Image Contrast Enhancement Using Classified Virtual Exposure Image Fusion

Chang-Hsing Lee, *Member,* IEEE, Ling-Hwei Chen, and Wei-Kang Wang

Abstract — *In our daily life, digital cameras and smart phones have been widely used to take pictures. However, digital cameras and smart phones have a limited dynamic range, which is much lower than that human eyes can perceive. Thus, the photographs taken in high dynamic range scenes often exhibit under-exposure or over-exposure artifacts in shadow or highlight regions. In this study, an image fusion based approach, called classified virtual exposure image fusion (CVEIF), is proposed for image enhancement. First, a function imitating the F-stop concept in photography is designed to generate several virtual images having different intensity. Then, a classified image fusion method, which blends pixels in distinct luminance classes using different fusion functions, is proposed to produce a fused image in which every image region is well exposed. Experimental results on four different kinds of generic images, including a normal image, a low-contrast images, a backlight image, and a dark scene image, have shown that the proposed CVEIF approach produced more pleasingly enhanced images than other methods*[1].

Index Terms — Classified virtual exposure image fusion, contrast enhancement, exposure fusion, image fusion.

I. INTRODUCTION

Digital cameras and smart phones have been widely used to take pictures in our daily life. However, digital cameras and smart phones have a limited dynamic range, which is much lower than that human eyes can perceive. As a result, people are not always pleasing with the photographs taken in high dynamic range scenes because they often exhibit under-exposure or over-exposure artifacts in shadow or highlight regions. The common defects found in real-life images include 1) a normal image with proper illumination/exposure but some regions are slightly under-exposed; 2) a backlight image with over-exposed and/or under-exposed regions; 3) a low-contrast image due to insufficient illumination/exposure; 4) a dark scene image which were taken in the night without using a photoflash. If the illumination/contrast of the acquired image is improper, a post-processing procedure using an image enhancement method is needed in order to produce an image having better quality. Many software or image enhancement methods were developed to cope with these problems. In general, image enhancement methods can be classified into four categories: histogram-based methods [1]-[12], transform-based methods [1], [13], [14], exposure-based methods [15], [16], and image fusion based methods [17]-[19].

Histogram equalization (HE) [1] is the most well-known technique for image enhancement. HE uses a non-linear mapping function to produce an enhanced image with its histogram approximating a uniform distribution. However, HE fails to produce pleasing pictures owing to three common drawbacks: 1) false contour; 2) amplified noises; 3) washed-out appearance. Pizer et al. [2] proposed a local HE method called adaptive histogram equalization. First, an image is divided into several non-overlapping blocks. Then, HE is applied on each block independently. Finally, the enhanced blocks are fused together using bilinear interpolation in order to reduce blocking artifacts. Some brightness preservation HE methods [3]-[12] tried to preserve the original brightness to some extent, which is essential for consumer electronic products. These methods first divide the histogram into two [3]-[8] or more [9]-[12] sub-histograms and then apply HE on each sub-histogram independently. The main drawback of brightness preservation methods is that sometimes they may produce unnatural artifacts because some regions may be enhanced excessively.

For transform-based methods [1], [13], [14], a transformation function (e.g., power-law or logarithmic function) is defined to map an input luminance value into an output one. These methods were widely provided in many consumer electronic products or software. Typically, some device-dependent parameters have to be specified in advance. The transform-based methods generally can produce a properly enhanced image for either under-exposed or over-exposed images by selecting appropriate parameters [1]. However, if an image has both under-exposed and over-exposed regions, the transform-based methods fail to produce appropriate contrast on both regions. Moroney [13] proposed an enhancement approach based on pixel-by-pixel gamma correction with a non-linear masking. The gamma correction of each pixel depends on the values of its neighboring pixels. Nevertheless, it may produce halo effects near edges. Thus, Schettini et al. [14] proposed a local and image-dependent exponential correction function for contrast enhancement in which the bilateral filter is used as the

[1] This work was supported in part by the National Science Council of R.O.C. under contract NSC-101-2221-E-216-033.

C. H. Lee is with the Department of Computer Science and Information Engineering, Chung Hua University, Taiwan (e-mail: chlee@chu.edu.tw).

L. H. Chen and W. K. Wang are with the Institute of Multimedia Engineering, National Chiao Tung University, Hsinchu, Taiwan (e-mail: lhchen@cc.nctu.edu.tw, bishop@debut.cis.nctu.edu.tw).

Contributed Paper
Manuscript received 08/13/12
Current version published 12/28/12
Electronic version published 12/28/12.

mask of the exponential correction function in order to reduce the halo effect. However, the global contrast of the whole image was reduced as well.

Exposure-based methods [15], [16] tried to adjust the exposure level of an image using a mapping function between the light values and the pixel values of interested objects. Battiato et al. [15] proposed a content dependent exposure correction approach using the camera response curve to adjust the exposure levels. Since this approach was specifically designed for interested regions, it can produce pleasing results in interested regions; whereas it may lead to poorer illumination in other regions. Safonov et al. [16] developed an enhancement method for global and local correction of various exposure defects. Their approach is based on contrast stretching and alpha-blending of the brightness of the original image and the estimated reflectance. The main problem with this approach is that it might exhibit insufficient illumination for some regions.

Image fusion based methods [17]-[19] aimed to combine relevant information from multiple images taken from the same scene in order to produce a fused image, which is more informative than each individual one. In this paradigm, several "pseudo images" or "virtual images" have to be generated from a single input image before image fusion. Hsieh et al. [17] used a linear function to fuse the input image and HE enhanced image to get a fused one. Pei et al. [18] generated two images, a HE enhanced image and a sharpened image using Laplacian operator, and then fused their discrete wavelet transform (DWT) coefficients to get a fused image having higher contrast and sharpness. Lim et al. [19] applied an intensity mapping function to the input image to generate multiple images having different exposure. The intensity mapping function can be either 1) estimated from a set of images taken by the same camera in order to imitate the camera response function or 2) expressed explicitly in terms of a power-law function. In the first case, several images taken by the same camera should be provided for learning the camera response function. For the second case, parameters of the power-law function should be chosen carefully in order to get a high contrast fused image.

In this study, an image fusion based approach, named classified virtual exposure image fusion, will be proposed for image contrast enhancement. The main contributions are as follows. First, a function imitating the F-stop concept in photography is designed to generate several virtual images having different intensity. Second, a classified image fusion method, which blends pixels in distinct luminance classes using different fusion functions, is proposed to produce a high-contrast image in which every image region is well exposed.

II. Proposed Classified Virtual Exposure Image Fusion Approach For Image Contrast Enhancement

In this study, an image fusion based approach, called classified virtual exposure image fusion (CVEIF), will be proposed for image contrast enhancement. Image fusion have been widely developed for producing high quality images in applications such as remote sensing [20]-[22], medical imaging [22], high dynamic range imaging (HDRI) [22]-[24], multi-focus imaging [22], [25], etc. In remote sensing or medical imaging, the input images captured from different sensors, having variant spatial and spectral properties, are combined to generate a high quality fused image. In HDRI, several input images taken with distinct exposure time, resulting in several images having different intensity, are blended together to produce a wide dynamic range image. In multi-focus image fusion, some input images captured using variant foci [25], with each one containing some objects in focus, are fused together to obtain an image in which all relevant objects are in focus. Acquisition of several images having different exposure or foci is a prerequisite for these applications. However, for image contrast enhancement, only one input image is given. Thus, several "pseudo images" or "virtual images" have to be generated from the input image to realize a image fusion system.

Since the proposed CVEIF approach works on luminance image, each input color image is first converted to the luminance image. In this study, the luminance value of each pixel is converted from the red, green, and blue color values using the following conversion function:

$$Y(x,y) = 0.299 \cdot R(x,y) + 0.587 \cdot G(x,y) + 0.114 \cdot B(x,y) \quad (1)$$

where $R(x, y)$, $G(x, y)$, and $B(x, y)$ denote the red, green, and blue color values of a pixel located at (x, y). Then, several virtual images having different intensity, realized by setting different imitative F-stops, will be generated. Meanwhile, a multilevel thresholding algorithm will be employed to classify all pixels in the input image different three luminance classes according to their luminance values. Then, a classified image fusion method, which fuses pixels in distinct luminance classes using different fusion functions, will be proposed to obtain a fused image with appropriate exposure on every image region. The block diagram of the proposed image contrast enhancement system is depicted in Fig. 1.

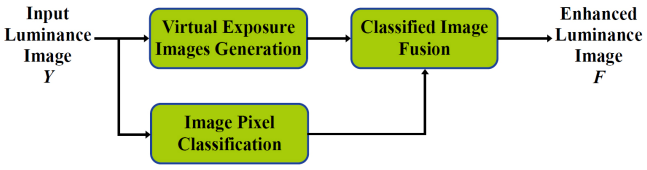

Fig. 1. Block diagram of the proposed CVEIF approach.

A. Generation of Virtual Exposure Images

In photography, exposure refers to how much light will reach the image sensors on digital cameras. To determine the correct exposure, we have to select an appropriate combination of shutter speed and F-stop. Shutter speed controls how long the shutter is open, determining the exposure time that the light of the scene reaches the image sensors. F-stop is used to control the size of the aperture, which is the hole the light of the scene passes through in a camera. Modern cameras use a standard F-stop scale, which is an approximately geometric sequence of

numbers that corresponds to the sequence of the powers of the square root of 2. Specifically, the standard F-stop number runs as follows: F1.4, F2, F2.8, F4, F5.6, F8, F11, F16, F22, and so on. Each stop represents a halving/doubling of the amount of light from its immediate predecessor/successor. For example, F1.4 allows the double of light through than F2, and 4 times more light through than F2.8. In this study, we take a half-step along this scale to make an exposure difference of "half a stop." That is, the generated luminance values associating with each pixel approximate a geometric sequence with common ratio $\sqrt{2}$. Let $Y(x, y)$ denote the luminance value of the input image $\mathbf{Y}$ at spatial location (x, y), and assume that each luminance value is an integer in the interval [0, 255]. The luminance value of each pixel in the k-th virtual image $\mathbf{Y}_k$ can be expressed as

$$Y_k(x, y) = \begin{cases} Y(x, y) \times \left(\sqrt{2}\right)^{-k}, & \text{if } Y(x, y) \times \left(\sqrt{2}\right)^{-k} < 255 \\ 255, & \text{otherwise} \end{cases} \quad (2)$$

In this study, N high exposure brighter images (with k = -N, -N+1, …, -1) and N low exposure darker images (with k = 1, 2, …, N) will be generated first (please see Fig. 2 for an example). From these generated virtual exposure images, we can see that, as the exposure increases, dark regions become more and more clear (see the central building) whereas brighter regions (see the sky area) become saturated.

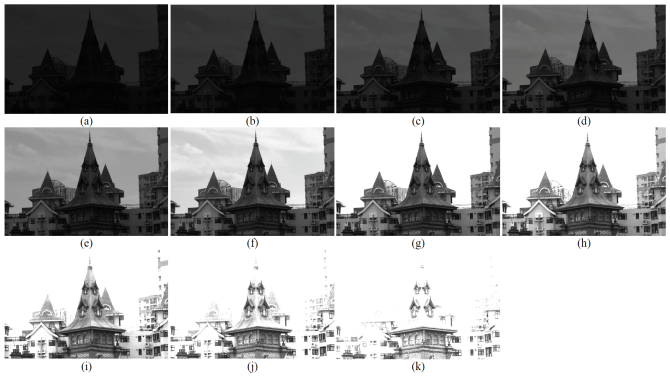

Fig. 2. Generated virtual exposure images (a) k = 5 (b) k = 4 (c) k = 3 (d) k = 2 (e) k = 1 (f) k = 0 (original image) (g) k = -1 (h) k = -2 (i) k = -3 (j) k = -4 (k) k = -5.

Among these 2N+1 virtual images, only those having some relevant informative regions will be selected for image fusion. That is, those images which are completely under-exposed or completely over-exposed will not be used in the image fusion process in an attempt to yield a high informative fused image. To this end, an anchor image (the image with the most proper exposure) among these 2N+1 virtual images will be selected first. The anchor image as well as its preceding M ($M \le N$) lower exposure images and succeeding M higher exposure images, resulting in an image set consisting of 2M+1 images, will be used for image fusion. The anchor image is found by evaluating the average luminance of each virtual exposure image. Let μ_k denote the average luminance of all pixels in $\mathbf{Y}_k$, k = -N, …, N. The exposure image with its average luminance closest to 128 (the middle value of the luminance interval [0, 255]) will be selected as the anchor image. That is, the index anc of the anchor image $\mathbf{Y}_{anc}$ can be determined by the following equation:

$$anc = \arg \min_{k=-N,\cdots,N} | \mu_k - 128 | \quad (3)$$

Besides $\mathbf{Y}_{anc}$, its preceding M lower exposure images and succeeding M higher exposure images, denoted by $\mathbf{Y}_{anc-M}$, …, $\mathbf{Y}_{anc}$, …, $\mathbf{Y}_{anc+M}$, will constitute the set of virtual images for image fusion.

B. Image Pixel Classification

In this study, the pixels in the input image will be classified into three classes: dim class (denoted by $\mathbf{Y}_L$), well-exposed class (denoted by $\mathbf{Y}_M$), and bright class (denoted by $\mathbf{Y}_H$), according to their luminance values. From the classification result, pixels in different classes will be blended using different fusion functions. To this end, the multilevel thresholding algorithm proposed by Liao et al. [26] is employed to find two thresholds, denoted by Thd_0 and Thd_1 ($Thd_0 < Thd_1$), such that the input image $\mathbf{Y}$ can be decomposed into three subimages:

$$\mathbf{Y} = \mathbf{Y}_L \cup \mathbf{Y}_M \cup \mathbf{Y}_H \quad (4)$$

where $\mathbf{Y}_L$, $\mathbf{Y}_M$, and $\mathbf{Y}_H$ respectively consist of pixels with luminance values smaller than Thd_0, in-between Thd_0 and Thd_1, and larger than Thd_1:

$$\mathbf{Y}_L = \{Y(x, y) \mid Y(x, y) < Thd_0, Y(x, y) \in \mathbf{Y}\} \quad (5)$$

$$\mathbf{Y}_M = \{Y(x, y) \mid Thd_0 \le Y(x, y) \le Thd_1, Y(x, y) \in \mathbf{Y}\} \quad (6)$$

$$\mathbf{Y}_H = \{Y(x, y) \mid Y(x, y) > Thd_1, Y(x, y) \in \mathbf{Y}\} \quad (7)$$

Fig. 3 shows the image pixel classification result of the input image shown in Fig. 2(f). We can see that the pixels in the sky region belong to bright class, most of the pixels in the central building and the windows in the other buildings belong to dim class, and the others are attributed to well-exposed class. To provide appropriate exposure on every region (particularly the central building), the proposed CVEIF approach aims to blend pixels in distinct classes using different fusion functions.

Fig. 3. Image pixel classification result of the image shown in Fig. 2(f).

C. Classified Image Fusion

In this study, a weighted average approach will be employed to blend together the 2M+1 virtual images with weights computed from the proposed quality measure. First, a weight map, indicating the contribution of each pixel to the

fused image, is generated for each virtual image to guide the fusion process. For each pixel, Mertens et al. [27] combined the information from different measures, including contrast, saturation, and well-exposedness, into a scalar weight value. Since the proposed classified image fusion method is conducted on the luminance image, the weight map considers only the contrast and well-exposedness measures. The contrast measure tries to preserve the detailed parts of an image such as edge or texture information. The well-exposedness measure is used to find proper exposure for every pixel. In this study, we combine the concept of just-noticeable-distortion (JND) model of the human visual system (HVS) in the contrast measure in order to prevent from amplifying noises. Further, for each pixel, a classified well-exposedness measure is proposed to find its proper luminance value.

1) JND-based Contrast Measure: In this study, the JND model of HVS is combined in the contrast measure to prevent from amplifying noises. For a pixel $\boldsymbol{p}$ located at (x, y) in virtual image Y_k, we first compute the maximum, minimum, and average luminance values of its eight neighbors within the 3×3 window centered at $\boldsymbol{p}$, denoted by $Y_k^{\max}(x, y)$, $Y_k^{\min}(x, y)$, and $Y_k^{\text{avg}}(x, y)$. Then, the difference between $Y_k^{\max}(x, y)$ and $Y_k^{\min}(x, y)$ is evaluated:

$$Y_k^{\text{dif}}(x, y) = Y_k^{\max}(x, y) - Y_k^{\min}(x, y) \tag{8}$$

This difference value provides a simple measure of the contrast value around pixel $\boldsymbol{p}$. If the difference value is smaller than the visibility threshold of HVS, indicating that there exist no visible edges of texture information around pixel $\boldsymbol{p}$, we set the contrast value 0. Otherwise, the contrast value is defined as $Y_k^{\text{dif}}(x, y)$. In addition, because different quality measures have distinct dynamic ranges and to prevent from the computed weight values being zero, we define the contrast value of pixel $\boldsymbol{p}$ as follows:

$$C_k^t(x, y) = \begin{cases} 1/256, & \text{if } Y_k^{\text{dif}}(x, y) < JND(Y_k^{\text{avg}}(x, y)) \\ (Y_k^{\text{dif}}(x, y) + 1)/256, & \text{otherwise} \end{cases} \tag{9}$$

where $C_k^t(x, y)$ denotes the contrast value of pixel $\boldsymbol{p}$, $JND(\cdot)$ is the visibility threshold function providing the JND that HVS can perceive. In this study, the JND model proposed by Chou and Li [28] is employed to design the visibility threshold function, which can be described by the following equation:

$$JND(g) = \begin{cases} T_0(1 - (g/127)^{0.5}) + 3, & \text{if } g \le 127 \\ \gamma(g - 127) + 3, & \text{otherwise} \end{cases} \tag{10}$$

where g is the luminance value in the interval [0, 255], the parameters T_0 and γ depend on the viewing distance between a tester and the monitor. In this study, T_0 and γ are set to be 17 and 3/128 according to the subjective experiments conducted by Chou and Li [28].

2) Classified Well-exposedness Measure: Well-exposedness assesses how well a pixel is exposed. Traditionally, a luminance value close to the middle value of the luminance interval is considered well-exposed, whereas a luminance value near the boundary of the luminance interval is regarded as poor-exposed. Thus, the well-exposed measure is generally defined by the following Gaussian function [27] [29]:

$$E_k^t(x, y) = \exp\left(-\frac{(Y_k(x, y) - 128)^2}{2\sigma^2}\right) \tag{11}$$

where 128 (viewed as the desired target luminance value) is the middle value of the luminance interval [0, 255], $E_k^t(x, y)$ denotes the well-exposedness value of the pixel located at (x, y), σ is the standard deviation of the Gaussian curve (set as 0.2×luminance range). This definition gives those pixels with luminance value close to 128 a larger well-exposedness value and a small weight value is assigned to those pixels with luminance values close to 0 or 255. Nevertheless, such a definition does not consider the original brightness of the pixels in the image. That is, the luminance values of both dark and bright pixels will be moved toward 128 in the fused image. As a result, the global contrast of the fused image will be reduced although all pixels are well exposed based on the well-exposedness measure defined in (11). To deal with this problem, the proposed classified well-exposedness measure defines distinct desired target luminance values for pixels belonging to different classes. Specifically, pixels in well-exposed class ($\mathbf{Y_M}$) have a desired target luminance value of 128, whereas pixels in bright class ($\mathbf{Y_H}$) and dim class ($\mathbf{Y_L}$) will be assigned a desired target luminance value larger than 128 and smaller than 128, respectively. Let μ_L and μ_H denote the average luminance values of all pixels in $\mathbf{Y_L}$ and $\mathbf{Y_H}$, respectively. The desired target luminance value for pixels in class $\mathbf{Y_L}$, denoted by Y_L^t, is defined by the following function:

$$Y_L^t = \begin{cases} 64, & \text{if } \mu_L > 64 \\ \mu_L, & \text{if } 32 \le \mu_L \le 64 \\ 64, & \text{if } \mu_L < 32 \text{ and } r_L > 0.5 \\ r_L \times 128, & \text{if } \mu_L < 32 \text{ and } 0.25 \le r_L \le 0.5 \\ 32, & \text{if } \mu_L < 32 \text{ and } r_L < 0.25 \end{cases} \tag{12}$$

where r_L is the proportion of the number of pixels in $\mathbf{Y_L}$, that is,

$$r_L = \frac{N_L}{N_L + N_M + N_H} \tag{13}$$

where N_L, N_M, and N_H are the number of pixels in $\mathbf{Y_L}$, $\mathbf{Y_M}$, and $\mathbf{Y_H}$. Note that the desired target luminance value for pixels in $\mathbf{Y_L}$ is defined in the interval [32, 64] according to the luminance distribution. If $\mu_L > 64$, indicating that the input image is a bright one, we set the desired target luminance value as 64. On the other hand, if $\mu_L < 32$, indicating that the input image is a dark one, the desired target luminance value for class $\mathbf{Y_L}$ will be defined according to the proportion of the number of pixels in $\mathbf{Y_L}$, r_L. If r_L is large, indicating that the input image consists of many dark pixels, a large desired target luminance level is chosen, and vice versa.

For pixels belonging to well-exposed class $\mathbf{Y}_{\boldsymbol{M}}$, the desired target luminance value is defined as 128, that is,

$$Y_M^t = 128 \tag{14}$$

Since the luminance values of dark pixels are often set toward higher ones, to preserve and even increase the global contrast of the fused image, the desired target luminance value for pixels in bright class $\mathbf{Y}_{\boldsymbol{H}}$, Y_H^t, is defined in a similar way. Specifically, Y_H^t is defined as follows:

$$Y_H^t = \begin{cases} \mu_H, & \text{if } \mu_H > 224 \\ 224, & \text{if } 192 \le \mu_H \le 224 \\ 192, & \text{if } \mu_H < 192 \end{cases} \tag{15}$$

According to the definition of desired target luminance values for different luminance classes, the classified well-exposedness measure is defined as follows:

$$E_k^t(x, y) = \begin{cases} \exp\left(-\dfrac{(Y_k(x, y) - Y_L^t)^2}{2\sigma_L^2}\right), & \text{if } Y(x, y) \in \mathbf{Y}_{\boldsymbol{L}} \\ \exp\left(-\dfrac{(Y_k(x, y) - Y_M^t)^2}{2\sigma_M^2}\right), & \text{if } Y(x, y) \in \mathbf{Y}_{\boldsymbol{M}} \\ \exp\left(-\dfrac{(Y_k(x, y) - Y_H^t)^2}{2\sigma_H^2}\right), & \text{if } Y(x, y) \in \mathbf{Y}_{\boldsymbol{H}} \end{cases} \tag{16}$$

where σ_L, σ_M, and σ_H are the desired target standard deviation of the Gaussian curve for $\boldsymbol{Y_L}$, $\boldsymbol{Y_M}$, and $\boldsymbol{Y_H}$ (in this study, we set $\sigma_L = 32$, $\sigma_M = 64$, and $\sigma_H = 32$).

3) Classified Weight Map Generation: Finally, the weight map, $\mathbf{W}_k^t$, associated with virtual image $\mathbf{Y}_k$ (k = *anc*-M, …, *anc*+M) is obtained by combining the information from both JND-based contrast measure and classified well-exposedness measure through multiplication:

$$W_k^t(x, y) = C_k^t(x, y) \times E_k^t(x, y) \tag{17}$$

To obtain a consistent result, we normalize the weight values of the 2M+1 weight maps such that for each pixel location the sum of the 2M+1 weight values equals one:

$$W_k(x, y) = \frac{W_k^t(x, y)}{\sum_{i=anc-M}^{anc+M} W_i^t(x, y)} \tag{18}$$

Fig. 4 and Fig. 5 show the weight maps generated by the proposed classified weight map generation method and that proposed by Mertens et al. [27]. From Fig. 5, we can observe that the weight values in the sky region of those low exposure images have larger weight values than those in high exposure ones. Consequently, the luminance value of the sky region will decrease in the fused image (see Fig. 6(a) to Fig. 6(e)). On the other hand, the proposed classified weight map generation method can produce proper luminance values for pixels in the sky region (see Fig. 6(j)).

4) Classified Image Fusion in the Discrete Wavelet Transform Domain: In this study, the proposed classified image fusion is performed in the discrete wavelet transform (DWT) domain to avoid annoying seams at pixels having sharp weight value transitions [29]. By using DWT, a multi-resolution representation can be constructed for each virtual image. Given a two-dimensional image, one-dimensional DWT can be successively applied to the rows and columns of the image. This process accomplishes one level of decomposition and results in four low-resolution subimages, denoted by LL, LH, HL, and HH. The subimage LL corresponds to a coarse approximate image, whereas the subimages LH, HL, and HH correspond to vertical, horizontal, and diagonal details. For multi-resolution wavelet decomposition, the subimage LL can further be decomposed into the other four subimages using the same decomposition procedure. Such a decomposition process is repeated until the desired number of levels determined by the application is reached. If L-level of wavelet decomposition is performed, we can obtain 3L+1 subimages. The reconstruction of the image can be carried out by reversing the above decomposition procedure level by level until the image is fully reconstructed.

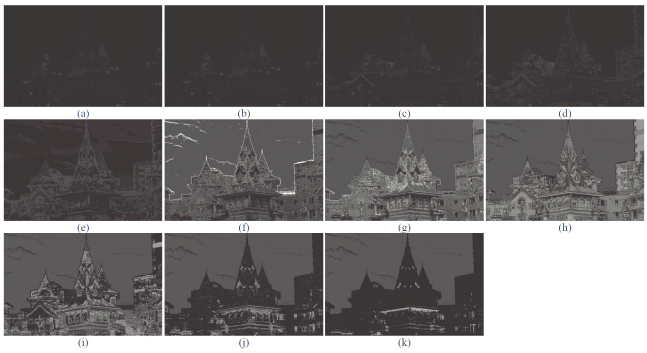

Fig. 4. Weight maps of different exposure images generated by using the proposed classified weight map generation method (a) Y_5 (b) Y_4 (c) Y_3 (d) Y_2 (e) Y_1 (f) Y_0 (g) Y_{-1} (h) Y_{-2} (i) Y_{-3} (j) Y_{-4} (k) Y_{-5}.

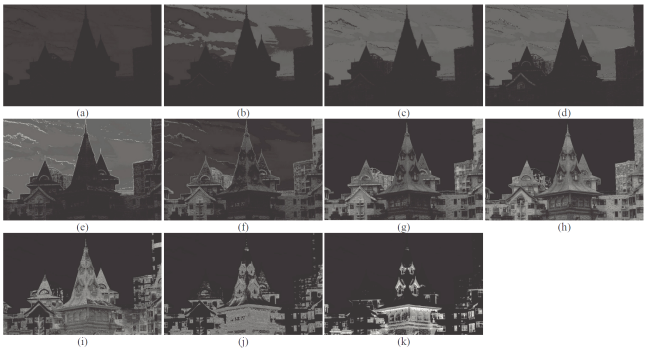

Fig. 5. Weight maps of different exposure images generated by using the method proposed by Mertens et al. [27] (a) Y_5 (b) Y_4 (c) Y_3 (d) Y_2 (e) Y_1 (f) Y_0 (g) Y_{-1} (h) Y_{-2} (i) Y_{-3} (j) Y_{-4} (k) Y_{-5}.

Let $\mathbf{Y}_k^{l,\theta}$ denote the wavelet subimage with direction θ ($\theta \in$ {LL, LH, HL, HH}) at level l ($1 \le l \le L$) of the wavelet transform of virtual image $\mathbf{Y}_k$. For each weight map, a corresponding Gaussian pyramid will be constructed. Let $\mathbf{W}_k^l$ denote the subimage at level l of the Gaussian pyramid of the weight map $\mathbf{W}_k$ associated with exposure image $\mathbf{Y}_k$. Blending is carried out independently for every wavelet subimage with direction θ and level l, with the subimage at the l-th level of

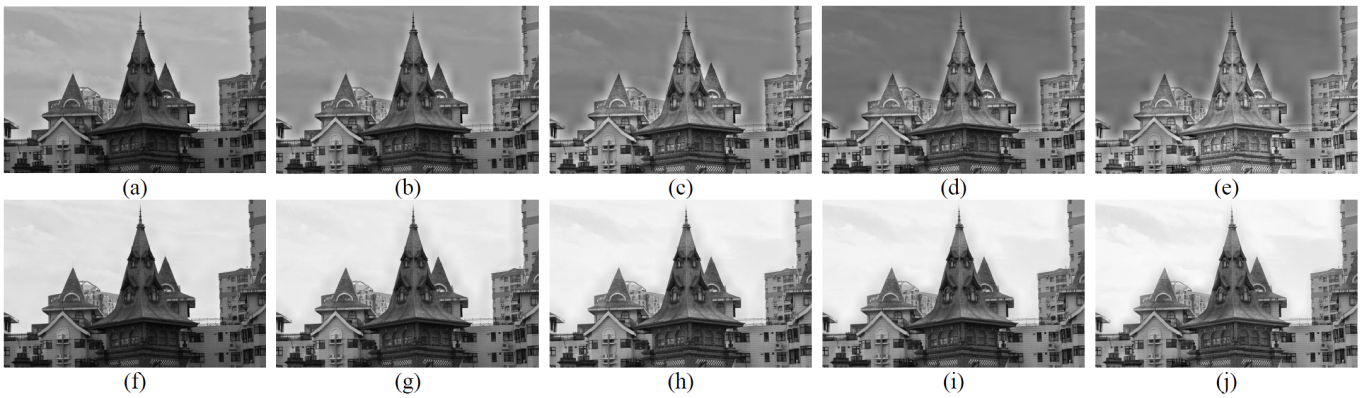

Fig. 6. Comparison of the fused images using wavelet based exposure fusion method proposed by Mertens et al. [27] ((a)–(e)) and the proposed CVEIF approach ((f)–(j)) by blending together 2*M*+1 exposure images (a) *M* = 1 (b) *M* = 2 (c) *M* = 3 (d) *M* = 4 (e) *M* = 5 (f) *M* = 1 (g) *M* = 2 (h) *M* = 3 (i) *M* = 4 (j) *M* = 5.

the Gaussian pyramid of the weight map serving as the weights:

$$F^{l,\theta}(x, y) = \sum_{k=anc-M}^{anc+M} Y_k^{l,\theta}(x, y) W_k^l(x, y) \quad (19)$$

where $F^{l,\theta}(x, y)$ is the fused wavelet coefficient located at (x, y) in the wavelet subimage with direction θ and level l. By applying inverse DWT on these wavelet subimages, we can reconstruct the fused image **F**. Fig. 6 compares the fused images obtained by blending different number of virtual images using the proposed CVEIF approach with wavelet based exposure fusion method proposed by Mertens et al. [27]. We can see that the proposed CVEIF approach yields better exposure on sky regions. In addition, improved contrast can be obtained as the number of exposure images used for image fusion increases. Thus, in the following experiments, we set the parameter $M = 5$.

D. Color Reconstruction

From the fused luminance image **F**, the R, G, B color values of each pixel can be reconstructed by using the following formula in order to prevent relevant hue shift and color desaturation [30]:

$$R^r(x, y) = \frac{1}{2}\left(\frac{F(x, y)}{Y(x, y)}\left(R(x, y) + Y(x, y)\right) + R(x, y) - Y(x, y)\right) \quad (20)$$

$$G^r(x, y) = \frac{1}{2}\left(\frac{F(x, y)}{Y(x, y)}\left(G(x, y) + Y(x, y)\right) + G(x, y) - Y(x, y)\right) \quad (21)$$

$$B^r(x, y) = \frac{1}{2}\left(\frac{F(x, y)}{Y(x, y)}\left(B(x, y) + Y(x, y)\right) + B(x, y) - Y(x, y)\right) \quad (22)$$

III. Experimental Results

In this study, four different kinds of generic images will be used for performance comparison: 1) a normal image with proper illumination/exposure but some regions are slightly under-exposed; 2) a low contrast image due to insufficient illumination/exposure; 3) a backlight image with over-exposed and/or under-exposed regions; 4) a dark scene image which was taken in the night without using a photoflash. The proposed CVEIF approach will be compared with some other methods, including 1) HE [1]; 2) Battiato's exposure correction (EC) algorithm [15]; 3) Schettini's local gamma correction (LCC) algorithm [14]; 4) Safonov's shadow correction (SC) algorithm [16]; 5) Pei's wavelet-based image fusion (WIF) method [18]; 6) Merten's exposure fusion (EF) method [27].

A. Experimental Results on a Normal Image

Fig. 7 shows an image with proper exposure and the enhanced images obtained using different enhancement methods. Fig. 7(a) shows the original image; we can see that the exposure of the whole image is appropriate, except that the central building is a little under-exposed. From Fig. 7(b), we can see that HE produces an image having sharper contrast than the original one. However, the tone of colors in the sky area is also changed to some extent. From Fig. 7(c), Fig. 7(d), and Fig. 7(f), we can see that the central building in the enhanced images using Battiato's EC algorithm, Schettini's LCC algorithm, and Pei's WIF method are clearer than that of original one, but the exposure is also insufficient. In Fig. 7(e), the central building has a better exposure but the global contrast of the image is not good enough. From Fig. 7(g), we can see that an unnatural, low-quality, and low-contrast image is obtained using Merten's EF method. By observing all enhanced images, we can find that the proposed CVEIF approach (see Fig. 7(h)) yields a better, high-contrast image.

B. Experimental Results on a Low Contrast Image

Fig. 8 compares the enhanced results of different methods on a low contrast image. Fig. 8(a) shows the original image; we can see that the exposure is deficient and thus a low-contrast dark image is acquired. From Fig. 8(b), we can see that HE produced a high-contrast image at the cost of desaturating the color of those bright regions. From Fig. 8(d), we can see that Schettini's LCC algorithm produced a washed-out appearance and thus the image looks unnatural. The contrast of the enhanced image using Safonov's SC algorithm is a little better than the original one (see Fig. 8(e)). By carefully observing the white wall in the background, we can see that it becomes gray in the enhanced image using

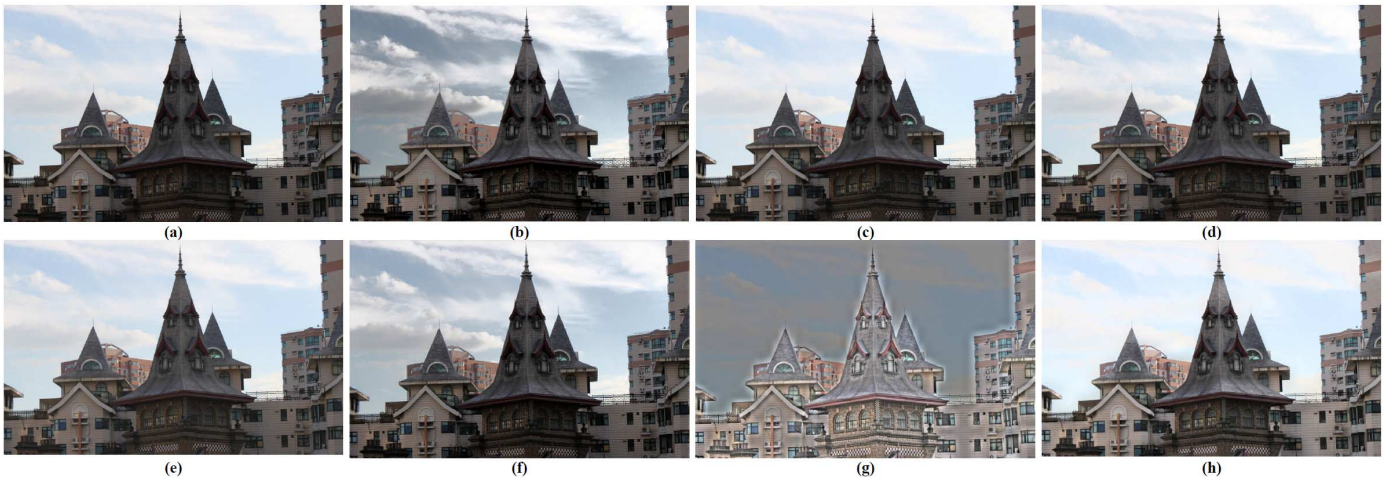

Fig. 7. Enhanced results of a normal image using different methods (a) Original image (b) HE (c) EC (d) LCC (e) SC (f) WIF (g) EF (h) Proposed CVEIF.

Fig. 8. Enhanced results of a low contrast image using different methods (a) Original image (b) HE (c) EC (d) LCC (e) SC (f) WIF (g) EF (h) Proposed CVEIF.

Merten's EF method (see Fig. 8(g)). From Fig. 8(c), Fig. 8(f), and Fig. 8(h), we can see that the proposed CVEIF approach, Pei's WIF method, and Battiato's EC algorithm yield comparably pleasing images with higher contrast.

C. Experimental Results on a Backlight Image

Fig. 9 shows a backlight image and the enhanced images obtained using different enhancement methods. Fig. 9(a) shows the original image; we can see that the foreground pagoda and trees are almost invisible. From Fig. 9(b), we can see that HE yields a clear image at the expense of a bit washed-out appearance (see the leaves near the sunlight). Battiato's EC algorithm, Schettini's LCC algorithm, and Safonov's SC algorithm can enhance the foreground pagoda and trees to some extent, but the global contrast is not acceptable (see Fig. 9(c), Fig. 9(d), and Fig. 9(e)). In Fig. 9(g), we can see that the sky region becomes gray in the enhanced image using Merten's EF method. From Fig. 9(f) and Fig. 9(h), we can see that the proposed CVEIF approach and Pei's WIF method produce images with better contrast. However, by carefully observing the enlarged detail parts of the enhanced images obtained using Pei's WIF method and the proposed CVEIF approach (see Fig. 9(i) and Fig. 9(j)), we can see that severe blocking artifacts exist in the enhanced image using Pei's WIF method.

D. Experimental Results on a Dark Scene Image

Generally, when images were taken in the night without using a photoflash and there exists some lighting in the background, the foreground objects will become dark and unclear due to relatively insufficient illumination in the foreground objects. Fig. 10 shows an image taken in the night and the enhanced images obtained using different enhancement methods. Fig. 10(a) shows the original image; we can see that the foreground subjects are dark and unclear due to the existence of the bright neon light and Chinese lanterns in the scene. By comparing all enhanced images, we can see that the faces in the foreground subjects become clear using HE, Pei's WIF method, and the proposed CVEIF approach (see Fig. 10(b), Fig. 10(f), and Fig. 10(h)). However, HE will change the tone of color a bit (please see the Chinese lanterns). Similarly, blocking artifacts also exist in the enhanced image produced using Pei's WIF method (see Fig. 11).

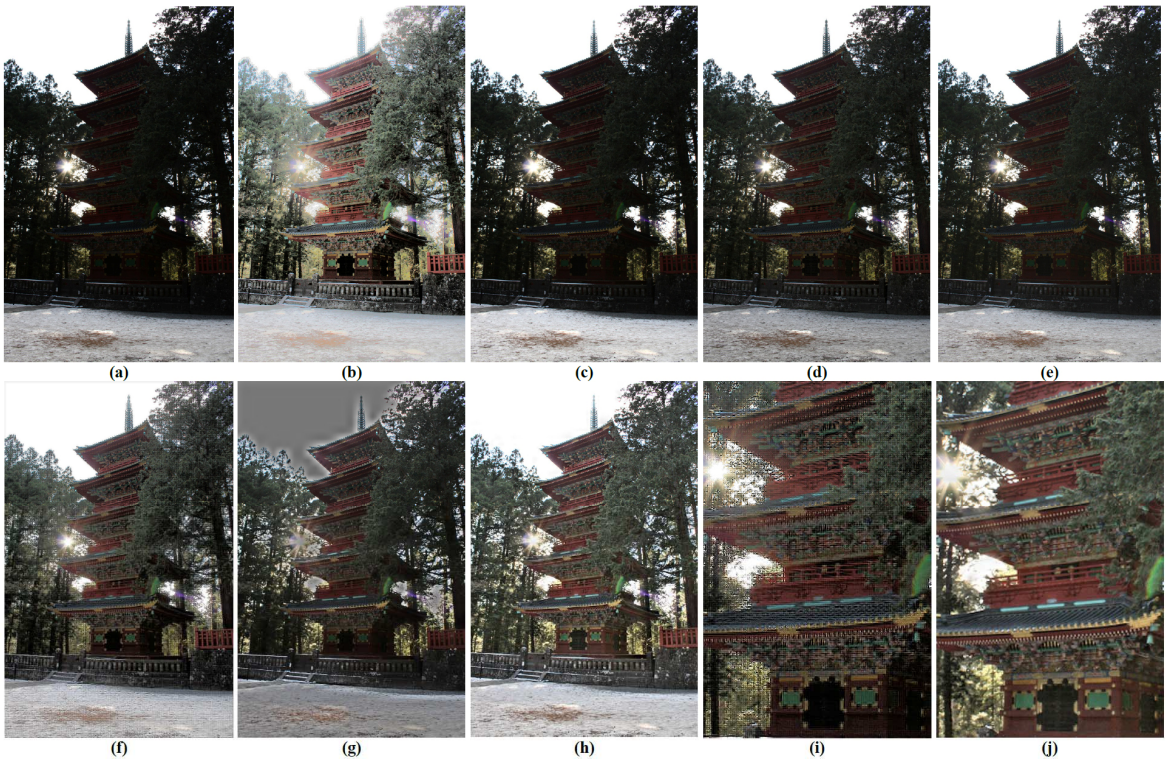

Fig. 9. Enhanced results of a backlight image using different methods (a) Original image (b) HE (c) EC (d) LCC (e) SC (f) WIF (g) EF (h) Proposed CVEIF (i) enlarged detail part of Fig. 9(f) (j) enlarged detail part of Fig. 9(h).

Fig. 10. Enhanced results of a dark scene image using different methods (a) Original image (b) HE (c) EC (d) LCC (e) SC (f) WIF (g) EF (h) Proposed CVEIF.

Fig. 11 Comparison of the enlarged detail parts of Fig. 10(f) and Fig. 10(h)) (a) WIF (b) Proposed CVEIF.

IV. Conclusion

In this study, an image fusion based approach, called classified virtual exposure image fusion (CVEIF), is proposed for image contrast enhancement. First, a function imitating the F-stop concept in photography is designed to produce several virtual images having different intensity. Then, a classified image fusion method, which blends pixels in distinct luminance classes using different fusion functions using the proposed JND-based contrast measure and classified well-exposedness measure, is designed to produce a fused image in which every region is well exposed. Experimental results on four different kinds of generic images have shown that the proposed CVEIF approach yielded more pleasingly enhanced

images than other methods, including HE, Battiato's EC algorithm, Schettini's LCC algorithm, Safonov's SC algorithm, Pei's WIF method, and Merten's EF method.

REFERENCES

[1] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 2nd ed., New Jersey: Prentice-Hall, 2002.

[2] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. H. Romeny, J. B. Zimmerman, and K. Zuiderveld, "Adaptive histogram equalization and its variations," *Comput. Vision, Graph., Image Process.*, vol. 39, no. 3, pp. 355-368, Sep. 1987.

[3] Y. T. Kim, "Contrast enhancement using brightness preserving bi-histogram equalization," *IEEE Trans. Consumer Electron.*, vol. 43, no. 1, pp. 1-8, Feb. 1997.

[4] Y. Wang, Q. Chen, and B. Zhang, "Image enhancement based on equal area dualistic sub-image histogram equalization method," *IEEE Trans. Consumer Electron.*, vol. 45, no. 1, pp. 68-75, Feb. 1999.

[5] S. D. Chen and A. R. Ramli, "Minimum mean brightness error bi-histogram equalization in contrast enhancement," *IEEE Trans. Consumer Electron.*, vol. 49, no. 4, pp. 1310-1319, Nov. 2003.

[6] C. Wang and Z. Ye, "Brightness preserving histogram equalization with maximum entropy: a variational perspective," *IEEE Trans. Consumer Electron.*, vol. 51, no. 4, pp. 1326-1334, Nov. 2005.

[7] C. H. Ooi, N. S. P. Kong, and H. Ibrahim, "Bi-histogram equalization with a plateau limit for digital image enhancement," *IEEE Trans. Consumer Electron.*, vol. 55, no. 4, pp. 2072-2080, Nov. 2009.

[8] N. Sengee, A. Sengee, and H. K. Choi, "Image contrast enhancement using bi-histogram equalization with neighborhood metrics," *IEEE Trans. Consumer Electron.*, vol. 56, no. 4, pp. 2727-2734, Nov. 2010.

[9] S. D. Chen and A. R. Ramli, "Contrast enhancement using recursive mean-separate histogram equalization for scalable brightness preservation," *IEEE Trans. Consumer Electron.*, vol. 49, no. 4, pp. 1301-1309, Nov. 2003.

[10] D. Menotti, L. Najman, J. Facon, and A. de A. Araújo, "Multi-histogram equalization methods for contrast enhancement and brightness preserving," *IEEE Trans. Consumer Electron.*, vol. 53, no. 3, pp. 1186-1194, Aug. 2007.

[11] M. Kim and M. G. Chung, "Recursively separated and weighted histogram equalization for brightness preservation and contrast enhancement," *IEEE Trans. Consumer Electron.*, vol. 54, no. 3, pp. 1389-1397, Aug. 2008.

[12] C. H. Ooi and N. A. M. Isa, "Adaptive contrast enhancement methods with brightness preserving," *IEEE Trans. Consumer Electron.*, vol. 56, no. 4, pp. 2543-2551, Nov. 2010.

[13] N. Moroney, "Local colour correction using non-linear masking," in *Proc. IS&T/SID Eighth Color Imaging Conf.*, 2000, pp. 108-111.

[14] R. Schettini, F. Gasparini, S. Corchs, F. Marini, A. Capra, and A. Castrorina, "Contrast image correction method," *J. Electron. Imaging*, vol. 19, no. 2, 023025, Apr.-June 2010.

[15] S. Battiato, A. Bosco, A. Castorina, and G. Messina, "Automatic image enhancement by content dependent exposure correction," *EURASIP J. Appl. Signal Process.*, vol. 2004, no. 12, pp. 1849-1860, 2004.

[16] I. V. Safonov, M. N. Rychagov, K. Kang, and S. H. Kim, "Automatic correction of exposure problems in photo printer," in *Proc. IEEE Tenth Int. Symp. Consumer Electron.*, 2006, pp. 1-6.

[17] C. H. Hsieh, B. C. Chen, C. M. Lin, and Q. Zhao, "Detail aware contrast enhancement with linear image fusion," in *Proc. 2nd Int. Symp. on Aware Computing*, 2010, pp. 1-5.

[18] L. Pei, Y. Zhao, and H. Luo, "Application of wavelet-based image fusion in image enhancement," in *Proc. 3rd Int. Cong. Image and Signal Process.*, 2010, pp. 649-653.

[19] B. R. Lim, R. H.Park, and S. Kim, "High dynamic range for contrast enhancement," *IEEE Trans. Consumer Electron.*, vol. 52, no. 4, pp. 1454-1462, Nov. 2006.

[20] Z. Wang, D. Ziou, C. Armenakis, D. Li, and Q. Li, "A comparative analysis of image fusion methods," *IEEE Trans. Geosci. Remore sensing*, vol. 43, no. 6, pp. 1391-1402, 2005.

[21] K. Amolins, Y. Zhang, and P. Dare, "Wavelet based image fusion techniques- An introduction, review and comparison," *ISPRS J. of Photogrammetry & Remote Sensing*, vol. 62, pp. 249-263, 2007.

[22] G. Pajares and J. M. de la Cruz, "A wavelet-based image fusion tutorial," *Pattern Recognition*, vol. 37, pp. 1855-1872, 2004.

[23] W. C. Kao, "High dynamic range imaging by fusing multiple raw images and tone reproduction," *IEEE Trans. Consumer Electron.*, vol. 54, no. 1, pp. 10-15, Feb. 2008.

[24] E. Reinhard, G. Ward, S. Pattanaik, and P. Debevec, *High Dynamic Range Imaging: Acquisition, Display and Image-Based Lighting*, Morgan Kaufmann Publishing, 2005.

[25] S. Li and B. Yang, "Multifocus image fusion using region segmentation and spatial frequency," *Image and Vision Comput.*, vol. 26, pp. 971-979, 2008.

[26] P. S. Liao and T. S. Chen and P. C. Chung, "A fast algorithm for multilevel thresholding," *J. Inform. Sci. Eng.*, vol. 17, no. 5, pp. 713-727, 2001.

[27] T. Mertens, J. Kautz, and F. V. Reeth, "Exposure fusion," in *Proc. 15th Pacific Conf. on Comput. Graph. and Applicat.*, 2007, pp. 382-390.

[28] C. H. Chou and Y. C. Li, "A perceptually tuned subband image coder based on the measure of just-noticeable-distortion profile," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 5, no. 6, pp. 467-476, Dec. 1995.

[29] M. H. Malik, S. Asif., and M. Gilani, "Wavelet based exposure fusion," in *Proc. World. Cong. on Eng.*, 2008, pp. 688-693.

[30] S. Sakaue, M. Nakayama, A. Tamura, and S. Maruno, "Adaptive gamma processing of the video cameras for the expansion of the dynamic range," *IEEE Trans. Consumer Electron.*, vol. 41, no. 3, pp. 555-562, Aug. 1995.

BIOGRAPHIES

Chang-Hsing Lee (M'11) was born in Tainan, Taiwan, in 1968. He received the B.S. and Ph.D. degrees from Computer and Information Science, National Chiao Tung University, Hsinchu, Taiwan in 1991 and 1995, respectively. He is currently an Associate Professor in the Department of Computer Science and Information Engineering, Chung Hua University, Hsinchu, Taiwan. His main research interests include audio/music classification, image processing, multimedia information retrieval, and pattern recognition.

Ling-Hwei Chen was born in Changhua, Taiwan, in 1954. She received the B.S. degree from Mathematics and the M.S. degree from Applied Mathematics, National Tsing Hua University, Hsinchu, Taiwan in 1975 and 1977, respectively, and the Ph.D. degree from Computer Engineering, National Chiao Tung University, Hsinchu, Taiwan in 1987. From August 1977 to April 1979 she worked as a research assistant in the Chung-Shan Institute of Science and Technology, Taoyan, Taiwan, From May 1979 to February 1981 she worked as a research associate in the Electronic Research and Service Organization, Industry Technology Research Institute, Hsinchu, Taiwan. From March 1981 to August 1983 she worked as an engineer in the Institute of Information Industry, Taipei, Taiwan. She is now a Professor in the Department of Computer Science, National Chiao Tung University. Her current research interests include image processing, pattern recognition, multimedia compression, content-based retrieval and multimedia steganography.

Wei-Kang Wang was born in Taipei, Taiwan, in 1988. He received the B.S. degree from Computer Science, National Chiao Tung University, Hsinchu, Taiwan in 2010. He is currently a graduate student in the Institute of Multimedia Engineering, National Chiao Tung University, Hsinchu, Taiwan. His main research interest is image processing.