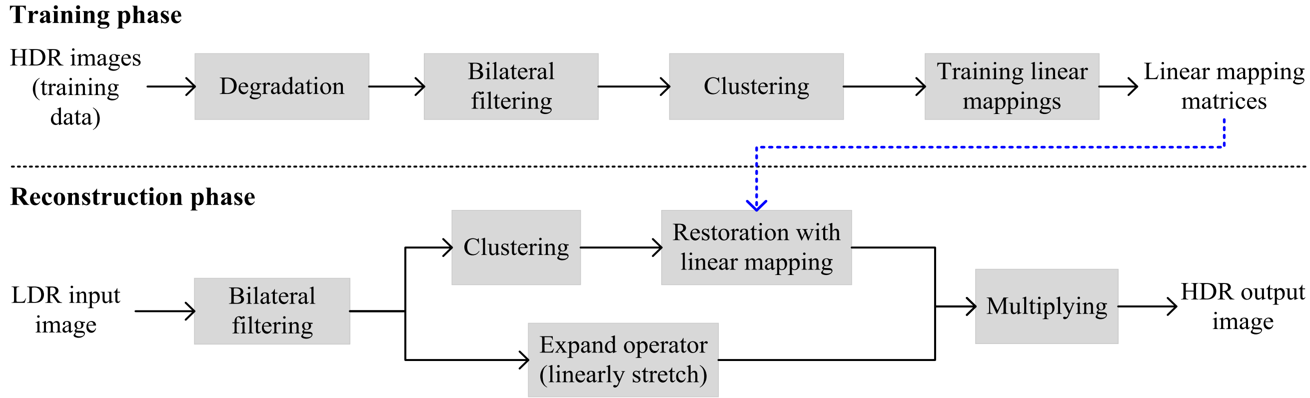

Fig. 1. Overview of the proposed iTM algorithm.

values in a particular region of an LDR image. However, in HDR restoration, the restoration of lost details have not been addressed as a major topic. Therefore, in this paper, we propose a machine learning based iTM method that can not only expand the dynamic range of the LDR image but also restore the lost details of the LDR image.

Bilateral Filtering

In order to restore the lost details of LDR images, we separate each input LDR image into a base layer and a detail layer. By doing so, we can extract the detailed components of the input LDR image that can be handled independently to the background luminance. For the separation into the detail layer and the base layer, we used a bilateral filter [17]. The bilateral filter is an edge-preserving filter that preserves relatively large edges while blurring the details. The bilateral filter is defined as

$$Bil\left(\mathrm{I}_L^p(\mathbf{x})\right)=\frac{1}{W_p}\sum_{\mathbf{y}\in\Omega(\mathbf{x})}\mathrm{I}_L^p(\mathbf{y})G_{\sigma_r}\left(\left\|\mathrm{I}_L^p(\mathbf{y})-\mathrm{I}_L^p(\mathbf{x})\right\|\right)G_{\sigma_s}\left(\left\|\mathbf{y}-\mathbf{x}\right\|\right) \quad (1)$$

where $\mathrm{I}_L^p(\mathbf{x})$ is an LDR pixel intensity in location $\mathbf{x}=(i,j)$ and $\mathrm{I}_L^p(\mathbf{y})$ is an LDR pixel intensity in location $\mathbf{y}=(m,n)$. In (1), we have $i-N/2\le m\le i+N/2$ and $j-N/2\le n\le j+N/2$ with a support region of size *N*. $G_{\sigma_r}(\cdot)$ and $G_{\sigma_s}(\cdot)$ are called a range kernel and a spatial kernel, respectively. They are often defined as Gaussian functions [17]. In this paper, $\sigma_r=0.15$ and $\sigma_s=0.005\times\max(w,h)$ are empirically selected and used where w and h are the width and the height of an LDR image. And W_p is a normalization factor, and is defined as

$$W_p=\sum_{\mathbf{y}\in\Omega(\mathbf{x})}G_{\sigma_r}\left(\left\|\mathrm{I}_L^p(\mathbf{y})-\mathrm{I}_L^p(\mathbf{x})\right\|\right)G_{\sigma_s}\left(\left\|\mathbf{y}-\mathbf{x}\right\|\right). \quad (2)$$

We can obtain a base layer of an LDR image, $\mathbf{B}_L^p$, by applying the bilateral filter to an LDR image. Once we obtain a base layer of an LDR image, we can obtain its detail layer by

$$\mathbf{D}_L^p=\mathbf{I}_L^p/\mathbf{B}_L^p \quad (3)$$

where $\mathbf{D}_L^p$ is the detail layer of an LDR image.

Method

To restore the lost details in LDR image is an ill-posed problem. In order to solve this, we adopt a learning-based iTM method that finds linear mappings from LDR details to HDR ones. Our proposed iTM method consists of two phases: training phase and reconstruction phase. In the training phase, a linear mapping function, which represents the relation between the detail layer of the original HDR image and that of the corresponding LDR image, is learned. In the reconstruction phase, the input LDR image is decomposed into a base layer and a detail layer by applying a bilateral filter and then, the base layer is linearly stretched, and the detail layer is mapped to that of the HDR image through the linear mapping function learned in the training phase. Finally one HDR image is reconstructed by the stretched base and linearly mapped detail layer. Fig. 1 shows the entire procedure of our proposed iTM method. A more detailed explanation for our proposed iTM method in Fig. 1 is provided in the following sections.

Clustering

As aforementioned, we learn the relation between the detail layers of LDR images and those of HDR images in the training phase and apply the learnt mapping functions to the detail layers of the LDR input images in the reconstruction phase. In order to learn this relation, we should cluster images into some categories depending on their features. Since images consists of nonstationary local texture regions, one single universal mapping may not be sufficient to describe the LDR-HDR relation of detail layers. So, we use patch-wise linear mapping from the detail layer of an LDR patch

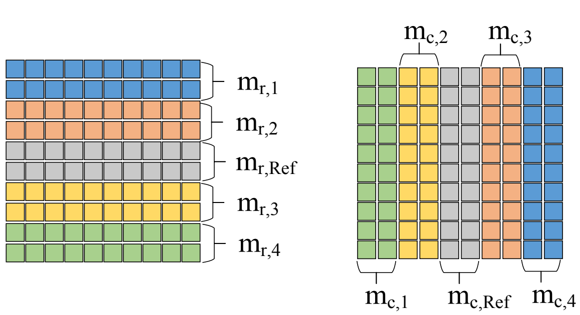

Fig. 2. Pixel variation types.

to that of its corresponding HDR patch. For this, we first split each training image into small-sized local patches (e.g., size of 10×10). All patches of the training images are clustered into 5 categories according to their background luminance values. Then, each clustered patch-group is sub-clustered into 256 categories depending on the pixel variation types. Fig. 2 describes how to determine the pixel variation type for each patch. In Fig. 2, $m_{r,1}$, $m_{r,2}$, $m_{r,ref}$, $m_{r,3}$ and $m_{r,4}$ indicate the means of pixel values in the first, second, third, fourth and fifth two rows of an LDR image's detail layer, respectively. Similarly, $m_{c,1}$, $m_{c,2}$, $m_{c,ref}$, $m_{c,3}$ and $m_{c,4}$ indicate the means of pixel values in the first, second, third, fourth and fifth two columns of the LDR image's detail layer, respectively. For the vertical pixel variation, we compare $m_{r,1}$, $m_{r,2}$, $m_{r,3}$ and $m_{r,4}$ with $m_{r,ref}$, and determine whether each mean is larger or smaller than $m_{r,ref}$. For the horizontal pixel variation, we repeat the same procedure for $m_{c,1}$, $m_{c,2}$, $m_{c,ref}$, $m_{c,3}$ and $m_{c,4}$. From this, each image in a patch group can be characterized as one of 256 categories. So, all patches are clustered into 1,280 (= 5 × 256) categories. For each category, one single linear mapping is learned for the LDR-to-HDR mapping in detail layers.

Training Phase

As shown in Fig. 1, the HDR images to be used for training is degraded by a tone mapping operator. In this paper, Reinhard's TMO [15] is used as a degradation model. Then, we apply a bilateral filter to each training LDR and HDR image pair so that we obtain the detail layer of the HDR image, $\mathbf{D}_H^p$ and the detail layer of the LDR image, $\mathbf{D}_L^p$. $\mathbf{D}_L^p$ and $\mathbf{D}_H^p$ are then split into 10×10-sized training patches and they form HDR-LDR training patch pairs. Here we assume that for each pair, there is linear relation between the *i*-th HDR-LDR detail layer patch pair as

$$\mathbf{h}_i = \mathbf{M}_i \mathbf{l}_i \tag{4}$$

where $\mathbf{h}_i$ is a vectorized *i*-th HDR detail layer patch and $\mathbf{l}_i$ is its corresponding vectorized LDR detail layer patch. $\mathbf{M}_i$ is a linear mapping matrix from the *i*-th LDR detail layer to the *i*-th HDR detail layer. In addition, we assume that the LDR and HDR detail layer patch pairs in a same category share an identical linear mapping function. That is, for the LDR and HDR patch pairs in category *c*, the relation between the *i*-th LDR and HDR detail layer patch pair can be represented as

$$\mathbf{h}_i^c = \mathbf{M}_c \mathbf{l}_i^c \tag{5}$$

where $\mathbf{h}_i^c$ is a vectorized *i*-th HDR detail layer patch in category *c*, $\mathbf{l}_i^c$ is its corresponding vectorized LDR detail layer patch, and $\mathbf{M}_c$ is a linear mapping matrix for the patch pairs in category *c*.

In training phase, we calculate the linear mapping matrices for each category from a training set. For this, we define batch matrices as

$$\mathbf{H}_c = \begin{bmatrix} | & & | \\ \mathbf{h}_1^c & \cdots & \mathbf{h}_n^c \\ | & & | \end{bmatrix} \text{ and } \mathbf{L}_c = \begin{bmatrix} | & & | \\ \mathbf{l}_1^c & \cdots & \mathbf{l}_n^c \\ | & & | \end{bmatrix} \tag{6}$$

where $\mathbf{H}_c$ and $\mathbf{L}_c$ are the batch matrices of vectorized HDR and LDR detail layer patches in category *c*, and *n* is the number of HDR-LDR detail layer patch pairs in category *c*. For category *c* in a training set, (5) is rewritten as

$$\mathbf{H}_c = \mathbf{M}_c \mathbf{L}_c . \tag{7}$$

Although we assume that the patch-pairs in a same category share an identical mapping function, it is obvious that the difference between the left side and the right side does exist. So we want find the optimal $\mathbf{M}_c$ which can best estimate $\mathbf{H}_c$ by formulating a kernel ridge regression problem such as

$$\mathbf{M}_c = \arg\min_{\mathbf{M}} \|\mathbf{H}_c - \mathbf{M}\mathbf{L}_c\|_2^2 + \lambda \|\mathbf{M}\|_F^2 \tag{8}$$

where λ is a regularization parameter. We can obtain the optimal solution to (8) as [18]

$$\mathbf{M}_C = \mathbf{H}_C \mathbf{L}_C^T \left(\mathbf{L}_C \mathbf{L}_C^T + \lambda \mathbf{I}\right)^{-1} \tag{9}$$

where $\mathbf{I}$ is the identity matrix. From (6) and (9), we can learn a linear mapping matrix for category *c* from a training set. We apply this procedure for all other categories.

Reconstruction Phase

As shown in Fig. 1, the reconstruction phase is applied to the base layer and the detail layer of each input LDR image independently. When the input LDR image is given, the image is separated into a base layer and a detail layer by applying a bilateral filter. The detailed procedures for each layer are described in the following sections.

Reconstruction for detail layers

For the detail layer of the input LDR image to be reconstructed, we first split the detail layer of the input LDR image into 10×10-sized patches as done in the training phase. Then each is assigned into one of the 1,280 categories. Then the detail layer of its corresponding HDR image is obtained by multiplying the corresponding category's linear mapping matrix to the detail layer of the LDR image.

Reconstruction for base layers

For the base layer of the input LDR image to be reconstructed, it should be expanded to fit the dynamic range of the target HDR display device. It is known that a simple linearly-stretching ensures sufficient subjective image quality [2]. Thus, in this paper, we just linearly stretch the dynamic range of the base layer of the input LDR image. More specifically, we first linearize the pixel intensity values in the base layer by applying a de-gamma curve [19], and then the pixel values becomes linearized luminance values. The linearized luminance value of each pixel is then stretched linearly to fit into the dynamic range of the target HDR display device. Then, the luminance values are transformed via PQ-OETF [20], so that it can be input for the target HDR display device.

After completing the procedures for the two layers described above, the two layers are multiplied and finally reconstructed into one single output HDR image as

Fig. 3. Comparison of HDR output images: Top - conventional method; bottom - proposed method.

$$\mathbf{I}_H^p = \mathbf{D}_H^p \mathbf{B}_H^p \tag{10}$$

where $\mathbf{I}_H^p$, $\mathbf{D}_H^p$ and $\mathbf{B}_H^p$ are the reconstructed HDR image, the detail layer and base layer of the reconstructed HDR image, respectively.

Results

To show the effectiveness of our proposed iTM method, some experiments are performed and subjectively tested on a commercial HDR TV, Samsung SUHDTV ™. Because the HDR TV can support the maximum luminance up to 1,000 cd/m^2, the maximum luminance of the iTM methods for comparison is set to 1,000cd/m^2. That is, for the proposed iTM method, the base layer is stretched to that luminance. For the detail layer, we use a training set of 90 HDR images and their corresponding LDR images which are tone mapped by Reinhard's TMO method [15]. Figs. 3-8 show the comparison of our proposed algorithm against the conventional method [3]. As shown in Figs. 3-8, our proposed iTM method can restore more details in the reconstructed HDR images.

Conclusion

In this paper, we proposed a learning-based linear mapping iTM method. It should be noted that most of the conventional iTM methods have focused their attentions on the dynamic range expansion of input LDR images, not the restoration of lost details. However, we presented an elaborate iTM method that can restore the lost details of the input LDR images. To get the detail layers, we used a bilateral filter. And the base layer is expanded by linear stretching, independent of the detail layer. From this, the suppressed dynamic ranges of LDR images can appropriately be expanded and

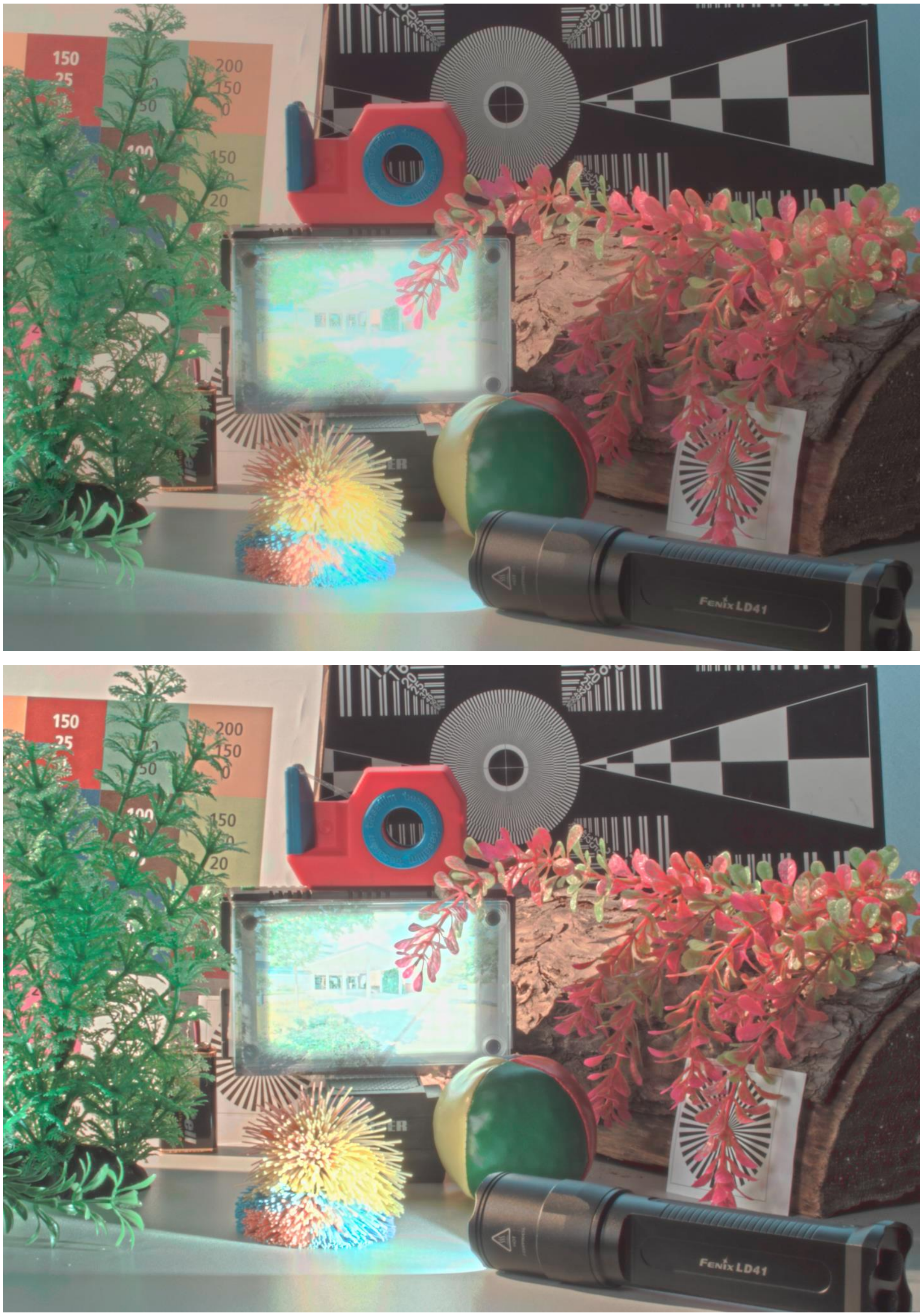

Fig. 5. Comparison of HDR output images: Top - conventional method; bottom - proposed method.

the lost details of LDR images can be revived into HDR reconstruction.

Acknowledgment

This work was supported by the National Research Foundation (NRF) of Korea grant funded by the Korea government (No. 2014R1A2A2A01006642).

References

[1] F. Banterle, P. Ledda, K. Debattista and A. Chalmers, "Inverse tone mapping," *Proc. 4th Int. Conf. Comput. Graph. Interact. Techn*., pp. 349-356, Nov. 2006.

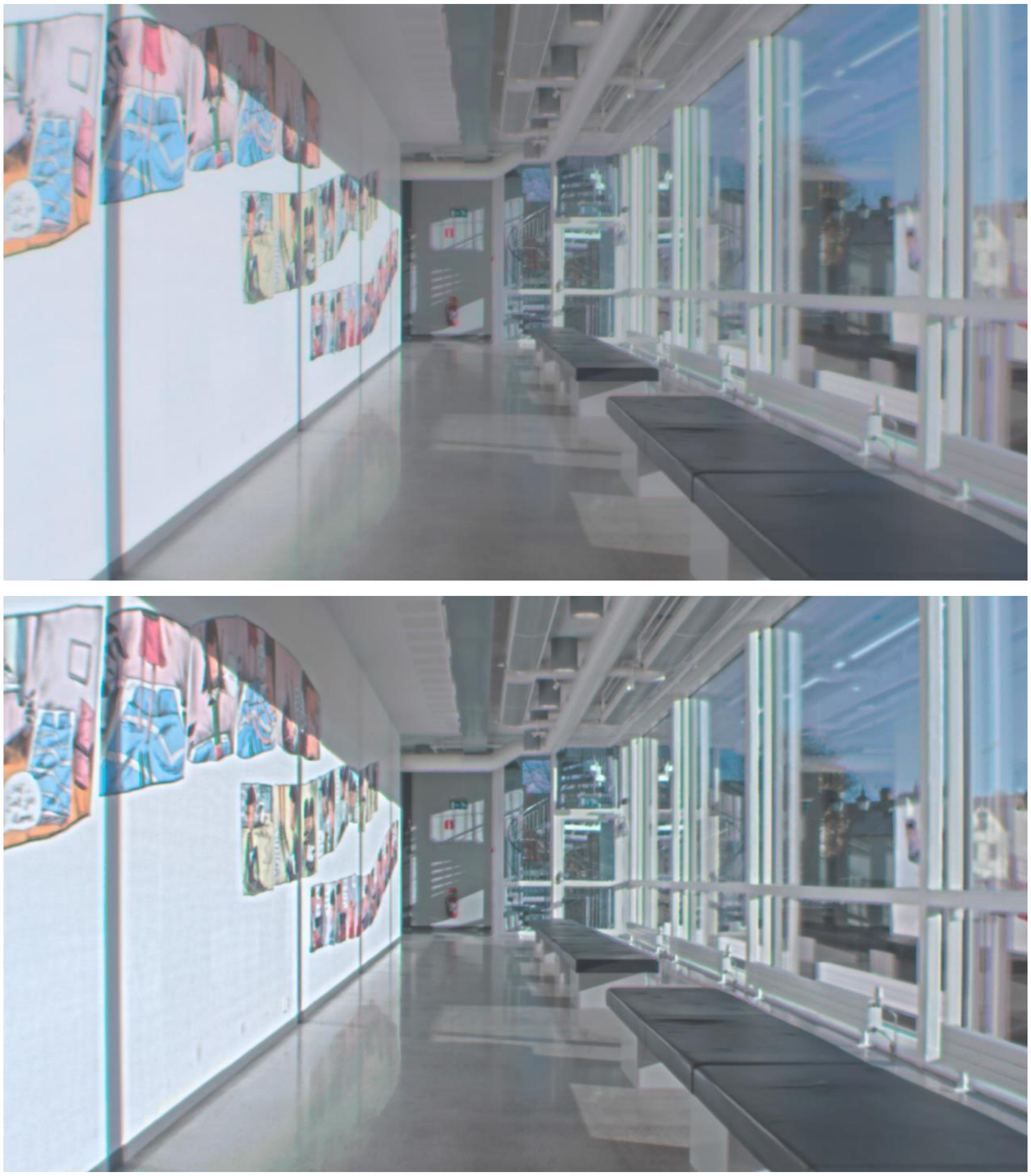

Fig. 4. Comparison of HDR output images: Top - conventional method; bottom - proposed method.

[2] A. O. Akyüz, R. Fleming, B. E. Riecke, E. Reinhard and H. H. Bülthoff, "Do HDR displays support LDR content?: A psychophysical evaluation," *ACM Trans. Graph.*, vol. 26, no. 3, pp. 1-7, July 2007.

[3] L. Meylan, S. Daly and S. Süsstrunk, "The reproduction of specular highlights on high dynamic *range* displays," *Proc. IST/SID*, pp. 333-338, Nov. 2006.

[4] A. G. Rempel, M. Trentacoste, H. Seetzen, H. D. Young, W. Heidrich, L. Whitehead and G. Ward, "Ldr2Hdr: On-the-fly reverse tone mapping of legacy video and photographs," *ACM Trans. Graph.*, vol. 26, no. 3, pp. 39-1 – 39-5, July 2007.

[5] F. Banterle, A. Chalmers and R. Scopigno, "Real-time high fidelity inverse tone mapping for low dynamic range content," *Proc. Eurograph. Short Papers*, pp. 349-356, May 2013.

[6] B. Masia, S. Agustin, R. W. Fleming, O. Sorkine and D. Gutierrez, "Evaluation of reverse tone mapping through varying exposure conditions," *ACM Trans. Graph.*, vol. 28, no. 5, pp. 1-8, Dec. 2009.

[7] Y. Huo, F. Yang and V. Brost, "Dodging and burning inspired inverse tone mapping algorithm," *J. Comput. Inf. Syst.*, vol. 9, no. 9, pp. 3461-3468, May 2013.

[8] P.-H. Kuo, C.-S. Tang and S.-Y. Chien, "Content-adaptive inverse tone mapping," *Proc. Visual Communications and Image Processing, 2012*, pp.1-6, Nov. 2012

[9] C.-R. Chen and C.-T. Chiu, "Curve-based and image-based JND contrast analysis for inverse tone mapping operators," *Proc. 18th ICIP*, pp. 345-348, Sept. 2011.

[10] Y. Huo, F. Yang, L. Dong and V. Brost, "Physiological inverse tone mapping based on retina response," *The Visual Computer*, vol. 30, no. 5, pp. 507-517, May 2014.

[11] R. P. Kovaleski and M. M. Oliveira, “High-quality reverse tone mapping for a wide range of exposures,” *Proc. 27th SIBGRAPI Conference on Graphics, Patterns and Images*, pp. 49-56, Aug. 2014.

[12] F. Banterle, A. Artusi, K. Debattista and A. Chalmers, Advanced High Dynamic Range Imaging: Theory and Practice, 2011, CRC Press.

[13] E. Reinhard, G. Ward, S. Pattanaik and P. Debevec, High Dynamic Range Imaging: Acquisition Display and Image-based Lighting, Morgan Kaufmann, 2010.

[14] K. Myszkowski, R. Mantiuk and G. Krawczyk, High Dynamic Range Video, *Synthesis Lectures in Computer Graphics and Animation*, 2008.

[15] E. Reinhard, M. Stark, P. Shirley and J. Ferwerda, “Photographic tone reproduction for digital images,” *ACM Trans. Graph.*, vol. 21, no. 3, pp. 267-276, Jul. 2002.

[16] T. O. Aydin, R. Mantiuk, K. Myszkowski and H.-P. Seidel, “Dynamic range independent image quality assessment,” *ACM Trans. Graph.*, vol. 27, no. 3, pp. 69:1-10, Aug. 2008.

[17] C. Tomasi and R. Manduchi, “Bilateral Filtering for Gray and Color Images,” *Proc. Sixth Int'l Conf. Computer Vision*, pp. 839-846, Jan. 1998.

[18] H. Hyötyniemi, “Multivariate regression—Techniques and tools,” Control Eng. Lab., Helsinki Univ. Technol., Helsinki, Finland, Tech. Rep. 125, 2001.

[19] BT. 1886, “Reference electro-optical transfer function for flat panel displays used in HDTV studio production,” Recommendation ITU-R, BT. Series Broadcasting service, Mar. 2011.

[20] SMPTE ST2084:2014, “High Dynamic Range Electro-Optical Transfer Function of Mastering Reference Displays,” the Society of Motion Picture and Television Engineers Journal, Aug. 2014.

Author Biography

Dae-Eun Kim received the B.S. degree in electrical engineering from Korea University, Seoul, South Korea, in 2011, and the M.S. degree in electrical engineering from KAIST, Daejeon, South Korea, in 2014. He is currently pursuing the Ph.D. degree in electrical engineering at Korea Advanced Institute of Science and Technology (KAIST). His research interests include image processing, video coding, and pattern recognition.

Munchurl Kim received the B.E. degree in electronics from Kyungpook National University, Daegu, Korea, in 1989, and the M.E. and Ph.D. degrees in Electrical and Computer Engineering from the University of Florida, Gainesville, in 1992 and 1996, respectively. He is a Full Professor of the School of Electrical Engineering, KAIST. His current research interests include high performance video coding, perceptual video coding, visual quality assessments on 3-D/UHD video, computational photography, machine learning and pattern recognition.