SMART TONE REPRODUCTION

Dun-Yu Hsiao and Hong-Yuan Mark Liao
Institute of Information Science
Academia Sinica, Taipei, Taiwan 115
+886-2-27883799 Ext 1550
silverseraph@gmail.com, liao@iis.sinica.edu.tw

ABSTRACT

In this paper, we propose an effective scheme to enhance the visual details at the minimal cost of user adjustments. The uprising importance of automatic tone reproduction comes from the increasing population of digital archive programs, which contains a large number of images/videos either old, irreproducible, or poorly captured. We attempt to solve above issues by a new local normalization step and an adaptive contrast assessment process. With those two processes, our method can effectively enhance poor quality regions and simultaneously preserving good quality ones with default parameter settings. The experimental results demonstrate that our method is superior to many existing algorithms when applied to aid digital archiving issues.

***Index Terms*—**Image enhancement

1. INTRODUCTION

The lack of sufficient dynamic range and proper shading condition are still challenging issues of modern sensor technology and photography. As a result, important details in a captured image/video can be suppressed and become imperceptible. To recover such details while simultaneously retaining well-defined parts, a number of tone reproduction algorithms have been developed in recent years [1-4]. Based on the strategies employed in the literature, we can classify the existing techniques into two categories: global tone reproduction algorithms and regional tone reproduction algorithms. There are many existing global tone reproduction approaches. Fattal *et al.* [2] proposed a method to suppress the magnitude of large luminance gradients and then preserve fine details by identifying changes in intensity. In [1], Durand *et al.* proposed a bilateral filter to decompose an image into two layers: a large-scale variation base layer and a visibility preserving detail layer. The two layers are produced by performing bilateral filtering, and the relative contrast is subsequently reduced in the large-scale variation layer. However, in the existing global approaches, high contrast regions are inevitably suppressed.

As to regional tone reproduction algorithms, Krawczyk *et al.* [3] proposed decomposing an image into areas of consistent luminance and calculating local lightness values. Since region contrast reproduction schemes process each image region differently, the main drawback of such approaches is that they will indispensably produce unnatural boundaries. No matter global or regional tone reproduction approaches, a common drawback is that the quality of their results heavily depends on how parameters are set.

Different from above approaches, we attempt to use two primary components in our proposed algorithm—local normalization and adaptive contrast assessment—to avoid previously mentioned drawbacks. In order to perform local normalization, our approach starts from acquiring the target image's local maximum and minimum surfaces. By thinking an image as a 3D surface, the surface patches that include the local maxima and the local minima can enclose the entire 3D surface from top and bottom. By normalizing the image toward the local maxima and minima, we can expand an image signal to utilize most part of its dynamic range. Then, we design an adaptive contrast assessment process in our tone reproduction scheme. With local normalization and adaptive contrast assessment working together, our method can directly enhance poor quality regions and at the same time preserving good quality ones like what photographers do in manual image enhancement and requires intrinsically no user parameter adjustments. This merit is important when applying our algorithm to aid digital archiving issues.

2. OUR METHOD

2.1. The Concepts

At present, an image model $I(x,y)$ is commonly regarded as a product of reflectance $R(x,y)$ and luminance $L(x,y)$ at each point of coordinates (x,y), i.e.,

$$I(x,y) = R(x,y) \times L(x,y) \quad (1)$$

Thus, we assume the enhanced image $I'(x,y)$ becomes:

$$I'(x,y) = R'(x,y) \times L'(x,y) \quad (2)$$

For homomorphic filtering on an image, the filter suppresses the luminance part of an image. This filtering makes the overall contrast reduced and the corresponding histogram converges toward the center of a dynamic range. On the other hand, we performed manual image enhancement over more than 150 different poorly captured images. Through the set of experiments, we found that the average distributions of histograms of those manually enhanced images were with low values at both ends of a histogram. The above two observations concur with each other. Under these circumstances, we have to avoid the image luminance being extreme. We express this characteristic as follows:

$$G_{\min} + \delta < L'(x,y) < G_{\max} - \phi, \tag{3}$$

where $G_{\min}$ and $G_{\max}$ are the minimum and maximum value of a dynamic range. δ and ϕ are two constants to prevent extreme luminance values. On the other hand, homomorphic filtering enhances image details by increasing the ratio of reflectance $R(x,y)$ to luminance $L(x,y)$. Accordingly, through the experiments we conducted, we also found the resultant contrasts became larger in the enhanced images than in the originals. Furthermore, Weber's Law also points out a similar concept: a contrast is perceivable only when it is greater than a predetermined threshold. Therefore, we can conclude that the details of an image are visible only when its local contrast is large enough. Having the above mentioned characteristics, we can derive:

$$R'(x,y)/L'(x,y) > R(x,y)/L(x,y), \tag{4}$$

$$Contrast(Region_k(x,y)) > threshold, \tag{5}$$

where $Region_k(.)$ is the kth region of image I that exhibits meaningful details to human beings. Equations (3), (4), and (5) together can work as guidelines for designing image enhancement algorithms. To follow these guidelines in designing our algorithm, we first start from Equations (4) and (5). In order to extend reflectance $R(x,y)$, we want our algorithm to make an image signal utilize its available dynamic range as much as possible to ensure Equations (4) and (5) are satisfied. Hence, we include a normalization stage in our design. This normalization process can guarantee the increase of dynamic range usage. Furthermore, the characteristic of Equation (3) can be satisfied by modifying the normalization stage into a local normalization one. For example, a signal can be considered as S piecewise connected dense line segments, which can be s as follows:

$$\begin{aligned} &I(x,y) = \{I_k(x,y),\ k = 1,2,...,S\}, \\ &I_k(x,y) = I_{\min,k} + Rand_k(\phi_k) \end{aligned} . \tag{6}$$

For each line segment I_k , $I_{\min,k}$ is the local minimum of I_k , $Rand_k(\phi_k)$ is a random variable that varies from *0* to ϕ_k . Hence after applying local normalization $LN(.)$, $I_{\min,k}$ will be removed and the normalized signal is shown as follows:

$$\begin{aligned} LN(I(x,y)) &= LN(\{I_k(x,y),\ k = 1,2,...,S\}) \\ &= LN(\{Rand_k(\phi_k),\ k = 1,2,...,S\}) \end{aligned} . \tag{7}$$

Therefore, the mean of $LN(I(x,y))$ is shown as follows:

$$Mean(LN(I(x,y))) = Sum(LN(\{Rand_k(\phi_k),\ k = 1,2,...,S\}))/S, \tag{8}$$

where $Sum(.)$ represents the summation process. For S is large, based on the central limit theorem, $Sum(LN(\{Rand_k(\phi_k),\ k = 1,2,...,S\}))$ will become normally distributed. Therefore, the mean of the normalized signal $LN(I(x,y))$ will be close to the center of dynamic range. On the other hand, $L'(x,y)$ can be considered as a low-pass filtered version of the normalized signal $LN(I(x,y))$. Hence, it is certain that $L'(x,y)$ will satisfy the characteristic of Equation (3). Note that although $LN(I(x,y))$ can satisfy the constraints set in Equations (3), (4), and (5), the original $L(x,y)$ is seriously removed. In addition, eliminating the entire $L(x,y)$ will completely remove the tone of the original. This outcome is definitely not what we want in our results. Therefore, we added an addition attenuation ratio T to multiply the lower bound intensity $I_{\min,k}$ in the local normalization process. This is to ensure that $L(x,y)$ can be preserved, but it will be suppressed to $1-T$. The effect of adding this addition attenuation ratio T is that the value of $Mean(LN(I(x,y)))$ will be somewhat shifted away from the center of the dynamic range. However, such shift will not move $Mean(LN(I(x,y)))$ to either of the two extreme values of the dynamic range. Consequently, the characteristic of Equation (3) still holds.

Other than checking Equations (3), (4), and (5) to see whether an image is well reproduced, we have to consider other possible side effects. One notable side effect often discussed in tone reproduction techniques is halo effect. Halo effect usually occurs when low-pass filtering is introduced in the enhancement process. However, halo effect is not exhibited among our results because our local normalization scheme does not use low-pass filtering to enhance an image.

Furthermore, in order to ensure the improvement of image quality, it is necessary to include an assessment mechanism in the design. With this mechanism, we can retain the regions of the original that are better than those of the locally normalized one. To achieve this, we add an adaptive contrast assessment portion in our tone reproduction algorithm. The adaptive contrast assessment part will compute an exponent factor that biases between the original image and the locally normalized one.

2.2. The Proposed Tone Reproduction Scheme

Following the concepts in the previous section, we propose a tone reproduction scheme as follows:

$$I'(x,y) = I(x,y) \times E(x,y)^{C(x,y)}, \tag{9}$$

where $I'(x,y)$ is an enhanced image signal. $I(x,y) \times E(x,y)$ is the locally normalized version of $I(x,y)$, hence $E(x,y)$ is the local normalization ratio kernel. In our design, for image $I(x,y)$, $E(x,y)$ is designed as follows:

$$E(x,y) = \frac{I(x,y) - I_{\min} \times T}{I_{\max} - I_{\min} \times T + \varepsilon} \times \frac{G}{I(x,y)}. \tag{10}$$

As described before, the normalization equation is not an ordinary one because we added an attenuation ratio T (ranging from of 0 to 1) in our tone reproduction scheme. $I_{\max}$ and $I_{\min}$ are the local maxima and minima of an image signal respectively. G is the full size of the dynamic range, and ε is an offset to avoid the divide-by-zero situation. As to $C(x, y)$, it's the adaptive contrast assessment factor:

$$C(x, y) = Gaussian(\arg\{Lap(I), Lap(I')\}), \quad (11)$$

where $\arg\{a, b\} = 0$ when $a > b$, $\arg\{a, b\} = 1$ when $a \le b$; $Lap(.)$ is the Laplacian operator. With $C(x, y)$, our proposed mechanism can enhance the low contrast regions of an image, while preserving the details of high contrast regions. For local normalization, we partition an image into piecewise connected segments such as in Equation (6) to evaluate the effectiveness of our algorithm. Thus, $E(x, y) = \{E_k(x, y),\ k = 1, 2, ..., S\}$. From Equations (1) and (10):

$$E_k(x, y) = \left(\frac{G}{G_{L,k}}\right) \times \frac{\left[1 - \frac{I_{\min,k} \times T}{I_k(x, y)}\right]}{(1 + \varepsilon / G_{L,k})},$$

where $G_{L,k} = I_{\max,k} - I_{\min,k} \times T$ .Let $\eta_k(x, y)$ be the intensity difference between $I_k(x, y)$ and $I_{\min,k}$, then we have

$$I_k(x, y) = I_{\min,k} + \eta_k(x, y),$$

$$I'_k(x, y) = I_k(x, y) \times E_k(x, y)^{C_k(x,y)}$$

$$= I_k(x, y) \times \{\frac{G}{G_{L,k}} \times \frac{1 - T \times \left(1 + \eta_k(x, y) / I_{\min,k}\right)^{-1}}{(1 + \varepsilon / G_{L,k})}\}^{C_k(x,y)}.$$

Now considering the following three ordinary target region cases for enhancing: 1) Region with low local contrast and high local luminance. 2) Region with low local contrast and medium local luminance. 3) Region with low local contrast and low local luminance. Because all three cases are in the condition of low local contrast, $G_{L,k}$ will be smaller in comparison with G . According to the statistics of Weber's Law [8], $G_{L,k}$ that falls below about $G/13$ will be hard to be perceptible. Now considering the first case, the local luminance is high so that $I_{\min,k}$ can be larger in comparison with $\eta_k(x, y)$. Therefore,

$$V'_{Case1}(x, y) = \frac{I'_{Case1}(x, y)}{I_{Case1}(x, y)} = [\frac{G \times (1 - T)}{G_{L,k} + \varepsilon}]^{C_k(x,y)}.$$

For the second case, $I_{\min,k}$ is comparable with $\eta_k(x, y)$,

$$V'_{Case2}(x, y) = [\frac{G \times \left(1 - T / \left(1 + \eta_k(x, y) / I_{\min,k}\right)\right)}{G_{L,k} + \varepsilon}]^{C_k(x,y)}.$$

For the third case, $\eta_k(x, y)$ is more dominant than $I_{\min,k}$,

$$V'_{Case3}(x, y) = \left[G \times \left(1 - T \times I_{\min,k} / \eta_k(x, y)\right) / (G_{L,k} + \varepsilon)\right]^{C_k(x,y)}.$$

By careful selection of T and ε , our proposed algorithm can guarantee to have V greater than 1 in all cases. In the scheme, ε is a constant offset we define which is no more than *G/35*. For example, if we set $T = 0.8$ and ε =7. When $V > 1$, one can interpret the above three cases as follows: for the first case, since the local luminance is high, $L_1(x,y)$ tends to decrease. However, in order to follow $V'_{case1} > 1$, $R_1(x,y)$ would tend to be greater than 1. As to the second case, the resulting V'_{case2} will be quite large with our settings. Even with a medium local luminance, $V'_{case2}/L_2(x,y) = R_2(x,y)$ is still greatly larger than 1. The third case is similar. With a dominating $\eta_k(x, y)$ would make V'_{case3} a large value comparing to 1, it is reasonable to have $L_3(x,y)>1$ *and* $R_3(x,y)>1$ to ensure a large V'_{case3}. To sum up, it is obvious that the proposed image enhancing process would follow the characteristics of Equations (3), (4), and (5). That is, the details and visibility of a processed image are both restored.

3. EXPERIMENT RESULTS

To test the effectiveness of our method, we conducted experiments on a set of images acquired under various shading/lighting conditions. Figures 1 (a), (b), (c), and (d) show four test images in which the complexity ranges from simple to complex. Figures 1 (e), (f), (g), and (h) show the tone reproduced results. It is apparent that the details and chromatic information of the tone reproduced images shown in Figures 1 (f) and (g) were both greatly enhanced. As to the tone reproduced images shown in Figures 1 (e) and (h), the visual details of their heavily shaded areas were recovered. One thing to be noted is that the halo effect was greatly reduced due to the avoidance of using low-pass filter in the local normalization process. Low-pass filtering is a major cause of the halo effect in most image enhancement algorithms.

To explain the effectiveness of our method, we introduce the concept of Weber's Law. An approximation of the gamma corrected Weber's Law function is addressed in [4]. With this curve, we can estimate the percentage of a photo where image details are visible to a human observer and accordingly obtain the contrast distribution of an image. Higher perceptible percentage corresponds to more image-detail-perceivable regions of an image. For example, in this set of experiments, our method restored the highest amount of image detail regions thus received the highest percentage (90.58%). Under above evaluation metric, Fattal *et al.*'s method generally yielded the second best results among all our test sequences, which contained more than 150 images. Hence in Figure 2, we show a more detailed comparison between Fattal *et al.*'s results and ours. Since their tone reproduction algorithms require a good parameter to obtain visually pleasing results, we fine tuned their parameters to obtain a near optimal set of results in this comparison. However, our method does not need to do manual parameter selection.

4. CONCLUSION

We have proposed an automatic tone reproduction scheme based on the image model. With careful design, our algorithm is proved to produce results that fit the goals of tone reproduction. From our test data, which includes a range of lighting conditions and shading effects, our tone reproduction algorithm can achieve excellent detail reproduction without any user parameter adjustments.

5. REFERENCES

[1] F. Durand, and J. Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images," *Proc. SIGGRAPH 2002*, pp. 257–266, 2002.

[2] R. Fattal, D. Lischinski, M. Werman, "Gradient domain high dynamic range compression," *Proc. SIGGRAPH 2002*, pp. 249-256, 2002.

[3] G. Krawczyk, K. Myszkowski, and H.-P. Seidel, "Lightness perception in tone reproduction for high dynamic range images," *Proc. of EUROGRAPHICS 2005,* vol. 24, pp. 635-645, 2005.

[4] 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, pp. 467-476, 1995.

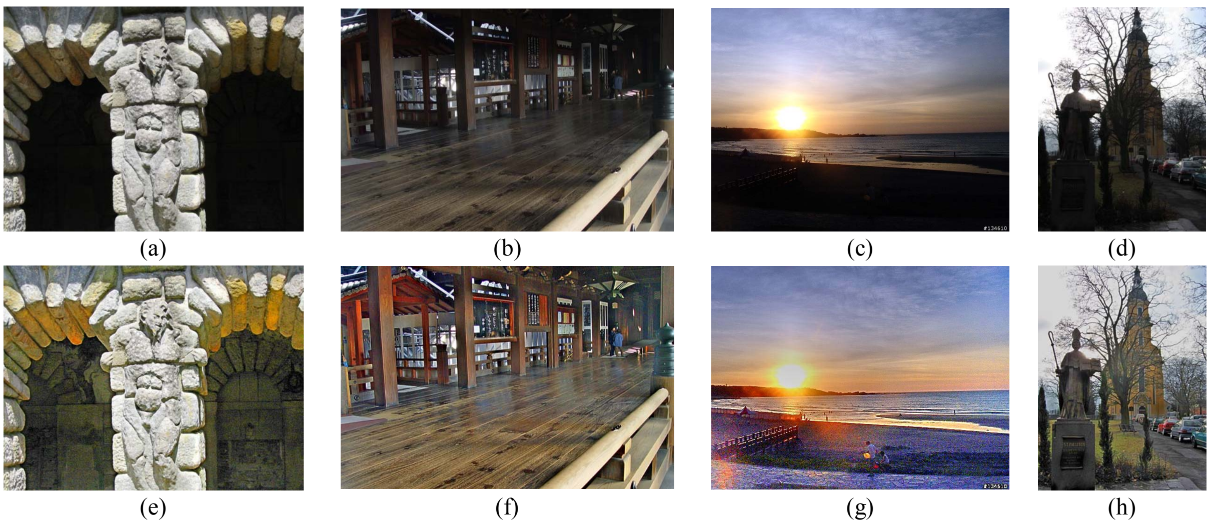

Figure 1. Test images and results with lighting/shading conditions: (a)-(d), original images; (e)-(h), enhanced images.

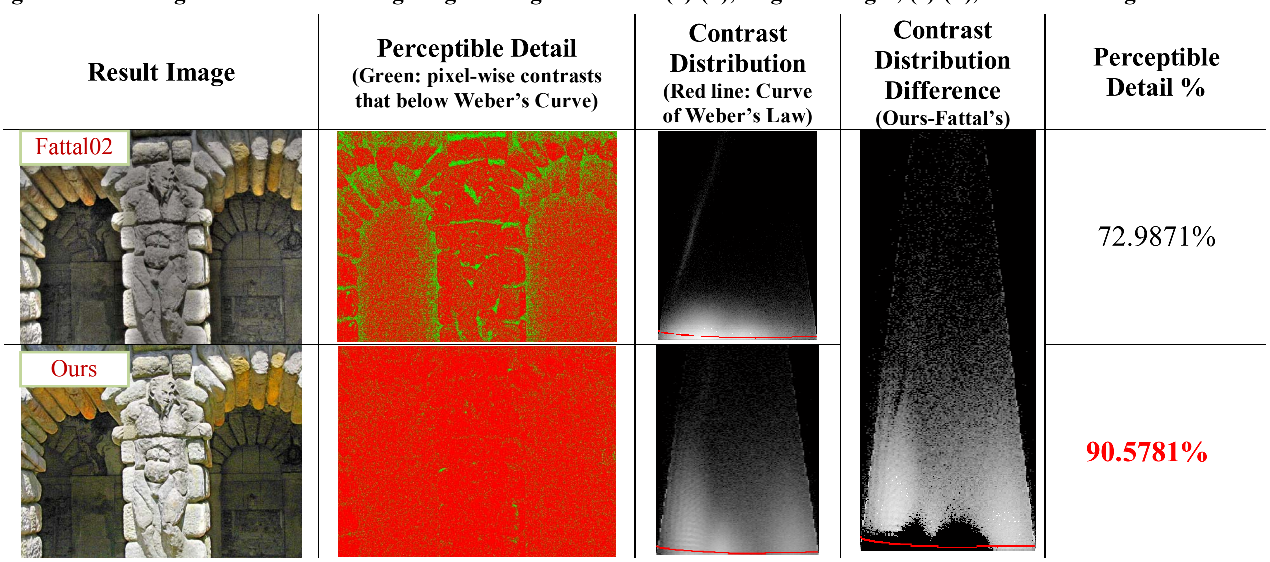

Figure 2. Perceptible percentage comparison over two tone reproduction algorithms. Noted the contrast distribution of our method spreads significantly wider.