Image-dependent Quality and Preference Control

Hiroaki Kotera
Department of Information and Image Sciences, Chiba University, Japan

Abstract

Most of image processing systems have been designed independent of image contents. Now, evolving computation power is making the "***image-dependent***" elegant algorithms possible to get the better color image renditions.

This paper introduces our recent approaches to quality and pleasant imaging based on *image-dependence*. The key technologies are represented by ***segmentation*** and ***multi-scale***. The source colors are transformed to the preferred reference colors in the ***segmented*** sub-spaces better than in the entire space. The spatial filtering works better if the ***multi-scale*** kernel is selected than a fixed single kernel. The typical applications of these key technologies are presented.

Introduction

The concept of **"*device-independent*"** has unified the stage of color interchange between the different devices, media, and systems. But, so far, a variety of image processing systems have been designed independent of image contents. Now, evolving computation power is making it possible to apply **"*image-dependent*"**[1] elegant algorithms to get the better color renditions.

This paper introduces our recent approaches to high quality and pleasant imaging based on the concept of image-dependence. To process the image dependent of its content, we apply the common key technologies represented by segmentation and multi-scale. For example, the source colors are matched to the destination colors or transformed to the preferred reference colors in the segmented sub-spaces better than in the entire space. As well, spatial filtering works better if the multi-scale kernel is selected than a fixed single kernel. Fig.1 shows a conceptual view of the system.

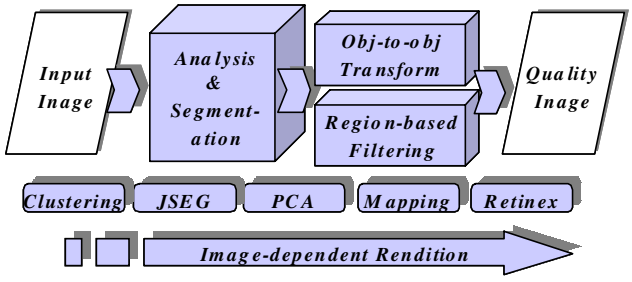


Figure 1. Concept of Image-dependent Rendition

The following typical applications are presented in this paper.

- Image-dependent preferred color transform:
- Image-dependent sharpness control
- Image-dependent 3D gamut mapping
- Image-dependent dynamic range control

Image-dependent Preferred Color Transform

A color image is composed of different objects distinguished from others that are distributed in clusters in 3D color space. The intent of preferred color reproduction is to colored object. Fig. 2 shows an object-to-object color transform system.[2] In the preferred color reproduction approach, the key color areas such as flesh tint, blue sky, or green grass are extracted by image segmentation and automatically matched to the desirable reference colors. In this model, PCA matching applies K sets of transform matrices dependent on the colored objects.

Extraction of Key Color Areas by Segmentation

The key color areas are extracted through the following two steps segmentation process.

[Step1: Classification of Color Cluster]

First, the clustered color objects are separated by the statistical classifier. For example, Bayesian decision rule classifies each pixel color vector $\boldsymbol{X}$ into the class $k=c$ group with the maximum likelihood to minimize the following discrimination function.

$$ {}_k d(Bayes) = -\log\{p(k)\} + \log\left(\left|{}_k\sum_X\right|\right) + \frac{1}{2}\left(X-{}_k\mu\right)^t {}_k\sum_X^{-1}\left(X-{}_k\mu\right) $$

$$ p(k)\text{: occurrence probability of class } k \qquad (1) $$

[Step2: Symbolic Segmentation by JSEG]

The ***step1*** gives rough segmentation of key objects but clustering errors happen in mixed color areas with texture. The ***step2*** JSEG algorithm[2] makes the ambiguous boundaries clear by merging the closest classes based on symbolic distance measure. JSEG converts the labeled image by class

number into J-image. J-image is composed of J values characterized by the relative symbolic distance from the overall center to each class center as follows.

$$J = (S_T - S_W)/S_W \tag{2}$$

$$S_T = \sum_{z \in Z} \|z - m\|^2, \quad S_W = \sum_{k=1}^{C} \sum_{z \in Z_k} \|z - m_k\|^2 \tag{3}$$

Where, z , m and m_k denote the geometric coordinates of symbolized class z, overall mean center, and each class center, respectively.

[Step3: Object-to-Object PC Matching]

Finally, a segmented key color area is transformed into the preferred color for their principal component (PC) to be matched to the trained reference color. The PC matching is done for typical memory colors such as skin, blue sky, or green grass, by operating the following matrix[3] ${}_k\boldsymbol{M}_C$

$${}_k\boldsymbol{M}_C = \left(\boldsymbol{A}_{REF}^{-1}\right)\left({}_k\boldsymbol{S}\right)\left(\boldsymbol{A}_{ORG}\right) \tag{4}$$

Where ${}_k\boldsymbol{A}_{ORG}$ and ${}_k\boldsymbol{A}_{REF}$ denote the eigen matrices for original segment in class k and corresponding reference color and ${}_k\boldsymbol{S}$ denotes the scaling matrix.

$${}_k\boldsymbol{S} = \begin{bmatrix} \sqrt{{}_k\lambda_{1REF} / {}_k\lambda_{1ORG}} & 0 & 0 \\ 0 & \sqrt{{}_k\lambda_{2REF} / {}_k\lambda_{2ORG}} & 0 \\ 0 & 0 & \sqrt{{}_k\lambda_{3REF} / {}_k\lambda_{3ORG}} \end{bmatrix} \tag{5}$$

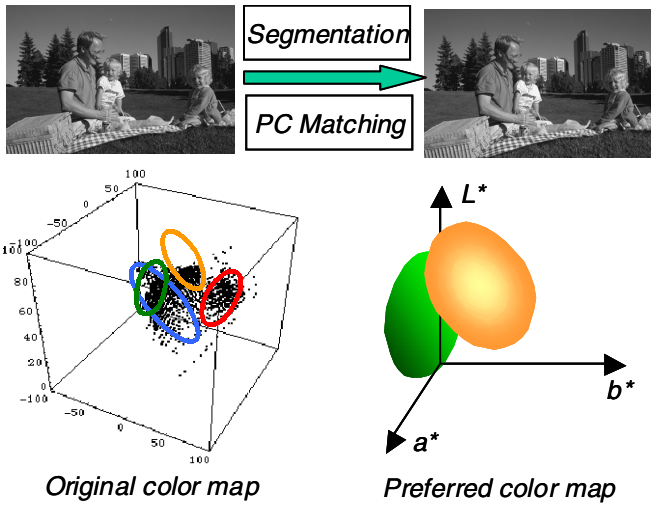


Figure 2. Concept of Object-to-object PC Color Matching

Transformation to Preferred Reference Colors

Fig. 3 (a) shows a preferred color transform result for the test image "Palace garden" including dull blue sky, dark green grass, drab red and yellow flowers. These key color areas are almost well segmented through ***step1*** and ***step2***, then transformed to the clear blue sky, bright green grass, and vivid red flower trained beforehand. Fig. 3(b) is another example for portrait, where blue sky and skin color areas are transformed to the trained reference colors. In this sample, dull blue sky is converted into clear sky very well. Although the facial skin color is also changed into fresh tint, but looks unnatural just like as doll's face. The reason may be in the unsuitable training for preferred skin color. The trained reference color samples must be distributed in narrow deviation of principal components axes like as doll's face.

Any interactive selection system using LUT will be necessary for the better setting of reference colors.

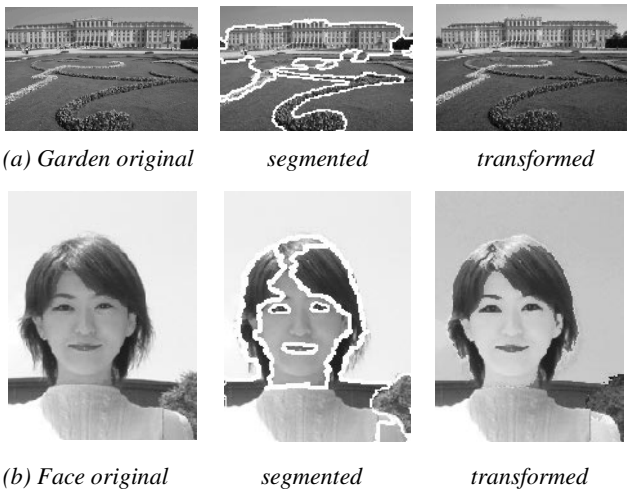


Figure 3. Color transform to trained preferred references

Image-dependent Sharpening

The color image has the different edge profiles depending on the contents of the scene. In the conventional edge enhancement method, a single sharpness filter such as digital Laplacian or un-sharp mask operator is applied to the entire image independent of the objects. The non-adaptive single sharpness filter has the drawbacks such as enhancement of random noise in flat area or insensitivity to the dull edges. The conventional sharpness operator such as digital Laplacian can't create the natural edge sharpness, because it has local edge responses different from receptive field in human vision. In our image- dependent sharpening system, the multiple sharpening filters are applied to work adaptive to the different edge slopes.

Edge Detection Operator

A variety of simple cell receptive field models for human vision have been considered such as Gaussian Derivative (GD), Gabor, DOG, DOOG, or DODOG. Young[4] and others reported GD is the best to minimize the joint space-spatial frequency uncertainty $\Delta x \bullet \Delta \omega$. The proposed model also used GD-based operator. The basic Gaussian distribution function in two dimensions is defined by

$$G(r) = \frac{1}{2\pi\sigma^2} exp\left(-\frac{r^2}{2\sigma^2}\right); \quad r^2 = x^2 + y^2 \tag{6}$$

Its second derivative is given by

$$\nabla^2 G = \frac{1}{\pi\sigma^4}\left(\frac{r^2}{2\sigma^2} - 1\right) exp\left(-\frac{r^2}{2\sigma^2}\right) \tag{7}$$

The edge signals are extracted from an image $f(x, y)$ by the two-dimensional convolution operation as

$$\delta_m(x,y) = \nabla^2 G_m <*> f(x,y) \tag{8}$$

Where, symbol <*> denotes the convolution operation and m means filter kernel size. The edge sharpness is measured by operating the pre-scan filter $\nabla^2 G_s$ with appropriate standard deviation σ_s.

Edge-adaptive Multi-Scale Sharpening

Figure 4 illustrates the sharpening process in the proposed system.[5] First, the *RGB* image is transformed into luminance-chrominance image such as YC_rC_b or *YIQ*. The edge enhancement filter is applied only to the luminance *Y* image to keep the gray balance on the edges. After pre-scanning the *Y* image by sharp GD filter $\nabla^2 G_s$, the edge profiles are classified into hard, medium, soft, and flat types by analyzing the histogram of $\delta_s(x, y)$. Thus the zone mask $M(x, y)$ is generated to discriminate these edge types. Next, the multiple GD operators $\nabla^2 G$ with different standard deviations, σ_1, σ_2, and σ_3 are applied to *Y* image and the corresponding edge signals, that is, hard edge δ_1(x, y), medium edge δ_2(x, y), or soft edge δ_3(x, y) are detected in response to the different edge slopes in the image.

These edge signals are selectively activated by looking up the zone mask plane $M(x, y)$ to discriminate the edge types. Thirdly, the *Y* image is sharpened by subtracting this edge adaptive GD signal δ(x, y) from image $f(x, y)$ as follows.

$$f'(x, y) = f(x, y) - \delta(x, y) \tag{9}$$

Except these three types of filtering, the flat/gentle slope areas are left unprocessed to keep original or intentionally ***smoothed*** by operating the normal Gaussian filter. This brought dramatic improvement in the suppression of background random noises.

Finally, the original *Y* image is replaced by sharpened luminance image *Y'* and converted into *R'G'B'* primary color image from $Y'C_rC_b$ image by inverse transform.

Figure 5 shows a result of the proposed multi-scale sharpening method. The blurred input image is sharpened by switching the multiple GD filters quickly looking up the zone mask plane. The zone mask plane is colored by red=hard, green= medium, blue=soft edges, and black=flat area. As compared with conventional single filtering, the facial areas recognized flat are very smoothly rendered with the reduction of background noises and other edge areas are naturally sharpened adaptive to the edge profiles.

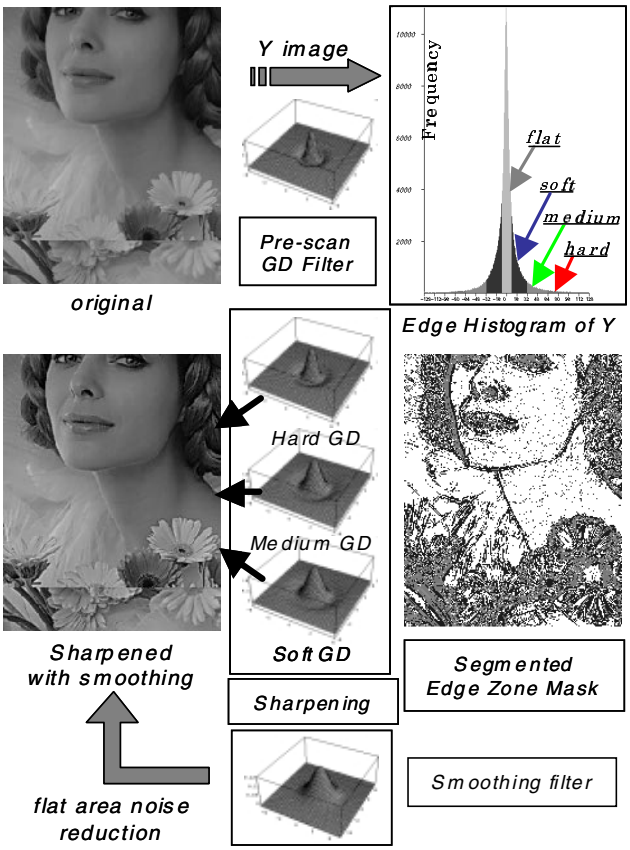


Figure 4. Edge-dependent multi-scale sharpening

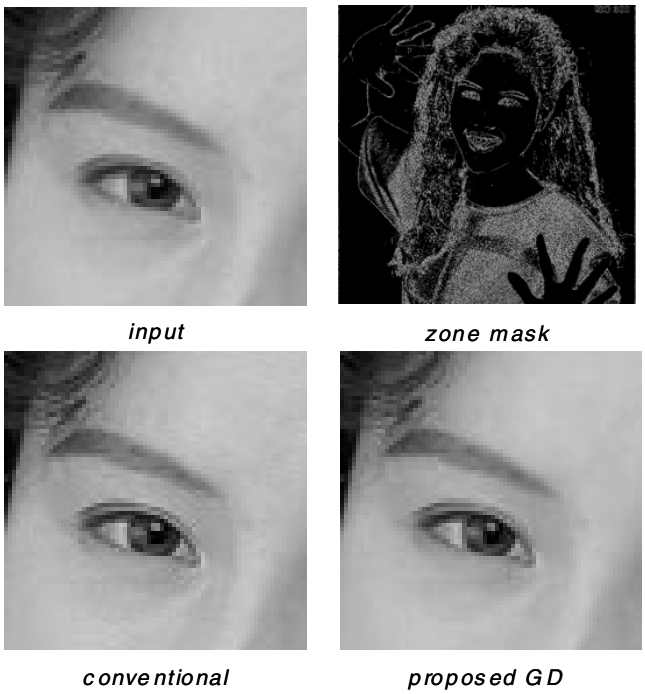


Figure 5. Adaptive edge sharpening result

The kernel size of GD filters would be $\pm 3\sigma_s \sim \pm 4\sigma_s$, and the standard deviation σ_s must be decided corresponding to the edge profiles of segmented zone mask. The automatic classification of edge type based on edge histogram and the decision of optimum σ_s are under development.

Image-dependent Gamut Mapping

So far most of GMA[6] were designed to work in segmented 2D L-C planes based on Device-to- Device (D-D) concept, not Image-to- Device (I-D). They don't reflect the image gamut exactly. In the current 2D ***D-D*** GMA, the source color ***s*** is mapped to the destination ***t*** in relation to the monitor gamut boundary ***m*** vs. printer's boundary ***o*** toward a focal point ***p***. However, the saturation and gradation losses will happen after the mapping, because the image color distributions don't always fill the entire monitor gamut. While, the ***I-D*** GMA[7] uses the image gamut boundary ***i*** instead of ***m***, then it can suppress such losses in minimum as shown in Fig. 6. Clearly, I-D GMA is better than D-D and 3D than 2D. We developed the seamless 3D I-D GMA[8,9] based on image-dependent. In the I-D GMA, the image gamut boundary plays an important role. Figure 7 (a) shows a simple way to extract the 3D image gamut boundary by dividing the entire color space into small segment by constant radial angle ($\Delta\theta$, $\Delta\varphi$) and extracting the maximum radial vector in each segment as follows.

$$r_i = \left[\left(L_i^* - L_0^*\right)^2 + \left(a_i^* - a_0^*\right)^2 + \left(b_i^* - b_0^*\right)^2 \right]^{1/2}; \quad 1 \le i \le n$$

$$\theta_j = tan^{-1}\left[\frac{b_j^* - b_0^*}{a_j^* - a_0^*} \right]; \quad 0 \le \theta_j \le 2\pi \qquad (10)$$

$$\varphi_k = (\pi/2) + tan^{-1}\left(\frac{L_k^* - L_0^*}{\left\{ \left(a_k^* - a_0^*\right)^2 + \left(b_k^* - b_0^*\right)^2 \right\}^{1/2}} \right); 0 \le \varphi_k \le \pi$$

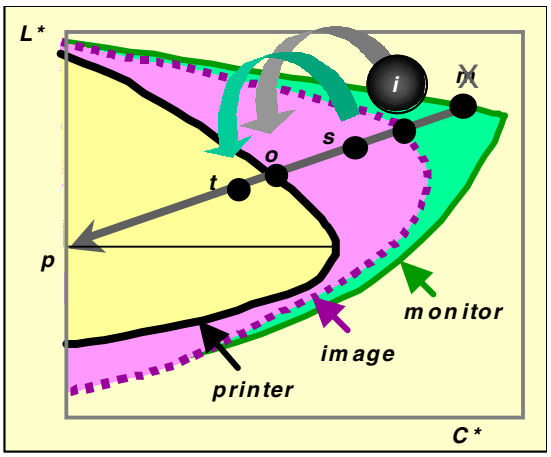
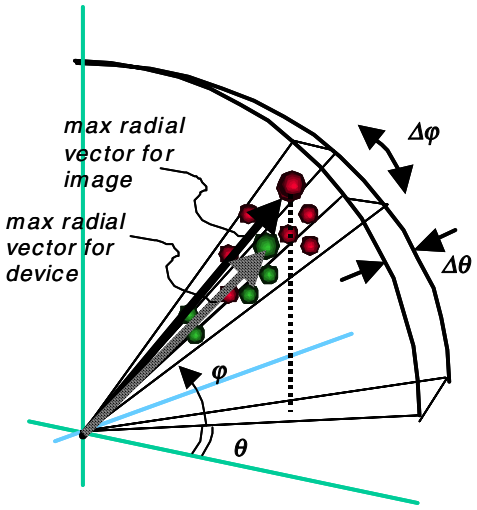
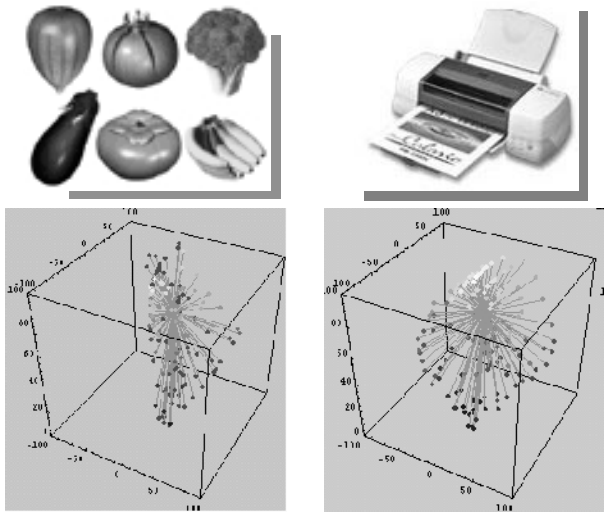


Figure 6. Basic concept of I-D GMA

The 3D I-D mapping is performed in efficient by comparing the extracted radial vectors between image and device. Figure 7 (b) shows an example of 3D gamut shapes by radial vectors for a CG image and Epson inkjet printer. The magnitudes of the maximum radial vectors are rearranged in θ–φ coordinate and converted to 2D gray scale image called ***r-image***[10] It can be compactly compressed by conventional picture coding system and attached to original image used for I-D GMA at user side.



(a) I-D gamut extraction in Δθ-Δφ segment



(b) Image vs Device gamut by maximum radial vectors

Figure 7. 3D gamut boundary description by radial vector

Figure 8 shows the ***r-image*** for test image "wool" compressed to only 200 bytes by wavelet coding and its reconstructed gamut shell shape. When the 3D gamut shell shape has the smoothed surface, the ***r-image*** will be highly

correlated in row and column. In such case, it can be compactly compressed by applying the SVD (Singular Value Decomposition) or conventional transform coding methods. Indeed, the complicated gamut shell shapes could be well described less than 500 bytes.

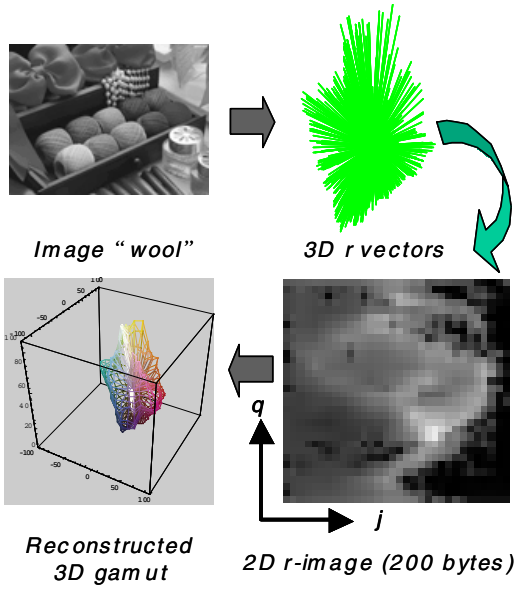


Figure 8. Compact Image gamut description by ***r-image***

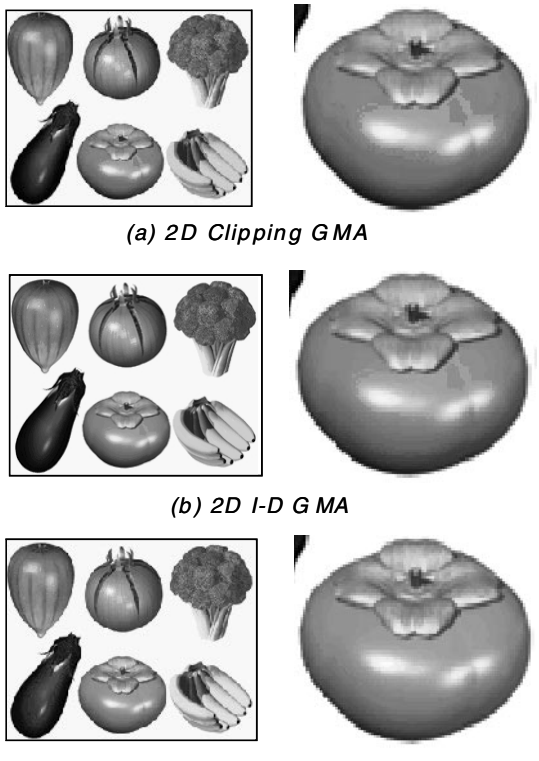


Figure 9. Gamut compression by ***3D I-D GMA***

Figure 9(c) is a gamut compression result for CG image by our 3D I-D GMA compared with conventional clipping GMA in (a) and 2D I-D GMA in (b). As shown clearly in the right side "persimmon", unacceptable heavy artifacts due to the gradation loss in (a) and (b) are naturally reproduced in (c).

Automatic Dynamic Range Control

The last topic is addressed to image-dependent dynamic range control. Human vision perceives very wide range of lightness by adaptation mechanism. But the picture taken by electronic camera under the heavy change in highlight and shadow often looks different and unnatural from our visual perception. Land and McCann[11] introduced Retinex model in human visual perception based on the relative lightness ratio product algorithm. It works to correct the scene lightness not by "pixel-to-pixel" but by "spatial-to-pixel" process. Jobson,[12] Rahman,[13] Funt,[14] and others are developing the single-scale retinex (SSR) into multi-scale retinex (MSR) based on the center/ surround model.

To apply these models for automatic dynamic range control, the following processes were introduced.[15]

- To avoid the instability in the logarithmic function of conventional Retinex, the linear lightness ratio is introduced to the pixel-to-surround computation.
- The surround lightness field is extracted from region adaptive multiple Gaussian spatial filters.

Here the retinex output $R_i(x, y)$ for channel $i=R, G, B$ input $I_i(x, y)$ is given by

$$R_i(x,y) = A\frac{I_i(x,y)}{SEL\langle[G_1 * Y(x,y)][G_2 * Y(x,y)]\cdots[G_M * Y(x,y)]\rangle}$$

$$Y(x,y) = 0.299R(x,y) + 0.587G(x,y) + 0.114B(x,y)$$

$$G_m = Kexp\left\{-\sigma_m{}^2/\left(x^2+y^2\right)\right\},\ \iint G_m dxdy = 1 \qquad (11)$$

G_m denotes Gaussian averaging filter with standard deviation σ_m for pixel-to-surround field and the symbol * denotes convolution operation. Each *RGB* channel output is calculated by the pixel-to-surround ratio, where the surround is selected from the multi-scale convolutions with common luminance $Y(x, y)$ to keep color balance.

$SEL\langle\bullet\rangle$ means to select the appropriate kernel size for multi-scale Gaussian convolution adaptive to the spatial lightness distribution of the image in attention.

Figure 11 illustrates the proposed adaptive MSR model.

The linear ratio process works robust in the dark signal levels with random noises and makes it easy to normalize the dynamic range. The proposed model synthesizes surround lightness field adaptive to the complexity of local areas by applying the multi-resolution Gaussian filters.

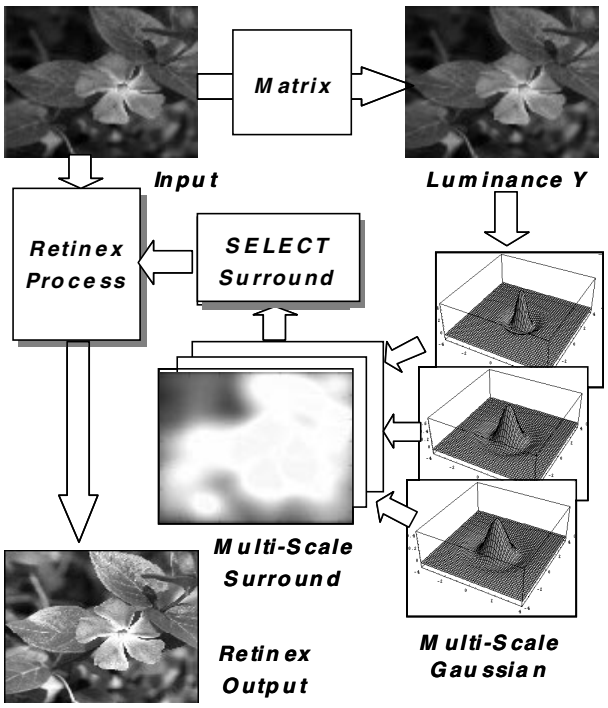


Figure 10. Multi-scale Linear Retinex Model

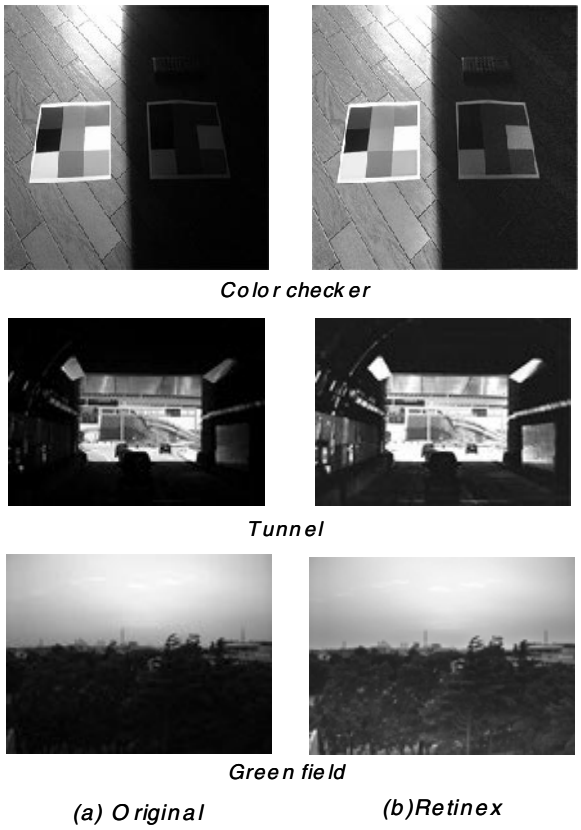


Figure 11. Appearance improvement by MSR retinex

The proposed MSR worked well to compress the dynamic range without loosing the details and to enhance the visibility in shadow areas keeping the color balance and saturation as shown in Fig. 11. The visibilities of color checker placed in sunny and shade, or dark tunnel scene are dramatically improved. As well, the heavily de-saturated green field is vividly recovered.

Conclusion and Discussion

The image-dependent segmentation and multi-scale filtering techniques were introduced to control the quality and preference toward comfortable color image rendition. The key concept lies in that individual image should be processed dependent of its content and to communicate its intent. Here, four typical applications are individually processed according to each objective. The algorithms worked well in some part, but failed in other part.

The first trial, ***image-dependent preferred color transform*** tells us the difficulties in accurate segmentation of colored objects in a natural scene. The small segmentation errors on the region boundary sometimes result in the fatal defect.

The second application, ***image-dependent sharpening*** was most successful multi-scale processing. It also uses segmentation to classify the edge types, but the miss selection of edge type is very much out of question.

The third, ***image-dependent gamut mapping*** will bring the more flexible cross-media color management at user side.

The last, ***automatic dynamic range control*** is the most familiar process to be applied to digital photography. The advances in ***retinex model*** is expected to reproduce the natural scene just as human vision do so.

References

1. H. Kotera et al, *Proc. PICS,* 288(2001)
2. H. Kotera et al, *Proc. IS&T's NIP14*, 310(1998)
3. M. Suzuki and H. Kotera, *Proc. IS&T's NIP17*, 450(2001)
4. H. Kotera et al, *J. Elect. Imag.,*10(4), 977(2001)
5. R. A. Young, *Proc. SPIE*, 1453, 92(1991)
6. K. Shimo and H. Kotera, Proc. IS&T's NIP16, 814 (2000)
7. J. Morovic and M. R. Luo, *J. Imag. Sci. & Tech.,* 45, 3, 283(2001)
8. H. S. Chen et al, *J. IS&T*, 45, 2, 141 (2000)
9. H. S. Chen and H. Kotera, *Proc. IS&T's NIP16*, 783 (2000)
10. H. Kotera and R. Saito, *Proc.9th CIC*, 56 (2001)
11. E. H. Land and J. J. McCann, *J. O. S. A.,*61, 1, 1 (1971)
12. D. J. Jobson and Z. Rahman, *IEEE Trans.,* 6, 3, 451 (1977)
13. Z. Rahman et al, *Proc. SPIE,* 2825 (1996)
14. B. Funt et al, *Proc.8th CIC*, 112 (2001)
15. M. Kobayashi and H. Kotera, *Proc. Color Forum Japan,* 151 (2001)

Biography

Hiroaki Kotera received his B.S degree from Nagoya Institute of Technology and Doctorate from University of Tokyo. He joined Matsushita Electric Industrial Co in 1963. Since 1973, he has been working in digital color image processing at Matsushita Research Institute Tokyo, Inc. In 1996, he moved to Chiba University. He is a professor at Dept. of Information and Image Sciences. He received Johann Gutenberg prize from SID in 1995 and journal awards from IS&T in 1993, from IIEEJ in 1990 and 2000.