Intelligent Image Processing

Hiroaki Kotera, Professor, Chiba University, Chiba, Japan

Abstract

Seeing 600th anniversary of Gutenberg's birth in 2000 A.D., we should look back the historical significance of letterpress technology and take a step forward into color imaging new age. Now digital imaging technology plays a leading role in visual communication, but meets severe assessment to satisfy human vision. Software on "What's human vision seeing?" is essential to capture, store, transmit, and reproduce a truly realistic image just as human vision seeing. Not only advances in high precision and high definition digital media, but also Intelligent Image Processing technologies will be helpful for more aesthetic and pleasant imaging. We are approaching towards this direction from a stance of engineering to use the scientific results in human vision research. This paper introduces just a little "intelligent" processing to "image sharpening", "local contrast enhancement", and "color transform" by region-based, spatially-variant, and scene-referred approaches.

Introduction

During past 10 years, ***CMS*** (*Color Management System*)[31] has developed to communicate the device-independent colors across multi-media and now is introducing some human visual aspects into standardization. Image processing technologies would be also requested to take the ***vision-based*** and ***intelligent*** approaches to contribute to the next age color imaging. Figure 1 illustrates a trend on vision-based ***CMS*** and a mission for image processing.

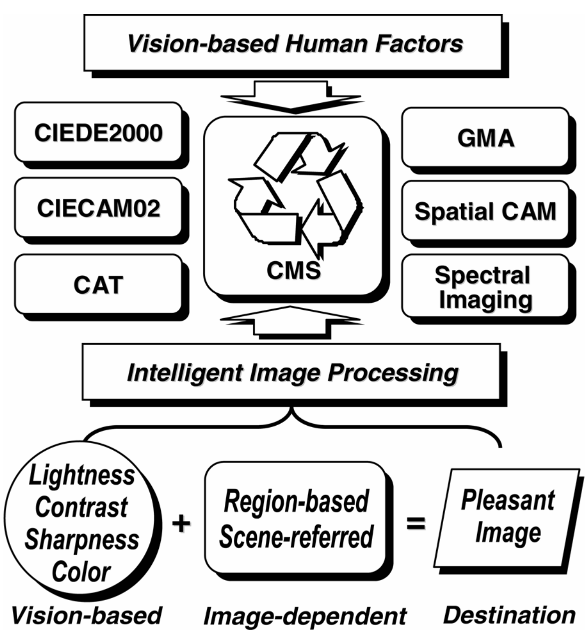

Figure 1. Vision-based CMS and Mission of Image Processing

INTELLIGENT Center/Surround Processing

The ***Center/Surround*** (***C/S***)[10] response is the first step in human vision, which represents many aspects of visual perception. The ***C/S*** model has hinted to *sharpness, lightness, or contrast* improvements in image processing.

Image Sharpening

The center/surround cell is well known to respond to the edge and not to diffuse light. The edge response in retina is modeled by a second *Gaussian Derivative* (***GD***) as illustrated in Fig. 1.

$$\nabla^2 G(x,y,\sigma) = \partial^2 G(r)/\partial x^2 + \partial^2 G(r)/\partial y^2;\ r^2 = x^2 + y^2$$
$$= \frac{1}{\pi\sigma^4}\left(\frac{r^2}{2\sigma^2}-1\right)\left(\frac{r^2}{2\sigma^2}-1\right)exp\left(-\frac{r^2}{2\sigma^2}\right) \tag{1}$$

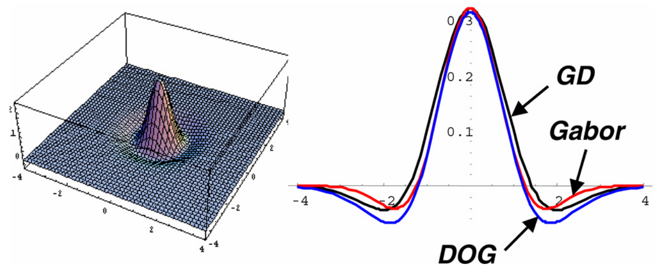

Figure 2. Edge response of center/surround cell in retina

The edge signal is extracted from a blurred image $g(x, y)$ by

$$z(x,y) = -\nabla^2 G(x,y,\sigma) \otimes g(x,y)\ ;\ \otimes : convolution \tag{2}$$

As well-known as *USM* (*Unsharp Mask*) or *Laplacian* operator, image sharpening process is generalized by

$$f(x,y) = g(x,y) + \lambda z(x,y) \tag{3}$$

Since a simple linear *USM* controls the edge strength by only a constant λ, it enhances unwanted background noises together with signals. To optionally sharpen the edge with suppressing noise, many *intelligent* methods have been developed. For instance, ***Adaptive USM***[4] operates a *spatially-variant* coefficients $\lambda_x(x, y)$ and $\lambda_y(x, y)$ separately to a horizontal and a vertical component $zx(x, y)$ and $z_y(x, y)$ as follows.

$$f(x,y) = g(x,y) + \boldsymbol{\Lambda}^t \boldsymbol{Z}(x,y)$$
$$\boldsymbol{Z}(x,y) = \begin{bmatrix} 2g(x,y) - g(x-1,y) - g(x+1,y) \\ 2g(x,y) - g(x,y-1) - g(x,y+1 \end{bmatrix}$$
$$\boldsymbol{\Lambda}^t = \left[\lambda_x(x,y),\ \lambda_y(x,y)\right];\ \ t : transpose \tag{4}$$

Different from a conventional *USM*, Λ works sensitive to edges but insensitive to slow gradients. As well, ***Cubic USM***[5] introduced a cubic function coupling a horizontal and a vertical luminance difference with a quadratic function sensitive to a high-gradient but less sensitive to a slow-gradient. ***Rational USM***[6] extended ***Cubic USM*** by enhancing the horizontal and vertical edges independently depending on a local activity. Many other ideas to merge sharpening and smoothing are reported such as ***Lower-Upper-Middle USM*** [7]or ***Fuzzy*** filters and so on.[8, 9]

Multi-Scale Image Sharpening Method

The above methods work insensitive to the flat area noise but don't suppress it perfectly. To suppress the flat area noise completely we developed a *multi-scale sharpening* method[11, 12] as shown in Fig. 3. Here, the image is clearly separated into the edge and non-edge areas by pre-scanning ***GD*** filter, generating the ***edge map*** $M(x, y)$. A non-edge flat area is assigned to $M(x, y)=0$, and edge areas are classified to $M(x, y)=1$: *soft edge*, $M(x, y)=2$: *medium edge*, and $M(x, y)=3$: *hard edge*. Thus, multiple ***GD*** operators with different deviations, σ_1, σ_2, and σ_3 are selectively applied by looking up the ***edge map*** $M(x, y)$. The flat areas with $M(x, y)=0$, are intentionally ***smoothed*** by operating the normal Gaussian filter as follows.

$$f(x,y)=\begin{cases} g(x,y)+z_M(x,y) & ;\ \text{for edge area } M(x,y)\neq 0 \\ G(x,y)\otimes g(x,y) & ;\ \text{for flat area } M(x,y)=0 \end{cases}$$
$$z_M(x,y)=-\nabla^2 G(x,y,\sigma_M)\otimes g(x,y) \tag{5}$$

where, $z_M(x, y)$ denotes a multi-scale ***GD*** signal with σ_M fit to enhance the *soft, medium,* and *hard* edges for M=1,2, and 3. The proposed method resulted in natural image sharpening with smoothing the background noises as shown in Fig. 4.

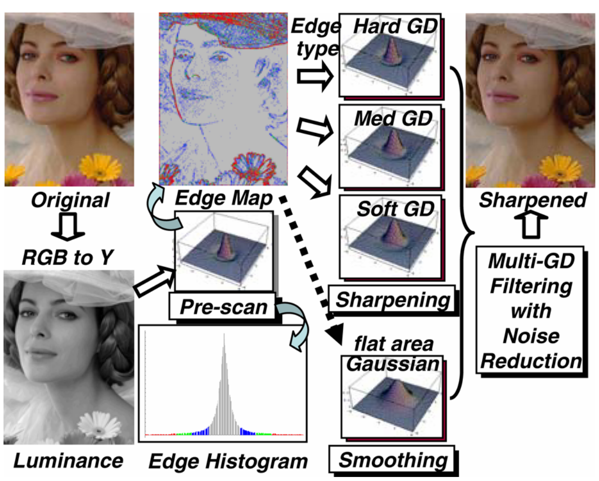

Figure 3. Proposed Multi-Scale Image Sharpening method

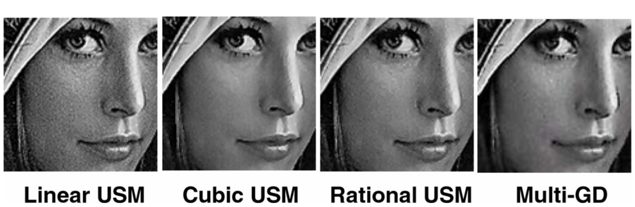

Figure 4. Comparison in adaptive image sharpening methods

Retinex and its Extension

Recently, *spatially-**invariant*** point process has been evolving into *spatially-**variant*** process. ***Retinex***[13] is a root of *spatially-**variant*** vision model. Basically it is a model to remove the non-uniformity of illumination as shown in Fig. 5. Simply, the image ***I*** captured by camera is equivalent to the product of the reflectance ***R*** and illuminant distribution ***L***. According to $\boldsymbol{R} \cong \boldsymbol{I}/\boldsymbol{L}$, the reflectance ***R*** is restored from Image ***I*** by inferring the illumination ***L***.[14-17]

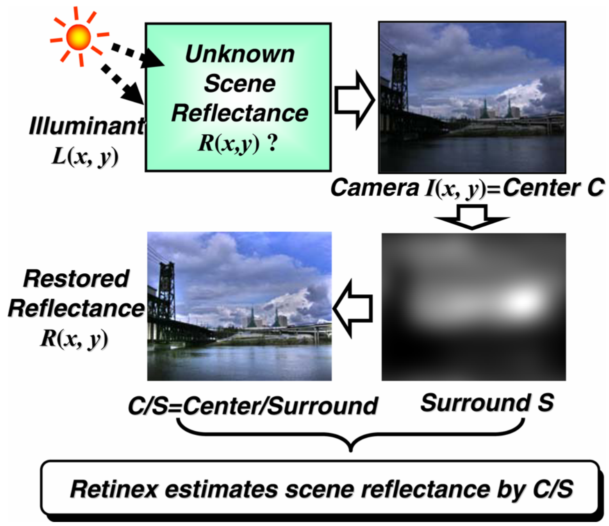

*Figure 5. Removal of Illumination in **C/S Retinex***

A typical ***C/S MSR*** (*Center/Surround Multi-Scale Retinex*) by *NASA*[18] has been effectively applied to improve the shadow appearance in digital camera images and now into ***HDR*** (*High-Dynamic Range*) images. However, in practice, it needs some technical skills, such as regulation in wide-spread output range due to the use of Log space, or correction in color imbalance caused by independent RGB channel process. In particular, the decision of weights for integrating multiple scales of ***SSR*** is a most troublesome problem. We proposed *Adaptive Scale-gain*[19,20] ***MSR*** model1 to solve these questions by, [1] Linear ***C/S*** process without *Log*, [2] Use of Luminance Y to keep color balance, and [3] Automatic weights setting, as given by the following equations.

Adaptive Scale-Gain MSR Algorithm

$$R_i(x,y,\sigma_m)=C\sum_{m=1}^{M} A(\sigma_m)\left\{\frac{I_i(x,y)}{S_m(x,y,\sigma_m)}\right\};\ i=R,G,B \tag{6}$$

$$S_m(x,y,\sigma_m)=G_m(x,y)\otimes Y(x,y) \tag{7}$$

$$G_m=K\exp\left\{-(x^2+y^2)/\sigma_m{}^2\right\},\ \sigma_m=2^m \tag{8}$$

$$A(\sigma_m)=\Sigma_{C/S}(\sigma_m)\Big/\sum_{m=1}^{M}\Sigma_{C/S}(\sigma_m) \tag{9}$$

$$\Sigma_{C/S}(\sigma_m)=\text{Standard deviation of } Y_{C/S} \tag{10}$$

$$Y_{C/S}(x,y,\sigma_m)=Y(x,y)/S_m(x,y,\sigma_m) \tag{11}$$

Fast MSR by Gaussian Pyramid

A fast algorithm to reduce the expensive ***MSR*** is introduced.

Our *adaptive scale-gain MSR* works in full automatic and nice in color, but it takes too much time to compute a large kernel size of surround. ***Gaussian Pyramid*** proved to be very effective to save the time by substituting it for small kernel size as shown in Fig. 6.

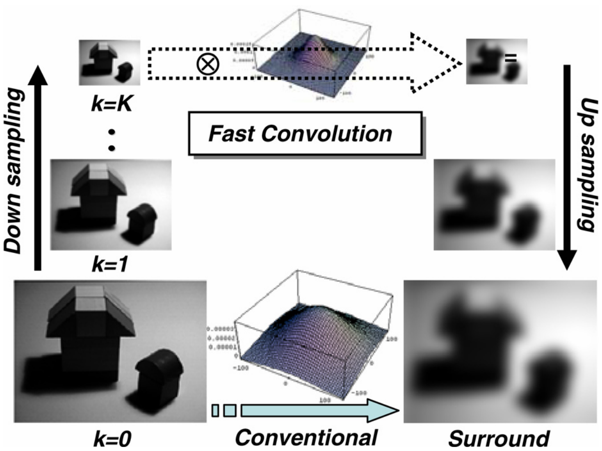

Figure 6. Fast convolution in MSR by Gaussian Pyramid

Retinex aims to reproduce the unbiased visual image, but the original is usually unknown unless we see the same scene by our naked eye. To seek the optimum parameters we synthesized a test target visually matched to the "color block" [q)] in our Lab by try and error using Photoshop as illustrated in Fig. 7. The color differences between this visual target and the processed images are evaluated. The *adaptive scale-gain MSR* resulted in the better color rendition than *NASA* with half in ΔE_{ab}^{*} even using the *Gaussian Pyramid.*

Local Contrast Enhancement

Retinex discounts a spatial non-uniformity of illumination based on ***C/S*** process, where the surround ***S*** reflects a global average in the image lightness. To recreate the viewer's sensation of the captured scene, a high dynamic range (***HDR***) has to be compressed to a low dynamic range (***LDR***) of the display devices. Recently ***HDR*** to ***LDR*** tone-mapping methods have been developed actively. *Spatially-invariant Tone Reproduction Curve* (***TRC***) operates point-wise on the image based on the global adaptation of human vision, as reported by Tumblin and Rushmeier[21] or Ward Larson.[23] While *spatially-variant Tone Reproduction Operator* (***TRO***) proposed by Chiu,[22] Pattanaik,[24] Fattal,[26] or Fairchild[27] attempts to preserve a local image contrast. However, most of ***TRO*** algorithms haven't clarified the relationship between the *input/output* ***TRC*** and the local image contrast. *Monobe* et al[28,29] introduced anew criterion to a local contrast. They proposed ***LCRT*** (*Local Contrast Range Transform*) operator to preserve a local contrast between input and output images under the given ***TRC*** as illustrated in Fig. 8.

The condition to preserve the local contrast is simply described by

$$\frac{g(x,y)}{g_{ave}(x,y)} = \frac{f(x,y)}{f_{ave}(x,y)} \tag{12}$$

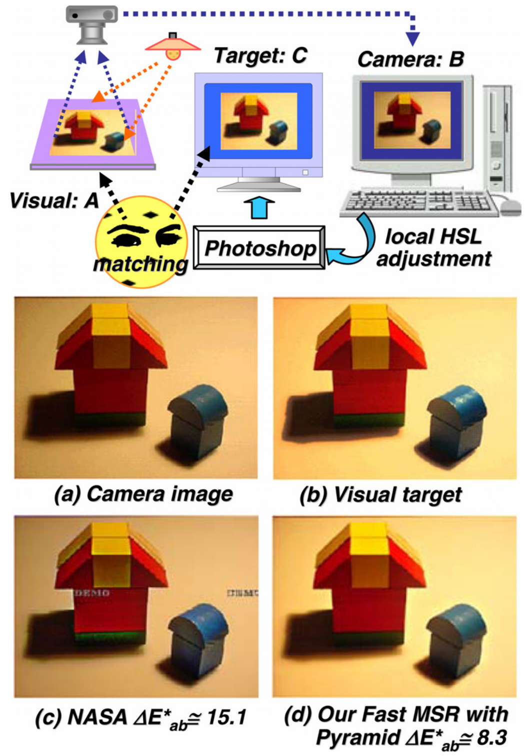

Figure 7. Synthesis of visual test target and evaluation.

Here, f(x,y) and fave(x,y) denote the input luminance level and its local average, g(x,y) and gave(x,y) denote the output luminance level and the local average of each pixel at (x, y), respectively. Taking the *Log* and denoting the variables in capital letters,

$$G(x,y) - G_{ave}(x,y) = F(x,y) - F_{ave}(x,y) \tag{13}$$

In *Log space,* taking a first-order Taylor expansion,

$$P(F_{ave}(x,y)) \cong P(F(x,y)) + \left\{\frac{dP(F(x,y))}{dF(x,y)}\right\} \cdot (F_{ave}(x,y) - F(x,y)) \tag{14}$$

where, $P\{F(x\ ,\ y)\}$ denotes an arbitrary ***TRC*** in *Log space*. Finally, the basic formula of ***LCRT*** is described in linear space as

$$g(x,y) = p(f(x,y)) \times \left(\frac{f(x,y)}{f_{ave}(x,y)}\right)^{\left\{1 - \frac{f(x,y)}{p(f(x,y))} \cdot \frac{dp(f(x,y))}{df(x,y)}\right\}} \tag{15}$$

Application of LCRT to Digital Video Camera

Digital video camera (*DVC*) is required to reproduce the main objects in shadow to middle luminance range as nice as possible through the dynamic range compression. For this reason, *DVC* product has adopted a "*knee curve*" to dynamic range compression, but it is difficult to preserve the highlight contrast. We applied ***LCRT*** to *DVC* and improved the visibility in a highlight range.[29] Because the *knee curve* = $p\{f(x\ ,y)\}$ in Equation (3) must be differentiable, it's approximated by a continuous cubic function and a local average is taken by convolving the luminance image *Y* with a *Gaussian* filter as illustrated in Fig. 9.

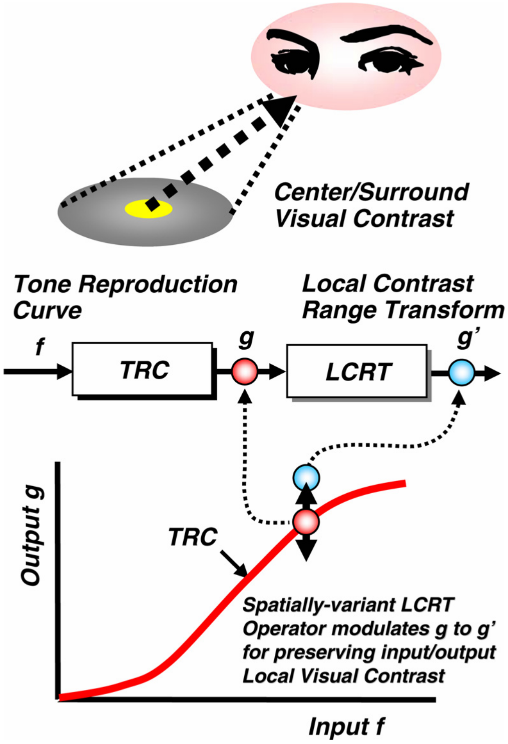

Figure 8. Local Contrast Enhancement by LCRT

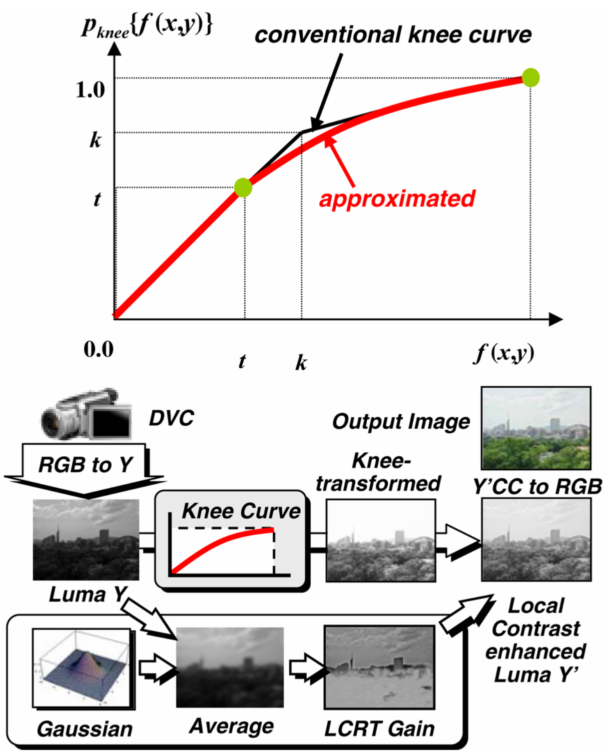

Figure 9 Application of LCRT to contrast enhancement in DVC

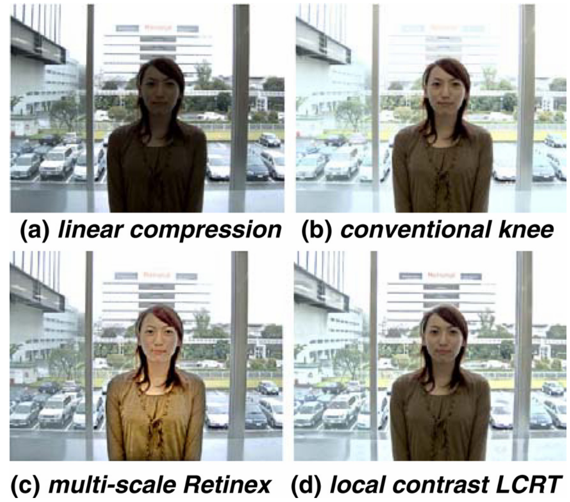

Figure 10. Improved highlight contrast in background by LCRT

$$f_{ave}(x,y) = G_m(x,y) \otimes Y(x,y);\ G_m = K exp\left\{-(x^2 + y^2)/\sigma_m{}^2\right\} \qquad (16)$$

Figure 10 shows how a background contrast is improved in the highlight *of DVC* image. ***Retinex*** also enhances the contrast but reproduces a girl just as taken under uniform illumination, while ***LCRT*** preserves it as same as original under the natural light.

Application of LCRT to Video Projector

LCRT is also applied to improve the visibility in display devices under ambient light. Figure 11 illustrates a *"fadeless"* projection system[28] by ***LCRT***. The system works to keep the visual contrast on screen in light room as same as that in dark room. Figure 12 shows a sample after vs. before ***LCRT*** process.

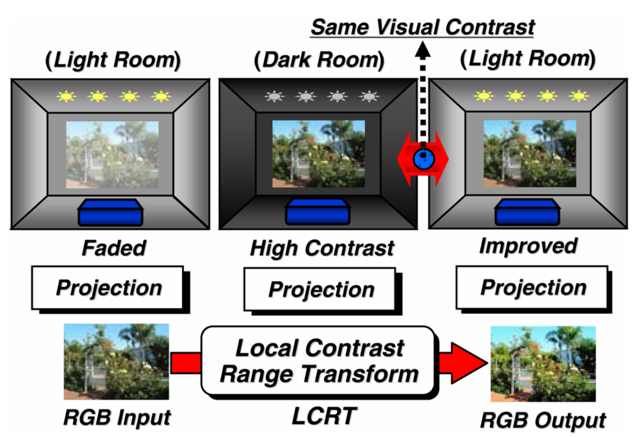

Figure 11. Fadeless video projection by LCRT

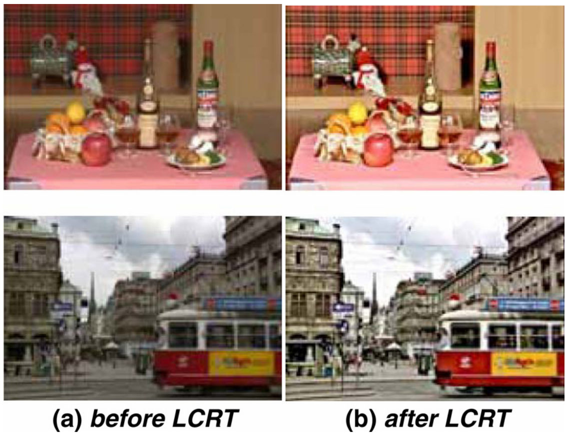

Figure 12. Enhanced contrasts for video projector by LCRT

INTELLIGENT Color Transform

Color reproduction systems have advanced as follows.

1st generation: ***Colorimetric Color Reproduction*** based on device-dependent closed system.

2nd generation: ***Cross-media Color Reproduction*** based on device-independent open system.

3rd generation: ***Color Appearance Reproduction*** based on human visual ***CAM*** (*Color Appearance Model*).

Now which way should we step forward to the next age? With the advent of various *FPD* (*Flat Panel Display*) raised a curtain on the new display age in place of CRT. A novel ***CMS*** for *FPD* is necessary. Since new devices such as LCD or PDP create a gamut space different from CRT, ***GMA*** (*Gamut Mapping Algorithm*) is one of the key technologies[31] to make use of the display gamut for pleasant color imaging. Towards "pleasant color imaging", we have developed

[1] ***Versatile GMA39*** that maps from "wide" to "narrow" or "narrow" to "wide" gamut in bi-direction.

[2] ***Scene-referred color transfer model***[37] that produces a pleasant image modeled on a reference target image.

Scene-referred Color Reproduction

CMS transfers a color via ***PCS*** with device profile created by measuring any standard color chart as a colorimetric reference. A "*scene-referred*" model tries to reproduce a color most to his/her taste learning from a reference target scene without using test chart as illustrated in Fig. 13.

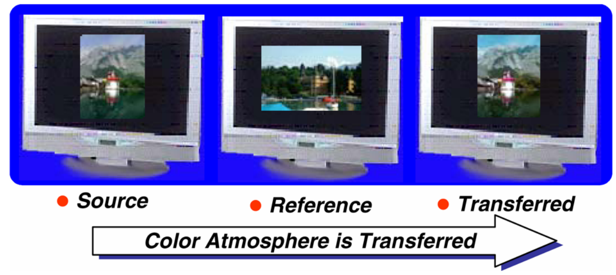

Figure 13. Concept of scene-referred color transfer system

Reinhard et al[32] tried to transfer the scene color from one to another using vision-based $l\alpha\beta$ color model. *Zhang* et al[33] applied it to correct the color imbalance between the right and left image in the same panoramic scene. These approaches are addressed to transfer the color atmosphere from one scene to another but the model doesn't work well for the scenes with color dissimilarity. Although $l\alpha\beta$ is a vision-based de-correlated color space, its major axes don't always correspond to the principal components of individual image. Hence, $l\alpha\beta$ works better for the scenes with color similarity but not for color dissimilarity. We introduced ***PC*** (*Principal Component*) matching model[34-36] to interchange the object colors between a pair of clusters with color dissimilarity. Different from $l\alpha\beta$ model, a ***PC*** of source color cluster is matched to that of target through the matching matrix $\boldsymbol{M}$ as shown in Fig. 14. The matching matrix ${}_{jk}\boldsymbol{M}_c$ is given by

$$ {}_{jk}\boldsymbol{M}_C = \left({}_k\boldsymbol{A}_{DST}^{-1}\right)\left({}_{jk}\boldsymbol{S}\right)\left({}_j\boldsymbol{A}_{org}\right) \tag{17} $$

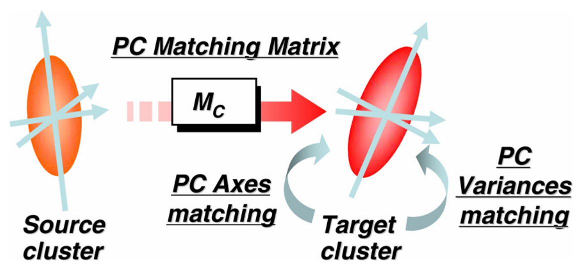

Figure 14. PC matching model between color clusters

It has two basic functions, "*Axes matching*" by rotating its cluster along the eigenvectors and "*Variance matching*" by scaling the color distribution along the ***PC*** axes. ${}_jAORG$ and ${}_kADST$ denote the eigen matrix for a source cluster *j* and a target cluster *k*. The scaling matrix ${}_{jk}\boldsymbol{S}$ is a diagonal matrix with the entries of *eigen values*' ratio. When the scene color distribution can be handled as a single cluster, its total scene color is easily interchanged with that of another scene. Figure 15 shows a scene color interchange sample. While, if the scene is composed of discrete multiple clusters, the ***PC*** matching should be done between each pair of separated color clusters after image segmentation. Figure 16 is our *local color interchange model* combining ***PC*** matching and *image segmentation*, where unsupervised [***k-means*** + ***Baysian***] classifier is introduced to segment the multi-clustered color objects.

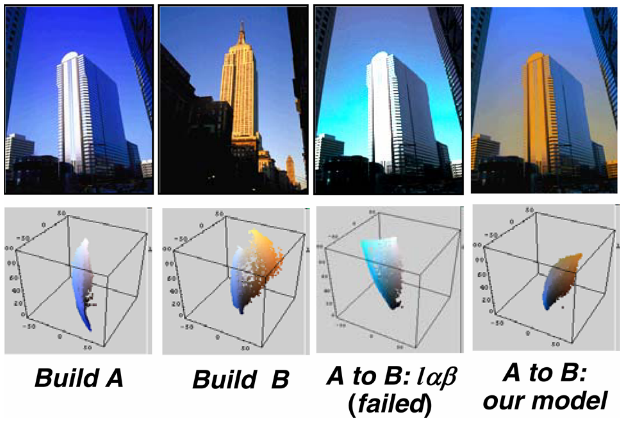

Figure 15. Color interchange sample for dissimilar color scenes

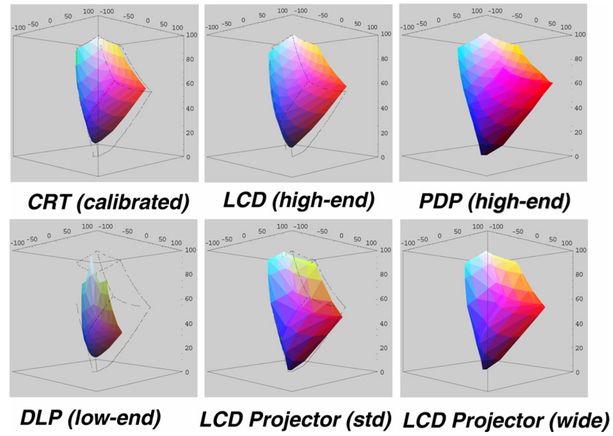

Figure 17. Color gamut of typical display devices

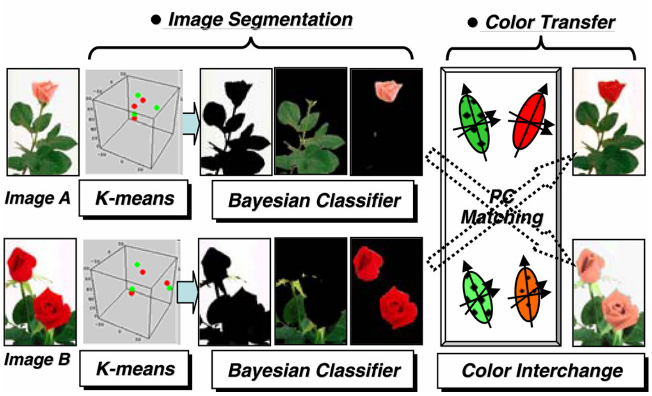

Figure 16. Local scene color interchange model for multiple clusters

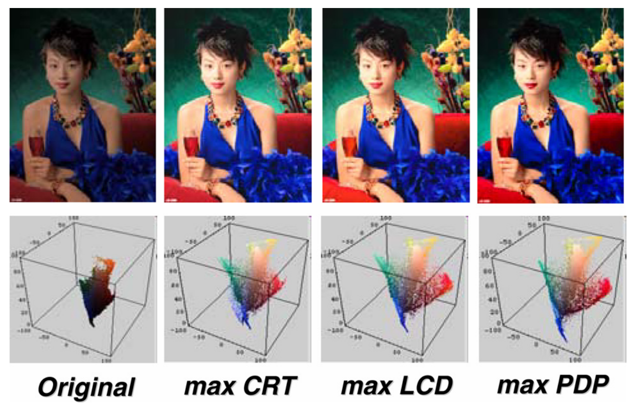

Figure 18. Most colorful images on display by gamut expansion GMA

To What Extent Is Image Gamut Expandable?

The up-to-date imaging devices are going ahead towards *"realistic color reproduction"* with ***HDR*** and *wide* ***Gamut*** besides *high resolution*. Since a source image has its own color gamut, "pleasant" color is reproduced by applying an adaptive ***GMA***[38, 39] depending on the image gamut is wider or narrower than the display's. A digital camera image in our daily life rarely fulfills the display gamut so that "*gamut expansion*" is useful to render the more vivid colors. Figure 17 shows a gamut comparison of recent display devices. Although CRT gamut can be calculated by a model-based additive mixture of RGB primary, those of LCD or PDP don't obey such linear model because the matrix cells are coupled with cross-talks. Hence the gamut surfaces in Fig. 17 are created based on spectral measurement of color chips on display screen. Figure 17 introduces a sample image by our *gamut expansion* ***GMA*** expanded up to the gamut boundary of each device. It tells how higher chromatic images are possible to reproduce.

Conclusions

The paper introduced our activities on *color image processing* by *Vision-based* and *Region-based* spatially-variant image transform. Taking the *scientific* fruits on vision research into *engineering,* a new field of applications for pleasant and realistic color imaging will be expected to open.[1-3]

References

1. Hiroaki Kotera: Image-dependent Quality…., PICS '02, 53-59 (2002)
2. Hiroaki Kotera: Intelligent Image Processing, ICIS'02, 606-607 (2002)
3. Hiroaki Kotera: A Comfortable Imaging …, IMID'04, 435-439 (2004)
4. Polesel, et al: Adaptive Unsharp Masking …, ICIP-97, 267-270 (1997)
5. G. Ramponi: A Cubic Unsharp…, Signal Process., 67, 211-222 (1998)
6. G. Ramponi: A Rational Unsharp ….., J. E. I., 7, 2, 333-338 (1998)
7. R. C. Hardie and C. Boncelet: LUM filters...… for smoothing and sharpening, IEEE Trans., Signal Processing, 41, 3, 1061-1076 (1993)
8. F. Russo: An image enhancement …combining sharpening and noise reduction, IEEE Trans., Instrument and Measure, 51,4, 824-828 (2002)
9. H. S. Kam, et al: An Improved Image Enhancement …. Smoothing & Sharpening, IEEE Conf., Image Processing, Tencon-745 (2003)
10. D. Marr and E. Hildreth: Theory of edge detection, Proc. R. Soc. Lond., B207, pp.187-217 (1980)
11. H. Kotera and H. Wang: Multi-Scale Image Sharp …, 10th CIC (2002)
12. H. Kotera and H. Wang: Multi-Scale Image ...…, J. E. I., 14(1), (2004)
13. Land and McCann: …the Retinex Theory, J. O. S. A.., 61, 1–11(1971)

14. Brainard and Wandell: Analysis of the Retinex Theory of Color Vision, J. O. S. A., A-3, 1651-661 (1986)
15. B. Funt et al, Retinex in MATLAB, 8th CIC (2000)
16. R. Kimmel et al: A Variational…, Elect. Imag., SPIE, 4662-49 (2002)
17. J. J. McCann, Retinex at 40, J. E. I., 13(1), 6-145 (2004).
18. D. J. Jobson et al: Properties and Performance of the Center/Surround Retinex, IEEE Trans., Image Proc., 6, 3, 451-462 (1997)
19. H. Kotera and M. Fujita: … Adaptive Scale-Gain ....,10th CIC (2002)
20. Y. Yoda and H. Kotera: …Adaptive Linear Retinex .…NIP21 (2004)
21. Tumblin and Rushmeier: Tone reproduction for realistic images, IEEE Computer Graphics and Application, 13, 42-48 (1993)
22. K.Chiu, et al: Spatially nonuniform…,Graphics Interface'93, 245-253
23. G. W. Larson, et al: A Visibility Matching … for High Dynamic Range Scenes, IEEE Trans. Vis. Comp. Graph., 3, 4, 291-306 (1997)
24. Pattanaik, et al: A multiscale model..., SIGGRAPH 98, 287-298 (1998)
25. J. Tumblin, et al: Two Methods for Display of High Contrast Images, ACM Trans. Graphics, vol. 18, 1, 56-94, (1999)
26. R. Fattal, et al: Gradient Domain…, SIGGRAPH 02, 249-256 (2002)
27. M. D. Fairchild and G. M. Johnson: Meet iCAM…, 10th CIC (2002)
28. Y. Monobe,…and H. Kotera: Fadeless Image Projection Preserving Local Contrast Under Ambient Light, 12th CIC (2004)
29. Y. Monobe, …and H. Kotera: HDR Compression for Digital Video Camera Using Local Contrast Enhancement, ICCE'05, 5.3.4 (2005)
30. J. Morovic and M. R. Luo: The Fundamentals of Gamut Mapping, A survey, Jour. IS&T, 45, 3, 283-290 (2001)
31. N. Katoh et al: Three-D gamut mapping.., J. E. I., 8(4), 365-379(1999)
32. E. Reinhard, et al: Color transfer between images, IEEE Computer Graphics and Applications, 34-40 (2001)
33. M. Zhang and N. D. Georganas: Fast color correction…, Real-Time Imaging 10, 23-30 (2004)
34. H. Kotera, et al: Object-Oriented Color Matching …, 6th CIC (1998)
35. H. Kotera, et al, Object-to-object …, J. E. I.,10(4)/1, 977-987 (2001)
36.\ H. Kotera and T. Horiuchi: Unsupervised …, NIP21, 622-626 (2004)
37. H. Kotera and T. Horiuchi: "Automatic Interchange ..., 12th CIC (2004)
38. H. Kotera and R. Saito: Compact …., J. E. I.,12 (4), 660-668 (2003)
39. R. Saito and H. Kotera: A versatile 3D Gamut..…, NIP21 (2004)

Author Biography

Hiroaki Kotera received his B.S degree from Nagoya Inst. Tech. and Doctorate from Univ. of Tokyo. In 1963, he joined Matsushita Electric Industrial Co. Since 1973, he worked in image processing at Matsushita Res. Inst. Tokyo. In 1996, he moved to Chiba University. He is a professor at Dept. 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, and from ISJ in 2003. He is a Fellow of IS&T.